Financial Risk Assessment

Team : John Mitton, Kamran Madatov, Bao Thai
Mentor: Dr. Ted Ahn, Dr. Goldwasser

❑ General Overview

- Project Goal
- What is Risk Assessment
- Application
- Current Timeline and Future Planning
- Functional Accomplishments
  - Monte Carlo (Concept, Process, Results and Analysis)
  - Sentiment Scoring (Concept, Process, Results and Analysis)
  - Database
  - Front-End (login, registration, stock portfolio, stock search/data api, database)
- Improvements
- Project Overview

❑ Our Goal

- The objective of this project is to **help investors make appropriate decisions** with their portfolio to reduce loss when holding, buying, and selling investments.

- A **mobile application** will be created to **track listed investments** and **inform the user** based on news, history of stock prices, trend volume, social media, and market stock exchange.

- By using **statistical data** and **probability calculation**, we would like to ensure accuracy of estimates and allow users to be more informed on changes with a stock.

❑ Financial Risk Assessment

- To estimate and predict changes with particular investments due to statistical data or news information that would be essential to the outcome of investments
- **Variable at Risk**
  - Statistical method to measure and quantify level of risk of particular investment opportunities

❑ VaR Application

- Determine the probability and rate of a stock’s growth for a certain time period
- Manage an aggregation of portfolio performance for an investment firm
- Can also be applied across different markets besides stocks (real estates, etc.)

❑ Functional Accomplishment

- Monte Carlo Simulation
  - Using Historical Data

Shifted to Apache Spark

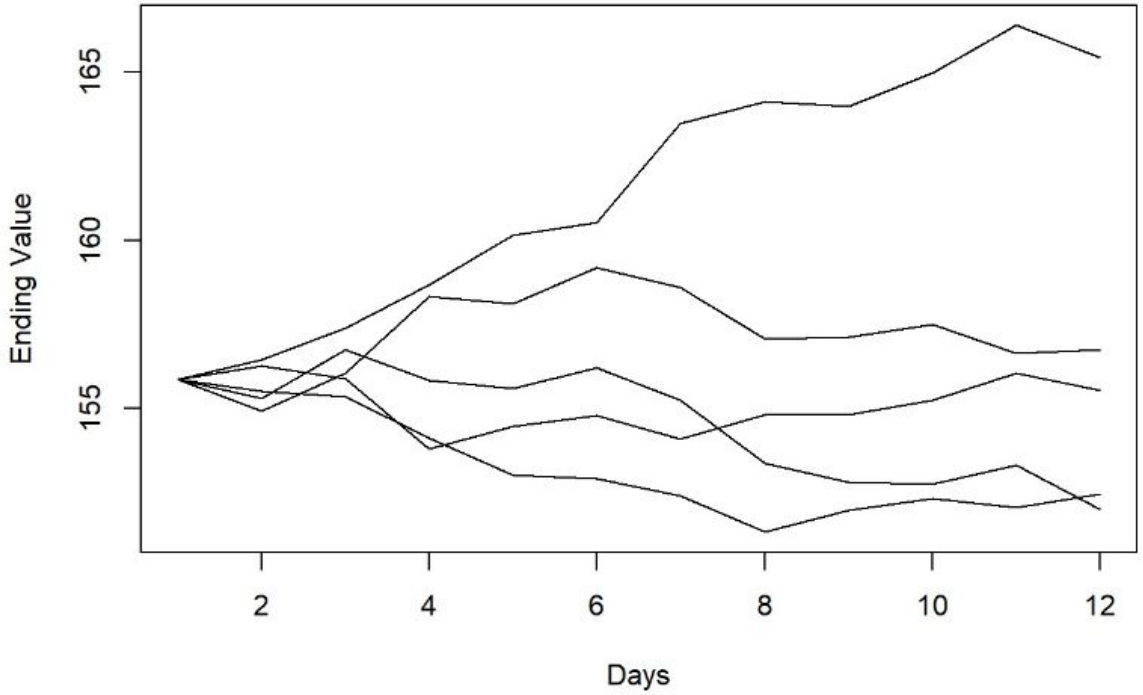
- Web Crawler
  - Article Parsings (Bloomberg, Motley Fool, etc)
- Sentiment Score
  - Loughran Mcdonald Dictionary

