

Programming Assignment

Problem Statement:

Benzinga is a stock market site and is looking to understand which ticker symbols are used the most and which drive the most traffic. The primary goal of this assignment is to combine the two datasets, tagged stock tickers and page views, and deliver a report that showcases key metrics and insights.

The report will include include visualizations of the following statistics derived from the two data sets:

- Overall number of different tickers per author
- Authors with the most overall pageviews
- Tickers with the most pageviews over time
- Most engaging tickers (most time spent viewing articles with that ticker)

Additional insights beyond these may also be provided. The report should contain some readable text explaining the results as well as graphical visualizations to illustrate the results.

The report should be generated within a Python 3 Jupyter Notebook. The output format can be contained within the notebook file, and the notebook file will be the deliverable. Any other publicly available libraries may be used.

The data is supplied in 2 files in the same directory as the notebook, posts.csv and tickers.csv. posts.csv contains all of the pageview statistics about each article, as well as the ID, author and title, while tickers.csv contains the ID and a space-separated list of tickers for that story. They share the first column as the ID of the story to join the data sets by.