**This file describes and provides codebook for the datafiles provided in the source\_ancestry\_survey\_files folder in addition to the Ancestry Survey file (Ancestry\_survey\_set\_original.dta).**

**Ancestry\_survey\_set\_original.dta** (Becker, S., I. Grosfeld, P. Grosjean, N. Voigtländer, and E. Zhuravskaya (2016). Ancestry Survey. http://www.parisschoolofeconomics.com/zhuravskaya-ekaterina/erc-economics-of-prejudice/ (accessed December 11, 2019)) - original Ancestry survey data with located places of ancestors (collected by authors).

Source: survey data collected by authors.

The questionnaire is found here: <http://www.parisschoolofeconomics.com/zhuravskaya-ekaterina/erc-economics-of-prejudice/> (accessed December 11, 2019)

The codebook is found in the source\_ancestry\_files folder: codebook\_Ancestry.pdf.

**gminy\_code\_coordinates.dta** - coordinates of gminy centeroid.

Source: <http://www.diva-gis.org/gdata> (accessed December 11, 2019)

Codebook :

Variable: latitude\_gminy\_cent - Label : Latitude gminy centroid : latitude of the centroid of the gminy

Variable: longitude\_gminy\_cent – Label : Longitude gminy centroid : longitude of the centroid of the gminy

Variable: gminy\_code – Label : Gminy code : Numeric code of the gminy

**gminy\_code\_map\_id\_dr\_1933.dta** – key to merging ids from 1933 WR map to gminy codes.

Source: original file generated by authors from digital maps provided in GIS MAPS folder (see this folder for the information on the sources).

Codebook:

Variable: gminy\_code – Label: Gminy code: Numeric code of the gminy

Variable: id\_dr\_1933 – Label: ID in 1933 WR: ID in 1933 Weimar Republic Census and map

Variable: SPR Second Polish Republic: 1 if gminy in the Second Polish Republic

Variable: gdansk Gdansk: 1 if gminy in Gdansk

**gminy\_code\_1900\_id\_for\_kresy\_survey\_locations.dta** – key to merging gminy to 1900 Prussian data.

Source: Kaiserliches Statistisches Amt (Editor): Statistik des Deutschen Reichs, Band 150: Die Volkszählung im Deutschen Reich am 1. Dezember 1900. Berlin; Puttkammer & Mühlbrecht 1903 and see GIS MAPS folder for maps.

Codebook:

Variable: gminy\_code – Label: Gminy code: Numeric code of the gminy

Variable: name Name in 1900 German Census – Name of district in 1900 German Census

Variable: id id in 1900 German Census – District id in 1900 German Census

**WeimarR\_1925\_1933\_Census.dta** – Weimar Republic Census (at KRNR level for those KRNR which were the places of origin for some of ancestors of our respondents and for those KRNR which coincided with WT).

Sources:

For 1925:

Statistisches Reichsamt (Hrsg.): Statistik des Deutschen Reichs, Band 401: Volks-, Berufs- und Betriebszählung vom 16. Juni 1925. Volkszählung. Die Bevölkerung des Deutschen Reichs nach den Ergebnissen  
der Volkszählung 1925. Berlin: Verlag von Reimar Hobbing 1928.

For 1933:

Statistisches Reichsamt (Hrsg.): Statistik des Deutschen Reichs, Band 451: Volks-, Berufs- und Betriebszählung vom 16. Juni 1933.Volkszählung. Die Bevölkerung des Deutschen Reichs nach den Ergebnissen der Volkszählung 1933, Berlin: Verlag für Sozialpolitik, Wirtschaft und Statistik, Paul Schmitt 1936.