Shifted to Front End, Automated Script Ready

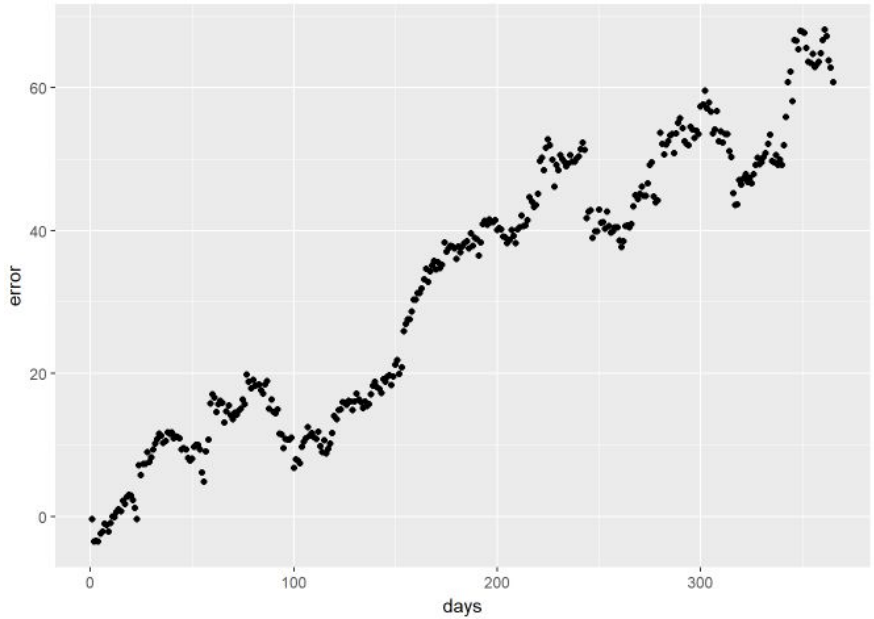
❑ Monte Carlo Simulation

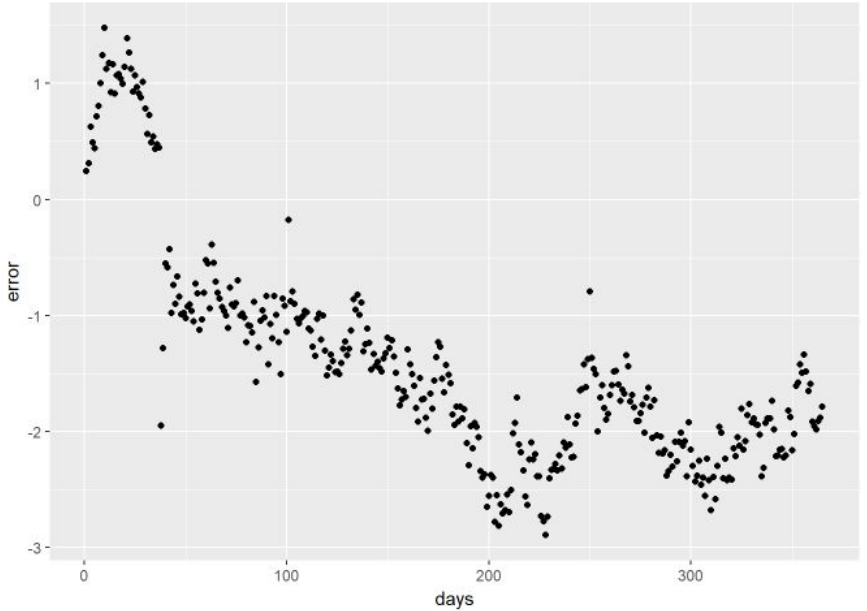
- Generate random samples that follow the problem constraints
- Find the mean of the samples
- This mean can be used to **predict values** and **determine risks**

❑ Monte Carlo Simulation Model



❑ Monte Carlo Simulation Testing





Amount of error over one year of tested data using ten years worth of historical data.

❑ Monte Carlo Simulation Accuracy

- Most accurate to one week
  - Error Values for 1 week: AAPL: -2.3 CNC: -0.34 JNP: 0.52 TSLA: 14.9
- Simulating for a year produces large mean error
  - Error Values for 1 year: AAPL: 31.9 CNC: 0.71 JNP: -1.4 TSLA: 64.5
- Better on less volatile stocks
  - E.g. ones that do not rely on the next big thing

