

Title: Trends In Library Usage

Team members: Taylore Prinsen, Peter Minwegen, Sky Johnson

Description: We intend to use a data set with weather statistics in Seattle as well as a data set with statistics related to library check outs. We'd like to explore whether or not people check out books more or less when the weather is rainy, sunny, etc. We would also like to see if there is a correlation between winter and summer months and the topics of books that people check out.

Prior Work: There's been a bit of work, [with this article](#) detailing the usage of the Seattle public library, however they remain to the specific Seattle Public Library dataset, and don't combine it with anything.

Datasets:

Seattle Checkouts by Title:

This dataset includes a monthly count of Seattle Public Library checkouts by title for physical and electronic items. The dataset begins with checkouts that occurred in April 2005.

<https://www.kaggle.com/city-of-seattle/seattle-checkouts-by-title>

We also plan to gather weather data to see if there's a significant increase within seattle By mixing two data sets, the <https://www.weather.gov/asos/> for seattle airports area, and <https://www.wunderground.com/history/daily/us/wa/seattle/KSEA/date/2019-5-26>

We may also be able to incorporate traffic data within the city using the Seattle city's API. Located here <https://data.seattle.gov/resource/38vd-gyvt.json>

Proposed Work: We will need to do a lot of data preprocessing and integration to make a usable data set to run our models on. Currently we will need to do a few things,

1. Data Integration on all the datasets to match them by date,
2. Perform a lot of data cleaning,
3. Reduce the data into specific aspects we are looking into
4. Then run some models to look for significant correlations within the dataset
- 5.

List of Tools: Python, most likely the *Pandas*, *Scikitlearn* packages

Evaluation: We can evaluate our results by following standard statistical practices like saying there's a 95% confidence that two aspects of the data set correlate.