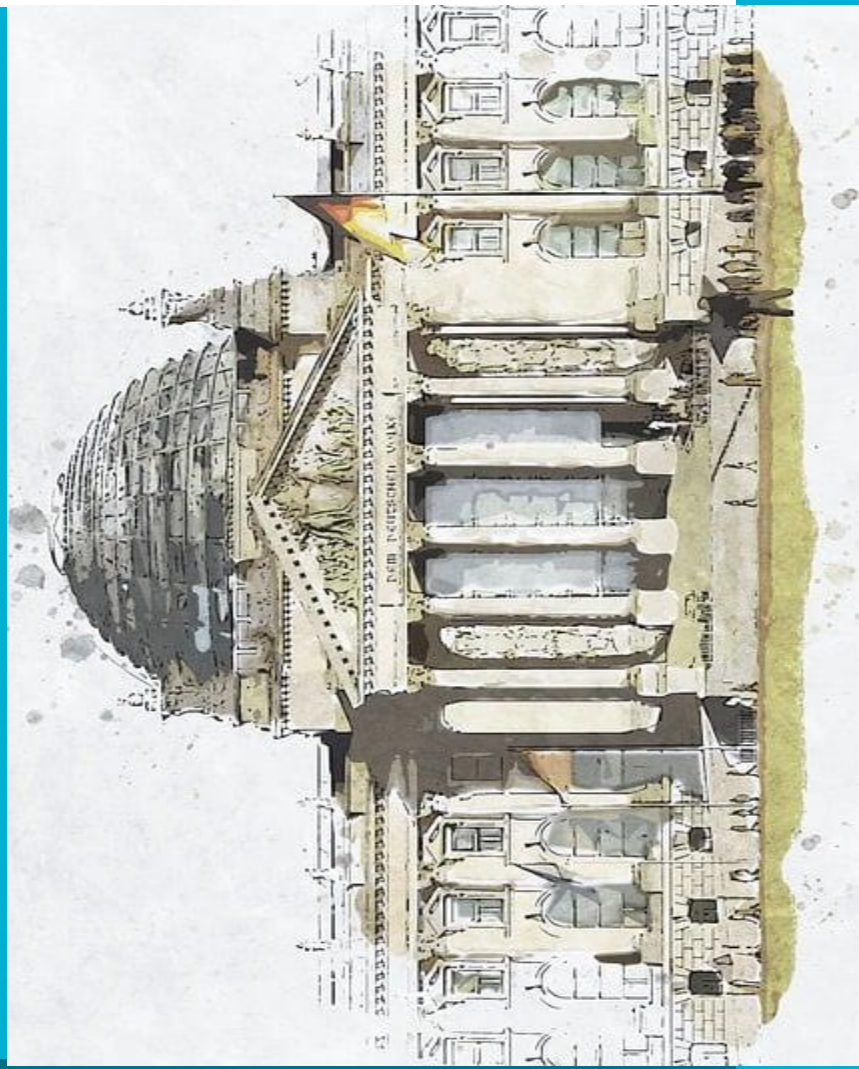
Project One: Making Sense of the German Bundestag

Paula Boks



Scope: Who sits in the German Bundestag?
(and getting to know their API...)

