# LINEARE REGRESSIONSANALYSE MIT SPSS

## Anwendungsbeispiel mit Daten von AUSSDA – The Austrian Social Science Data Archive

**Erstellt von:**

Otto Bodi-Fernandez

Universität Graz

Kontakt:

otto.bodi@uni-graz.at

Version: 1.0

DOI 10.5281/zenodo.3612747

# Lineare Regressionsanalyse

In diesem Anwendungsbespiel wird Schritt für Schritt eine lineare Regressionsanalyse mit der Statistiksoftware SPSS (Version 25) durchgeführt. Die Datenbasis ist der österreichische Teil des International Social Survey Programme (ISSP) aus dem Jahr 2012. Der Datensatz ist bei AUSSDA – The Austrian Social Science Data Archive unter https://data.aussda.at für Forschung und Lehre frei zugänglich. Die Ergebnisse dieses Beispiels können durch das Ausführen der dokumentierten Syntaxbefehle repliziert werden.

## Datensatz

Höllinger, Franz; Eder, Anja; Haring-Mosbacher, Sabine, 2019, "ISSP-2012 Austria - Family and gender roles (with supplementary questions on leisure, social contact, and health issues) (SUF edition)", doi:10.11587/WMBC7S, AUSSDA Dataverse, V1

Der Datensatz repräsentiert eine Mehrthemenumfrage, die repräsentativ für die österreichische Wohnbevölkerung ab 18 Jahren ist. Frage 4 im Fragbogen enthält eine Batterie zur Einstellung zu Heirat und Ehe:

**4. Ich habe hier noch einige Aussagen. Inwieweit stimmen Sie diesen zu oder nicht zu? (Karte, Durchfragen)**

| | **stimme voll und ganz zu** | **stimme eher zu** | **weder noch** | **stimme eher nicht zu** | **stimme überhaupt nicht zu** | **kann ich nicht sagen** |
|---|---|---|---|---|---|---|
| Verheiratete Menschen sind im Allgemeinen glücklicher als Menschen, die nicht verheiratet sind. | 1 | 2 | 3 | 4 | 5 | 8 |
| Menschen, die Kinder wollen, sollten heiraten. | 1 | 2 | 3 | 4 | 5 | 8 |
| Es ist in Ordnung, dass ein Paar zusammenlebt, ohne die Absicht zu heiraten. | 1 | 2 | 3 | 4 | 5 | 8 |
| Eine Scheidung ist im Allgemeinen die beste Lösung, wenn ein Paar seine Eheprobleme nicht lösen kann. | 1 | 2 | 3 | 4 | 5 | 8 |

Wir wollen mittels einer linearen Regressionsanalyse mögliche Einflussfaktoren auf die Einstellung zur Heirat untersuchen. Dazu soll als abhängige Variable (Kriterium) ein Index aus den vier Items gebildet werden. Zuvor wird die Reliabilitätsanalyse durchgeführt, indem die Itemkonsistenz mittels Chronsbach's Alpha geprüft wird. Die Reliabilitätsanalyse erfordert, dass alle vier Items inhaltliche in die gleiche Richtung gepolt sind. Im Fragebogen sind allerdings die ersten beiden und die letzten beiden Aussagen in Bezug auf die Einstellung zur Heirat gegensätzlich gepolt. Es müssen daher zunächst das dritte und vierte Item vor der Durchführung der Reliabilitätsanalyse umgepolt werden.

## Vorbereitung

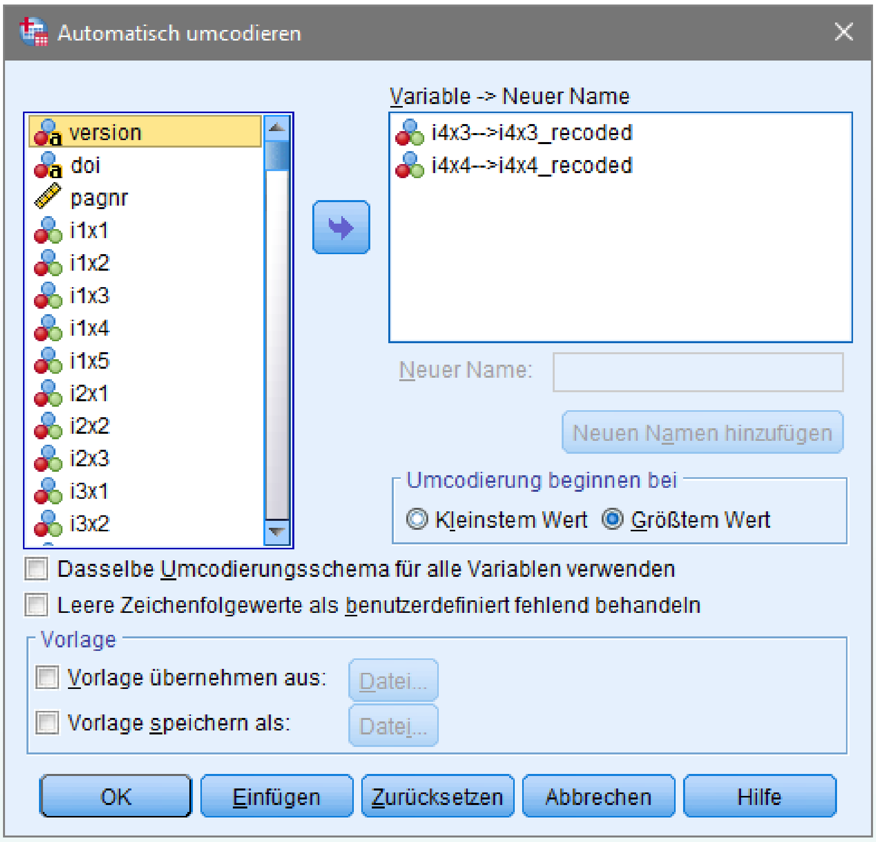
Zunächst wird die Gewichtungsvariable (weight1) aktiviert. Weiters muss die Kategorie 8 der Items (kann ich nicht sagen) als fehlender Wert definiert werden. Wir verwenden dazu folgende Syntax-Befehle:

*Syntax:*

```
WEIGHT BY weight1.
missing values i4x1 i4x2 i4x3 i4x4(8).
```

### Umpolung der Items:

Befehlskette:
***Transformieren → Automatisches umkodieren***




Für die Umpolung werden die Variablen **i4x3** und **i4x4** ausgewählt. Für die neu zu konstruierenden Variablen werden die Namen **i4x3_recoded** und **i4x4_recoded** gewählt. Die Umcodierung beginnt beim größten Wert. Mit dem Button ‚Einfügen' können die entsprechenden Befehle in der Syntaxdatei gespeichert werden.

