

# Corpus des Deutschen Bundesrechts (C-DBR-Source)

## COMPILATION REPORT

Version 2021-01-05

License MIT-0

DOI: [10.5281/zenodo.4384767](https://doi.org/10.5281/zenodo.4384767)

|                  |   |
|------------------|---|
| <b>Titel</b>     | Source Code des »Corpus des Deutschen Bundesrechts«   |
| <b>Abkürzung</b> | C-DBR-Source  |
| <b>Autor</b>     | Seán Fobbe  |
| <b>Version</b>   | 2021-01-05  |
| <b>Download</b>  | <a href="https://doi.org/10.5281/zenodo.4384767">https://doi.org/10.5281/zenodo.4384767</a> |
| <b>Lizenz</b>    | MIT No Attribution (MIT-0)  |

### Zitiervorschlag

*Seán Fobbe* (2021). Source Code des »Corpus des Deutschen Bundesrechts« (C-DBR-Source). Version 2021-01-05. Zenodo. DOI: 10.5281/zenodo.4384767.

### Digital Object Identifier (DOI): Concept DOI und Version DOI

Soweit nicht anders angegeben ist die DOI immer eine »Version DOI« und bezieht sich nur auf eine bestimmte Version der Software. Sie verlinkt daher nur Version 2021-01-05. Für das Gesamtkonzept der Software steht eine »Concept DOI« zur Verfügung, die auf der Zenodo-Seite jeder Version unter »Cite all versions?« zu finden ist. Sie lautet 10.5281/zenodo.4072934. Die »Concept DOI« verlinkt immer die aktuellste Version.

### Lizenz: MIT No Attribution (MIT-0)

Copyright — 2021— Seán Fobbe

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the »Software«), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so.

THE SOFTWARE IS PROVIDED »AS IS«, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

### Disclaimer

Dieser Datensatz ist eine private wissenschaftliche Initiative und steht in keiner Verbindung zu Behörden, Gerichten oder anderen amtlichen Stellen der Bundesrepublik Deutschland.

# Inhaltsverzeichnis

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Einleitung</b>  | <b>8</b>  |
| 1.1      | Überblick . . . . .  | 8         |
| 1.2      | Endprodukte . . . . .  | 8         |
| 1.3      | Kompilierung . . . . .   | 9         |
| 1.3.1    | Datensatz . . . . .  | 9         |
| 1.3.2    | Codebook . . . . .   | 9         |
| 1.4      | Systemanforderungen . . . . .  | 10        |
| <b>2</b> | <b>Parameter</b>   | <b>11</b> |
| 2.1      | Name des Datensatzes . . . . .   | 11        |
| 2.2      | DOI des Datensatz-Konzeptes . . . . .                                  | 11        |
| 2.3      | DOI der konkreten Version . . . . .                                    | 11        |
| 2.4      | Verzeichnis für Analyse-Ergebnisse . . . . .                           | 11        |
| 2.5      | Optionen: Quanteda . . . . .   | 11        |
| 2.6      | Optionen: Knitr . . . . .  | 11        |
| 2.6.1    | Ausgabe-Format . . . . .   | 11        |
| 2.6.2    | DPI für Raster-Grafiken . . . . .                                      | 11        |
| 2.6.3    | Ausrichtung von Grafiken im Compilation Report . . . . .               | 11        |
| 2.7      | Frequenztabellen: Liste zu prüfender Variablen . . . . .               | 12        |
| <b>3</b> | <b>Vorbereitung</b>  | <b>13</b> |
| 3.1      | Datumsstempel . . . . .  | 13        |
| 3.2      | Datum und Uhrzeit (Beginn) . . . . .                                   | 13        |
| 3.3      | Ordner für Analyse-Ergebnisse erstellen . . . . .                      | 13        |
| 3.4      | Packages Laden . . . . .   | 13        |
| 3.5      | Zusätzliche Funktionen einlesen . . . . .                              | 15        |
| 3.6      | Quanteda-Optionen setzen . . . . .                                     | 15        |
| 3.7      | Knitr Optionen setzen . . . . .  | 15        |
| 3.8      | Vollzitate statistischer Software . . . . .                            | 15        |
| 3.9      | Parallelisierung aktivieren . . . . .                                  | 15        |
| 3.9.1    | Logische Kerne . . . . .   | 16        |
| 3.9.2    | Quanteda . . . . .   | 16        |
| 3.9.3    | Data.table . . . . .   | 16        |
| 3.9.4    | DoParallel . . . . .   | 16        |
| <b>4</b> | <b>Funktionen definieren</b>   | <b>17</b> |
| 4.1      | Transformation von Gliederungs-Metadaten . . . . .                     | 17        |
| 4.2      | NA in leere Listen-Elemente einsetzen . . . . .                        | 17        |
| <b>5</b> | <b>Download vorbereiten</b>  | <b>18</b> |
| 5.1      | XML-Inhaltsverzeichnis einlesen . . . . .                              | 18        |
| 5.2      | Links zu XML-Dateien aus XML-Inhaltsverzeichnis extrahieren . . . . .  | 18        |
| 5.3      | Links zu HTML Landing Pages generieren . . . . .                       | 18        |
| 5.4      | Funktion anzeigen . . . . .  | 18        |
| 5.5      | Links aus HTML Landing Pages extrahieren . . . . .                     | 18        |
| 5.6      | Dateinamen von PDF und EPUB-Dateien in separate Vektoren sortieren . . | 19        |
| 5.7      | Vektor der Langtitel erstellen . . . . .                               | 19        |
| 5.7.1    | Namen bereinigen und kürzen . . . . .                                  | 19        |

|          |   |           |
|----------|---|-----------|
| 5.7.2    | Index des 1. AEG bestimmen . . . . .                      | 19        |
| 5.7.3    | Erstes AEG umbenennen . . . . .                           | 19        |
| 5.8      | Vektor der Kurztitel erstellen . . . . .                  | 19        |
| 5.9      | Vektoren der Titel vereinigen . . . . .                   | 20        |
| 5.10     | Prüfung auf Namens-Kollisionen . . . . .                  | 20        |
| 5.11     | Bereinigung von Namens-Kollisionen . . . . .              | 20        |
| 5.12     | Dateierweiterungen hinzufügen . . . . .                   | 20        |
| 5.13     | Links zu EPUB-Dateien erstellen . . . . .                 | 21        |
| 5.14     | Links zu PDF-Dateien erstellen . . . . .                  | 21        |
| 5.15     | Data Table für Download vorbereiten . . . . .             | 21        |
| 5.16     | Abkürzungsverzeichnis erstellen . . . . .                 | 21        |
| 5.17     | Download Table als CSV speichern . . . . .                | 21        |
| 5.18     | Abkürzungsverzeichnis als CSV speichern . . . . .         | 22        |
| 5.19     | Anzahl herunterzuladender Dateien . . . . .               | 22        |
| 5.19.1   | Pro Format . . . . .                                      | 22        |
| 5.19.2   | Insgesamt . . . . .                                       | 22        |
| <b>6</b> | <b>Verarbeitung der DTD und XML-Dateien mit Anlagen</b>   | <b>23</b> |
| 6.1      | Document Type Definition herunterladen . . . . .          | 23        |
| 6.2      | Download der XML-Dateien . . . . .                        | 23        |
| 6.3      | Download-Ergebnis . . . . .                               | 23        |
| 6.3.1    | Anzahl herunterzuladender Dateien . . . . .               | 23        |
| 6.3.2    | Anzahl heruntergeladener Dateien . . . . .                | 23        |
| 6.3.3    | Fehlbetrag . . . . .                                      | 23        |
| 6.3.4    | Fehlende Dateien . . . . .                                | 24        |
| 6.4      | Extrahieren der XML-Dateien und ihrer Anlagen . . . . .   | 24        |
| 6.5      | XML Dateien auflisten und Dateigrößen speichern . . . . . | 24        |
| 6.6      | Zu extrahierende Metadaten-Variablen definieren . . . . . | 24        |
| 6.7      | Korpus erstellen: Einzelnormen . . . . .                  | 25        |
| 6.7.1    | Beginn XML Parsing . . . . .                              | 25        |
| 6.7.2    | XML Parsen . . . . .                                      | 25        |
| 6.7.3    | Liste in Data Table umwandeln . . . . .                   | 28        |
| 6.7.4    | Ende XML Parsing . . . . .                                | 28        |
| 6.7.5    | Dauer XML Parsing . . . . .                               | 28        |
| 6.7.6    | Variable »doc_id« erstellen . . . . .                     | 28        |
| 6.7.7    | Variablen-Name für Ausfertigungsdatum anpassen . . . . .  | 28        |
| 6.7.8    | Variable »fundstellentyp« anpassen . . . . .              | 28        |
| 6.7.9    | Variable »bulddate_iso« erstellen . . . . .               | 29        |
| 6.7.10   | Variable »ausfertigung_jahr« hinzufügen . . . . .         | 29        |
| 6.7.11   | Variable »doi_concept« hinzufügen . . . . .               | 29        |
| 6.7.12   | Variable »doi_version« hinzufügen . . . . .               | 29        |
| 6.7.13   | Variable »version« hinzufügen . . . . .                   | 29        |
| 6.8      | Stichprobe für Qualitätsprüfung ziehen . . . . .          | 29        |
| 6.9      | Korpus erstellen: Rechtsakte . . . . .                    | 30        |
| 6.9.1    | Variablen definieren . . . . .                            | 30        |
| 6.9.2    | Einzelnormen zu Rechtsakten vereinigen . . . . .          | 30        |
| 6.9.3    | Variable »dateiname« in »doc_id« umbenennen . . . . .     | 30        |
| 6.10     | Datensatz erstellen: XML-Metadaten . . . . .              | 32        |
| 6.10.1   | Beginn XML Parsing . . . . .                              | 32        |

|           |  |            |
|-----------|--|------------|
| 6.10.2    | XML Parsen . . . . .                                     | 32         |
| 6.10.3    | Liste in Data Table umwandeln . . . . .                  | 33         |
| 6.10.4    | Ende XML Parsing . . . . .                               | 33         |
| 6.10.5    | Dauer XML Parsing . . . . .                              | 33         |
| 6.10.6    | Variablen-Name für Ausfertigungsdatum anpassen . . . . . | 34         |
| 6.10.7    | Variable »fundstellentyp« anpassen . . . . .             | 34         |
| 6.10.8    | Variable »builddate_iso« erstellen . . . . .             | 34         |
| 6.10.9    | Variable »ausfertigung_jahr« hinzufügen . . . . .        | 34         |
| 6.10.10   | Variable »doi_concept« hinzufügen . . . . .              | 34         |
| 6.10.11   | Variable »doi_version« hinzufügen . . . . .              | 34         |
| 6.10.12   | Variable »version« hinzufügen . . . . .                  | 34         |
| 6.11      | Wiederverpacken der XML-Dateien . . . . .                | 34         |
| 6.11.1    | XML-Dateien verpacken . . . . .                          | 35         |
| 6.11.2    | Anhänge zu XML-Dateien verpacken . . . . .               | 35         |
| <b>7</b>  | <b>Frequenztabellen erstellen: Einzelnormen</b>          | <b>36</b>  |
| 7.1       | Funktion anzeigen . . . . .                              | 36         |
| 7.2       | Liste zu prüfender Variablen . . . . .                   | 37         |
| 7.3       | Frequenztabellen erstellen . . . . .                     | 37         |
| <b>8</b>  | <b>Frequenztabellen erstellen: Rechtsakte</b>            | <b>96</b>  |
| 8.1       | Variablen ignorieren . . . . .                           | 96         |
| 8.2       | Liste zu prüfender Variablen . . . . .                   | 96         |
| 8.3       | Frequenztabellen erstellen . . . . .                     | 96         |
| <b>9</b>  | <b>Frequenztabellen erstellen: XML-Metadaten</b>         | <b>105</b> |
| 9.1       | Liste zu prüfender Variablen . . . . .                   | 105        |
| 9.2       | Frequenztabellen erstellen . . . . .                     | 105        |
| <b>10</b> | <b>Frequenztabellen visualisieren</b>                    | <b>115</b> |
| 10.1      | Präfixe erstellen . . . . .                              | 115        |
| 10.2      | Tabellen für Einzelnormen einlesen . . . . .             | 115        |
| 10.3      | Tabellen für Rechtsakte einlesen . . . . .               | 115        |
| 10.4      | Tabellen für XML-Metadaten einlesen . . . . .            | 115        |
| 10.5      | Periodikum . . . . .                                     | 116        |
| 10.5.1    | Einzelnormen . . . . .                                   | 116        |
| 10.5.2    | Rechtsakte . . . . .                                     | 118        |
| 10.5.3    | XML-Metadaten . . . . .                                  | 120        |
| 10.6      | Ausfertigungsjahr . . . . .                              | 122        |
| 10.6.1    | Einzelnormen . . . . .                                   | 122        |
| 10.6.2    | Rechtsakte . . . . .                                     | 123        |
| 10.6.3    | XML-Metadaten . . . . .                                  | 124        |
| <b>11</b> | <b>Korpus-Analytik</b>                                   | <b>126</b> |
| 11.1      | Quanteda-Korpora erstellen . . . . .                     | 126        |
| 11.2      | Berechnung linguistischer Kennwerte . . . . .            | 126        |
| 11.2.1    | Einzelnormen . . . . .                                   | 126        |
| 11.2.2    | Rechtsakte . . . . .                                     | 126        |
| 11.3      | Variablen-Namen anpassen . . . . .                       | 127        |
| 11.3.1    | Einzelnormen . . . . .                                   | 127        |

|           |  |            |
|-----------|--|------------|
| 11.3.2    | Rechtsakte . . . . .                             | 127        |
| 11.4      | Kennwerte den Gesamtkorpora hinzufügen . . . . . | 127        |
| 11.4.1    | Einzelnormen . . . . .                           | 127        |
| 11.4.2    | Rechtsakte . . . . .                             | 127        |
| 11.5      | Anzahl Variablen der Datensätze . . . . .        | 127        |
| 11.6      | Alle Variablen-Namen der Datensätze . . . . .    | 128        |
| 11.7      | Quantitative Variablen . . . . .                 | 129        |
| 11.7.1    | Ausfertigungsdatum . . . . .                     | 129        |
| 11.7.2    | Ausfertigungsjahr . . . . .                      | 130        |
| 11.8      | Linguistische Kennwerte: Einzelnormen . . . . .  | 131        |
| 11.8.1    | Zusammenfassungen berechnen . . . . .            | 131        |
| 11.8.2    | Zusammenfassungen anzeigen . . . . .             | 132        |
| 11.8.3    | Zusammenfassungen speichern . . . . .            | 132        |
| 11.9      | Linguistische Kennwerte: Rechtsakte . . . . .    | 133        |
| 11.9.1    | Zusammenfassungen berechnen . . . . .            | 133        |
| 11.9.2    | Zusammenfassungen anzeigen . . . . .             | 134        |
| 11.9.3    | Zusammenfassungen speichern . . . . .            | 134        |
| 11.10     | Density . . . . .                                | 135        |
| 11.10.1   | Density (Tokens) . . . . .                       | 135        |
| 11.10.2   | Density (Typen) . . . . .                        | 137        |
| 11.10.3   | Density (Sätze) . . . . .                        | 139        |
| <b>12</b> | <b>CSV-Dateien erstellen</b>                     | <b>141</b> |
| 12.1      | Einzelnormen (Korpus) . . . . .                  | 141        |
| 12.1.1    | Name für CSV definieren . . . . .                | 141        |
| 12.1.2    | Datensatz speichern . . . . .                    | 141        |
| 12.2      | Einzelnormen (Metadaten) . . . . .               | 141        |
| 12.2.1    | Name für CSV definieren . . . . .                | 141        |
| 12.2.2    | Datensatz speichern . . . . .                    | 141        |
| 12.3      | Rechtsakte (Korpus) . . . . .                    | 141        |
| 12.3.1    | Name für CSV definieren . . . . .                | 141        |
| 12.3.2    | Datensatz speichern . . . . .                    | 142        |
| 12.4      | Rechtsakte (Metadaten) . . . . .                 | 142        |
| 12.4.1    | Name für CSV definieren . . . . .                | 142        |
| 12.4.2    | Datensatz speichern . . . . .                    | 142        |
| 12.5      | XML-Metadaten . . . . .                          | 142        |
| 12.5.1    | Name für CSV definieren . . . . .                | 142        |
| 12.5.2    | Datensatz speichern . . . . .                    | 142        |
| <b>13</b> | <b>Download der PDF-Dateien</b>                  | <b>143</b> |
| 13.1      | Download durchführen . . . . .                   | 143        |
| 13.2      | Download-Ergebnis . . . . .                      | 143        |
| 13.2.1    | Anzahl herunterzuladender Dateien . . . . .      | 143        |
| 13.2.2    | Anzahl heruntergeladener Dateien . . . . .       | 143        |
| 13.2.3    | Fehlbetrag . . . . .                             | 143        |
| 13.2.4    | Fehlende Dateien . . . . .                       | 143        |
| <b>14</b> | <b>TXT-Dateien erstellen</b>                     | <b>144</b> |
| 14.1      | Anzahl zu extrahierender Dateien . . . . .       | 144        |
| 14.2      | Seiten zählen: Funktion anzeigen . . . . .       | 144        |

|           |  |            |
|-----------|--|------------|
| 14.3      | Anzahl zu extrahierender Seiten . . . . .                        | 144        |
| 14.4      | PDF extrahieren: Funktion anzeigen . . . . .                     | 145        |
| 14.5      | Text Extrahieren . . . . .                                       | 145        |
| <b>15</b> | <b>Download der EPUB-Dateien</b>                                 | <b>146</b> |
| 15.1      | Download durchführen . . . . .                                   | 146        |
| 15.2      | Download-Ergebnis . . . . .                                      | 146        |
| 15.2.1    | Anzahl herunterzuladender Dateien . . . . .                      | 146        |
| 15.2.2    | Anzahl heruntergeladener Dateien . . . . .                       | 146        |
| 15.2.3    | Fehlbetrag . . . . .   | 146        |
| 15.2.4    | Fehlende Dateien . . . . .                                       | 146        |
| <b>16</b> | <b>Dateigrößen analysieren</b>                                   | <b>147</b> |
| 16.1      | Gesamtgröße . . . . .  | 147        |
| 16.1.1    | PDF-Dateien (MB) . . . . .                                       | 147        |
| 16.1.2    | EPUB-Dateien (MB) . . . . .                                      | 147        |
| 16.1.3    | XML-Dateien (MB) . . . . .                                       | 147        |
| 16.1.4    | TXT-Dateien (MB) . . . . .                                       | 147        |
| 16.1.5    | Korpus-Objekte in RAM (MB) . . . . .                             | 148        |
| 16.2      | Verteilung der Dateigrößen (PDF) . . . . .                       | 149        |
| 16.3      | Verteilung der Dateigrößen (EPUB) . . . . .                      | 151        |
| 16.4      | Verteilung der Dateigrößen (XML) . . . . .                       | 153        |
| 16.5      | Verteilung der Dateigrößen (TXT) . . . . .                       | 155        |
| <b>17</b> | <b>ZIP-Archive erstellen</b>                                     | <b>157</b> |
| 17.1      | Verpacken der CSV-Dateien . . . . .                              | 157        |
| 17.2      | Verpacken der PDF-Dateien . . . . .                              | 157        |
| 17.3      | Verpacken der TXT-Dateien . . . . .                              | 157        |
| 17.4      | Verpacken der EPUB-Dateien . . . . .                             | 158        |
| 17.5      | Verpacken der Analyse-Dateien . . . . .                          | 158        |
| 17.6      | Verpacken der Source-Dateien . . . . .                           | 158        |
| <b>18</b> | <b>Kryptographische Hashes</b>                                   | <b>159</b> |
| 18.1      | Liste der ZIP-Archive erstellen . . . . .                        | 159        |
| 18.2      | Funktion anzeigen . . . . .                                      | 159        |
| 18.3      | Hashes berechnen . . . . .                                       | 160        |
| 18.4      | In Data Table umwandeln . . . . .                                | 160        |
| 18.5      | Index hinzufügen . . . . .                                       | 160        |
| 18.6      | Hashes in CSV-Datei speichern . . . . .                          | 160        |
| 18.7      | Leerzeichen hinzufügen um Zeilenumbruch zu ermöglichen . . . . . | 160        |
| 18.8      | In Bericht anzeigen . . . . .                                    | 161        |
| <b>19</b> | <b>Abschluss</b>   | <b>164</b> |
| 19.1      | Cluster stoppen . . . . .  | 164        |
| 19.2      | Datum und Uhrzeit (Ende) . . . . .                               | 164        |
| 19.3      | Laufzeit des gesamten Skripts . . . . .                          | 164        |
| 19.4      | Warnungen . . . . .  | 164        |
| <b>20</b> | <b>Parameter für strenge Replikationen</b>                       | <b>165</b> |
|           | <b>Literaturverzeichnis</b>                                      | <b>166</b> |

# 1 Einleitung

## 1.1 Überblick

Dieses Skript wertet das amtliche Internetangebot »Gesetze im Internet« (<https://www.gesetze-im-internet.de>) der Bundesrepublik Deutschland vollständig aus und kompiliert es in einen reichhaltigen menschen- und maschinenlesbaren Korpus. Es ist die Grundlage des **Corpus des Deutschen Bundesrechts (C-DBR)**.

Alle mit diesem Skript erstellten Datensätze werden dauerhaft kostenlos und urheberrechtsfrei auf Zenodo, dem wissenschaftlichen Archiv des CERN, veröffentlicht. Alle Versionen sind mit einem persistenten Digital Object Identifier (DOI) versehen. Die neueste Version des Datensatzes ist immer über den Link der Concept DOI erreichbar: <https://doi.org/10.5281/zenodo.3832111>

## 1.2 Endprodukte

Primäre Endprodukte des Skripts sind folgende ZIP-Archive:

1. Der volle Datensatz im CSV-Format, unterteilt in Einzelnormen; nur Rechtsakte mit veröffentlichtem Normtext sind erfasst
2. Die Metadaten aller Einzelnormen im CSV-Format (wie 1, nur ohne Normtexte)
3. Der volle Datensatz im CSV-Format, unterteilt in Rechtsakte; nur Rechtsakte mit veröffentlichtem Normtext sind erfasst
4. Die Metadaten aller Rechtsakte im CSV-Format (wie 3, nur ohne Normtexte)
5. Die Metadaten aller auf »Gesetze im Internet« als XML veröffentlichten Rechtsakte, im CSV-Format, unabhängig davon ob sie Normtext enthalten oder nicht
6. Der volle Datensatz im XML-Format, unterteilt in Rechtsakte; Grundlage für die CSV-Varianten
7. Alle Anlagen zu den XML-Dateien im jeweiligen Original-Format
8. Alle Rechtstexte im TXT-Format, unterteilt in Rechtsakte (deutlich reduzierter Umfang an Metadaten)
9. Alle Rechtstexte im PDF-Format, unterteilt in Rechtsakte (deutlich reduzierter Umfang an Metadaten)
10. Alle Rechtstexte im EPUB-Format, unterteilt in Gesetze (deutlich reduzierter Umfang an Metadaten)
11. Alle Analyse-Ergebnisse (Tabellen als CSV, Grafiken als PDF und PNG)

Zusätzlich werden für alle ZIP-Archive kryptographische Signaturen (SHA2-256 und SHA3-512) berechnet und in einer CSV-Datei hinterlegt. Die Analyse-Ergebnisse werden zum Ende hin nicht gelöscht, damit sie für die Codebook-Erstellung verwendet werden können. Weiterhin kann optional ein PDF-Bericht erstellt werden (siehe unter »Kompilierung«).



## 1.3 Kompilierung

Mit der Funktion `render()` von **rmarkdown** können der **vollständige Datensatz** und das **Codebook** kompiliert und die Skripte mitsamt ihrer Rechenergebnisse in ein gut lesbares PDF-Format überführt werden.

Alle Kommentare sind im roxygen2-Stil gehalten. Die beiden Skripte können daher auch **ohne** `render()` regulär als R-Skripte ausgeführt werden. Es wird in diesem Fall kein PDF-Bericht erstellt und Diagramme werden nicht abgespeichert.

### 1.3.1 Datensatz

Um den **vollständigen Datensatz** zu kompilieren und einen PDF-Bericht zu erstellen, kopieren Sie bitte alle im Source-Archiv bereitgestellten Dateien in einen leeren Ordner und führen mit R diesen Befehl aus:

```
rmarkdown::render(input = "C-DBR_Source_CorpusCreation.R",
                  output_file = paste0("C-DBR_",
                                       Sys.Date(),
                                       "_CompilationReport.pdf"),
                  envir = new.env())
```

### 1.3.2 Codebook

Um das **Codebook** zu kompilieren und einen PDF-Bericht zu erstellen, kopieren Sie bitte alle im Source-Archiv bereitgestellten Dateien in einen leeren Ordner und führen im Anschluss an die Kompilierung des Datensatzes (!) untenstehenden Befehl mit R aus.

Bei der Prüfung der GPG-Signatur wird ein Fehler auftreten und im Codebook dokumentiert, weil die Daten nicht mit meiner Original-Signatur versehen sind. Dieser Fehler hat jedoch keine Auswirkungen auf die Funktionalität und hindert die Kompilierung nicht.

```
rmarkdown::render(input = "C-DBR_Source_CodebookCreation.R",
                  output_file = paste0("C-DBR_",
                                       Sys.Date(),
                                       "_Codebook.pdf"),
                  envir = new.env())
```

## 1.4 Systemanforderungen

Das Skript in seiner veröffentlichten Form kann nur unter Linux ausgeführt werden, da es Linux-spezifische Optimierungen (z.B. Fork Cluster) und Shell-Kommandos (z.B. OpenSSL) nutzt. Das Skript wurde unter Fedora Linux entwickelt und getestet. Die zur Kompilierung benutzte Version entnehmen Sie bitte dem **sessionInfo()**-Ausdruck am Ende dieses Berichts.

In der Standard-Einstellung wird das Skript vollautomatisch die maximale Anzahl an Rechenkernen/Threads auf dem System zu nutzen. Wenn die Anzahl Threads (Variable »fullCores«) auf 1 gesetzt wird, ist die Parallelisierung deaktiviert.

Auf der Festplatte sollten 4 GB Speicherplatz vorhanden sein.

Um die PDF-Berichte kompilieren zu können benötigen Sie das R package **rmarkdown**, eine vollständige Installation von L<sup>A</sup>T<sub>E</sub>X und alle in der Präambel-TEX-Datei angegebenen L<sup>A</sup>T<sub>E</sub>X Packages.

## 2 Parameter

### 2.1 Name des Datensatzes

```
datasetname <- "C-DBR"
```

### 2.2 DOI des Datensatz-Konzeptes

```
doi.concept <- "10.5281/zenodo.3832111" # checked
```

### 2.3 DOI der konkreten Version

```
doi.version <- "10.5281/zenodo.4384771" # checked
```

### 2.4 Verzeichnis für Analyse-Ergebnisse

Muss mit einem Schrägstrich enden!

```
outputdir <- paste0(getwd(), "/ANALYSE/")
```

### 2.5 Optionen: Quanteda

```
tokens_locale <- "de_DE"
```

### 2.6 Optionen: Knitr

#### 2.6.1 Ausgabe-Format

```
dev <- c("pdf", "png")
```

#### 2.6.2 DPI für Raster-Grafiken

```
dpi <- 150
```

#### 2.6.3 Ausrichtung von Grafiken im Compilation Report

```
fig.align <- "center"
```

## 2.7 Frequenztabellen: Liste zu prüfender Variablen

**Hinweis:** Nur diese Variablen werden bei der Erstellung der Frequenztabellen berücksichtigt.

```
vars.freqtable <- c("periodikum",  
                    "fundstellentyp",  
                    "check_neuf",  
                    "check_aufh",  
                    "check_sonst",  
                    "check_hinweis",  
                    "check_stand",  
                    "gliederungskennzahl",  
                    "ausfertigung_jahr",  
                    "doi_concept",  
                    "doi_version",  
                    "version")
```

## 3 Vorbereitung

### 3.1 Datumsstempel

Dieser Datumsstempel wird in alle Dateinamen eingefügt. Er wird am Anfang des Skripts gesetzt, für den Fall, dass die Laufzeit die Datumsbarriere durchbricht.

```
datestamp <- Sys.Date()
print(datestamp)
```

```
## [1] "2021-01-05"
```

### 3.2 Datum und Uhrzeit (Beginn)

```
begin.script <- Sys.time()
print(begin.script)
```

```
## [1] "2021-01-05 08:29:10 CET"
```

### 3.3 Ordner für Analyse-Ergebnisse erstellen

```
dir.create(outputdir)
```

### 3.4 Packages Laden

```
library(rvest)          # HTML/XML-Extraktion
```

```
## Loading required package: xml2
```

```
library(knitr)           # Professionelles Reporting
library(kableExtra)      # Verbesserte Kable Tabellen
library(pdftools)        # Extrahieren von PDF-Dateien
```

```
## Using poppler version 0.84.0
```

```
library(doParallel)    # Parallelisierung
```

```
## Loading required package: foreach
```

```
## Loading required package: iterators
```

```
## Loading required package: parallel
```

```
library(ggplot2)       # Fortgeschrittene Datenvisualisierung  
library(data.table)    # Fortgeschrittene Datenverarbeitung
```

```
## data.table 1.13.4 using 8 threads (see ?getDTthreads). Latest news: r-  
datatable.com
```

```
library(quanteda)      # Fortgeschrittene Computerlinguistik
```

```
## Package version: 2.1.2
```

```
## Parallel computing: 2 of 16 threads used.
```

```
## See https://quanteda.io for tutorials and examples.
```

```
##  
## Attaching package: 'quanteda'
```

```
## The following object is masked from 'package:utils':  
##  
##     View
```

```
library(scales)        # Skalierung von Diagrammen
```

### 3.5 Zusätzliche Funktionen einlesen

**Hinweis:** Die hieraus verwendeten Funktionen werden jeweils vor der ersten Benutzung in vollem Umfang angezeigt um den Lesefluss zu verbessern.

```
source("General_Source_Functions.R")
```

### 3.6 Quanteda-Optionen setzen

```
quanteda_options(tokens_locale = tokens_locale)
```

### 3.7 Knitr Optionen setzen

```
knitr::opts_chunk$set(fig.path = outputdir,  
                        dev = dev,  
                        dpi = dpi,  
                        fig.align = fig.align)
```

### 3.8 Vollzitate statistischer Software

```
knitr::write_bib(c(.packages()), "packages.bib")
```

```
## tweaking foreach
```

### 3.9 Parallelisierung aktivieren

Parallelisierung wird zur Beschleunigung des XML-Parsings, der Konvertierung von PDF zu TXT und der Datenanalyse mittels **quanteda** und **data.table** verwendet. Die Anzahl threads wird automatisch auf das verfügbare Maximum des Systems gesetzt, kann aber auch nach Belieben auf das eigene System angepasst werden. Die Parallelisierung kann deaktiviert werden, indem die Variable **fullCores** auf 1 gesetzt wird.

Der Download der Dateien von <https://www.gesetze-im-internet.de> ist bewusst nicht parallelisiert, damit das Skript nicht versehentlich als DoS-Tool verwendet wird.

Die hier verwendete Funktion **makeForkCluster()** ist viel schneller, funktioniert aber nur auf Unix-basierten Systemen (Linux, MacOS). Bei einer Ausführung unter Windows sollten Sie **makecluster()** verwenden.

### 3.9.1 Logische Kerne

```
fullCores <- detectCores()
print(fullCores)
```

```
## [1] 16
```

### 3.9.2 Quanteda

```
quanteda_options(threads = fullCores)
```

### 3.9.3 Data.table

```
setDTthreads(threads = fullCores)
```

### 3.9.4 DoParallel

```
cl <- makeForkCluster(fullCores)
registerDoParallel(cl)
```



## 4 Funktionen definieren

### 4.1 Transformation von Gliederungs-Metadaten

Wird bei der Umwandlung der Metadaten aus dem XML-Format benötigt. Konkret werden hierdurch Werte die nur einmal pro Abschnitt (z.B. Gliederungsüberschriften) hochgerechnet, damit jede Norm die ihr zugehörigen Abschnitts-Metadaten zugewiesen erhält.

```
f.heading.transform <- function(inputvec){  
  
  rep.text <- c("NA", inputvec[is.na(inputvec) == FALSE])  
  
  which <- c(1, which(is.na(inputvec) == FALSE), length(inputvec)+1)  
  rep.count <- diff(which)  
  
  rep <- data.table(rep.text, rep.count)  
  
  replist <- vector("list", rep[,.N])  
  
  for (i in 1:rep[,.N]){  
    replist[[i]]<- rep(rep.text[i], rep.count[i])  
  }  
  
  outvec <- unlist(replist)  
  return(outvec)  
}
```

### 4.2 NA in leere Listen-Elemente einsetzen

```
f.zero.NA <- function(x) if (length(x) == 0){  
  NA_character_  
}else{  
  paste(x, collapse = " ")}
```

## 5 Download vorbereiten

### 5.1 XML-Inhaltsverzeichnis einlesen

```
URL <- "https://www.gesetze-im-internet.de/gii-toc.xml"

XML <- read_xml(URL)
```

### 5.2 Links zu XML-Dateien aus XML-Inhaltsverzeichnis extrahieren

```
links <- xml_nodes(XML, "link")

links.xml <- xml_text(links)
```

### 5.3 Links zu HTML Landing Pages generieren

```
links.html <- gsub("/xml.zip",
                  "/index.html",
                  links.xml)
```

### 5.4 Funktion anzeigen

```
print(f.linkextract)
```

```
## function(URL){
##   tryCatch({
##     read_html(URL) %>%
##       html_nodes("a")%>%
##       html_attr('href')},
##     error=function(cond) {
##       return(NA)}
##   )
## }
```

### 5.5 Links aus HTML Landing Pages extrahieren

```
links.list <- lapply(links.html,
                    f.linkextract)

links.raw <- unlist(links.list)
```

## 5.6 Dateinamen von PDF und EPUB-Dateien in separate Vektoren sortieren

```
filenames.pdf <- grep (".pdf$",  
                        links.raw,  
                        ignore.case = TRUE,  
                        value = TRUE)  
  
filenames.epub <- grep (".epub$",  
                         links.raw,  
                         ignore.case = TRUE,  
                         value = TRUE)
```

## 5.7 Vektor der Langtitel erstellen

**Hinweis:** Das »Allgemeine Eisenbahngesetz« kommt zweimal mit demselben Namen vor, obwohl es sich um zwei unterschiedliche Gesetze handelt. Das erste AEG wird daher mit seiner Jahreszahl ergänzt um es von seinem Nachfolger abzugrenzen und den Dateinamen einzigartig zu machen.

```
longtitle.raw <- xml_nodes(XML, "title") %>% xml_text()
```

### 5.7.1 Namen bereinigen und kürzen

```
longtitle <- gsub(" ", "", longtitle.raw)  
longtitle <- gsub("[:punct:]", "", longtitle)
```

### 5.7.2 Index des 1. AEG bestimmen

```
AEGindex <- grep("AllgemeinesEisenbahngesetz", longtitle)[1]
```

### 5.7.3 Erstes AEG umbenennen

```
longtitle[AEGindex] <- "AllgemeinesEisenbahngesetz1994"
```

## 5.8 Vektor der Kurztitel erstellen

```
shorttitle <- filenames.pdf  
  
shorttitle <- gsub(".pdf",  
                  "",  
                  shorttitle)
```

```
shorttitle <- gsub("_",  
                  "",  
                  shorttitle)
```

## 5.9 Vektoren der Titel vereinigen

Die Kurz- und Langtitel werden zu einem Vektor zusammengefügt. Dieser wird dann auf maximal 200 Zeichen gekürzt, damit keine Probleme für Windows-User entstehen.

```
title <- paste(shorttitle,  
              longtitle,  
              sep="_")  
  
title <- strtrim(title, 200)
```

## 5.10 Prüfung auf Namens-Kollisionen

Kollidierende Namen anzeigen. Wenn Namens-Kollisionen bestehen (wie oben beim AEG) müssen diese unbedingt bereinigt werden, weil ansonsten beim Herunterladen ein Rechtsakt alle anderen mit dem gleichen Namen überschreibt.

```
title[duplicated(title)]
```

```
## character(0)
```

## 5.11 Bereinigung von Namens-Kollisionen

Eine manuelle Bereinigung von Kollisionen ist bevorzugt. Falls keine manuelle Bereinigung stattgefunden hat wird in diesem Schritt eine automatische Bereinigung durchgeführt.

```
title <- make.unique(title,  
                    sep = "-")
```

## 5.12 Dateierweiterungen hinzufügen

```
title.xml <- paste0(title, ".zip")  
title.epub <- paste0(title, ".epub")  
title.pdf <- paste0(title, ".pdf")
```

### 5.13 Links zu EPUB-Dateien erstellen

```
prelinks.epub <- gsub("xml.zip",  
  "",  
  links.xml)  
  
links.epub <- paste0(prelinks.epub,  
  filenames.epub)
```

### 5.14 Links zu PDF-Dateien erstellen

```
prelinks.pdf <- gsub("xml.zip",  
  "",  
  links.xml)  
  
links.pdf <- paste0(prelinks.pdf,  
  filenames.pdf)
```

### 5.15 Data Table für Download vorbereiten

```
download <- data.table(title.xml,  
  links.xml,  
  title.epub,  
  links.epub,  
  title.pdf,  
  links.pdf)
```

### 5.16 Abkürzungsverzeichnis erstellen

```
conctable <- data.table(shorttitle,  
  longtitle.raw)  
  
colnames(conctable) <- c("Kurztitel",  
  "Langtitel")
```

### 5.17 Download Table als CSV speichern

```
fwrite(download,  
  paste0(outputdir,  
    datasetname,  
    "_02_Links.csv"),  
  na = "NA")
```

## 5.18 Abkürzungsverzeichnis als CSV speichern

```
fwrite(conctable,  
      paste0(datasetname,  
             "-",  
             datestamp,  
             "_DE_Abkürzungsverzeichnis.csv"),  
      na = "NA")
```

## 5.19 Anzahl herunterzuladender Dateien

### 5.19.1 Pro Format

```
download[, .N]
```

```
## [1] 6577
```

### 5.19.2 Insgesamt

```
download[, .N] * 3
```

```
## [1] 19731
```

## 6 Verarbeitung der DTD und XML-Dateien mit Anlagen

### 6.1 Document Type Definition herunterladen

Die Document Type Definition (DTD) »definiert den Aufbau des XML-Formats zur Veröffentlichung der aktuellen Bundesgesetze und Rechtsverordnungen ueber [www.gesetze-im-internet.de](http://www.gesetze-im-internet.de)« (Zitat aus dem Inhalt der Datei).

```
download.file("http://www.gesetze-im-internet.de/dtd/1.01/gii-norm.dtd",
              paste0(datasetsname,
                    "_",
                    datestamp,
                    "_DE_XML_DocumentTypeDefinition_v1-01.dtd"))
```

### 6.2 Download der XML-Dateien

```
mapply(download.file,
        download$links.xml,
        download$title.xml)
```

### 6.3 Download-Ergebnis

#### 6.3.1 Anzahl herunterzuladender Dateien

```
download[,.N]
```

```
## [1] 6577
```

#### 6.3.2 Anzahl heruntergeladener Dateien

```
files.zip <- list.files(pattern = "\\\\.zip")
length(files.zip)
```

```
## [1] 6577
```

#### 6.3.3 Fehlbetrag

```
N.missing <- download[,.N] - length(files.zip)
print(N.missing)
```

```
## [1] 0
```

### 6.3.4 Fehlende Dateien

```
missing <- setdiff(download$title.xml,  
                    files.zip)  
print(missing)
```

```
## character(0)
```

## 6.4 Extrahieren der XML-Dateien und ihrer Anlagen

XML-Dateien und ihre Anlagen sind einzeln nach Rechtsakten in ZIP-Archiven verpackt. Diese werden nun extrahiert und die ZIP-Archive im Anschluss gelöscht.

```
files.zip <- list.files(pattern="\\.zip",  
                        ignore.case = TRUE)  
  
for (file in files.zip){  
  unzip(file)  
}  
  
unlink(files.zip)
```

## 6.5 XML Dateien auflisten und Dateigrößen speichern

```
files.xml <- list.files(pattern="\\.xml",  
                        ignore.case = TRUE)  
  
xml.MB <- file.size(files.xml) / 10^6
```

## 6.6 Zu extrahierende Metadaten-Variablen definieren

```
varlist <- c("jurabk",  
             "amtabk",  
             "ausfertigung-datum",  
             "periodikum",  
             "zitstelle",  
             "langue",  
             "kurzue")
```



## 6.7 Korpus erstellen: Einzelnormen

**Wichtiger Hinweis:** Es werden nur Rechtsakte ausgewertet, bei denen mindestens eine Einzelnorm mit Text-Inhalt vorhanden ist!

