CS 122 Project: Mapping Trends

Team Name: Twitter Trolls

Members: Pooja Barai, Wesley Hsu, Kunal Mahajan, Leah Wang

Project description:

We are planning to use live Twitter geo-related data on location and trends, and map our findings using location data onto graphical information system (GIS) to show a heatmap of the trend in the United States. We will be filtering live tweets by hashtag, and exporting it onto a GIS map that can be displayed on a website. The users will also be able to see how the trend spreads throughout the trending period, where the trend started, which areas followed it first, and whether this pattern is consistent for all hashtags trending at that moment. Our geographical information will come from two sources: the profile geo-tag and the tweet geo-tag. Sources can be overlayed (if data is sufficient) or displayed individually. The goal of our project is to broaden the information we can obtain from Twitter trends and potentially investigate the trending information other than the ones displayed to the users. This will be helpful for businesses investigating potential areas to expand into, or for politicians on which areas display a higher interest in certain issues.

The data sources:

- Twitter geographical data: https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location
- Twitter Trend data: https://developer.twitter.com/en/docs/trends/locations-with-trending-topics/overview
- GIS Display Packages: Geopandas, GDAL, Shapely

Tasks:

Task	Title	Description	Deadline
1	Data Scraping and Collecting	(1) Download map (+city to state/county listings) (2) Collect a set of preliminary data (Preliminary data could be analysed without geotagging, just using home cities mapping to a state level with dictionaries or something) (3) Decide the trends and hashtags that we are interested in, and also the boundaries of the map that we will be looking at	Week 5 Monday
2	Data Cleaning	(1) Filter the hashtag data relevant to our question/interest - not all of it would be relevant despite the hashtag (2) Clean the data so that it is formatted for the mapping	Week 6 Monday
3	Data Analysis	(1) Making a list of questions that we are interested in looking at using our findings - the relevance of the trends we are analysing and how the patterns in location/trend can be useful (2) Analysing the data to see if/how	Week 7 Monday

		to answer these questions	
4	Intermediate coding	(1) Debugging, making sure our pipeline works on different trends (2) Add complexity to our project if needed	Week 8 check in
5	Data Representation	(1) Making a user interface for collecting data (2) Picking out interesting data trends and using a specific visual representation style to present the data in a useful manner	Week 9 Friday
6	Making the project presentation	(1) Making the powerpoint presentation (2) Assigning slides and preparing for the presentation itself	Week 10 Monday