Codebook:

Variable: id – Label: District id for 1925 1933 Germany: District id number for 1925 1933 Germany

Variable: TotalLabor1925 – Label: Labor Force in 1925: Labor Force in 1925

Variable: AgricLabor1925 – Label: Agricultural Employment in 1925 : Agricultural Employment in 1925

Variable: IndLabor1925 – Label: Industrial and Trade Employment in 1925 : Industrial and Trade Employment in 1925

Variable: pop1925 – Label: Population in 1925: Total population in 1925

Variable: protestant1925 – Label: Protestants in 1925: Number of Protestants in 1925

Variable: catholic1925 – Label: Catholics in 1925: Number of Catholics in 1925

Variable: jews1925 – Label: Jews in 1925: Number of Jews in 1925

Variable: pop1933 – Label: Population in 1933: Total Population in 1933

Variable: TotalLabor1933 – Label: Total Labor Force in 1933: Total Labor Force in 1933

Variable: Unemployed1933 – Label: Total Unemployed in 1933: Total Unemployed in 1933

Variable: AgricLabor1933 – Label: Agricultural Employment in 1933: Agricultural Employment in 1933

Variable: IndLabor1933 – Label: Industrial Employment in 1933: Industrial Employment in 1933

**germany1900\_religion.dta** – Prussian Census 1900 (at kreis level for those kreis that coincided with WT).

Source: Kaiserliches Statistisches Amt: Statistik des Deutschen Reichs, Band 150: Die Volkszählung im Deutschen Reich am 1. Dezember 1900. Berlin; Puttkammer & Mühlbrecht 1903

Codebook:

Variable: kreis – Label: Name of district in 1900 Census: Name of district in 1900 Census

Variable: pop – Label: Population in 1900 : Population in 1900

Variable: relprot – Label: Number of Protestants in 1900: Number of Protestants in 1900

Variable: relcath – Label: Number of Catholics in 1900: Number of Catholics in 1900

Variable: relotherchristian – Label: Number of Other Christian in 1900 : Number of Other Christian in 1900

Variable: reljew – Label: Number of Jews in 1900 : Number of Jews in 1900

Variable: relother – Label: Number of Other religion in 1900: Number of Other religion in 1900

Variable: id – Label: id in 1900 German Census: Numeric id code for 1900 German Census

**germany1900\_language.dta**– Prussian Census 1900 (at kreis level for those kreis that coincided with WT).

Source: Kaiserliches Statistisches Amt: Statistik des Deutschen Reichs, Band 150: Die Volkszählung im Deutschen Reich am 1. Dezember 1900. Berlin; Puttkammer & Mühlbrecht 1903

Codebook:

Variable: id - Label (and meaning): id in 1900 German Census

Variable: lan\_german1900 - Label (and meaning): German lang in 1900, % [lang stands for langague]

Variable: lan\_polish1900 - Label (and meaning): Polish lang in 1900, %

Variable: lan\_masurish1900 - Label (and meaning): Masurish lang in 1900, %

Variable: lan\_kassubian1900 - Label (and meaning): Kassubian lang in 1900, %

Variable: lan\_mahrish1900 - Label (and meaning): Mahrish lang in 1900, %

Variable: lan\_czech1900 - Label (and meaning): Czech lang in 1900, %

Variable: lan\_russian1900 - Label (and meaning): Russian lang in 1900, %

Variable: lan\_lithuanian1900 - Label (and meaning): Lithuanian lang in 1900, %

Variable: lan\_not\_german1900 - Label (and meaning): did not speak German in 1900, %

**1931\_Polish\_census.dta** - Polish census 1931 (at powiat level for all powiat in SPR).

Source: *Drugi Powszechny Spis Ludnosci z dn. 9.XII.1931 R.* 1938. Warszawa: Glowny Urzad Statystyczny

Codebook:

Variable: code\_map\_1931pow - Label (and meaning): ID code for 1931 Census

Variable: map\_powiaty\_code\_1931- Label (and meaning): ID code for 1931 Census

Variable: r\_rcatholic\_1931pow- Label (and meaning): Share Roman Catholics in powiat 1931

Variable: r\_gcatholic\_1931pow - Label (and meaning): Share Greek Catholics in powiat 1931

Variable: r\_orthodox\_1931pow- Label (and meaning): Share Orthodox in powiat 1931