*Syntax:*

```
AUTORECODE VARIABLES=i4x3 i4x4
 /INTO i4x3_recoded i4x4_recoded
 /DESCENDING
 /PRINT.
```

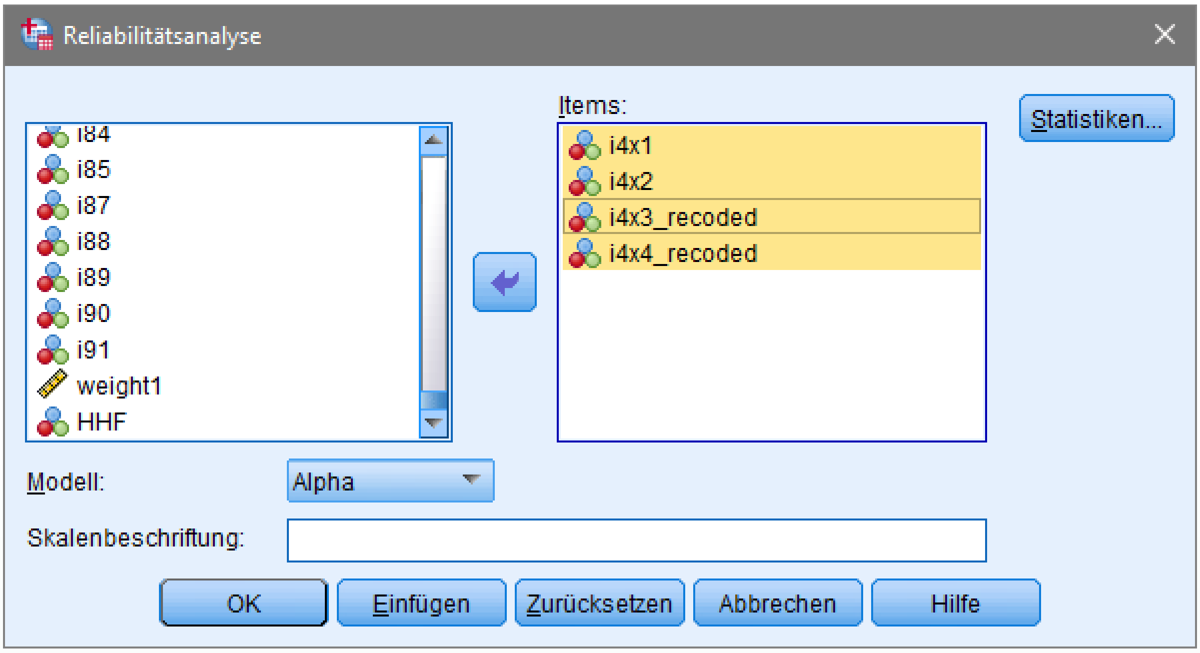
Im Ausgabefenster kann die automatische Umcodierung nachvollzogen werden.

```
i4x3 into i4x3_recoded (Es ist in Ordnung, dass ein Paar zusammenlebt, ohne
    die Absicht zu heiraten.)
                 Old Value   New Value   Value Label

                         5           1   stimme überhaupt nicht zu
                         4           2   stimme eher nicht zu
                         3           3   weder noch
                         2           4   stimme eher zu
                         1           5   stimme voll und ganz zu
                         8M          6M kann ich nicht sagen


i4x4 into i4x4_recoded (Eine Scheidung ist im Allgemeinen die beste Lösung,
    wenn ein Paar seine Eheprobleme nicht lösen kann.)
                 Old Value   New Value   Value Label

                         5           1   stimme überhaupt nicht zu
                         4           2   stimme eher nicht zu
                         3           3   weder noch
                         2           4   stimme eher zu
                         1           5   stimme voll und ganz zu
                         8M          6M kann ich nicht sagen
```

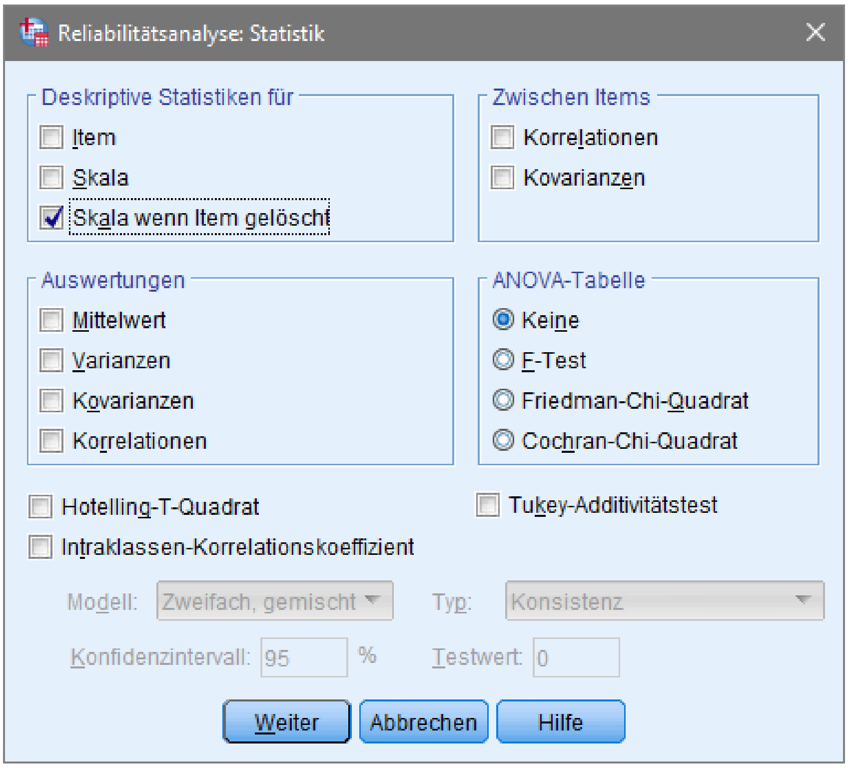
## Reliabilitätsprüfung

Zur Prüfung der Reliabilität wird nun Chronbach's Alpha als Maß der internen Konsistenz der ausgewählten Items **i4x3, i4x4, i4x3_recoded** und **i4x4_recoded** berechnet.

