



Research Institute
Social Cohesion

RDC

SOEP

SCP Documentation

German Social Cohesion Panel

SCP 2021/22 W1 Codebook PPATH: Person-Related Meta-Dataset (German)



German Social Cohesion Panel

Established in 2021, the German Social Cohesion Panel (SCP) is a wide-ranging representative longitudinal study of private households in Germany, carried out in collaboration of the Research Institute Social Cohesion (RISC) and the German Socio-Economic Panel (SOEP).

The aim of the SCP Documentation is to thoroughly document the survey's data collection and data processing.

Recommended Citation

Groh-Samberg, O., Axenfeld, J. B., Gerlitz, J.-Y., Cornesse, C., Kroh, M., Lengfeld, H., Liebig, S., Minkus, L., Reinecke, J., Richter, D., Teichler, N., Traummüller, R., & Zinn, S. (2024). SCP 2021/22 W1 - Codebook PPATH: Person-Related Meta-Dataset (German). *German Social Cohesion Panel 2021/22 - Wave 1*. Bremen and Berlin: RDC-RISC/SOCIUM, SOEP/DIW Berlin. doi:10.60532/scp.2021_22.w1.v1

- ▶ **Authors:** Olaf Groh-Samberg, Julian B. Axenfeld, Jean-Yves Gerlitz, Carina Cornesse, Martin Kroh, Holger Lengfeld, Stefan Liebig, Lara Minkus, Jost Reinecke, David Richter, Nils Teichler, Richard Traummüller, Sabine Zinn
- ▶ **Contributors:** Cosima Adams, Anton Bochert, Martin Gerike, Josefine Kuhrmeier, Anna-Tabea Müller, Eric Nissen, Rainer Siegers, Hans Walter Steinhauer, Knut Wenzig, Julia Witton (Project Members), infas (Data Collector)
- ▶ **Publisher:** RDC-RISC
SOCIUM, University of Bremen
P.O. Box 330 440
28334 Bremen
Germany

SOEP
DIW Berlin
German Socio-Economic Panel (SOEP)
Mohrenstr. 58
10117 Berlin
Germany
- ▶ **DOI:** 10.60532/scp.2021_22.w1.v1
- ▶ **Website:** www.fgz-risc-data.de
www.diw.de



The text of this publication is published under the Creative Commons license CC BY-SA 4.0 Attribution-ShareAlike 4.0 International. The exact wording of the license CC BY-SA 4.0 can be found here:

<https://creativecommons.org/licenses/by-sa/4.0/>

SCP Documentation

German Social Cohesion Panel

SCP 2021/22 W1 Codebook PPATH: Person-Related Meta-Dataset (German)

Inhaltsverzeichnis

1	Allgemeine Informationen	2
2	Identifikatoren und Stichprobeninformationen	2
	pid – Unveränderliche Personen-ID	2
	cid – Ursprüngliche Haushalts-ID	3
	psample_scp – Stichprobe	3
	hid21_1 – Aktuelle Haushalts-ID 2021 Teil 1	4
	hid21_2 – Aktuelle Haushalts-ID 2021 Teil 2	4
3	Studienhistorie	5
	eintritt – Jahr des ersten Kontaktes	5
	austritt – Jahr des letzten Kontaktes	5
4	Demographie	6
	sex_scp – Geschlecht	6
	sex_scp_flag – Flag Geschlecht	6
	gebjahr – Geburtsjahr	6
	gebjahr_flag – Flag Geburtsjahr	7
	germborn – In Deutschland geboren	7
	immiyear – Jahr der Einwanderung nach Deutschland	8
	corigin_scp – Herkunftsland	9
	migback_scp – Migrationshintergrund	10
	sampreg_scp21_1 – Akt. Stichprobenregion 2021 Teil 1	10
	sampreg_scp21_2 – Akt. Stichprobenregion 2021 Teil 2	11

1 Allgemeine Informationen

Der PPATH-Datensatz verknüpft zur Unterstützung von Längsschnittanalysen persönliche Informationen aus verschiedenen Erhebungswellen des SCP miteinander. Jede Person, die an mindestens einer Erhebungswelle teilgenommen hat, ist auch in PPATH enthalten. PPATH enthält auch zwei Personen, die die Rekrutierungserhebung angefangen, aber keine Fragen beantwortet haben und deshalb aus den Datensätzen P, PGEN und PHRF ausgeschlossen wurden.