Variable: r\_protestant\_lutherian\_1931pow- Label (and meaning): Share Lutherian in powiat 1931

Variable: r\_protestant\_reformed\_1931pow- Label (and meaning): Share Reformed Protestants in powiat 1931

Variable: r\_jewish\_1931pow- Label (and meaning): Share Jewish in powiat 1931

Variable: r\_rcatholic\_1931rural- Label (and meaning): Share Roman Catholics in powiat 1931 - rural areas

Variable: r\_gcatholic\_1931rural- Label (and meaning): Share Greek Catholics in powiat 1931 - rural areas

Variable: r\_orthodox\_1931rural- Label (and meaning): Share Orthodox in powiat 1931 - rural areas

Variable: r\_protestant\_lutherian\_1931rural- Label (and meaning): Share Lutherian in powiat 1931 - rural areas

Variable: r\_protestant\_reformed\_1931rural- Label (and meaning): Share Reformed Protestants in powiat 1931 - rural areas

Variable: r\_jewish\_1931rural- Label (and meaning): Share Jewish in powiat 1931 - rural areas

Variable: r\_rcatholic\_1931urban- Label (and meaning): Share Roman Catholics in powiat 1931 - urban areas

Variable: r\_gcatholic\_1931urban- Label (and meaning): Share Greek Catholics in powiat 1931 - urban areas

Variable: r\_orthodox\_1931urban - Label (and meaning): Share Orthodox in powiat 1931 - urban areas

Variable: r\_protestant\_lutherian\_1931urban- Label (and meaning): Share Lutherian in powiat 1931 - urban areas

Variable: r\_protestant\_reformed\_1931urban- Label (and meaning): Share Reformed Protestants in powiat 1931 - urban areas

Variable: r\_jewish\_1931urban - Label (and meaning): Share Jewish in powiat 1931 - urban areas

Variable: r\_protestant\_1931pow - Label (and meaning): Share All Protestants in powiat 1931

Variable: r\_protestant\_1931rural- Label (and meaning): Share All Protestants in powiat 1931- rural areas

Variable: r\_protestant\_1931urban- Label (and meaning): Share All Protestants in powiat 1931- urban areas

Variable: Lang\_polish\_1931pow - Label (and meaning): Share speaks Polish in powiat 1931

Variable: Lang\_ukranian\_1931pow- Label (and meaning): Share speaks Ukrainian in powiat 1931

Variable: Lang\_ruthenian\_1931pow- Label (and meaning): Share speaks Ruthenian in powiat 1931

Variable: Lang\_bel\_1931pow- Label (and meaning): Share speaks Belarusian in powiat 1931

Variable: Lang\_russian\_1931pow- Label (and meaning): Share speaks Russian in powiat 1931

Variable: Lang\_lit\_1931pow- Label (and meaning): Share speaks Lithuanian in powiat 1931

Variable: Lang\_german\_1931pow - Label (and meaning): Share speaks German powiat 1931

Variable: Lang\_polish\_1931rural - Label (and meaning): Share speaks Polish in powiat 1931 - rural areas

Variable: Lang\_ukranian\_1931rural- Label (and meaning): Share speaks Ukrainian in powiat 1931 - rural areas

Variable: Lang\_ruthenian\_1931rural- Label (and meaning): Share speaks Ruthenian in powiat 1931 - rural areas

Variable: Lang\_bel\_1931rural- Label (and meaning): Share speaks Belarusian in powiat 1931 - rural areas

Variable: Lang\_russian\_1931rural- Label (and meaning): Share speaks Russian in powiat 1931 - rural areas

Variable: Lang\_lit\_1931rural - Label (and meaning): Share speaks Lithuanian in powiat 1931 - rural areas

Variable: Lang\_german\_1931rural- Label (and meaning): Share speaks German powiat 1931 - rural areas

Variable: Lang\_polish\_1931urban- Label (and meaning): Share speaks Polish in powiat 1931- urban areas

