

Meaningful Data Counts: A survey on data citation and reuse practices

Codebook and Survey Overview

This study investigates how researchers reuse and cite data in their work. Data reuse refers to using data (for any purpose) which others have created. This study is funded by the Alfred P. Sloan Foundation and is conducted by researchers at the Scholarly Communications Lab at the University of Ottawa, Canada as part of the Meaningful Data Counts project, a collaboration involving the non-profit organization DataCite.

The survey was delivered in 2021 to researchers from a diversity of communities who reuse and cite data, as well as from those who do not.

Codebook

Variable: RespondentID

Label: Respondent identifier

Question text: NA

Universe: All respondents

Variable: Q1

Label: Ever reused data

Question text: Have you ever reused data which other people have created, for any purpose?

Universe: All respondents

Response categories	Code	Frequency
No	1	466
Yes	2	2026

Variable: Q2

Label:

Question text: Please describe the type of secondary data that you reuse (optional).

Universe: All respondents

Note : Not included in dataset

Variable: Q3

Label: Data type

Question text: How would you classify the type of data that you reuse?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Qualitative data	1	136
Quantitative data	2	913
Both	3	945
Other	4	32
Missing	999	465
N		2026

Variable: Q4.1

Label: As the basis for a study

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	175
Rarely	2	330
Sometimes	3	636
Often	4	642
Always	5	243
Missing	999	466
N		2026

Variable: Q4.2

Label: To prepare for a new project/proposal or to generate new ideas

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	131
Rarely	2	293
Sometimes	3	705
Often	4	668

Always	5	229
Missing	999	466
N		2026

Variable: Q4.3

Label: To integrate with other data

Question text: How frequently do you reuse secondary data for the following purposes?

Note: To integrate with other data (e.g. to create a new dataset)

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	272
Rarely	2	379
Sometimes	3	604
Often	4	612
Always	5	159
Missing	999	466
N		2026

Variable: Q4.4

Label: As model, algorithm or system inputs

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	560
Rarely	2	394
Sometimes	3	477
Often	4	463
Always	5	132
Missing	999	466
N		2026

Variable: Q4.5

Label: To calibrate instruments or models

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	790
Rarely	2	426
Sometimes	3	436
Often	4	280
Always	5	94
Missing	999	466
N		2026

Variable: Q4.6

Label: To verify my own data

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	316
Rarely	2	296
Sometimes	3	638
Often	4	593
Always	5	183
Missing	999	466
N		2026

Variable: Q4.7

Label: To identify trends, to make comparisons or to make predictions

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	145
Rarely	2	220
Sometimes	3	558
Often	4	834
Always	5	269
Missing	999	466
N		2026

Variable: Q4.8

Label: To create visualizations or summaries

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	284
Rarely	2	358
Sometimes	3	641
Often	4	600
Always	5	143
Missing	999	466
N		2026

Variable: Q4.9

Label: For teaching/training

Question text: How frequently do you reuse secondary data for the following purposes?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	193
Rarely	2	244
Sometimes	3	542
Often	4	736
Always	5	311
Missing	999	466
N		2026

Variable: Q5.1

Label: The data themselves

Question text: When you reuse data, what do you usually cite or reference?

Note: The data themselves (e.g. a particular dataset or record)

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	194
Rarely	2	238
Sometimes	3	309

Often	4	354
Always	5	847
N/A	6	84
Missing	999	466
N		2026

Variable: Q5.2

Label: The source of the data

Question text: When you reuse data, what do you usually cite or reference?

Note: The source of the data (e.g. a data repository or database)

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	135
Rarely	2	130
Sometimes	3	272
Often	4	307
Always	5	1123
N/A	6	59
Missing	999	466
N		2026

Variable: Q5.3

Label: A collection of related data

Question text: When you reuse data, what do you usually cite or reference?

