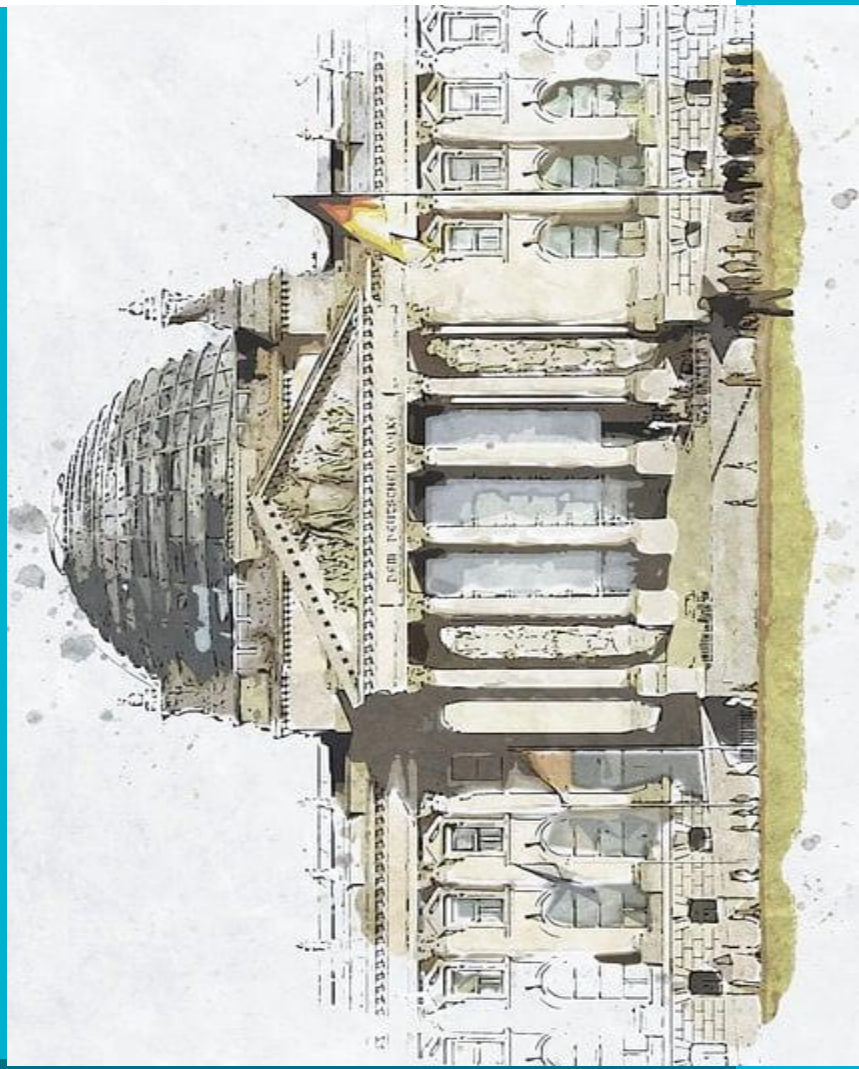
Project One: Making sense of the German Bundestag