Variable: Lang\_ukranian\_1931urban - Label (and meaning): Share speaks Ukrainian in powiat 1931 - urban areas

Variable: Lang\_ruthenian\_1931urban- Label (and meaning): Share speaks Ruthenian in powiat 1931 - urban areas

Variable: Lang\_bel\_1931urban- Label (and meaning): Share speaks Belarusian in powiat 1931 - urban areas

Variable: Lang\_russian\_1931urban - Label (and meaning): Share speaks Russian in powiat 1931 - urban areas

Variable: Lang\_lit\_1931urban - Label (and meaning): Share speaks Lithuanian in powiat 1931 - urban areas

Variable: Lang\_german\_1931urban- Label (and meaning): Share speaks German powiat 1931 - urban areas

Variable: Lang\_jews\_1931\_pow - Label (and meaning): Share speaks Yiddish powiat 1931- Label (and meaning):

Variable: Lang\_jews\_1931\_rural- Label (and meaning): Share speaks Yiddish powiat 1931 - rural areas

Variable: Lang\_jews\_1931\_urban- Label (and meaning): Share speaks Yiddish powiat 1931 - urban areas

Variable: powiaty\_code\_1931- Label (and meaning): Powiaty code 1931

Variable: urbanization\_1931- Label (and meaning): Urbanisation powiat 1931

Variable: literate\_1931pow- Label (and meaning): Share literate powiat 1931

Variable: readonly\_1931pow - Label (and meaning): Share can read only powiat 1931

Variable: notliterate\_1931pow- Label (and meaning): Share illiterate powiat 1931

Variable: literate\_1931rural- Label (and meaning): Share literate powiat 1931 - rural areas

Variable: readonly\_1931rural- Label (and meaning): Share can read only powiat 1931 - rural areas

Variable: notliterate\_1931rural - Label (and meaning): Share literate powiat 1931 - rural areas

Variable: literate\_1931urban- Label (and meaning): Share literate powiat 1931 - urban areas

Variable: readonly\_1931urban- Label (and meaning): Share can read only powiat 1931 - urban areas

Variable: notliterate\_1931urban - Label (and meaning): Share literate powiat 1931 - urban areas

Variable: in\_kresy- Label: In Kresy : Powiat in Kresy. 0 : No ; 1 : Yes.

**1921\_census\_pow\_for\_kresy\_survey.dta** – Polish Census 1921(at powiat level for all powiat in SPR).

Source: *Pierwszy Powszechny Spis Rzeczypospolitej Polskiej z dnia 30 wrzesnia 1921 roku*. 1926, 1927, 1928. Warszawa: Glowny Urzad Statystyczny

Codebook :

Variable: powiaty\_code1931- Label (and meaning) Powiaty code 1931

Variable: total\_pop\_u- Label (and meaning) Total population - urban areas

Variable: total\_pop\_t- Label (and meaning) Total population

Variable: total\_pop\_r- Label (and meaning) Total population - rural areas

Variable: literate\_1921\_t- Label (and meaning) Share literate in total pop, 1921

Variable: literate\_1921\_u- Label (and meaning) Share literate in total pop, 1921 - urban area

Variable: literate\_1921\_r- Label (and meaning) Share literate in total pop, 1921 - rural area

Variable: lit\_rcath\_1921\_t- Label (and meaning) Share literate among Roman Catholics, 1921

Variable: lit\_rcath\_1921\_u- Label (and meaning) Share literate among Roman Catholics, 1921 - urban areas

Variable: lit\_rcath\_1921\_r- Label (and meaning) Share literate among Roman Catholics, 1921 - rural areas

Variable: r\_catholic\_1921\_t- Label (and meaning) Share roman catholics in tot pop, 1921

Variable: r\_catholic\_1921\_u- Label (and meaning) Share roman catholics in tot pop, 1921- urban areas

Variable: r\_catholic\_1921\_r- Label (and meaning) Share roman catholics in tot pop, 1921 - rural areas

**1897\_RussCens\_for\_kresy\_survey.dta**– Russian Empire Census 1897 (at province level for those provinces that were the places of origin of few ancestors of our respondents who lived in 1939 to the east of SPR).