Note: A collection of related data (e.g. a series of datasets)

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	399
Rarely	2	362
Sometimes	3	438
Often	4	252
Always	5	333
N/A	6	242
Missing	999	466
N		2026

Variable: Q5.4

Label: An article or publication analyzing the data

Question text: When you reuse data, what do you usually cite or reference?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	70
Rarely	2	93
Sometimes	3	258
Often	4	457
Always	5	1102
N/A	6	46
Missing	999	466
N		2026

Variable: Q5.5

Label: A data paper

Question text: When you reuse data, what do you usually cite or reference?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	462
Rarely	2	357
Sometimes	3	347
Often	4	231
Always	5	430
N/A	6	199
Missing	999	466
N		2026

Variable: Q6.1

Label: Including a citation to a data in reference lists

Question text: When you reuse data, how do you cite or reference them?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	185

Rarely	2	141
Sometimes	3	218
Often	4	417
Always	5	1065
Missing	999	466
N		2026

Variable: Q6.2

Label: Including a citation to a related paper in reference lists

Question text: When you reuse data, how do you cite or reference them?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	124
Rarely	2	114
Sometimes	3	307
Often	4	496
Always	5	985
Missing	999	466
N		2026

Variable: Q6.3

Label: Mentioning data in a footnote

Question text: When you reuse data, how do you cite or reference them?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	842
Rarely	2	438
Sometimes	3	413
Often	4	181
Always	5	152
Missing	999	466
N		2026

Variable: Q6.4

Label: Mentioning data in body of text

Question text: When you reuse data, how do you cite or reference them?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	174
Rarely	2	175
Sometimes	3	457
Often	4	575
Always	5	645
Missing	999	466
N		2026

Variable: Q6.5

Label: Mentioning data in acknowledgements

Question text: When you reuse data, how do you cite or reference them?

Note: Mentioning data in acknowledgements, appendix, etc.

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	570
Rarely	2	471
Sometimes	3	489
Often	4	275
Always	5	221
Missing	999	466
N		2026

Variable: Q6.6

Label: Mentioning data in captions, figures or tables

Question text: When you reuse data, how do you cite or reference them?

Universe: Responded yes to Q1

Response categories	Code	Frequency
Never	1	273
Rarely	2	287
Sometimes	3	549
Often	4	497
Always	5	420
Missing	999	466
N		2026

Variable: Q7_Q15

Label: Want another person to cite/reference your data

Question text: If another person were to reuse data which you have created, would you want them to cite/reference your data?

Universe: All respondents

Response categories	Code	Frequency
No	1	37
Yes	2	2455
N		2492

Variable: Q8_Q16.1

Label: The data themselves

Question text: What would you prefer that other people cite/reference when they use your data?
Note: The data themselves (e.g. a particular dataset or record)

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	1282
Selected	1	1173
Missing	999	37
N		2455

Variable: Q8_Q16.2

Label: The source of the data

Question text: What would you prefer that other people cite/reference when they use your data?
Note: The source of the data (e.g. a data repository or database)

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	1100
Selected	1	1355
Missing	999	37
N		2455

Variable: Q8_Q16.3

Label: A collection of related data

Question text: What would you prefer that other people cite/reference when they use your data?

Note: A collection of related data (e.g. a series or database)

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	2227
Selected	1	228
Missing	999	37
N		2455

Variable: Q8_Q16.4

Label: An article or publication analyzing the data

Question text: What would you prefer that other people cite/reference when they use your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	385
Selected	1	2070
Missing	999	37
N		2455

Variable: Q8_Q16.5

Label: A data paper

Question text: What would you prefer that other people cite/reference when they use your data?