Befehlskette:
***Analysieren → Skala → Relialbilitätsanalyse***




Im Dialogfeld zur Reliabilitätsanalyse werden die entsprechenden Items ausgewählt. Unter ‚Statistiken' erreicht man ein weiteres Dialogfeld mit mehreren Optionen. Von Interesse ist in unserem Fall die Statistik für ‚Skala wenn Item gelöscht'. Damit kann der entsprechende Chronbachsche Alpha-Wert angegeben werden, den die Skala hätte, wenn ein bestimmtes Item weggelassen würde. Ist dieser Alpha-Wert höher als jener der Gesamtskala mit allen Items, so trägt dieses Item zu wenig zur Reliabilität im Sinne der Itemkonsistenz bei, da die Reliabilität der Skala ohne dem entsprechendem Item sogar höher wäre als mit dem Item. In diesem Fall sollte das entsprechende Item aus der Skala ausgeschlossen werden.




*Syntax:*

```
RELIABILITY
 /VARIABLES=i4x1 i4x2 i4x3_recoded i4x4_recoded
 /SCALE('ALL VARIABLES') ALL
 /MODEL=ALPHA
 /SUMMARY=TOTAL.
```

**Reliabilitätsstatistiken**

| Cronbachs Alpha | Anzahl der Items |
|---|---|
| ,686 | 4 |

**Item-Skala-Statistiken**

| | Skalenmittelwert, wenn Item weggelassen | Skalenvarianz, wenn Item weggelassen | Korrigierte Item-Skala-Korrelation | Cronbachs Alpha, wenn Item weggelassen |
|---|---|---|---|---|
| i4x1 Verheiratete Menschen sind im Allgemeinen glücklicher als Menschen, die nicht verheiratet sind. | 10,84 | 6,960 | ,511 | ,593 |
| i4x2 Menschen, die Kinder wollen, sollten heiraten. | 11,05 | 6,259 | ,589 | ,534 |
| i4x3_recoded Es ist in Ordnung, dass ein Paar zusammenlebt, ohne die Absicht zu heiraten. | 10,02 | 7,761 | ,555 | ,574 |
| i4x4_recoded Eine Scheidung ist im Allgemeinen die beste Lösung, wenn ein Paar seine Eheprobleme nicht lösen kann. | 10,16 | 9,607 | ,251 | ,737 |

Der ermittelte Chronbach's Alpha-Wert beträgt 0,686 was für sozialwissenschaftliche Analysen einen akzeptablen Wert bedeutet (Faustregel: ab 0,6 akzeptabel; ab 0,8 gut[1]). Aus der zweiten Tabelle lassen sich noch weitere Kennwerte ablesen, wie etwa die Trennschärfekoeffizienten (Korrigierte Item-Skala-Korrelation). Diese gibt die Korrelation zwischen dem einzelnen Item und der Gesamtskala wieder. Unter der Annahme, dass mehrere Items die gesuchte Dimension besser messen und ein einzelnes Item diese Dimension gut abbildet, müsste die Trennschärfekorrelation hoch sein. Ein niedriger Trennschärfekoeffizient wiest darauf hin, dass das entsprechende Item nicht gut zu den anderen Items passt. In diesem Fall ist bei Item i4x4 ein niedriger Trennschärfekoeffizient zu beobachten. In der letzten Spalte der Tabelle ist abzulesen, dass der Chronbach's Alpha-Wert der Gesamtskala auf 0,737 steigen würden, wenn Item **i4x4** weggelassen werden würde. Dies spricht dafür, dieses Item auszuschließen. Wir machen daher eine erneute Reliabilitätsanalyse mit den verbleibenden drei Items.

[1] Die Faustregel ist in der empirischen Sozialforschung gebräuchlich. Für Skalenbildungen zum Zweck psychologischer Testungen wären strengere Reliabilitätskriterien anzuwenden.

*Syntax:*

```
RELIABILITY
  /VARIABLES=i4x1 i4x2 i4x3_recoded
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /SUMMARY=TOTAL.
```