Paula Boks



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

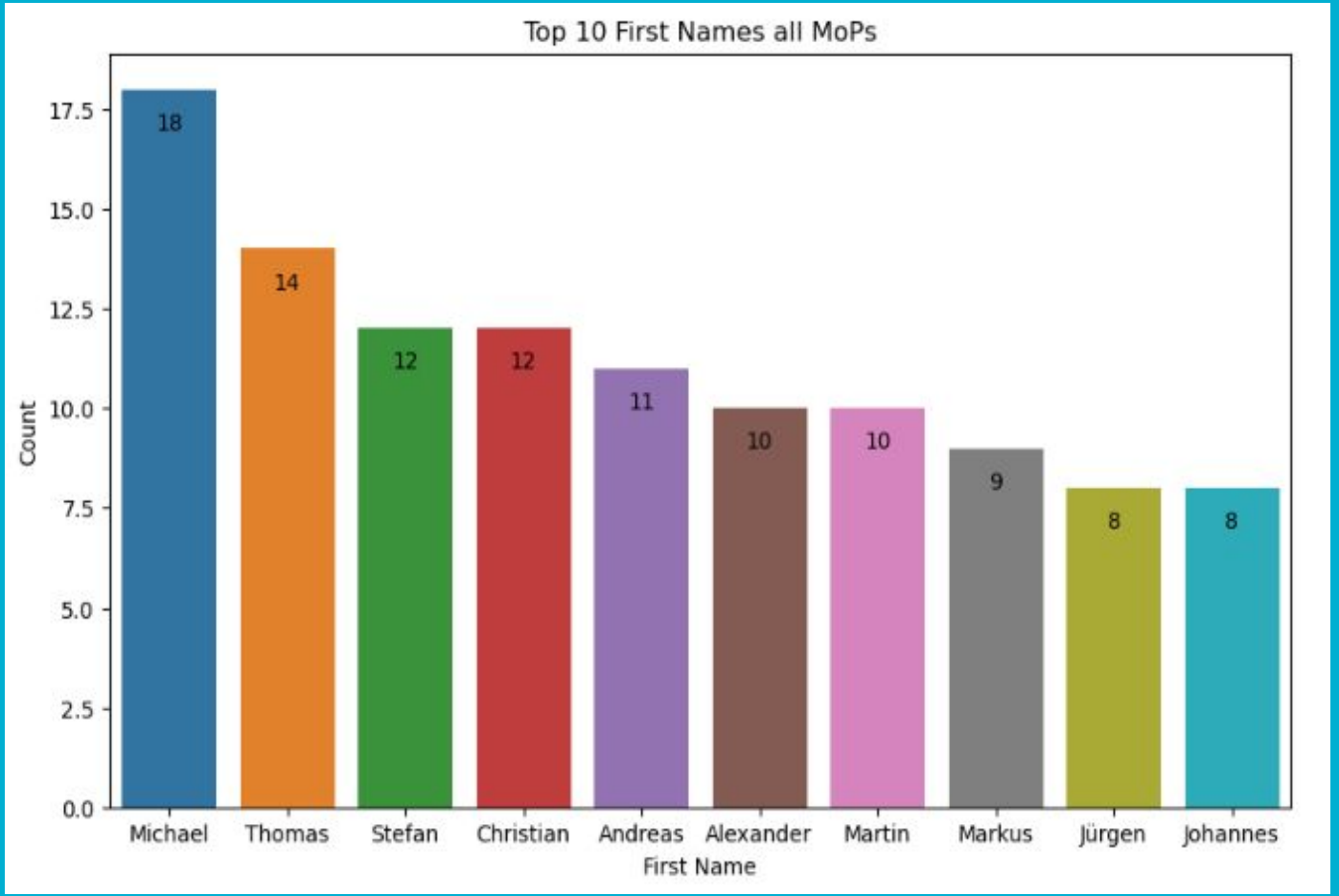
Hypotheses:

1. Distribution of **first names** generally and in regards to parties
2. Distribution of **seats** in parliament like in 2021
3. **Academical grade** disproportionately higher than in population