An einigen Stellen in der Dokumentation und in den Daten werden Jahreszahlen z. B. für die Bezeichnung von Variablen und des Fragebogeninstruments verwendet. Diese Jahreszahlen orientieren sich stets am Feldstart der Datenerhebung der entsprechenden Erhebungswelle. Die Daten in diesem Datensatz können aus verschiedenen Gründen fehlende Werte aufweisen. Fehlende Angaben werden mit „[-1] Keine Angabe“ codiert, wenn die befragte Person zu einer Variable keine gültigen Angaben macht. Fehlende Angaben werden mit „[-2] Trifft nicht zu“ codiert, wenn aufgrund des Studiendesigns Informationen fehlen.

2 Identifikatoren und Stichprobeninformationen

pid – Unveränderliche Personen-ID

2110000301	1
2110000302	1
2110000901	1
2110000902	1
2110001001	1
2110001201	1
2110001601	1
2110002001	1
2110002101	1
2110003201	1
2110003701	1
2110003901	1
2110004401	1
2110004402	1
2110004403	1
... (16999 Zeilen unterdrückt)	16999
2113796701	1
2113796702	1
2113796703	1
2113797101	1
2113797102	1
2113797201	1
2113797301	1
2113797601	1
2113797801	1
2113797901	1
2113798501	1
2113798701	1
2113799101	1
2113799102	1

211380001 1

Die PID ist die unveränderliche Kennziffer einer Person, die über die Zeit und über alle Datensätze identisch gehalten wird.

cid - Ursprüngliche Haushalts-ID

21100003	2
21100009	2
21100010	1
21100012	1
21100016	1
21100020	1
21100021	1
21100032	1
21100037	1
21100039	1
21100044	3
21100045	1
21100049	1
21100050	1
21100058	1
... (13025 Zeilen unterdrückt)	16989
21137960	1
21137961	1
21137963	2
21137964	2
21137967	3
21137971	2
21137972	1
21137973	1
21137976	1
21137978	1
21137979	1
21137985	1
21137987	1
21137991	2
21138000	1

Diese Kennziffer gruppiert Individuen in ihre Ursprungshaushalte zu Beginn des Panels. Das bedeutet, dass die CID einer Person zeitkonstant gehalten wird und sie immer mit dem Haushalt verbunden wird, zu dem sie initial gehört hat, selbst wenn sie seitdem den Haushalt gewechselt hat.

psample_scp - Stichprobe

1 [1] A 2021 Ausgangstichprobe 17029

Unveränderliche Information über die Stichprobe

hid21_1 - Aktuelle Haushalts-ID 2021 Teil 1

21100003	2
21100009	2
21100010	1
21100012	1
21100016	1
21100020	1
21100021	1
21100032	1
21100037	1
21100039	1
21100044	3
21100045	1
21100049	1
21100050	1
21100058	1
... (13025 Zeilen unterdrückt)	16989
21137960	1
21137961	1
21137963	2
21137964	2
21137967	3
21137971	2
21137972	1
21137973	1
21137976	1
21137978	1
21137979	1
21137985	1
21137987	1
21137991	2
21138000	1

Diese Kennziffer gruppiert Individuen in ihre zugehörigen Haushalte zum Zeitpunkt der aktuellen Erhebungswelle. Das bedeutet, dass die HID einer Person sich über die Zeit verändern kann, zum Beispiel wenn ein erwachsenes Kind aus dem elterlichen Haushalt auszieht und einen eigenen Haushalt eröffnet.

hid21_2 - Aktuelle Haushalts-ID 2021 Teil 2

21100003	2
21100009	2
21100010	1

21100012	1
21100016	1
21100020	1
21100021	1
21100032	1
21100037	1
21100039	1
21100044	3
21100045	1
21100049	1
21100050	1
21100058	1
... (13025 Zeilen unterdrückt)	16989
21137960	1
21137961	1
21137963	2
21137964	2
21137967	3
21137971	2
21137972	1
21137973	1
21137976	1
21137978	1
21137979	1
21137985	1
21137987	1
21137991	2
21138000	1

Diese Kennziffer gruppiert Individuen in ihre zugehörigen Haushalte zum Zeitpunkt der aktuellen Erhebungswelle. Das bedeutet, dass die HID einer Person sich über die Zeit verändern kann, zum Beispiel wenn ein erwachsenes Kind aus dem elterlichen Haushalt auszieht und einen eigenen Haushalt eröffnet.

