



Helge Marahrens & Anne Kavalierchik
Indiana University Bloomington
Department of Sociology
Email correspondence to: hmarahre@iu.edu

Workshop in Methods (WIM)

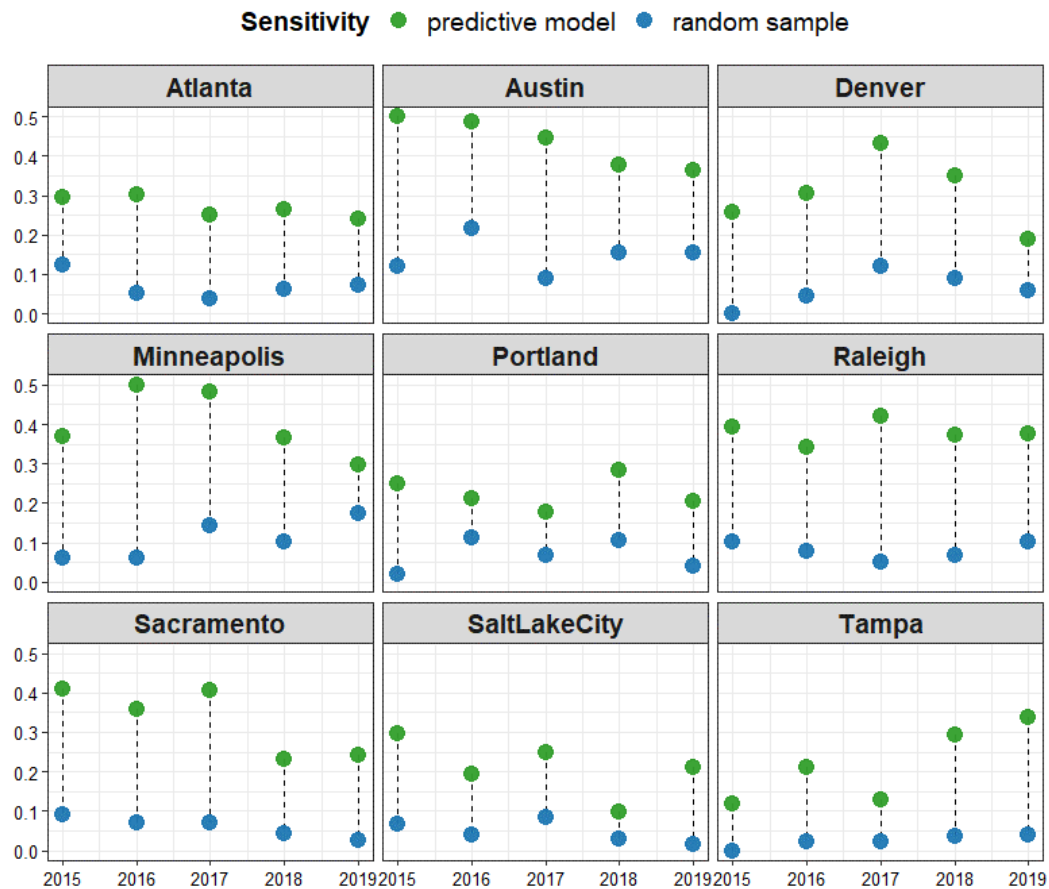
Introduction to Python for Social Scientists



- ❑ demystify Python
- ❑ get you started
 - basics
 - dealing with messy data



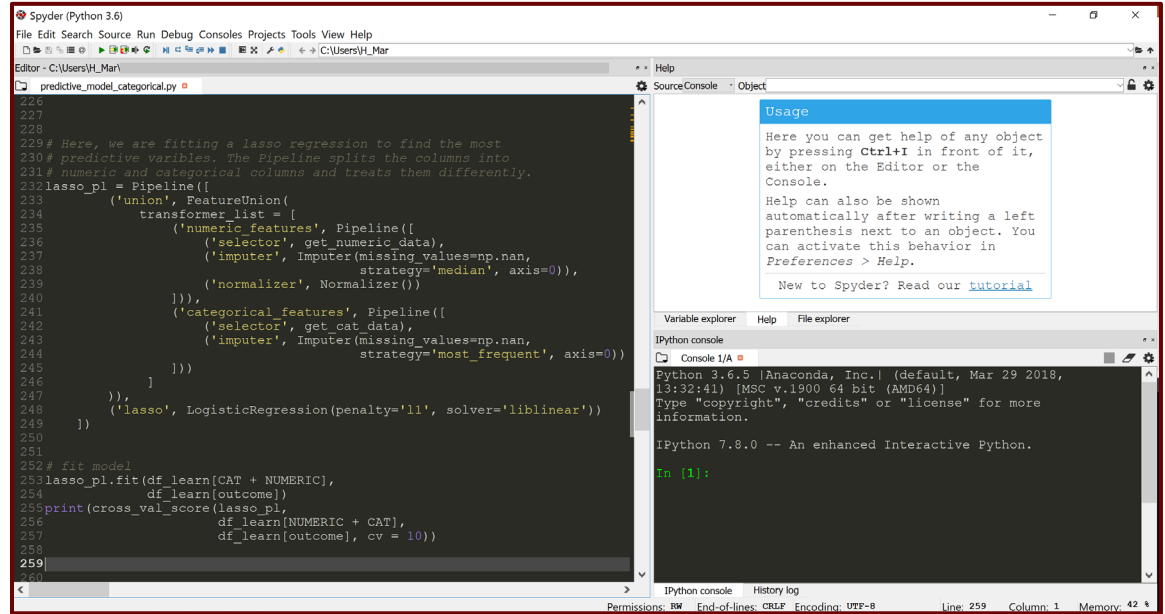
- digital life/ digital trace data
- large data
- textual data
- predictive modelling



- Python 3
- IDE → IDLE

Other IDE's:

- Spyder
- Jupyter Notebook
- ...





- data types
- function vs. method
- indentation is key (4 spaces)
- counting begins at zero



Coding languages, coding as language

- There are many different ways to express the same thought. For example, all of the following mean essentially the same thing:
 - Are we finished?
 - Is this over?
 - Has this ended?
 - Can I leave?
- In Python, there are many different ways to do a given task



Part 2: Data Wrangling

- ProPublica API
- `congress_house.xlsx`
 - Excel spreadsheet of the current House of Representatives
- `ProPublica_Members-Bills.pkl`
 - “pickle” file (Python module) of 1000 bills
- How combine the two to see **how many bills** members of the House of Representative have passed?





INDIANA UNIVERSITY BLOOMINGTON