
Uncertainty Now!

A Data Driven Analysis of Coping Strategies in the COVID-19 Pandemic.


Tim Ruhe

Data Science Kolloquium, June 17th 2021

tim.ruhe@tu-dortmund.de



Jan-Hendryk de Boer, Daniel Ullrich , Timo Lüke, Isabelle Borucki, Elena Weber, Alexander Unser, Jithender Timothy, Maren Freudenberg, Cristian Cercel



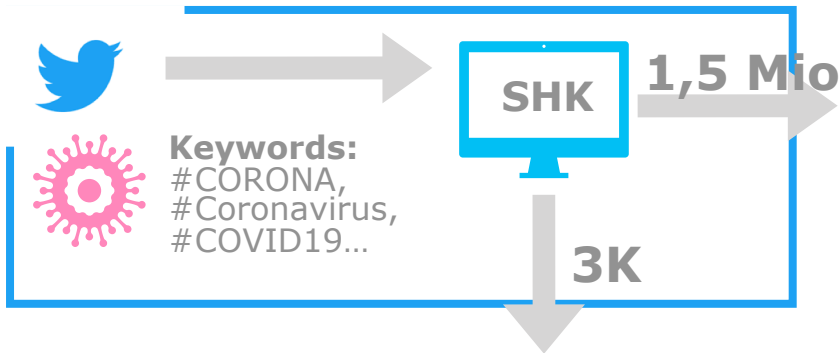
Premise: There are certain coping strategies that people apply to deal with situations of inherent uncertainty (e.g. the COVID-19 pandemic, the European refugee crisis, the worldwide banking crisis...)

Question: Which of these strategies are actually applied (expressed in social media) and how does that correlate for example with elections, regions or socio-economic factors (net income)?

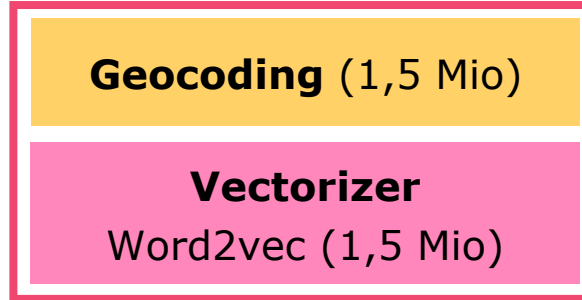
Jan-Hendryk de Boer, [Daniel Ullrich](#) , Timo Lüke, [Isabelle Borucki](#), Elena Weber, [Alexander Unser](#), [Jithender Timothy](#), Maren Freudenberg, Cristian Cercel

WORKFLOW

DATA DOWNLOAD, CLEANING



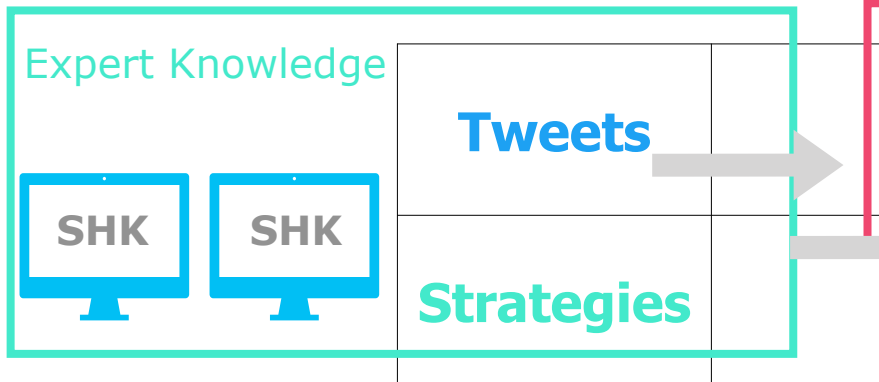
AI/DATA (Tim/Jithender)



ANALYSE



Encoding (Isabelle, Alexander)