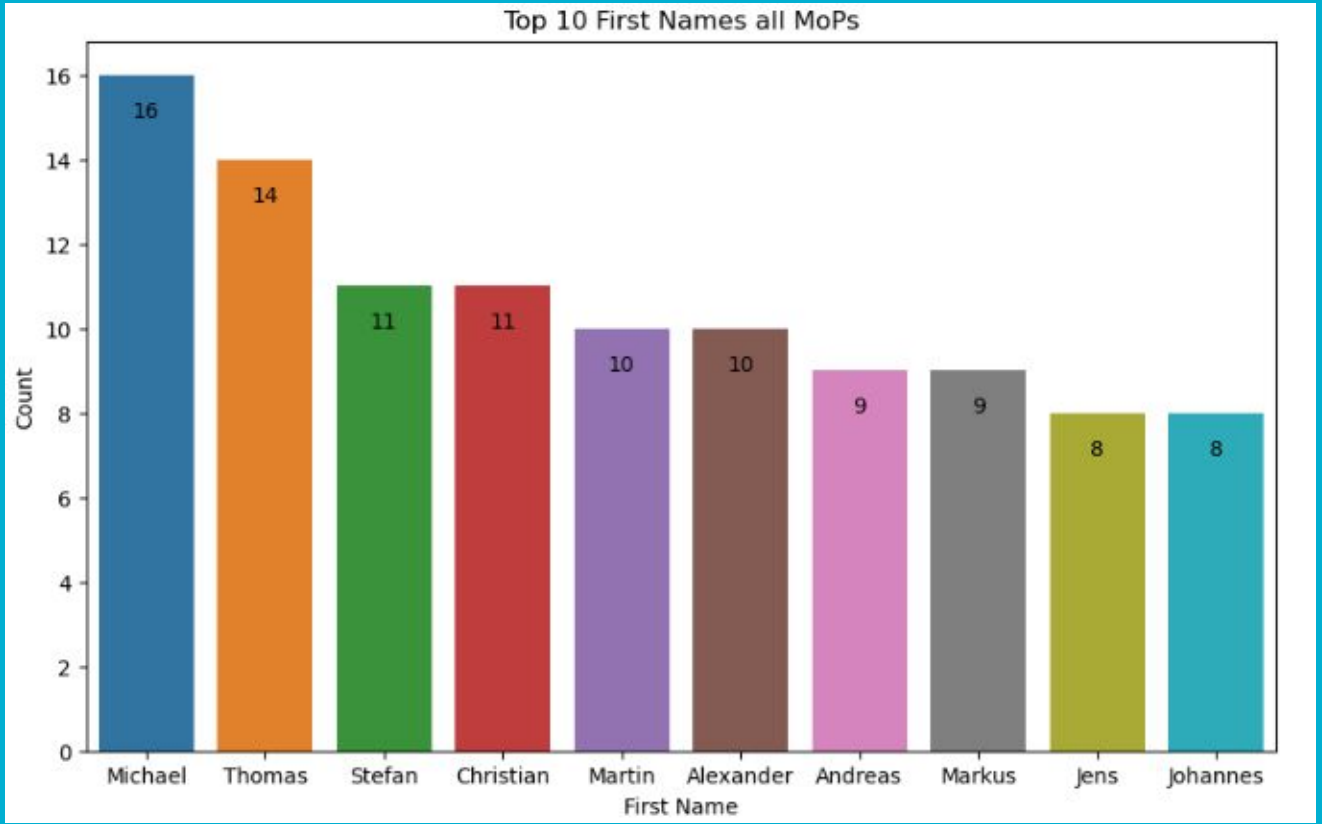
Hypotheses:

1. Distribution of **first names** generally and in regards to parties
2. Distribution of **seats** in parliament
3. **Academical grade** disproportionately higher than in population

Data Sources

- **API** of the German Bundestag
- **Webscraping** German Bundestag using **Selenium**

