Deliverable:

Reputation Monitoring Demo

•    Query [reddit.com](http://reddit.com) for mentions of a public facing company (For Example, Starbucks)

•    Transform and save data in a MySQL database

•    Determine the most common subjects discussed using natural language processing algorithms.

•    Plot the prevalence of these topics over time to see if there are any trends.

•    Set up recurring polling to detect mentions of the company that occur going forward to identify any emergent issues that need a response.

Assumptions

1. Sampling from Reddit is an adequate way to measure sentiment
2. Arrest controversy in April 12th 2018 would have an impact on customer sentiment
3. Most redditors are un U.S. and would comment on incident. Other scandals around the world would not impact a greatly
4. Customer sentiment would return to previous level after some months

Sentiment

Subjectivity

Polarity

Topic is arrest – LDA

All Comments

Arrest Related Comments

Synonyms for arrent

DATA

Comments from Reddit.com where there is any mention of starbucks in the discussion thread.

Who am I

Thank Don, Owen for giving me the opportunity to speak with you today.

Goal: Is it possible to gauge public sentiment about a company by looking at online discourse on Reddit.com

Why should you care?

Structured Data vs Unstructured data

Detect issues as they arise

Assumptions

Data

Reddit.com data from Google BigQuery

Analysis

Open source tools (SpaCy, NLTK, TextBlob)

Sentiment analysis

Examples of Sentiment Analysis

Negative

Positive

Timeline of Starbucks Controversies

April 12th

Synopsis:

Incident

Response

How long people are talking about it?

Did Sentiment change?

How many reddit users are in U.S. North America?

As of February 2018, Reddit had 542 million monthly visitors (234 million unique users), ranking as the #3 most visited website in U.S. and #6 in the world, according to [Alexa Internet](https://en.wikipedia.org/wiki/Alexa_Internet), with 57.4% of its user base coming from the [United States](https://en.wikipedia.org/wiki/United_States), followed by the [United Kingdom](https://en.wikipedia.org/wiki/United_Kingdom) at 7.5% and [Canada](https://en.wikipedia.org/wiki/Canada) at 6.3%.[[4]](https://en.wikipedia.org/wiki/Reddit#cite_note-4)

https://en.wikipedia.org/wiki/Reddit