> STATISTIK

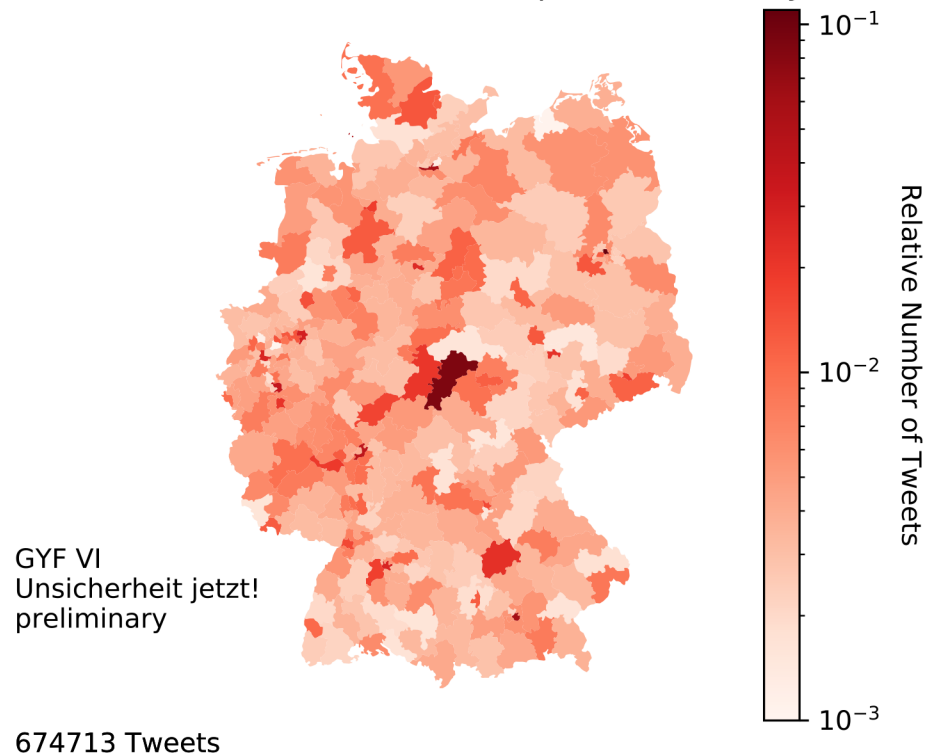


Distribution of Tweets and Sanity Check

Top 10 Constituencies:

1. Berlin-Mitte 75 (23,5%)
2. Eisenach – Wartburgkreis – Unstrut-Hainich-Kreis (8,4%)
3. Hamburg-Altona 19 (5,5%)
4. München-Ost 218 (4,5%)
5. Frankfurt am Main II 183 (4,0%)
6. Köln I 93 (2,5%)
7. Hannover-Land I (1,6%)
8. Dresden I (1,6%)
9. Mainz (1,6%)
10. Stuttgart I (1,6%)