Note: A data paper (A data paper is a short article with in-depth descriptions about the particular data)

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	1706
Selected	1	749
Missing	999	37
N		2455

Variable: Q8_Q16.6

Label: What do you want people to cite: Other

Question text: What would you prefer that other people cite/reference when they use your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	708
Selected	1	1747
Missing	999	37
N		2455

Variable: Q9_Q17.1

Label: Including a citation to data in reference lists

Question text: How would you prefer other people to cite/reference your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	708
Selected	1	1747
Missing	999	37
N		2455

Variable: Q9_Q17.2

Label: Including a citation to a related paper in reference lists

Question text: How would you prefer other people to cite/reference your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	759
Selected	1	1696
Missing	999	37
N		2455

Variable: Q9_Q17.3

Label: Mentioning data in a footnote

Question text: How would you prefer other people to cite/reference your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	2057
Selected	1	398

Missing	999	37
N		2455

Variable: Q9_Q17.4

Label: Mentioning data in body of text

Question text: How would you prefer other people to cite/reference your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	1292
Selected	1	1163
Missing	999	37
N		2455

Variable: Q9_Q17.5

Label: Mentioning data in acknowledgements

Question text: How would you prefer other people to cite/reference your data?

Note: Mentioning data in acknowledgements, appendix, etc.

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	1879
Selected	1	576
Missing	999	37
N		2455

Variable: Q9_Q17.6

Label: Mentioning data in captions, figures or tables

Question text: How would you prefer other people to cite/reference your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	1464
Selected	1	991
Missing	999	37
N		2455

Variable: Q9_Q17.7

Label: How do you want people to cite your data: Other

Question text: How would you prefer other people to cite/reference your data?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
Not selected	0	2388
Selected	1	67
Missing	999	37
N		2455

Variable: Q10

Label: Do you cite secondary data

Question text: Do you cite secondary data (by including a data citation in a reference list) in any way for any purpose?

Universe: Responded yes to Q7_Q15

Response categories	Code	Frequency
No	1	150
Yes	2	1407
Sometimes	3	469
Missing	999	37
N		2455

Variable: Q11.1

Label: As a way of showing intellectual debt to the data creator/data provider

Question text: Why do you cite secondary data?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	404
Selected	1	1472
Missing	999	616
N		1876

Variable: Q11.2

Label: As a way of supporting the validity of your claims

Question text: Why do you cite secondary data?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	542
Selected	1	1334
Missing	999	616
N		1876

Variable: Q11.3

Label: As a way of helping others to locate and access the data you used

Question text: Why do you cite secondary data?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	450
Selected	1	1426
Missing	999	616
N		1876

Variable: Q11.4

Label: As a way of rewarding the data creators/data providers

Question text: Why do you cite secondary data?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	847
Selected	1	1029
Missing	999	616
N		1876

Variable: Q11.5

Label: As a way of indicating that you have used the data in some way

Question text: Why do you cite secondary data?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	650
Selected	1	1226
Missing	999	616
N		1876

Variable: Q11.6

Label: Because you were advised to cite data

Question text: Why do you cite secondary data?

Note: Because you were advised to cite data (e.g. by a journal editor)

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	1721
Selected	1	155
Missing	999	616
N		1876

Variable: Q11.7

Label: Why do you cite: Other

Question text: Why do you cite secondary data?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	1817
Selected	1	59
Missing	999	616
N		1876

Variable: Q12.1

Label: To correct your own data

Question text: Do you ever cite data for the following reasons?

Note: To correct your own data (you cite your own data)

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	1482
Selected	1	394
Missing	999	616
N		1876

Variable: Q12.2

Label: To build on or use data you have created

Question text: Do you ever cite data for the following reasons?

Note: To build on or use data you have created (you cite your own data)

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	823
Selected	1	1053
Missing	999	616
N		1876

Variable: Q12.3

Label: To criticize or correct the data of others

Question text: Do you ever cite data for the following reasons?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	1376
Selected	1	500
Missing	999	616
N		1876

Variable: Q12.4

Label: To bring recognition to your own data

Question text: Do you ever cite data for the following reasons?

Note: To bring recognition to your own data (you cite your own data)

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	1286
Selected	1	590
Missing	999	616
N		1876

Variable: Q12.5

Label: None of the above

Question text: Do you ever cite data for the following reasons?

