

Python Data Analytics Tutorials

The bulk of my research involves some degree of 'Big Data' – such as datasets with a million or more tweets. Getting these data prepped for analysis can involve massive amounts of data manipulation – anything from aggregating data to the daily or organizational level, to merging in additional variables, to generating data required for social network analysis. For all such steps I now almost exclusively use Python's PANDAS library ('Python Data Analysis Library'). In conjunction with the Jupyter Notebook interactive computing framework and packages such as NetworkX, you will have a powerful set of analysis tools at your disposal. This page contains links to the tutorials I have created to help you learn data analytics in Python. I also have a page with shorter (typically one-liner) data analytic code bytes.

Data Collection

- Python: Where to Start?
- Running your first code
- Four ways to run your code
- Setting up Your Computer to Use My Python Code
- Setting up Access to the Twitter API
- Using Your Twitter API Key
- Using Python to Grab Twitter User Data