# Leveraging Twitter to Predict COVID Caseload

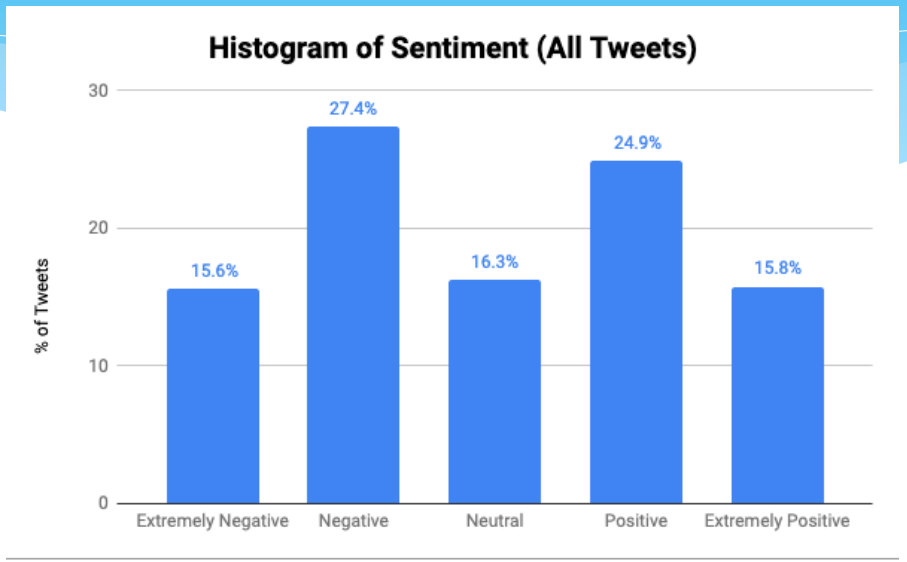
By Will Carnevale

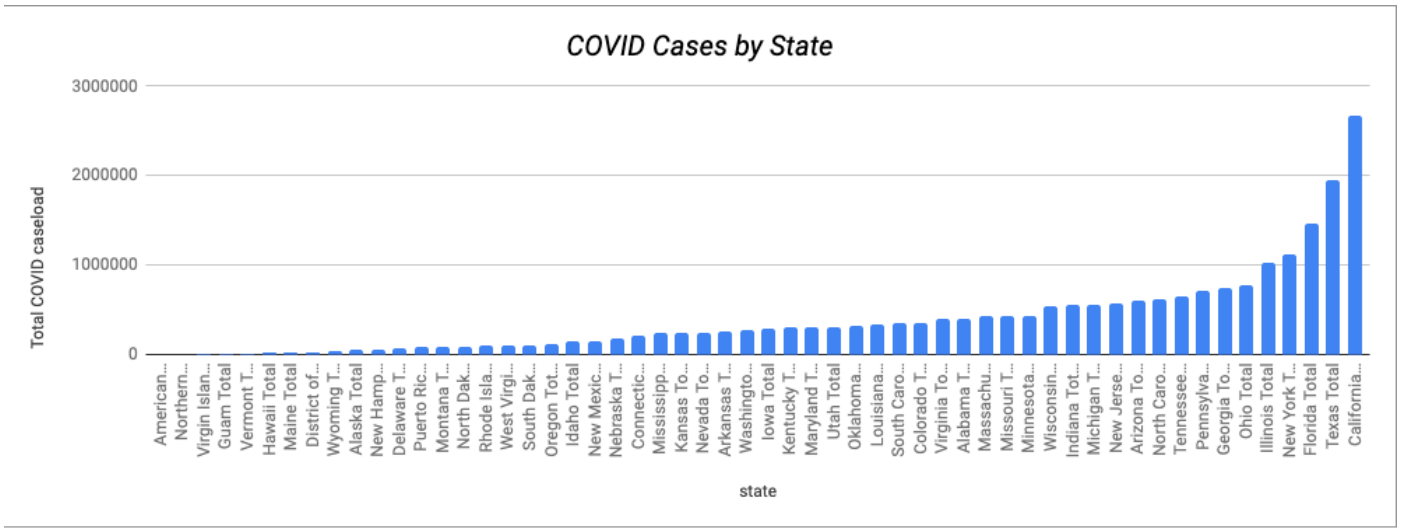
# Opportunity


* Twitter Sentiment: Tweets classified into (extremely) positive, (extremely) negative, and Neutral
* **Impact Hypothesis:** We can use twitter sentiment as a predictive engine to decrease the time taken to assess proper allocation of COVID aid
    i. decreases the likelihood of large-scale outbreaks occurring