3 Studienhistorie

eintritt – Jahr des ersten Kontaktes

2021	16666
2022	363

austritt – Jahr des letzten Kontaktes

2021	5218
2022	11811

4 Demographie

sex_scp – Geschlecht

? Welches Geschlecht haben Sie? *(from: scp/2021_2/Q50:pgender[])*

1	[1] Männlich	8281
2	[2] Weiblich	8518
3	[3] Divers	40
-1	[-1] Keine Angabe	190
-2	[-2] Trifft nicht zu	0
-3	[-3] Nicht valide	0
-4	[-4] Unzulässige Mehrfachantwort	0
-5	[-5] In Fragebogenversion nicht enthalten	0
-6	[-6] Fragebogenversion mit geänderter Filterführung	0
-7	[-7] Nur in weniger eingeschränkter Edition verfügbar	0
-8	[-8] Frage in diesem Jahr nicht Teil des Frageprogramms	0
-9	[-9] Möchte ich nicht beantworten [nur CAWI]	0
-10	[-10] Abbruch [nur CAWI]	0

Fehlende Werte werden mit validen Werten aus der nachfolgenden Befragungswelle ersetzt.

sex_scp_flag – Flag Geschlecht

0	[0] Nicht geändert seit erster Inform.	16904
1	[1] Geändert seit erster Inform.	125

Diese Variable zeigt an, ob sich ein valider Wert in PGENDER mindestens einmal in einer folgenden Befragungswelle im Vergleich zur ersten validen Angabe verändert hat. Eine Veränderung von einem validen Wert zu einem fehlenden Wert oder umgekehrt wird nicht angezeigt.

gebjahr – Geburtsjahr

? Wann sind Sie geboren? // Jahr *(from: scp/2021_1/Q6:pbirthyear[])*

? In welchem Jahr sind Sie geboren? // Geburtsjahr [...] *(from: scp/2021_2/Q49:pbirthyear[])*

1921	[1921] 1921 und früher	2
1922	[1922] 1922	1
1923	[1923] 1923	3
1924	[1924] 1924	1
1925	[1925] 1925	5
1926	[1926] 1926	10
1927	[1927] 1927	7
1928	[1928] 1928	12
1929	[1929] 1929	13
1930	[1930] 1930	20
1931	[1931] 1931	29

1932	[1932] 1932	26
1933	[1933] 1933	36
1934	[1934] 1934	59
1935	[1935] 1935	81
...	(65 Zeilen unterdrückt)	15850
2001	[2001] 2001	243
2002	[2002] 2002	231
2003	[2003] 2003	160
2004	[2004] 2004	2
2005	[2005] 2005	2
-1	[-1] Keine Angabe	236
-2	[-2] Trifft nicht zu	0
-3	[-3] Nicht valide	0
-4	[-4] Unzulässige Mehrfachantwort	0
-5	[-5] In Fragebogenversion nicht enthalten	0
-6	[-6] Fragebogenversion mit geänderter Filterführung	0
-7	[-7] Nur in weniger eingeschränkter Edition verfügbar	0
-8	[-8] Frage in diesem Jahr nicht Teil des Frageprogramms	0
-9	[-9] Möchte ich nicht beantworten [nur CAWI]	0
-10	[-10] Abbruch [nur CAWI]	0

Geburtsjahre 1921 und früher wurden aus Datenschutzgründen zu einer Kategorie zusammengefasst. Fehlende Werte werden mit validen Werten aus der nachfolgenden Befragungswelle ersetzt.

gebjahr_flag – Flag Geburtsjahr

0	[0] Nicht geändert seit erster Inform.	16837
1	[1] Geändert seit erster Inform.	192

Diese Variable zeigt an, ob sich ein valider Wert in PBIRTHYEAR mindestens einmal in einer folgenden Befragungswelle im Vergleich zur ersten validen Angabe verändert hat. Eine Veränderung von einem validen Wert zu einem fehlenden Wert oder umgekehrt wird nicht angezeigt.

germborn – In Deutschland geboren

? In welchem Land sind Sie geboren? (*from: scp/2021_1/Q15a:porigin[]*)

? Seit wann leben Sie hauptsächlich in Deutschland? Seit dem Jahr (*from: scp/2021_1/Q15c:pgermstay[]*)

1	[1] In Deutschland geboren oder vor 1949 eingewandert	15289
2	[2] Nach 1948 eingewandert	1481
-1	[-1] Keine Angabe	259
-2	[-2] Trifft nicht zu	0
-3	[-3] Nicht valide	0
-4	[-4] Unzulässige Mehrfachantwort	0
-5	[-5] In Fragebogenversion nicht enthalten	0