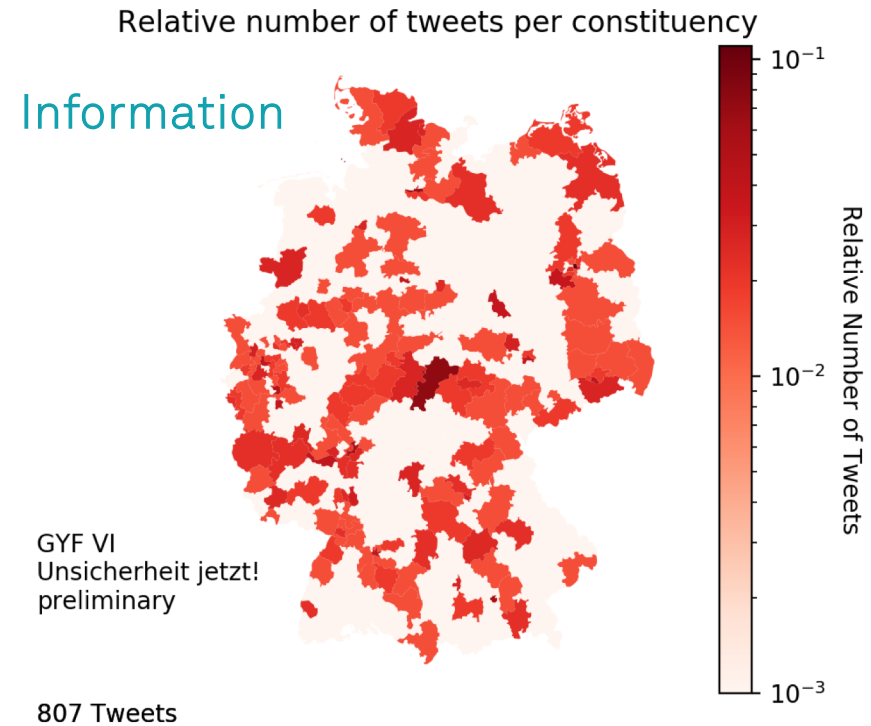
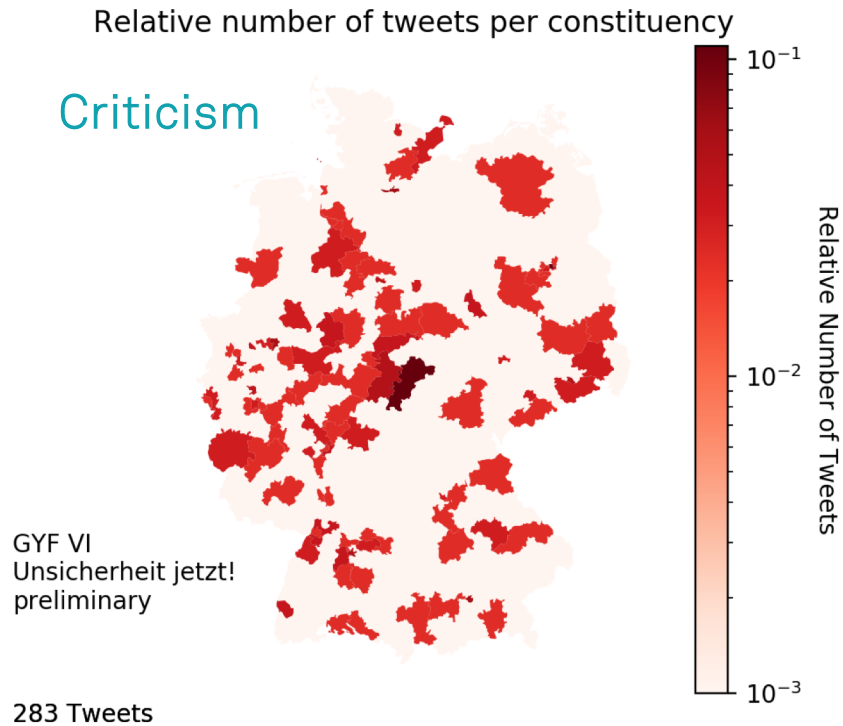
Relative number of tweets per constituency



Null Hypothesis:

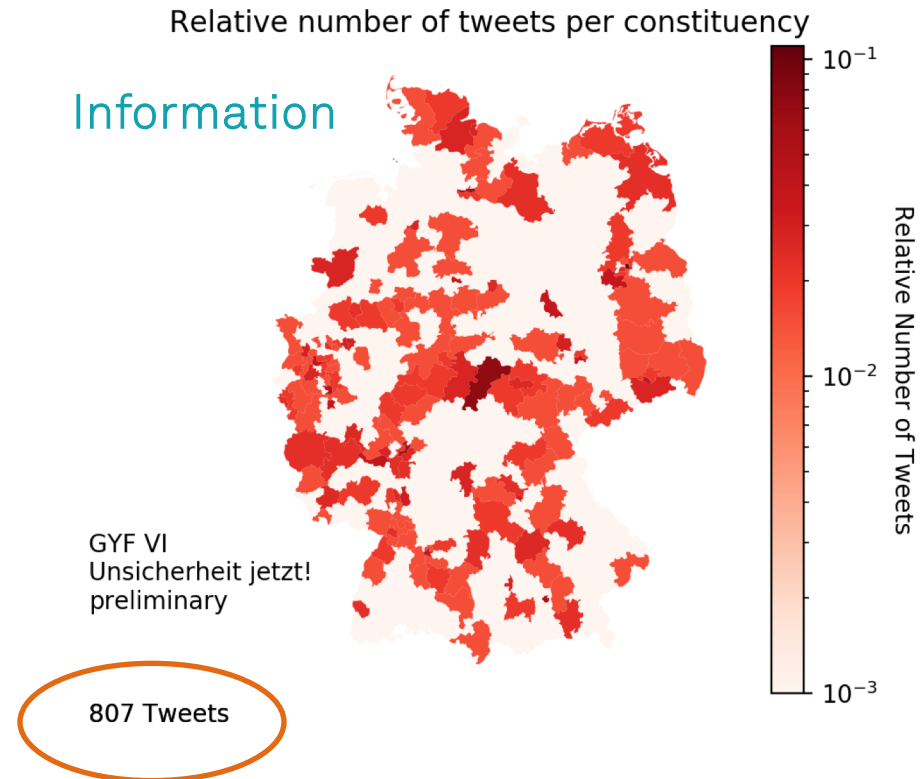
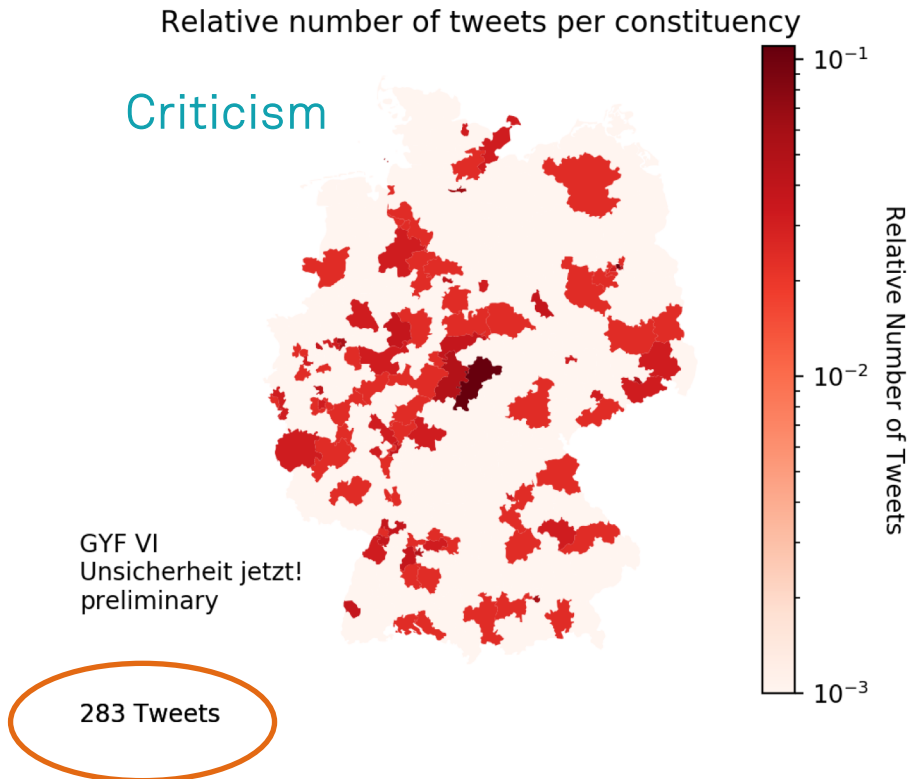
Tweets expressing certain coping strategies are distributed in the same way as all tweets.