# Methodology

* Twitter Data: Kaggle
  i. Tweet, Location, date, sentiment
* COVID Data: Kaggle
* Cleaning/EDA: Google Sheets
* Data Visualization: Tableau

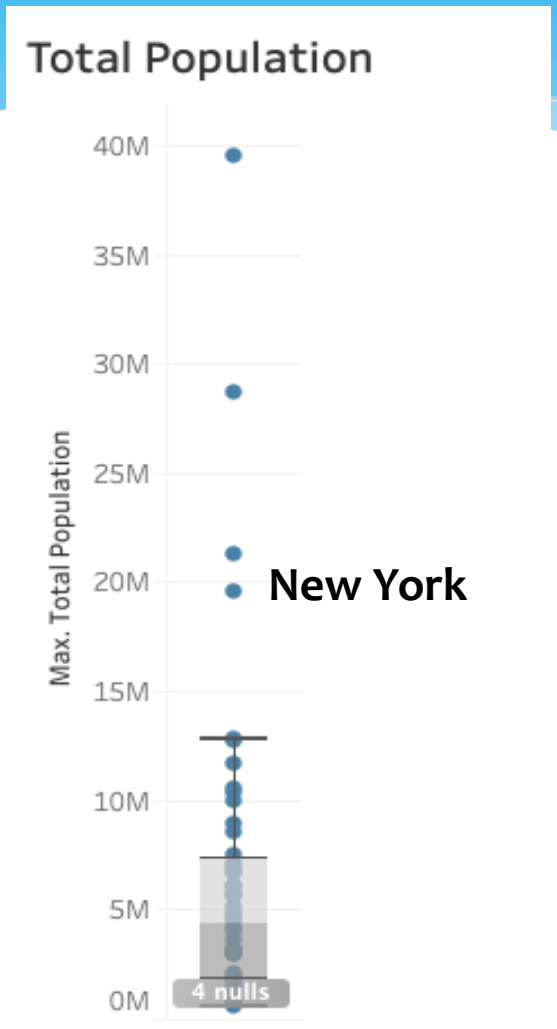


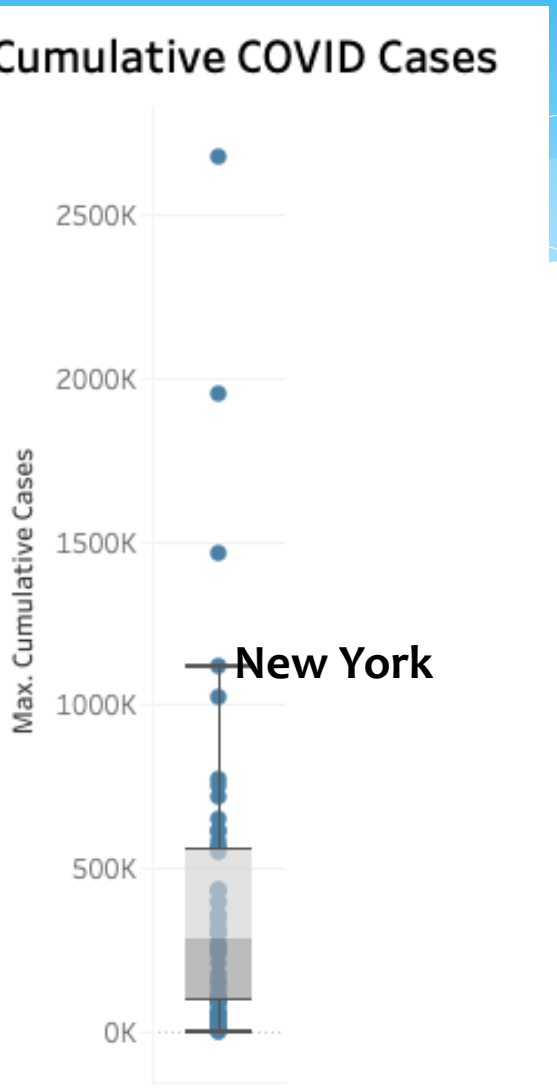