-6	[-6] Fragebogenversion mit geänderter Filterführung	0
-7	[-7] Nur in weniger eingeschränkter Edition verfügbar	0
-8	[-8] Frage in diesem Jahr nicht Teil des Frageprogramms	0
-9	[-9] Möchte ich nicht beantworten [nur CAWI]	0
-10	[-10] Abbruch [nur CAWI]	0

Die Informationen dieser Variable stammen aus PORIGIN. Befragte, die angegeben haben, dass sie vor 1949 nach Deutschland eingewandert sind (PGERMSTAY), bekommen den Wert '[1] In Deutschland geboren oder vor 1949 eingewandert' zugewiesen. Für Befragte mit dem Wert '[-1] Keine Angabe' konnten keine validen Informationen generiert werden.

immiyear – Jahr der Einwanderung nach Deutschland

? In welchem Land sind Sie geboren? (from: scp/2021_1/Q15a:porigin[])

? Seit wann leben Sie hauptsächlich in Deutschland? Seit dem Jahr (from: scp/2021_1/Q15c:pgermstay[])

1921	[1921] 1921 und früher	0
1922	[1922] 1922	0
1923	[1923] 1923	0
1924	[1924] 1924	0
1925	[1925] 1925	0
1926	[1926] 1926	0
1927	[1927] 1927	0
1928	[1928] 1928	0
1929	[1929] 1929	0
1930	[1930] 1930	0
1931	[1931] 1931	0
1932	[1932] 1932	0
1933	[1933] 1933	0
1934	[1934] 1934	0
1935	[1935] 1935	0
...	(82 Zeilen unterdrückt)	1316
2018	[2018] 2018	54
2019	[2019] 2019	58
2020	[2020] 2020	41
2021	[2021] 2021	12
2022	[2022] 2022	0
-1	[-1] Keine Angabe	259
-2	[-2] Trifft nicht zu	15289
-3	[-3] Nicht valide	0
-4	[-4] Unzulässige Mehrfachantwort	0
-5	[-5] In Fragebogenversion nicht enthalten	0
-6	[-6] Fragebogenversion mit geänderter Filterführung	0
-7	[-7] Nur in weniger eingeschränkter Edition verfügbar	0
-8	[-8] Frage in diesem Jahr nicht Teil des Frageprogramms	0
-9	[-9] Möchte ich nicht beantworten [nur CAWI]	0
-10	[-10] Abbruch [nur CAWI]	0

Die Informationen dieser Variable stammen aus PGERMSTAY. Befragte, die '[1] In Deutschland geboren oder vor 1949 eingewandert' (GERMBORN) sind, bekommen den Wert '[-2] Trifft nicht zu' zugewiesen. Für Befragte mit dem Wert '[-1] Keine Angabe' konnten keine validen Informationen generiert werden.

corigin_scp – Herkunftsland

- ? In welchem Land sind Sie geboren? (from: scp/2021_1/Q15a:porigin[])
- ? In welchem anderen Land sind Sie geboren? (from: scp/2021_1/Q15b:pothcount[])
- ? Seit wann leben Sie hauptsächlich in Deutschland? Seit dem Jahr (from: scp/2021_1/Q15c:pgermstay[])
- ? Welche andere Staatsbürgerschaft haben Sie? (from: scp/2021_1/Q16b:pnaoth[])

4	[4] Afghanistan	22
8	[8] Albanien	11
10	[10] Antarktis	0
12	[12] Algerien	8
16	[16] Amerikanisch-Samoa	0
20	[20] Andorra	0
24	[24] Angola	0
28	[28] Antigua und Barbuda	0
31	[31] Aserbaidshan	2
32	[32] Argentinien	4
36	[36] Australien	4
40	[40] Österreich	38
44	[44] Bahamas	0
48	[48] Bahrain	0
50	[50] Bangladesch	0
...	(230 Zeilen unterdrückt)	16641
876	[876] Wallis und Futuna	0
882	[882] Samoa	0
887	[887] Jemen	0
894	[894] Sambia	0
900	[900] Kosovo	13
-1	[-1] Keine Angabe	286
-2	[-2] Trifft nicht zu	0
-3	[-3] Nicht valide	0
-4	[-4] Unzulässige Mehrfachantwort	0
-5	[-5] In Fragebogenversion nicht enthalten	0
-6	[-6] Fragebogenversion mit geänderter Filterführung	0
-7	[-7] Nur in weniger eingeschränkter Edition verfügbar	0
-8	[-8] Frage in diesem Jahr nicht Teil des Frageprogramms	0
-9	[-9] Möchte ich nicht beantworten [nur CAWI]	0
-10	[-10] Abbruch [nur CAWI]	0