Universe: Responded yes or sometimes to Q10

Response categories	Code	Frequency
Not selected	0	1296
Selected	1	580
Missing	999	616
N		1876

Variable: Q13.1

Label: Reusing data is not relevant to your research methods

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	257
Selected	1	209
Missing	999	2026
N		466

Variable: Q13.2

Label: Reusing data is not normal in your research community

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	321
Selected	1	145
Missing	999	2026
N		466

Variable: Q13.3

Label: There are no available relevant data for your research topic

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	343
Selected	1	123
Missing	999	2026
N		466

Variable: Q13.4

Label: You get more credit for your work if you create your own data

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	349
Selected	1	117
Missing	999	2026
N		466

Variable: Q13.5

Label: You do not trust data that others create

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	413
Selected	1	53
Missing	999	2026
N		466

Variable: Q13.6

Label: You cannot find the data you need

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	388
Selected	1	78
Missing	999	2026
N		466

Variable: Q13.7

Label: You do not know how to give credit for data created by other people

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	430
Selected	1	36
Missing	999	2026
N		466

Variable: Q13.8

Label: You did not know that you could reuse data created by other people

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	417
Selected	1	49
Missing	999	2026
N		466

Variable: Q13.9

Label: Reasons for not reusing data: Other

Question text: What are your reasons for not reusing data created by other people?

Universe: Responded no to Q1

Response categories	Code	Frequency
Not selected	0	409
Selected	1	57
Missing	999	2026
N		466

Variable: Q18.1

Label: Datacite

Question text: Are you aware of and do you use recommendations/standards for citing data from the following bodies?

Universe: All respondents

Response categories	Code	Frequency
I am aware of and use this recommendation	1	448
I am aware of and do not use this recommendation	2	201
I am unaware of this recommendation	3	1843
N		2492

Variable: Q18.2

Label : Data repositories

Question text : Are you aware of and do you use recommendations/standards for citing data from the following bodies?

Universe: All respondents

Response categories	Code	Frequency
I am aware of and use this recommendation	1	1059
I am aware of and do not use this recommendation	2	361
I am unaware of this recommendation	3	1072
N		2492

Variable: Q18.3

Label : Scientific societies

Question text : Are you aware of and do you use recommendations/standards for citing data from the following bodies?

Note: Scientific societies (e.g. the American Geophysical Union)

Universe: All respondents

Response categories	Code	Frequency
I am aware of and use this recommendation	1	684
I am aware of and do not use this recommendation	2	305
I am unaware of this recommendation	3	1503
N		2492

Variable: Q18.4

Label : Journals/publishers

Question text : Are you aware of and do you use recommendations/standards for citing data from the following bodies?

Note : Journals or publishers

Universe: All respondents

Response categories	Code	Frequency
I am aware of and use this recommendation	1	1913
I am aware of and do not use this recommendation	2	147
I am unaware of this recommendation	3	432
N		2492

Variable: Q18.5

Label : Citation style guides

Question text : Are you aware of and do you use recommendations/standards for citing data from the following bodies?

Note : Citation style guides (e.g. APA, MLA, etc.)

Universe: All respondents

Response categories	Code	Frequency
I am aware of and use this recommendation	1	1487
I am aware of and do not use this recommendation	2	214
I am unaware of this recommendation	3	791
N		2492

Variable: Q19.1

Label: Having data be recognized as a standalone research output, separate from publications

Question text: How important are the following to you?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	105
Unimportant	2	289
Neither important nor unimportant	3	632
Important	4	1075
Extremely important	5	391
N		2492

Variable: Q19.2

Label: Having the creation of good data documentation and workflows be rewarded

Question text: How important are the following to you?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	74
Unimportant	2	114
Neither important nor unimportant	3	368
Important	4	1190
Extremely important	5	746
N		2492

Variable: Q20

Label: Ever shared own data

Question text: Have you ever shared your own research data?

Universe: All respondents

Response categories	Code	Frequency
No	1	485
Yes	2	2007
N		2492

Variable: Q21

Label: Used own data multiple times

Question text: Have you ever used your own data multiple times (e.g. for new studies, or to produce multiple publications?)