Criticism and Information



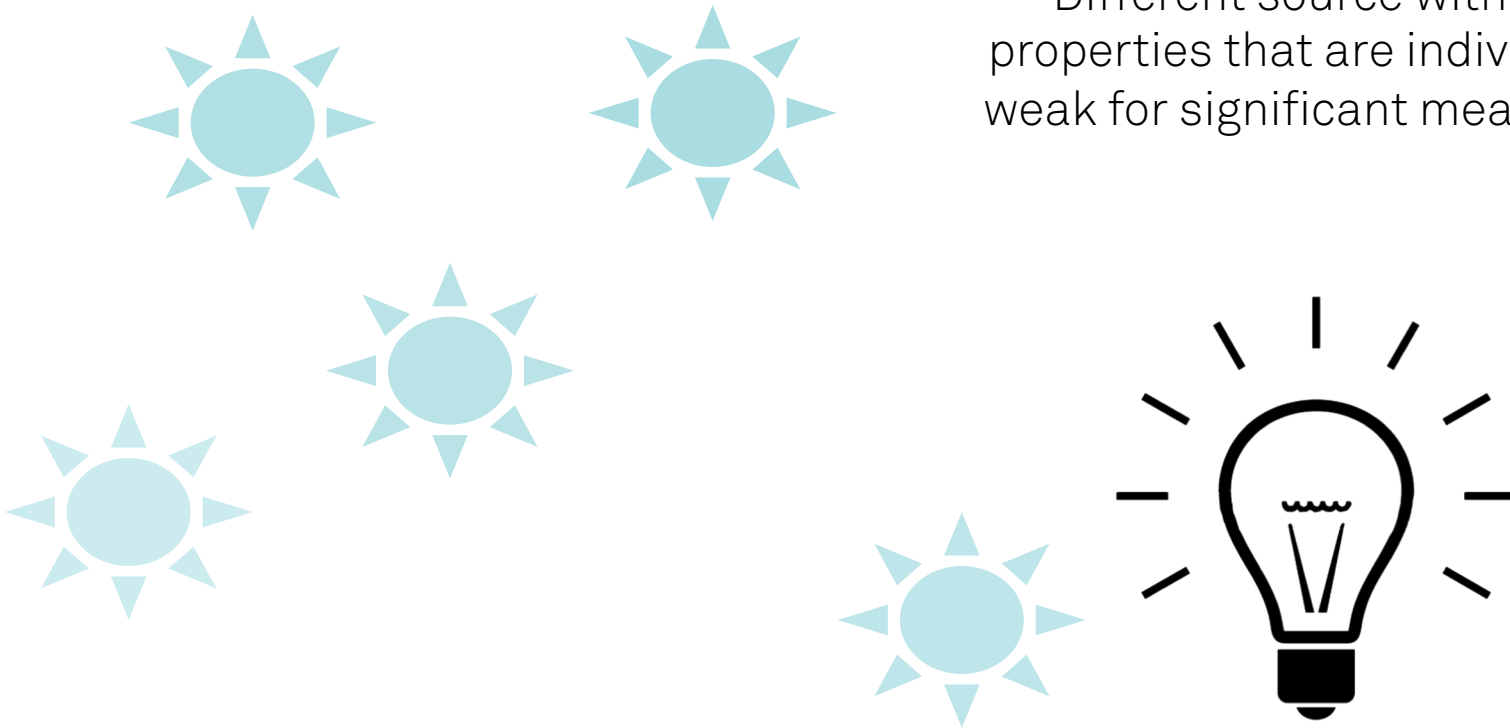
Criticism and Information

Meaningful conclusions are challenging, due to the small numbers.

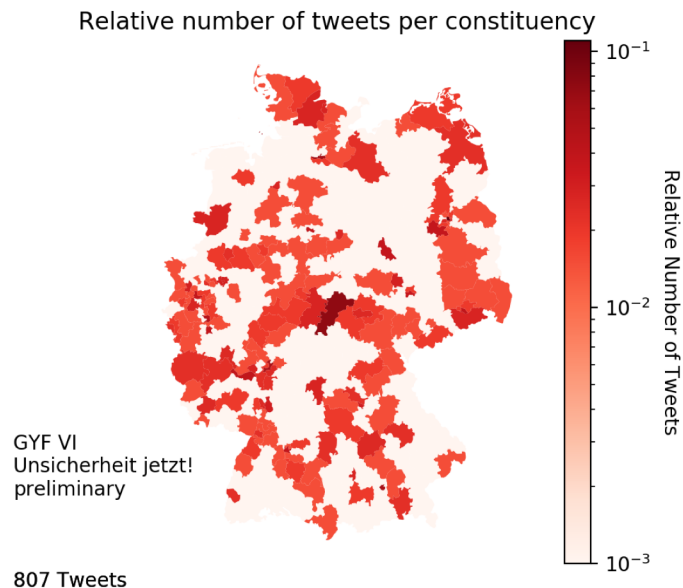


Source Stacking in Astroparticle-Physics

Different source with similar properties that are individually too weak for significant measurement.