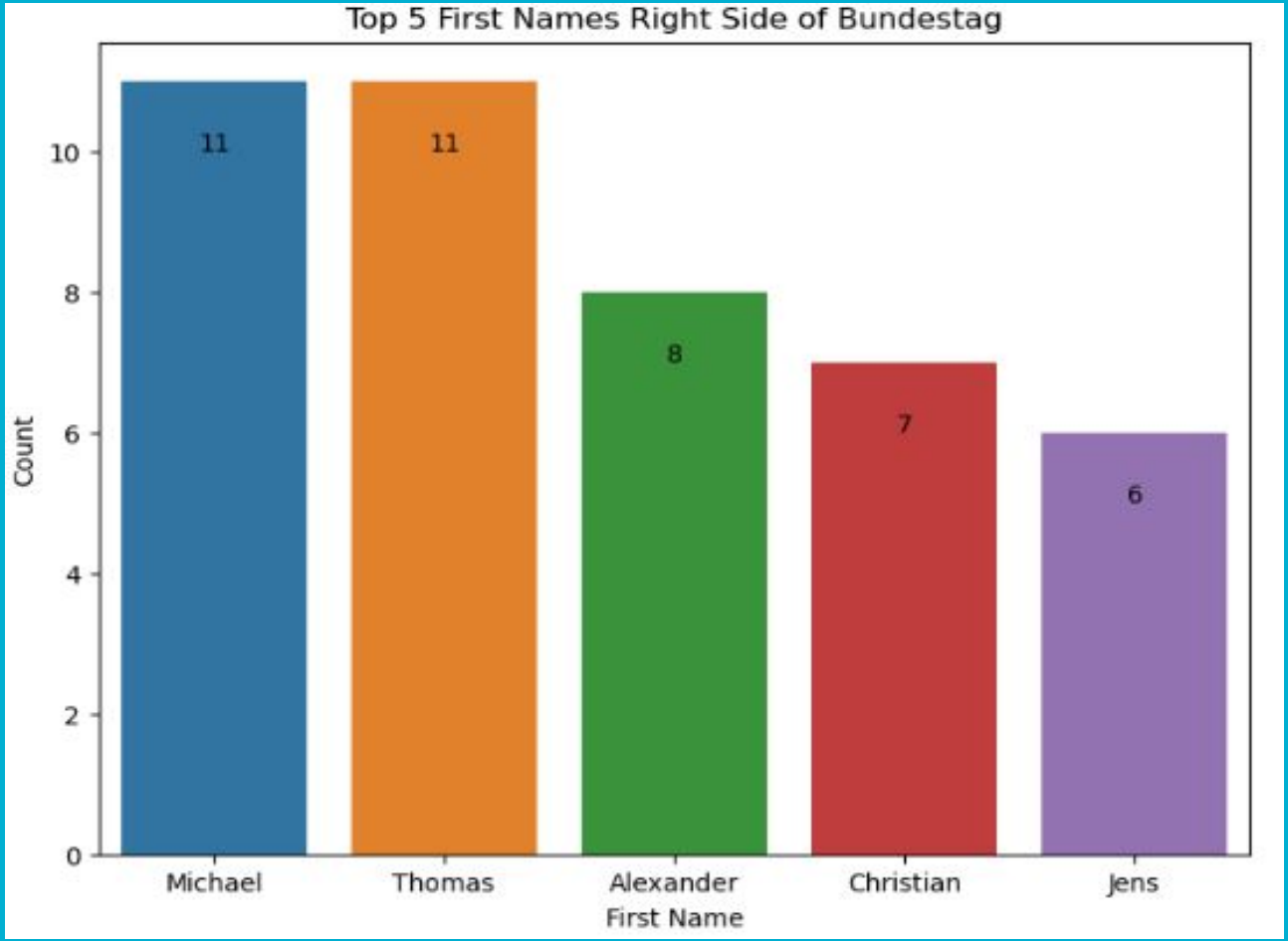
Data sources

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

First Names

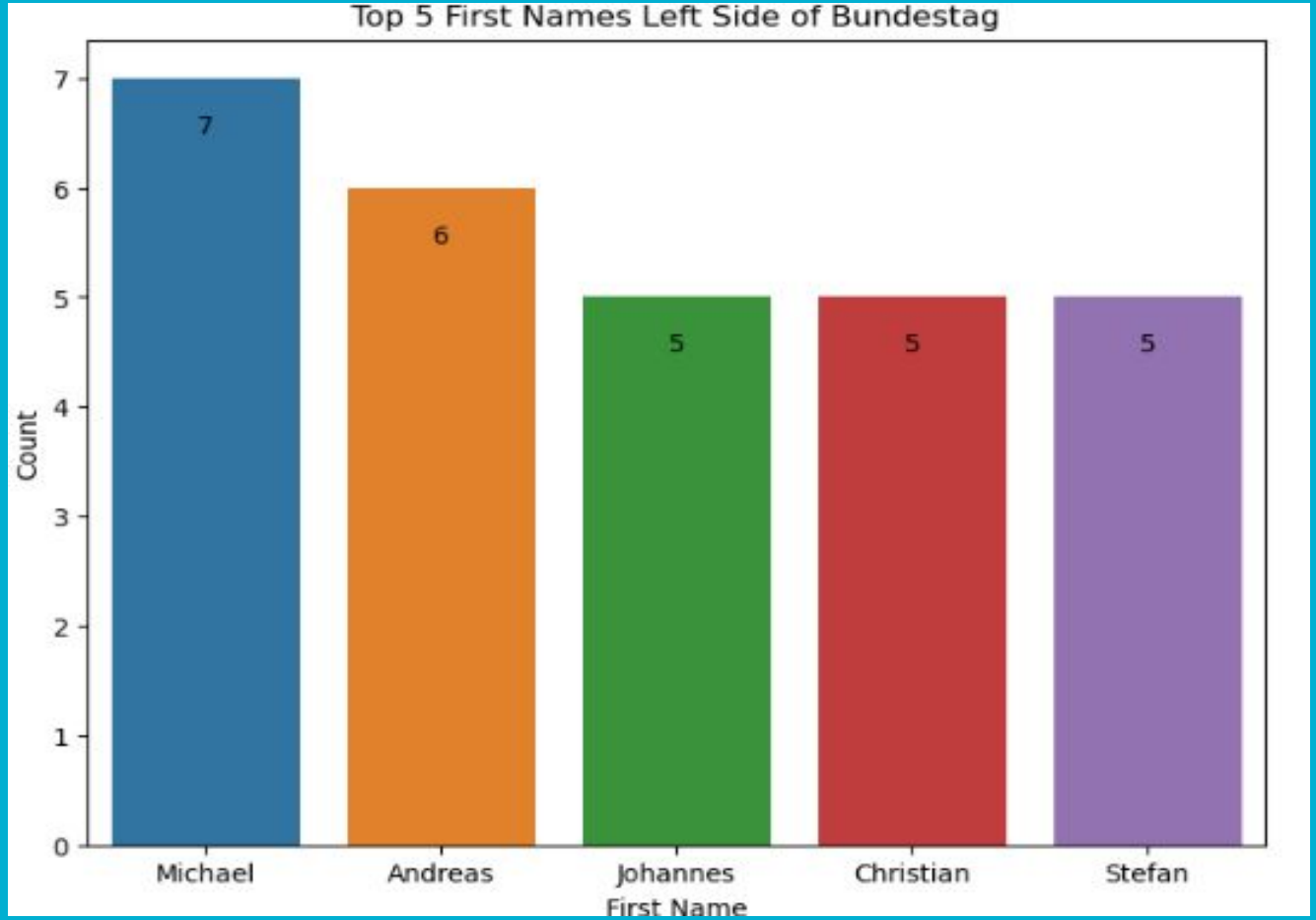

Top 10 First Names all MoPs
Count
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
18
14
12
12
11
10
10
9
8
8
Michael
Thomas
Stefan
Christian
Andreas
Alexander
Martin
Markus
Jürgen
Johannes
First Name