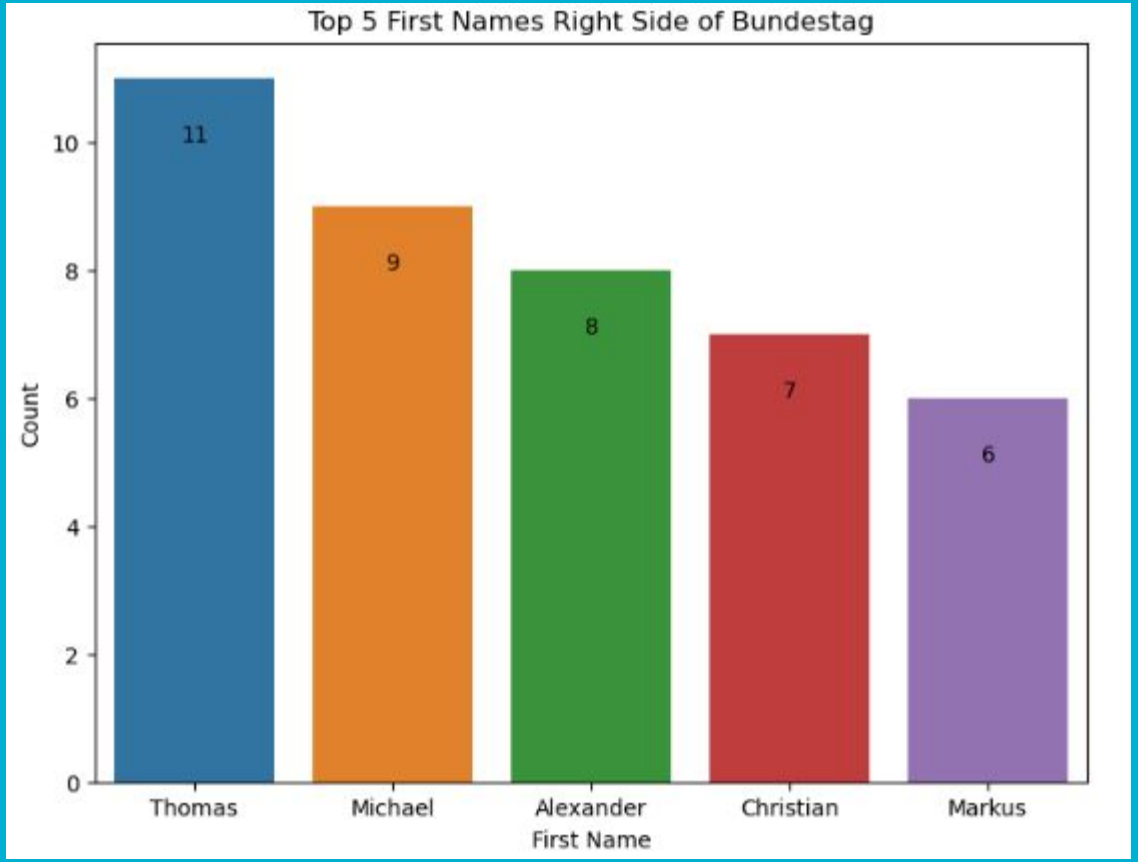
First Names


Top 10 First Names all MoPs
Count
First Name
0
2
4
6
8
10
12
14
16
16
14
11
11
10
10
9
9
8
8
Michael
Thomas
Stefan
Christian
Martin
Alexander
Andreas
Markus
Jens
Johannes