Universe: All respondents

Response categories	Code	Frequency
No	1	671
Yes	2	1821
N		2492

Variable: Q22.1

Label: Assess the reach and influence of your own data

Question text: How important is it to you to do the following?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	44
Unimportant	2	157
Neither important nor unimportant	3	495
Important	4	1292
Extremely important	5	438
N/A	6	66
N		2492

Variable: Q22.2

Label: Get credit specifically for sharing your own data

Question text: How important is it to you to do the following?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	69
Unimportant	2	225

Neither important nor unimportant	3	561
Important	4	1111
Extremely important	5	452
N/A	6	74
N		2492

Variable: Q22.3

Label: Assess the reach and influence of other people's data which you reuse

Question text: How important is it to you to do the following?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	72
Unimportant	2	286
Neither important nor unimportant	3	690
Important	4	1037
Extremely important	5	297
N/A	6	110
N		2492

Variable: Q22.4

Label: Get credit for reusing other people's data

Question text: How important is it to you to do the following?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	179
Unimportant	2	557
Neither important nor unimportant	3	873
Important	4	554
Extremely important	5	189
N/A	6	140
N		2492

Variable: Q23.1

Label: The number of citations the data have received

Question text: How important would it be for you to know the following information about others' data which you may potentially reuse?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	180
Unimportant	2	489
Neither important nor unimportant	3	604
Important	4	1008
Extremely important	5	211
N		2492

Variable: Q23.2

Label: The number of times the data have been downloaded

Question text: How important would it be for you to know the following information about others' data which you may potentially reuse?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	216
Unimportant	2	601
Neither important nor unimportant	3	772
Important	4	783
Extremely important	5	120
N		2492

Variable: Q23.3

Label: The number of times the data have been viewed

Question text: How important would it be for you to know the following information about others' data which you may potentially reuse?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	243
Unimportant	2	676
Neither important nor unimportant	3	876
Important	4	612
Extremely important	5	85
N		2492

Variable: Q23.4

Label: Information about where the data were used

Question text: How important would it be for you to know the following information about others' data which you may potentially reuse?

Note: Information about where the data were used (e.g., in which country, at which institution)

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	181
Unimportant	2	438
Neither important nor unimportant	3	688
Important	4	972
Extremely important	5	213
N		2492

Variable: Q23.5

Label: Descriptions or a narrative providing details about how the data were used

Question text: How important would it be for you to know the following information about others' data which you may potentially reuse?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	115
Unimportant	2	297
Neither important nor unimportant	3	557
Important	4	1156
Extremely important	5	367
N		2492

Variable: Q23.6

Label: Information about who has used the data

Question text: How important would it be for you to know the following information about others' data which you may potentially reuse?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	187
Unimportant	2	504
Neither important nor unimportant	3	809
Important	4	844
Extremely important	5	148
N		2492

Variable: Q23.7

Label: If the data have received recognition outside the scholarly system

Question text: How important would it be for you to know the following information about others' data which you may potentially reuse?

Note: If the data have received recognition outside the scholarly system, e.g. in newspapers

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	273
Unimportant	2	589
Neither important nor unimportant	3	792
Important	4	697
Extremely important	5	141
N		2492

Variable: Q24.1

Label: The number of citations your data have received

Question text: How important would it be for you to know the following information about your data?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	54
Unimportant	2	126
Neither important nor unimportant	3	277
Important	4	1441
Extremely important	5	594
N		2492

Variable: Q24.2

Label: The number of times your data have been downloaded

Question text: How important would it be for you to know the following information about your data?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	70
Unimportant	2	209
Neither important nor unimportant	3	473
Important	4	1364
Extremely important	5	376
N		2492

Variable: Q24.3

Label: The number of times your data have been viewed

Question text: How important would it be for you to know the following information about your data?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	75
Unimportant	2	270
Neither important nor unimportant	3	687
Important	4	1204
Extremely important	5	256
N		2492

Variable: Q24.4

Label: Information about where your data were used

Question text: How important would it be for you to know the following information about your data?