First Names on the "right" side



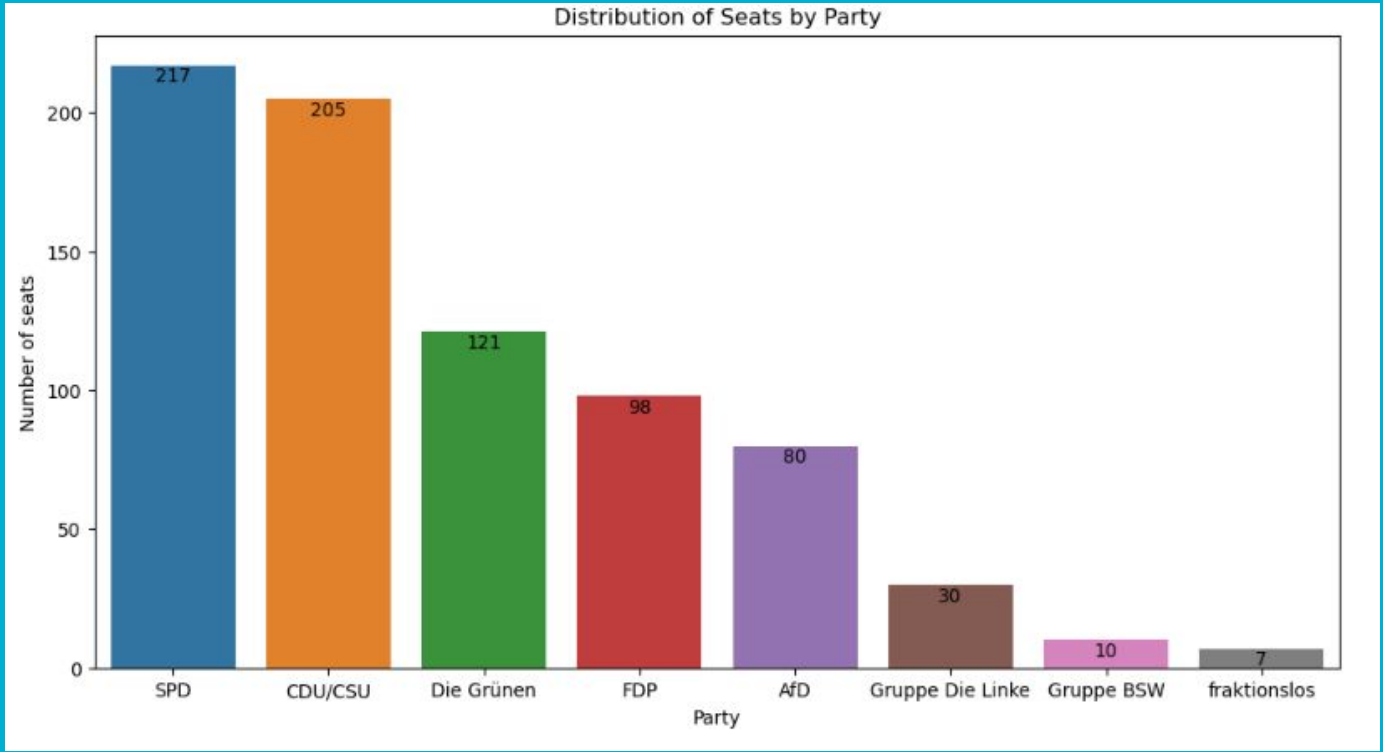
(AfD, CDU/CSU, FDP)

First Names on the "left" side



(SPD, Die Grünen, "Die Linke")

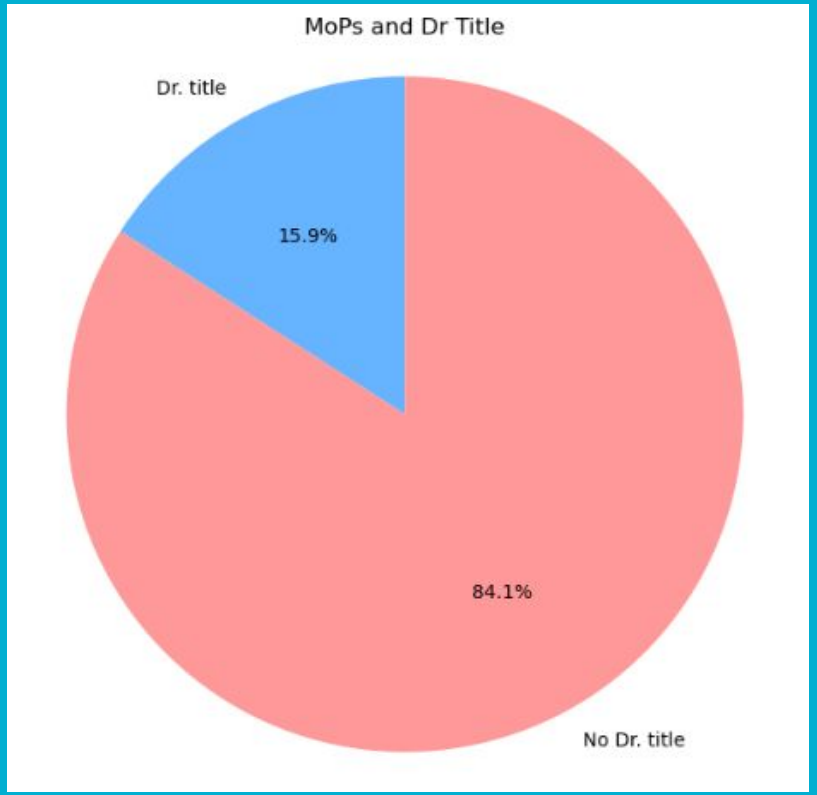
Seat distribution: what went wrong?