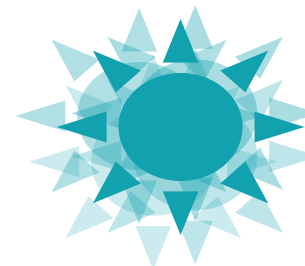
Source Stacking in Astroparticle-Physics



Idea: Pile individual sources on top of each other to obtain a significant measurement.



Stack constituencies for example by direct mandate

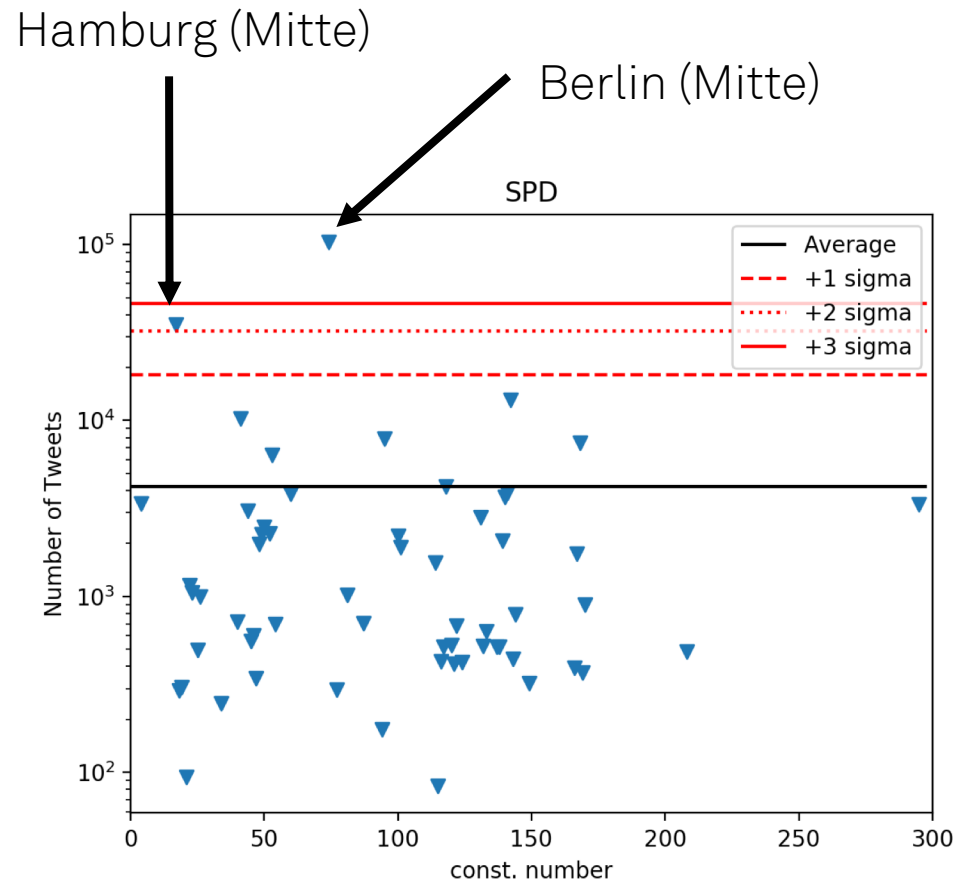
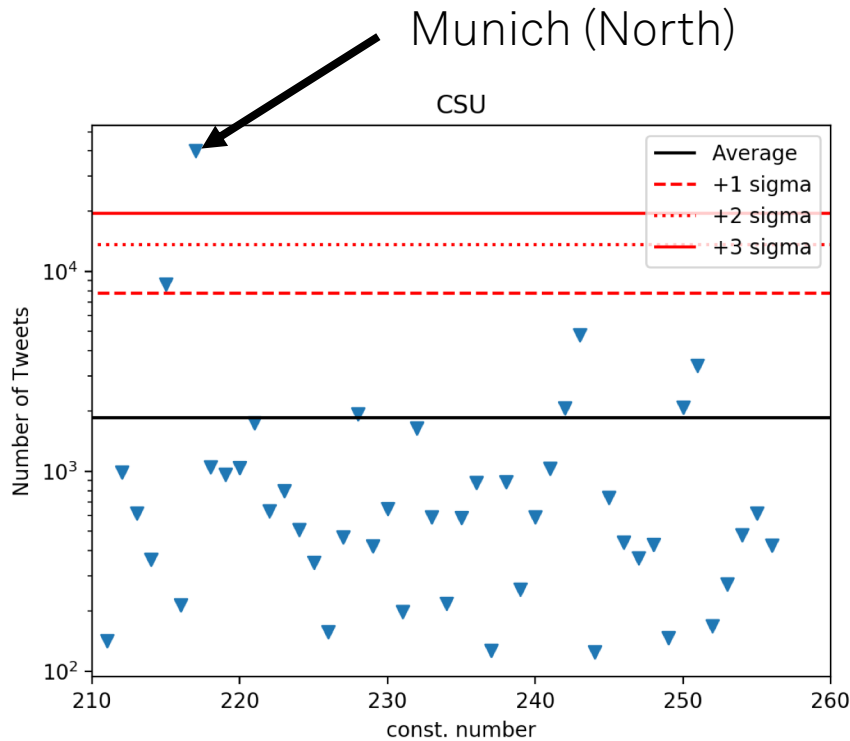


