**Trump Tweet and Market Analysis project documentation**

**Section 1 Tweet Data processing**

**About reading the tweets on demand**

1. Authentication

TO read the tweets on demand, we need to sign up a tweeter account, create the token and secret for the tweeter API for later used for reading the tweets on demand.

4 piece of information are needed.

# Consumer keys and access tokens, used for OAuth

CONSUMER\_KEY = '7GrVM16QYpKYu6GO11lntwTCR'

CONSUMER\_SECRET = 'tMsYF2wOW5l7oSDGLNDbg0Fsa4MfxQKQwItBmjYyoW9Q7Ku0R0'

# Access:

ACCESS\_TOKEN = '1184651600566067201-s6Cu1jM9YcIY9HMS8PpkQtMBBbeTQL'

ACCESS\_SECRET = 'roCUtyk7072aDQ1W09WPFKNHb2tZX9sdVwXlDhH4bjtgg'

1. Package use for reading

Tweetpy

By default, Twitter API allow only 200 tweets per page to be read, we need to page down to read all the tweets

**About tweet data filtering and aggregation**

1. Filtering 1

Data will be filtered to use only those starting from Trump in office

1. Filtering 2

Data will be filtered to discard those marked as retweet

1. Aggregation

Data will be aggregate per day by summing up the retweet count, favorite count, etc.

For later sentiment and trade tag scores, each tweet is evaluate the score first then aggregate by day

**About tweet sentiment/polarity processing and scoring**

1. VADER Sentiment analysis

Using VADER, we score and record the sentiment/polarity for each tweet and later we aggregate into day numbers.

>>   
VADER Sentiment Analysis. VADER (Valence Aware Dictionary and sEntiment Reasoner) is a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media, and works well on texts from other domains.

<<

1. Reference source

<https://www.kaggle.com/nltkdata/vader-lexicon>

**About tweet content tagging**

1. initiatives

We are trying to tag the tweets into 4 main categories: foreign policy, domestic affairs/politics/partisan, trade/China trade, and other.

1. How tagging is done

Basic tagging using the key word/phrases we deem important to use

**About trade related tweet scoring**

1. initiatives

we try to score the trade especially related tweets so as to give a boosting or multiplier with those looked after keywords/key phrase and later used it to relate to the sentiment analysis and market impact analysis.

1. How scoring works

Scores will used the keyword/key phrasing occurrence multiply a rank score and later aggregate into daily score.

**Section 2 SPX Data processing**

1. SPX data will be read from yahoo finance page and the format of the data will look like this.