**Reliabilitätsstatistiken**

| Cronbachs Alpha | Anzahl der Items |
|---|---|
| ,739 | 3 |

**Item-Skala-Statistiken**

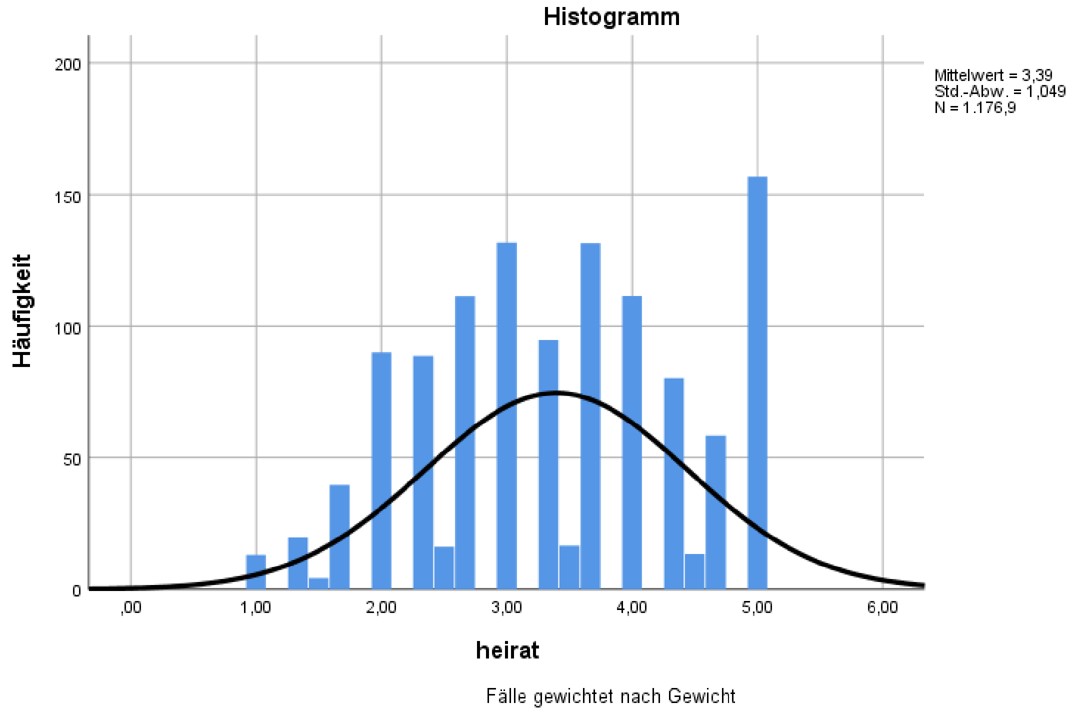
| | Skalenmittelwert, wenn Item weggelassen | Skalenvarianz, wenn Item weggelassen | Korrigierte Item-Skala-Korrelation | Cronbachs Alpha, wenn Item weggelassen |
|---|---|---|---|---|
| i4x1 Verheiratete Menschen sind im Allgemeinen glücklicher als Menschen, die nicht verheiratet sind. | 6,96 | 4,703 | ,551 | ,672 |
| i4x2 Menschen, die Kinder wollen, sollten heiraten. | 7,17 | 4,016 | ,659 | ,533 |
| i4x3_recoded Es ist in Ordnung, dass ein Paar zusammenlebt, ohne die Absicht zu heiraten. | 6,15 | 5,811 | ,506 | ,724 |

Mit Blick auf die letzte Spalte wird ersichtlich, dass sich der Reliabilitätskoeffizient Chronbach's Alpha von 0,739 durch das entfernen weiterer Items nicht mehr erhöhen lässt. Somit bilden wir eine Skala zur Einstellung zur Heirat aus den drei Items i4x1, i4x2 und i4x3_recoded. Wir verwenden dazu folgende Syntax:

*Syntax:*

```
compute heirat=mean(i4x1, i4x2, i4x3_recoded).
exe.
```

Die neu konstruierte Skala trägt den Variablennamen ‚heirat' und weißt folgende Verteilung auf:




Inhaltlich stehen in dieser Skala, die Werte von 1 bis 5 annehmen kann, höhere Werte für einen liberalere Einstellung zur Heirat.

Wir wollen nun untersuchen, inwieweit verschiedene Einflussfaktoren die Einstellung zur Heirat beeinflussen. Dazu möchten wir ein lineares Regressionsmodell mit folgenden Merkmalen als unabhängige Variablen (Prädiktoren) erstellen:

- Geschlecht
- Alter
- Bildung
- Religiosität
- Geschlechterrolleneinstellung

**Geschlecht (i56)** und **Alter (i57)** wurden im Fragebogen direkt abgefragt.

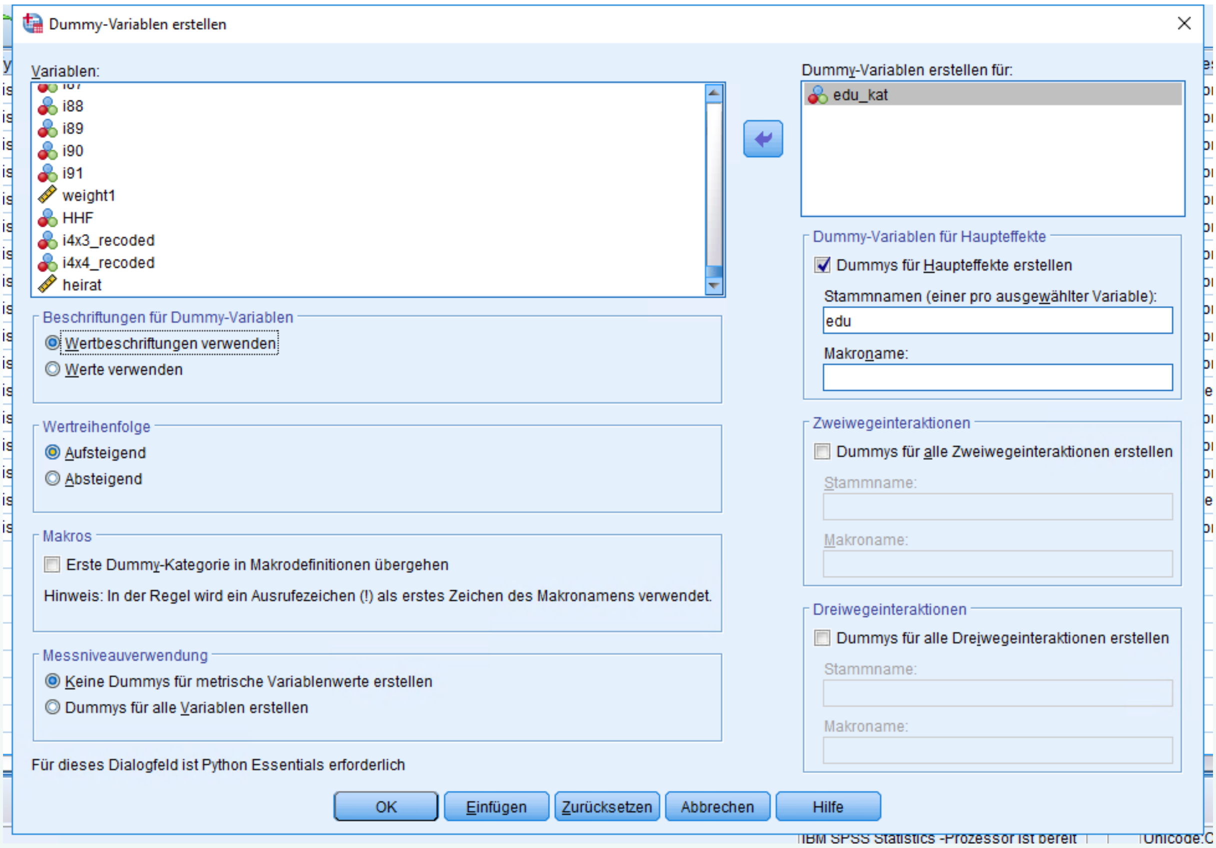
Bildung kann über den **höchsten Schulabschluss (i58)** operationalisiert werden. Dazu werden mit der folgenden Syntax die sieben Kategorien der Variable **i58** zu vier ordinal skalierten Kategorien zusammengefasst (Variable: **edu_kat**).