Die XML-Daten enthalten keine Leerzeichen zwischen den Tags, sowie zwischen den Tags und ihrem Inhalt. Damit beim Entfernen der Tags keine Inhalte zusammengefügt werden, wird die XML-Datei zunächst als Character-Vektor eingelesen, Leerzeichen hinzugefügt und im Anschluss erst die XML-Struktur eingelesen. Zwischen dem Anfang des Dokuments und dem ersten Tag darf kein Leerzeichen sein, dieses wird einzeln nachkorrigiert. Zusätzlicher whitespace ist bei späterer Text-Verarbeitung unschädlich und wird im Rahmen der Tokenisierung praktisch immer entfernt.

Falls dieser Schritt nicht vorgenommen wird können Ergebnisse so aussehen: »Zollkodex,d)alle Verfahren«

### 6.7.1 Beginn XML Parsing

```
begin.parse <- Sys.time()
```

### 6.7.2 XML Parsen

```
limit <- length(files.xml)

out <- foreach(z = 1:limit, .errorhandling = 'pass') %dopar% {

  ## XML als Character-Vektor einlesen
  xml.char <- readChar(files.xml[z],
                      file.info(files.xml[z])$size)

  ## Leerzeichen einfügen
  xml.char <- gsub(">", "> ", xml.char)
  xml.char <- gsub("<", " <", xml.char)
  xml.char <- sub(" <", "<", xml.char)

  ## XML-Struktur lesen
  XML <- read_xml(xml.char)

  ## Schleife vorbereiten
  nodes <- xml_nodes(XML, xpath = "//norm")
  scope <- seq_along(nodes)

  ## Inhaltsdaten extrahieren
  text.temp <- vector("list", max(scope))
  enbez.temp <- vector("list", max(scope))
  g.kennzahl.temp <- vector("list", max(scope))
  g.bez.temp <- vector("list", max(scope))
  g.titel.temp <- vector("list", max(scope))

  for (i in scope){
    text.temp[[i]] <- xml_nodes(nodes[i],
```

```

                                xpath = "textdaten//text//Content") %>% xml_
text(trim = TRUE)
    enbez.temp[[i]] <- xml_nodes(nodes[i],
                                xpath = "metadaten//enbez") %>% xml_text(
trim = TRUE)
    g.kennzahl.temp[[i]] <- xml_nodes(nodes[i],
                                xpath = "metadaten//gliederungseinheit/
/gliederungskennzahl") %>% xml_text(trim = TRUE)
    g.bez.temp[[i]] <- xml_nodes(nodes[i],
                                xpath = "metadaten//gliederungseinheit//
gliederungsbez") %>% xml_text(trim = TRUE)
    g.titel.temp[[i]] <- xml_nodes(nodes[i],
                                xpath = "metadaten//gliederungseinheit//
gliederungstitel") %>% xml_text(trim = TRUE)
}

## Leere Elemente mit NA kennzeichnen
enbez <- sapply(enbez.temp, f.zero.NA)
text <- sapply(text.temp, f.zero.NA)
g.kennzahl.pos <- sapply(g.kennzahl.temp, f.zero.NA)
g.bez.pos <- sapply(g.bez.temp, f.zero.NA)
g.titel.pos <- sapply(g.titel.temp, f.zero.NA)

## Gliederungsinformationen transformieren
gliederungskennzahl <- f.heading.transform(g.kennzahl.pos)
gliederungsbez <- f.heading.transform(g.bez.pos)
gliederungstitel <- f.heading.transform(g.titel.pos)
builddate_original <- xml_attr(nodes, attr = "builddate")

content.out <- data.table(builddate_original,
                          gliederungskennzahl,
                          gliederungsbez,
                          gliederungstitel,
                          enbez,
                          text)

content.out <- content.out[text != ""]

## Allgemeine Metadaten extrahieren
meta <- vector("list", length(varlist))

for (i in 1:length(varlist)){
  temp <- xml_node(XML, varlist[i]) %>% xml_text(trim = TRUE)
  meta[[i]] <- rep(temp,
                  content.out[, .N])
}

setDT(meta)
setnames(meta, new = varlist)

meta$fundstellentyp <- rep(xml_node(XML, "fundstelle") %>% xml_attr(attr = "
typ"),
                          content.out[, .N])

meta$dateiname <- rep(files.xml[z],

```

```

        content.out[,.N])

## Standangaben extrahieren
standtyp <- xml_nodes(XML, "standtyp") %>% xml_text(trim = TRUE)
standkommentar <- xml_nodes(XML, "standkommentar") %>% xml_text(trim = TRUE)
standcheck <- xml_nodes(XML, "standangabe") %>% xml_attr(attr = "checked")

dt.stand <- data.table(standtyp,
                      standkommentar,
                      standcheck)

if (dt.stand[,.N] > 0){

  ## Standkommentar
  dt.typ <- dt.stand[,
                    lapply(list(standkommentar),
                           function(x)paste(x, collapse = "  ")),
                    keyby = c("standtyp")]

  setnames(dt.typ,
           "V1",
           "standkommentar")

  dt.typ <- transpose(dt.typ,
                     make.names = "standtyp")

  setnames(dt.typ,
           names(dt.typ),
           tolower(names(dt.typ)))

  ## Standcheck
  dt.check <- dt.stand[,lapply(.SD, as.factor)][, .(standtyp, standcheck)]
  dt.check <- dt.check[, lapply(list(standtyp), unique), keyby = "
standcheck"]
  setnames(dt.check,
           "V1",
           "standtyp")

  dt.check <- transpose(dt.check, make.names = "standtyp")

  setnames(dt.check,
           names(dt.check),
           paste0("check_",
                 tolower(names(dt.check))))

  dt.stand.all <- cbind(dt.typ, dt.check)

  dt.stand.all.rep <- dt.stand.all[rep(dt.stand.all[, .I],
                                     content.out[,.N])]

  out.dt <- cbind(meta,
                 dt.stand.all.rep,
                 content.out)
}else{
  out.dt <- cbind(meta,

```

```

        content.out)
    }

    return(out.dt)
}

```

### 6.7.3 Liste in Data Table umwandeln

```

dt.normen <- rbindlist(out,
                      use.names = TRUE,
                      fill = TRUE)

```

### 6.7.4 Ende XML Parsing

```

end.parse <- Sys.time()

```

### 6.7.5 Dauer XML Parsing

```

end.parse - begin.parse

```

```

## Time difference of 46.50689 secs

```

### 6.7.6 Variable »doc\_id« erstellen

Eine einzigartige doc\_id wird benötigt um z.B. einen Quanteda-Korpus erstellen zu können. Diese wird aus dem Dateinamen zusammen mit einer Kollisionsnummer gebildet.

```

dt.normen$doc_id <- make.unique(dt.normen$dateiname)

```

### 6.7.7 Variablen-Name für Ausfertigungsdatum anpassen

```

setnames(dt.normen,
        "ausfertigung-datum",
        "ausfertigung_datum")

```

### 6.7.8 Variable »fundstellentyp« anpassen

```

dt.normen[grep("amtlich",
              dt.normen$fundstellentyp,
              invert = TRUE)]$fundstellentyp <- "nichtamtlich"

```

### 6.7.9 Variable »builddate\_iso« erstellen

```
dt.normen$builddate_iso <- as.POSIXct(dt.normen$builddate_original,  
                                     format = "%Y%m%d%H%M%S")
```

### 6.7.10 Variable »ausfertigung\_jahr« hinzufügen

```
dt.normen$ausfertigung_jahr <- year(dt.normen$ausfertigung_datum)
```

### 6.7.11 Variable »doi\_concept« hinzufügen

```
dt.normen$doi_concept <- rep(doi.concept, dt.normen[, .N])
```

### 6.7.12 Variable »doi\_version« hinzufügen

```
dt.normen$doi_version <- rep(doi.version, dt.normen[, .N])
```

### 6.7.13 Variable »version« hinzufügen

```
dt.normen$version <- as.character(rep(datestamp, dt.normen[, .N]))
```

## 6.8 Stichprobe für Qualitätsprüfung ziehen

```
idx <- sample(dt.normen[, .N], 300)  
  
check <- dt.normen[idx]  
  
fwrite(check, "QA_Stichprobe_Normen.csv")
```

## 6.9 Korpus erstellen: Rechtsakte

### 6.9.1 Variablen definieren

```
varlist.r1 <- gsub("ausfertigung-datum",  
                  "ausfertigung_datum",  
                  varlist)  
  
standvars <- c("stand",  
               "aufh",  
               "neuf",  
               "hinweis",  
               "sonst")  
  
standvars <- c(standvars,  
               paste0("check_",  
                      standvars))  
  
varlist.r2 <- c(varlist.r1,  
                standvars,  
                "fundstellentyp",  
                "ausfertigung_jahr",  
                "doi_concept",  
                "doi_version",  
                "version")
```

### 6.9.2 Einzelnormen zu Rechtsakten vereinigen

```
text.rechtsakte <- dt.normen[,  
                             lapply(list(text),  
                                     function(x)paste(x, collapse = " ")),  
                             keyby = dateiname]  
  
setnames(text.rechtsakte,  
         "V1",  
         "text")  
  
meta.rechtsakte <- dt.normen[,  
                             lapply(.SD, unique),  
                             .SDcols = varlist.r2,  
                             keyby = dateiname]  
  
dt.rechtsakte <- text.rechtsakte[meta.rechtsakte,  
                                 on = "dateiname"]
```

### 6.9.3 Variable »dateiname« in »doc\_id« umbenennen

```
setnames(dt.rechtsakte,  
         "dateiname",  
         "doc_id")
```

## 6.10 Datensatz erstellen: XML-Metadaten

An dieser Stelle werden Metadaten für alle Rechtsakte von »Gesetze im Internet« erhoben, unabhängig davon ob die Rechtsakte Text enthalten oder nur mit Überschrift nachgewiesen sind.

### 6.10.1 Beginn XML Parsing

```
begin.parse <- Sys.time()
```

### 6.10.2 XML Parsen

```
out <- foreach(z = 1:limit, .errorhandling = 'pass') %dopar% {

  XML <- read_xml(files.xml[z])
  nodes <- xml_nodes(XML, xpath = "//norm//metadaten")
  scope <- 1:length(nodes)

  meta <- vector("list", length(varlist))

  for (i in 1:length(varlist)){
    meta[[i]] <- xml_node(XML, varlist[i]) %>% xml_text()
  }

  setDT(meta)
  setnames(meta, new = varlist)

  meta$fundstellentyp <- xml_node(XML, "fundstelle") %>% xml_attr(attr = "typ")

  meta$doc_id <- files.xml[z]

  meta$builddate_original <- xml_attr(XML, attr = "builddate")

  ## Standangaben extrahieren
  standtyp <- xml_nodes(XML, "standtyp") %>% xml_text(trim = TRUE)
  standkommentar <- xml_nodes(XML, "standkommentar") %>% xml_text(trim = TRUE)
  standcheck <- xml_nodes(XML, "standangabe") %>% xml_attr(attr = "checked")

  dt.stand <- data.table(standtyp,
                        standkommentar,
                        standcheck)

  if (dt.stand[, .N] > 0){

    ## Standkommentar
    dt.typ <- dt.stand[,
                      lapply(list(standkommentar),
                             function(x) paste(x, collapse = " ")),
                      keyby = c("standtyp")]

    setnames(dt.typ,
```



```

        "V1",
        "standkommentar")

dt.typ <- transpose(dt.typ,
                    make.names = "standtyp")

setnames(dt.typ,
         names(dt.typ),
         tolower(names(dt.typ)))

## Standcheck
dt.check <- dt.stand[,lapply(.SD, as.factor)][, .(standtyp, standcheck)]
dt.check <- dt.check[, lapply(list(standtyp), unique), keyby = "
standcheck"]
setnames(dt.check,
        "V1",
        "standtyp")

dt.check <- transpose(dt.check, make.names = "standtyp")

setnames(dt.check,
        names(dt.check),
        paste0("check_",
                tolower(names(dt.check))))

dt.stand.all <- cbind(dt.typ, dt.check)

meta <- cbind(meta,
              dt.stand.all)
}

return(meta)
}

```

### 6.10.3 Liste in Data Table umwandeln

```

dt.meta <- rbindlist(out,
                    use.names = TRUE,
                    fill = TRUE)

```

### 6.10.4 Ende XML Parsing

```

end.parse <- Sys.time()

```

### 6.10.5 Dauer XML Parsing

```

end.parse - begin.parse

```

```
## Time difference of 11.13108 secs
```

#### 6.10.6 Variablen-Name für Ausfertigungsdatum anpassen

```
setnames(dt.meta,  
         "ausfertigung-datum",  
         "ausfertigung_datum")
```

#### 6.10.7 Variable »fundstellentyp« anpassen

```
dt.meta[grep("amtlich", dt.meta$fundstellentyp, invert = TRUE)]$fundstellentyp <-  
  "nichtamtlich"
```

#### 6.10.8 Variable »builddate\_iso« erstellen

```
dt.meta$builddate_iso <- as.POSIXct(dt.meta$builddate_original,  
                                   format = "%Y%m%d%H%M%S")
```

#### 6.10.9 Variable »ausfertigung\_jahr« hinzufügen

```
dt.meta$ausfertigung_jahr <- year(as.IDate(dt.meta$ausfertigung_datum))
```

#### 6.10.10 Variable »doi\_concept« hinzufügen

```
dt.meta$doi_concept <- rep(doi.concept, dt.meta[,.N])
```

#### 6.10.11 Variable »doi\_version« hinzufügen

```
dt.meta$doi_version <- rep(doi.version, dt.meta[,.N])
```

#### 6.10.12 Variable »version« hinzufügen

```
dt.meta$version <- as.character(rep(datestamp, dt.meta[,.N]))
```

### 6.11 Wiederverpacken der XML-Dateien

Wiederverpacken der gesammelten XML-Dateien in ein einziges Archiv. Wiederverpacken der Anlagen in ein separates Archiv. Die Roh-Daten werden im Anschluss jeweils gelöscht.

### 6.11.1 XML-Dateien verpacken

```
zip(paste(datasetname,
          datestamp,
          "DE_XML_Datensatz.zip",
          sep = "_"),
    files.xml)

unlink(files.xml)
```

### 6.11.2 Anhänge zu XML-Dateien verpacken

```
attachments <- list.files(pattern="(\\.jpg)|\\.gif)|\\.pdf)|\\.png)",
                          ignore.case = TRUE)

zip(paste(datasetname,
          datestamp,
          "DE_XML_Anlagen.zip",
          sep = "_"),
    attachments)

unlink(attachments)
```

## 7 Frequenztabellen erstellen: Einzelnormen

### 7.1 Funktion anzeigen

```
print(f.fast.freqtable)
```

```
function(x, varlist = names(x), sumrow = TRUE, output.list = TRUE, output.kable = FALSE, output.csv = FALSE, outputdir = »./«, prefix = ){
```

```
## Begin List
freqtable.list <- vector("list", length(varlist))

## Calculate Frequency Table
for (i in seq_along(varlist)){

  varname <- varlist[i]

  freqtable <- x[, .N, keyby=c(paste0(varname))]

  freqtable[, c("exactpercent",
               "roundedpercent",
               "cumulpercent") := {
    exactpercent <- N/sum(N)*100
    roundedpercent <- round(exactpercent, 2)
    cumulpercent <- round(cumsum(exactpercent), 2)
    list(exactpercent,
         roundedpercent,
         cumulpercent)}]

  ## Calculate Summary Row
  if (sumrow == TRUE){
    colsums <- cbind("Total",
                    freqtable[, lapply(.SD, function(x){round(sum(x))}),
                      .SDcols = c("N",
                                   "exactpercent",
                                   "roundedpercent")
                    ], round(max(freqtable$cumulpercent)))

    colnames(colsums)[c(1,5)] <- c(varname, "cumulpercent")
    freqtable <- rbind(freqtable, colsums)
  }

  ## Add Frequency Table to List
  freqtable.list[[i]] <- freqtable

  ## Write CSV
  if (output.csv == TRUE){

    fwrite(freqtable,
           paste0(outputdir,
                  prefix,
                  varname,
```

```

        ".csv"),
        na = "NA")

}

## Output Kable
if (output.kable == TRUE){

  cat("\n-----\n")
  cat(paste0("Frequency Table for Variable:  ", varname, "\n"))
  cat("-----\n")
  cat(paste0("\n ",
             x[, .N, keyby=c(paste0(varname))][, .N],
             " unique value(s) detected.\n\n"))

  print(kable(freqtable,
              format = "latex",
              align = "r",
              booktabs=TRUE,
              longtable=TRUE) %>% kable_styling(latex_options = "repeat_
header"))
}
}

## Return List of Frequency Tables
if (output.list == TRUE){
  return(freqtable.list)
}

}

```

## 7.2 Liste zu prüfender Variablen

```
print(vars.freqtable)
```

```
## [1] "periodikum"      "fundstellentyp"  "check_neuf"
## [4] "check_aufh"      "check_sonst"     "check_hinweis"
## [7] "check_stand"     "gliederungskennzahl" "ausfertigung_jahr"
## [10] "doi_concept"     "doi_version"     "version"
```

## 7.3 Frequenztabelle erstellen

```
prefix <- paste0(datasetname,
                  "_01_Einzelnormen_Frequenztabelle_var-")
```

```
f.fast.freqtable(dt.normen,
```

```

varlist = vars.freqtable,
sumrow = TRUE,
output.list = FALSE,
output.kable = TRUE,
output.csv = TRUE,
outputdir = outputdir,
prefix = prefix)

```

---

Frequency Table for Variable: periodikum

---

21 unique value(s) detected.

| periodikum | N     | exactpercent | roundedpercent | cumulpercent |
|------------|-------|--------------|----------------|--------------|
| BAGVBl     | 1     | 0.0010139    | 0.00           | 0.00         |
| BAnz       | 1156  | 1.1721047    | 1.17           | 1.17         |
| BGBI       | 2286  | 2.3178472    | 2.32           | 3.49         |
| BGBI I     | 82671 | 83.8227242   | 83.82          | 87.31        |
| BGBI II    | 3265  | 3.3104861    | 3.31           | 90.62        |
| GBl DDR    | 233   | 0.2362460    | 0.24           | 90.86        |
| GBl DDR I  | 622   | 0.6306653    | 0.63           | 91.49        |
| GBl DDR II | 43    | 0.0435991    | 0.04           | 91.53        |
| GVBl BE    | 9     | 0.0091254    | 0.01           | 91.54        |
| NV         | 15    | 0.0152090    | 0.02           | 91.56        |
| RAnz       | 38    | 0.0385294    | 0.04           | 91.60        |
| RGBI       | 5634  | 5.7124896    | 5.71           | 97.31        |
| RGBI I     | 1657  | 1.6800844    | 1.68           | 98.99        |
| RGBI II    | 347   | 0.3518342    | 0.35           | 99.34        |
| RMBI       | 230   | 0.2332042    | 0.23           | 99.58        |
| VOBl BrZ   | 19    | 0.0192647    | 0.02           | 99.59        |
| VkBl       | 105   | 0.1064628    | 0.11           | 99.70        |
| WiGBI      | 122   | 0.1236996    | 0.12           | 99.82        |
| ZBl        | 80    | 0.0811145    | 0.08           | 99.91        |
| eBAnz      | 52    | 0.0527244    | 0.05           | 99.96        |
| Öff Anz    | 41    | 0.0415712    | 0.04           | 100.00       |
| Total      | 98626 | 100.0000000  | 100.00         | 100.00       |

(continued)

| periodikum | N | exactpercent | roundedpercent | cumulpercent |
|------------|---|--------------|----------------|--------------|
|------------|---|--------------|----------------|--------------|

Frequency Table for Variable: fundstellentyp

1 unique value(s) detected.

| fundstellentyp | N     | exactpercent | roundedpercent | cumulpercent |
|----------------|-------|--------------|----------------|--------------|
| amtlich        | 98626 | 100          | 100            | 100          |
| Total          | 98626 | 100          | 100            | 100          |

Frequency Table for Variable: check\_neuf

2 unique value(s) detected.

| check_neuf | N     | exactpercent | roundedpercent | cumulpercent |
|------------|-------|--------------|----------------|--------------|
| NA         | 74602 | 75.64131     | 75.64          | 75.64        |
| ja         | 24024 | 24.35869     | 24.36          | 100.00       |
| Total      | 98626 | 100.00000    | 100.00         | 100.00       |

Frequency Table for Variable: check\_aufh

2 unique value(s) detected.

| check_aufh | N     | exactpercent | roundedpercent | cumulpercent |
|------------|-------|--------------|----------------|--------------|
| NA         | 96969 | 98.319916    | 98.32          | 98.32        |
| ja         | 1657  | 1.680084     | 1.68           | 100.00       |
| Total      | 98626 | 100.000000   | 100.00         | 100.00       |

Frequency Table for Variable: check\_sonst

2 unique value(s) detected.

| check_sonst | N     | exactpercent | roundedpercent | cumulpercent |
|-------------|-------|--------------|----------------|--------------|
| NA          | 82810 | 83.96366     | 83.96          | 83.96        |
| ja          | 15816 | 16.03634     | 16.04          | 100.00       |
| Total       | 98626 | 100.00000    | 100.00         | 100.00       |

Frequency Table for Variable: check\_hinweis

2 unique value(s) detected.

| check_hinweis | N     | exactpercent | roundedpercent | cumulpercent |
|---------------|-------|--------------|----------------|--------------|
| NA            | 78531 | 79.62505     | 79.63          | 79.63        |
| ja            | 20095 | 20.37495     | 20.37          | 100.00       |
| Total         | 98626 | 100.00000    | 100.00         | 100.00       |

Frequency Table for Variable: check\_stand

2 unique value(s) detected.

| check_stand | N     | exactpercent | roundedpercent | cumulpercent |
|-------------|-------|--------------|----------------|--------------|
| NA          | 22202 | 22.51131     | 22.51          | 22.51        |
| ja          | 76424 | 77.48869     | 77.49          | 100.00       |
| Total       | 98626 | 100.00000    | 100.00         | 100.00       |

Frequency Table for Variable: gliederungskennzahl

1443 unique value(s) detected.

| gliederungskennzahl | N  | exactpercent | roundedpercent | cumulpercent |
|---------------------|----|--------------|----------------|--------------|
| 000                 | 77 | 0.0780727    | 0.08           | 0.08         |
| 000010              | 4  | 0.0040557    | 0.00           | 0.08         |



(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 000020              | 8    | 0.0081115    | 0.01           | 0.09         |
| 000030              | 7    | 0.0070975    | 0.01           | 0.10         |
| 000040              | 7    | 0.0070975    | 0.01           | 0.10         |
| 000050              | 8    | 0.0081115    | 0.01           | 0.11         |
| 000060              | 5    | 0.0050697    | 0.01           | 0.12         |
| 000070              | 3    | 0.0030418    | 0.00           | 0.12         |
| 000080              | 2    | 0.0020279    | 0.00           | 0.12         |
| 000090              | 1    | 0.0010139    | 0.00           | 0.12         |
| 000100              | 2    | 0.0020279    | 0.00           | 0.13         |
| 000110              | 3    | 0.0030418    | 0.00           | 0.13         |
| 000120              | 5    | 0.0050697    | 0.01           | 0.13         |
| 001000              | 2    | 0.0020279    | 0.00           | 0.14         |
| 001010              | 1    | 0.0010139    | 0.00           | 0.14         |
| 001020              | 3    | 0.0030418    | 0.00           | 0.14         |
| 001030              | 1    | 0.0010139    | 0.00           | 0.14         |
| 001040              | 1    | 0.0010139    | 0.00           | 0.14         |
| 001051              | 1    | 0.0010139    | 0.00           | 0.14         |
| 001060              | 2    | 0.0020279    | 0.00           | 0.14         |
| 001080              | 1    | 0.0010139    | 0.00           | 0.15         |
| 001090              | 1    | 0.0010139    | 0.00           | 0.15         |
| 001100              | 1    | 0.0010139    | 0.00           | 0.15         |
| 001101              | 1    | 0.0010139    | 0.00           | 0.15         |
| 001110              | 1    | 0.0010139    | 0.00           | 0.15         |
| 001120              | 5    | 0.0050697    | 0.01           | 0.16         |
| 010                 | 5998 | 6.0815606    | 6.08           | 6.24         |
| 010010              | 886  | 0.8983432    | 0.90           | 7.14         |
| 010010000010        | 1    | 0.0010139    | 0.00           | 7.14         |
| 010010000020        | 1    | 0.0010139    | 0.00           | 7.14         |
| 010010010           | 118  | 0.1196439    | 0.12           | 7.26         |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 010010020           | 117 | 0.1186300    | 0.12           | 7.38         |
| 010010020010        | 4   | 0.0040557    | 0.00           | 7.38         |
| 010010020010010     | 37  | 0.0375155    | 0.04           | 7.42         |
| 010010020010020     | 24  | 0.0243344    | 0.02           | 7.44         |
| 010010020020        | 13  | 0.0131811    | 0.01           | 7.45         |
| 010010020030        | 1   | 0.0010139    | 0.00           | 7.46         |
| 010010030           | 29  | 0.0294040    | 0.03           | 7.48         |
| 010010030010        | 7   | 0.0070975    | 0.01           | 7.49         |
| 010010030020        | 17  | 0.0172368    | 0.02           | 7.51         |
| 010010030030        | 13  | 0.0131811    | 0.01           | 7.52         |
| 010010030040        | 13  | 0.0131811    | 0.01           | 7.54         |
| 010010030050        | 1   | 0.0010139    | 0.00           | 7.54         |
| 010010040           | 35  | 0.0354876    | 0.04           | 7.57         |
| 010010050           | 9   | 0.0091254    | 0.01           | 7.58         |
| 010010060           | 6   | 0.0060836    | 0.01           | 7.59         |
| 010010070010        | 10  | 0.0101393    | 0.01           | 7.60         |
| 010010070020        | 5   | 0.0050697    | 0.01           | 7.60         |
| 010020              | 777 | 0.7878247    | 0.79           | 8.39         |
| 010020010           | 136 | 0.1378947    | 0.14           | 8.53         |
| 010020010000030     | 1   | 0.0010139    | 0.00           | 8.53         |
| 010020010000040     | 1   | 0.0010139    | 0.00           | 8.53         |
| 010020010000050     | 1   | 0.0010139    | 0.00           | 8.53         |
| 010020010000060     | 1   | 0.0010139    | 0.00           | 8.53         |
| 010020020           | 131 | 0.1328250    | 0.13           | 8.67         |
| 010020020000070     | 1   | 0.0010139    | 0.00           | 8.67         |
| 010020020000080     | 1   | 0.0010139    | 0.00           | 8.67         |
| 010020020000090     | 1   | 0.0010139    | 0.00           | 8.67         |
| 010020020000100     | 1   | 0.0010139    | 0.00           | 8.67         |
| 010020020000110     | 1   | 0.0010139    | 0.00           | 8.67         |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 010020020000120     | 1   | 0.0010139    | 0.00           | 8.67         |
| 010020020010        | 17  | 0.0172368    | 0.02           | 8.69         |
| 010020020020        | 21  | 0.0212926    | 0.02           | 8.71         |
| 010020020030        | 9   | 0.0091254    | 0.01           | 8.72         |
| 010020020040        | 15  | 0.0152090    | 0.02           | 8.73         |
| 010020020050        | 18  | 0.0182508    | 0.02           | 8.75         |
| 010020030           | 101 | 0.1024071    | 0.10           | 8.85         |
| 010020030000130     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000140     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000170     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000171     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000172     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000190     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000200     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000210     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000220     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000230     | 1   | 0.0010139    | 0.00           | 8.86         |
| 010020030000240     | 1   | 0.0010139    | 0.00           | 8.87         |
| 010020030010        | 3   | 0.0030418    | 0.00           | 8.87         |
| 010020030020        | 2   | 0.0020279    | 0.00           | 8.87         |
| 010020030030        | 2   | 0.0020279    | 0.00           | 8.87         |
| 010020040           | 65  | 0.0659055    | 0.07           | 8.94         |
| 010020040000250     | 1   | 0.0010139    | 0.00           | 8.94         |
| 010020040000260     | 1   | 0.0010139    | 0.00           | 8.94         |
| 010020050           | 58  | 0.0588080    | 0.06           | 9.00         |
| 010020050020000380  | 1   | 0.0010139    | 0.00           | 9.00         |
| 010020050020000390  | 1   | 0.0010139    | 0.00           | 9.00         |
| 010020050020000400  | 1   | 0.0010139    | 0.00           | 9.00         |
| 010020050020000410  | 1   | 0.0010139    | 0.00           | 9.00         |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 010020050020000420  | 1   | 0.0010139    | 0.00           | 9.00         |
| 010020060           | 48  | 0.0486687    | 0.05           | 9.05         |
| 010020060000430     | 1   | 0.0010139    | 0.00           | 9.05         |
| 010020060000440     | 1   | 0.0010139    | 0.00           | 9.06         |
| 010020060000450     | 1   | 0.0010139    | 0.00           | 9.06         |
| 010020060000460     | 1   | 0.0010139    | 0.00           | 9.06         |
| 010020070           | 31  | 0.0314319    | 0.03           | 9.09         |
| 010020070010000461  | 1   | 0.0010139    | 0.00           | 9.09         |
| 010020070020000462  | 1   | 0.0010139    | 0.00           | 9.09         |
| 010020070020000463  | 1   | 0.0010139    | 0.00           | 9.09         |
| 010020070030000464  | 1   | 0.0010139    | 0.00           | 9.09         |
| 010020070040000465  | 1   | 0.0010139    | 0.00           | 9.09         |
| 010020080           | 41  | 0.0415712    | 0.04           | 9.14         |
| 010020090           | 1   | 0.0010139    | 0.00           | 9.14         |
| 010030              | 518 | 0.5252165    | 0.53           | 9.66         |
| 010030000470        | 1   | 0.0010139    | 0.00           | 9.66         |
| 010030000480        | 1   | 0.0010139    | 0.00           | 9.66         |
| 010030010           | 134 | 0.1358668    | 0.14           | 9.80         |
| 010030010010        | 8   | 0.0081115    | 0.01           | 9.81         |
| 010030010020        | 14  | 0.0141950    | 0.01           | 9.82         |
| 010030010030        | 20  | 0.0202786    | 0.02           | 9.84         |
| 010030010040        | 12  | 0.0121672    | 0.01           | 9.85         |
| 010030010050        | 4   | 0.0040557    | 0.00           | 9.86         |
| 010030010060        | 8   | 0.0081115    | 0.01           | 9.87         |
| 010030010070        | 8   | 0.0081115    | 0.01           | 9.87         |
| 010030010080        | 4   | 0.0040557    | 0.00           | 9.88         |
| 010030010380        | 2   | 0.0020279    | 0.00           | 9.88         |
| 010030010400        | 4   | 0.0040557    | 0.00           | 9.88         |
| 010030010440        | 1   | 0.0010139    | 0.00           | 9.89         |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 010030010450        | 3   | 0.0030418    | 0.00           | 9.89         |
| 010030020           | 85  | 0.0861842    | 0.09           | 9.98         |
| 010030020010        | 86  | 0.0871981    | 0.09           | 10.06        |
| 010030020020        | 40  | 0.0405573    | 0.04           | 10.10        |
| 010030020030        | 12  | 0.0121672    | 0.01           | 10.11        |
| 010030020040        | 12  | 0.0121672    | 0.01           | 10.13        |
| 010030020050        | 9   | 0.0091254    | 0.01           | 10.14        |
| 010030020060        | 18  | 0.0182508    | 0.02           | 10.15        |
| 010030030           | 70  | 0.0709752    | 0.07           | 10.23        |
| 010030030010        | 10  | 0.0101393    | 0.01           | 10.24        |
| 010030030020        | 10  | 0.0101393    | 0.01           | 10.25        |
| 010030030030        | 7   | 0.0070975    | 0.01           | 10.25        |
| 010030030040        | 2   | 0.0020279    | 0.00           | 10.25        |
| 010030030050        | 1   | 0.0010139    | 0.00           | 10.26        |
| 010030030060        | 1   | 0.0010139    | 0.00           | 10.26        |
| 010030040           | 39  | 0.0395433    | 0.04           | 10.30        |
| 010030050           | 40  | 0.0405573    | 0.04           | 10.34        |
| 010030060           | 6   | 0.0060836    | 0.01           | 10.34        |
| 010030060010        | 16  | 0.0162229    | 0.02           | 10.36        |
| 010030060020        | 8   | 0.0081115    | 0.01           | 10.37        |
| 010030060030        | 3   | 0.0030418    | 0.00           | 10.37        |
| 010030060040        | 3   | 0.0030418    | 0.00           | 10.37        |
| 010030060050        | 2   | 0.0020279    | 0.00           | 10.38        |
| 010030070           | 17  | 0.0172368    | 0.02           | 10.39        |
| 010031              | 2   | 0.0020279    | 0.00           | 10.39        |
| 010040              | 340 | 0.3447367    | 0.34           | 10.74        |
| 010040010           | 80  | 0.0811145    | 0.08           | 10.82        |
| 010040020           | 64  | 0.0648916    | 0.06           | 10.89        |
| 010040020010        | 3   | 0.0030418    | 0.00           | 10.89        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 010040020020        | 5   | 0.0050697    | 0.01           | 10.89        |
| 010040020030        | 2   | 0.0020279    | 0.00           | 10.90        |
| 010040030           | 41  | 0.0415712    | 0.04           | 10.94        |
| 010040040           | 21  | 0.0212926    | 0.02           | 10.96        |
| 010040040010        | 4   | 0.0040557    | 0.00           | 10.96        |
| 010040040020        | 9   | 0.0091254    | 0.01           | 10.97        |
| 010040040030        | 4   | 0.0040557    | 0.00           | 10.98        |
| 010040040040        | 9   | 0.0091254    | 0.01           | 10.98        |
| 010040040050        | 1   | 0.0010139    | 0.00           | 10.99        |
| 010040040060        | 3   | 0.0030418    | 0.00           | 10.99        |
| 010040040070        | 8   | 0.0081115    | 0.01           | 11.00        |
| 010040120010        | 1   | 0.0010139    | 0.00           | 11.00        |
| 010040120020        | 1   | 0.0010139    | 0.00           | 11.00        |
| 010040120030        | 1   | 0.0010139    | 0.00           | 11.00        |
| 010040120040        | 1   | 0.0010139    | 0.00           | 11.00        |
| 010040120050        | 1   | 0.0010139    | 0.00           | 11.00        |
| 010040120060        | 1   | 0.0010139    | 0.00           | 11.00        |
| 010041              | 2   | 0.0020279    | 0.00           | 11.01        |
| 010050              | 163 | 0.1652708    | 0.17           | 11.17        |
| 010050010           | 59  | 0.0598220    | 0.06           | 11.23        |
| 010050020           | 85  | 0.0861842    | 0.09           | 11.32        |
| 010050020020        | 2   | 0.0020279    | 0.00           | 11.32        |
| 010050030           | 49  | 0.0496826    | 0.05           | 11.37        |
| 010050030010        | 2   | 0.0020279    | 0.00           | 11.37        |
| 010050030020        | 1   | 0.0010139    | 0.00           | 11.37        |
| 010050030030        | 2   | 0.0020279    | 0.00           | 11.37        |
| 010050040           | 3   | 0.0030418    | 0.00           | 11.38        |
| 010060              | 179 | 0.1814937    | 0.18           | 11.56        |
| 010060010           | 46  | 0.0466408    | 0.05           | 11.60        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 010060020           | 73  | 0.0740170    | 0.07           | 11.68        |
| 010060020010        | 11  | 0.0111532    | 0.01           | 11.69        |
| 010060020020        | 10  | 0.0101393    | 0.01           | 11.70        |
| 010060020030        | 5   | 0.0050697    | 0.01           | 11.70        |
| 010060020040        | 14  | 0.0141950    | 0.01           | 11.72        |
| 010060020050        | 1   | 0.0010139    | 0.00           | 11.72        |
| 010060030           | 57  | 0.0577941    | 0.06           | 11.78        |
| 010060030810        | 7   | 0.0070975    | 0.01           | 11.78        |
| 010060030820        | 8   | 0.0081115    | 0.01           | 11.79        |
| 010060030830        | 3   | 0.0030418    | 0.00           | 11.80        |
| 010060030840        | 1   | 0.0010139    | 0.00           | 11.80        |
| 010060040           | 14  | 0.0141950    | 0.01           | 11.81        |
| 010060050           | 33  | 0.0334597    | 0.03           | 11.84        |
| 010060060           | 23  | 0.0233204    | 0.02           | 11.87        |
| 010060070           | 5   | 0.0050697    | 0.01           | 11.87        |
| 010070              | 219 | 0.2220510    | 0.22           | 12.10        |
| 010070010           | 8   | 0.0081115    | 0.01           | 12.10        |
| 010070010010        | 11  | 0.0111532    | 0.01           | 12.11        |
| 010070010020        | 6   | 0.0060836    | 0.01           | 12.12        |
| 010070020           | 14  | 0.0141950    | 0.01           | 12.13        |
| 010070030           | 14  | 0.0141950    | 0.01           | 12.15        |
| 010080              | 127 | 0.1287693    | 0.13           | 12.28        |
| 010080010           | 9   | 0.0091254    | 0.01           | 12.29        |
| 010080010010        | 2   | 0.0020279    | 0.00           | 12.29        |
| 010080010020        | 11  | 0.0111532    | 0.01           | 12.30        |
| 010080020           | 17  | 0.0172368    | 0.02           | 12.32        |
| 010080030           | 8   | 0.0081115    | 0.01           | 12.33        |
| 010090              | 60  | 0.0608359    | 0.06           | 12.39        |
| 010090010           | 15  | 0.0152090    | 0.02           | 12.40        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 010090020           | 12   | 0.0121672    | 0.01           | 12.41        |
| 010090030           | 5    | 0.0050697    | 0.01           | 12.42        |
| 010091              | 5    | 0.0050697    | 0.01           | 12.42        |
| 010092              | 1    | 0.0010139    | 0.00           | 12.42        |
| 010100              | 25   | 0.0253483    | 0.03           | 12.45        |
| 010100010           | 2    | 0.0020279    | 0.00           | 12.45        |
| 010100020           | 1    | 0.0010139    | 0.00           | 12.45        |
| 010110              | 37   | 0.0375155    | 0.04           | 12.49        |
| 010110010           | 1    | 0.0010139    | 0.00           | 12.49        |
| 010110020           | 1    | 0.0010139    | 0.00           | 12.49        |
| 010110030           | 1    | 0.0010139    | 0.00           | 12.49        |
| 010120              | 4    | 0.0040557    | 0.00           | 12.50        |
| 010130              | 9    | 0.0091254    | 0.01           | 12.51        |
| 010140              | 5    | 0.0050697    | 0.01           | 12.51        |
| 011                 | 11   | 0.0111532    | 0.01           | 12.52        |
| 012                 | 2    | 0.0020279    | 0.00           | 12.53        |
| 014                 | 2    | 0.0020279    | 0.00           | 12.53        |
| 020                 | 6174 | 6.2600126    | 6.26           | 18.79        |
| 020000500           | 1    | 0.0010139    | 0.00           | 18.79        |
| 020000510           | 1    | 0.0010139    | 0.00           | 18.79        |
| 020000520           | 1    | 0.0010139    | 0.00           | 18.79        |
| 020000530           | 1    | 0.0010139    | 0.00           | 18.79        |
| 020000531           | 1    | 0.0010139    | 0.00           | 18.79        |
| 020010              | 1611 | 1.6334435    | 1.63           | 20.43        |
| 020010010           | 293  | 0.2970819    | 0.30           | 20.72        |
| 020010020           | 230  | 0.2332042    | 0.23           | 20.96        |
| 020010020010        | 25   | 0.0253483    | 0.03           | 20.98        |
| 020010020020        | 8    | 0.0081115    | 0.01           | 20.99        |
| 020010020020010     | 4    | 0.0040557    | 0.00           | 20.99        |