NOOOOO

≠

Academical grade



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

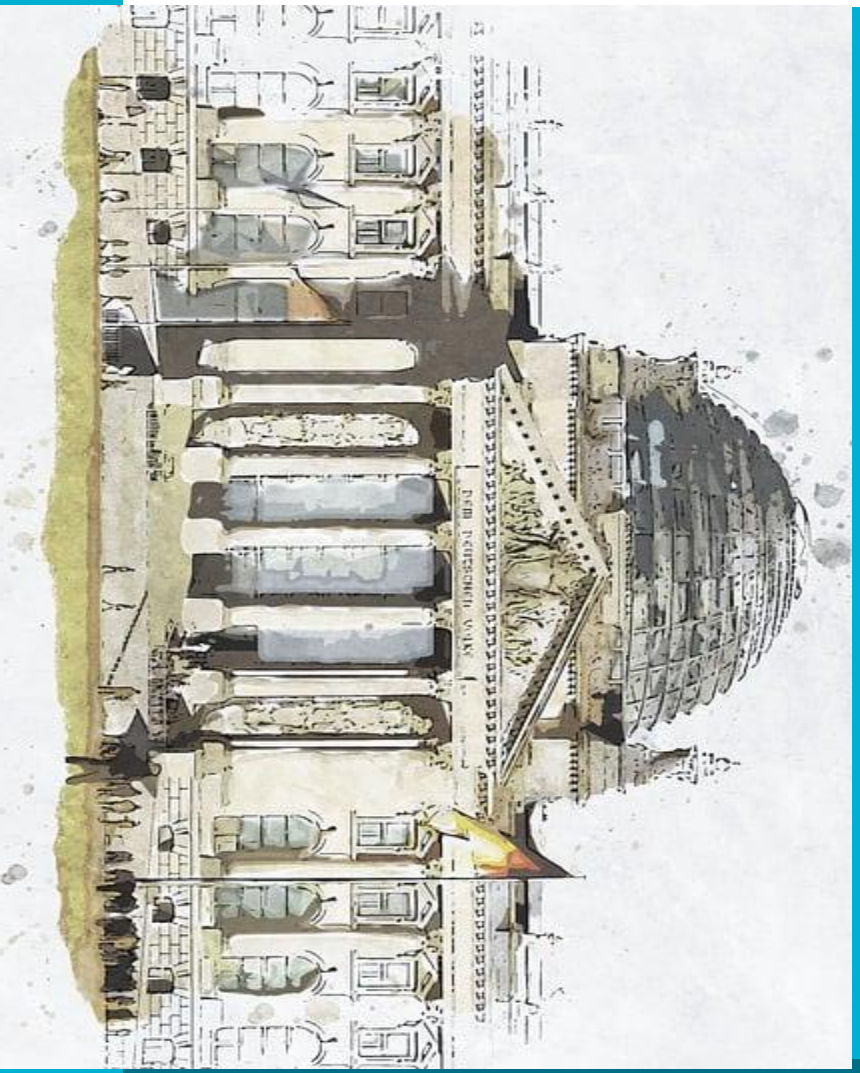
(https://www.academics.de/ratgeber/promotion-statistik [21.10.24])

Making sense of the data: API

- Embedded dictionaries en masse
- What do I have here?
  - What are all these duplicates?
  - Why is there only 398 MoPs?? (**733!!!**)

New data: Webscraping

- Validating names and function
- Selenium
- But: Why are there 768 MoPs? (733!!!)



… Thank you!