Syntax:

```
RECODE i58 (7=4) (1 thru 2=1) (3 thru 4=2) (5 thru 6=3) INTO edu_kat.
VARIABLE LABELS  edu_kat 'höchster Schulabschluss'.
EXECUTE.

value labels edu_kat
1 'max. Pflichtschule'
2 'Lehre/BMS'
3 'Matura'
4 'Uni/FH'
```

Die Kategorien der Variable **edu_kat (höchster Schulabschluss**) sollen als Dummy-Variablen in das Regressionsmodell eingehen. Wird der niedrigste Schulabschuss (max. Pflichtschule) als Referenzkategorie gewählt, sind drei Dummy-Variablen für die restlichen Kategorien nötig.

Dummies können in SPSS automatisch erstellt werden.
Befehlskette: ***Transformieren → Dummies-Variablen erstellen***



Im Dialogfeld kann die Variable **edu_kat** ausgewählt werden. Als Stammname für die Dummies wird ***edu*** eingetragen. Die Prozedur erstellt für jede Kategorie der Variable **edu_kat** jeweils eine Dummy-Variable mit den Ausprägungen 0 (trifft nicht zu) und 1 (trifft zu). Diese erhalten die Variablenbezeichnungen:

- **edu_1** (max. Pflichtschule)
- **edu_2** (Lehre/BMS)
- **edu_3** (Matura)
- **edu_4** (Uni/FH)

Da die Kategorie ‚max. Pflichtschule' als Referenzkategorie dienen soll, wird diese in der Analyse nicht benötigt.

*Syntax:*

```
SPSSINC CREATE DUMMIES VARIABLE=edu_kat
ROOTNAME1=edu
/OPTIONS ORDER=A USEVALUELABELS=YES USEML=YES OMITFIRST=NO.
```

Die Variable **i80** (Häufigkeit des Gottesdienstbesuches) soll als Proxy für Religiosität dienen. Die Kategorie 98=kann ich nicht sagen/weiß nicht muss als fehlend gesetzt werden:

*Syntax:*

```
missing values i80 (98).
```

Die Variable Geschlechterrolleneinstellung (genderrole) soll als Index aus den Items der Frage 2 im Fragebogen gebildet werden:

**2. Und inwieweit stimmen Sie den folgenden Aussagen zu oder nicht zu? (Durchfragen)**

| | stimme voll und ganz zu | stimme eher zu | weder noch | stimme eher nicht zu | stimme überhaupt nicht zu | kann ich nicht sagen |
|---|---|---|---|---|---|---|
| Der Mann und die Frau sollten beide zum Haushaltseinkommen beitragen. | 1 | 2 | 3 | 4 | 5 | 8 |
| Die Aufgabe des Mannes ist es, Geld zu verdienen, die der Frau, sich um Haushalt und Familie zu kümmern. | 1 | 2 | 3 | 4 | 5 | 8 |
| Auch wenn beide Eltern voll berufstätig sind, ist es besser, wenn die Verantwortung für den Haushalt und die Kinder hauptsächlich bei der Frau liegt. | 1 | 2 | 3 | 4 | 5 | 8 |

Da Item i2x1 gegenüber den anderen beiden negativ gepolt ist, muss sie zunächst wieder umcodiert werden. Zuvor muss auch die Kategorie 8 (kann ich nicht sagen) als fehlender Wert definiert werden.

*Syntax:*

```
missing values i2x1 i2x2 i2x3 (8).

AUTORECODE VARIABLES=i2x1
  /INTO i2x1_recoded
  /DESCENDING
  /PRINT.
```

Danach wird zur Analyse der Reliabilität wieder der Chonbachsche Alpha-Wert errechnet:

*Syntax:*

```
RELIABILITY
  /VARIABLES=i2x1_recoded i2x2 i2x3
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /SUMMARY=TOTAL.
```

Wir erhalten folgenden Ergebnisoutput:

**Reliabilitätsstatistiken**

| Cronbachs Alpha | Anzahl der Items |
|---|---|
| ,654 | 3 |

**Item-Skala-Statistiken**

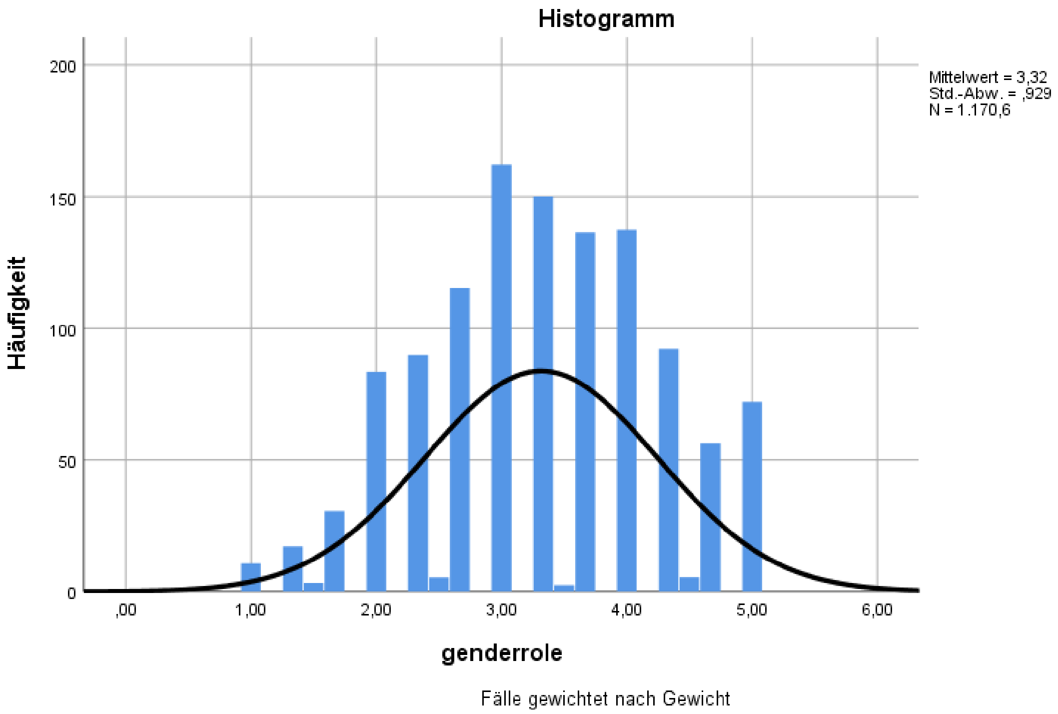
| | Skalenmittelwert, wenn Item weggelassen | Skalenvarianz, wenn Item weggelassen | Korrigierte Item-Skala-Korrelation | Cronbachs Alpha, wenn Item weggelassen |
|---|---|---|---|---|
| Der Mann und die Frau sollten beide zum Haushaltseinkommen beitragen. | 6,11 | 4,656 | ,423 | ,618 |
| Die Aufgabe des Mannes ist es, Geld zu verdienen, die der Frau, sich um Haushalt und Familie zu kümmern. | 6,78 | 3,146 | ,583 | ,379 |
| Auch wenn beide Eltern voll berufstätig sind, ist es besser, wenn die Verantwortung für den Haushalt und die Kinder hauptsächlich bei der Frau liegt. | 7,03 | 3,899 | ,414 | ,631 |