(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 020010020020020     | 3   | 0.0030418    | 0.00           | 21.00        |
| 020010020020030     | 3   | 0.0030418    | 0.00           | 21.00        |
| 020010020020040     | 3   | 0.0030418    | 0.00           | 21.00        |
| 020010020020050     | 1   | 0.0010139    | 0.00           | 21.00        |
| 020010020030        | 12  | 0.0121672    | 0.01           | 21.02        |
| 020010020040        | 6   | 0.0060836    | 0.01           | 21.02        |
| 020010020050        | 5   | 0.0050697    | 0.01           | 21.03        |
| 020010020060        | 4   | 0.0040557    | 0.00           | 21.03        |
| 020010030           | 144 | 0.1460061    | 0.15           | 21.18        |
| 020010030010        | 11  | 0.0111532    | 0.01           | 21.19        |
| 020010030020        | 7   | 0.0070975    | 0.01           | 21.20        |
| 020010030030010     | 4   | 0.0040557    | 0.00           | 21.20        |
| 020010030030020010  | 5   | 0.0050697    | 0.01           | 21.20        |
| 020010030030020020  | 8   | 0.0081115    | 0.01           | 21.21        |
| 020010030040        | 4   | 0.0040557    | 0.00           | 21.22        |
| 020010040           | 91  | 0.0922678    | 0.09           | 21.31        |
| 020010040010        | 5   | 0.0050697    | 0.01           | 21.31        |
| 020010040020        | 3   | 0.0030418    | 0.00           | 21.32        |
| 020010040030        | 6   | 0.0060836    | 0.01           | 21.32        |
| 020010041           | 5   | 0.0050697    | 0.01           | 21.33        |
| 020010050           | 92  | 0.0932817    | 0.09           | 21.42        |
| 020010060           | 32  | 0.0324458    | 0.03           | 21.45        |
| 020010070           | 38  | 0.0385294    | 0.04           | 21.49        |
| 020010070010        | 4   | 0.0040557    | 0.00           | 21.50        |
| 020010070020        | 7   | 0.0070975    | 0.01           | 21.50        |
| 020010070030        | 7   | 0.0070975    | 0.01           | 21.51        |
| 020010080           | 19  | 0.0192647    | 0.02           | 21.53        |
| 020010090           | 31  | 0.0314319    | 0.03           | 21.56        |
| 020010100           | 11  | 0.0111532    | 0.01           | 21.57        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 020010110           | 7    | 0.0070975    | 0.01           | 21.58        |
| 020010120           | 9    | 0.0091254    | 0.01           | 21.59        |
| 020011              | 3    | 0.0030418    | 0.00           | 21.59        |
| 020011042           | 1    | 0.0010139    | 0.00           | 21.59        |
| 020020              | 1516 | 1.5371200    | 1.54           | 23.13        |
| 020020010           | 146  | 0.1480340    | 0.15           | 23.28        |
| 020020010010        | 31   | 0.0314319    | 0.03           | 23.31        |
| 020020010010010     | 2    | 0.0020279    | 0.00           | 23.31        |
| 020020010010020     | 1    | 0.0010139    | 0.00           | 23.31        |
| 020020010010030     | 3    | 0.0030418    | 0.00           | 23.32        |
| 020020010010040     | 1    | 0.0010139    | 0.00           | 23.32        |
| 020020010010050     | 3    | 0.0030418    | 0.00           | 23.32        |
| 020020010020        | 11   | 0.0111532    | 0.01           | 23.33        |
| 020020010030        | 6    | 0.0060836    | 0.01           | 23.34        |
| 020020010040        | 10   | 0.0101393    | 0.01           | 23.35        |
| 020020010050        | 2    | 0.0020279    | 0.00           | 23.35        |
| 020020010060        | 6    | 0.0060836    | 0.01           | 23.35        |
| 020020010070        | 3    | 0.0030418    | 0.00           | 23.36        |
| 020020010080        | 2    | 0.0020279    | 0.00           | 23.36        |
| 020020020           | 220  | 0.2230649    | 0.22           | 23.58        |
| 020020020010        | 33   | 0.0334597    | 0.03           | 23.62        |
| 020020020020        | 29   | 0.0294040    | 0.03           | 23.65        |
| 020020020030        | 25   | 0.0253483    | 0.03           | 23.67        |
| 020020020040        | 9    | 0.0091254    | 0.01           | 23.68        |
| 020020020050        | 19   | 0.0192647    | 0.02           | 23.70        |
| 020020020060        | 1    | 0.0010139    | 0.00           | 23.70        |
| 020020020070010     | 3    | 0.0030418    | 0.00           | 23.70        |
| 020020020070020     | 7    | 0.0070975    | 0.01           | 23.71        |
| 020020020070030     | 2    | 0.0020279    | 0.00           | 23.71        |

(continued)

| gliederungskennzahl | N  | exactpercent | roundedpercent | cumulpercent |
|---------------------|----|--------------|----------------|--------------|
| 020020030           | 93 | 0.0942956    | 0.09           | 23.81        |
| 020020030010        | 19 | 0.0192647    | 0.02           | 23.83        |
| 020020030020        | 23 | 0.0233204    | 0.02           | 23.85        |
| 020020030030        | 19 | 0.0192647    | 0.02           | 23.87        |
| 020020030040        | 17 | 0.0172368    | 0.02           | 23.89        |
| 020020030050        | 5  | 0.0050697    | 0.01           | 23.89        |
| 020020030060        | 2  | 0.0020279    | 0.00           | 23.89        |
| 020020030070        | 16 | 0.0162229    | 0.02           | 23.91        |
| 020020030080        | 4  | 0.0040557    | 0.00           | 23.91        |
| 020020030090        | 3  | 0.0030418    | 0.00           | 23.92        |
| 020020030100        | 1  | 0.0010139    | 0.00           | 23.92        |
| 020020040           | 64 | 0.0648916    | 0.06           | 23.98        |
| 020020043           | 1  | 0.0010139    | 0.00           | 23.98        |
| 020020050           | 19 | 0.0192647    | 0.02           | 24.00        |
| 020020050010        | 24 | 0.0243344    | 0.02           | 24.03        |
| 020020050020        | 9  | 0.0091254    | 0.01           | 24.04        |
| 020020060           | 44 | 0.0446130    | 0.04           | 24.08        |
| 020020060010        | 1  | 0.0010139    | 0.00           | 24.08        |
| 020020060020        | 6  | 0.0060836    | 0.01           | 24.09        |
| 020020060030        | 5  | 0.0050697    | 0.01           | 24.09        |
| 020020070           | 12 | 0.0121672    | 0.01           | 24.11        |
| 020020080           | 2  | 0.0020279    | 0.00           | 24.11        |
| 020020080010        | 6  | 0.0060836    | 0.01           | 24.11        |
| 020020080020        | 11 | 0.0111532    | 0.01           | 24.12        |
| 020020080030        | 4  | 0.0040557    | 0.00           | 24.13        |
| 020020080040        | 2  | 0.0020279    | 0.00           | 24.13        |
| 020020090           | 4  | 0.0040557    | 0.00           | 24.13        |
| 020020100           | 13 | 0.0131811    | 0.01           | 24.15        |
| 020021              | 23 | 0.0233204    | 0.02           | 24.17        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 020030              | 1068 | 1.0828788    | 1.08           | 25.25        |
| 020030010           | 148  | 0.1500618    | 0.15           | 25.40        |
| 020030010010        | 11   | 0.0111532    | 0.01           | 25.42        |
| 020030010020        | 5    | 0.0050697    | 0.01           | 25.42        |
| 020030010020010     | 2    | 0.0020279    | 0.00           | 25.42        |
| 020030010020020     | 7    | 0.0070975    | 0.01           | 25.43        |
| 020030010020030     | 2    | 0.0020279    | 0.00           | 25.43        |
| 020030010020040     | 1    | 0.0010139    | 0.00           | 25.43        |
| 020030010030        | 8    | 0.0081115    | 0.01           | 25.44        |
| 020030010040        | 6    | 0.0060836    | 0.01           | 25.45        |
| 020030020           | 221  | 0.2240788    | 0.22           | 25.67        |
| 020030020010        | 19   | 0.0192647    | 0.02           | 25.69        |
| 020030020020        | 2    | 0.0020279    | 0.00           | 25.69        |
| 020030030           | 106  | 0.1074767    | 0.11           | 25.80        |
| 020030040           | 48   | 0.0486687    | 0.05           | 25.85        |
| 020030040010        | 6    | 0.0060836    | 0.01           | 25.85        |
| 020030040020        | 2    | 0.0020279    | 0.00           | 25.86        |
| 020030040030        | 5    | 0.0050697    | 0.01           | 25.86        |
| 020030040040        | 1    | 0.0010139    | 0.00           | 25.86        |
| 020030050           | 39   | 0.0395433    | 0.04           | 25.90        |
| 020030050010        | 9    | 0.0091254    | 0.01           | 25.91        |
| 020030050020        | 16   | 0.0162229    | 0.02           | 25.93        |
| 020030060           | 6    | 0.0060836    | 0.01           | 25.93        |
| 020030070           | 7    | 0.0070975    | 0.01           | 25.94        |
| 020030080           | 9    | 0.0091254    | 0.01           | 25.95        |
| 020030090           | 3    | 0.0030418    | 0.00           | 25.95        |
| 020030100           | 1    | 0.0010139    | 0.00           | 25.95        |
| 020030100010        | 1    | 0.0010139    | 0.00           | 25.95        |
| 020030100020        | 11   | 0.0111532    | 0.01           | 25.97        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 020030100030        | 1   | 0.0010139    | 0.00           | 25.97        |
| 020030110           | 3   | 0.0030418    | 0.00           | 25.97        |
| 020030120           | 2   | 0.0020279    | 0.00           | 25.97        |
| 020030130           | 7   | 0.0070975    | 0.01           | 25.98        |
| 020030140           | 2   | 0.0020279    | 0.00           | 25.98        |
| 020030150           | 8   | 0.0081115    | 0.01           | 25.99        |
| 020031              | 18  | 0.0182508    | 0.02           | 26.01        |
| 020040              | 735 | 0.7452396    | 0.75           | 26.75        |
| 020040010           | 56  | 0.0567802    | 0.06           | 26.81        |
| 020040020           | 57  | 0.0577941    | 0.06           | 26.87        |
| 020040030           | 35  | 0.0354876    | 0.04           | 26.90        |
| 020040030010        | 1   | 0.0010139    | 0.00           | 26.90        |
| 020040030020        | 6   | 0.0060836    | 0.01           | 26.91        |
| 020040040           | 22  | 0.0223065    | 0.02           | 26.93        |
| 020040050           | 5   | 0.0050697    | 0.01           | 26.94        |
| 020040060           | 8   | 0.0081115    | 0.01           | 26.95        |
| 020040070           | 1   | 0.0010139    | 0.00           | 26.95        |
| 020040160           | 1   | 0.0010139    | 0.00           | 26.95        |
| 020041              | 6   | 0.0060836    | 0.01           | 26.95        |
| 020050              | 418 | 0.4238233    | 0.42           | 27.38        |
| 020050010           | 38  | 0.0385294    | 0.04           | 27.42        |
| 020050010010        | 5   | 0.0050697    | 0.01           | 27.42        |
| 020050010020        | 4   | 0.0040557    | 0.00           | 27.42        |
| 020050010030        | 3   | 0.0030418    | 0.00           | 27.43        |
| 020050020           | 37  | 0.0375155    | 0.04           | 27.47        |
| 020050030           | 5   | 0.0050697    | 0.01           | 27.47        |
| 020050030010        | 7   | 0.0070975    | 0.01           | 27.48        |
| 020050030020        | 14  | 0.0141950    | 0.01           | 27.49        |
| 020050030030        | 6   | 0.0060836    | 0.01           | 27.50        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 020050030040        | 2   | 0.0020279    | 0.00           | 27.50        |
| 020050030050        | 12  | 0.0121672    | 0.01           | 27.51        |
| 020050040           | 7   | 0.0070975    | 0.01           | 27.52        |
| 020050050           | 1   | 0.0010139    | 0.00           | 27.52        |
| 020050180           | 1   | 0.0010139    | 0.00           | 27.52        |
| 020050190           | 1   | 0.0010139    | 0.00           | 27.52        |
| 020050200           | 1   | 0.0010139    | 0.00           | 27.52        |
| 020050210           | 1   | 0.0010139    | 0.00           | 27.52        |
| 020051              | 7   | 0.0070975    | 0.01           | 27.53        |
| 020051211           | 1   | 0.0010139    | 0.00           | 27.53        |
| 020051212           | 1   | 0.0010139    | 0.00           | 27.53        |
| 020051213           | 1   | 0.0010139    | 0.00           | 27.53        |
| 020052              | 1   | 0.0010139    | 0.00           | 27.54        |
| 020053              | 1   | 0.0010139    | 0.00           | 27.54        |
| 020060              | 295 | 0.2991098    | 0.30           | 27.84        |
| 020060010           | 31  | 0.0314319    | 0.03           | 27.87        |
| 020060020           | 17  | 0.0172368    | 0.02           | 27.88        |
| 020060020010        | 11  | 0.0111532    | 0.01           | 27.90        |
| 020060020020010     | 2   | 0.0020279    | 0.00           | 27.90        |
| 020060020020020     | 2   | 0.0020279    | 0.00           | 27.90        |
| 020060020020030     | 1   | 0.0010139    | 0.00           | 27.90        |
| 020060020020040     | 1   | 0.0010139    | 0.00           | 27.90        |
| 020060020020050     | 1   | 0.0010139    | 0.00           | 27.90        |
| 020060030           | 10  | 0.0101393    | 0.01           | 27.91        |
| 020060030010        | 2   | 0.0020279    | 0.00           | 27.91        |
| 020060030020        | 2   | 0.0020279    | 0.00           | 27.92        |
| 020060030030        | 12  | 0.0121672    | 0.01           | 27.93        |
| 020060030040        | 6   | 0.0060836    | 0.01           | 27.93        |
| 020060030050        | 1   | 0.0010139    | 0.00           | 27.94        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 020060040           | 9   | 0.0091254    | 0.01           | 27.94        |
| 020060050           | 4   | 0.0040557    | 0.00           | 27.95        |
| 020060050010        | 2   | 0.0020279    | 0.00           | 27.95        |
| 020060050020        | 14  | 0.0141950    | 0.01           | 27.97        |
| 020060060           | 6   | 0.0060836    | 0.01           | 27.97        |
| 020060070           | 9   | 0.0091254    | 0.01           | 27.98        |
| 020060080010        | 5   | 0.0050697    | 0.01           | 27.99        |
| 020060080020        | 14  | 0.0141950    | 0.01           | 28.00        |
| 020060090           | 11  | 0.0111532    | 0.01           | 28.01        |
| 020060100           | 1   | 0.0010139    | 0.00           | 28.01        |
| 020060110           | 3   | 0.0030418    | 0.00           | 28.01        |
| 020060220           | 1   | 0.0010139    | 0.00           | 28.02        |
| 020060230           | 2   | 0.0020279    | 0.00           | 28.02        |
| 020061              | 3   | 0.0030418    | 0.00           | 28.02        |
| 020070              | 182 | 0.1845355    | 0.18           | 28.21        |
| 020070010           | 17  | 0.0172368    | 0.02           | 28.22        |
| 020070020           | 5   | 0.0050697    | 0.01           | 28.23        |
| 020070020010        | 10  | 0.0101393    | 0.01           | 28.24        |
| 020070020020        | 23  | 0.0233204    | 0.02           | 28.26        |
| 020070020020010     | 1   | 0.0010139    | 0.00           | 28.26        |
| 020070020020020     | 1   | 0.0010139    | 0.00           | 28.26        |
| 020070020030010     | 15  | 0.0152090    | 0.02           | 28.28        |
| 020070020030020010  | 3   | 0.0030418    | 0.00           | 28.28        |
| 020070020030020020  | 6   | 0.0060836    | 0.01           | 28.29        |
| 020070020030020030  | 1   | 0.0010139    | 0.00           | 28.29        |
| 020070020030020040  | 2   | 0.0020279    | 0.00           | 28.29        |
| 020070020030020050  | 1   | 0.0010139    | 0.00           | 28.29        |
| 020070020030030     | 1   | 0.0010139    | 0.00           | 28.29        |
| 020070020040        | 1   | 0.0010139    | 0.00           | 28.29        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 020070020050        | 2   | 0.0020279    | 0.00           | 28.30        |
| 020070020060        | 5   | 0.0050697    | 0.01           | 28.30        |
| 020070020069        | 1   | 0.0010139    | 0.00           | 28.30        |
| 020070020070        | 3   | 0.0030418    | 0.00           | 28.30        |
| 020070020080        | 1   | 0.0010139    | 0.00           | 28.31        |
| 020070020090        | 1   | 0.0010139    | 0.00           | 28.31        |
| 020070030           | 4   | 0.0040557    | 0.00           | 28.31        |
| 020070030010        | 10  | 0.0101393    | 0.01           | 28.32        |
| 020070030020        | 6   | 0.0060836    | 0.01           | 28.33        |
| 020070030030        | 14  | 0.0141950    | 0.01           | 28.34        |
| 020070030040        | 3   | 0.0030418    | 0.00           | 28.34        |
| 020070030050        | 1   | 0.0010139    | 0.00           | 28.35        |
| 020070040           | 5   | 0.0050697    | 0.01           | 28.35        |
| 020070240           | 1   | 0.0010139    | 0.00           | 28.35        |
| 020080              | 137 | 0.1389086    | 0.14           | 28.49        |
| 020080010           | 5   | 0.0050697    | 0.01           | 28.50        |
| 020080010010        | 23  | 0.0233204    | 0.02           | 28.52        |
| 020080010020010     | 2   | 0.0020279    | 0.00           | 28.52        |
| 020080010020020     | 7   | 0.0070975    | 0.01           | 28.53        |
| 020080010020030     | 11  | 0.0111532    | 0.01           | 28.54        |
| 020080010030        | 6   | 0.0060836    | 0.01           | 28.55        |
| 020080010040        | 1   | 0.0010139    | 0.00           | 28.55        |
| 020080020           | 18  | 0.0182508    | 0.02           | 28.56        |
| 020080030           | 3   | 0.0030418    | 0.00           | 28.57        |
| 020080030010010     | 3   | 0.0030418    | 0.00           | 28.57        |
| 020080030010020     | 24  | 0.0243344    | 0.02           | 28.59        |
| 020080030020        | 3   | 0.0030418    | 0.00           | 28.60        |
| 020080030030        | 1   | 0.0010139    | 0.00           | 28.60        |
| 020080030040        | 1   | 0.0010139    | 0.00           | 28.60        |



(continued)

| gliederungskennzahl | N  | exactpercent | roundedpercent | cumulpercent |
|---------------------|----|--------------|----------------|--------------|
| 020080030050        | 2  | 0.0020279    | 0.00           | 28.60        |
| 020080030060        | 2  | 0.0020279    | 0.00           | 28.60        |
| 020080040           | 20 | 0.0202786    | 0.02           | 28.62        |
| 020080050           | 1  | 0.0010139    | 0.00           | 28.63        |
| 020080050010        | 19 | 0.0192647    | 0.02           | 28.64        |
| 020080050020010     | 7  | 0.0070975    | 0.01           | 28.65        |
| 020080050020011     | 6  | 0.0060836    | 0.01           | 28.66        |
| 020080050020020010  | 4  | 0.0040557    | 0.00           | 28.66        |
| 020080050020020011  | 4  | 0.0040557    | 0.00           | 28.67        |
| 020080050020020020  | 16 | 0.0162229    | 0.02           | 28.68        |
| 020080050020030     | 5  | 0.0050697    | 0.01           | 28.69        |
| 020080050020040     | 14 | 0.0141950    | 0.01           | 28.70        |
| 020080050020050010  | 5  | 0.0050697    | 0.01           | 28.71        |
| 020080050020050020  | 9  | 0.0091254    | 0.01           | 28.72        |
| 020080050020050030  | 2  | 0.0020279    | 0.00           | 28.72        |
| 020080050020050040  | 3  | 0.0030418    | 0.00           | 28.72        |
| 020080050020060     | 2  | 0.0020279    | 0.00           | 28.72        |
| 020080050030        | 5  | 0.0050697    | 0.01           | 28.73        |
| 020080050040        | 8  | 0.0081115    | 0.01           | 28.74        |
| 020080050050        | 31 | 0.0314319    | 0.03           | 28.77        |
| 020080060           | 10 | 0.0101393    | 0.01           | 28.78        |
| 020080070           | 7  | 0.0070975    | 0.01           | 28.78        |
| 020080080           | 3  | 0.0030418    | 0.00           | 28.79        |
| 020080080010        | 24 | 0.0243344    | 0.02           | 28.81        |
| 020080080020        | 8  | 0.0081115    | 0.01           | 28.82        |
| 020080090010010     | 25 | 0.0253483    | 0.03           | 28.85        |
| 020080090010020     | 8  | 0.0081115    | 0.01           | 28.85        |
| 020080090010030     | 6  | 0.0060836    | 0.01           | 28.86        |
| 020080090010040     | 1  | 0.0010139    | 0.00           | 28.86        |

(continued)

| gliederungskennzahl | N  | exactpercent | roundedpercent | cumulpercent |
|---------------------|----|--------------|----------------|--------------|
| 020080090020        | 5  | 0.0050697    | 0.01           | 28.87        |
| 020080090030        | 2  | 0.0020279    | 0.00           | 28.87        |
| 020080090040        | 25 | 0.0253483    | 0.03           | 28.89        |
| 020080100010        | 4  | 0.0040557    | 0.00           | 28.90        |
| 020080100020        | 5  | 0.0050697    | 0.01           | 28.90        |
| 020080100030        | 1  | 0.0010139    | 0.00           | 28.90        |
| 020080100040        | 4  | 0.0040557    | 0.00           | 28.91        |
| 020080110           | 6  | 0.0060836    | 0.01           | 28.91        |
| 020080120010        | 13 | 0.0131811    | 0.01           | 28.93        |
| 020080120020        | 3  | 0.0030418    | 0.00           | 28.93        |
| 020080120030010     | 3  | 0.0030418    | 0.00           | 28.93        |
| 020080120030020     | 4  | 0.0040557    | 0.00           | 28.94        |
| 020080120030030010  | 4  | 0.0040557    | 0.00           | 28.94        |
| 020080120030030020  | 7  | 0.0070975    | 0.01           | 28.95        |
| 020080120030030030  | 10 | 0.0101393    | 0.01           | 28.96        |
| 020080130           | 11 | 0.0111532    | 0.01           | 28.97        |
| 020080140           | 13 | 0.0131811    | 0.01           | 28.98        |
| 020080150           | 5  | 0.0050697    | 0.01           | 28.99        |
| 020080160           | 36 | 0.0365015    | 0.04           | 29.02        |
| 020080170           | 18 | 0.0182508    | 0.02           | 29.04        |
| 020080180           | 3  | 0.0030418    | 0.00           | 29.05        |
| 020080190           | 3  | 0.0030418    | 0.00           | 29.05        |
| 020080200           | 14 | 0.0141950    | 0.01           | 29.06        |
| 020080210           | 1  | 0.0010139    | 0.00           | 29.06        |
| 020080220           | 3  | 0.0030418    | 0.00           | 29.07        |
| 020080230           | 10 | 0.0101393    | 0.01           | 29.08        |
| 020080240           | 16 | 0.0162229    | 0.02           | 29.09        |
| 020080250           | 4  | 0.0040557    | 0.00           | 29.10        |
| 020080260           | 11 | 0.0111532    | 0.01           | 29.11        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 020080270           | 32  | 0.0324458    | 0.03           | 29.14        |
| 020090              | 107 | 0.1084907    | 0.11           | 29.25        |
| 020090010           | 15  | 0.0152090    | 0.02           | 29.26        |
| 020090020           | 3   | 0.0030418    | 0.00           | 29.27        |
| 020090030           | 12  | 0.0121672    | 0.01           | 29.28        |
| 020100              | 81  | 0.0821284    | 0.08           | 29.36        |
| 020100010           | 8   | 0.0081115    | 0.01           | 29.37        |
| 020100020           | 7   | 0.0070975    | 0.01           | 29.38        |
| 020100030           | 3   | 0.0030418    | 0.00           | 29.38        |
| 020100040           | 2   | 0.0020279    | 0.00           | 29.38        |
| 020110              | 106 | 0.1074767    | 0.11           | 29.49        |
| 020110010           | 1   | 0.0010139    | 0.00           | 29.49        |
| 020110020           | 2   | 0.0020279    | 0.00           | 29.49        |
| 020110030           | 2   | 0.0020279    | 0.00           | 29.49        |
| 020110040           | 2   | 0.0020279    | 0.00           | 29.50        |
| 020110050           | 3   | 0.0030418    | 0.00           | 29.50        |
| 020120              | 51  | 0.0517105    | 0.05           | 29.55        |
| 020120010           | 2   | 0.0020279    | 0.00           | 29.55        |
| 020120020           | 4   | 0.0040557    | 0.00           | 29.56        |
| 020120030           | 3   | 0.0030418    | 0.00           | 29.56        |
| 020130              | 70  | 0.0709752    | 0.07           | 29.63        |
| 020130010           | 1   | 0.0010139    | 0.00           | 29.63        |
| 020130020           | 1   | 0.0010139    | 0.00           | 29.63        |
| 020130030           | 7   | 0.0070975    | 0.01           | 29.64        |
| 020130040           | 4   | 0.0040557    | 0.00           | 29.64        |
| 020140              | 95  | 0.0963235    | 0.10           | 29.74        |
| 020150              | 43  | 0.0435991    | 0.04           | 29.78        |
| 020160              | 53  | 0.0537384    | 0.05           | 29.84        |
| 020170              | 39  | 0.0395433    | 0.04           | 29.88        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 020180              | 48   | 0.0486687    | 0.05           | 29.93        |
| 020190              | 40   | 0.0405573    | 0.04           | 29.97        |
| 020200              | 37   | 0.0375155    | 0.04           | 30.00        |
| 020210              | 37   | 0.0375155    | 0.04           | 30.04        |
| 020220              | 42   | 0.0425851    | 0.04           | 30.08        |
| 020230              | 46   | 0.0466408    | 0.05           | 30.13        |
| 020240              | 34   | 0.0344737    | 0.03           | 30.17        |
| 020250              | 43   | 0.0435991    | 0.04           | 30.21        |
| 020260              | 35   | 0.0354876    | 0.04           | 30.24        |
| 020270              | 36   | 0.0365015    | 0.04           | 30.28        |
| 020280              | 37   | 0.0375155    | 0.04           | 30.32        |
| 020290              | 13   | 0.0131811    | 0.01           | 30.33        |
| 020300              | 25   | 0.0253483    | 0.03           | 30.36        |
| 021                 | 24   | 0.0243344    | 0.02           | 30.38        |
| 021010              | 5    | 0.0050697    | 0.01           | 30.39        |
| 021020              | 2    | 0.0020279    | 0.00           | 30.39        |
| 021030              | 4    | 0.0040557    | 0.00           | 30.39        |
| 030                 | 5291 | 5.3647111    | 5.36           | 35.76        |
| 030000550           | 1    | 0.0010139    | 0.00           | 35.76        |
| 030000560           | 1    | 0.0010139    | 0.00           | 35.76        |
| 030000590           | 1    | 0.0010139    | 0.00           | 35.76        |
| 030000600           | 1    | 0.0010139    | 0.00           | 35.76        |
| 030000610           | 1    | 0.0010139    | 0.00           | 35.76        |
| 030000620           | 1    | 0.0010139    | 0.00           | 35.76        |
| 030000630           | 1    | 0.0010139    | 0.00           | 35.76        |
| 030000640           | 1    | 0.0010139    | 0.00           | 35.77        |
| 030000660           | 1    | 0.0010139    | 0.00           | 35.77        |
| 030000670           | 1    | 0.0010139    | 0.00           | 35.77        |
| 030000680           | 1    | 0.0010139    | 0.00           | 35.77        |

(continued)

| gliederungskennzahl | N | exactpercent | roundedpercent | cumulpercent |
|---------------------|---|--------------|----------------|--------------|
| 030000690           | 1 | 0.0010139    | 0.00           | 35.77        |
| 030000730           | 1 | 0.0010139    | 0.00           | 35.77        |
| 030000740           | 1 | 0.0010139    | 0.00           | 35.77        |
| 030000760           | 1 | 0.0010139    | 0.00           | 35.77        |
| 030000770           | 1 | 0.0010139    | 0.00           | 35.77        |
| 030000780           | 1 | 0.0010139    | 0.00           | 35.77        |
| 030000790           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000800           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000810           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000820           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000830           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000840           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000850           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000860           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000890           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000900           | 1 | 0.0010139    | 0.00           | 35.78        |
| 030000910           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030000930           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030000940           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030000960           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030000970           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030000980           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030000990           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030001000           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030001010           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030001020           | 1 | 0.0010139    | 0.00           | 35.79        |
| 030001040           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001050           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001060           | 1 | 0.0010139    | 0.00           | 35.80        |

(continued)

| gliederungskennzahl | N | exactpercent | roundedpercent | cumulpercent |
|---------------------|---|--------------|----------------|--------------|
| 030001070           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001080           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001090           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001100           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001110           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001120           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001130           | 1 | 0.0010139    | 0.00           | 35.80        |
| 030001140           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001150           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001160           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001170           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001180           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001190           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001200           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001210           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001220           | 1 | 0.0010139    | 0.00           | 35.81        |
| 030001230           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001240           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001250           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001260           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001270           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001280           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001290           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001300           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001310           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001320           | 1 | 0.0010139    | 0.00           | 35.82        |
| 030001330           | 1 | 0.0010139    | 0.00           | 35.83        |
| 030001370           | 1 | 0.0010139    | 0.00           | 35.83        |
| 030001380           | 1 | 0.0010139    | 0.00           | 35.83        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 030001390           | 1    | 0.0010139    | 0.00           | 35.83        |
| 030001400           | 1    | 0.0010139    | 0.00           | 35.83        |
| 030001430           | 1    | 0.0010139    | 0.00           | 35.83        |
| 030001440           | 1    | 0.0010139    | 0.00           | 35.83        |
| 030001470           | 1    | 0.0010139    | 0.00           | 35.83        |
| 030001520           | 1    | 0.0010139    | 0.00           | 35.83        |
| 030010              | 1078 | 1.0930181    | 1.09           | 36.93        |
| 030010010           | 90   | 0.0912538    | 0.09           | 37.02        |
| 030010020           | 113  | 0.1145743    | 0.11           | 37.13        |
| 030010020010        | 4    | 0.0040557    | 0.00           | 37.14        |
| 030010020020        | 6    | 0.0060836    | 0.01           | 37.14        |
| 030010020030        | 6    | 0.0060836    | 0.01           | 37.15        |
| 030010030           | 67   | 0.0679334    | 0.07           | 37.22        |
| 030010030010        | 15   | 0.0152090    | 0.02           | 37.23        |
| 030010030020        | 8    | 0.0081115    | 0.01           | 37.24        |
| 030010030030        | 4    | 0.0040557    | 0.00           | 37.24        |
| 030010030040        | 6    | 0.0060836    | 0.01           | 37.25        |
| 030010030050        | 1    | 0.0010139    | 0.00           | 37.25        |
| 030010031           | 5    | 0.0050697    | 0.01           | 37.26        |
| 030010040           | 35   | 0.0354876    | 0.04           | 37.29        |
| 030010050           | 19   | 0.0192647    | 0.02           | 37.31        |
| 030010060           | 11   | 0.0111532    | 0.01           | 37.32        |
| 030010070           | 3    | 0.0030418    | 0.00           | 37.32        |
| 030010080           | 3    | 0.0030418    | 0.00           | 37.33        |
| 030010150           | 3    | 0.0030418    | 0.00           | 37.33        |
| 030020              | 1021 | 1.0352240    | 1.04           | 38.37        |
| 030020010           | 75   | 0.0760449    | 0.08           | 38.44        |
| 030020010010        | 11   | 0.0111532    | 0.01           | 38.45        |
| 030020010020        | 13   | 0.0131811    | 0.01           | 38.47        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 030020010030        | 6   | 0.0060836    | 0.01           | 38.47        |
| 030020010050        | 4   | 0.0040557    | 0.00           | 38.48        |
| 030020010060        | 7   | 0.0070975    | 0.01           | 38.48        |
| 030020020           | 103 | 0.1044349    | 0.10           | 38.59        |
| 030020020010        | 4   | 0.0040557    | 0.00           | 38.59        |
| 030020020020        | 2   | 0.0020279    | 0.00           | 38.59        |
| 030020020030        | 3   | 0.0030418    | 0.00           | 38.60        |
| 030020020040        | 7   | 0.0070975    | 0.01           | 38.60        |
| 030020020050        | 3   | 0.0030418    | 0.00           | 38.61        |
| 030020020060        | 1   | 0.0010139    | 0.00           | 38.61        |
| 030020020070        | 2   | 0.0020279    | 0.00           | 38.61        |
| 030020020080        | 2   | 0.0020279    | 0.00           | 38.61        |
| 030020020090        | 5   | 0.0050697    | 0.01           | 38.62        |
| 030020020100        | 1   | 0.0010139    | 0.00           | 38.62        |
| 030020030           | 71  | 0.0719891    | 0.07           | 38.69        |
| 030020040           | 29  | 0.0294040    | 0.03           | 38.72        |
| 030020050           | 11  | 0.0111532    | 0.01           | 38.73        |
| 030020060           | 8   | 0.0081115    | 0.01           | 38.74        |
| 030020060010        | 5   | 0.0050697    | 0.01           | 38.74        |
| 030020060020        | 4   | 0.0040557    | 0.00           | 38.75        |
| 030020070010        | 1   | 0.0010139    | 0.00           | 38.75        |
| 030020070020        | 5   | 0.0050697    | 0.01           | 38.75        |
| 030020070030        | 3   | 0.0030418    | 0.00           | 38.76        |
| 030020080           | 7   | 0.0070975    | 0.01           | 38.76        |
| 030020090           | 6   | 0.0060836    | 0.01           | 38.77        |
| 030020170           | 2   | 0.0020279    | 0.00           | 38.77        |
| 030020180           | 2   | 0.0020279    | 0.00           | 38.77        |
| 030020190           | 5   | 0.0050697    | 0.01           | 38.78        |
| 030021              | 2   | 0.0020279    | 0.00           | 38.78        |