Note: Information about where your data were used (e.g., in which country, at which institution)

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	80
Unimportant	2	199
Neither important nor unimportant	3	550
Important	4	1287
Extremely important	5	376
N		2492

Variable: Q24.5

Label: Descriptions or a narrative providing details about how your data were used

Question text: How important would it be for you to know the following information about your data?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	55
Unimportant	2	164
Neither important nor unimportant	3	542
Important	4	1274
Extremely important	5	457
N		2492

Variable: Q24.6

Label: Information about who has used your data

Question text: How important would it be for you to know the following information about your data?

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	63
Unimportant	2	227
Neither important nor unimportant	3	600
Important	4	1251
Extremely important	5	351
N		2492

Variable: Q24.7

Label: If your data have received recognition outside the scholarly system

Question text: How important would it be for you to know the following information about your data?

Note: If your data have received recognition outside the scholarly system, e.g. in newspapers

Universe: All respondents

Response categories	Code	Frequency
Extremely unimportant	1	63
Unimportant	2	227
Neither important nor unimportant	3	600
Important	4	1251
Extremely important	5	351
N		2492

Variable: Q25

Label: Years of professional experience

Question text: How many years of professional experience do you have in your field?

Universe: All respondents

Response categories	Code	Frequency
0-5	1	157
6-15	2	811
16-30	3	904
30+	4	620
N		2492

Variable: Q26_FOS

Label: OECD Fields of Research and Technology classification

Question text: With which disciplinary domain do you most identify?

Universe: All respondents

Response categories	Code	Frequency
Natural sciences	1.00	5
Mathematics	1.01	60
Computer and information sciences	1.02	100
Physical sciences	1.03	141
Chemical sciences	1.04	85
Earth and related environmental sciences	1.05	199
Biological sciences	1.06	408
Other natural sciences	1.07	29
Engineering and technology	2.00	8
Civil engineering	2.01	33
Electrical engineering, Electronic engineering, Information engineering	2.02	69
Mechanical engineering	2.03	46
Chemical engineering	2.04	24
Materials engineering	2.05	40
Medical engineering	2.06	11
Environmental engineering	2.07	31
Environmental biotechnology	2.08	3
Industrial biotechnology	2.09	3
Nano-technology	2.10	12
Other engineering and technologies	2.11	39
Medical and Health sciences	3.00	14
Basic medicine	3.01	31
Clinical medicine	3.02	162
Health sciences	3.03	210
Medical biotechnology	3.04	7
Other medical sciences	3.05	60
Agricultural sciences	4.00	5

Agriculture, forestry, and fisheries	4.01	44
Animal and dairy sciences	4.02	12
Veterinary sciences	4.03	21
Agricultural biotechnology	4.04	8
Other agricultural sciences	4.05	16
Social sciences	5.00	3
Psychology	5.01	98
Economics and business	5.02	128
Education sciences	5.03	47
Sociology	5.04	51
Law	5.05	5
Political science	5.06	34
Social and economic geography	5.07	18
Media and communications	5.08	12
Other social sciences	5.09	64
Humanities	6.00	1
History, archaeology and related studies	6.01	30
Languages and literature	6.02	33
Philosophy	6.03	15
Arts (arts, history of arts, performing arts, music), architecture and design	6.04	3
Other humanities	6.05	14
N		2492

Variable: Q26_FOS_Refined

Label: OECD refined levels of Fields of Research and Technology classification

Question text: With which disciplinary domain do you most identify?

Universe: All respondents

Response categories	Code	Frequency
Natural Sciences	1	1027
Engineering and Technology	2	319

Medical and Health Sciences	3	484
Agriculture Sciences	4	106
Social Sciences	5	460
Humanities	6	96
N		2492

Variable: Q27

Label: Research methods

Question text: Which approach best describes the research methods which you use?