Die Reliabilität liegt mit einem Chronbach's Alpha Wert von 0,654 im akzeptablen Bereich und könnte nicht durch Weglassen eines der drei Items erhöht werden. Der Index wird somit auf Basis aller drei Items gebildet.

*Syntax:*

```
compute genderrole = mean (i2x1_recoded, i2x2, i2x3).
exe.
```

Die Variable **genderrole** weist folgende Verteilung auf: Höhere Werte stehen für egalitärere Geschlechterrolleneinstellungen.

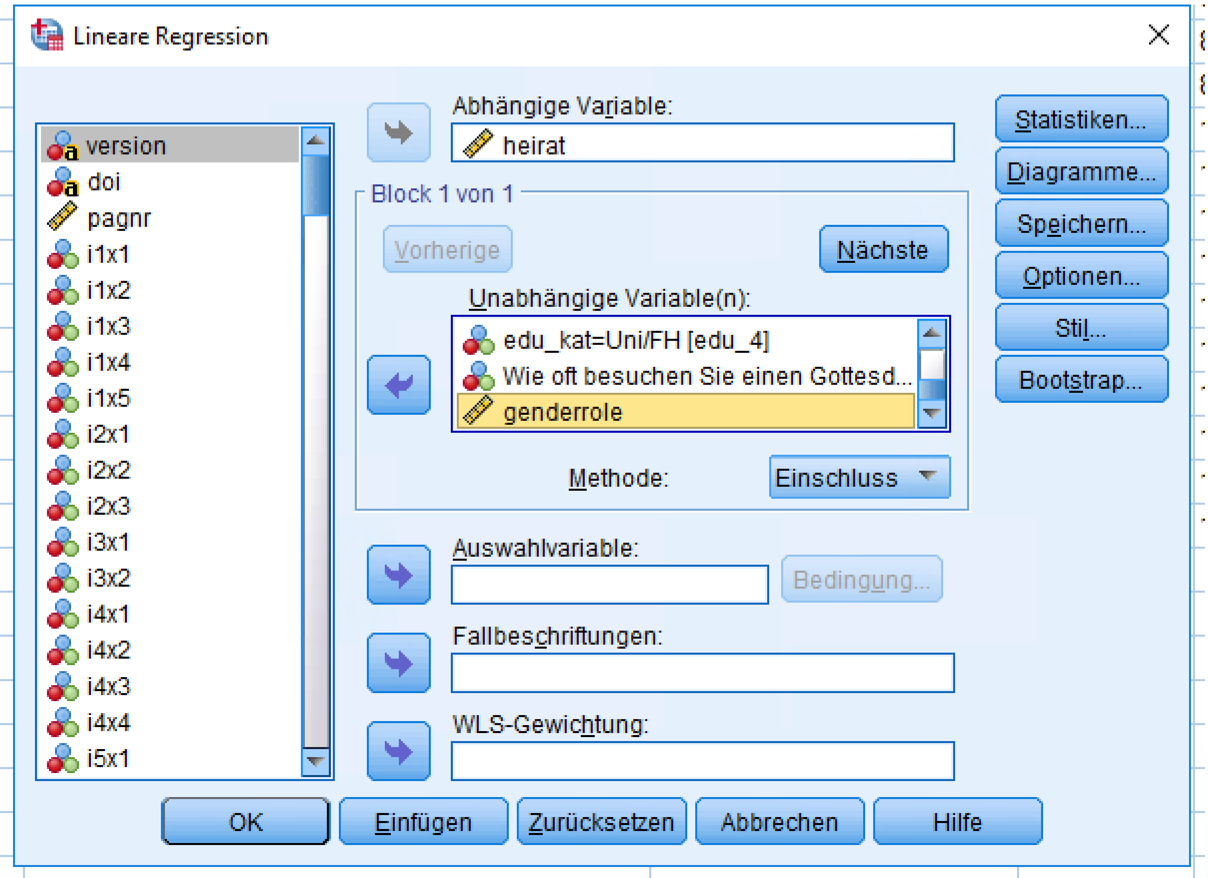



## Durchführung der Regressionsanalyse

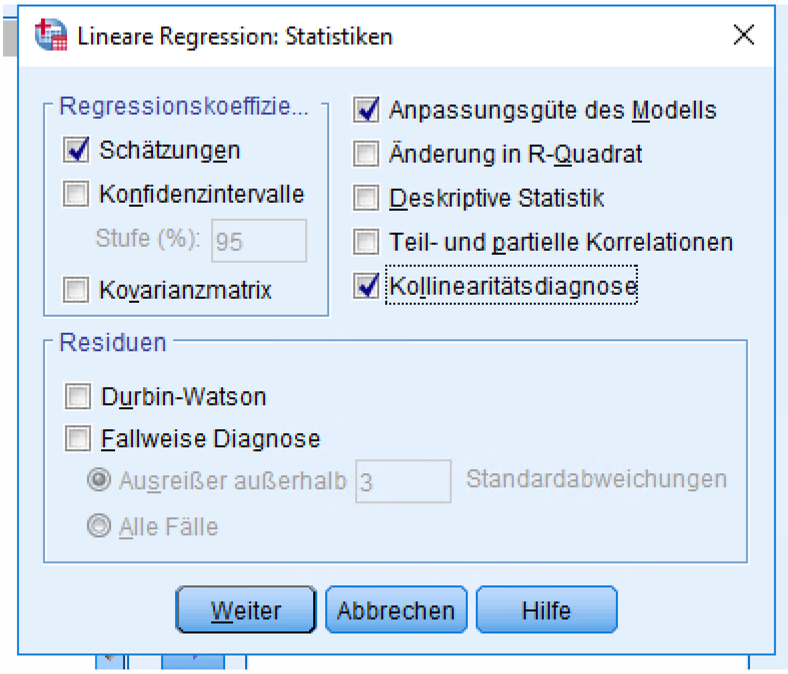
Mit der Befehlskette ***Analysieren → Regression → Linear*** wird das Dialogfeld aufgerufen. Variable **heirat** bildet die Abhängige Variable (Kriterium). Als unabhängige Variablen werden nun ausgewählt:

- **i56**
- **i57**
- **edu_2** (Dummy: Lehre/BMS)
- **edu_3** (Dummy: Matura)
- **edu_4** (Dummy: Uni/FH)
- **i80** (Häufigkeit Gottesdienstbesuch)
- **genderole** (Geschlechterrolleneinstellung)

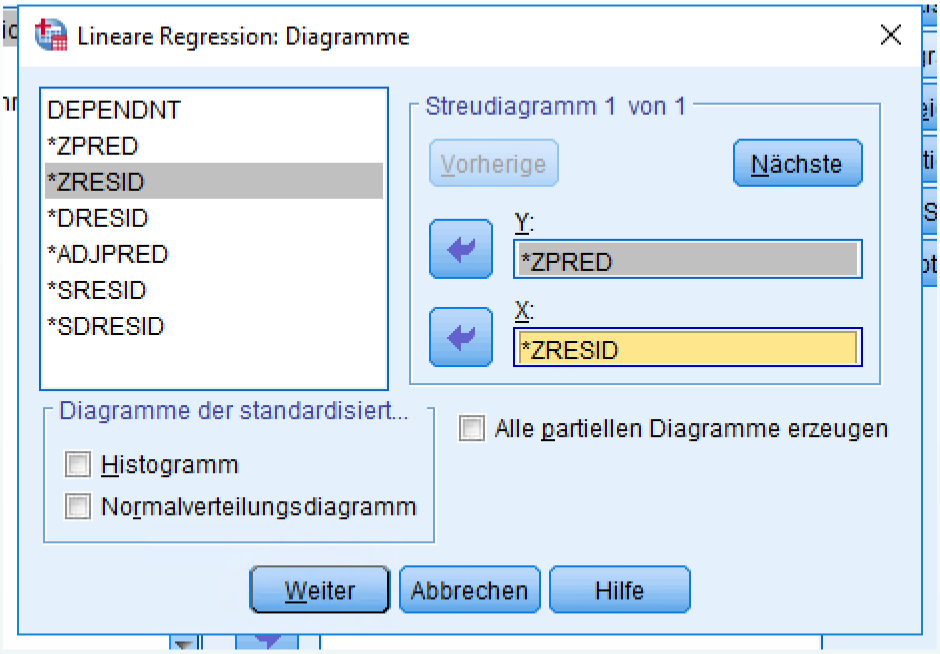
Mit der Methode ‚Einschuss' (Voreinstellung in SPSS) werden alle unabhängigen Variable gleichzeitig ins Modell aufgenommen.