Simulation

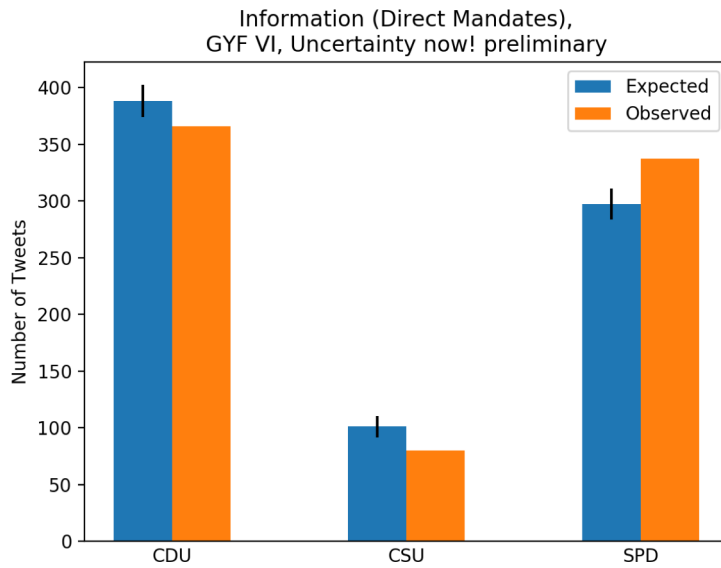
- Obtain probability p of a tweet to stem from a certain constituency
- Obtain number of tweets n observed for a certain coping strategy
- Sample n tweets according to probability p
- Repeat n_{trial} time

- Group tweets by criterion, e.g. direct mandate, compute expectation
- Compare expectation and observation

