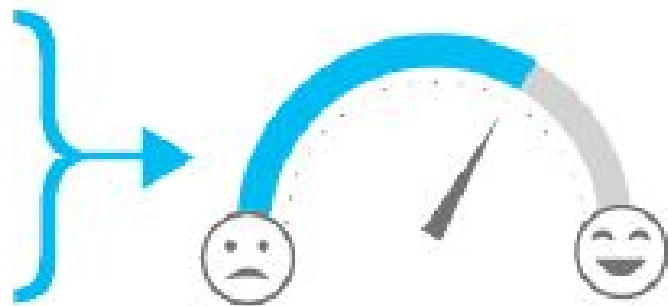


# Bitcoin Bias

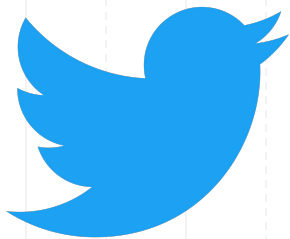


Robert Deng  
Andrew Mamroth  
Ian Bettinger

# The Goal

The goal of this project is to use existing API's to gather data on various cryptocurrencies and twitter data, channel it to a database and from there build a trading strategy.





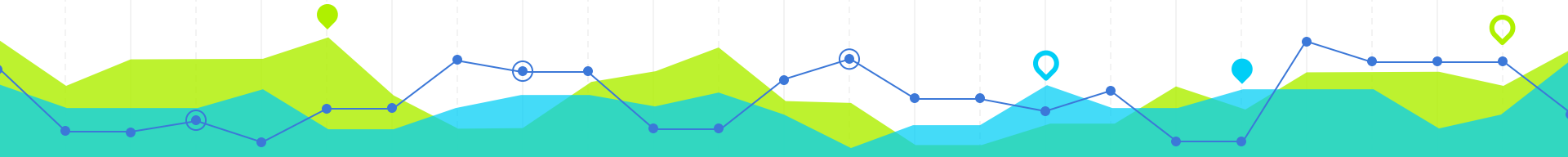
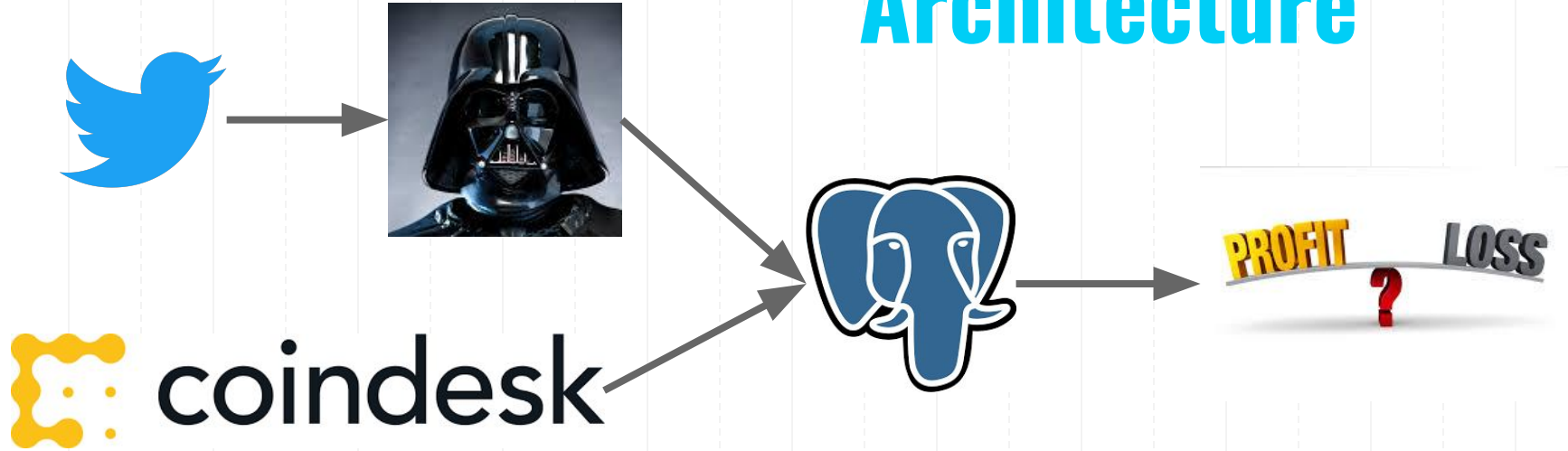
# The Steel Thread



The steel thread of the project is to gather data from coinmarketcap.com and twitter, place that data in a Postgres database, and output a trading strategy.