Bei der Berechnung des Regressionsmodells sollten auch etwaige Verletzungen der Modellvoraussetzungen geprüft werden. Solch eine Verletzung würde etwa vorliegen, wenn einzelne oder mehrere Prädiktoren im Modell stark miteinander korreliert sind, sodass ein Prädiktor weitestgehend durch die weiteren Prädiktoren im Modell erklärt werden kann. In diesem Fall spricht man von Multikollinearität. Zur Überprüfung von Multikollinearität wird üblicherweise der Kennwert ‚Toleranz' herangezogen. Dieser errechnet sich aus $1 - R^2$ für die jeweilige Vorhersage des Prädiktors durch die verbleibenden Prädiktoren. Niedrige Toleranzwerte weisen somit auf hohe Multikollinearität hin. Ein weiterer Kennwert für Multikollinearität ist der ‚Varianzinflationsfaktor' (VIF), der als Kehrwert der Toleranz errechnet wird. Beide Kennwerte können im SPSS-Dialogfeld angefordert werden, indem unter dem Punkt ‚Statistiken' die Option ‚Kollinearitätsdiagnose' ausgewählt wird.




Eine weitere zu prüfende Voraussetzung ist das Vorliegen von Homoskedastizität. Dies bedeutet, dass die Varianz der Residuen unabhängig von den x-Werten sein sollte. Ist dies der Fall spricht man von Heteroskedastizität. Die Überprüfung erfolgt durch visuelle Inspektion, indem die standardisierten Residuen gegen die standardisierten vorhergesagten Werte geplottet werden. Typisches Merkmal für Heteroskedastizität ist eine Trichterform im Streudiagramm. Das Diagramm kann im Dialogfeld unter dem Punkt ‚Diagramme' angewählt werden.



