**Capstone 1 Proposals**

# Alternate data and it’s relation to a company’s stock price (and/or earnings surprise)

Q: Does alternate data have any explanatory power with regards to earnings surprise and/or stock price?

Questions:

* Do negative reviews for a new product extrapolate out to earnings?
* Does the impact of negative press due to a one-off event have a lasting effect on stock price? Does this impact retail driven securities in a more drastic manner?
* Does a large earnings surprise extrapolate out to downstream industries?

Difficulties:

* Determining to what extent the qualitative data impacts price/earnings
* Autocorrelation in lagged earnings variable ie. Earnings surprise downstream could be more of a function of overall economic growth rather than a function of a specific industry/company.

Data Sources:

* Earnings Calendar - <http://www.earningscalendar.net/>
* Stock Price - <https://www.alphavantage.co/>

# Spotify’s top 200 daily streamed songs

Q: Is there a meaningful extrapolation of Spotify’s top 200 to provide insight on trends in music or specific artists? Can variables explain future Spotify stream numbers or stock price?

Questions:

* Do certain genres of music tend to be streamed more often than others?
* Does featuring another artist have any impact on streaming numbers?
* Does time of year impact music streaming figures?
* Are stream numbers a good proxy for stock price?

Difficulties:

* No API to pull top 200 streamed songs. Data requires some scraping.

Data Sources:

* + Spotify Charts - <https://spotifycharts.com/regional>
  + Last.fm API - <https://www.last.fm/api> (for meta data)
  + musiXmatch API - <https://rapidapi.com/musixmatch.com/api/musixmatch> (lyrics)

# Accidental drug related deaths in the State of Connecticut

Q: How have accidental drug related deaths changed in the State of Connecticut in the past 6 years? Are there variables that could potentially highlight future trends? (# of deaths, location, etc.)

Questions:

* Is there a correlation between the mixture of certain drugs and accidental death?
* Are certain age or ethnic groups more prone to accidental drug overdose?
* How does income or geographical location relate to drug overdoses?
* Do individuals have a higher propensity to overdose on certain days of the year?

Data Sources:

* State of Connecticut open source data - <https://catalog.data.gov/dataset/accidental-drug-related-deaths-january-2012-sept-2015>