**Open Ended Capstone Project Idea 1:**

**Integrating StockTwits with Sentiment Analysis for better Prediction of Stock Price Movement**

Sentiment Analysis is new way of machine learning to extract opinion orientation (positive, negative, neutral) from a text segment written for any product, organization, person or any other entity. Sentiment Analysis can be used to predict the mood of people that have impact on stock prices, therefore it can help in prediction of actual stock movement. In order to exploit the benefits of sentiment analysis in stock market industry

Stock Market Data to be fetched from IEX Cloud:

<https://iexcloud.io/core-data-catalog#?utm_source=intercom&utm_medium=email&utm_content=welcomedatacatalog&utm_campaign=onboarding>

**Few Articles:**

Equities Data

<https://appsource.microsoft.com/en-us/product/office/WA200001834?tab=Overview>

Combining with Social Media Data to check the effect on Stock Prices

<https://blogs.lse.ac.uk/businessreview/2018/09/19/microblogging-sentiment-and-stock-returns/>

**Capstone Project Idea 2:**

## Data Source

#### Street Parking Occupancy data (Processed data)

[Data Source Link](https://data.seattle.gov/Transportation/2019-Paid-Parking-Occupancy-Year-to-date-/qktt-2bsy)

<https://data.seattle.gov/Transportation/2020-Paid-Parking-Occupancy-Year-to-date-/wtpb-jp8d>

* Seattle.gov update the data with a week delay.
* Granularity of the data is by minute.
* About 290 million records in a year. (~45GB)
* Approx total of 1.4 billion records since 2012. (~320GB)

**Capstone Project Idea 3:**

# EDGAR Log File Data Set

The Division of Economic and Risk Analysis (DERA) has assembled information on internet search traffic for EDGAR filings through SEC.gov generally covering the period February 14, 2003 through June 30, 2017.  The data is intended to provide insight into the usage of publicly accessible EDGAR company filings in a simple but extensive manner.

**https://www.sec.gov/dera/data/edgar-log-file-data-set.html**