Source: Troynitsky, N. (1899). The First General Census of Russian Empire in 89 Volumes. The Central Statistical Bureau of the Ministry of Internal Affairs of the Russian Empire, St. Petersburg,1899-1905.

Codebook:

Variable: prov- Label (and meaning) Province

Variable: literate\_r- Label (and meaning) Share literate, 1897 - rural areas

Variable: not\_literate\_r- Label (and meaning) Share not literate, 1897 - rural areas

Variable: higher\_educate\_r- Label (and meaning) Share higher education, 1897 - rural areas

Variable: literate\_t- Label (and meaning) Share literate, 1897

Variable: not\_literate\_t- Label (and meaning) Share not literate, 1897

Variable: higher\_educate\_t- Label (and meaning) Share higher education, 1897

Variable: literate\_u- Label (and meaning) Share literate, 1897 - urban areas

Variable: not\_literate\_u- Label (and meaning) Share not literate, 1897 - urban areas

Variable: higher\_educate\_u- Label (and meaning) Share higher education, 1897 - urban areas

Variable: greek\_catholic\_r- Label (and meaning) Share Greek Catholics, 1897 - rural areas

Variable: jewish\_r- Label (and meaning) Share Jewish, 1897 - rural areas

Variable: orthodox\_r- Label (and meaning) Share Orthodox, 1897 - rural areas

Variable: protestant\_r- Label (and meaning) Share Protestants, 1897 - rural areas

Variable: roman\_catholic\_r- Label (and meaning) Share Roman Catholics, 1897 - rural areas

Variable: greek\_catholic\_t- Label (and meaning) Share Greek Catholics, 1897

Variable: jewish\_t- Label (and meaning) Share Jewish, 1897

Variable: orthodox\_t- Label (and meaning) Share Orthodox, 1897

Variable: protestant\_t- Label (and meaning) Share Protestants, 1897

Variable: roman\_catholic\_t- Label (and meaning) Share Roman Catholics, 1897

Variable: greek\_catholic\_u- Label (and meaning) Share Greek Catholics, 1897 - urban areas

Variable: jewish\_u- Label (and meaning) Share Jewish, 1897 - urban areas

Variable: orthodox\_u- Label (and meaning) Share Orthodox, 1897 - urban areas

Variable: protestant\_u- Label (and meaning) Share Protestants, 1897 - urban areas

Variable: roman\_catholic\_u- Label (and meaning) Share Roman Catholics, 1897 - urban areas

Variable: polish\_r- Label (and meaning) Share Polish, 1897 - rural areas

Variable: ukrainian\_r- Label (and meaning) Share Ukrainian, 1897 - rural areas

Variable: belorussian\_r- Label (and meaning) Share Belorusian, 1897 - rural areas

Variable: russian\_r- Label (and meaning) Share Russian, 1897 - rural areas

Variable: lithuanian\_r- Label (and meaning) Share Lithuanian, 1897 - rural areas

Variable: german\_r- Label (and meaning) Share German, 1897 - rural areas

Variable: jew\_r- Label (and meaning) Share Jewish, 1897 - rural areas

Variable: polish\_t- Label (and meaning) Share Polish, 1897

Variable: ukrainian\_t- Label (and meaning) Share Ukrainian, 1897

Variable: belorussian\_t- Label (and meaning) Share Belorusian, 1897

Variable: russian\_t- Label (and meaning) Share Russian, 1897

Variable: lithuanian\_t- Label (and meaning) Share Lithuanian, 1897

Variable: german\_t - Label (and meaning) Share German, 1897

Variable: jew\_t- Label (and meaning) Share Jewish, 1897

Variable: polish\_u- Label (and meaning) Share Polish, 1897 - urban areas

Variable: ukrainian\_u- Label (and meaning) Share Ukrainian, 1897 - urban areas

Variable: belorussian\_u- Label (and meaning) Share Belorusian, 1897 - urban areas