Universe: All respondents

Response categories	Code	Frequency
Mostly qualitative	1	230
Mostly quantitative	2	1185
Mixed methods (quantitative and qualitative)	3	1050
Other	4	27
N		2492

Variable: Q28

Label: Organization

Question text: What type of organization do you work for?

Universe: All respondents

Response categories	Code	Frequency
Corporate	1	74
Government agency	2	105
Research institution	3	367
University of college	4	1789
Other	5	157
N		2492

Variable: Q29

Label: Country

Question text: In what country do you currently work?

Universe: All respondents

Response categories	Code	Frequency
Albania	008	1
Algeria	012	7
Antigua and Barbuda	028	1
Argentina	032	15
Australia	036	88
Austria	040	21
Bahrain	048	2
Bangladesh	050	4
Belgium	056	17
Bolivia (Plurinational State of)	068	1
Bosnia and Herzegovina	070	1
Botswana	072	2
Brazil	076	78
Bulgaria	100	8
Cameroon	120	3
Canada	124	160
Chile	152	14
China	156	19
Taiwan (Province of China)	158	17
Colombia	170	21
Costa Rica	188	3
Croatia	191	14
Cuba	192	1
Cyprus	196	6
Czechia	203	18
Denmark	208	28
Dominican Republic (the)	214	1
Ecuador	218	2
Ethiopia	231	12
Estonia	233	6
Finland	246	14
France	250	65
Georgia	268	1
Palestine, State of	275	2
Germany	276	102
Ghana	288	9
Greece	300	13
Guatemala	320	1
Hong Kong	344	1
Hungary	348	11
Iceland	352	19
India	356	94
Indonesia	360	9

Iran (Islamic Republic of)	364	19
Iraq	368	4
Israel	376	20
Italy	380	126
Côte d'Ivoire	384	1
Jamaica	388	1
Japan	392	24
Jordan	400	9
Kenya	404	7
Korea (the Republic of)	410	13
Kuwait	414	1
Lebanon	422	2
Latvia	428	3
Lithuania	440	4
Luxembourg	442	1
Malawi	454	1
Malaysia	458	16
Mali	466	2
Mauritius	480	1
Mexico	484	48
Morocco	504	3
Mozambique	508	1
Oman	512	2
Namibia	516	2
Nepal	524	4
Netherlands (the)	528	31
New Caledonia	540	1
New Zealand	554	25
Nigeria	566	14
Norway	578	40
Pakistan	586	17
Paraguay	600	1
Peru	604	4
Philippines (the)	608	4
Poland	616	35
Portugal	620	22
Puerto Rico	630	3
Qatar	634	2
Romania	642	19
Russian Federation (the)	643	25
Saudi Arabia	682	11
Serbia	688	19
Singapore	702	9
Slovakia	703	8
Viet Nam	704	4
Slovenia	705	8

South Africa	710	24
Spain	724	97
Sudan (the)	729	1
Sweden	752	37
Switzerland	756	19
Thailand	764	7
Togo	768	1
Trinidad and Tobago	780	1
United Arab Emirates (the)	784	7
Tunisia	788	8
Turkey	792	30
Uganda	800	6
Ukraine	804	4
North Macedonia	807	3
Egypt	818	15
United Kingdom of Great Britain and Northern Ireland (the)	826	114
Tanzania, the United Republic of	834	2
United States of America (the)	840	581
Uruguay	858	2
Venezuela (Bolivarian Republic of)	862	3
Zambia	894	1
N		2492

Variable: Q30

Label: Gender

Question text: How do you identify your gender

Universe: All respondents

Response categories	Code	Frequency
Man	1	1631
Woman	2	805
Non-binary	3	5
Option not listed	4	1
Prefer not to say	5	50
N		2492

Variable: Weight

Label: Weight discipline

Universe: All respondents

Variable: Weight2

Label: Weight Discipline by career stage

Universe: All respondents