Das Regressionsmodell sowie die angewählten Optionen können auch mit folgenden Syntax-Befehlen aufgerufen werden:

*Syntax:*

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT heirat
  /METHOD=ENTER i56 i57 edu_2 edu_3 edu_4 i80 genderrole
  /SCATTERPLOT=(*ZRESID ,*ZPRED).
```

## Ergebnisse

Die erste Tabelle im Ergebnisoutput mit dem Titel ‚Modellzusammenfassung' informiert über die Güte der Vorhersage. Das Bestimmtheitsmaß $R^2$ besagt in unserem Fall, dass 31,2% der Varianz der abhängigen Variable durch die Prädiktoren im Modell erklärt werden kann. Da es mehrere Prädiktoren im Modell gibt, ist das korrigierte $R^2$ zur Interpretation heranzuziehen, welches aber nahezu gleich hoch ist.

**Modellzusammenfassung[b]**

| Modell | R | R-Quadrat | Korrigiertes R-Quadrat | Standardfehler des Schätzers |
|---|---|---|---|---|
| 1 | ,559[a] | ,312 | ,308 | ,87375 |

a. Einflußvariablen : (Konstante), genderrole, Geschlecht, edu_kat=Lehre/BMS, Wie oft besuchen Sie einen Gottesdienst?, Sagen Sie mir bitte, wie alt Sie sind?, edu_kat=Uni/FH, edu_kat=Matura

b. Abhängige Variable: heirat

In der Tabelle „ANOVA“ wird das Bestimmtheitsmaßes $R^2$ auf Signifikanz geprüft. Die letzte Spalte mit der Überschrift „Sig.“ enthält die Irrtumswahrscheinlichkeit p beim Verwerfen der Nullhypothese (Die Nullhypothese lautet: $R^2$ in der Grundgesamtheit ist 0). Da diese Irrtumswahrscheinlichkeit geringer als 0,001 und somit kleiner als 1% ausfällt, kann die Nullhypothese verworfen werden.

**ANOVA[a]**

| Modell | | Quadratsumme | df | Mittel der Quadrate | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 395,648 | 7 | 56,521 | 74,035 | ,000[b] |
| | Nicht standardisierte Residuen | 871,410 | 1141 | ,763 | | |
| | Gesamt | 1267,058 | 1148 | | | |

a. Abhängige Variable: heirat

b. Einflußvariablen : (Konstante), genderrole, Geschlecht, edu_kat=Lehre/BMS, Wie oft besuchen Sie einen Gottesdienst?, Sagen Sie mir bitte, wie alt Sie sind?, edu_kat=Uni/FH, edu_kat=Matura

Aus der Tabelle ‚Koeffizienten‘ können schließlich die Regressionskoeffizienten sowie die zugehörige Signifikanzüberprüfung abgelesen werden. Die erste Spalte enthält die unstandardisierten Regressionskoeffizienten (B). Diese geben an, um wie viel sich die y-Werte ändern, wenn sich die entsprechenden x-Werte um eine Einheit ändern. zB: Mit jedem Lebensjahr, erhöht sich der Wert zur Einstellung zur Heirat um -0,008 Skalenpunkte. Da die abhängige Variable eine Einstellungsmessung anhand einer konstruierten Skala ist und auch die unabhängigen Variablen in ihren Skaleneinheiten nicht vergleichbar sind, können die unstandardisierte Regressionskoeffizienten kaum sinnvoll interpretiert werden. Zur Interpretation können aber die standardisierten Regressionskoeffizienten (Beta-Werte) herangezogen werden. Diese geben an, um wie viele Standardabweichungen sich y ändert, wenn sich die entsprechenden x-Werte um eine Standardabweichung ändern. Aufgrund der Standardisierung können die Beta-Werte untereinander verglichen werden und ihre Höhe als Einflussstärke auf die abhängige Variable interpretiert werden. Die Spalte ‚Sign.‘ enthält die p-Werte, also die Irrtumswahrscheinlichkeit beim Verwerfen der Nullhypothese.

Demnach zeigt sich, dass die Geschlechterrolleneinstellung mit einem Beta-Wert von 0,36 die deutlich stärkste Vorhersagekraft für die Einstellung zur Heirat im Modell hat. Je traditioneller die Einstellung zu Geschlechterrollen, desto traditioneller auch die Einstellung zur Heirat. Unabhängig von der Geschlechterrolleneinstellung zeigen auch Religiosität und Alter signifikanten Einfluss auf die Einstellung zu Heirat. Die Einstellung ist je traditioneller, je häufiger Gottesdienste besucht werden und je älter die Personen sind. Sowohl Geschlecht als auch Bildung zeigen hingegen keinen signifikanten Einfluss auf die Einstellung zur Heirat.

**Koeffizienten[a]**

| Modell | | Nicht standardisierte Koeffizienten RegressionskoeffizientB | Std.-Fehler | Standardisierte Koeffizienten Beta | T | Sig. | Kollinearitätsstatistik Toleranz | VIF |
|---|---|---|---|---|---|---|---|---|
| 1 | (Konstante) | 1,509 | ,182 | | 8,290 | ,000 | | |
| | Geschlecht | ,102 | ,052 | ,048 | 1,949 | ,052 | ,977 | 1,024 |
| | Sagen Sie mir bitte, wie alt Sie sind? | -,008 | ,002 | -,134 | -4,967 | ,000 | ,828 | 1,208 |
| | edu_kat=Lehre/BMS | ,031 | ,077 | ,015 | ,400 | ,689 | ,449 | 2,228 |
| | edu_kat=Matura | ,024 | ,097 | ,008 | ,243 | ,808 | ,510 | 1,961 |
| | edu_kat=Uni/FH | -,062 | ,102 | -,020 | -,607 | ,544 | ,568 | 1,762 |
| | Wie oft besuchen Sie einen Gottesdienst? | ,124 | ,012 | ,274 | 10,417 | ,000 | ,873 | 1,145 |
| | genderrole | ,404 | ,030 | ,360 | 13,568 | ,000 | ,858 | 1,165 |

a. Abhängige Variable: heirat

Die letzten beiden Spalten in der Tabelle geben die Toleranz- und VIF-Werte zur Prüfung von Multikollinearität an. Als Faustregel werden Toleranzwerte unter 0,1 und VIF-Werte ab 10 als kritisch angesehen. In unserem Fall ist keine Gefahr von Multikollinearität gegeben.

Das Streudiagramm dient schließlich zur visuellen Überprüfung der Verteilung der Residuen um die Regressionsgerade. Diese sollten möglichst homogen um die vorhergesagten Werte streuen. Die visuelle Überprüfung beinhaltet immer einen gewissen Interpretationsspielraum. In unserem Fall ist keine Trichterform zu erkennen, sodass von Homoskedastizität und somit von keiner Verletzung der Modellvoraussetzungen ausgegangen werden kann.