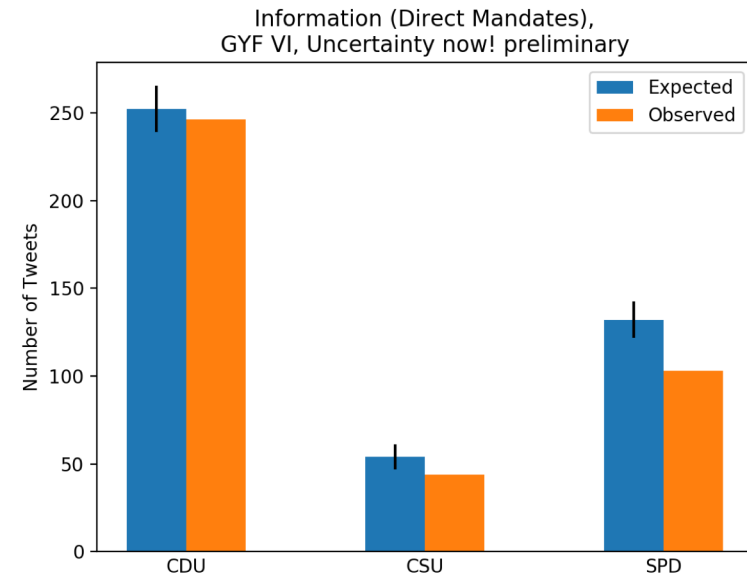
Possible Bias towards larger cities



Selected (and preliminary) Results



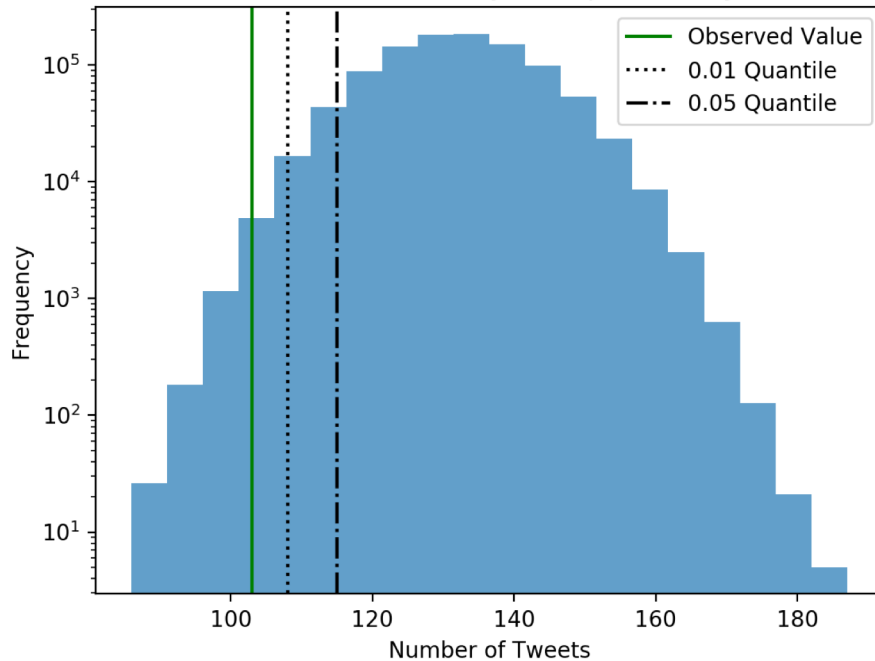
All constituencies



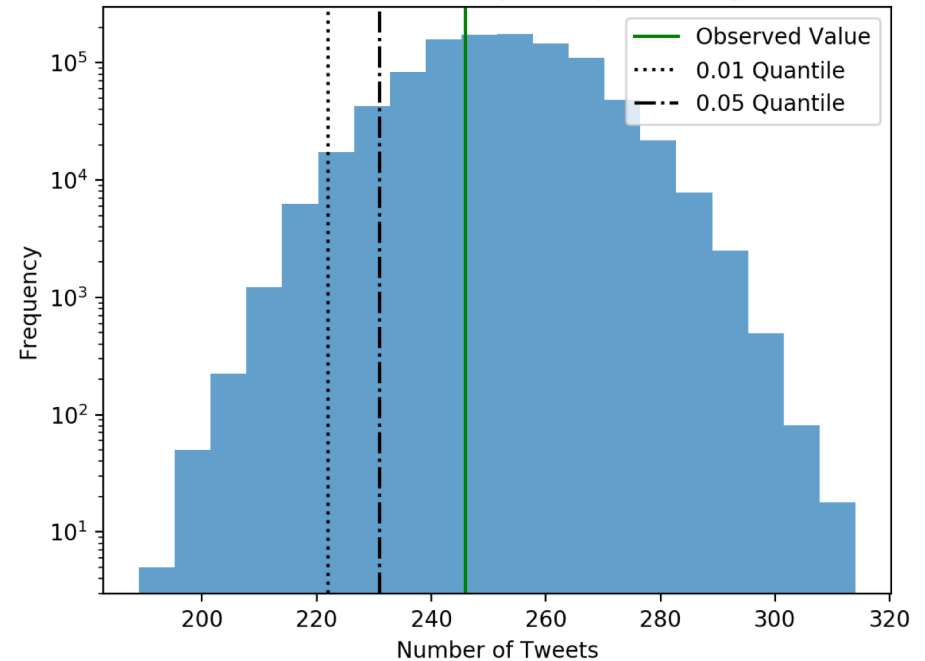
Constituencies above 2 sigma excluded

More Results on Information

SPD
GYF VI, Uncertainty now! preliminary



CDU
GYF VI, Uncertainty now! preliminary



Summary & Outlook