(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 030030              | 580 | 0.5880802    | 0.59           | 39.37        |
| 030030010           | 86  | 0.0871981    | 0.09           | 39.46        |
| 030030020           | 71  | 0.0719891    | 0.07           | 39.53        |
| 030030030           | 37  | 0.0375155    | 0.04           | 39.57        |
| 030030030010        | 10  | 0.0101393    | 0.01           | 39.58        |
| 030030030020        | 9   | 0.0091254    | 0.01           | 39.59        |
| 030030030030        | 7   | 0.0070975    | 0.01           | 39.59        |
| 030030030040        | 5   | 0.0050697    | 0.01           | 39.60        |
| 030030030050        | 7   | 0.0070975    | 0.01           | 39.61        |
| 030030030060        | 20  | 0.0202786    | 0.02           | 39.63        |
| 030030040           | 40  | 0.0405573    | 0.04           | 39.67        |
| 030030050           | 9   | 0.0091254    | 0.01           | 39.68        |
| 030030270           | 2   | 0.0020279    | 0.00           | 39.68        |
| 030030280           | 2   | 0.0020279    | 0.00           | 39.68        |
| 030030290           | 5   | 0.0050697    | 0.01           | 39.68        |
| 030030330           | 4   | 0.0040557    | 0.00           | 39.69        |
| 030031              | 5   | 0.0050697    | 0.01           | 39.69        |
| 030040              | 357 | 0.3619735    | 0.36           | 40.06        |
| 030040010           | 34  | 0.0344737    | 0.03           | 40.09        |
| 030040010010        | 1   | 0.0010139    | 0.00           | 40.09        |
| 030040010020        | 4   | 0.0040557    | 0.00           | 40.09        |
| 030040010030        | 3   | 0.0030418    | 0.00           | 40.10        |
| 030040010040        | 1   | 0.0010139    | 0.00           | 40.10        |
| 030040010050        | 2   | 0.0020279    | 0.00           | 40.10        |
| 030040010060        | 1   | 0.0010139    | 0.00           | 40.10        |
| 030040010070        | 1   | 0.0010139    | 0.00           | 40.10        |
| 030040010080        | 3   | 0.0030418    | 0.00           | 40.11        |
| 030040020           | 26  | 0.0263622    | 0.03           | 40.13        |
| 030040020010        | 44  | 0.0446130    | 0.04           | 40.18        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 030040020020        | 18  | 0.0182508    | 0.02           | 40.20        |
| 030040020030        | 8   | 0.0081115    | 0.01           | 40.20        |
| 030040020040        | 4   | 0.0040557    | 0.00           | 40.21        |
| 030040020050        | 2   | 0.0020279    | 0.00           | 40.21        |
| 030040020060        | 1   | 0.0010139    | 0.00           | 40.21        |
| 030040020070        | 1   | 0.0010139    | 0.00           | 40.21        |
| 030040020080        | 4   | 0.0040557    | 0.00           | 40.22        |
| 030040030           | 6   | 0.0060836    | 0.01           | 40.22        |
| 030040030010        | 2   | 0.0020279    | 0.00           | 40.22        |
| 030040030020        | 5   | 0.0050697    | 0.01           | 40.23        |
| 030040030030        | 2   | 0.0020279    | 0.00           | 40.23        |
| 030040040           | 12  | 0.0121672    | 0.01           | 40.24        |
| 030040100           | 2   | 0.0020279    | 0.00           | 40.24        |
| 030041              | 2   | 0.0020279    | 0.00           | 40.25        |
| 030050              | 186 | 0.1885912    | 0.19           | 40.44        |
| 030050010           | 59  | 0.0598220    | 0.06           | 40.50        |
| 030050020           | 44  | 0.0446130    | 0.04           | 40.54        |
| 030050030           | 13  | 0.0131811    | 0.01           | 40.55        |
| 030050040           | 11  | 0.0111532    | 0.01           | 40.56        |
| 030050050           | 3   | 0.0030418    | 0.00           | 40.57        |
| 030060              | 137 | 0.1389086    | 0.14           | 40.71        |
| 030060010010        | 1   | 0.0010139    | 0.00           | 40.71        |
| 030060010020        | 3   | 0.0030418    | 0.00           | 40.71        |
| 030060010030        | 4   | 0.0040557    | 0.00           | 40.71        |
| 030060010040        | 1   | 0.0010139    | 0.00           | 40.72        |
| 030060010050        | 1   | 0.0010139    | 0.00           | 40.72        |
| 030060010060        | 1   | 0.0010139    | 0.00           | 40.72        |
| 030060020           | 3   | 0.0030418    | 0.00           | 40.72        |
| 030070              | 75  | 0.0760449    | 0.08           | 40.80        |

(continued)

| gliederungskennzahl | N  | exactpercent | roundedpercent | cumulpercent |
|---------------------|----|--------------|----------------|--------------|
| 030070010           | 83 | 0.0841563    | 0.08           | 40.88        |
| 030070010010        | 4  | 0.0040557    | 0.00           | 40.88        |
| 030070010020        | 3  | 0.0030418    | 0.00           | 40.89        |
| 030070020           | 11 | 0.0111532    | 0.01           | 40.90        |
| 030070020010        | 8  | 0.0081115    | 0.01           | 40.91        |
| 030070020020        | 5  | 0.0050697    | 0.01           | 40.91        |
| 030070030           | 3  | 0.0030418    | 0.00           | 40.92        |
| 030070030010        | 2  | 0.0020279    | 0.00           | 40.92        |
| 030070030020        | 8  | 0.0081115    | 0.01           | 40.93        |
| 030070030030        | 2  | 0.0020279    | 0.00           | 40.93        |
| 030070030040        | 1  | 0.0010139    | 0.00           | 40.93        |
| 030070040           | 6  | 0.0060836    | 0.01           | 40.93        |
| 030080              | 38 | 0.0385294    | 0.04           | 40.97        |
| 030080010           | 56 | 0.0567802    | 0.06           | 41.03        |
| 030080020           | 24 | 0.0243344    | 0.02           | 41.05        |
| 030081              | 2  | 0.0020279    | 0.00           | 41.06        |
| 030090              | 57 | 0.0577941    | 0.06           | 41.11        |
| 030100              | 46 | 0.0466408    | 0.05           | 41.16        |
| 030100010           | 6  | 0.0060836    | 0.01           | 41.17        |
| 030100020           | 4  | 0.0040557    | 0.00           | 41.17        |
| 030100030           | 4  | 0.0040557    | 0.00           | 41.17        |
| 030110              | 29 | 0.0294040    | 0.03           | 41.20        |
| 030120              | 36 | 0.0365015    | 0.04           | 41.24        |
| 030130              | 31 | 0.0314319    | 0.03           | 41.27        |
| 030130010           | 13 | 0.0131811    | 0.01           | 41.29        |
| 030130020           | 12 | 0.0121672    | 0.01           | 41.30        |
| 030130030           | 2  | 0.0020279    | 0.00           | 41.30        |
| 030130040           | 2  | 0.0020279    | 0.00           | 41.30        |
| 030140              | 26 | 0.0263622    | 0.03           | 41.33        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 030150              | 48   | 0.0486687    | 0.05           | 41.38        |
| 030160              | 1    | 0.0010139    | 0.00           | 41.38        |
| 030170              | 4    | 0.0040557    | 0.00           | 41.38        |
| 030220              | 1    | 0.0010139    | 0.00           | 41.38        |
| 030260              | 1    | 0.0010139    | 0.00           | 41.38        |
| 030280              | 14   | 0.0141950    | 0.01           | 41.40        |
| 030970              | 55   | 0.0557662    | 0.06           | 41.45        |
| 030971              | 5    | 0.0050697    | 0.01           | 41.46        |
| 030980              | 1    | 0.0010139    | 0.00           | 41.46        |
| 030990              | 1    | 0.0010139    | 0.00           | 41.46        |
| 030991              | 1    | 0.0010139    | 0.00           | 41.46        |
| 030992              | 1    | 0.0010139    | 0.00           | 41.46        |
| 030993              | 1    | 0.0010139    | 0.00           | 41.46        |
| 031                 | 10   | 0.0101393    | 0.01           | 41.47        |
| 040                 | 4155 | 4.2128850    | 4.21           | 45.69        |
| 040001570           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001630           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001640           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001650           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001660           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001670           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001680           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001690           | 1    | 0.0010139    | 0.00           | 45.69        |
| 040001700           | 1    | 0.0010139    | 0.00           | 45.70        |
| 040001710           | 1    | 0.0010139    | 0.00           | 45.70        |
| 040001720           | 1    | 0.0010139    | 0.00           | 45.70        |
| 040001730           | 1    | 0.0010139    | 0.00           | 45.70        |
| 040001740           | 1    | 0.0010139    | 0.00           | 45.70        |
| 040001750           | 1    | 0.0010139    | 0.00           | 45.70        |

*(continued)*

| gliederungskennzahl | N | exactpercent | roundedpercent | cumulpercent |
|---------------------|---|--------------|----------------|--------------|
| 040001760           | 1 | 0.0010139    | 0.00           | 45.70        |
| 040001770           | 1 | 0.0010139    | 0.00           | 45.70        |
| 040001780           | 1 | 0.0010139    | 0.00           | 45.70        |
| 040001790           | 1 | 0.0010139    | 0.00           | 45.70        |
| 040001800           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001810           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001820           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001830           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001840           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001850           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001860           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001870           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001880           | 1 | 0.0010139    | 0.00           | 45.71        |
| 040001890           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001900           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001910           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001920           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001930           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001940           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001950           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001960           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001970           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001980           | 1 | 0.0010139    | 0.00           | 45.72        |
| 040001990           | 1 | 0.0010139    | 0.00           | 45.73        |
| 040002000           | 1 | 0.0010139    | 0.00           | 45.73        |
| 040002010           | 1 | 0.0010139    | 0.00           | 45.73        |
| 040002020           | 1 | 0.0010139    | 0.00           | 45.73        |
| 040002030           | 1 | 0.0010139    | 0.00           | 45.73        |
| 040002070           | 1 | 0.0010139    | 0.00           | 45.73        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 040002080           | 1   | 0.0010139    | 0.00           | 45.73        |
| 040002090           | 1   | 0.0010139    | 0.00           | 45.73        |
| 040002100           | 1   | 0.0010139    | 0.00           | 45.73        |
| 040002120           | 1   | 0.0010139    | 0.00           | 45.73        |
| 040002130           | 1   | 0.0010139    | 0.00           | 45.74        |
| 040002140           | 1   | 0.0010139    | 0.00           | 45.74        |
| 040002150           | 1   | 0.0010139    | 0.00           | 45.74        |
| 040002170           | 1   | 0.0010139    | 0.00           | 45.74        |
| 040002180           | 1   | 0.0010139    | 0.00           | 45.74        |
| 040010              | 465 | 0.4714781    | 0.47           | 46.21        |
| 040010010           | 54  | 0.0547523    | 0.05           | 46.27        |
| 040010020           | 74  | 0.0750309    | 0.08           | 46.34        |
| 040010020010        | 7   | 0.0070975    | 0.01           | 46.35        |
| 040010020020        | 20  | 0.0202786    | 0.02           | 46.37        |
| 040010020030        | 1   | 0.0010139    | 0.00           | 46.37        |
| 040010020040        | 3   | 0.0030418    | 0.00           | 46.37        |
| 040010030           | 28  | 0.0283901    | 0.03           | 46.40        |
| 040010030010        | 8   | 0.0081115    | 0.01           | 46.41        |
| 040010030020        | 4   | 0.0040557    | 0.00           | 46.41        |
| 040010030030        | 7   | 0.0070975    | 0.01           | 46.42        |
| 040010040           | 16  | 0.0162229    | 0.02           | 46.44        |
| 040010050           | 25  | 0.0253483    | 0.03           | 46.46        |
| 040010060           | 1   | 0.0010139    | 0.00           | 46.46        |
| 040010060010        | 28  | 0.0283901    | 0.03           | 46.49        |
| 040010060020010     | 6   | 0.0060836    | 0.01           | 46.50        |
| 040010060020020     | 1   | 0.0010139    | 0.00           | 46.50        |
| 040010060020030010  | 7   | 0.0070975    | 0.01           | 46.50        |
| 040010060020030020  | 28  | 0.0283901    | 0.03           | 46.53        |
| 040010060020030030  | 20  | 0.0202786    | 0.02           | 46.55        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 040010060020030040  | 12  | 0.0121672    | 0.01           | 46.57        |
| 040010060020030050  | 36  | 0.0365015    | 0.04           | 46.60        |
| 040010060020040     | 1   | 0.0010139    | 0.00           | 46.60        |
| 040010060030        | 6   | 0.0060836    | 0.01           | 46.61        |
| 040010070           | 1   | 0.0010139    | 0.00           | 46.61        |
| 040010070010        | 5   | 0.0050697    | 0.01           | 46.62        |
| 040010070011        | 2   | 0.0020279    | 0.00           | 46.62        |
| 040010070020010     | 1   | 0.0010139    | 0.00           | 46.62        |
| 040010070020020     | 13  | 0.0131811    | 0.01           | 46.63        |
| 040010070020030     | 4   | 0.0040557    | 0.00           | 46.64        |
| 040010070020040     | 4   | 0.0040557    | 0.00           | 46.64        |
| 040010070020050     | 3   | 0.0030418    | 0.00           | 46.64        |
| 040010070030        | 1   | 0.0010139    | 0.00           | 46.64        |
| 040010080           | 3   | 0.0030418    | 0.00           | 46.65        |
| 040020              | 397 | 0.4025308    | 0.40           | 47.05        |
| 040020010           | 135 | 0.1368807    | 0.14           | 47.19        |
| 040020010010        | 5   | 0.0050697    | 0.01           | 47.19        |
| 040020010020        | 13  | 0.0131811    | 0.01           | 47.20        |
| 040020010030        | 8   | 0.0081115    | 0.01           | 47.21        |
| 040020010040        | 11  | 0.0111532    | 0.01           | 47.22        |
| 040020010050        | 6   | 0.0060836    | 0.01           | 47.23        |
| 040020010060        | 7   | 0.0070975    | 0.01           | 47.24        |
| 040020010070        | 4   | 0.0040557    | 0.00           | 47.24        |
| 040020010080        | 1   | 0.0010139    | 0.00           | 47.24        |
| 040020020           | 108 | 0.1095046    | 0.11           | 47.35        |
| 040020020010        | 29  | 0.0294040    | 0.03           | 47.38        |
| 040020020020        | 8   | 0.0081115    | 0.01           | 47.39        |
| 040020020030        | 12  | 0.0121672    | 0.01           | 47.40        |
| 040020020040        | 15  | 0.0152090    | 0.02           | 47.42        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 040020020050        | 7   | 0.0070975    | 0.01           | 47.42        |
| 040020030           | 61  | 0.0618498    | 0.06           | 47.49        |
| 040020030010        | 28  | 0.0283901    | 0.03           | 47.51        |
| 040020030020        | 5   | 0.0050697    | 0.01           | 47.52        |
| 040020030030        | 3   | 0.0030418    | 0.00           | 47.52        |
| 040020030040        | 1   | 0.0010139    | 0.00           | 47.52        |
| 040020040           | 37  | 0.0375155    | 0.04           | 47.56        |
| 040020050           | 81  | 0.0821284    | 0.08           | 47.64        |
| 040020060           | 16  | 0.0162229    | 0.02           | 47.66        |
| 040020070           | 11  | 0.0111532    | 0.01           | 47.67        |
| 040020070010        | 27  | 0.0273761    | 0.03           | 47.70        |
| 040020070020        | 6   | 0.0060836    | 0.01           | 47.70        |
| 040020080           | 7   | 0.0070975    | 0.01           | 47.71        |
| 040020090           | 5   | 0.0050697    | 0.01           | 47.72        |
| 040030              | 370 | 0.3751546    | 0.38           | 48.09        |
| 040030010           | 25  | 0.0253483    | 0.03           | 48.12        |
| 040030010010        | 38  | 0.0385294    | 0.04           | 48.15        |
| 040030010020        | 52  | 0.0527244    | 0.05           | 48.21        |
| 040030010030        | 21  | 0.0212926    | 0.02           | 48.23        |
| 040030010040        | 3   | 0.0030418    | 0.00           | 48.23        |
| 040030010050        | 7   | 0.0070975    | 0.01           | 48.24        |
| 040030010060        | 14  | 0.0141950    | 0.01           | 48.25        |
| 040030020           | 61  | 0.0618498    | 0.06           | 48.31        |
| 040030020010        | 9   | 0.0091254    | 0.01           | 48.32        |
| 040030020020        | 2   | 0.0020279    | 0.00           | 48.33        |
| 040030030           | 48  | 0.0486687    | 0.05           | 48.37        |
| 040030040           | 11  | 0.0111532    | 0.01           | 48.39        |
| 040030040010        | 1   | 0.0010139    | 0.00           | 48.39        |
| 040030040020        | 1   | 0.0010139    | 0.00           | 48.39        |



(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 040030050           | 2   | 0.0020279    | 0.00           | 48.39        |
| 040040              | 180 | 0.1825077    | 0.18           | 48.57        |
| 040040010           | 71  | 0.0719891    | 0.07           | 48.64        |
| 040040020           | 14  | 0.0141950    | 0.01           | 48.66        |
| 040040030           | 5   | 0.0050697    | 0.01           | 48.66        |
| 040050              | 65  | 0.0659055    | 0.07           | 48.73        |
| 040050010           | 5   | 0.0050697    | 0.01           | 48.73        |
| 040050020           | 3   | 0.0030418    | 0.00           | 48.74        |
| 040050030           | 7   | 0.0070975    | 0.01           | 48.74        |
| 040060              | 42  | 0.0425851    | 0.04           | 48.79        |
| 040060010           | 6   | 0.0060836    | 0.01           | 48.79        |
| 040060020           | 4   | 0.0040557    | 0.00           | 48.80        |
| 040060050           | 5   | 0.0050697    | 0.01           | 48.80        |
| 040070              | 10  | 0.0101393    | 0.01           | 48.81        |
| 040080              | 17  | 0.0172368    | 0.02           | 48.83        |
| 040090              | 33  | 0.0334597    | 0.03           | 48.86        |
| 040100              | 16  | 0.0162229    | 0.02           | 48.88        |
| 040110              | 5   | 0.0050697    | 0.01           | 48.88        |
| 040120              | 6   | 0.0060836    | 0.01           | 48.89        |
| 040130              | 5   | 0.0050697    | 0.01           | 48.90        |
| 040140              | 2   | 0.0020279    | 0.00           | 48.90        |
| 040160              | 14  | 0.0141950    | 0.01           | 48.91        |
| 040170              | 6   | 0.0060836    | 0.01           | 48.92        |
| 040190              | 1   | 0.0010139    | 0.00           | 48.92        |
| 040200              | 1   | 0.0010139    | 0.00           | 48.92        |
| 040390              | 1   | 0.0010139    | 0.00           | 48.92        |
| 040400              | 8   | 0.0081115    | 0.01           | 48.93        |
| 040410              | 1   | 0.0010139    | 0.00           | 48.93        |
| 040420              | 1   | 0.0010139    | 0.00           | 48.93        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 040430              | 1    | 0.0010139    | 0.00           | 48.93        |
| 040440              | 1    | 0.0010139    | 0.00           | 48.93        |
| 040450              | 1    | 0.0010139    | 0.00           | 48.93        |
| 040460              | 1    | 0.0010139    | 0.00           | 48.94        |
| 040470              | 1    | 0.0010139    | 0.00           | 48.94        |
| 040480              | 1    | 0.0010139    | 0.00           | 48.94        |
| 040500              | 1    | 0.0010139    | 0.00           | 48.94        |
| 040510              | 1    | 0.0010139    | 0.00           | 48.94        |
| 040520              | 1    | 0.0010139    | 0.00           | 48.94        |
| 041                 | 31   | 0.0314319    | 0.03           | 48.97        |
| 041010              | 9    | 0.0091254    | 0.01           | 48.98        |
| 041020              | 5    | 0.0050697    | 0.01           | 48.99        |
| 041030              | 2    | 0.0020279    | 0.00           | 48.99        |
| 041040              | 10   | 0.0101393    | 0.01           | 49.00        |
| 041451              | 3    | 0.0030418    | 0.00           | 49.00        |
| 041452              | 3    | 0.0030418    | 0.00           | 49.00        |
| 041454              | 14   | 0.0141950    | 0.01           | 49.02        |
| 042                 | 2    | 0.0020279    | 0.00           | 49.02        |
| 042010              | 1    | 0.0010139    | 0.00           | 49.02        |
| 042020              | 9    | 0.0091254    | 0.01           | 49.03        |
| 042030              | 1    | 0.0010139    | 0.00           | 49.03        |
| 042040              | 11   | 0.0111532    | 0.01           | 49.04        |
| 050                 | 3014 | 3.0559893    | 3.06           | 52.10        |
| 050002190           | 1    | 0.0010139    | 0.00           | 52.10        |
| 050002200           | 1    | 0.0010139    | 0.00           | 52.10        |
| 050002210           | 1    | 0.0010139    | 0.00           | 52.10        |
| 050002220           | 1    | 0.0010139    | 0.00           | 52.10        |
| 050002230           | 1    | 0.0010139    | 0.00           | 52.10        |
| 050002231           | 1    | 0.0010139    | 0.00           | 52.10        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 050002240           | 3   | 0.0030418    | 0.00           | 52.11        |
| 050002250           | 1   | 0.0010139    | 0.00           | 52.11        |
| 050002260           | 1   | 0.0010139    | 0.00           | 52.11        |
| 050002270           | 1   | 0.0010139    | 0.00           | 52.11        |
| 050002280           | 1   | 0.0010139    | 0.00           | 52.11        |
| 050002290           | 53  | 0.0537384    | 0.05           | 52.17        |
| 050010              | 338 | 0.3427088    | 0.34           | 52.51        |
| 050010010           | 34  | 0.0344737    | 0.03           | 52.54        |
| 050010020           | 38  | 0.0385294    | 0.04           | 52.58        |
| 050010030           | 16  | 0.0162229    | 0.02           | 52.60        |
| 050010040           | 42  | 0.0425851    | 0.04           | 52.64        |
| 050010050           | 34  | 0.0344737    | 0.03           | 52.67        |
| 050010050010        | 1   | 0.0010139    | 0.00           | 52.68        |
| 050010050020        | 1   | 0.0010139    | 0.00           | 52.68        |
| 050010050030        | 2   | 0.0020279    | 0.00           | 52.68        |
| 050010050040        | 2   | 0.0020279    | 0.00           | 52.68        |
| 050010060           | 5   | 0.0050697    | 0.01           | 52.69        |
| 050010070           | 3   | 0.0030418    | 0.00           | 52.69        |
| 050010080           | 4   | 0.0040557    | 0.00           | 52.69        |
| 050010090           | 5   | 0.0050697    | 0.01           | 52.70        |
| 050010100010        | 3   | 0.0030418    | 0.00           | 52.70        |
| 050010100020        | 3   | 0.0030418    | 0.00           | 52.70        |
| 050010100030        | 1   | 0.0010139    | 0.00           | 52.71        |
| 050010110010        | 1   | 0.0010139    | 0.00           | 52.71        |
| 050010110020        | 19  | 0.0192647    | 0.02           | 52.73        |
| 050010110030        | 13  | 0.0131811    | 0.01           | 52.74        |
| 050010110040        | 8   | 0.0081115    | 0.01           | 52.75        |
| 050010110050        | 10  | 0.0101393    | 0.01           | 52.76        |
| 050010110060        | 1   | 0.0010139    | 0.00           | 52.76        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 050010120           | 9   | 0.0091254    | 0.01           | 52.77        |
| 050011              | 5   | 0.0050697    | 0.01           | 52.77        |
| 050020              | 290 | 0.2940401    | 0.29           | 53.07        |
| 050020010           | 72  | 0.0730031    | 0.07           | 53.14        |
| 050020010010        | 12  | 0.0121672    | 0.01           | 53.15        |
| 050020010010010     | 17  | 0.0172368    | 0.02           | 53.17        |
| 050020010010020     | 15  | 0.0152090    | 0.02           | 53.18        |
| 050020010010030     | 14  | 0.0141950    | 0.01           | 53.20        |
| 050020010020        | 12  | 0.0121672    | 0.01           | 53.21        |
| 050020020           | 38  | 0.0385294    | 0.04           | 53.25        |
| 050020020010        | 5   | 0.0050697    | 0.01           | 53.25        |
| 050020020020        | 15  | 0.0152090    | 0.02           | 53.27        |
| 050020020030        | 31  | 0.0314319    | 0.03           | 53.30        |
| 050020020040        | 28  | 0.0283901    | 0.03           | 53.33        |
| 050020020050        | 4   | 0.0040557    | 0.00           | 53.33        |
| 050020030           | 54  | 0.0547523    | 0.05           | 53.39        |
| 050020040           | 53  | 0.0537384    | 0.05           | 53.44        |
| 050020040010        | 28  | 0.0283901    | 0.03           | 53.47        |
| 050020040020        | 16  | 0.0162229    | 0.02           | 53.49        |
| 050020040030        | 8   | 0.0081115    | 0.01           | 53.49        |
| 050020050           | 17  | 0.0172368    | 0.02           | 53.51        |
| 050020060           | 10  | 0.0101393    | 0.01           | 53.52        |
| 050020070           | 4   | 0.0040557    | 0.00           | 53.53        |
| 050020080           | 2   | 0.0020279    | 0.00           | 53.53        |
| 050020090           | 2   | 0.0020279    | 0.00           | 53.53        |
| 050020100           | 2   | 0.0020279    | 0.00           | 53.53        |
| 050020110           | 9   | 0.0091254    | 0.01           | 53.54        |
| 050030              | 196 | 0.1987306    | 0.20           | 53.74        |
| 050030010           | 47  | 0.0476548    | 0.05           | 53.79        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 050030020           | 49   | 0.0496826    | 0.05           | 53.84        |
| 050030030           | 54   | 0.0547523    | 0.05           | 53.89        |
| 050030040           | 48   | 0.0486687    | 0.05           | 53.94        |
| 050030050           | 5    | 0.0050697    | 0.01           | 53.95        |
| 050030060           | 32   | 0.0324458    | 0.03           | 53.98        |
| 050030070           | 23   | 0.0233204    | 0.02           | 54.00        |
| 050030080           | 9    | 0.0091254    | 0.01           | 54.01        |
| 050040              | 112  | 0.1135603    | 0.11           | 54.12        |
| 050040010           | 4    | 0.0040557    | 0.00           | 54.13        |
| 050040020           | 14   | 0.0141950    | 0.01           | 54.14        |
| 050040030           | 8    | 0.0081115    | 0.01           | 54.15        |
| 050050              | 69   | 0.0699613    | 0.07           | 54.22        |
| 050060              | 21   | 0.0212926    | 0.02           | 54.24        |
| 050070              | 23   | 0.0233204    | 0.02           | 54.26        |
| 050080              | 22   | 0.0223065    | 0.02           | 54.29        |
| 050090              | 16   | 0.0162229    | 0.02           | 54.30        |
| 050460              | 6    | 0.0060836    | 0.01           | 54.31        |
| 050660              | 1    | 0.0010139    | 0.00           | 54.31        |
| 051                 | 37   | 0.0375155    | 0.04           | 54.35        |
| 052                 | 7    | 0.0070975    | 0.01           | 54.35        |
| 053                 | 3    | 0.0030418    | 0.00           | 54.36        |
| 060                 | 2336 | 2.3685438    | 2.37           | 56.73        |
| 060002300           | 1    | 0.0010139    | 0.00           | 56.73        |
| 060002310           | 10   | 0.0101393    | 0.01           | 56.74        |
| 060002320           | 12   | 0.0121672    | 0.01           | 56.75        |
| 060002330010        | 13   | 0.0131811    | 0.01           | 56.76        |
| 060002330020        | 7    | 0.0070975    | 0.01           | 56.77        |
| 060002340           | 12   | 0.0121672    | 0.01           | 56.78        |
| 060002350           | 2    | 0.0020279    | 0.00           | 56.78        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 060002360           | 3   | 0.0030418    | 0.00           | 56.79        |
| 060002370           | 2   | 0.0020279    | 0.00           | 56.79        |
| 060010              | 188 | 0.1906191    | 0.19           | 56.98        |
| 060010010           | 9   | 0.0091254    | 0.01           | 56.99        |
| 060010020           | 23  | 0.0233204    | 0.02           | 57.01        |
| 060010030           | 11  | 0.0111532    | 0.01           | 57.02        |
| 060010040           | 12  | 0.0121672    | 0.01           | 57.04        |
| 060010050           | 5   | 0.0050697    | 0.01           | 57.04        |
| 060010060           | 3   | 0.0030418    | 0.00           | 57.04        |
| 060010070           | 6   | 0.0060836    | 0.01           | 57.05        |
| 060020              | 188 | 0.1906191    | 0.19           | 57.24        |
| 060020010           | 23  | 0.0233204    | 0.02           | 57.26        |
| 060020020           | 23  | 0.0233204    | 0.02           | 57.29        |
| 060020030           | 1   | 0.0010139    | 0.00           | 57.29        |
| 060020030010        | 4   | 0.0040557    | 0.00           | 57.29        |
| 060020030020        | 24  | 0.0243344    | 0.02           | 57.32        |
| 060020030030        | 13  | 0.0131811    | 0.01           | 57.33        |
| 060020040           | 2   | 0.0020279    | 0.00           | 57.33        |
| 060020050           | 3   | 0.0030418    | 0.00           | 57.33        |
| 060020060           | 1   | 0.0010139    | 0.00           | 57.34        |
| 060021              | 4   | 0.0040557    | 0.00           | 57.34        |
| 060030              | 116 | 0.1176160    | 0.12           | 57.46        |
| 060030010           | 23  | 0.0233204    | 0.02           | 57.48        |
| 060030020           | 21  | 0.0212926    | 0.02           | 57.50        |
| 060030030           | 6   | 0.0060836    | 0.01           | 57.51        |
| 060030040           | 2   | 0.0020279    | 0.00           | 57.51        |
| 060030050           | 12  | 0.0121672    | 0.01           | 57.52        |
| 060030060           | 1   | 0.0010139    | 0.00           | 57.52        |
| 060040              | 58  | 0.0588080    | 0.06           | 57.58        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 060040010           | 7    | 0.0070975    | 0.01           | 57.59        |
| 060040020           | 6    | 0.0060836    | 0.01           | 57.60        |
| 060050              | 3    | 0.0030418    | 0.00           | 57.60        |
| 060101              | 6    | 0.0060836    | 0.01           | 57.60        |
| 060102              | 5    | 0.0050697    | 0.01           | 57.61        |
| 060110              | 3    | 0.0030418    | 0.00           | 57.61        |
| 060670              | 1    | 0.0010139    | 0.00           | 57.61        |
| 061                 | 14   | 0.0141950    | 0.01           | 57.63        |
| 070                 | 1529 | 1.5503011    | 1.55           | 59.18        |
| 070002390           | 1    | 0.0010139    | 0.00           | 59.18        |
| 070002400           | 7    | 0.0070975    | 0.01           | 59.19        |
| 070002420           | 2    | 0.0020279    | 0.00           | 59.19        |
| 070002430           | 1    | 0.0010139    | 0.00           | 59.19        |
| 070002440           | 1    | 0.0010139    | 0.00           | 59.19        |
| 070002460           | 1    | 0.0010139    | 0.00           | 59.19        |
| 070002461           | 4    | 0.0040557    | 0.00           | 59.20        |
| 070002462           | 2    | 0.0020279    | 0.00           | 59.20        |
| 070002463           | 1    | 0.0010139    | 0.00           | 59.20        |
| 070002470           | 19   | 0.0192647    | 0.02           | 59.22        |
| 070002471           | 2    | 0.0020279    | 0.00           | 59.22        |
| 070002480010        | 2    | 0.0020279    | 0.00           | 59.22        |
| 070002480020        | 9    | 0.0091254    | 0.01           | 59.23        |
| 070002480030        | 6    | 0.0060836    | 0.01           | 59.24        |
| 070002480040        | 4    | 0.0040557    | 0.00           | 59.24        |
| 070002490           | 3    | 0.0030418    | 0.00           | 59.24        |
| 070002500           | 10   | 0.0101393    | 0.01           | 59.25        |
| 070002510           | 2    | 0.0020279    | 0.00           | 59.26        |
| 070002520           | 1    | 0.0010139    | 0.00           | 59.26        |
| 070002530           | 3    | 0.0030418    | 0.00           | 59.26        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 070010              | 194 | 0.1967027    | 0.20           | 59.46        |
| 070010010           | 10  | 0.0101393    | 0.01           | 59.47        |
| 070010020           | 10  | 0.0101393    | 0.01           | 59.48        |
| 070010030           | 7   | 0.0070975    | 0.01           | 59.48        |
| 070010040           | 7   | 0.0070975    | 0.01           | 59.49        |
| 070010050           | 5   | 0.0050697    | 0.01           | 59.50        |
| 070010060           | 1   | 0.0010139    | 0.00           | 59.50        |
| 070010070           | 1   | 0.0010139    | 0.00           | 59.50        |
| 070011              | 1   | 0.0010139    | 0.00           | 59.50        |
| 070020              | 174 | 0.1764241    | 0.18           | 59.68        |
| 070020010           | 13  | 0.0131811    | 0.01           | 59.69        |
| 070020020           | 20  | 0.0202786    | 0.02           | 59.71        |
| 070020020010        | 3   | 0.0030418    | 0.00           | 59.71        |
| 070020020020        | 6   | 0.0060836    | 0.01           | 59.72        |
| 070020020030        | 1   | 0.0010139    | 0.00           | 59.72        |
| 070020030           | 13  | 0.0131811    | 0.01           | 59.73        |
| 070020040           | 1   | 0.0010139    | 0.00           | 59.73        |
| 070030              | 83  | 0.0841563    | 0.08           | 59.82        |
| 070030010           | 7   | 0.0070975    | 0.01           | 59.82        |
| 070030020           | 3   | 0.0030418    | 0.00           | 59.83        |
| 070030020010        | 8   | 0.0081115    | 0.01           | 59.84        |
| 070030020020        | 2   | 0.0020279    | 0.00           | 59.84        |
| 070030030           | 1   | 0.0010139    | 0.00           | 59.84        |
| 070030040           | 1   | 0.0010139    | 0.00           | 59.84        |
| 070040              | 34  | 0.0344737    | 0.03           | 59.87        |
| 070050              | 29  | 0.0294040    | 0.03           | 59.90        |
| 070060              | 6   | 0.0060836    | 0.01           | 59.91        |
| 070080              | 5   | 0.0050697    | 0.01           | 59.92        |
| 070090              | 6   | 0.0060836    | 0.01           | 59.92        |



(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 070680              | 1    | 0.0010139    | 0.00           | 59.92        |
| 070690              | 1    | 0.0010139    | 0.00           | 59.92        |
| 071                 | 3    | 0.0030418    | 0.00           | 59.93        |
| 080                 | 1002 | 1.0159593    | 1.02           | 60.94        |
| 080010              | 234  | 0.2372600    | 0.24           | 61.18        |
| 080010010           | 8    | 0.0081115    | 0.01           | 61.19        |
| 080010020           | 24   | 0.0243344    | 0.02           | 61.21        |
| 080010030           | 14   | 0.0141950    | 0.01           | 61.23        |
| 080010040           | 7    | 0.0070975    | 0.01           | 61.23        |
| 080010050           | 6    | 0.0060836    | 0.01           | 61.24        |
| 080020              | 146  | 0.1480340    | 0.15           | 61.39        |
| 080020010           | 15   | 0.0152090    | 0.02           | 61.40        |
| 080020020           | 8    | 0.0081115    | 0.01           | 61.41        |
| 080020020010        | 6    | 0.0060836    | 0.01           | 61.42        |
| 080020020020        | 25   | 0.0253483    | 0.03           | 61.44        |
| 080020020030        | 55   | 0.0557662    | 0.06           | 61.50        |
| 080020030           | 14   | 0.0141950    | 0.01           | 61.51        |
| 080020040           | 12   | 0.0121672    | 0.01           | 61.52        |
| 080020050           | 6    | 0.0060836    | 0.01           | 61.53        |
| 080020060           | 8    | 0.0081115    | 0.01           | 61.54        |
| 080030              | 69   | 0.0699613    | 0.07           | 61.61        |
| 080030010           | 18   | 0.0182508    | 0.02           | 61.63        |
| 080030020           | 4    | 0.0040557    | 0.00           | 61.63        |
| 080030020010        | 3    | 0.0030418    | 0.00           | 61.63        |
| 080030020020        | 3    | 0.0030418    | 0.00           | 61.64        |
| 080030020030        | 2    | 0.0020279    | 0.00           | 61.64        |
| 080030020040        | 1    | 0.0010139    | 0.00           | 61.64        |
| 080030020050        | 1    | 0.0010139    | 0.00           | 61.64        |
| 080030030           | 4    | 0.0040557    | 0.00           | 61.65        |

(continued)

| gliederungskennzahl | N    | exactpercent | roundedpercent | cumulpercent |
|---------------------|------|--------------|----------------|--------------|
| 080030040           | 1    | 0.0010139    | 0.00           | 61.65        |
| 080040              | 54   | 0.0547523    | 0.05           | 61.70        |
| 080050              | 52   | 0.0527244    | 0.05           | 61.75        |
| 080060              | 26   | 0.0263622    | 0.03           | 61.78        |
| 080060010           | 4    | 0.0040557    | 0.00           | 61.78        |
| 080060020           | 3    | 0.0030418    | 0.00           | 61.79        |
| 080060030           | 5    | 0.0050697    | 0.01           | 61.79        |
| 080060040           | 2    | 0.0020279    | 0.00           | 61.79        |
| 080070              | 3    | 0.0030418    | 0.00           | 61.80        |
| 080080              | 7    | 0.0070975    | 0.01           | 61.80        |
| 081                 | 5    | 0.0050697    | 0.01           | 61.81        |
| 090                 | 1076 | 1.0909902    | 1.09           | 62.90        |
| 090010              | 43   | 0.0435991    | 0.04           | 62.94        |
| 090010010           | 25   | 0.0253483    | 0.03           | 62.97        |
| 090010020           | 12   | 0.0121672    | 0.01           | 62.98        |
| 090020              | 48   | 0.0486687    | 0.05           | 63.03        |
| 090020010           | 9    | 0.0091254    | 0.01           | 63.04        |
| 090020020           | 20   | 0.0202786    | 0.02           | 63.06        |
| 090020030           | 7    | 0.0070975    | 0.01           | 63.07        |
| 090030              | 69   | 0.0699613    | 0.07           | 63.14        |
| 090040              | 32   | 0.0324458    | 0.03           | 63.17        |
| 090040010           | 11   | 0.0111532    | 0.01           | 63.18        |
| 090040020           | 3    | 0.0030418    | 0.00           | 63.18        |
| 090050              | 20   | 0.0202786    | 0.02           | 63.20        |
| 090060              | 1    | 0.0010139    | 0.00           | 63.20        |
| 091                 | 11   | 0.0111532    | 0.01           | 63.22        |
| 100                 | 582  | 0.5901081    | 0.59           | 63.81        |
| 100010              | 13   | 0.0131811    | 0.01           | 63.82        |
| 100010010           | 5    | 0.0050697    | 0.01           | 63.82        |

