

Übung 05

Hypothesen testen

European Social Survey

INFI-IS

5xHWII

Albert Greinöcker

January 18, 2023



1 Hypothesen testen

So wie im Student Performance - Datensatz besprochen soll nun ein etwas größerer Fragebogen (European Social Survey¹) kennengelernt und teilweise ausgeweitet werden. Ein Download des Datensatzes plus einem PDF in dem genaue Beschreibungen zu den Variablen zu finden sind, befindet als zip-File in Moodle bei dieser Aufgabenstellung.

1.1 Importieren und Kennenlernen des Datensatzes

Die entsprechende Datei entpacken und die Datei ESS8e02.1_F1.csv so wie im gemeinsamen Skript zum European Social Survey (ex_10_ess.py) beschrieben, importieren.

```
df = pd.read_csv('../path.../ESS8e02.1_F1.csv', sep=',')
```

Im Datensatz sind die einzelnen Variablen in Zahlen kodiert (so wie es sich gehört). Will man diesen Zahlen Beschriftungen zuordnen, macht man das auf diese Art:

```
df['gndr'] = pd.cut(df['gndr'], [0,1,2,9], labels=['Male', 'Female', 'No Answer'])
```

Listing 1: Hier wird z.B. der Variable gndr (Geschlecht) die Werte "Male", "Female", "No Answer" für die Zahlen 1, 2, 9 zugeordnet.

Hier wird für 1 Male zugeordnet, für 2 Female und für 9 No Answer.

In einzelnen Auswertungen wird der Vergleich zwischen 2 Ländern oder die Situation in nur einem Land untersucht. Aus diesem Grund müssen Subsets mit nur diesen Ländern gebildet werden. Das geht z.B. so:

¹<https://www.europeansocialsurvey.org>

```

1 df_de = df.loc[df.cntry == 'AT']
2 df_de_it = df.loc[df.cntry.isin(['AT', 'IT'])]

```

Listing 2: Hier wird 1x ein Datensatz aller Österreicher und 1x ein Datensatz mit Österreichern und Italienern erzeugt.

1.2 Test-Übersicht

Folgende Tests sind bei den einzelnen Situationen anzuwenden (entsprechende Beispiele dazu gibt es im Projekt auf Github):

- 2 Variablen gegenübergestellt mit Messniveau
 - Nominalskala: χ^2
 - Ordinalskala: Korrelation (Spearman)
 - Intervall/Ratioskala: Korrelation (Pearson)
- 2 unabhängige Gruppen (z.B. Männlich/Weiblich) gegenübergestellt Variablen mit Messniveau
 - Ordinalskala: Mann-Whitney-U-Test
 - Intervall/Ratioskala: t-Test (haben wir nicht besprochen, kommt in diesem Fragebogen nicht vor)

1.3 Aufgabenstellen

In den eigentlichen Aufgabenstellungen sollen bestimmte Hypothesen getestet werden. Es soll bei jeder Hypothese...

- a. eine Aufbereitung der Daten gemacht werden,
- b. eine grafische Veranschaulichung (sofern möglich) gemacht werden,
- c. ein passender Hypothesentest ausgewählt und ausgeführt werden und
- d. eine Interpretation der Ergebnisse erstellt werden.

Hier die Hypothesen (in Klammern die betroffenen Variablen):

- a. Männer haben mehr Glauben in die Polizei (trstplc, gender).
- b. Es besteht ein negativer Zusammenhang bei "mehr Strom aus nuklearer Energie" und "mehr Strom aus Solarenergie" (elgnuc, elgsun).
- c. In Österreich ist der Eindruck, dass sich der Klimawandel schlecht auf die Menschen auswirkt, stärker als in Ungarn (ccgdbd).
- d. Frauen stimmen einem bedingungslosen Grundeinkommen eher zu (basinc).

Bitte noch 2 Hypothesen selbst wählen und überprüfen!