First Names on the "Right" Side

364 MoPs

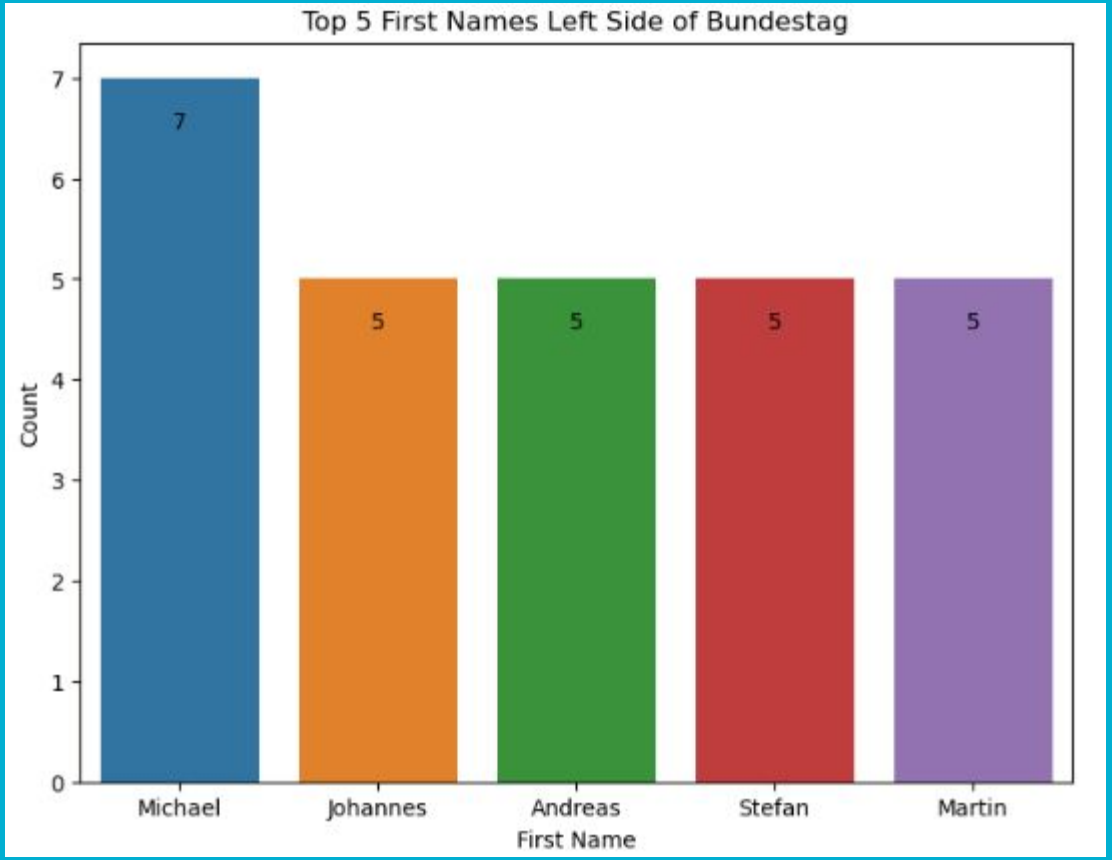
(AfD, CDU/CSU, FDP)