(continued)

| gliederungskennzahl | N  | exactpercent | roundedpercent | cumulpercent |
|---------------------|----|--------------|----------------|--------------|
| 100010020           | 10 | 0.0101393    | 0.01           | 63.83        |
| 100020              | 25 | 0.0253483    | 0.03           | 63.86        |
| 100020010           | 20 | 0.0202786    | 0.02           | 63.88        |
| 100020020           | 13 | 0.0131811    | 0.01           | 63.89        |
| 100020030           | 3  | 0.0030418    | 0.00           | 63.90        |
| 100030              | 27 | 0.0273761    | 0.03           | 63.92        |
| 100030010           | 4  | 0.0040557    | 0.00           | 63.93        |
| 100030020           | 4  | 0.0040557    | 0.00           | 63.93        |
| 100040              | 6  | 0.0060836    | 0.01           | 63.94        |
| 100050              | 14 | 0.0141950    | 0.01           | 63.95        |
| 100051              | 1  | 0.0010139    | 0.00           | 63.95        |
| 100060              | 8  | 0.0081115    | 0.01           | 63.96        |
| 100070              | 1  | 0.0010139    | 0.00           | 63.96        |
| 100080              | 2  | 0.0020279    | 0.00           | 63.96        |
| 100090              | 4  | 0.0040557    | 0.00           | 63.97        |
| 100100              | 1  | 0.0010139    | 0.00           | 63.97        |
| 101                 | 11 | 0.0111532    | 0.01           | 63.98        |
| 102                 | 11 | 0.0111532    | 0.01           | 63.99        |
| 102010              | 1  | 0.0010139    | 0.00           | 63.99        |
| 102020              | 2  | 0.0020279    | 0.00           | 63.99        |
| 102030010           | 10 | 0.0101393    | 0.01           | 64.00        |
| 102030020010        | 4  | 0.0040557    | 0.00           | 64.01        |
| 102030020020        | 6  | 0.0060836    | 0.01           | 64.01        |
| 102030020030        | 1  | 0.0010139    | 0.00           | 64.02        |
| 102030030           | 5  | 0.0050697    | 0.01           | 64.02        |
| 103                 | 1  | 0.0010139    | 0.00           | 64.02        |
| 103010              | 1  | 0.0010139    | 0.00           | 64.02        |
| 103020              | 1  | 0.0010139    | 0.00           | 64.02        |
| 103030              | 1  | 0.0010139    | 0.00           | 64.02        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 103040              | 1   | 0.0010139    | 0.00           | 64.03        |
| 103050              | 1   | 0.0010139    | 0.00           | 64.03        |
| 103060              | 1   | 0.0010139    | 0.00           | 64.03        |
| 103070              | 1   | 0.0010139    | 0.00           | 64.03        |
| 103080              | 1   | 0.0010139    | 0.00           | 64.03        |
| 103090              | 1   | 0.0010139    | 0.00           | 64.03        |
| 103100              | 1   | 0.0010139    | 0.00           | 64.03        |
| 103130              | 1   | 0.0010139    | 0.00           | 64.03        |
| 104                 | 1   | 0.0010139    | 0.00           | 64.03        |
| 105                 | 1   | 0.0010139    | 0.00           | 64.03        |
| 105010              | 1   | 0.0010139    | 0.00           | 64.04        |
| 106                 | 1   | 0.0010139    | 0.00           | 64.04        |
| 107                 | 1   | 0.0010139    | 0.00           | 64.04        |
| 108                 | 1   | 0.0010139    | 0.00           | 64.04        |
| 109                 | 1   | 0.0010139    | 0.00           | 64.04        |
| 110                 | 411 | 0.4167258    | 0.42           | 64.46        |
| 110010              | 61  | 0.0618498    | 0.06           | 64.52        |
| 110010010           | 5   | 0.0050697    | 0.01           | 64.52        |
| 110010020           | 3   | 0.0030418    | 0.00           | 64.53        |
| 110010030           | 5   | 0.0050697    | 0.01           | 64.53        |
| 110020              | 19  | 0.0192647    | 0.02           | 64.55        |
| 110020010           | 14  | 0.0141950    | 0.01           | 64.57        |
| 110020020           | 5   | 0.0050697    | 0.01           | 64.57        |
| 110020030           | 5   | 0.0050697    | 0.01           | 64.58        |
| 110030              | 25  | 0.0253483    | 0.03           | 64.60        |
| 110040              | 8   | 0.0081115    | 0.01           | 64.61        |
| 110040010           | 3   | 0.0030418    | 0.00           | 64.61        |
| 110040020           | 5   | 0.0050697    | 0.01           | 64.62        |
| 110050              | 8   | 0.0081115    | 0.01           | 64.62        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 110050010           | 10  | 0.0101393    | 0.01           | 64.64        |
| 110050020           | 4   | 0.0040557    | 0.00           | 64.64        |
| 110050020010        | 4   | 0.0040557    | 0.00           | 64.64        |
| 110050020020        | 6   | 0.0060836    | 0.01           | 64.65        |
| 110050020030        | 2   | 0.0020279    | 0.00           | 64.65        |
| 110050020040        | 2   | 0.0020279    | 0.00           | 64.65        |
| 110050030           | 2   | 0.0020279    | 0.00           | 64.66        |
| 110050040           | 5   | 0.0050697    | 0.01           | 64.66        |
| 110050050           | 2   | 0.0020279    | 0.00           | 64.66        |
| 110050060           | 2   | 0.0020279    | 0.00           | 64.66        |
| 110050070           | 1   | 0.0010139    | 0.00           | 64.67        |
| 110050080           | 1   | 0.0010139    | 0.00           | 64.67        |
| 110060              | 8   | 0.0081115    | 0.01           | 64.67        |
| 110060010           | 9   | 0.0091254    | 0.01           | 64.68        |
| 110060020           | 5   | 0.0050697    | 0.01           | 64.69        |
| 110070              | 5   | 0.0050697    | 0.01           | 64.69        |
| 110070010           | 2   | 0.0020279    | 0.00           | 64.70        |
| 110070020           | 6   | 0.0060836    | 0.01           | 64.70        |
| 110080              | 13  | 0.0131811    | 0.01           | 64.72        |
| 111                 | 1   | 0.0010139    | 0.00           | 64.72        |
| 112                 | 2   | 0.0020279    | 0.00           | 64.72        |
| 120                 | 181 | 0.1835216    | 0.18           | 64.90        |
| 120010              | 43  | 0.0435991    | 0.04           | 64.95        |
| 120020              | 24  | 0.0243344    | 0.02           | 64.97        |
| 120030              | 25  | 0.0253483    | 0.03           | 65.00        |
| 130                 | 137 | 0.1389086    | 0.14           | 65.13        |
| 130010              | 17  | 0.0172368    | 0.02           | 65.15        |
| 130020              | 24  | 0.0243344    | 0.02           | 65.18        |
| 130030              | 10  | 0.0101393    | 0.01           | 65.19        |

(continued)

| gliederungskennzahl | N   | exactpercent | roundedpercent | cumulpercent |
|---------------------|-----|--------------|----------------|--------------|
| 130040              | 9   | 0.0091254    | 0.01           | 65.19        |
| 130050              | 26  | 0.0263622    | 0.03           | 65.22        |
| 130060              | 2   | 0.0020279    | 0.00           | 65.22        |
| 130070              | 3   | 0.0030418    | 0.00           | 65.23        |
| 130080              | 5   | 0.0050697    | 0.01           | 65.23        |
| 140                 | 104 | 0.1054489    | 0.11           | 65.34        |
| 150                 | 93  | 0.0942956    | 0.09           | 65.43        |
| 150010              | 8   | 0.0081115    | 0.01           | 65.44        |
| 150020              | 8   | 0.0081115    | 0.01           | 65.45        |
| 150030              | 1   | 0.0010139    | 0.00           | 65.45        |
| 160                 | 108 | 0.1095046    | 0.11           | 65.56        |
| 160010              | 12  | 0.0121672    | 0.01           | 65.57        |
| 160020              | 8   | 0.0081115    | 0.01           | 65.58        |
| 160030              | 10  | 0.0101393    | 0.01           | 65.59        |
| 170                 | 26  | 0.0263622    | 0.03           | 65.61        |
| 171                 | 5   | 0.0050697    | 0.01           | 65.62        |
| 172                 | 2   | 0.0020279    | 0.00           | 65.62        |
| 174                 | 2   | 0.0020279    | 0.00           | 65.62        |
| 180                 | 27  | 0.0273761    | 0.03           | 65.65        |
| 180010              | 35  | 0.0354876    | 0.04           | 65.69        |
| 180020              | 4   | 0.0040557    | 0.00           | 65.69        |
| 180030              | 5   | 0.0050697    | 0.01           | 65.70        |
| 180040              | 1   | 0.0010139    | 0.00           | 65.70        |
| 180050              | 1   | 0.0010139    | 0.00           | 65.70        |
| 180060              | 1   | 0.0010139    | 0.00           | 65.70        |
| 180070              | 1   | 0.0010139    | 0.00           | 65.70        |
| 180080              | 1   | 0.0010139    | 0.00           | 65.70        |
| 180090              | 1   | 0.0010139    | 0.00           | 65.70        |
| 180100              | 1   | 0.0010139    | 0.00           | 65.70        |

(continued)

| gliederungskennzahl | N  | exactpercent | roundedpercent | cumulpercent |
|---------------------|----|--------------|----------------|--------------|
| 180110              | 1  | 0.0010139    | 0.00           | 65.70        |
| 180120              | 1  | 0.0010139    | 0.00           | 65.70        |
| 180130              | 1  | 0.0010139    | 0.00           | 65.71        |
| 180140              | 3  | 0.0030418    | 0.00           | 65.71        |
| 180160              | 1  | 0.0010139    | 0.00           | 65.71        |
| 180170              | 1  | 0.0010139    | 0.00           | 65.71        |
| 180180              | 1  | 0.0010139    | 0.00           | 65.71        |
| 180190              | 2  | 0.0020279    | 0.00           | 65.71        |
| 190                 | 21 | 0.0212926    | 0.02           | 65.74        |
| 200                 | 13 | 0.0131811    | 0.01           | 65.75        |
| 210                 | 21 | 0.0212926    | 0.02           | 65.77        |
| 220                 | 22 | 0.0223065    | 0.02           | 65.79        |
| 230                 | 23 | 0.0233204    | 0.02           | 65.82        |
| 231                 | 1  | 0.0010139    | 0.00           | 65.82        |
| 240                 | 14 | 0.0141950    | 0.01           | 65.83        |
| 250                 | 13 | 0.0131811    | 0.01           | 65.84        |
| 260                 | 5  | 0.0050697    | 0.01           | 65.85        |
| 270                 | 3  | 0.0030418    | 0.00           | 65.85        |
| 280                 | 4  | 0.0040557    | 0.00           | 65.86        |
| 281                 | 16 | 0.0162229    | 0.02           | 65.87        |
| 290                 | 5  | 0.0050697    | 0.01           | 65.88        |
| 300                 | 3  | 0.0030418    | 0.00           | 65.88        |
| 301                 | 1  | 0.0010139    | 0.00           | 65.88        |
| 310                 | 5  | 0.0050697    | 0.01           | 65.89        |
| 320                 | 2  | 0.0020279    | 0.00           | 65.89        |
| 330                 | 16 | 0.0162229    | 0.02           | 65.90        |
| 340                 | 6  | 0.0060836    | 0.01           | 65.91        |
| 350                 | 6  | 0.0060836    | 0.01           | 65.92        |
| 360                 | 2  | 0.0020279    | 0.00           | 65.92        |

(continued)

| gliederungskennzahl | N | exactpercent | roundedpercent | cumulpercent |
|---------------------|---|--------------|----------------|--------------|
| 370                 | 2 | 0.0020279    | 0.00           | 65.92        |
| 371                 | 1 | 0.0010139    | 0.00           | 65.92        |
| 380                 | 4 | 0.0040557    | 0.00           | 65.93        |
| 390                 | 2 | 0.0020279    | 0.00           | 65.93        |
| 400                 | 2 | 0.0020279    | 0.00           | 65.93        |
| 410                 | 3 | 0.0030418    | 0.00           | 65.93        |
| 420                 | 4 | 0.0040557    | 0.00           | 65.94        |
| 421                 | 1 | 0.0010139    | 0.00           | 65.94        |
| 430                 | 2 | 0.0020279    | 0.00           | 65.94        |
| 440                 | 5 | 0.0050697    | 0.01           | 65.95        |
| 450                 | 1 | 0.0010139    | 0.00           | 65.95        |
| 460                 | 5 | 0.0050697    | 0.01           | 65.95        |
| 460460              | 3 | 0.0030418    | 0.00           | 65.95        |
| 460490              | 2 | 0.0020279    | 0.00           | 65.96        |
| 460510              | 8 | 0.0081115    | 0.01           | 65.96        |
| 460590              | 4 | 0.0040557    | 0.00           | 65.97        |
| 460620              | 1 | 0.0010139    | 0.00           | 65.97        |
| 470                 | 2 | 0.0020279    | 0.00           | 65.97        |
| 480                 | 3 | 0.0030418    | 0.00           | 65.97        |
| 491                 | 3 | 0.0030418    | 0.00           | 65.98        |
| 510                 | 2 | 0.0020279    | 0.00           | 65.98        |
| 520                 | 6 | 0.0060836    | 0.01           | 65.99        |
| 530                 | 2 | 0.0020279    | 0.00           | 65.99        |
| 550                 | 1 | 0.0010139    | 0.00           | 65.99        |
| 560                 | 2 | 0.0020279    | 0.00           | 65.99        |
| 580                 | 1 | 0.0010139    | 0.00           | 65.99        |
| 590                 | 1 | 0.0010139    | 0.00           | 65.99        |
| 600                 | 2 | 0.0020279    | 0.00           | 65.99        |
| 610                 | 2 | 0.0020279    | 0.00           | 66.00        |



(continued)

| gliederungskennzahl | N | exactpercent | roundedpercent | cumulpercent |
|---------------------|---|--------------|----------------|--------------|
| 620                 | 8 | 0.0081115    | 0.01           | 66.00        |
| 630                 | 7 | 0.0070975    | 0.01           | 66.01        |
| 640                 | 1 | 0.0010139    | 0.00           | 66.01        |
| 650                 | 1 | 0.0010139    | 0.00           | 66.01        |
| 680                 | 1 | 0.0010139    | 0.00           | 66.02        |
| 681                 | 2 | 0.0020279    | 0.00           | 66.02        |
| 690                 | 2 | 0.0020279    | 0.00           | 66.02        |
| 700                 | 2 | 0.0020279    | 0.00           | 66.02        |
| 700010              | 1 | 0.0010139    | 0.00           | 66.02        |
| 700020              | 1 | 0.0010139    | 0.00           | 66.02        |
| 700030              | 1 | 0.0010139    | 0.00           | 66.02        |
| 700040              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700050              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700060              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700070              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700080              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700090              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700100              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700110              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700120              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700130              | 1 | 0.0010139    | 0.00           | 66.03        |
| 700140              | 1 | 0.0010139    | 0.00           | 66.04        |
| 700150              | 1 | 0.0010139    | 0.00           | 66.04        |
| 700160              | 1 | 0.0010139    | 0.00           | 66.04        |
| 700170              | 1 | 0.0010139    | 0.00           | 66.04        |
| 700180              | 1 | 0.0010139    | 0.00           | 66.04        |
| 710                 | 2 | 0.0020279    | 0.00           | 66.04        |
| 720                 | 7 | 0.0070975    | 0.01           | 66.05        |
| 730                 | 1 | 0.0010139    | 0.00           | 66.05        |

(continued)

| gliederungskennzahl | N     | exactpercent | roundedpercent | cumulpercent |
|---------------------|-------|--------------|----------------|--------------|
| 730740              | 5     | 0.0050697    | 0.01           | 66.05        |
| 731                 | 6     | 0.0060836    | 0.01           | 66.06        |
| 740                 | 6     | 0.0060836    | 0.01           | 66.07        |
| 780                 | 2     | 0.0020279    | 0.00           | 66.07        |
| 800                 | 1     | 0.0010139    | 0.00           | 66.07        |
| 810                 | 1     | 0.0010139    | 0.00           | 66.07        |
| 820                 | 1     | 0.0010139    | 0.00           | 66.07        |
| 830                 | 5     | 0.0050697    | 0.01           | 66.08        |
| 840                 | 4     | 0.0040557    | 0.00           | 66.08        |
| 870                 | 1     | 0.0010139    | 0.00           | 66.08        |
| 880                 | 5     | 0.0050697    | 0.01           | 66.09        |
| 911                 | 1     | 0.0010139    | 0.00           | 66.09        |
| 920                 | 1     | 0.0010139    | 0.00           | 66.09        |
| 930                 | 1     | 0.0010139    | 0.00           | 66.09        |
| 950                 | 1     | 0.0010139    | 0.00           | 66.09        |
| 960                 | 2     | 0.0020279    | 0.00           | 66.09        |
| 980                 | 1     | 0.0010139    | 0.00           | 66.09        |
| 990                 | 1     | 0.0010139    | 0.00           | 66.10        |
| 991                 | 3     | 0.0030418    | 0.00           | 66.10        |
| NA                  | 33436 | 33.9018109   | 33.90          | 100.00       |
| Total               | 98626 | 100.0000000  | 99.00          | 100.00       |

Frequency Table for Variable: ausfertigung\_jahr

122 unique value(s) detected.

| ausfertigung_jahr | N   | exactpercent | roundedpercent | cumulpercent |
|-------------------|-----|--------------|----------------|--------------|
| 1869              | 146 | 0.1480340    | 0.15           | 0.15         |
| 1871              | 558 | 0.5657737    | 0.57           | 0.71         |

(continued)

| ausfertigung_jahr | N    | exactpercent | roundedpercent | cumulpercent |
|-------------------|------|--------------|----------------|--------------|
| 1877              | 97   | 0.0983513    | 0.10           | 0.81         |
| 1884              | 10   | 0.0101393    | 0.01           | 0.82         |
| 1886              | 3    | 0.0030418    | 0.00           | 0.83         |
| 1887              | 4    | 0.0040557    | 0.00           | 0.83         |
| 1889              | 214  | 0.2169813    | 0.22           | 1.05         |
| 1892              | 115  | 0.1166021    | 0.12           | 1.16         |
| 1894              | 4    | 0.0040557    | 0.00           | 1.17         |
| 1895              | 74   | 0.0750309    | 0.08           | 1.24         |
| 1896              | 2863 | 2.9028856    | 2.90           | 4.14         |
| 1897              | 1131 | 1.1467564    | 1.15           | 5.29         |
| 1898              | 3    | 0.0030418    | 0.00           | 5.29         |
| 1899              | 2    | 0.0020279    | 0.00           | 5.30         |
| 1901              | 51   | 0.0517105    | 0.05           | 5.35         |
| 1903              | 1    | 0.0010139    | 0.00           | 5.35         |
| 1906              | 1    | 0.0010139    | 0.00           | 5.35         |
| 1907              | 26   | 0.0263622    | 0.03           | 5.38         |
| 1908              | 7    | 0.0070975    | 0.01           | 5.38         |
| 1909              | 118  | 0.1196439    | 0.12           | 5.50         |
| 1910              | 11   | 0.0111532    | 0.01           | 5.51         |
| 1911              | 47   | 0.0476548    | 0.05           | 5.56         |
| 1913              | 45   | 0.0456269    | 0.05           | 5.61         |
| 1919              | 99   | 0.1003792    | 0.10           | 5.71         |
| 1920              | 7    | 0.0070975    | 0.01           | 5.72         |
| 1921              | 40   | 0.0405573    | 0.04           | 5.76         |
| 1922              | 248  | 0.2514550    | 0.25           | 6.01         |
| 1923              | 15   | 0.0152090    | 0.02           | 6.02         |
| 1924              | 2    | 0.0020279    | 0.00           | 6.02         |
| 1925              | 2    | 0.0020279    | 0.00           | 6.03         |
| 1926              | 31   | 0.0314319    | 0.03           | 6.06         |

(continued)

| ausfertigung_jahr | N    | exactpercent | roundedpercent | cumulpercent |
|-------------------|------|--------------|----------------|--------------|
| 1927              | 32   | 0.0324458    | 0.03           | 6.09         |
| 1928              | 8    | 0.0081115    | 0.01           | 6.10         |
| 1929              | 8    | 0.0081115    | 0.01           | 6.11         |
| 1930              | 14   | 0.0141950    | 0.01           | 6.12         |
| 1931              | 46   | 0.0466408    | 0.05           | 6.17         |
| 1933              | 196  | 0.1987306    | 0.20           | 6.37         |
| 1934              | 512  | 0.5191329    | 0.52           | 6.89         |
| 1935              | 157  | 0.1591872    | 0.16           | 7.04         |
| 1936              | 284  | 0.2879565    | 0.29           | 7.33         |
| 1937              | 320  | 0.3244581    | 0.32           | 7.66         |
| 1938              | 48   | 0.0486687    | 0.05           | 7.71         |
| 1939              | 90   | 0.0912538    | 0.09           | 7.80         |
| 1940              | 215  | 0.2179953    | 0.22           | 8.02         |
| 1941              | 25   | 0.0253483    | 0.03           | 8.04         |
| 1942              | 20   | 0.0202786    | 0.02           | 8.06         |
| 1943              | 24   | 0.0243344    | 0.02           | 8.09         |
| 1944              | 12   | 0.0121672    | 0.01           | 8.10         |
| 1947              | 19   | 0.0192647    | 0.02           | 8.12         |
| 1948              | 73   | 0.0740170    | 0.07           | 8.19         |
| 1949              | 314  | 0.3183745    | 0.32           | 8.51         |
| 1950              | 2087 | 2.1160749    | 2.12           | 10.62        |
| 1951              | 584  | 0.5921359    | 0.59           | 11.22        |
| 1952              | 445  | 0.4511995    | 0.45           | 11.67        |
| 1953              | 1815 | 1.8402855    | 1.84           | 13.51        |
| 1954              | 209  | 0.2119117    | 0.21           | 13.72        |
| 1955              | 303  | 0.3072212    | 0.31           | 14.03        |
| 1956              | 700  | 0.7097520    | 0.71           | 14.74        |
| 1957              | 990  | 1.0037921    | 1.00           | 15.74        |
| 1958              | 149  | 0.1510758    | 0.15           | 15.89        |

(continued)

| ausfertigung_jahr | N    | exactpercent | roundedpercent | cumulpercent |
|-------------------|------|--------------|----------------|--------------|
| 1959              | 750  | 0.7604486    | 0.76           | 16.65        |
| 1960              | 1097 | 1.1122828    | 1.11           | 17.77        |
| 1961              | 1292 | 1.3099994    | 1.31           | 19.08        |
| 1962              | 245  | 0.2484132    | 0.25           | 19.32        |
| 1963              | 338  | 0.3427088    | 0.34           | 19.67        |
| 1964              | 330  | 0.3345974    | 0.33           | 20.00        |
| 1965              | 1462 | 1.4823677    | 1.48           | 21.48        |
| 1966              | 564  | 0.5718573    | 0.57           | 22.06        |
| 1967              | 330  | 0.3345974    | 0.33           | 22.39        |
| 1968              | 478  | 0.4846592    | 0.48           | 22.87        |
| 1969              | 722  | 0.7320585    | 0.73           | 23.61        |
| 1970              | 457  | 0.4633667    | 0.46           | 24.07        |
| 1971              | 776  | 0.7868108    | 0.79           | 24.86        |
| 1972              | 823  | 0.8344656    | 0.83           | 25.69        |
| 1973              | 239  | 0.2423296    | 0.24           | 25.93        |
| 1974              | 900  | 0.9125383    | 0.91           | 26.85        |
| 1975              | 927  | 0.9399144    | 0.94           | 27.79        |
| 1976              | 2492 | 2.5267171    | 2.53           | 30.31        |
| 1977              | 747  | 0.7574068    | 0.76           | 31.07        |
| 1978              | 429  | 0.4349766    | 0.43           | 31.50        |
| 1979              | 528  | 0.5353558    | 0.54           | 32.04        |
| 1980              | 1127 | 1.1427007    | 1.14           | 33.18        |
| 1981              | 580  | 0.5880802    | 0.59           | 33.77        |
| 1982              | 542  | 0.5495508    | 0.55           | 34.32        |
| 1983              | 244  | 0.2473993    | 0.25           | 34.57        |
| 1984              | 356  | 0.3609596    | 0.36           | 34.93        |
| 1985              | 623  | 0.6316793    | 0.63           | 35.56        |
| 1986              | 569  | 0.5769270    | 0.58           | 36.14        |
| 1987              | 456  | 0.4623527    | 0.46           | 36.60        |

(continued)

| ausfertigung_jahr | N    | exactpercent | roundedpercent | cumulpercent |
|-------------------|------|--------------|----------------|--------------|
| 1988              | 1268 | 1.2856650    | 1.29           | 37.89        |
| 1989              | 1142 | 1.1579097    | 1.16           | 39.04        |
| 1990              | 2438 | 2.4719648    | 2.47           | 41.52        |
| 1991              | 798  | 0.8091173    | 0.81           | 42.32        |
| 1992              | 1030 | 1.0443494    | 1.04           | 43.37        |
| 1993              | 1003 | 1.0169732    | 1.02           | 44.39        |
| 1994              | 3875 | 3.9289842    | 3.93           | 48.31        |
| 1995              | 747  | 0.7574068    | 0.76           | 49.07        |
| 1996              | 958  | 0.9713463    | 0.97           | 50.04        |
| 1997              | 1872 | 1.8980796    | 1.90           | 51.94        |
| 1998              | 1492 | 1.5127857    | 1.51           | 53.45        |
| 1999              | 802  | 0.8131730    | 0.81           | 54.27        |
| 2000              | 705  | 0.7148216    | 0.71           | 54.98        |
| 2001              | 1725 | 1.7490317    | 1.75           | 56.73        |
| 2002              | 2339 | 2.3715856    | 2.37           | 59.10        |
| 2003              | 1226 | 1.2430799    | 1.24           | 60.35        |
| 2004              | 2728 | 2.7660049    | 2.77           | 63.11        |
| 2005              | 2030 | 2.0582808    | 2.06           | 65.17        |
| 2006              | 2282 | 2.3137915    | 2.31           | 67.48        |
| 2007              | 1667 | 1.6902237    | 1.69           | 69.17        |
| 2008              | 2238 | 2.2691785    | 2.27           | 71.44        |
| 2009              | 2840 | 2.8795652    | 2.88           | 74.32        |
| 2010              | 1455 | 1.4752702    | 1.48           | 75.80        |
| 2011              | 2399 | 2.4324215    | 2.43           | 78.23        |
| 2012              | 1507 | 1.5279946    | 1.53           | 79.76        |
| 2013              | 3435 | 3.4828544    | 3.48           | 83.24        |
| 2014              | 1611 | 1.6334435    | 1.63           | 84.88        |
| 2015              | 2054 | 2.0826151    | 2.08           | 86.96        |
| 2016              | 3109 | 3.1523128    | 3.15           | 90.11        |

(continued)

| ausfertigung_jahr | N     | exactpercent | roundedpercent | cumulpercent |
|-------------------|-------|--------------|----------------|--------------|
| 2017              | 3530  | 3.5791779    | 3.58           | 93.69        |
| 2018              | 1640  | 1.6628475    | 1.66           | 95.35        |
| 2019              | 1721  | 1.7449760    | 1.74           | 97.10        |
| 2020              | 2863  | 2.9028856    | 2.90           | 100.00       |
| Total             | 98626 | 100.0000000  | 100.00         | 100.00       |

Frequency Table for Variable: doi\_concept

1 unique value(s) detected.

| doi_concept            | N     | exactpercent | roundedpercent | cumulpercent |
|------------------------|-------|--------------|----------------|--------------|
| 10.5281/zenodo.3832111 | 98626 | 100          | 100            | 100          |
| Total                  | 98626 | 100          | 100            | 100          |

Frequency Table for Variable: doi\_version

1 unique value(s) detected.

| doi_version            | N     | exactpercent | roundedpercent | cumulpercent |
|------------------------|-------|--------------|----------------|--------------|
| 10.5281/zenodo.4384771 | 98626 | 100          | 100            | 100          |
| Total                  | 98626 | 100          | 100            | 100          |

Frequency Table for Variable: version

1 unique value(s) detected.

| version    | N     | exactpercent | roundedpercent | cumulpercent |
|------------|-------|--------------|----------------|--------------|
| 2021-01-05 | 98626 | 100          | 100            | 100          |
| Total      | 98626 | 100          | 100            | 100          |

## 8 Frequenztabellen erstellen: Rechtsakte

### 8.1 Variablen ignorieren

Folgende Variablen sind wegen der geringeren Auflösung der Metadaten (nur Rechtsaktebene, nicht Normebene) nicht mehr nutzbar:

```
varremove <- c("gliederungskennzahl")

vars.freqtable.rechtsakte <- grep(paste(varremove,
                                         collapse="|"),
                                vars.freqtable,
                                invert = TRUE,
                                value = TRUE)
```

### 8.2 Liste zu prüfender Variablen

```
print(vars.freqtable.rechtsakte)
```

```
## [1] "periodikum"      "fundstellentyp"  "check_neuf"
## [4] "check_aufh"      "check_sonst"     "check_hinweis"
## [7] "check_stand"     "ausfertigung_jahr" "doi_concept"
## [10] "doi_version"     "version"
```

### 8.3 Frequenztabellen erstellen

```
prefix <- paste0(datasetname,
                  "_01_Rechtsakte_Frequenztafel_var-")
```

```
f.fast.freqtable(dt.rechtsakte,
                 varlist = vars.freqtable.rechtsakte,
                 sumrow = TRUE,
                 output.list = FALSE,
                 output.kable = TRUE,
                 output.csv = TRUE,
                 outputdir = outputdir,
                 prefix = prefix)
```

---

Frequency Table for Variable: periodikum

---

21 unique value(s) detected.



| periodikum | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| BAGVBl     | 1    | 0.0179565    | 0.02           | 0.02         |
| BAnz       | 133  | 2.3882205    | 2.39           | 2.41         |
| BGBI       | 21   | 0.3770874    | 0.38           | 2.78         |
| BGBI I     | 4795 | 86.1016340   | 86.10          | 88.88        |
| BGBI II    | 375  | 6.7337044    | 6.73           | 95.62        |
| GBl DDR    | 15   | 0.2693482    | 0.27           | 95.89        |
| GBl DDR I  | 44   | 0.7900880    | 0.79           | 96.68        |
| GBl DDR II | 3    | 0.0538696    | 0.05           | 96.73        |
| GVBl BE    | 1    | 0.0179565    | 0.02           | 96.75        |
| NV         | 1    | 0.0179565    | 0.02           | 96.77        |
| RAnz       | 2    | 0.0359131    | 0.04           | 96.80        |
| RGBI       | 51   | 0.9157838    | 0.92           | 97.72        |
| RGBI I     | 48   | 0.8619142    | 0.86           | 98.58        |
| RGBI II    | 45   | 0.8080445    | 0.81           | 99.39        |
| RMBI       | 4    | 0.0718262    | 0.07           | 99.46        |
| VOBl BrZ   | 1    | 0.0179565    | 0.02           | 99.48        |
| VkBl       | 10   | 0.1795655    | 0.18           | 99.66        |
| WiGBI      | 7    | 0.1256958    | 0.13           | 99.78        |
| ZBl        | 2    | 0.0359131    | 0.04           | 99.82        |
| eBAnz      | 6    | 0.1077393    | 0.11           | 99.93        |
| Öff Anz    | 4    | 0.0718262    | 0.07           | 100.00       |
| Total      | 5569 | 100.0000000  | 100.00         | 100.00       |

---

Frequency Table for Variable: fundstellentyp

---

1 unique value(s) detected.

| fundstellentyp | N    | exactpercent | roundedpercent | cumulpercent |
|----------------|------|--------------|----------------|--------------|
| amtlich        | 5569 | 100          | 100            | 100          |
| Total          | 5569 | 100          | 100            | 100          |

---

Frequency Table for Variable: check\_neuf

---

2 unique value(s) detected.

| check_neuf | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| NA         | 5150 | 92.476208    | 92.48          | 92.48        |
| ja         | 419  | 7.523792     | 7.52           | 100.00       |
| Total      | 5569 | 100.000000   | 100.00         | 100.00       |

---

Frequency Table for Variable: check\_aufh

---

2 unique value(s) detected.

| check_aufh | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| NA         | 5413 | 97.198779    | 97.2           | 97.2         |
| ja         | 156  | 2.801221     | 2.8            | 100.0        |
| Total      | 5569 | 100.000000   | 100.0          | 100.0        |

---

Frequency Table for Variable: check\_sonst

---

2 unique value(s) detected.

| check_sonst | N    | exactpercent | roundedpercent | cumulpercent |
|-------------|------|--------------|----------------|--------------|
| NA          | 5110 | 91.757946    | 91.76          | 91.76        |
| ja          | 459  | 8.242054     | 8.24           | 100.00       |
| Total       | 5569 | 100.000000   | 100.00         | 100.00       |

---

Frequency Table for Variable: check\_hinweis

---

2 unique value(s) detected.

| check_hinweis | N    | exactpercent | roundedpercent | cumulpercent |
|---------------|------|--------------|----------------|--------------|
| NA            | 5286 | 94.918298    | 94.92          | 94.92        |
| ja            | 283  | 5.081702     | 5.08           | 100.00       |
| Total         | 5569 | 100.000000   | 100.00         | 100.00       |

Frequency Table for Variable: check\_stand

2 unique value(s) detected.

| check_stand | N    | exactpercent | roundedpercent | cumulpercent |
|-------------|------|--------------|----------------|--------------|
| NA          | 2806 | 50.38607     | 50.39          | 50.39        |
| ja          | 2763 | 49.61393     | 49.61          | 100.00       |
| Total       | 5569 | 100.00000    | 100.00         | 100.00       |

Frequency Table for Variable: ausfertigung\_jahr

122 unique value(s) detected.

| ausfertigung_jahr | N | exactpercent | roundedpercent | cumulpercent |
|-------------------|---|--------------|----------------|--------------|
| 1869              | 1 | 0.0179565    | 0.02           | 0.02         |
| 1871              | 3 | 0.0538696    | 0.05           | 0.07         |
| 1877              | 3 | 0.0538696    | 0.05           | 0.13         |
| 1884              | 1 | 0.0179565    | 0.02           | 0.14         |
| 1886              | 2 | 0.0359131    | 0.04           | 0.18         |
| 1887              | 1 | 0.0179565    | 0.02           | 0.20         |
| 1889              | 2 | 0.0359131    | 0.04           | 0.23         |
| 1892              | 1 | 0.0179565    | 0.02           | 0.25         |
| 1894              | 2 | 0.0359131    | 0.04           | 0.29         |
| 1895              | 2 | 0.0359131    | 0.04           | 0.32         |
| 1896              | 2 | 0.0359131    | 0.04           | 0.36         |
| 1897              | 6 | 0.1077393    | 0.11           | 0.47         |

(continued)

| ausfertigung_jahr | N | exactpercent | roundedpercent | cumulpercent |
|-------------------|---|--------------|----------------|--------------|
| 1898              | 1 | 0.0179565    | 0.02           | 0.48         |
| 1899              | 2 | 0.0359131    | 0.04           | 0.52         |
| 1901              | 2 | 0.0359131    | 0.04           | 0.56         |
| 1903              | 1 | 0.0179565    | 0.02           | 0.57         |
| 1906              | 1 | 0.0179565    | 0.02           | 0.59         |
| 1907              | 1 | 0.0179565    | 0.02           | 0.61         |
| 1908              | 1 | 0.0179565    | 0.02           | 0.63         |
| 1909              | 3 | 0.0538696    | 0.05           | 0.68         |
| 1910              | 2 | 0.0359131    | 0.04           | 0.72         |
| 1911              | 2 | 0.0359131    | 0.04           | 0.75         |
| 1913              | 2 | 0.0359131    | 0.04           | 0.79         |
| 1919              | 4 | 0.0718262    | 0.07           | 0.86         |
| 1920              | 1 | 0.0179565    | 0.02           | 0.88         |
| 1921              | 4 | 0.0718262    | 0.07           | 0.95         |
| 1922              | 7 | 0.1256958    | 0.13           | 1.08         |
| 1923              | 2 | 0.0359131    | 0.04           | 1.11         |
| 1924              | 2 | 0.0359131    | 0.04           | 1.15         |
| 1925              | 2 | 0.0359131    | 0.04           | 1.19         |
| 1926              | 4 | 0.0718262    | 0.07           | 1.26         |
| 1927              | 5 | 0.0897827    | 0.09           | 1.35         |
| 1928              | 3 | 0.0538696    | 0.05           | 1.40         |
| 1929              | 3 | 0.0538696    | 0.05           | 1.45         |
| 1930              | 3 | 0.0538696    | 0.05           | 1.51         |
| 1931              | 5 | 0.0897827    | 0.09           | 1.60         |
| 1933              | 8 | 0.1436524    | 0.14           | 1.74         |
| 1934              | 7 | 0.1256958    | 0.13           | 1.87         |
| 1935              | 3 | 0.0538696    | 0.05           | 1.92         |
| 1936              | 7 | 0.1256958    | 0.13           | 2.05         |
| 1937              | 9 | 0.1616089    | 0.16           | 2.21         |

(continued)

| ausfertigung_jahr | N  | exactpercent | roundedpercent | cumulpercent |
|-------------------|----|--------------|----------------|--------------|
| 1938              | 10 | 0.1795655    | 0.18           | 2.39         |
| 1939              | 6  | 0.1077393    | 0.11           | 2.50         |
| 1940              | 7  | 0.1256958    | 0.13           | 2.62         |
| 1941              | 2  | 0.0359131    | 0.04           | 2.66         |
| 1942              | 2  | 0.0359131    | 0.04           | 2.69         |
| 1943              | 1  | 0.0179565    | 0.02           | 2.71         |
| 1944              | 1  | 0.0179565    | 0.02           | 2.73         |
| 1947              | 1  | 0.0179565    | 0.02           | 2.75         |
| 1948              | 5  | 0.0897827    | 0.09           | 2.84         |
| 1949              | 12 | 0.2154785    | 0.22           | 3.05         |
| 1950              | 21 | 0.3770874    | 0.38           | 3.43         |
| 1951              | 32 | 0.5746094    | 0.57           | 4.00         |
| 1952              | 14 | 0.2513916    | 0.25           | 4.26         |
| 1953              | 41 | 0.7362184    | 0.74           | 4.99         |
| 1954              | 17 | 0.3052613    | 0.31           | 5.30         |
| 1955              | 26 | 0.4668702    | 0.47           | 5.76         |
| 1956              | 33 | 0.5925660    | 0.59           | 6.36         |
| 1957              | 46 | 0.8260011    | 0.83           | 7.18         |
| 1958              | 20 | 0.3591309    | 0.36           | 7.54         |
| 1959              | 27 | 0.4848267    | 0.48           | 8.03         |
| 1960              | 33 | 0.5925660    | 0.59           | 8.62         |
| 1961              | 52 | 0.9337403    | 0.93           | 9.55         |
| 1962              | 31 | 0.5566529    | 0.56           | 10.11        |
| 1963              | 26 | 0.4668702    | 0.47           | 10.58        |
| 1964              | 28 | 0.5027833    | 0.50           | 11.08        |
| 1965              | 45 | 0.8080445    | 0.81           | 11.89        |
| 1966              | 25 | 0.4489136    | 0.45           | 12.34        |
| 1967              | 39 | 0.7003053    | 0.70           | 13.04        |
| 1968              | 41 | 0.7362184    | 0.74           | 13.77        |