Variable: russian\_u- Label (and meaning) Share Russian, 1897 - urban areas

Variable: lithuanian\_u- Label (and meaning) Share Lithuanian, 1897 - urban areas

Variable: german\_u - Label (and meaning) Share German, 1897 – urban areas

Variable: jew\_u- Label (and meaning) Share Jewish, 1897 - urban areas

Var : rn – Label : Respondent number : Respondent number for merge with Ancestry survey

Variable: Rus\_urbanization- Label (and meaning) Urbanization, 1897

**AH\_1910\_religions\_languages.dta** – Census of Austro-Hungarian empire, religion and languages, 1910 (at local level for 4 places of origin of ancestors of our respondents who lived in 1939 in the territory of Austro-Hungarian empire).

Source: Die Ergebnisse der Volkszählung vom 31. Dezember 1910 in den im Reichsrate vertretenen Königreichen und Ländern. (=Neue Folge Österreichische Statistik); Publisher: Wien Karl Gerolds Sohn (1912); Volume 1, issue 1, Table III: pages 80-93; and Volume 1, issue 2, Table VIII, pages 66-71.

Codebook:

Variable: row\_austrian\_data\_file Respondent number: Respondent number for merge with ancestor in Ancestry survey

Variable: AH\_rel\_jewish\_1910- Label (and meaning): Share Jewish, 1910

Variable: AH\_rel\_rcatholic\_1910- Label (and meaning): Share Roman Catholics, 1910

Variable: AH\_rel\_protestant\_1910- Label (and meaning): Share Protestants, 1910

Variable: AH\_rel\_gcatholic\_1910- Label (and meaning): Share Greek Catholics, 1910

Variable: AH\_rel\_orthodox\_1910- Label (and meaning): Share Orthodox, 1910

Variable: AH\_Lang\_polish\_1910- Label (and meaning): Share speaks Polish, 1910

Variable: AH\_Lang\_ukranian\_1910- Label (and meaning): Share speaks Ukrainian, 1910

Variable: AH\_Lang\_bel\_1910- Label (and meaning): Share speaks Belarusian, 1910

Variable: AH\_Lang\_russian\_1910- Label (and meaning): Share speaks Russian, 1910

Variable: AH\_Lang\_lit\_1910- Label (and meaning): Share speaks Lithuanian, 1910

Variable: AH\_Lang\_german\_1910- Label (and meaning): Share speaks German, 1910

Variable: AH\_Lang\_jews\_1910- Label (and meaning): Share speaks Yiddish, 1910

**AH\_1910\_literacy.dta**– Census of Austro-Hungarian empire, literacy, 1910 (at local level for 4 places of origin of ancestors of our respondents who lived in 1939 in the territory of Austro-Hungarian empire).

Source: Die Ergebnisse der Volkszählung vom 31. Dezember 1910 in den im Reichsrate vertretenen Königreichen und Ländern. (=Neue Folge Österreichische Statistik); Publisher: Wien Karl Gerolds Sohn (1912); Volume 1, issue 1, Table III: pages 80-93; and Volume 1, issue 2, Table VIII, pages 66-71.

Codebook:

Variable: runningnr- Label (and meaning): Respondent number

Variable: AH\_literate\_1910- Label (and meaning): Share literate, 1910

Variable: AH\_ililiterate\_1910- Label (and meaning): Share illiterate, 1910

Variable: AH\_readonly\_1910- Label (and meaning): Share can read only, 1910

**woj\_1931\_centeroids.dta**– coordinates of the centroid of the 1931 wojewodztwo. From Poland\_1931\_woj.shp in GIS\_MAPS subfolder.

Source: MPIDR [Max Planck Institute for Demographic Research] and CGG [Chair for Geodesy and Geoinformatics, University of Rostock] 2011: MPIDR Population History GIS Collection – Rostock.