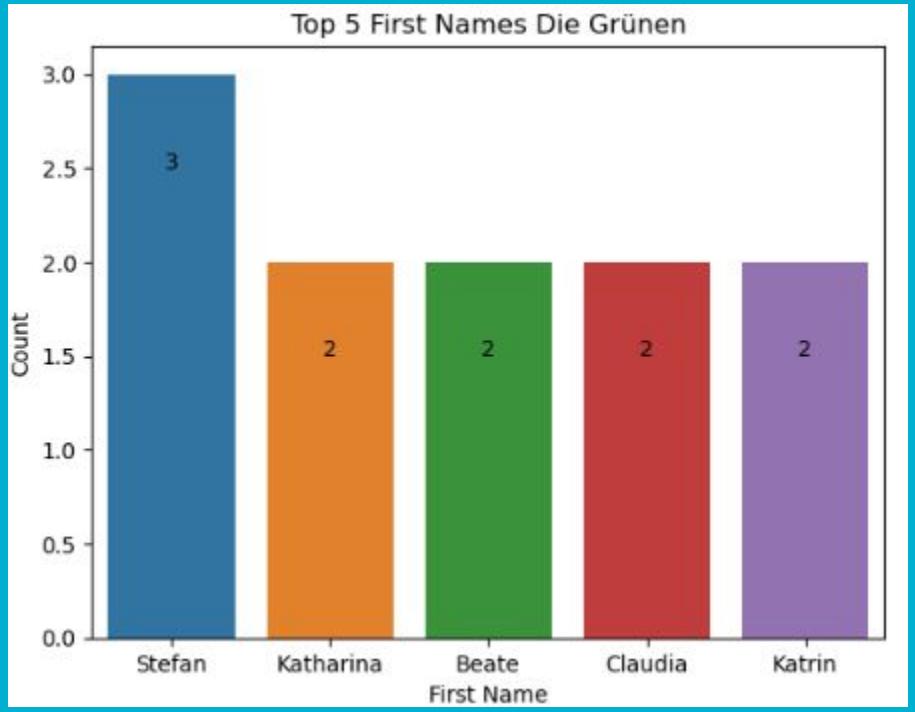
First Names on the "Left" Side

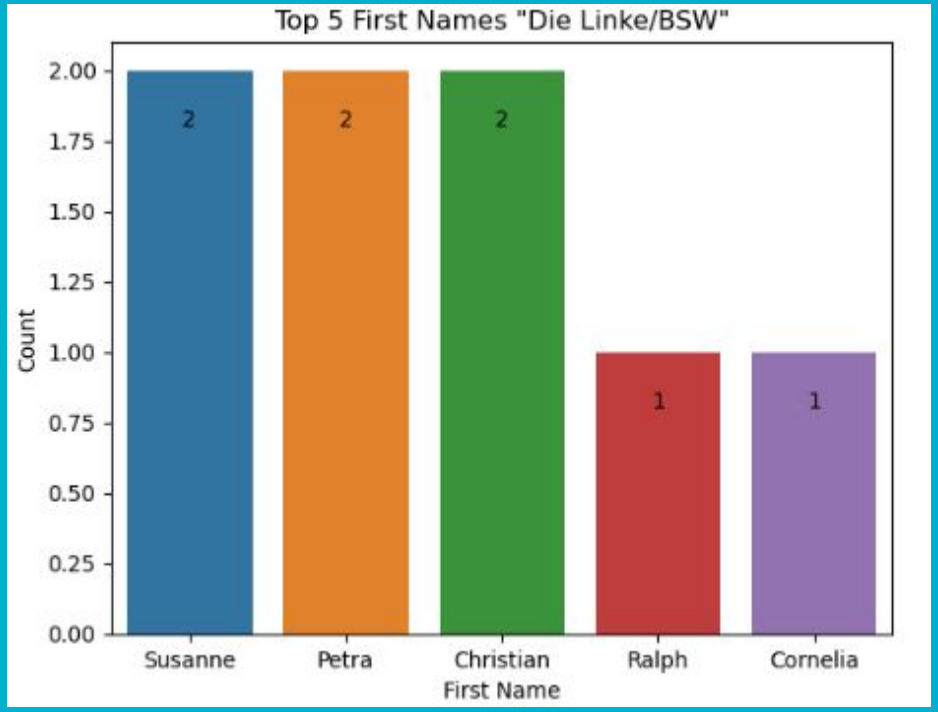
324 MoPs

(SPD, Die Grünen, "Die Linke")