(continued)

| ausfertigung_jahr | N   | exactpercent | roundedpercent | cumulpercent |
|-------------------|-----|--------------|----------------|--------------|
| 1969              | 54  | 0.9696534    | 0.97           | 14.74        |
| 1970              | 46  | 0.8260011    | 0.83           | 15.57        |
| 1971              | 54  | 0.9696534    | 0.97           | 16.54        |
| 1972              | 53  | 0.9516969    | 0.95           | 17.49        |
| 1973              | 33  | 0.5925660    | 0.59           | 18.08        |
| 1974              | 54  | 0.9696534    | 0.97           | 19.05        |
| 1975              | 65  | 1.1671754    | 1.17           | 20.22        |
| 1976              | 74  | 1.3287843    | 1.33           | 21.55        |
| 1977              | 59  | 1.0594362    | 1.06           | 22.61        |
| 1978              | 45  | 0.8080445    | 0.81           | 23.42        |
| 1979              | 40  | 0.7182618    | 0.72           | 24.13        |
| 1980              | 55  | 0.9876100    | 0.99           | 25.12        |
| 1981              | 41  | 0.7362184    | 0.74           | 25.86        |
| 1982              | 42  | 0.7541749    | 0.75           | 26.61        |
| 1983              | 29  | 0.5207398    | 0.52           | 27.13        |
| 1984              | 32  | 0.5746094    | 0.57           | 27.71        |
| 1985              | 41  | 0.7362184    | 0.74           | 28.44        |
| 1986              | 52  | 0.9337403    | 0.93           | 29.38        |
| 1987              | 38  | 0.6823487    | 0.68           | 30.06        |
| 1988              | 48  | 0.8619142    | 0.86           | 30.92        |
| 1989              | 50  | 0.8978273    | 0.90           | 31.82        |
| 1990              | 159 | 2.8550907    | 2.86           | 34.67        |
| 1991              | 79  | 1.4185671    | 1.42           | 36.09        |
| 1992              | 79  | 1.4185671    | 1.42           | 37.51        |
| 1993              | 83  | 1.4903932    | 1.49           | 39.00        |
| 1994              | 164 | 2.9448734    | 2.94           | 41.95        |
| 1995              | 70  | 1.2569582    | 1.26           | 43.20        |
| 1996              | 88  | 1.5801760    | 1.58           | 44.78        |
| 1997              | 109 | 1.9572634    | 1.96           | 46.74        |

(continued)

| ausfertigung_jahr | N    | exactpercent | roundedpercent | cumulpercent |
|-------------------|------|--------------|----------------|--------------|
| 1998              | 124  | 2.2266116    | 2.23           | 48.97        |
| 1999              | 79   | 1.4185671    | 1.42           | 50.39        |
| 2000              | 61   | 1.0953493    | 1.10           | 51.48        |
| 2001              | 119  | 2.1368289    | 2.14           | 53.62        |
| 2002              | 153  | 2.7473514    | 2.75           | 56.37        |
| 2003              | 100  | 1.7956545    | 1.80           | 58.16        |
| 2004              | 142  | 2.5498294    | 2.55           | 60.71        |
| 2005              | 136  | 2.4420901    | 2.44           | 63.15        |
| 2006              | 132  | 2.3702640    | 2.37           | 65.52        |
| 2007              | 105  | 1.8854372    | 1.89           | 67.41        |
| 2008              | 130  | 2.3343509    | 2.33           | 69.74        |
| 2009              | 165  | 2.9628300    | 2.96           | 72.71        |
| 2010              | 123  | 2.2086551    | 2.21           | 74.91        |
| 2011              | 122  | 2.1906985    | 2.19           | 77.11        |
| 2012              | 110  | 1.9752200    | 1.98           | 79.08        |
| 2013              | 188  | 3.3758305    | 3.38           | 82.46        |
| 2014              | 103  | 1.8495242    | 1.85           | 84.31        |
| 2015              | 134  | 2.4061771    | 2.41           | 86.71        |
| 2016              | 147  | 2.6396121    | 2.64           | 89.35        |
| 2017              | 182  | 3.2680912    | 3.27           | 92.62        |
| 2018              | 99   | 1.7776980    | 1.78           | 94.40        |
| 2019              | 122  | 2.1906985    | 2.19           | 96.59        |
| 2020              | 190  | 3.4117436    | 3.41           | 100.00       |
| Total             | 5569 | 100.0000000  | 100.00         | 100.00       |

---

Frequency Table for Variable: doi\_concept

---

1 unique value(s) detected.

| doi_concept            | N    | exactpercent | roundedpercent | cumulpercent |
|------------------------|------|--------------|----------------|--------------|
| 10.5281/zenodo.3832111 | 5569 | 100          | 100            | 100          |
| Total                  | 5569 | 100          | 100            | 100          |

---

Frequency Table for Variable: doi\_version

---

1 unique value(s) detected.

| doi_version            | N    | exactpercent | roundedpercent | cumulpercent |
|------------------------|------|--------------|----------------|--------------|
| 10.5281/zenodo.4384771 | 5569 | 100          | 100            | 100          |
| Total                  | 5569 | 100          | 100            | 100          |

---

Frequency Table for Variable: version

---

1 unique value(s) detected.

| version    | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| 2021-01-05 | 5569 | 100          | 100            | 100          |
| Total      | 5569 | 100          | 100            | 100          |



## 9 Frequenztabellen erstellen: XML-Metadaten

### 9.1 Liste zu prüfender Variablen

```
print(vars.freqtable.rechtsakte)
```

```
## [1] "periodikum"      "fundstellentyp"  "check_neuf"  
## [4] "check_aufh"      "check_sonst"     "check_hinweis"  
## [7] "check_stand"     "ausfertigung_jahr" "doi_concept"  
## [10] "doi_version"     "version"
```

### 9.2 Frequenztabellen erstellen

```
prefix <- paste0(datasetname,  
                  "_01_Meta_Frequenztabelle_var-")
```

```
f.fast.freqtable(dt.meta,  
                 varlist = vars.freqtable.rechtsakte,  
                 sumrow = TRUE,  
                 output.list = FALSE,  
                 output.kable = TRUE,  
                 output.csv = TRUE,  
                 outputdir = outputdir,  
                 prefix = prefix)
```

---

Frequency Table for Variable: periodikum

---

36 unique value(s) detected.

| periodikum | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| NA         | 5    | 0.0760225    | 0.08           | 0.08         |
| ABl EU     | 1    | 0.0152045    | 0.02           | 0.09         |
| BAGVBl     | 2    | 0.0304090    | 0.03           | 0.12         |
| BAnz       | 248  | 3.7707161    | 3.77           | 3.89         |
| BGBI       | 23   | 0.3497035    | 0.35           | 4.24         |
| BGBI I     | 4807 | 73.0880341   | 73.09          | 77.33        |
| BGBI II    | 1186 | 18.0325376   | 18.03          | 95.36        |

(continued)

| periodikum | N  | exactpercent | roundedpercent | cumulpercent |
|------------|----|--------------|----------------|--------------|
| BGBI III   | 1  | 0.0152045    | 0.02           | 95.38        |
| BremGBI    | 4  | 0.0608180    | 0.06           | 95.44        |
| GBI BW     | 1  | 0.0152045    | 0.02           | 95.45        |
| GBI DDR    | 15 | 0.2280675    | 0.23           | 95.68        |
| GBI DDR I  | 45 | 0.6842025    | 0.68           | 96.37        |
| GBI DDR II | 3  | 0.0456135    | 0.05           | 96.41        |
| GV NW      | 2  | 0.0304090    | 0.03           | 96.44        |
| GVBI BB    | 1  | 0.0152045    | 0.02           | 96.46        |
| GVBI BE    | 1  | 0.0152045    | 0.02           | 96.47        |
| GVBI BY    | 4  | 0.0608180    | 0.06           | 96.53        |
| GVBI HE    | 5  | 0.0760225    | 0.08           | 96.61        |
| GVBI RP I  | 2  | 0.0304090    | 0.03           | 96.64        |
| Mtbl BAA   | 1  | 0.0152045    | 0.02           | 96.66        |
| NV         | 4  | 0.0608180    | 0.06           | 96.72        |
| Nds GVBI   | 1  | 0.0152045    | 0.02           | 96.73        |
| RAnz       | 2  | 0.0304090    | 0.03           | 96.76        |
| RBesBI     | 1  | 0.0152045    | 0.02           | 96.78        |
| RGBI       | 56 | 0.8514520    | 0.85           | 97.63        |
| RGBI I     | 55 | 0.8362475    | 0.84           | 98.46        |
| RGBI II    | 48 | 0.7298160    | 0.73           | 99.19        |
| RMBI       | 4  | 0.0608180    | 0.06           | 99.25        |
| RegBI WB   | 4  | 0.0608180    | 0.06           | 99.32        |
| RegBI WH   | 1  | 0.0152045    | 0.02           | 99.33        |
| VOBI BrZ   | 3  | 0.0456135    | 0.05           | 99.38        |
| VkBI       | 22 | 0.3344990    | 0.33           | 99.71        |
| WiGBI      | 7  | 0.1064315    | 0.11           | 99.82        |
| ZBI        | 2  | 0.0304090    | 0.03           | 99.85        |
| eBAnz      | 6  | 0.0912270    | 0.09           | 99.94        |
| Öff Anz    | 4  | 0.0608180    | 0.06           | 100.00       |

(continued)

| periodikum | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| Total      | 6577 | 100.0000000  | 100.00         | 100.00       |

Frequency Table for Variable: fundstellentyp

2 unique value(s) detected.

| fundstellentyp | N    | exactpercent | roundedpercent | cumulpercent |
|----------------|------|--------------|----------------|--------------|
| amtlich        | 6572 | 99.9239775   | 99.92          | 99.92        |
| nichtamtlich   | 5    | 0.0760225    | 0.08           | 100.00       |
| Total          | 6577 | 100.0000000  | 100.00         | 100.00       |

Frequency Table for Variable: check\_neuf

2 unique value(s) detected.

| check_neuf | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| NA         | 6158 | 93.629314    | 93.63          | 93.63        |
| ja         | 419  | 6.370686     | 6.37           | 100.00       |
| Total      | 6577 | 100.000000   | 100.00         | 100.00       |

Frequency Table for Variable: check\_aufh

2 unique value(s) detected.

| check_aufh | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| NA         | 6412 | 97.491257    | 97.49          | 97.49        |
| ja         | 165  | 2.508743     | 2.51           | 100.00       |
| Total      | 6577 | 100.000000   | 100.00         | 100.00       |

---

Frequency Table for Variable: check\_sonst

---

2 unique value(s) detected.

| check_sonst | N    | exactpercent | roundedpercent | cumulpercent |
|-------------|------|--------------|----------------|--------------|
| NA          | 6109 | 92.884294    | 92.88          | 92.88        |
| ja          | 468  | 7.115706     | 7.12           | 100.00       |
| Total       | 6577 | 100.000000   | 100.00         | 100.00       |

---

Frequency Table for Variable: check\_hinweis

---

2 unique value(s) detected.

| check_hinweis | N    | exactpercent | roundedpercent | cumulpercent |
|---------------|------|--------------|----------------|--------------|
| NA            | 6294 | 95.697126    | 95.7           | 95.7         |
| ja            | 283  | 4.302874     | 4.3            | 100.0        |
| Total         | 6577 | 100.000000   | 100.0          | 100.0        |

---

Frequency Table for Variable: check\_stand

---

2 unique value(s) detected.

| check_stand | N    | exactpercent | roundedpercent | cumulpercent |
|-------------|------|--------------|----------------|--------------|
| NA          | 3812 | 57.95956     | 57.96          | 57.96        |
| ja          | 2765 | 42.04044     | 42.04          | 100.00       |
| Total       | 6577 | 100.00000    | 100.00         | 100.00       |

---

Frequency Table for Variable: ausfertigung\_jahr

---

125 unique value(s) detected.

| ausfertigung_jahr | N | exactpercent | roundedpercent | cumulpercent |
|-------------------|---|--------------|----------------|--------------|
| NA                | 1 | 0.0152045    | 0.02           | 0.02         |
| 1869              | 1 | 0.0152045    | 0.02           | 0.03         |
| 1871              | 3 | 0.0456135    | 0.05           | 0.08         |
| 1877              | 3 | 0.0456135    | 0.05           | 0.12         |
| 1884              | 1 | 0.0152045    | 0.02           | 0.14         |
| 1886              | 2 | 0.0304090    | 0.03           | 0.17         |
| 1887              | 1 | 0.0152045    | 0.02           | 0.18         |
| 1889              | 2 | 0.0304090    | 0.03           | 0.21         |
| 1892              | 1 | 0.0152045    | 0.02           | 0.23         |
| 1894              | 2 | 0.0304090    | 0.03           | 0.26         |
| 1895              | 2 | 0.0304090    | 0.03           | 0.29         |
| 1896              | 2 | 0.0304090    | 0.03           | 0.32         |
| 1897              | 6 | 0.0912270    | 0.09           | 0.41         |
| 1898              | 1 | 0.0152045    | 0.02           | 0.43         |
| 1899              | 2 | 0.0304090    | 0.03           | 0.46         |
| 1901              | 2 | 0.0304090    | 0.03           | 0.49         |
| 1903              | 1 | 0.0152045    | 0.02           | 0.50         |
| 1906              | 1 | 0.0152045    | 0.02           | 0.52         |
| 1907              | 1 | 0.0152045    | 0.02           | 0.53         |
| 1908              | 1 | 0.0152045    | 0.02           | 0.55         |
| 1909              | 4 | 0.0608180    | 0.06           | 0.61         |
| 1910              | 2 | 0.0304090    | 0.03           | 0.64         |
| 1911              | 3 | 0.0456135    | 0.05           | 0.68         |
| 1913              | 2 | 0.0304090    | 0.03           | 0.71         |
| 1919              | 7 | 0.1064315    | 0.11           | 0.82         |
| 1920              | 1 | 0.0152045    | 0.02           | 0.84         |
| 1921              | 4 | 0.0608180    | 0.06           | 0.90         |
| 1922              | 7 | 0.1064315    | 0.11           | 1.00         |
| 1923              | 2 | 0.0304090    | 0.03           | 1.03         |
| 1924              | 2 | 0.0304090    | 0.03           | 1.06         |

(continued)

| ausfertigung_jahr | N  | exactpercent | roundedpercent | cumulpercent |
|-------------------|----|--------------|----------------|--------------|
| 1925              | 2  | 0.0304090    | 0.03           | 1.09         |
| 1926              | 4  | 0.0608180    | 0.06           | 1.16         |
| 1927              | 5  | 0.0760225    | 0.08           | 1.23         |
| 1928              | 4  | 0.0608180    | 0.06           | 1.29         |
| 1929              | 4  | 0.0608180    | 0.06           | 1.35         |
| 1930              | 3  | 0.0456135    | 0.05           | 1.40         |
| 1931              | 5  | 0.0760225    | 0.08           | 1.47         |
| 1933              | 9  | 0.1368405    | 0.14           | 1.61         |
| 1934              | 7  | 0.1064315    | 0.11           | 1.72         |
| 1935              | 3  | 0.0456135    | 0.05           | 1.76         |
| 1936              | 9  | 0.1368405    | 0.14           | 1.90         |
| 1937              | 9  | 0.1368405    | 0.14           | 2.04         |
| 1938              | 11 | 0.1672495    | 0.17           | 2.20         |
| 1939              | 11 | 0.1672495    | 0.17           | 2.37         |
| 1940              | 8  | 0.1216360    | 0.12           | 2.49         |
| 1941              | 2  | 0.0304090    | 0.03           | 2.52         |
| 1942              | 2  | 0.0304090    | 0.03           | 2.55         |
| 1943              | 1  | 0.0152045    | 0.02           | 2.57         |
| 1944              | 4  | 0.0608180    | 0.06           | 2.63         |
| 1945              | 1  | 0.0152045    | 0.02           | 2.65         |
| 1946              | 12 | 0.1824540    | 0.18           | 2.83         |
| 1947              | 11 | 0.1672495    | 0.17           | 3.00         |
| 1948              | 13 | 0.1976585    | 0.20           | 3.19         |
| 1949              | 18 | 0.2736810    | 0.27           | 3.47         |
| 1950              | 23 | 0.3497035    | 0.35           | 3.82         |
| 1951              | 41 | 0.6233845    | 0.62           | 4.44         |
| 1952              | 24 | 0.3649080    | 0.36           | 4.80         |
| 1953              | 47 | 0.7146115    | 0.71           | 5.52         |
| 1954              | 26 | 0.3953170    | 0.40           | 5.91         |

(continued)

| ausfertigung_jahr | N  | exactpercent | roundedpercent | cumulpercent |
|-------------------|----|--------------|----------------|--------------|
| 1955              | 34 | 0.5169530    | 0.52           | 6.43         |
| 1956              | 54 | 0.8210430    | 0.82           | 7.25         |
| 1957              | 56 | 0.8514520    | 0.85           | 8.10         |
| 1958              | 36 | 0.5473620    | 0.55           | 8.65         |
| 1959              | 48 | 0.7298160    | 0.73           | 9.38         |
| 1960              | 47 | 0.7146115    | 0.71           | 10.10        |
| 1961              | 71 | 1.0795195    | 1.08           | 11.18        |
| 1962              | 42 | 0.6385890    | 0.64           | 11.81        |
| 1963              | 37 | 0.5625665    | 0.56           | 12.38        |
| 1964              | 42 | 0.6385890    | 0.64           | 13.02        |
| 1965              | 62 | 0.9426790    | 0.94           | 13.96        |
| 1966              | 38 | 0.5777710    | 0.58           | 14.54        |
| 1967              | 50 | 0.7602250    | 0.76           | 15.30        |
| 1968              | 50 | 0.7602250    | 0.76           | 16.06        |
| 1969              | 71 | 1.0795195    | 1.08           | 17.14        |
| 1970              | 58 | 0.8818610    | 0.88           | 18.02        |
| 1971              | 67 | 1.0187015    | 1.02           | 19.04        |
| 1972              | 66 | 1.0034970    | 1.00           | 20.04        |
| 1973              | 69 | 1.0491105    | 1.05           | 21.09        |
| 1974              | 63 | 0.9578835    | 0.96           | 22.05        |
| 1975              | 75 | 1.1403375    | 1.14           | 23.19        |
| 1976              | 83 | 1.2619735    | 1.26           | 24.45        |
| 1977              | 72 | 1.0947240    | 1.09           | 25.54        |
| 1978              | 58 | 0.8818610    | 0.88           | 26.43        |
| 1979              | 57 | 0.8666565    | 0.87           | 27.29        |
| 1980              | 70 | 1.0643150    | 1.06           | 28.36        |
| 1981              | 48 | 0.7298160    | 0.73           | 29.09        |
| 1982              | 47 | 0.7146115    | 0.71           | 29.80        |
| 1983              | 40 | 0.6081800    | 0.61           | 30.41        |

(continued)

| ausfertigung_jahr | N   | exactpercent | roundedpercent | cumulpercent |
|-------------------|-----|--------------|----------------|--------------|
| 1984              | 53  | 0.8058385    | 0.81           | 31.21        |
| 1985              | 53  | 0.8058385    | 0.81           | 32.02        |
| 1986              | 75  | 1.1403375    | 1.14           | 33.16        |
| 1987              | 46  | 0.6994070    | 0.70           | 33.86        |
| 1988              | 63  | 0.9578835    | 0.96           | 34.82        |
| 1989              | 59  | 0.8970655    | 0.90           | 35.72        |
| 1990              | 173 | 2.6303786    | 2.63           | 38.35        |
| 1991              | 87  | 1.3227915    | 1.32           | 39.67        |
| 1992              | 95  | 1.4444276    | 1.44           | 41.11        |
| 1993              | 100 | 1.5204501    | 1.52           | 42.63        |
| 1994              | 190 | 2.8888551    | 2.89           | 45.52        |
| 1995              | 113 | 1.7181086    | 1.72           | 47.24        |
| 1996              | 108 | 1.6420861    | 1.64           | 48.88        |
| 1997              | 140 | 2.1286301    | 2.13           | 51.01        |
| 1998              | 139 | 2.1134256    | 2.11           | 53.12        |
| 1999              | 95  | 1.4444276    | 1.44           | 54.57        |
| 2000              | 73  | 1.1099285    | 1.11           | 55.68        |
| 2001              | 135 | 2.0526076    | 2.05           | 57.73        |
| 2002              | 165 | 2.5087426    | 2.51           | 60.24        |
| 2003              | 119 | 1.8093356    | 1.81           | 62.05        |
| 2004              | 155 | 2.3566976    | 2.36           | 64.41        |
| 2005              | 158 | 2.4023111    | 2.40           | 66.81        |
| 2006              | 141 | 2.1438346    | 2.14           | 68.95        |
| 2007              | 119 | 1.8093356    | 1.81           | 70.76        |
| 2008              | 154 | 2.3414931    | 2.34           | 73.10        |
| 2009              | 178 | 2.7064011    | 2.71           | 75.81        |
| 2010              | 140 | 2.1286301    | 2.13           | 77.94        |
| 2011              | 135 | 2.0526076    | 2.05           | 79.99        |
| 2012              | 125 | 1.9005626    | 1.90           | 81.89        |



(continued)

| ausfertigung_jahr | N    | exactpercent | roundedpercent | cumulpercent |
|-------------------|------|--------------|----------------|--------------|
| 2013              | 192  | 2.9192641    | 2.92           | 84.81        |
| 2014              | 104  | 1.5812681    | 1.58           | 86.39        |
| 2015              | 136  | 2.0678121    | 2.07           | 88.46        |
| 2016              | 152  | 2.3110841    | 2.31           | 90.77        |
| 2017              | 185  | 2.8128326    | 2.81           | 93.58        |
| 2018              | 101  | 1.5356546    | 1.54           | 95.12        |
| 2019              | 126  | 1.9157671    | 1.92           | 97.04        |
| 2020              | 195  | 2.9648776    | 2.96           | 100.00       |
| Total             | 6577 | 100.0000000  | 100.00         | 100.00       |

Frequency Table for Variable: doi\_concept

1 unique value(s) detected.

| doi_concept            | N    | exactpercent | roundedpercent | cumulpercent |
|------------------------|------|--------------|----------------|--------------|
| 10.5281/zenodo.3832111 | 6577 | 100          | 100            | 100          |
| Total                  | 6577 | 100          | 100            | 100          |

Frequency Table for Variable: doi\_version

1 unique value(s) detected.

| doi_version            | N    | exactpercent | roundedpercent | cumulpercent |
|------------------------|------|--------------|----------------|--------------|
| 10.5281/zenodo.4384771 | 6577 | 100          | 100            | 100          |
| Total                  | 6577 | 100          | 100            | 100          |

Frequency Table for Variable: version

1 unique value(s) detected.

| version    | N    | exactpercent | roundedpercent | cumulpercent |
|------------|------|--------------|----------------|--------------|
| 2021-01-05 | 6577 | 100          | 100            | 100          |
| Total      | 6577 | 100          | 100            | 100          |

## 10 Frequenztabellen visualisieren

### 10.1 Präfixe erstellen

```
prefix.normen <- paste0("ANALYSE/",  
                        datasetname,  
                        "_01_Einzelnormen_Frequenztafel_var-")  
  
prefix.rechtsakte <- paste0("ANALYSE/",  
                            datasetname,  
                            "_01_Rechtsakte_Frequenztafel_var-")  
  
prefix.meta <- paste0("ANALYSE/",  
                      datasetname,  
                      "_01_Meta_Frequenztafel_var-")
```

### 10.2 Tabellen für Einzelnormen einlesen

```
table.normen.periodikum <- fread(paste0(prefix.normen,  
                                         "periodikum.csv"))  
  
table.normen.ausjahr <- fread(paste0(prefix.normen,  
                                       "ausfertigung_jahr.csv"))
```

### 10.3 Tabellen für Rechtsakte einlesen

```
table.rechtsakte.periodikum <- fread(paste0(prefix.rechtsakte,  
                                              "periodikum.csv"))  
  
table.rechtsakte.ausjahr <- fread(paste0(prefix.rechtsakte,  
                                           "ausfertigung_jahr.csv"))
```

### 10.4 Tabellen für XML-Metadaten einlesen

```
table.meta.periodikum <- fread(paste0(prefix.meta,  
                                       "periodikum.csv"))  
  
table.meta.ausjahr <- fread(paste0(prefix.meta,  
                                    "ausfertigung_jahr.csv"))
```

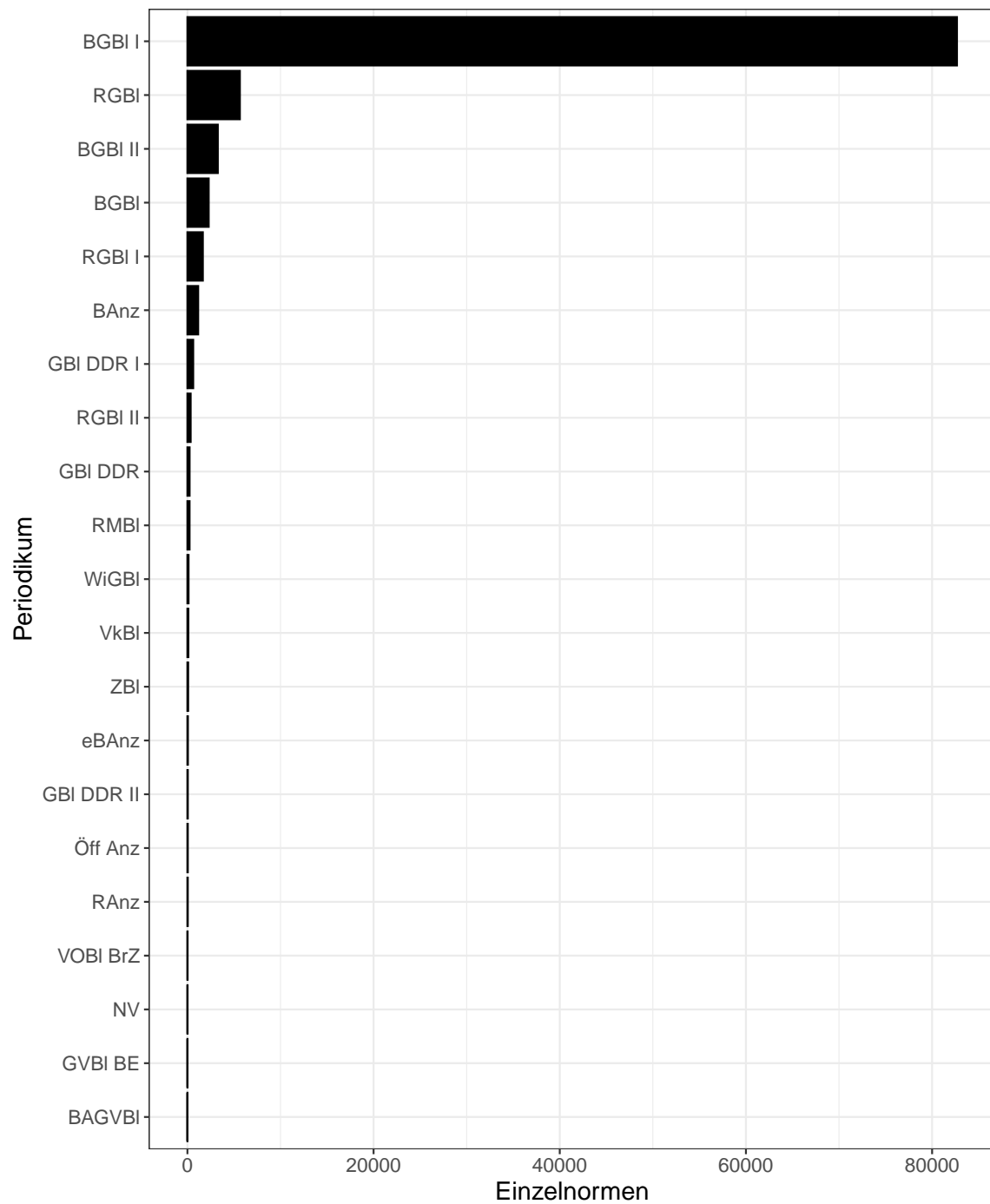
## 10.5 Periodikum

### 10.5.1 Einzelnormen

```
freqtable <- table.normen.periodikum[-.N]
```

```
ggplot(data = freqtable)+  
  geom_bar(aes(x = reorder(periodikum,  
                           N),  
              y = N),  
          stat = "identity",  
          fill = "black",  
          color = "black")+  
  coord_flip()+  
  theme_bw()+  
  labs(  
    title = paste(datasetname,  
                  "| Version",  
                  datestamp,  
                  "| Einzelnormen je Periodikum"),  
    caption = paste("DOI:",  
                    doi.version),  
    x = "Periodikum",  
    y = "Einzelnormen"  
  )+  
  theme(  
    text = element_text(size = 14),  
    plot.title = element_text(size = 14, face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```

# **C-DBR | Version 2021-01-05 | Einzelnormen je Periodikum**



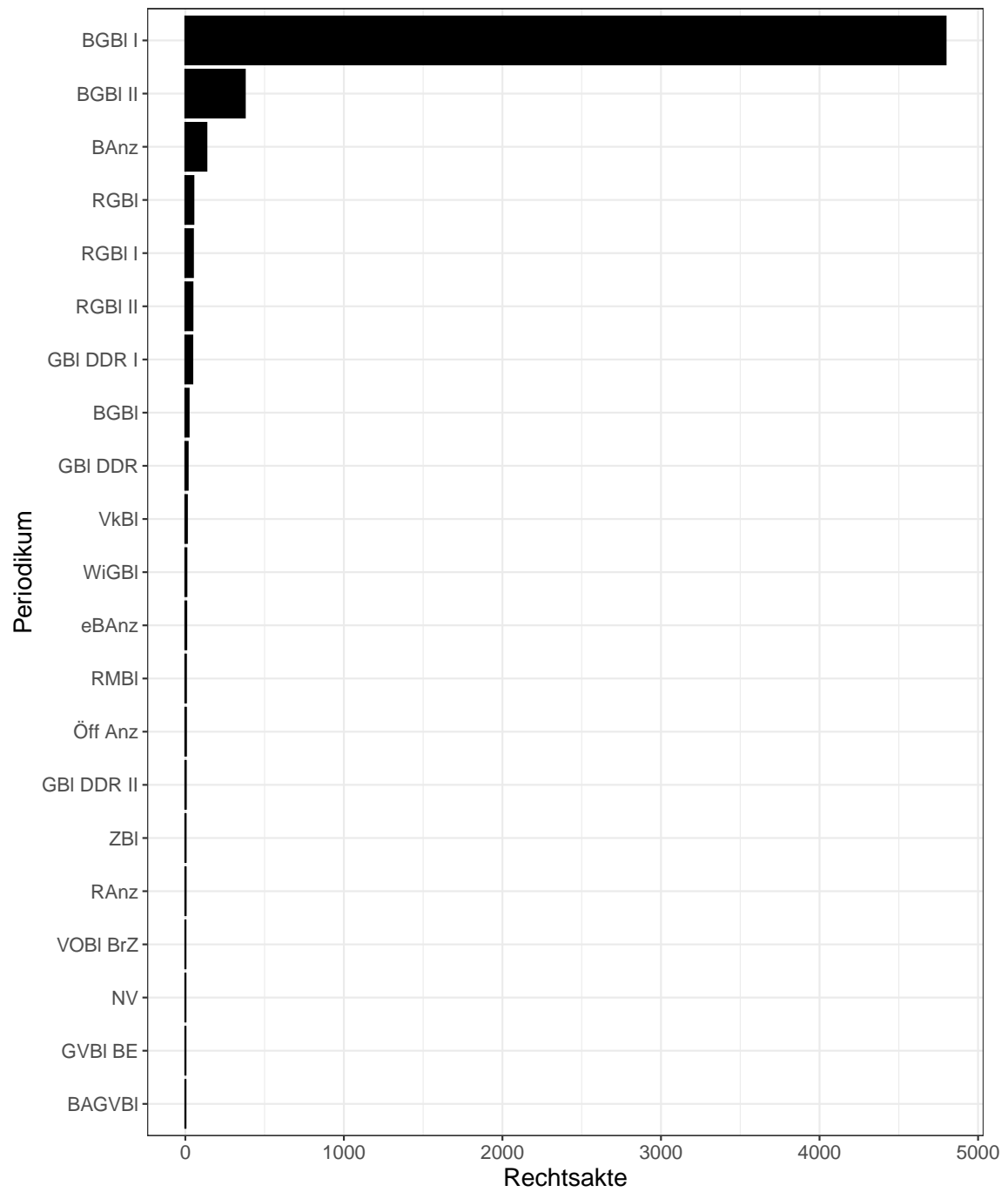
DOI: 10.5281/zenodo.4384771

### 10.5.2 Rechtsakte

```
freqtable <- table.rechtsakte.periodikum[-.N]
```

```
ggplot(data = freqtable) +  
  geom_bar(aes(x = reorder(periodikum,  
                           N),  
              y = N),  
          stat = "identity",  
          fill = "black",  
          color = "black") +  
  coord_flip() +  
  theme_bw() +  
  labs(  
    title = paste(datasetname,  
                  "| Version",  
                  datestamp,  
                  "| Rechtsakte mit Inhalt je Periodikum"),  
    caption = paste("DOI:",  
                   doi.version),  
    x = "Periodikum",  
    y = "Rechtsakte"  
  ) +  
  theme(  
    text = element_text(size = 14),  
    plot.title = element_text(size = 14, face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```

# C-DBR | Version 2021-01-05 | Rechtsakte mit Inhalt je Periodikum



DOI: 10.5281/zenodo.4384771

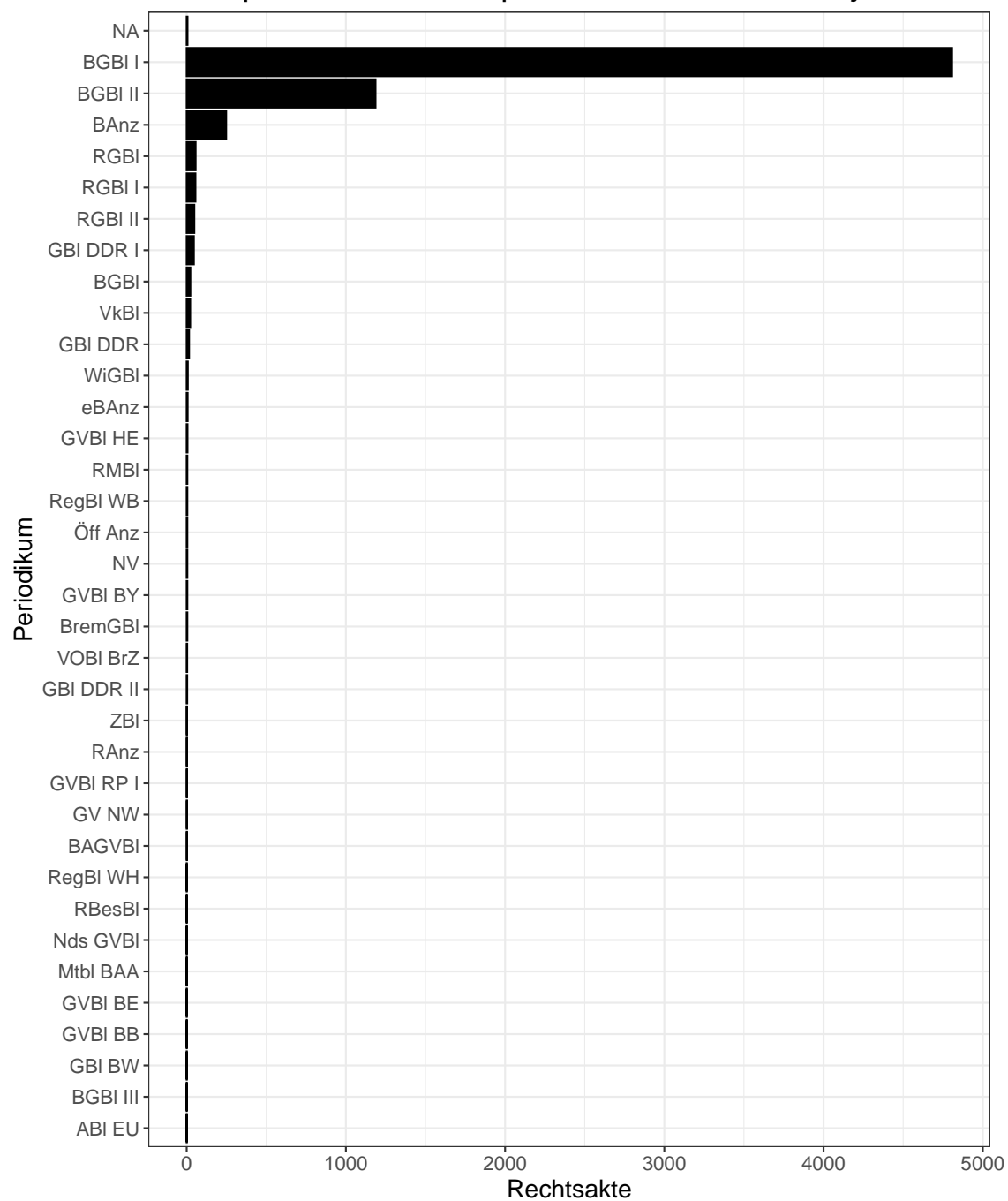
### 10.5.3 XML-Metadaten

```
frequetable <- table.meta.periodikum[-.N]
```

```
ggplot(data = frequetable) +  
  geom_bar(aes(x = reorder(periodikum,  
                           N),  
              y = N),  
          stat = "identity",  
          fill = "black",  
          color = "black") +  
  coord_flip() +  
  theme_bw() +  
  labs(  
    title = paste(datasetname,  
                  "| Version",  
                  datestamp,  
                  "| Rechtsakte nach Metadaten je Periodikum"),  
    caption = paste("DOI:",  
                    doi.version),  
    x = "Periodikum",  
    y = "Rechtsakte"  
  ) +  
  theme(  
    text = element_text(size = 14),  
    plot.title = element_text(size = 14, face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```



# C-DBR | Version 2021-01-05 | Rechtsakte nach Metadaten je Periodikum



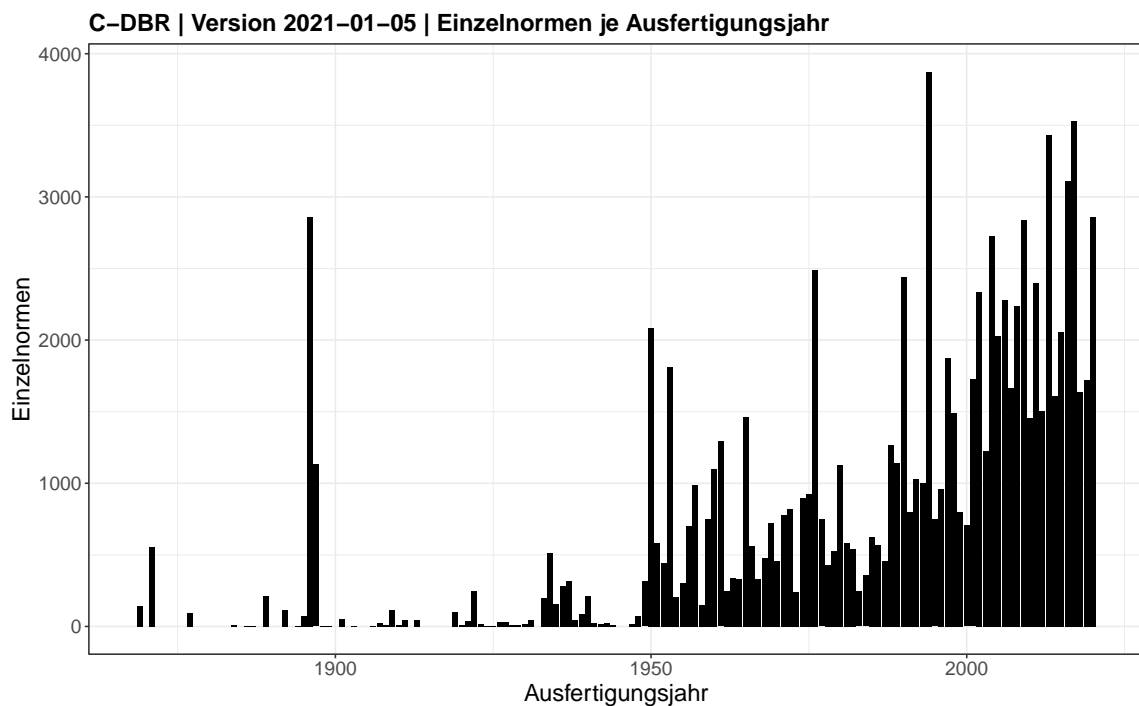
DOI: 10.5281/zenodo.4384771

## 10.6 Ausfertigungsjahr

### 10.6.1 Einzelnormen