# Architecture



# Tweepy

- Tweepy run continuously using filter for 'btc' and 'bitcoin' in tweets
- Code adjusted to prevent 420 Rate limit stoppages.

Date	Count
11-Dec	1,290,158
12-Dec	1,874,917
13-Dec	1,201,843
14-Dec	1,513,459
15-Dec	97,761
<b>Total</b>	<b>5,978,138</b>



# Coin Market Cap

Hourly ticker data pulled via cron, 6am - 6pm M-F

```
[crypto=# select * from ticker_data_12_9 limit 10;
```

index	last_updated	id	symbol	price_usd	available_supply	total_supply
0	2017-12-11 05:59:14	bitcoin	BTC	16590.0	16733500.0	16733500.0
1	2017-12-11 05:59:13	ethereum	ETH	463.026	96265981.0	96265981.0
2	2017-12-11 05:59:29	bitcoin-cash	BCH	1398.15	16849013.0	16849013.0
3	2017-12-11 05:59:25	iota	MIOTA	4.30883	2779530283.0	2779530283.0
4	2017-12-11 05:59:01	ripple	XRP	0.245261	38739144847.0	99993093880.0
5	2017-12-11 05:59:01	litecoin	LTC	161.909	54248158.0	54248158.0
6	2017-12-11 05:59:02	dash	DASH	719.419	7746051.0	7746051.0
7	2017-12-11 05:59:36	bitcoin-gold	BTG	264.747	16699774.0	16799774.0
8	2017-12-11 05:59:04	monero	XMR	257.679	15449232.0	15449232.0
9	2017-12-11 05:59:10	nem	XEM	0.41862	8999999999.0	8999999999.0



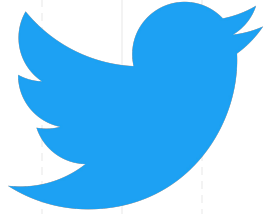
# Vader Sentiment Analysis

VADER - Valence Aware Dictionary and sEntiment Reasoner

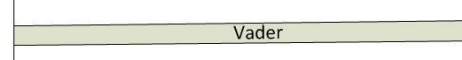
- Tunes for positive and negative sentiment (polarity) as well as the intensity (valence) for social media text
- Range of  $[-1, 1]$  for compound sentiment
- Takes in punctuation and slang into account



# ERD



tweets	
tweet	
created_at	
lang	
coordinates	
user_time_zone	
user_id	
user_name	



tweet_sentiment	
FK1	created_at_hr
	sentiment
	price_usd
	index



coindesk



ticker_data_12_9	
PK	<u>index</u>
last_updated	
id	
symbol	
price_usd	
available_supply	
total_supply	