Die Informationen dieser Variable stammen aus POTHCOUNT und PORIGIN. Befragte, die '[1] In Deutschland geboren oder vor 1949 eingewandert' (GERMBORN) sind, bekommen den Wert '[276] Deutschland' zugewiesen. Wenn es keine Informationen über das Geburtsland gibt und GERMBORN '[2] Nach 1948 eingewandert' ist, werden Informationen über

die Staatsbürgerschaft einer Person (PNATOTH) genutzt, um CORIGIN_SCP zu generieren. Für Befragte mit dem Wert '[-1] Keine Angabe' konnten keine validen Informationen generiert werden.

migback_scp – Migrationshintergrund

- ? In welchem Land sind Sie geboren? (from: scp/2021_1/Q15a:porigin[])
 ? Welche Staatsbürgerschaft haben Sie? // Eine andere (from: scp/2021_1/Q16a:pnation02[])
 ? Nun möchten wir etwas über die Herkunft Ihrer Eltern und Großeltern erfahren. // Sind Ihre Eltern beide in Deutschland geboren? (from: scp/2021_1/Q17:pmigback01[])
 ? Nun möchten wir etwas über die Herkunft Ihrer Eltern und Großeltern erfahren. // Sind Ihre Großeltern alle in Deutschland geboren? (from: scp/2021_1/Q17:pmigback02[])

1	[1] Kein Migrationshintergrund	12016
2	[2] Direkter Migrationshintergrund	1780
3	[3] Indirekter Migrationshintergrund	2292
4	[4] Migrationshintergrund, Typ unklar	6
-1	[-1] Keine Angabe	935
-2	[-2] Trifft nicht zu	0
-3	[-3] Nicht valide	0
-4	[-4] Unzulässige Mehrfachantwort	0
-5	[-5] In Fragebogenversion nicht enthalten	0
-6	[-6] Fragebogenversion mit geänderter Filterführung	0
-7	[-7] Nur in weniger eingeschränkter Edition verfügbar	0
-8	[-8] Frage in diesem Jahr nicht Teil des Frageprogramms	0
-9	[-9] Möchte ich nicht beantworten [nur CAWI]	0
-10	[-10] Abbruch [nur CAWI]	0

Diese Variable gibt Auskunft über den Migrationshintergrund. Sie berücksichtigt das Geburtsland und die Staatsangehörigkeit der Person sowie ihre Migrationsgeschichte. Direkter/indirekter Migrationshintergrund gibt an, ob die befragte Person selbst zugewandert ist oder eine nicht-deutsche Staatsbürgerschaft hat (direkter Migrationshintergrund) oder ob sie einen Migrationshintergrund hat, aber in Deutschland geboren ist (indirekter Migrationshintergrund). Befragte mit dem Code 4 haben ebenfalls einen Migrationshintergrund, aber wir wissen nicht, ob sie selbst eingewandert sind, da Angaben zu ihrem Geburtsland fehlen. Der Code '[1] Kein Migrationshintergrund' umfasst Befragte mit deutscher Staatsbürgerschaft, die in Deutschland geboren sind und deren Eltern und Großeltern ebenfalls in Deutschland geboren wurden. Für Befragte mit dem Wert '[-1] Keine Angabe' konnten keine validen Informationen generiert werden, meist aufgrund fehlender Angaben zum Geburtsland der Eltern oder Großeltern.

sampreg_scp21_1 – Akt. Stichprobenregion 2021 Teil 1

1	[1] Westdeutschland	11063
2	[2] Ostdeutschland	5212
3	[3] Berlin	754

Diese Variable gibt an, in welcher Stichprobenregion sich die Haushalte in der jeweiligen Erhebungswelle befanden: Westdeutschland, Ostdeutschland oder Berlin.

sampreg_scp21_2 - Akt. Stichprobenregion 2021 Teil 2

1	[1] Westdeutschland	11063
2	[2] Ostdeutschland	5212
3	[3] Berlin	754

Diese Variable gibt an, in welcher Stichprobenregion sich die Haushalte in der jeweiligen Erhebungswelle befanden: Westdeutschland, Ostdeutschland oder Berlin.