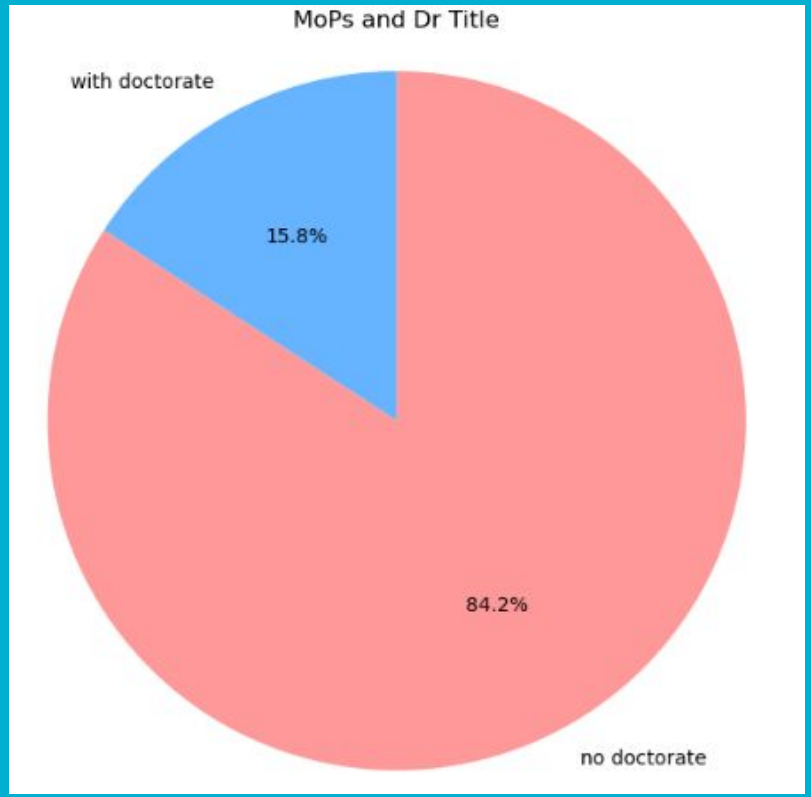


Looking for Women





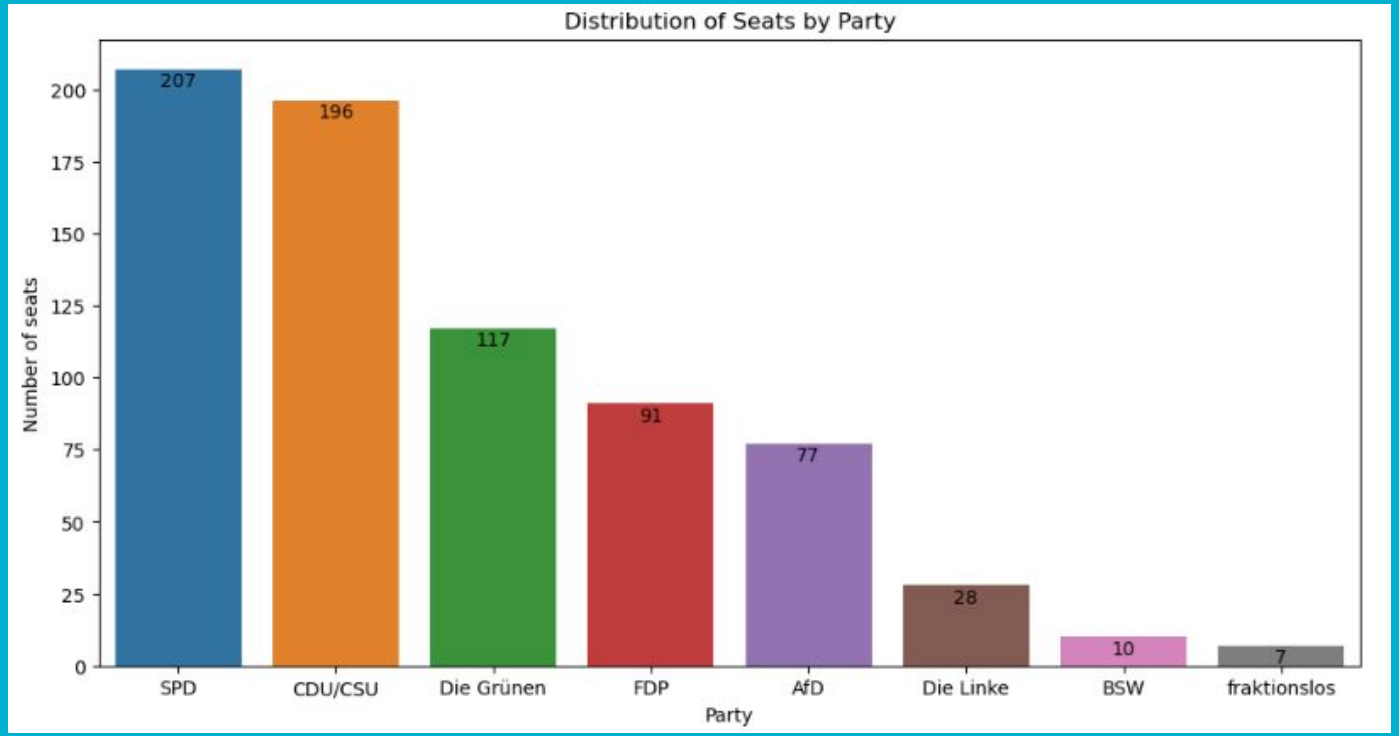
Academical Grade



Compared to **~1.2%** in the German population who have a doctorate

(https://www.academics.de/ratgeber/promotion-statistik [21.10.24])