# Proposed Solution Path

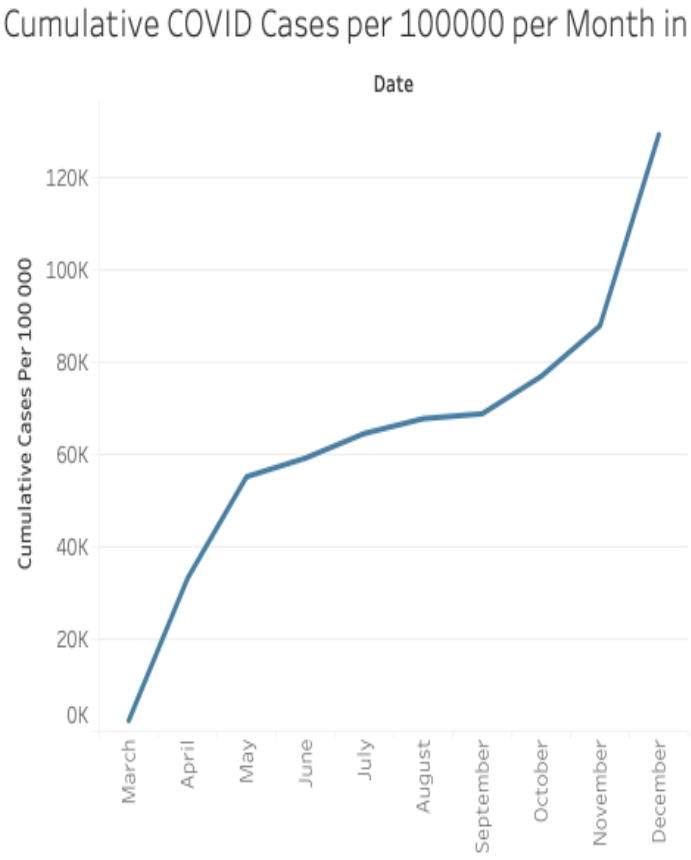
* Build an Natural Language Processing (NLP) model for topic analysis of COVID tweets
* Combine NLP model with Time series analysis of COVID caseload data