Codebook:

Variable: latitude\_cent\_woj\_1931 – Label (and meaning): Latitude wojewodztwo centroid

Variable: longitude\_cent\_woj\_1931 – Label (and meaning): Longitude wojewodztwo centroid

Variable: woj\_code – Label (and meaning): Wojewodztwo code

**WOJ\_1931\_census\_for\_kresy.dta**- Polish census 1931 (at wojewodztwo level for all wojewodztwo in SPR).

Source: *Drugi Powszechny Spis Ludnosci z dn. 9.XII.1931 R.* 1938. Warszawa: Glowny Urzad Statystyczny

Codebook:

Variable: woj\_code – Label : Wojewodztwo code 1931: Wojewodztwo numeric code

Variable: wr\_r\_catholic\_1931pow – Label (and meaning): Share Roman Catholics in wojewodztwo 1931

Variable: wr\_g\_catholic\_1931pow – Label (and meaning): Share Greek Catholics in wojewodztwo 1931

Variable: wr\_orthodox\_1931pow – Label (and meaning): Share Orthodox in wojewodztwo 1931

Variable: wr\_jewish\_1931pow – Label (and meaning): Share Jewish in wojewodztwo 1931

Variable: wr\_r\_catholic\_1931rural – Label (and meaning): Share Roman Catholics in wojewodztwo 1931 - rural areas

Variable: wr\_g\_catholic\_1931rural – Label (and meaning): Share Greek Catholics in wojewodztwo 1931 - rural areas

Variable: wr\_orthodox\_1931rural – Label (and meaning): Share Orthodox in wojewodztwo 1931 - rural areas

Variable: wr\_jewish\_1931rural – Label (and meaning): Share Jewish in wojewodztwo 1931 - rural areas

Variable: wr\_r\_catholic\_1931urban – Label (and meaning): Share Roman Catholics in wojewodztwo 1931 - urban areas

Variable: wr\_g\_catholic\_1931urban – Label (and meaning): Share Greek Catholics in wojewodztwo 1931 - urban areas

Variable: wr\_orthodox\_1931urban – Label (and meaning): Share Orthodox in wojewodztwo 1931 - urban areas

Variable: wr\_jewish\_1931urban – Label (and meaning): Share Jewish in wojewodztwo 1931 - urban areas

Variable: wr\_protestant\_1931pow – Label (and meaning): Share All Protestants in wojewodztwo 1931

Variable: wr\_protestant\_1931rural – Label (and meaning): Share All Protestants in wojewodztwo 1931- rural areas

Variable: wr\_protestant\_1931urban – Label (and meaning): Share All Protestants in wojewodztwo 1931- urban areas

Variable: wLang\_polish\_1931pow – Label (and meaning): Share speaks Polish in wojewodztwo 1931

Variable: wLang\_ukranian\_1931pow – Label (and meaning): Share speaks Ukrainian in wojewodztwo 1931

Variable: wLang\_russian\_1931pow – Label (and meaning): Share speaks Russian in wojewodztwo 1931

Variable: wL\_lithuanian\_1931pow – Label (and meaning): Share speaks Lithuanian in wojewodztwo 1931

Variable: wL\_belorussian\_1931pow – Label (and meaning): Share speaks Belarusian in wojewodztwo 1931

Variable: wL\_belorussian\_1931rural - Labe; (and meaning): Share speaks Belarusian in wojewodztwo 1931- rural areas

Variable: wL\_belorussian\_1931urban - Label (and meaning): Share speaks Belarusian in wojewodztwo 1931- urban areas

Variable: wLang\_german\_1931pow - Label (and meaning): Share speaks German wojewodztwo 1931

Variable: wLang\_polish\_1931rural - Label (and meaning): Share speaks Polish in wojewodztwo 1931 - rural areas

Variable: wLang\_ukranian\_1931rural - Labe; (and meaning): Share speaks Ukrainian in wojewodztwo 1931 - rural areas

