# Übersicht

## Variablen

29 Variablen soll es im Endprodukt geben:

|  |  |  |
| --- | --- | --- |
| item |  |  |
| all | Anzahl Gesamtnennung |  |
| west | Westdeutschland |  |
| east | Ostdeutschland |  |
| age1 | Alter in Jahren (14 – 29) |  |
| age2 | Alter in Jahren (30 – 44) |  |
| age3 | Alter in Jahren (45 – 59) |  |
| age4 | Alter in Jahren (60 +) |  |
| male | Männlich |  |
| female | Weiblich |  |
| edu1 | Abschluss: Haupt/Volksschule |  |
| edu2 | Abschluss: MR/POS |  |
| edu3 | Abschluss: Abitur/FH |  |
| employ1 | Berufstätig: Ja |  |
| employ0 | Berufstätig: Nein |  |
| income1 | Monatl. HH-Netto-EK < 1,5 |  |
| income2 | 1500 =< x < 3000 |  |
| income3 | 3000 < x |  |
| cdu | CDU/CSU |  |
| spd | SPD |  |
| leftists | Die Linke |  |
| greens | B90/Grüne |  |
| fdp | FDP |  |
| afd | AfD |  |
| elec\_else | Sonstige |  |
| elec\_none | Nicht/ungültig wählen |  |
| elec\_dk | Parteipräferenz: weiß nicht |  |
| sp\_1 | Survey period 1 (from…) |  |
| sp\_2 | Survey period 2 (to…) |  |

Pro Woche folgende Zeilen:

* Alle Items **Zusammenfassung (1. & 2. Nennung)**
* Alle Items **1. Nennung**
* Alle Items **2. Nennung**

# Tasks

### Überprüfung der Variablen zur Parteienpräferenz (Funktion)

Zu unterschiedlichen Zeitpunkten wurden nicht immer alle Parteien in der Umfrage abgefragt, sondern unter „Sonstige“ subsummiert. Zwecks mergings unseres Gesamtdatensatzes brauchen wir konsistente Variablen.

To-do: Funktion schreiben, welche überprüft welche Parteienvariablen im spezifischen Sheet existieren; nicht existente Variablen erstellen (NAs).

## Konsistente (Um-)Benennung der Variablen

Soziodemographische Variablen (und Parteienpräferenz, so noch nicht geschehen) konsistent benennen. Überflüssige Zeilen löschen.

## Identifizierung falsch zusammengefügter Zeilen (Funktion)

Im OCR Prozess wurden vereinzelt Zeilen irrtümlich zusammengefügt.

**To-do:** Funktion schreiben, welche Zeilen auf Korrektheit überprüft bzw. fälschlicherweise zusammengefügte Zeilen identifiziert.

## Merging pro Woche

## Erhebungszeitraum aus Info\_table extrahieren (Done)

Erhebungszeitraum aus Info\_table extrahieren und allen Items der zugehörigen Woche in Variablen sp\_1 und sp\_2 zuweisen (von… bis…).

## General merging

Produkt: 3 Tabellen (Aggr, first, scnd)