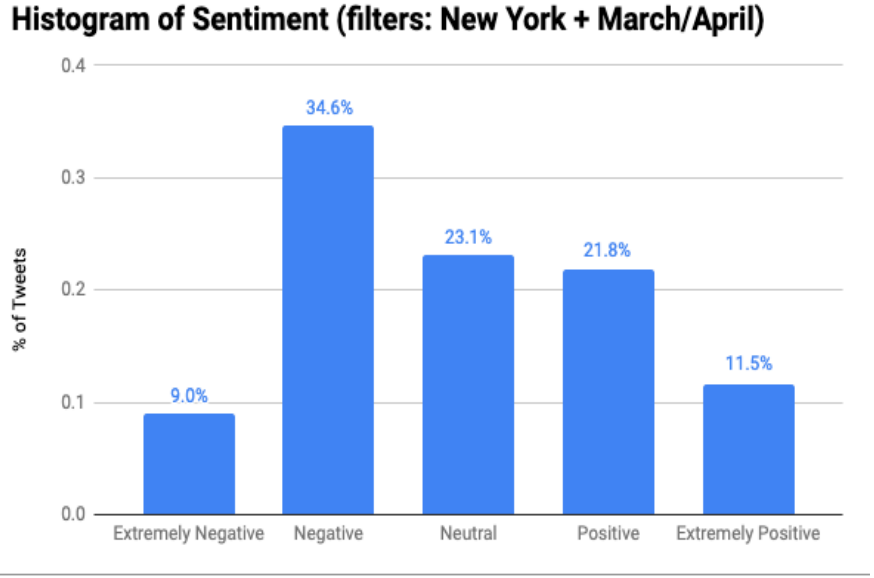
# Case Study: New York



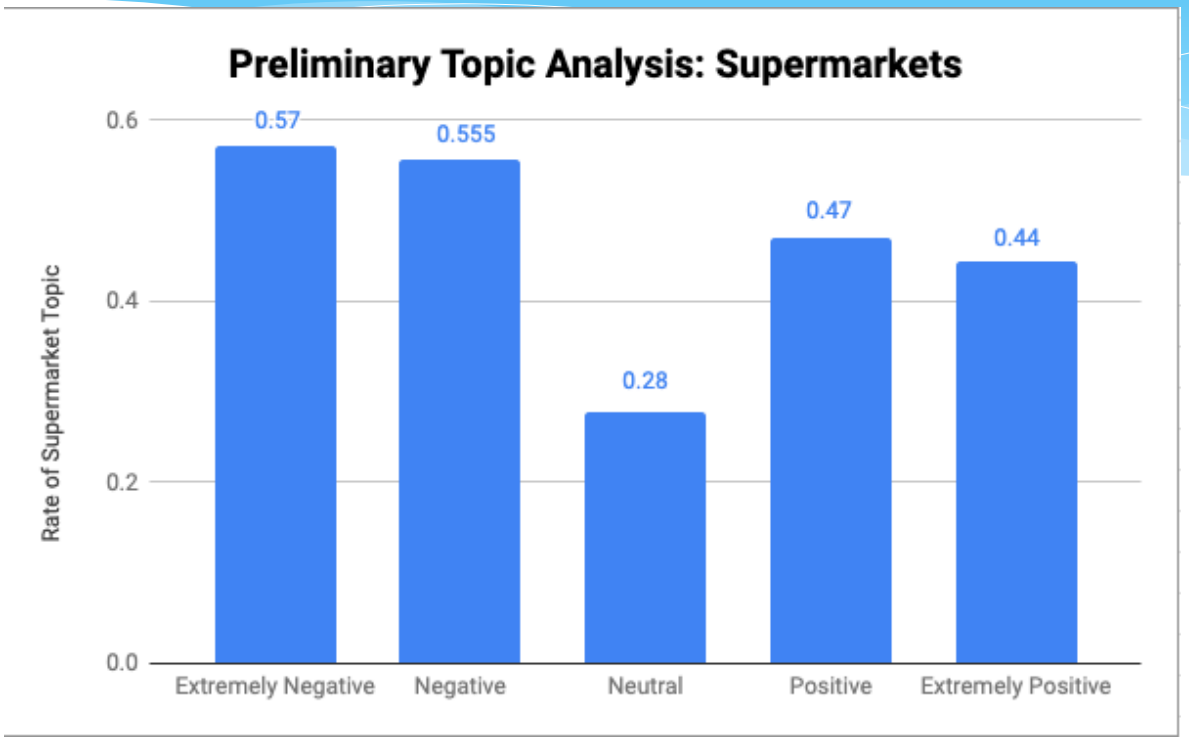



# Case Study: New York (2)



# Case Study: New York (3)




* Preliminary topic analysis: Grocery store and supermarket hoarding/scramble for goods

# Summary

* Case Study conclusions:

  1. New York outperformed in reducing COVID cases relative to population

  2. New Yorks twitter sentiment tracks more negative during sharp increase in COVID cases

  3. Tweets with Negative sentiment have a higher rate of mentioning supermarkets

* Future Work:

  i. Collection of more comprehensive twitter data (with accurate location data)

  ii. Control for outside biases such as mainstream media coverage and bot accounts