Variable: wLang\_russian\_1931rural - Labe; (and meaning): Share speaks Russian in wojewodztwo 1931 - rural areas

Variable: wL\_lithuanian\_1931rural - Label (and meaning): Share speaks Lithuanian in wojewodztwo 1931 - rural areas

Variable: wLang\_german\_1931rural - Labe; (and meaning): Share speaks German wojewodztwo 1931 - rural areas

Variable: wLang\_polish\_1931urban - Label (and meaning): Share speaks Polish in wojewodztwo 1931- urban areas

Variable: wLang\_ukranian\_1931urban - Label (and meaning): Share speaks Ukrainian in wojewodztwo 1931 - urban areas

Variable: wLang\_russian\_1931urban - Label (and meaning): Share speaks Russian in wojewodztwo 1931 - urban areas

Variable: wL\_lithuanian\_1931urban - Label (and meaning): Share speaks Lithuanian in wojewodztwo 1931 - urban areas

Variable: wLang\_german\_1931urban - Label (and meaning): Share speaks German wojewodztwo 1931 - urban areas

Variable: wLang\_jews\_1931pow - Label (and meaning): Share speaks Yiddish wojewodztwo 1931Label (and meaning):

Variable: wLang\_jews\_1931rural - Labe; (and meaning): Share speaks Yiddish wojewodztwo 1931 - rural areas

Variable: wLang\_jews\_1931urban - Label (and meaning): Share speaks Yiddish wojewodztwo 1931 - urban areas

Variable: wliterate\_1931pow – Label (and meaning): Share literate wojewodztwo 1931

Variable: wreadonly\_1931pow - Label (and meaning): Share can read only wojewodztwo 1931

Variable: wnotliterate\_1931pow – Label (and meaning): Share illiterate wojewodztwo 1931

Variable: wliterate\_1931rural - Label (and meaning): Share literate wojewodztwo 1931 - rural areas

Variable: wreadonly\_1931rural - Label (and meaning): Share can read only wojewodztwo 1931 - rural areas

Variable: wnotliterate\_1931rural - Label (and meaning): Share literate wojewodztwo 1931 - rural areas

Variable: wliterate\_1931urban - Label (and meaning): Share literate wojewodztwo 1931 - urban areas

Variable: wreadonly\_1931urban - Label (and meaning): Share can read only wojewodztwo 1931

Variable: wnotliterate\_1931urban - Label (and meaning): Share literate wojewodztwo 1931 - urban areas

**1921\_census\_woj\_for\_kresy\_survey.dta** - Polish census 1921 (at wojewodztwo level for all wojewodztwo in SPR). Source: *Pierwszy Powszechny Spis Rzeczypospolitej Polskiej z dnia 30 wrzesnia 1921 roku*. 1926, 1927, 1928. Warszawa: Glowny Urzad Statystyczny

Codebook:

Variable: woj\_code – Label : Wojewodztwo code 1931: Wojewodztwo numeric code

Variable: literate\_1921\_t - Label (and meaning) Share literate in total pop, 1921

Variable: literate\_1921\_u - Label (and meaning) Share literate in total pop, 1921 - urban area

Variable: literate\_1921\_r - Label (and meaning) Share literate in total pop, 1921 - rural area

Variable: lit\_rcath\_1921\_t - Label (and meaning) Share literate among Roman Catholics, 1921

Variable: lit\_rcath\_1921\_u - Label (and meaning) Share literate among Roman Catholics, 1921 - urban areas

Variable: lit\_rcath\_1921\_r - Label (and meaning) Share literate among Roman Catholics, 1921 - rural areas

Variable: r\_catholic\_1921\_t - Label (and meaning) Share roman catholics in tot pop, 1921

Variable: r\_catholic\_1921\_u - Label (and meaning) Share roman catholics in tot pop, 1921- urban areas

Variable: r\_catholic\_1921\_r - Label (and meaning) Share roman catholics in tot pop, 1921 - rural areas