```
frequetable <- table.normen.ausjahr[-.N][,lapply(.SD, as.numeric)]
```

```
ggplot(data = frequetable) +  
  geom_bar(aes(x = ausfertigung_jahr,  
               y = N),  
           stat = "identity",  
           fill = "black")+  
  theme_bw()+  
  labs(  
    title = paste(datasetname,  
                  "| Version",  
                  datestamp,  
                  "| Einzelnormen je Ausfertigungsjahr"),  
    caption = paste("DOI:",  
                    doi.version),  
    x = "Ausfertigungsjahr",  
    y = "Einzelnormen"  
  )+  
  theme(  
    text = element_text(size = 16),  
    plot.title = element_text(size = 16, face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```

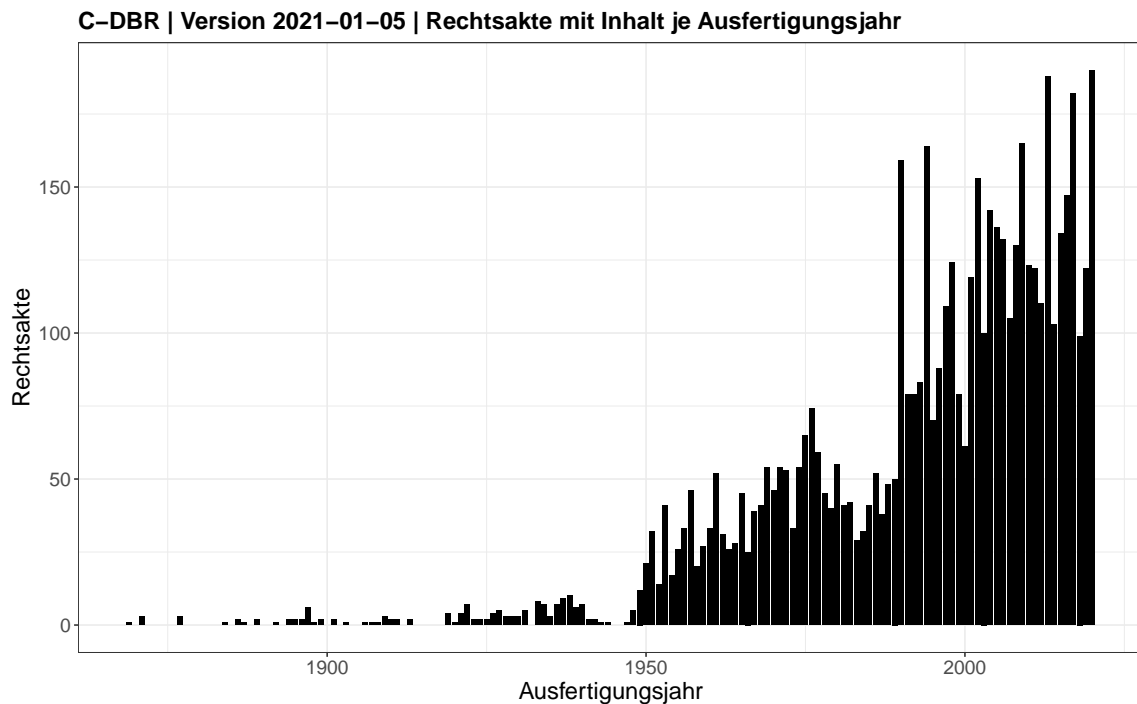


DOI: 10.5281/zenodo.4384771

## 10.6.2 Rechtsakte

```
freqtable <- table.rechtsakte.ausjahr[-.N][,lapply(.SD, as.numeric)]
```

```
ggplot(data = freqtable) +  
  geom_bar(aes(x = ausfertigung_jahr,  
               y = N),  
           stat = "identity",  
           fill = "black") +  
  theme_bw() +  
  labs(  
    title = paste(datasetname,  
                  "| Version",  
                  datestamp,  
                  "| Rechtsakte mit Inhalt je Ausfertigungsjahr"),  
    caption = paste("DOI:",  
                    doi.version),  
    x = "Ausfertigungsjahr",  
    y = "Rechtsakte"  
  ) +  
  theme(  
    text = element_text(size = 16),  
    plot.title = element_text(size = 16, face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```

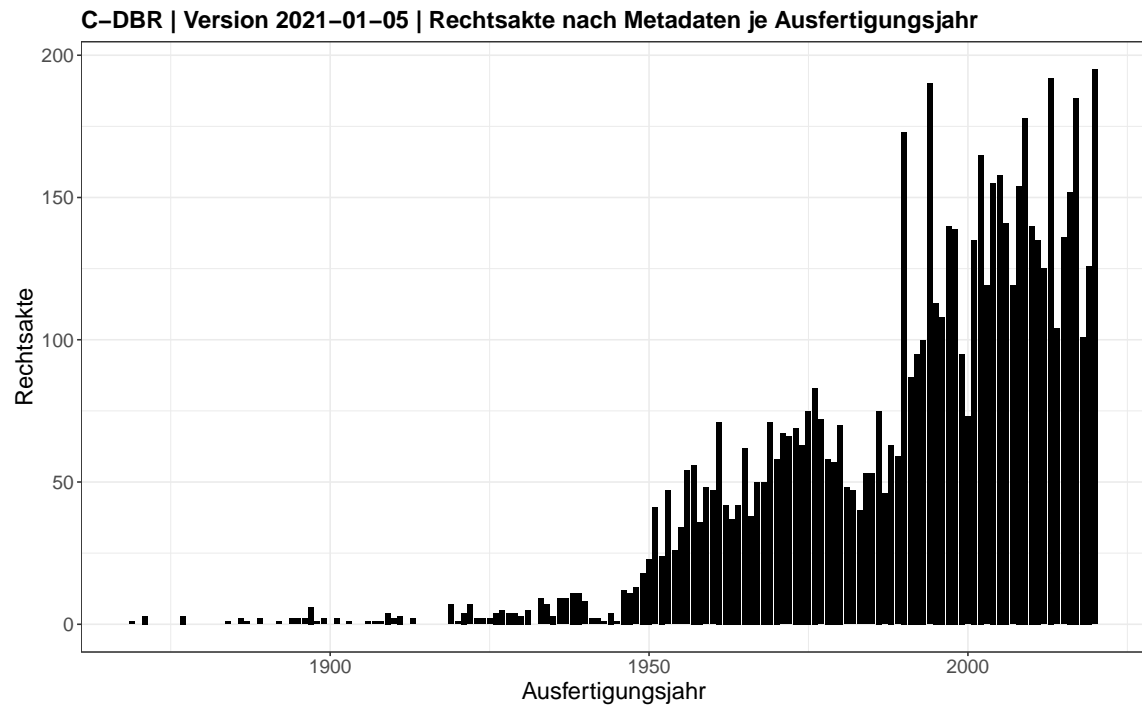


### 10.6.3 XML-Metadaten

```
freqtable <- table.meta.ausjahr[-.N][,lapply(.SD, as.numeric)]
```

```
ggplot(data = freqtable) +  
  geom_bar(aes(x = ausfertigung_jahr,  
               y = N),  
           stat = "identity",  
           fill = "black") +  
  theme_bw() +  
  labs(  
    title = paste(datasetname,  
                  "| Version",  
                  datestamp,  
                  "| Rechtsakte nach Metadaten je Ausfertigungsjahr"),  
    caption = paste("DOI:",  
                    doi.version),  
    x = "Ausfertigungsjahr",  
    y = "Rechtsakte"  
  ) +  
  theme(  
    text = element_text(size = 16),  
    plot.title = element_text(size = 16, face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```

```
## Warning: Removed 1 rows containing missing values (position_stack).
```



DOI: 10.5281/zenodo.4384771

## 11 Korpus-Analytik

### 11.1 Quanteda-Korpora erstellen

```
corpus.normen <- corpus(dt.normen)
corpus.rechtsakte <- corpus(dt.rechtsakte)
```

### 11.2 Berechnung linguistischer Kennwerte

An dieser Stelle werden für jedes Dokument die Anzahl Tokens, Typen und Sätze berechnet und mit den jeweiligen Metadaten verknüpft. Das Ergebnis ist grundsätzlich identisch mit dem eigentlichen Datensatz, nur ohne den Text der Entscheidungen.

#### 11.2.1 Einzelnormen

```
scope <- 1:ndoc(corpus.normen)
result <- foreach(i = scope,
                  .errorhandling = 'pass') %dopar% {
  temp <- summary(corpus.normen[i])
  return(temp)
}

summary.corpus.normen <- rbindlist(result)
```

#### 11.2.2 Rechtsakte

```
scope <- 1:ndoc(corpus.rechtsakte)
result <- foreach(i = scope,
                  .errorhandling = 'pass') %dopar% {
  temp <- summary(corpus.rechtsakte[i])
  return(temp)
}

summary.corpus.rechtsakte <- rbindlist(result)
```

## 11.3 Variablen-Namen anpassen

### 11.3.1 Einzelnormen

```
setnames(summary.corpus.normen,  
  old = c("Text",  
          "Tokens",  
          "Types",  
          "Sentences"),  
  new = c("doc_id",  
          "tokens",  
          "typen",  
          "saetze"))
```

### 11.3.2 Rechtsakte

```
setnames(summary.corpus.rechtsakte,  
  old = c("Text",  
          "Tokens",  
          "Types",  
          "Sentences"),  
  new = c("doc_id",  
          "tokens",  
          "typen",  
          "saetze"))
```

## 11.4 Kennwerte den Gesamtkorpora hinzufügen

### 11.4.1 Einzelnormen

```
dt.normen$tokens <- summary.corpus.normen$tokens  
dt.normen$typen <- summary.corpus.normen$typen  
dt.normen$saetze <- summary.corpus.normen$saetze
```

### 11.4.2 Rechtsakte

```
dt.rechtsakte$tokens <- summary.corpus.rechtsakte$tokens  
dt.rechtsakte$typen <- summary.corpus.rechtsakte$typen  
dt.rechtsakte$saetze <- summary.corpus.rechtsakte$saetze
```

## 11.5 Anzahl Variablen der Datensätze

```
length(dt.normen)
```

```
## [1] 34
```

```
length(dt.rechtsakte)
```

```
## [1] 27
```

```
length(dt.meta)
```

```
## [1] 25
```

## 11.6 Alle Variablen-Namen der Datensätze

```
names(dt.normen)
```

```
## [1] "jurabk"          "amtabk"          "ausfertigung_datum"
## [4] "periodikum"      "zitstelle"       "langue"
## [7] "kurzue"          "fundstellentyp"  "dateiname"
## [10] "builddate_original" "gliederungskennzahl" "gliederungsbez"
## [13] "gliederungstitel" "enbez"           "text"
## [16] "stand"           "check_stand"     "aufh"
## [19] "check_aufh"      "neuf"            "check_neuf"
## [22] "sonst"           "check_sonst"     "hinweis"
## [25] "check_hinweis"   "doc_id"          "builddate_iso"
## [28] "ausfertigung_jahr" "doi_concept"     "doi_version"
## [31] "version"         "tokens"          "typen"
## [34] "saetze"
```

```
names(dt.rechtsakte)
```

```
## [1] "doc_id"          "text"            "jurabk"
## [4] "amtabk"          "ausfertigung_datum" "periodikum"
## [7] "zitstelle"       "langue"          "kurzue"
## [10] "stand"           "aufh"            "neuf"
## [13] "hinweis"         "sonst"           "check_stand"
## [16] "check_aufh"      "check_neuf"      "check_hinweis"
## [19] "check_sonst"     "fundstellentyp"  "ausfertigung_jahr"
## [22] "doi_concept"     "doi_version"     "version"
## [25] "tokens"          "typen"           "saetze"
```



```
names(dt.meta)
```

```
## [1] "jurabk"      "amtabk"      "ausfertigung_datum"
## [4] "periodikum"  "zitstelle"   "langue"
## [7] "kurzue"      "fundstellentyp" "doc_id"
## [10] "bulldate_original" "stand"      "check_stand"
## [13] "aufh"        "check_aufh"  "neuf"
## [16] "check_neuf"   "sonst"       "check_sonst"
## [19] "hinweis"      "check_hinweis" "bulldate_iso"
## [22] "ausfertigung_jahr" "doi_concept"  "doi_version"
## [25] "version"
```

## 11.7 Quantitative Variablen

### 11.7.1 Ausfertigungsdatum

#### *Einzelnormen*

```
summary(as.IDate(dt.normen$ausfertigung_datum))
```

```
##           Min.          1st Qu.          Median          Mean          3rd Qu.          Max.
## "1869-06-21" "1972-01-15" "1996-12-16" "1987-04-04" "2010-08-12" "2020-12-29"
```

#### *Rechtsakte*

```
summary(as.IDate(dt.rechtsakte$ausfertigung_datum))
```

```
##           Min.          1st Qu.          Median          Mean          3rd Qu.          Max.
## "1869-06-21" "1980-10-28" "1999-10-19" "1994-03-22" "2011-01-24" "2020-12-29"
```

#### *XML-Metadaten*

```
summary(as.IDate(dt.meta$ausfertigung_datum))
```

```
##           Min.          1st Qu.          Median          Mean          3rd Qu.          Max.
## "1869-06-21" "1977-07-20" "1997-07-26" "1992-06-04" "2009-08-10" "2020-12-29"
##           NA's
##           "1"
```

## 11.7.2 Ausfertigungsjahr

### *Einzelnormen*

```
summary(dt.normen$ausfertigung_jahr)
```

|    |      |         |        |      |         |      |
|----|------|---------|--------|------|---------|------|
| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| ## | 1869 | 1972    | 1996   | 1987 | 2010    | 2020 |

### *Rechtsakte*

```
summary(dt.rechtsakte$ausfertigung_jahr)
```

|    |      |         |        |      |         |      |
|----|------|---------|--------|------|---------|------|
| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| ## | 1869 | 1980    | 1999   | 1994 | 2011    | 2020 |

### *XML-Metadaten*

```
summary(dt.meta$ausfertigung_jahr)
```

|    |      |         |        |      |         |      |      |
|----|------|---------|--------|------|---------|------|------|
| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | NA's |
| ## | 1869 | 1977    | 1997   | 1992 | 2009    | 2020 | 1    |

## 11.8 Linguistische Kennwerte: Einzelnormen

**Hinweis:** Typen sind definiert als einzigartige Tokens und werden für jedes Dokument gesondert berechnet. Daher ergibt es an dieser Stelle auch keinen Sinn die Typen zu summieren, denn bezogen auf den Gesamt-Korpus wäre der Kennwert ein anderer. Der Wert wird daher manuell auf »NA« gesetzt.

### 11.8.1 Zusammenfassungen berechnen

```
dt.summary.ling <- summary.corpus.normen[, lapply(.SD,
                                                function(x)unclass(summary(x))),
                                           .SDcols = c("tokens",
                                                       "saetze",
                                                       "typen")]

dt.sums.ling <- summary.corpus.normen[,
                                       lapply(.SD, sum),
                                       .SDcols = c("tokens",
                                                   "saetze",
                                                   "typen")]

dt.sums.ling$typen <- NA

dt.stats.ling <- rbind(dt.sums.ling, dt.summary.ling)
dt.stats.ling <- transpose(dt.stats.ling, keep.names = "names")

setnames(dt.stats.ling, c("Variable",
                          "Sum",
                          "Min",
                          "Quart1",
                          "Median",
                          "Mean",
                          "Quart3",
                          "Max"))
```

### 11.8.2 Zusammenfassungen anzeigen

```
kable(dt.stats.ling,  
      format.args = list(big.mark = ","),  
      format = "latex",  
      booktabs=TRUE,  
      longtable=TRUE)
```

| Variable | Sum        | Min | Quart1 | Median | Mean       | Quart3 | Max    |
|----------|------------|-----|--------|--------|------------|--------|--------|
| tokens   | 21,446,633 | 1   | 46     | 107    | 217.454150 | 230    | 56,584 |
| saetze   | 724,389    | 1   | 2      | 4      | 7.344808   | 8      | 2,381  |
| typen    | NA         | 1   | 35     | 66     | 91.997932  | 113    | 11,096 |

### 11.8.3 Zusammenfassungen speichern

```
fwrite(dt.stats.ling,  
       paste0(outputdir,  
               datasetname,  
               "_00_Einzelnormen_KorpusStatistik_ZusammenfassungLinguistisch.csv"),  
       ,  
       na = "NA")
```

## 11.9 Linguistische Kennwerte: Rechtsakte

**Hinweis:** Typen sind definiert als einzigartige Tokens und werden für jedes Dokument gesondert berechnet. Daher ergibt es an dieser Stelle auch keinen Sinn die Typen zu summieren, denn bezogen auf den Gesamt-Korpus wäre der Kennwert ein anderer. Der Wert wird daher manuell auf »NA« gesetzt.

### 11.9.1 Zusammenfassungen berechnen

```
dt.summary.ling <- summary.corpus.rechtsakte[, lapply(.SD,
                                                    function(x)unclass(summary(x))),
                                                    .SDcols = c("tokens",
                                                                "saetze",
                                                                "typen")]

dt.sums.ling <- summary.corpus.rechtsakte[,
                                           lapply(.SD, sum),
                                           .SDcols = c("tokens",
                                                       "saetze",
                                                       "typen")]

dt.sums.ling$typen <- NA

dt.stats.ling <- rbind(dt.sums.ling, dt.summary.ling)
dt.stats.ling <- transpose(dt.stats.ling, keep.names = "names")

setnames(dt.stats.ling, c("Variable",
                          "Sum",
                          "Min",
                          "Quart1",
                          "Median",
                          "Mean",
                          "Quart3",
                          "Max"))
```

### 11.9.2 Zusammenfassungen anzeigen

```
kable(dt.stats.ling,  
      format.args = list(big.mark = ","),  
      format = "latex",  
      booktabs=TRUE,  
      longtable=TRUE)
```

| Variable | Sum        | Min | Quart1 | Median | Mean       | Quart3 | Max     |
|----------|------------|-----|--------|--------|------------|--------|---------|
| tokens   | 21,446,633 | 17  | 263    | 923    | 3,851.0743 | 3,378  | 273,108 |
| saetze   | 712,908    | 1   | 13     | 36     | 128.0136   | 114    | 14,306  |
| typen    | NA         | 15  | 129    | 303    | 633.4236   | 803    | 14,622  |

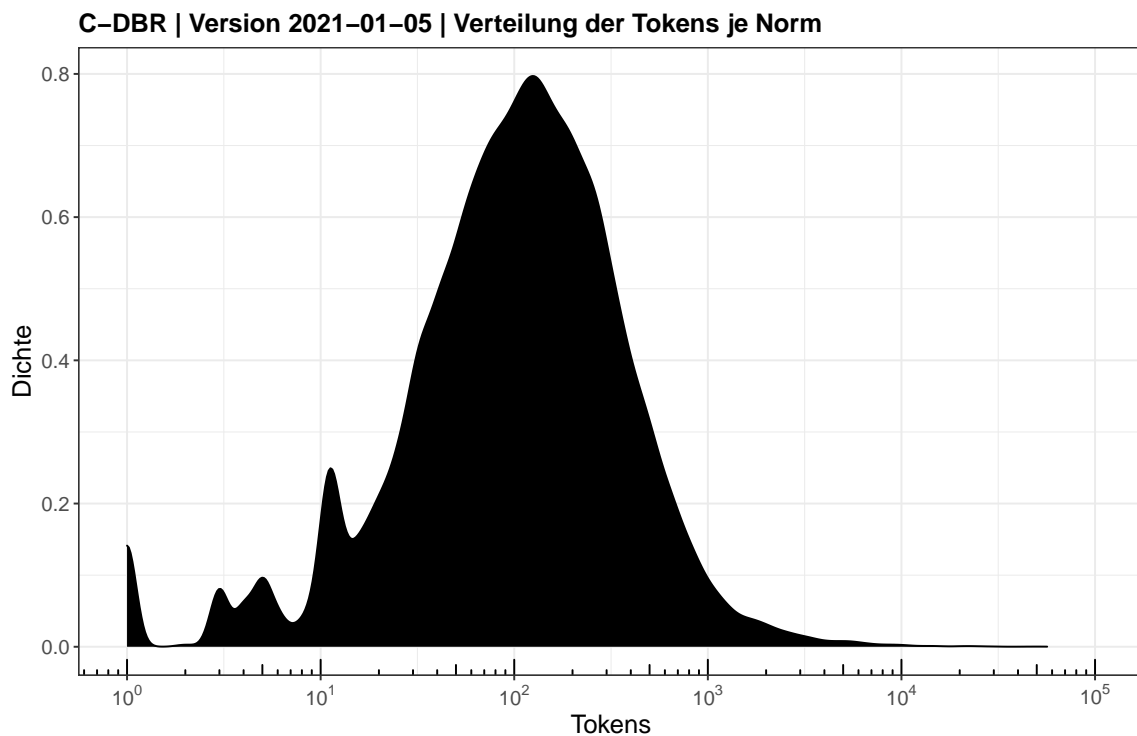
### 11.9.3 Zusammenfassungen speichern

```
fwrite(dt.stats.ling,  
       paste0(outputdir,  
               datasetname,  
               "_00_Rechtsakte_KorpusStatistik_ZusammenfassungLinguistisch.csv"),  
       na = "NA")
```

## 11.10 Density

### 11.10.1 Density (Tokens)

```
ggplot(data = summary.corpus.normen)+  
  geom_density(aes(x = tokens),  
               fill = "black") +  
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),  
               labels = trans_format("log10", math_format(10^.x)))+  
  annotation_logticks(sides = "b")+  
  coord_cartesian(xlim = c(1, 10^5))+  
  theme_bw()+  
  labs(  
    title = paste(datasetname,  
                  "| Version",  
                  datestamp,  
                  "| Verteilung der Tokens je Norm"),  
    caption = paste("DOI:",  
                    doi.version),  
    x = "Tokens",  
    y = "Dichte"  
  )+  
  theme(  
    text = element_text(size = 14),  
    plot.title = element_text(size = 14,  
                               face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```

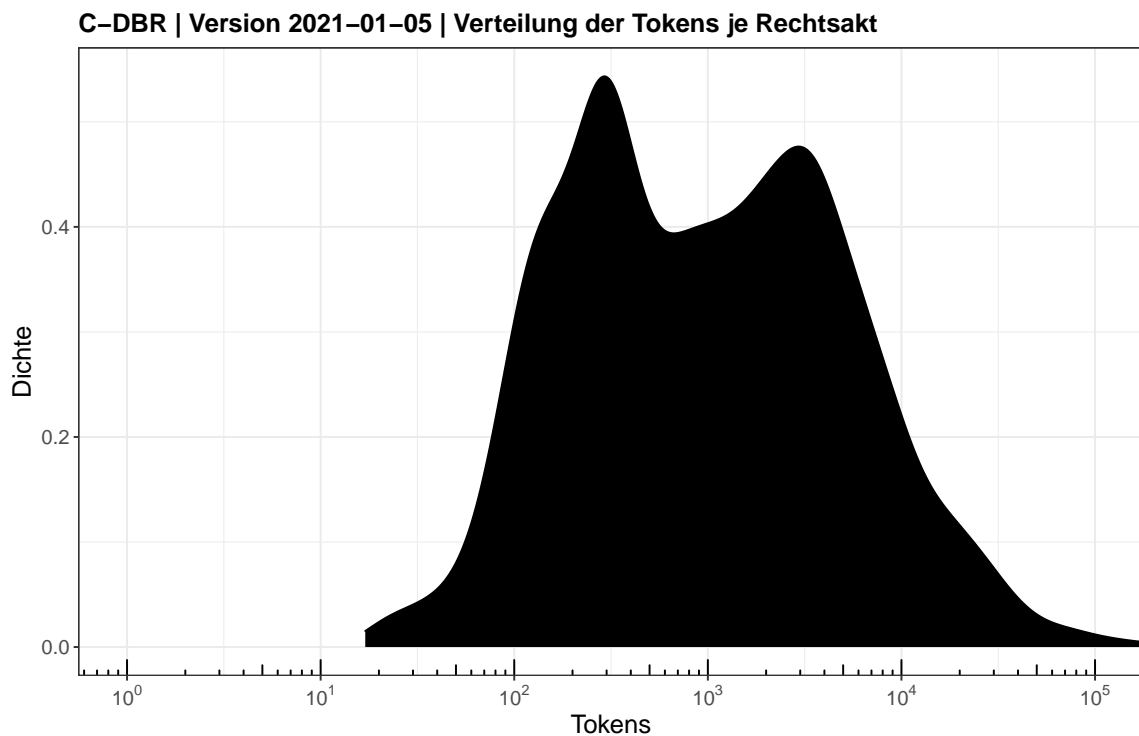


DOI: 10.5281/zenodo.4384771

```

ggplot(data = summary.corpus.rechtsakte)+
  geom_density(aes(x = tokens),
    fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
    labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  coord_cartesian(xlim = c(1, 10^5))+
  theme_bw() +
  labs(
    title = paste(datasetname,
      "| Version",
      datestamp,
      "| Verteilung der Tokens je Rechtsakt"),
    caption = paste("DOI:",
      doi.version),
    x = "Tokens",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
      face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )

```

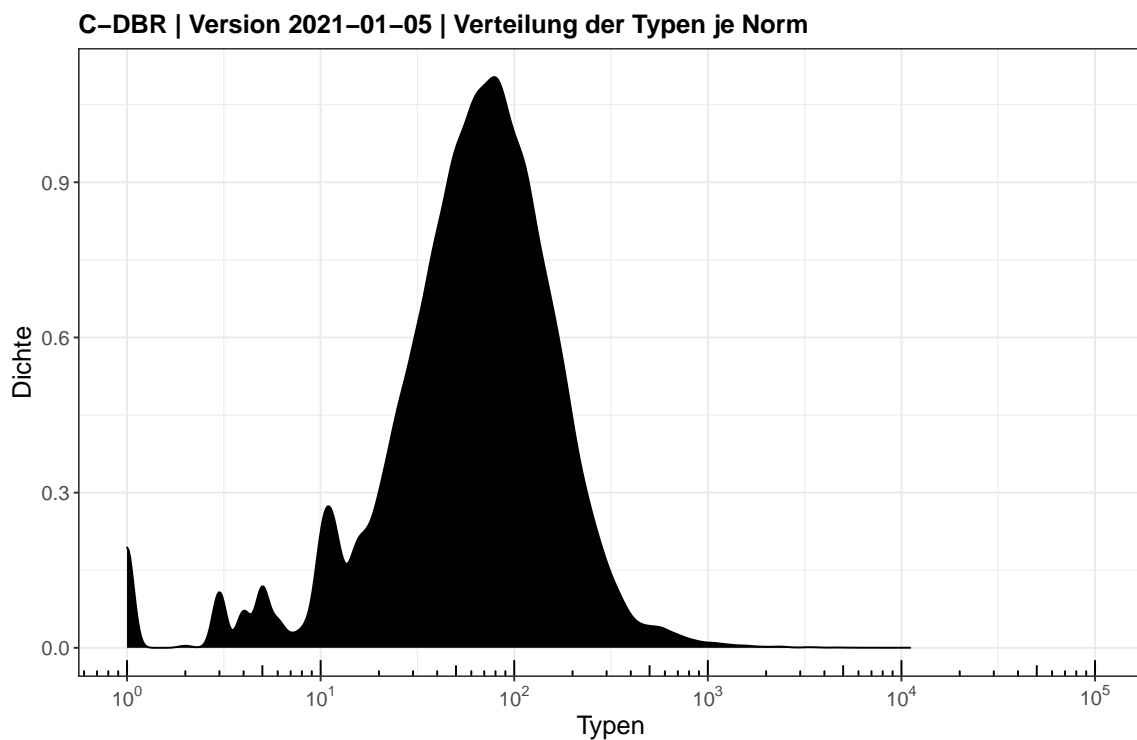


DOI: 10.5281/zenodo.4384771



### 11.10.2 Density (Typen)

```
ggplot(data = summary.corpus.normen)+  
  geom_density(aes(x = typen),  
    fill = "black")+  
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),  
    labels = trans_format("log10", math_format(10^.x)))+  
  annotation_logticks(sides = "b")+  
  coord_cartesian(xlim = c(1, 10^5))+  
  theme_bw()+  
  labs(  
    title = paste(datasetname,  
      "| Version",  
      datestamp,  
      "| Verteilung der Typen je Norm"),  
    caption = paste("DOI:",  
      doi.version),  
    x = "Typen",  
    y = "Dichte"  
  )+  
  theme(  
    text = element_text(size = 14),  
    plot.title = element_text(size = 14,  
      face = "bold"),  
    legend.position = "none",  
    plot.margin = margin(10, 20, 10, 10)  
  )
```

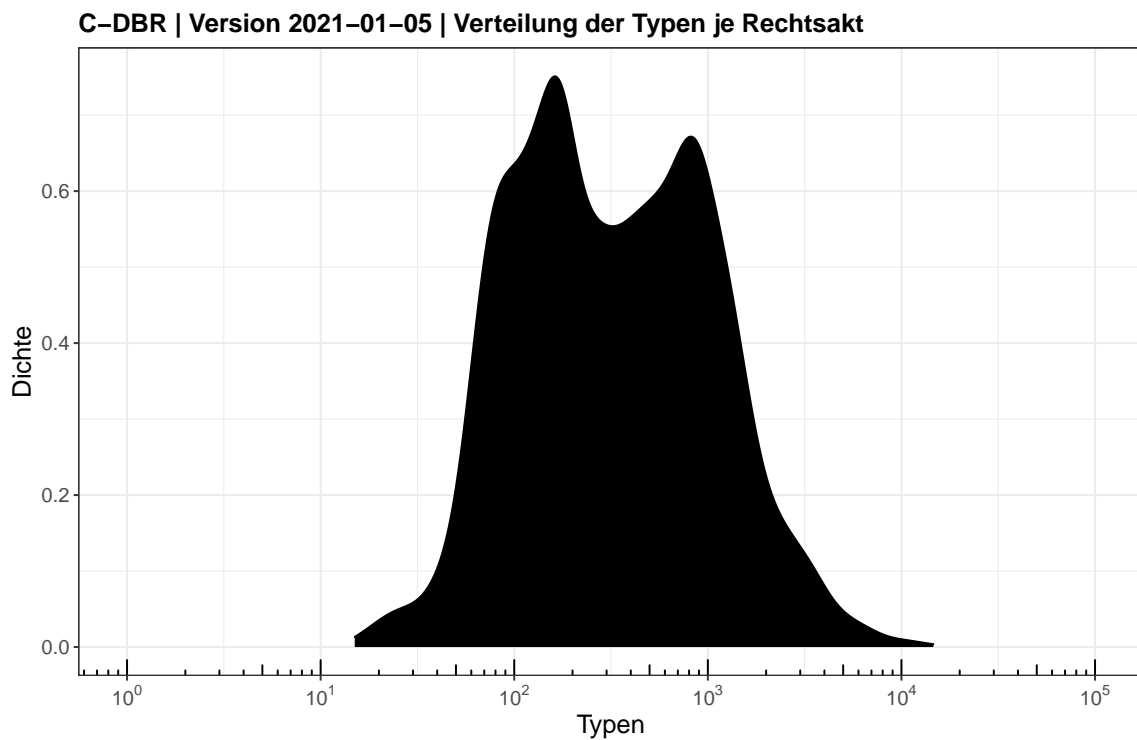


DOI: 10.5281/zenodo.4384771

```

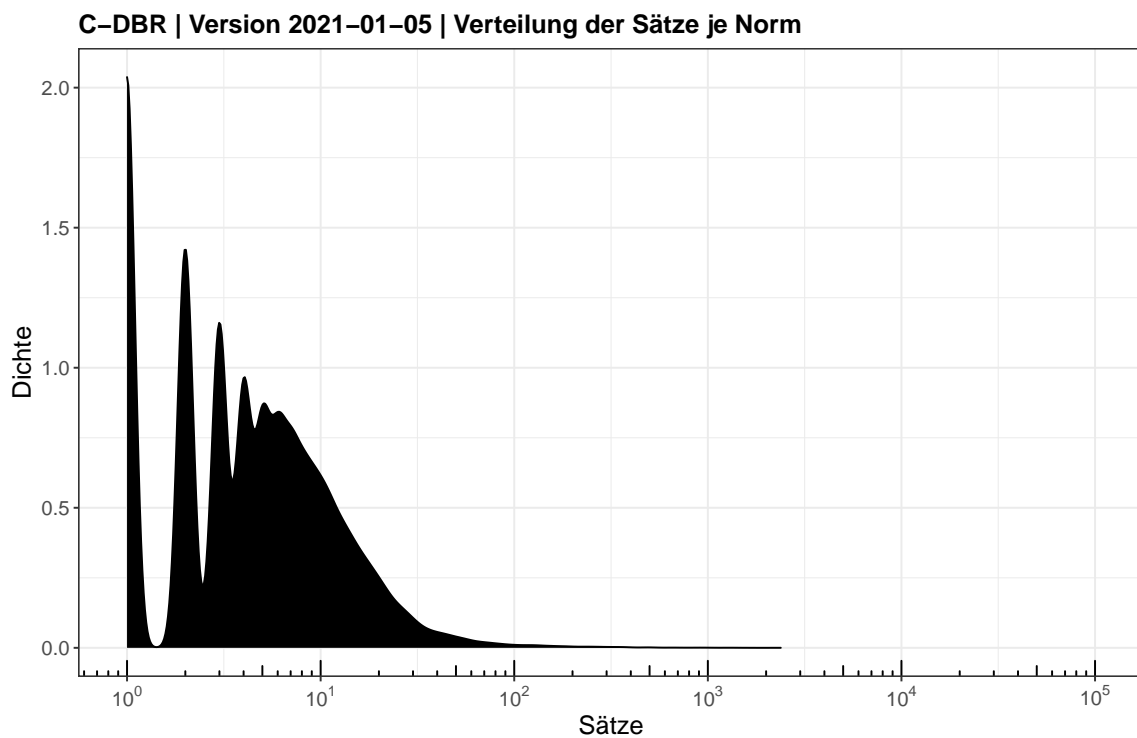
ggplot(data = summary.corpus.rechtsakte)+
  geom_density(aes(x = typen),
    fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
    labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  coord_cartesian(xlim = c(1, 10^5))+
  theme_bw()+
  labs(
    title = paste(datasetname,
      "| Version",
      datestamp,
      "| Verteilung der Typen je Rechtsakt"),
    caption = paste("DOI:",
      doi.version),
    x = "Typen",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
      face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )

```



### 11.10.3 Density (Sätze)

```
ggplot(data = summary.corpus.normen)+
  geom_density(aes(x = saetze),
    fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
    labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  coord_cartesian(xlim = c(1, 10^5))+
  theme_bw()+
  labs(
    title = paste(datasetname,
      "| Version",
      datestamp,
      "| Verteilung der Sätze je Norm"),
    caption = paste("DOI:",
      doi.version),
    x = "Sätze",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
      face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )
```

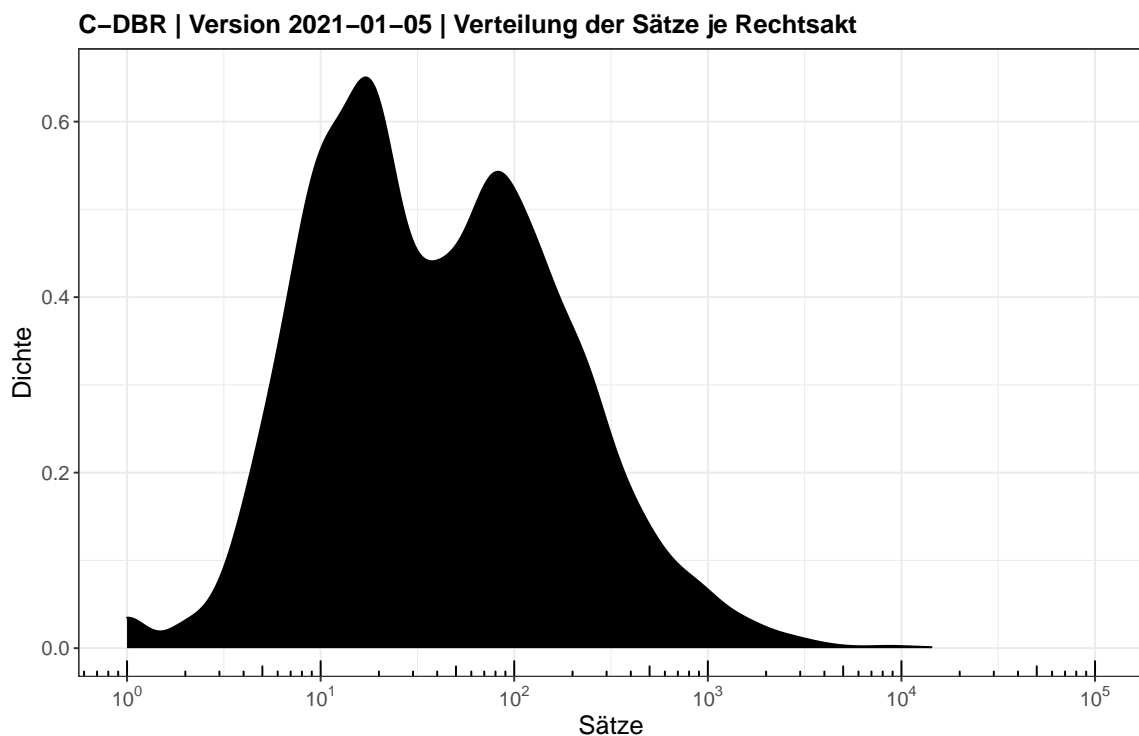


DOI: 10.5281/zenodo.4384771

```

ggplot(data = summary.corpus.rechtsakte)+
  geom_density(aes(x = saetze),
    fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
    labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  coord_cartesian(xlim = c(1, 10^5))+
  theme_bw()+
  labs(
    title = paste(datasetname,
      "| Version",
      datestamp,
      "| Verteilung der Sätze je Rechtsakt"),
    caption = paste("DOI:", doi.version),
    x = "Sätze",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
      face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )

```



DOI: 10.5281/zenodo.4384771

## 12 CSV-Dateien erstellen

### 12.1 Einzelnormen (Korpus)

#### 12.1.1 Name für CSV definieren

```
csvname.normen.gesamt <- paste(datasetname,  
                                datestamp,  
                                "DE_CSV_Einzelnormen_Datensatz.csv",  
                                sep = "_")
```

#### 12.1.2 Datensatz speichern

```
fwrite(dt.normen,  
       csvname.normen.gesamt,  
       na = "NA")
```

### 12.2 Einzelnormen (Metadaten)

#### 12.2.1 Name für CSV definieren

```
csvname.normen.meta <- paste(datasetname,  
                              datestamp,  
                              "DE_CSV_Einzelnormen_Metadaten.csv",  
                              sep = "_")
```

#### 12.2.2 Datensatz speichern

```
fwrite(summary.corpus.normen,  
       csvname.normen.meta,  
       na = "NA")
```

### 12.3 Rechtsakte (Korpus)

#### 12.3.1 Name für CSV definieren

```
csvname.rechtsakte.gesamt <- paste(datasetname,  
                                    datestamp,  
                                    "DE_CSV_Rechtsakte_Datensatz.csv",  
                                    sep = "_")
```

### 12.3.2 Datensatz speichern

```
fwrite(dt.rechtsakte,  
       csvname.rechtsakte.gesamt,  
       na = "NA")
```

## 12.4 Rechtsakte (Metadaten)

### 12.4.1 Name für CSV definieren

```
csvname.rechtsakte.meta <- paste(datasetname,  
                                  datestamp,  
                                  "DE_CSV_Rechtsakte_Metadaten.csv",  
                                  sep = "_")
```

### 12.4.2 Datensatz speichern

```
fwrite(summary.corpus.rechtsakte,  
       csvname.rechtsakte.meta,  
       na = "NA")
```

## 12.5 XML-Metadaten

Diese Datei unterscheidet sich von der Variante »DE\_CSV\_Rechtsakte\_Metadaten«, weil sie auch Rechtsakte enthält, die ohne Text veröffentlicht wurden. Die Differenz betrifft etwa 1000 Rechtsakte, ist also signifikant.

### 12.5.1 Name für CSV definieren

```
csvname.meta <- paste(datasetname,  
                      datestamp,  
                      "DE_CSV_MetadatenXML.csv",  
                      sep = "_")
```

### 12.5.2 Datensatz speichern

```
fwrite(dt.meta,  
       csvname.meta,  
       na = "NA")
```

## 13 Download der PDF-Dateien

### 13.1 Download durchführen

```
mapply(download.file,  
        download$links.pdf,  
        download$title.pdf)
```

### 13.2 Download-Ergebnis

#### 13.2.1 Anzahl herunterzuladender Dateien

```
download[, .N]
```

```
## [1] 6577
```

#### 13.2.2 Anzahl heruntergeladener Dateien

```
files.pdf <- list.files(pattern = "\\..pdf")  
length(files.pdf)
```

```
## [1] 6577
```

#### 13.2.3 Fehlbetrag

```
N.missing <- download[, .N] - length(files.pdf)  
print(N.missing)
```

```
## [1] 0
```

#### 13.2.4 Fehlende Dateien

```
missing <- setdiff(download$title.pdf,  
                   files.pdf)  
print(missing)
```

```
## character(0)
```

## 14 TXT-Dateien erstellen

An dieser Stelle wird der reine Text aus den PDF-Dateien extrahiert und ein zusätzliches Datei-Format (TXT) generiert. TXT-Dateien sind besonders für quantitative Analysten ohne XML-Kenntnisse ein lohnenswerter Einstieg und verringern die Hürde für die Arbeit mit dem Korpus.