Seat Distribution/ Reality Check



733 Sitze

Die Sitze verteilen sich wie folgt:

SPD	207
CDU/CSU	196
Bündnis 90/Die Grünen	117
FDP	91
AfD	76
Gruppe Die Linke	28
Gruppe BSW	10
fraktionslos	8

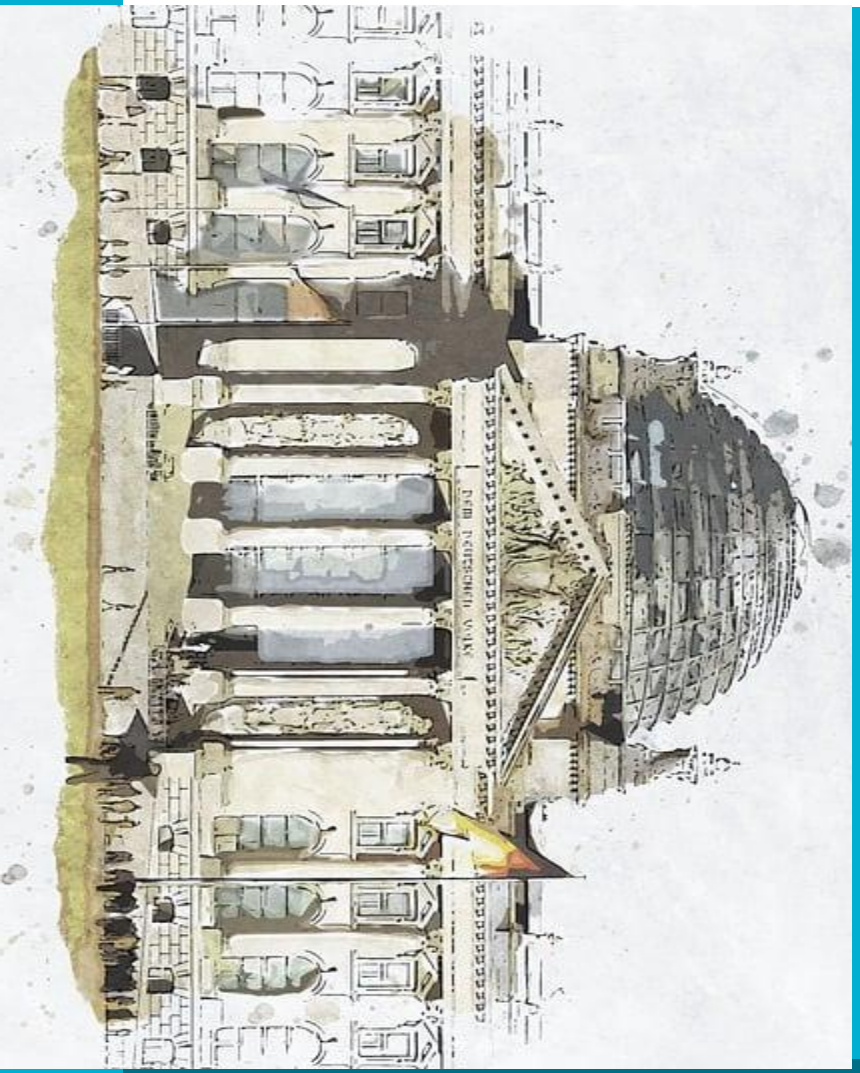
Source: https://www.bundestag.de/parlament/plenum/sitzverteilung_20wp [22.10.2024]

Making sense of the data: API

- Embedded dictionaries en masse
- What is this data?
  - What are all these duplicates?
  - Why is there only 398 MoPs?? (**733!!!**)

New data: Webscraping

- Validating names and function
- Selenium: how to deal with sliders
- Finally 733 MoPs



… Thank you!