

# Weather & Twitter Sentiment

Projects in Data Science: Python



# INTRO

- Can we predict Twitter sentiment based on the weather forecast?
- For given regions, combine the results of sentiment analysis of tweets with weather data
- Train a model to predict the sentiment of a tweet given the weather at its location
- Create a weekly 'emotion forecast' for Twitter based on the weather forecast

# DATA

## TWEETS

- Use tweepy streamer to stream tweets from specified locations – running on Google Cloud
- [how many tweets/cities?]

## SENTIMENT ANALYSIS

- 18500-word list with sentiment values (-1 – 1)

## WEATHER

- Get weather data for specific weather stations from NOAA (<ftp.ncdc.noaa.gov>), corresponding with specified locations for Tweets
- Includes temperature, windspeed, cloud coverage, precipitation

# METHODOLOGY

1. Stream in [] tweets for [] locations – write to json
2. Calculate sentiment score for each tweet using wordlist – add to json
3. Get weather data for each tweet for closest weather station/closest time – add to json
4. Train model on collected data to predict positive or negative sentiment for each tweet based on weather features, using Random Forest Classifier [?]
5. Use trained model to start predicting sentiment in regions given weather forecast

# PRELIMINARY RESULTS

# CONCLUSION & NEXT STEPS

QUESTIONS?