```
files.pdf <- list.files(pattern="\\.pdf",  
                        ignore.case = TRUE)
```

### 14.1 Anzahl zu extrahierender Dateien

```
length(files.pdf)
```

```
## [1] 6577
```

### 14.2 Seiten zählen: Funktion anzeigen

```
print(f.dopar.pagenums)
```

```
function(x){
```

```
  pagenums <- foreach(filename = x,  
                      .combine = 'c',  
                      .errorhandling = 'remove',  
                      .inorder = FALSE) %dopar% {  
    pdf_length(filename)  
  }  
  return(pagenums)
```

```
}
```

### 14.3 Anzahl zu extrahierender Seiten

```
sum(f.dopar.pagenums(files.pdf))
```

```
## [1] 53901
```



## 14.4 PDF extrahieren: Funktion anzeigen

```
print(f.dopar.pdfextract)
```

```
function(x){
```

```
  newnames <- gsub("\\.pdf",
                  "\\ .txt",
                  x)

  out <- foreach(i = seq_along(x),
                .errorhandling = 'pass') %dopar% {

    ## Extract text layer from PDF
    pdf.extracted <- pdf_text(x[i])

    ## Write TXT to Disk
    write.table(pdf.extracted,
                newnames[i],
                quote = FALSE,
                row.names = FALSE,
                col.names = FALSE)

  }
  return(length(out))
}
```

```
}
```

## 14.5 Text Extrahieren

```
f.dopar.pdfextract(files.pdf)
```

```
## [1] 6577
```

## 15 Download der EPUB-Dateien

### 15.1 Download durchführen

```
mapply(download.file,  
        download$links.epub,  
        download$title.epub)
```

### 15.2 Download-Ergebnis

#### 15.2.1 Anzahl herunterzuladender Dateien

```
download[, .N]
```

```
## [1] 6577
```

#### 15.2.2 Anzahl heruntergeladener Dateien

```
files.epub <- list.files(pattern = "\\\\.epub")  
length(files.epub)
```

```
## [1] 6577
```

#### 15.2.3 Fehlbetrag

```
N.missing <- download[, .N] - length(files.epub)  
print(N.missing)
```

```
## [1] 0
```

#### 15.2.4 Fehlende Dateien

```
missing <- setdiff(download$title.epub, files.epub)  
print(missing)
```

```
## character(0)
```

## 16 Dateigrößen analysieren

```
files.txt <- list.files(pattern = "\\..txt$", ignore.case = TRUE)
files.pdf <- list.files(pattern = "\\..pdf$", ignore.case = TRUE)
files.epub <- list.files(pattern = "\\..epub$", ignore.case = TRUE)

txt.MB <- file.size(files.txt) / 10^6
pdf.MB <- file.size(files.pdf) / 10^6
epub.MB <- file.size(files.epub) / 10^6
```

### 16.1 Gesamtgröße

#### 16.1.1 PDF-Dateien (MB)

```
sum(pdf.MB)
```

```
## [1] 618.9734
```

#### 16.1.2 EPUB-Dateien (MB)

```
sum(epub.MB)
```

```
## [1] 435.1321
```

#### 16.1.3 XML-Dateien (MB)

```
sum(xml.MB)
```

```
## [1] 258.1441
```

#### 16.1.4 TXT-Dateien (MB)

```
sum(txt.MB)
```

```
## [1] 190.8064
```

### 16.1.5 Korpus-Objekte in RAM (MB)

```
print(object.size(corpus.normen),  
      standard= "SI",  
      humanReadable=TRUE,  
      units ="MB")
```

```
## 197.7 MB
```

```
print(object.size(corpus.rechtsakte),  
      standard= "SI",  
      humanReadable=TRUE,  
      units ="MB")
```

```
## 145.2 MB
```

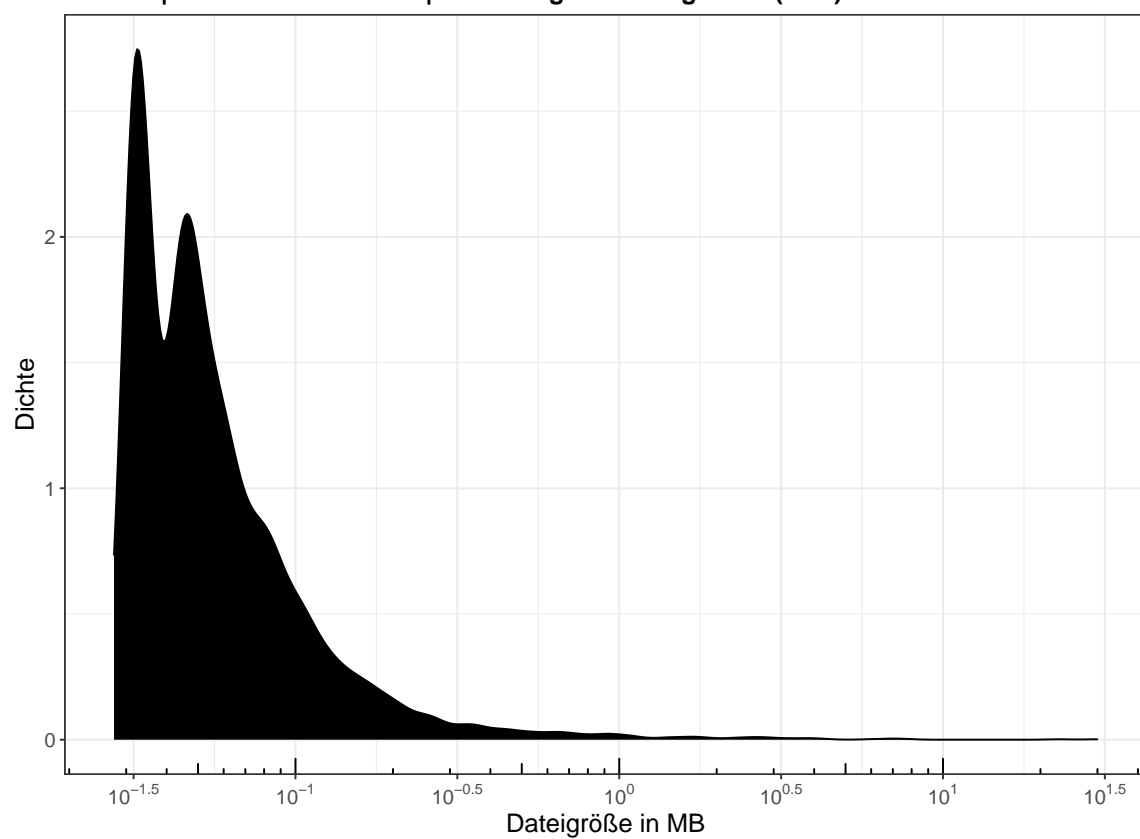
```
print(object.size(dt.meta),  
      standard= "SI",  
      humanReadable=TRUE,  
      units ="MB")
```

```
## 5.3 MB
```

## 16.2 Verteilung der Dateigrößen (PDF)

```
dt.plot <- data.table(pdf.MB)
```

```
ggplot(data = dt.plot,
  aes(x = pdf.MB))+
  geom_density(fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
    labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  theme_bw()+
  labs(
    title = paste(datasetname,
      "| Version",
      datestamp,
      "| Verteilung der Dateigrößen (PDF)",
    caption = paste("DOI:",
      doi.version),
    x = "Dateigröße in MB",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
      face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )
```



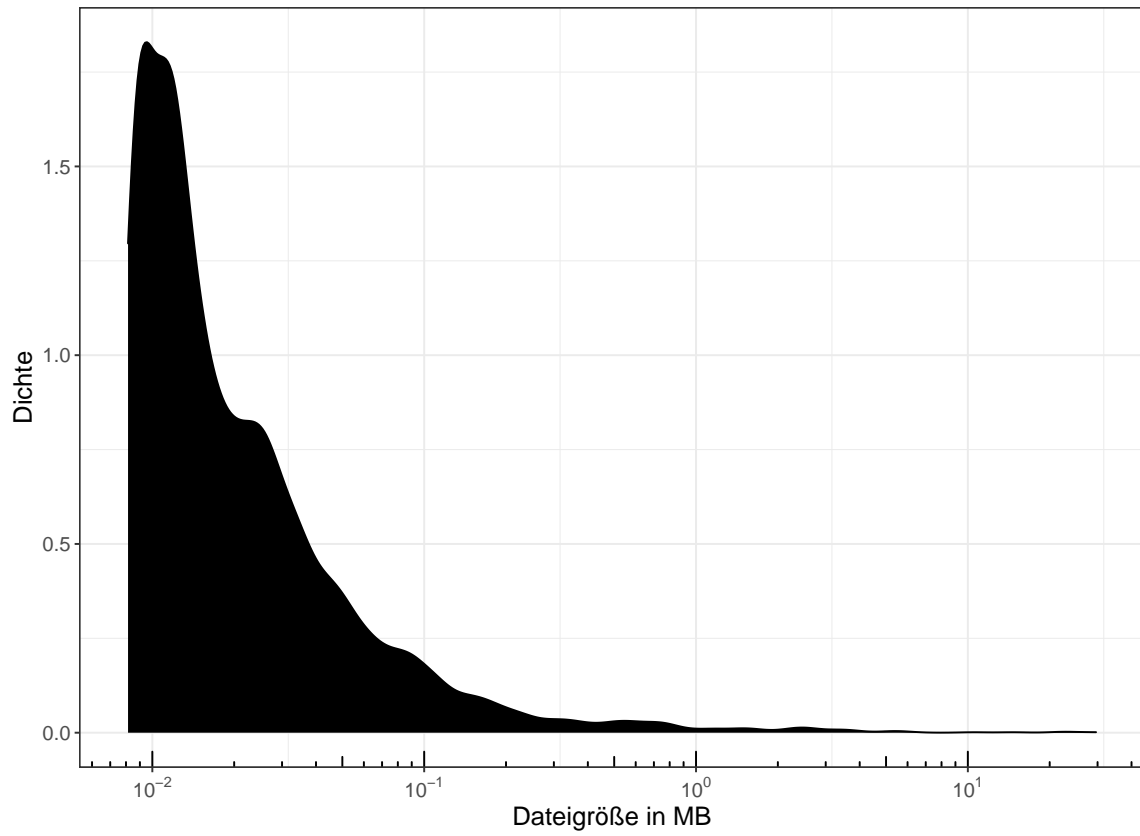
DOI: 10.5281/zenodo.4384771

## 16.3 Verteilung der Dateigrößen (EPUB)

```
dt.plot <- data.table(epub.MB)
```

```
ggplot(data = dt.plot,
  aes(x = epub.MB))+
  geom_density(fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
    labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  theme_bw()+
  labs(
    title = paste(datasetname,
      "| Version",
      datestamp,
      "| Verteilung der Dateigrößen (EPUB)" ),
    caption = paste("DOI:",
      doi.version),
    x = "Dateigröße in MB",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
      face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )
```

C-DBR | Version 2021-01-05 | Verteilung der Dateigrößen (EPUB)



DOI: 10.5281/zenodo.4384771

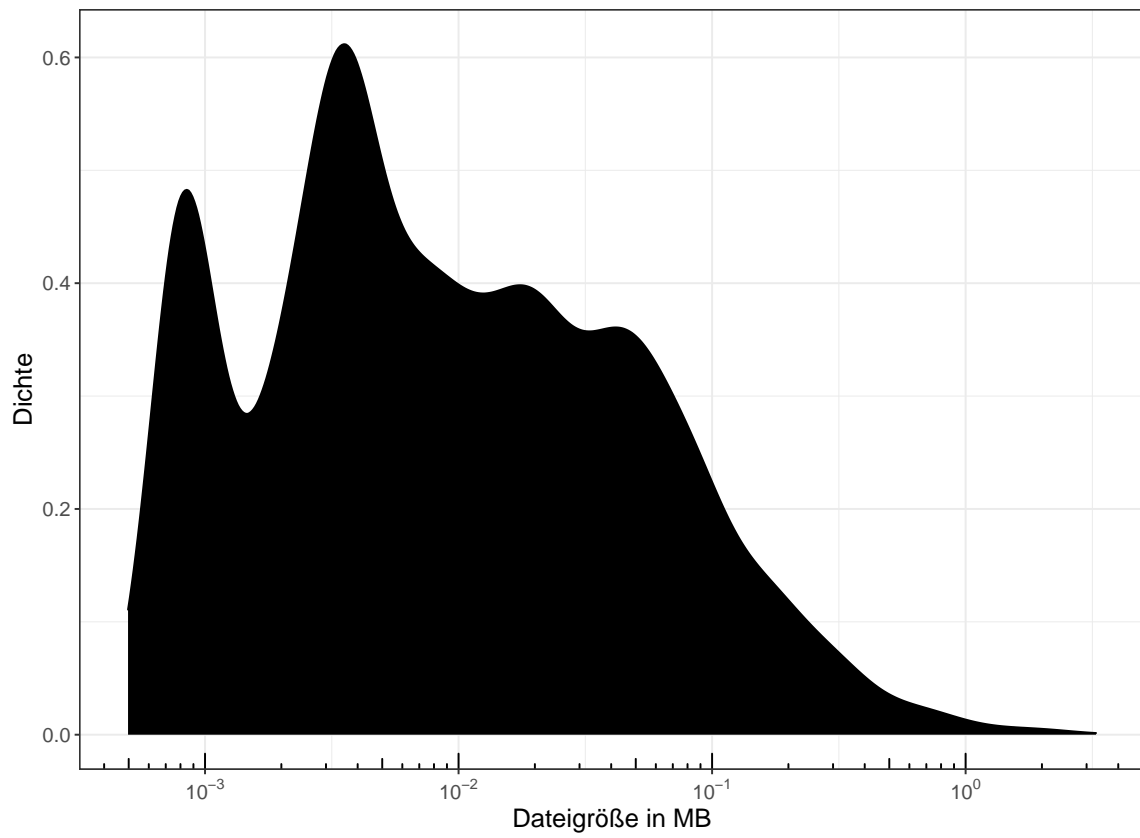


## 16.4 Verteilung der Dateigrößen (XML)

```
dt.plot <- data.table(xml.MB)
```

```
ggplot(data = dt.plot,
       aes(x = xml.MB))+
  geom_density(fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
               labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  theme_bw()+
  labs(
    title = paste(datasetname,
                  "| Version",
                  datestamp,
                  "| Verteilung der Dateigrößen (XML)" ),
    caption = paste("DOI:",
                    doi.version),
    x = "Dateigröße in MB",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
                               face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )
```

C-DBR | Version 2021-01-05 | Verteilung der Dateigrößen (XML)

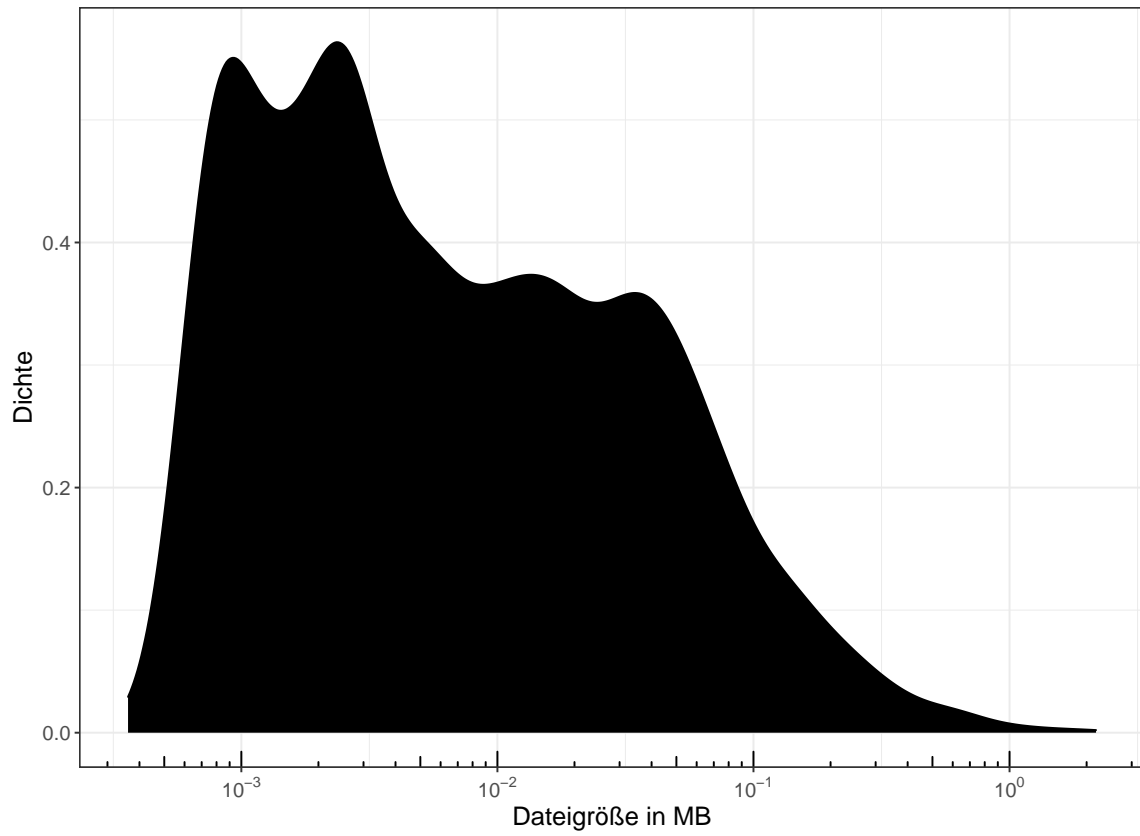


DOI: 10.5281/zenodo.4384771

## 16.5 Verteilung der Dateigrößen (TXT)

```
dt.plot <- data.table(txt.MB)
```

```
ggplot(data = dt.plot,
  aes(x = txt.MB))+
  geom_density(fill = "black")+
  scale_x_log10(breaks = trans_breaks("log10", function(x) 10^x),
    labels = trans_format("log10", math_format(10^.x)))+
  annotation_logticks(sides = "b")+
  theme_bw()+
  labs(
    title = paste(datasetname,
      "| Version",
      datestamp,
      "| Verteilung der Dateigrößen (TXT)",
    caption = paste("DOI:",
      doi.version),
    x = "Dateigröße in MB",
    y = "Dichte"
  )+
  theme(
    text = element_text(size = 14),
    plot.title = element_text(size = 14,
      face = "bold"),
    legend.position = "none",
    plot.margin = margin(10, 20, 10, 10)
  )
```



DOI: 10.5281/zenodo.4384771

## 17 ZIP-Archive erstellen

### 17.1 Verpacken der CSV-Dateien

```
files.csv <- c(csvname.normen.gesamt,  
              csvname.normen.meta,  
              csvname.rechtsakte.gesamt,  
              csvname.rechtsakte.meta,  
              csvname.meta)  
  
csvnames.zip <- gsub(".csv",  
                    ".zip",  
                    files.csv)  
  
for (i in seq_along(files.csv)){  
  zip(csvnames.zip[i],  
      files.csv[i])  
}  
  
unlink(csvname.normen.gesamt)  
unlink(csvname.rechtsakte.gesamt)
```

### 17.2 Verpacken der PDF-Dateien

```
files.pdf <- list.files(pattern="\\.pdf$",  
                       ignore.case = TRUE)  
  
zip(paste(datasetname,  
          datestamp,  
          "DE_PDF_Datensatz.zip",  
          sep = "_"),  
    files.pdf)  
  
unlink(files.pdf)
```

### 17.3 Verpacken der TXT-Dateien

```
files.txt <- list.files(pattern="\\.txt$",  
                       ignore.case = TRUE)  
  
zip(paste(datasetname,  
          datestamp,  
          "DE_TXT_Datensatz.zip",  
          sep = "_"),  
    files.txt)  
  
unlink(files.txt)
```

## 17.4 Verpacken der EPUB-Dateien

```
files.epub <- list.files(pattern="\\.epub$",
                        ignore.case = TRUE)

zip(paste(datasetname,
          datestamp,
          "DE_EPUB_Datensatz.zip",
          sep = "_"),
    files.epub)

unlink(files.epub)
```

## 17.5 Verpacken der Analyse-Dateien

```
zip(paste0(datasetname,
          "_",
          datestamp,
          "_DE_",
          basename(outputdir),
          ".zip"),
    basename(outputdir))
```

## 17.6 Verpacken der Source-Dateien

```
files.source <- c(list.files(pattern = "Source"),
                  "buttons")

files.source <- grep("spin",
                    files.source,
                    value = TRUE,
                    ignore.case = TRUE,
                    invert = TRUE)

zip(paste(datasetname,
          datestamp,
          "Source_Files.zip",
          sep = "_"),
    files.source)
```

## 18 Kryptographische Hashes

Dieses Modul berechnet für jedes ZIP-Archiv zwei Arten von Hashes: SHA2-256 und SHA3-512. Mit diesen kann die Authentizität der Dateien geprüft werden und es wird dokumentiert, dass sie aus diesem Source Code hervorgegangen sind. Die SHA-2 und SHA-3 Algorithmen gelten derzeit als sicher und ein SHA3-Hash mit 512 bit Länge ist nach derzeitigem Wissen auch gegenüber quantenkryptoanalytischen Verfahren hinreichend resistent.

### 18.1 Liste der ZIP-Archive erstellen

```
files.zip <- list.files(pattern= "\\\\.zip$",  
                        ignore.case = TRUE)
```

### 18.2 Funktion anzeigen

```
print(f.dopar.multihashes)
```

```
function(x){
```

```
  multihashes <- foreach(filename = x,  
    .errorhandling = 'pass',  
    .combine = 'rbind') %dopar% {  
  
    sha2.256 <- system2("openssl",  
                      paste("sha256",  
                            filename),  
                      stdout = TRUE)  
  
    sha2.256 <- gsub("^.*\\|= ",  
                  "",  
                  sha2.256)  
  
    sha3.512 <- system2("openssl",  
                      paste("sha3-512",  
                            filename),  
                      stdout = TRUE)  
  
    sha3.512 <- gsub("^.*\\|= ",  
                  "",  
                  sha3.512)  
  
    out <- data.frame(filename,  
                     sha2.256,  
                     sha3.512)  
  
    return(out)  
  }  
  return(multihashes)
```

```
}
```

### 18.3 Hashes berechnen

```
multihashes <- f.dopar.multihashes(files.zip)
```

### 18.4 In Data Table umwandeln

```
setDT(multihashes)
```

### 18.5 Index hinzufügen

```
multihashes$index <- seq_len(multihashes[,.N])
```

### 18.6 Hashes in CSV-Datei speichern

```
fwrite(multihashes,  
       paste(datasetname,  
             datestamp,  
             "KryptographischeHashes.csv",  
             sep = "_"),  
       na = "NA")
```

### 18.7 Leerzeichen hinzufügen um Zeilenumbruch zu ermöglichen

```
multihashes$sha3.512 <- paste(substr(multihashes$sha3.512, 1, 64),  
                             substr(multihashes$sha3.512, 65, 128))
```



## 18.8 In Bericht anzeigen

```
kable(multihashes[,.(index,filename)],  
      format = "latex",  
      align = c("p{1cm}", "p{13cm}"),  
      booktabs=TRUE,  
      longtable=TRUE)
```

| index | filename   |
|-------|--|
| 1     | C-DBR_2021-01-05_DE_ANALYSE.zip                    |
| 2     | C-DBR_2021-01-05_DE_CSV_Einzelnormen_Datensatz.zip |
| 3     | C-DBR_2021-01-05_DE_CSV_Einzelnormen_Metadaten.zip |
| 4     | C-DBR_2021-01-05_DE_CSV_MetadatenXML.zip           |
| 5     | C-DBR_2021-01-05_DE_CSV_Rechtsakte_Datensatz.zip   |
| 6     | C-DBR_2021-01-05_DE_CSV_Rechtsakte_Metadaten.zip   |
| 7     | C-DBR_2021-01-05_DE_EPUB_Datensatz.zip             |
| 8     | C-DBR_2021-01-05_DE_PDF_Datensatz.zip              |
| 9     | C-DBR_2021-01-05_DE_TXT_Datensatz.zip              |
| 10    | C-DBR_2021-01-05_DE_XML_Anlagen.zip                |
| 11    | C-DBR_2021-01-05_DE_XML_Datensatz.zip              |
| 12    | C-DBR_2021-01-05_Source_Files.zip                  |

```
kable(multihashes[,.(index,sha2.256)],
      format = "latex",
      align = c("c", "p{13cm}"),
      booktabs=TRUE,
      longtable=TRUE)
```

| index | sha2.256   |
|-------|--|
| 1     | d21ed0f4d5c9e660e7435fd6ae5913b3e942128f16590d79e4047396230b2a8c |
| 2     | e5bfa9269c143c78654ef4861bf591a8684d034bb32e61e4ea35d74f6d1e3cef |
| 3     | 1d13579d1996586e94f65808f78202512a14711c3f50a18ee94826814830cea1 |
| 4     | 0f0f7ecfc5eba83aa5de7c6c61aa86c59bcab0f219b3749a859d678f48f5a82d |
| 5     | 9cc495059251199f317ec35ce492ad61ea1a9cb85aaceacb8cd41ccc0eef84bf |
| 6     | e132827fc6ad40f8760483960dd4a7512e02799805d0ebcbb4b7aa648d35cfd4 |
| 7     | 4c70048097c358a9783183bafdfbf02110a1d803a4da647410c143e9167fec0d |
| 8     | 95e422fa5c5be43c237f972aa3d4560dc3def68ca3db5070bc8471d2db91fc45 |
| 9     | 5d18c9fb2ae70044bc518db76ddca43e8494de8fe3e02e2e926d5bac5c65afbd |
| 10    | e8ac0f49e74c47ece84f0f68a080ba1d35c3495cd035f822d2db16325fc355ac |
| 11    | 4c7d38095f83fa9a5cfe949e125f91b437c870800d241c7b96bd7fe0cdede0ab |
| 12    | 162975f0b400ce147679d07c2938526fe8bd35a0c36c73792c54212a83d55cca |

```
kable(multihashes[,.(index,sha3.512)],
      format = "latex",
      align = c("c", "p{13cm}"),
      booktabs=TRUE,
      longtable=TRUE)
```

| index | sha3.512  |
|-------|---|
| 1     | 5b454eacad6ab52f7e02b67e559462094d974969e770df718afcaf5ce0dedfb9bd00353c0c001e4466d899e8d202103f6bfc0afa9611ffc771eaaa838efb5c9b  |
| 2     | d7ee215c835a79cc86de00d0739a987644f653bc5027cbce480acd87b1b567c54ff5042e86110eadee215a9f57a401019bcf58cde8984f8ba85c7e2fd409aa52  |
| 3     | 024abf3aaa8d656d34361a824245896254c4b1e9b1e1835e06063cf70f776af2e00d64663c8e45af66e6dcaa9860f269631daaf095f2462debd0e8f576cb4529  |
| 4     | 380159fc7887da88a9866b40411d86edeee98d75a43e982e979108e84dd0851990a9bc01e60960e6fe5f7e6ed4aa63d126033eab14e25bc3fba364223712e4ad  |
| 5     | 9ab096bdb2db4f3fedf6725ec66baf4b70995150ff201b3d8a0ea1f735090787d3d0a82650ec805792ae407950c92a9d64108ef1af861bdd538c8d0e41f0512e  |
| 6     | 5f548c6a2a262a26fca213af7f4162df409d83cb7edc11d85aebcea6ad33bfff6d5242170044e97728ff13e8507e6cd3e6b0f5358ddf4d0889795730c6ab64ec7 |
| 7     | 3bec75847703cf95a21b10e248b32dca10a558329b3f5daf7ea3629103002962c6cff0066d307207d6ebf68d62c6690ce629fdb917f25a84b7dc9ca33c7bc5da  |
| 8     | 4aecf4f8c6dbb21d3d349f2c0b576c1bbcb01ebaca94a2da5f5cb5bc69682412e3ca2f3890b01eebc3eb43b58247b0261c08e4de0ec06e0bdf2dceb05e136218  |
| 9     | c27fe26e76fd2b232dd2c8159324f926ef091de351610f82c01caf503f8950e24912987c8dc4aa3b2207ca46211c3e77baabd305e0163d08cb7b59011833d9ef  |
| 10    | 74398de79f4a2391941e5d777a428fe7269476e86ccab8adbe856ffaf919ec5fe448ae5e7d93fcd926dcbae7ea1f0f79ff0c2e207ff967e799e7b1fc8c33c0f   |
| 11    | 1a0f16fc0cca7316ab55b8a5358aff7ef365a0f4e44984a4257b7f1597a1b86bec27c5c834aadffad953c9fa60f639beeaca2697539817106aa0ab0ca3af0438  |
| 12    | 09d5dad3947bc281b28d798d66afc8b709b1a039883fbd8c3b96b2d508d98c43a8fbfe3ef275c5f6a30d9881627e9b540663485156bda3ed6b2f8197f9472f0   |

## 19 Abschluss

### 19.1 Cluster stoppen

```
stopCluster(cl)
```

### 19.2 Datum und Uhrzeit (Ende)

```
end.script <- Sys.time()  
print(end.script)
```

```
## [1] "2021-01-05 09:53:23 CET"
```

### 19.3 Laufzeit des gesamten Skripts

```
print(end.script - begin.script)
```

```
## Time difference of 1.403702 hours
```

### 19.4 Warnungen

```
warnings()
```

## 20 Parameter für strenge Replikationen

```
system2("openssl", "version", stdout = TRUE)
```

```
## [1] "OpenSSL 1.1.1i FIPS 8 Dec 2020"
```

```
sessionInfo()
```

```
## R version 4.0.3 (2020-10-10)
## Platform: x86_64-redhat-linux-gnu (64-bit)
## Running under: Fedora 32 (Workstation Edition)
##
## Matrix products: default
## BLAS/LAPACK: /usr/lib64/libopenblas-r0.3.12.so
##
## locale:
##  [1] LC_CTYPE=en_US.utf8      LC_NUMERIC=C
##  [3] LC_TIME=en_US.utf8      LC_COLLATE=en_US.utf8
##  [5] LC_MONETARY=en_US.utf8  LC_MESSAGES=en_US.utf8
##  [7] LC_PAPER=en_US.utf8     LC_NAME=C
##  [9] LC_ADDRESS=C            LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.utf8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] parallel stats      graphics grDevices utils      datasets methods
## [8] base
##
## other attached packages:
##  [1] scales_1.1.1      quanteda_2.1.2      data.table_1.13.4 ggplot2_3.3.2
##  [5] doParallel_1.0.16 iterators_1.0.13   foreach_1.5.1      pdftools_2.3.1
##  [9] kableExtra_1.3.1  knitr_1.30          rvest_0.3.6        xml2_1.3.2
##
## loaded via a namespace (and not attached):
##  [1] qpdf_1.1          tidyselect_1.1.0    xfun_0.19          purrr_0.3.4
##  [5] lattice_0.20-41   colorspace_2.0-0    vctr_0.3.6         generics_0.1.0
##  [9] htmltools_0.5.0   usethis_2.0.0       viridisLite_0.3.0  yaml_2.2.1
## [13] rlang_0.4.9       pillar_1.4.7        glue_1.4.2         withr_2.3.0
## [17] selectr_0.4-2     lifecycle_0.2.0     stringr_1.4.0      munsell_0.5.0
## [21] gtable_0.3.0      codetools_0.2-18    evaluate_0.14      labeling_0.4.2
## [25] curl_4.3          highr_0.8           Rcpp_1.0.5         magick_2.5.2
## [29] RcppParallel_5.0.2 webshot_0.5.2       farver_2.0.3       fs_1.5.0
## [33] fastmatch_1.1-0   stopwords_2.1        askpass_1.1        digest_0.6.27
## [37] stringi_1.5.3     dplyr_1.0.2         grid_4.0.3         tools_4.0.3
## [41] magrittr_2.0.1    tibble_3.0.4        crayon_1.3.4       pkgconfig_2.0.3
## [45] ellipsis_0.3.1    Matrix_1.2-18       rmarkdown_2.5      httr_1.4.2
## [49] rstudioapi_0.13   R6_2.5.0            compiler_4.0.3
```

## Literaturverzeichnis

Analytics, Revolution, and Steve Weston. 2020. *Iterators: Provides Iterator Construct*. <https://github.com/RevolutionAnalytics/iterators>.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. “Quanteda: An R Package for the Quantitative Analysis of Textual Data.” *Journal of Open Source Software* 3 (30): 774. <https://doi.org/10.21105/joss.00774>.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, Akitaka Matsuo, Jiong Wei Lua, Jouni Kuha, and William Lowe. 2020. *Quanteda: Quantitative Analysis of Textual Data*. <https://quanteda.io>.

Corporation, Microsoft, and Steve Weston. 2020. *DoParallel: Foreach Parallel Adaptor for the Parallel Package*. <https://CRAN.R-project.org/package=doParallel>.

Dowle, Matt, and Arun Srinivasan. 2020. *Data.table: Extension of ‘Data.frame’*. <https://CRAN.R-project.org/package=data.table>.

Ooms, Jeroen. 2020. *Pdftools: Text Extraction, Rendering and Converting of Pdf Documents*. <https://CRAN.R-project.org/package=pdfutils>.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.

Revolution Analytics, and Steve Weston. n.d. *Foreach: Provides Foreach Looping Construct*.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.

———. 2020. *Rvest: Easily Harvest (Scrape) Web Pages*. <https://CRAN.R-project.org/package=rvest>.

Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2020. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. <https://CRAN.R-project.org/package=ggplot2>.

Wickham, Hadley, Jim Hester, and Jeroen Ooms. 2020. *Xml2: Parse Xml*. <https://CRAN.R-project.org/package=xml2>.

Wickham, Hadley, and Dana Seidel. 2020. *Scales: Scale Functions for Visualization*. <https://CRAN.R-project.org/package=scales>.

Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. <http://www.crcpress.com/product/isbn/9781466561595>.

———. 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. <https://yihui.org/knitr/>.

———. 2020. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. <https://yihui.org/knitr/>.

Zhu, Hao. 2020. *KableExtra: Construct Complex Table with Kable and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.