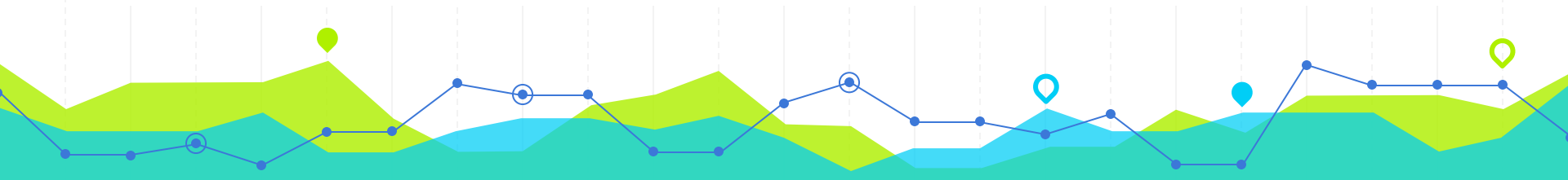
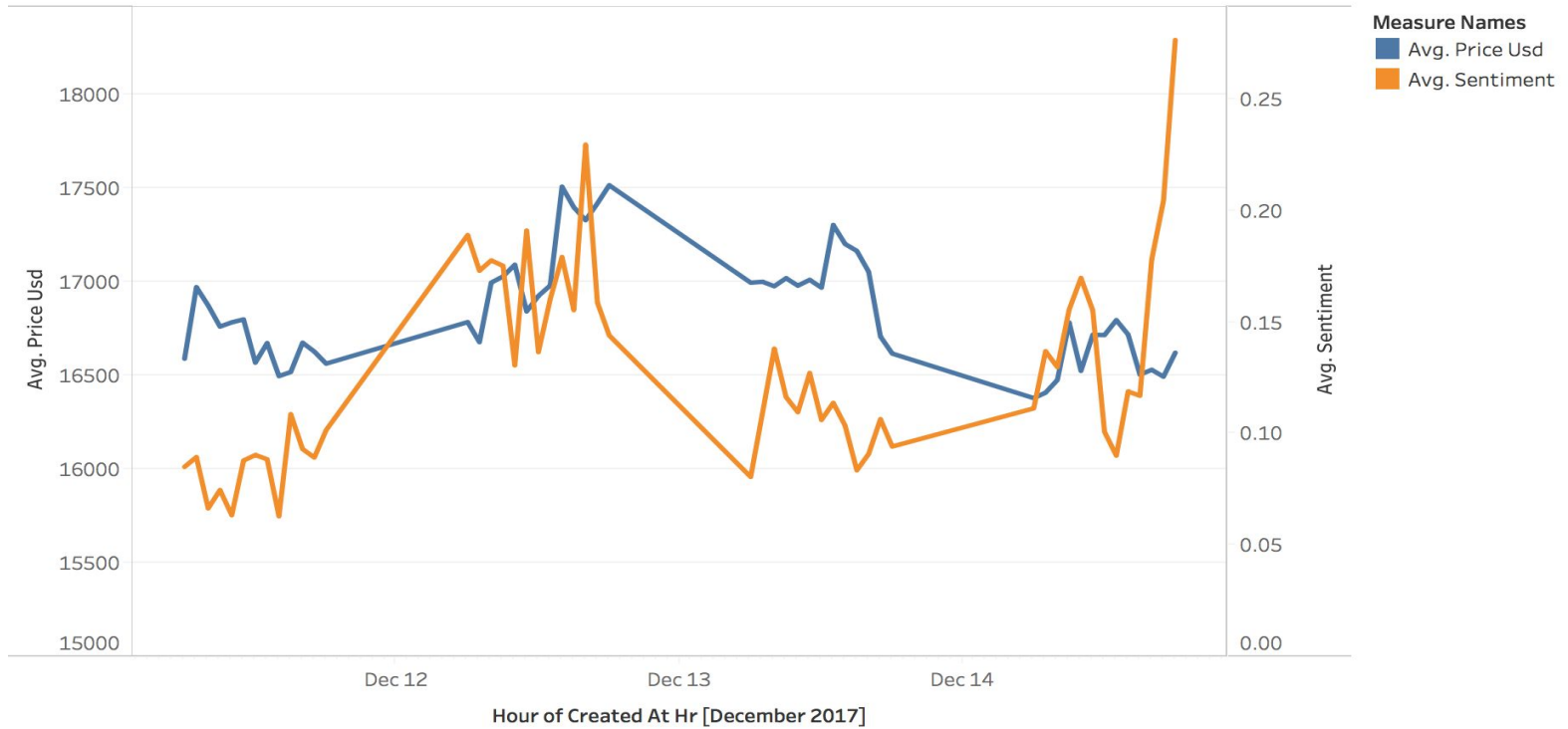


ticker_btc (view)	
	update_hour
	index
	last_updated
	id
	symbol
	price_usd
	available_supply
	total_supply





## Trending Sentiment vs Price



# A Simple Betting Strategy

Starting values were set at 500 in cash and 500 in bitcoin and adjusted each hour based on average sentiment.

```
w205@ip-172-31-81-83: ~/w205_fall_2017_groupProject
2017-12-13 06:00:00 Cash: 749.78 Bitcoin: 276.87, Total: 1026.65
2017-12-13 07:00:00 Cash: 483.32 Bitcoin: 543.41, Total: 1026.72
2017-12-13 08:00:00 Cash: 355.66 Bitcoin: 670.31, Total: 1025.97
2017-12-13 09:00:00 Cash: 461.11 Bitcoin: 566.55, Total: 1027.66
2017-12-13 10:00:00 Cash: 493.91 Bitcoin: 532.41, Total: 1026.33
2017-12-13 11:00:00 Cash: 414.79 Bitcoin: 612.52, Total: 1027.31
2017-12-13 12:00:00 Cash: 516.28 Bitcoin: 509.52, Total: 1025.81
2017-12-13 13:00:00 Cash: 478.96 Bitcoin: 556.89, Total: 1035.85
2017-12-13 14:00:00 Cash: 528.76 Bitcoin: 503.83, Total: 1032.59
2017-12-13 15:00:00 Cash: 626.28 Bitcoin: 405.22, Total: 1031.50
2017-12-13 16:00:00 Cash: 571.29 Bitcoin: 457.58, Total: 1028.88
2017-12-13 17:00:00 Cash: 473.37 Bitcoin: 546.23, Total: 1019.60
2017-12-13 18:00:00 Cash: 536.25 Bitcoin: 480.37, Total: 1016.62
2017-12-14 06:00:00 Cash: 438.88 Bitcoin: 570.87, Total: 1009.75
2017-12-14 07:00:00 Cash: 337.82 Bitcoin: 672.94, Total: 1010.76
2017-12-14 08:00:00 Cash: 372.82 Bitcoin: 640.71, Total: 1013.53
2017-12-14 09:00:00 Cash: 298.80 Bitcoin: 726.68, Total: 1025.48
2017-12-14 10:00:00 Cash: 271.53 Bitcoin: 742.77, Total: 1014.30
2017-12-14 11:00:00 Cash: 334.99 Bitcoin: 687.96, Total: 1022.95
2017-12-14 12:00:00 Cash: 576.54 Bitcoin: 446.34, Total: 1022.89
2017-12-14 13:00:00 Cash: 623.66 Bitcoin: 401.35, Total: 1025.02
2017-12-14 14:00:00 Cash: 424.66 Bitcoin: 598.52, Total: 1023.18
2017-12-14 15:00:00 Cash: 434.27 Bitcoin: 581.24, Total: 1015.51
2017-12-14 16:00:00 Cash: 207.65 Bitcoin: 808.77, Total: 1016.43
2017-12-14 17:00:00 Cash: 176.80 Bitcoin: 837.83, Total: 1014.63
2017-12-14 18:00:00 Cash: 114.52 Bitcoin: 906.59, Total: 1021.11
Bitcoin benchmark: 1001.89873418
[w205@ip-172-31-81-83 w205_fall_2017_groupProject]$
```

