Arielle Metaxas

SI 206

Final Project

I used the following APIs:

1. Instagram

* Accessed recent media posts from the authenticating user. Note: due to the nature of Instagram app registration, authenticating users in sandbox mode may only access the 20 most recent posts. It is for this reason that only 20, instead of 100 interactions were gathered.
  + Info gathered: time of post, like count on post, and comment count on post

2. Tumblr

* Accessed 100 most recent posts from my tumblr. The function I created can be applicable to any tumblr domain with more than 100 posts.
  + Info gathered: time of post, type of post, and note count

3. IMDB

* Accessed every episode from a list of Tv Shows. Note: since the length of each tv show varies, sometimes more than 100 episodes are gathered, and sometimes less than 100 episodes are gathered.
  + Info gathered: Title, release date, number of episodes, and rating of each episode

4. NYT

* Accessed the 100 most recent NYT posts that include the specified query, in this case, ‘The University of Michigan’.
  + Info gathered: Post headline, post date, document type, word count, and url.

5. OpenTable

* Accessed 100 restaurants matching a query (city and price). Note: this is an unofficial OpenTable API <https://opentable.herokuapp.com>
  + Info gathered: Name, Address, and Phone Number

Visualizations I used:

1. Bar Chart comparing my daily/hourly engagement of Instagram and Tumblr for my most recent posts.

2. Word Cloud representing the most common words from the headlines accessed from the NYT API.

3. Scatter plot showing the relationship between the number of episodes and the average rating of episodes for a variety of TV shows.