# The Three V's

## Velocity

- ❑ Coinmarket cap ~ 1.4k rows per hour, 12k per day
- ❑ Tweepy ~1M rows per day (BTC)

## Volume

- ❑ ~6M rows of twitter data
- ❑ ~70K rows of ticker data
- ❑ 1.5+ gigs of data in 1 week

## Variety

- ❑ Vader sentiment, Twitter text, Ticker time series

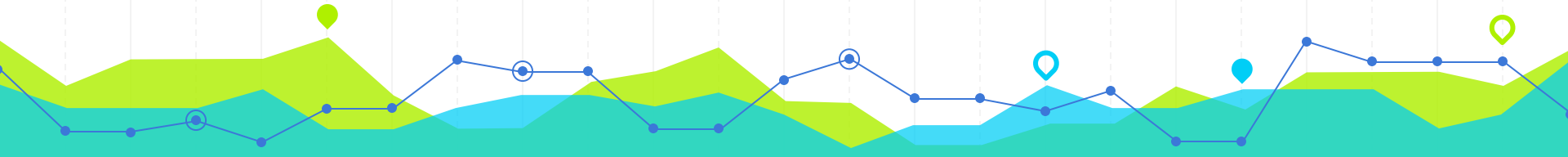


# Scale Up, Scale Out

The scalability of this project lies in both the size of the database and the number of trade strategies that will be output.

Initially, the intent to only hold data for a week to minimize the size of the database required. To scale it out, that data could be held (theoretically) forever.

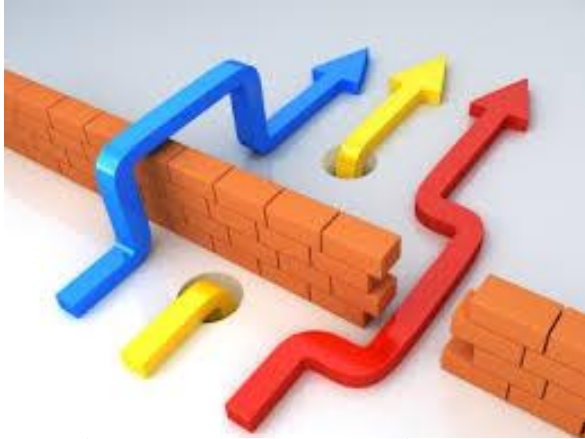
Additionally, the initial intent for trade strategies is the build one simply strategy. Scaling this up would involve build various trade strategies.



# Scale Up, Scale Out

Component	Demo	Scale out
Storage Engine	Postgres	Redshift Postgres
Processing Engine	Python/cron	Storm
Machine Learning	Python	Spark ML
Visualization	Tableau	Tableau

# Roadblocks



- API Pipelines to Database
- Cron Scheduling Cadence
- Computation at Scale (Vader)
- Postgres - Tableau Integration
- Large tweets table slowed down performance

# References

<https://coinmarketcap.com/api/>

<http://www.tweepy.org/>

<https://www.postgresql.org/docs/>

<http://initd.org/psycopg/>

<https://pypi.python.org/pypi/vaderSentiment>