❑ Spark Testing

Time to run one simulation on Hopper

```
[mittonjw@hopper csvget]$ time python monteCarlo.py
(229.60866234094658, 232.37271069552997, 235.13675905011337)

real    0m1.839s
user    0m2.271s
sys     0m3.609s
[mittonjw@hopper csvget]$ time python monteCarlo.py
(127.36229985967023, 129.10873267725515, 130.85516549484007)

real    0m1.440s
user    0m2.000s
sys     0m3.612s
[mittonjw@hopper csvget]$ time python monteCarlo.py
(9.2076127787427762, 9.6533378882849412, 10.099062997827106)

real    0m1.825s
user    0m2.259s
sys     0m3.540s
[mittonjw@hopper csvget]$ time python monteCarlo.py
(510.1790187071482, 520.28121949368722, 530.38342028022623)

real    0m1.611s
user    0m2.079s
sys     0m3.608s
```

Time to run one simulation in Spark on an AWS cluster

```
[hadoop@ip-172-31-41-182 ~]$ time spark-submit monteCarlo.py
350.341495613

real    0m2.689s
user    0m3.380s
sys     0m0.220s
[hadoop@ip-172-31-41-182 ~]$ time spark-submit monteCarlo.py
234.367435337

real    0m3.171s
user    0m3.712s
sys     0m0.232s
[hadoop@ip-172-31-41-182 ~]$ time spark-submit monteCarlo.py
39.0466213559

real    0m2.812s
user    0m3.724s
sys     0m0.228s
[hadoop@ip-172-31-41-182 ~]$ time spark-submit monteCarlo.py
1026.76373186

real    0m2.900s
user    0m4.016s
sys     0m0.196s
```

❑ Loughran McDonald Dictionary

The dictionary provides a means of determining which tokens (collections of characters) are actual words, which is important for consistency in word counts.

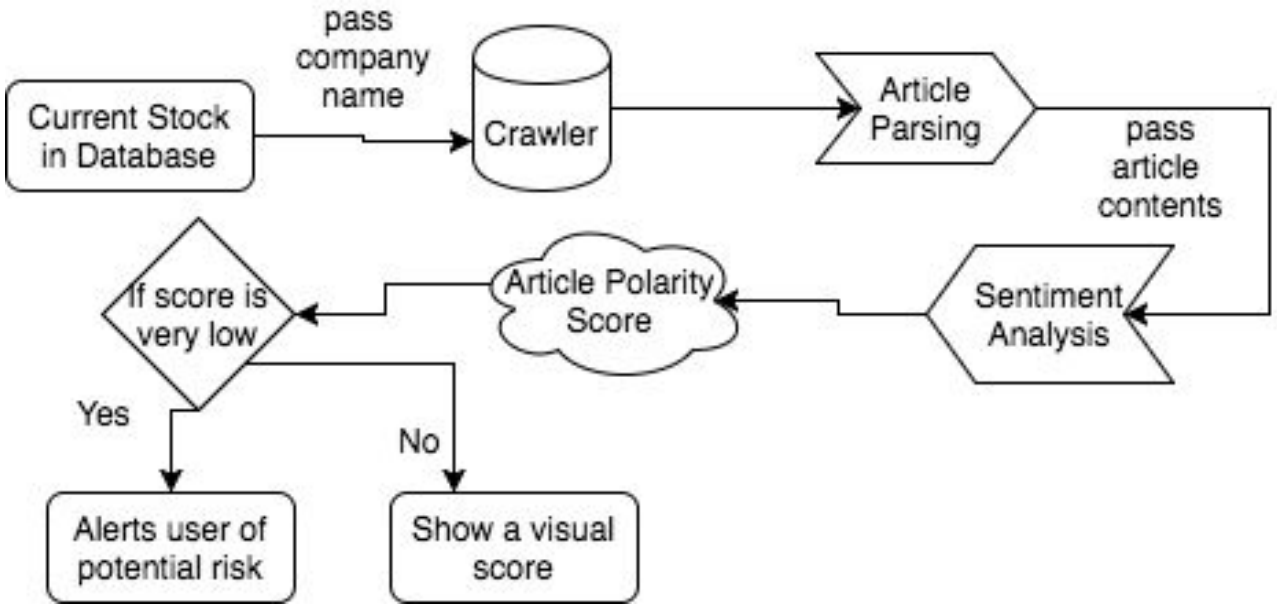
- statistics for word frequencies in all 10-Ks from 1994-2014 (including 10-X variants)
- 85,000+ proper nouns and abbreviations (negative, positive, uncertainty, litigious, modal, constraining)
- 354 Positive
- 2329 Negative

❑ Sentiment Scoring

1. Parse through the article
2. Count total number of actual words (LM dictionary)
3. Count Number of positive of words and negative
4. % Positive Words = Total Positive Words / Total Actual Words
5. % Negative Words = Total Negative Words / Total Actual Words

Sentiment Score: %Positive - %Negative

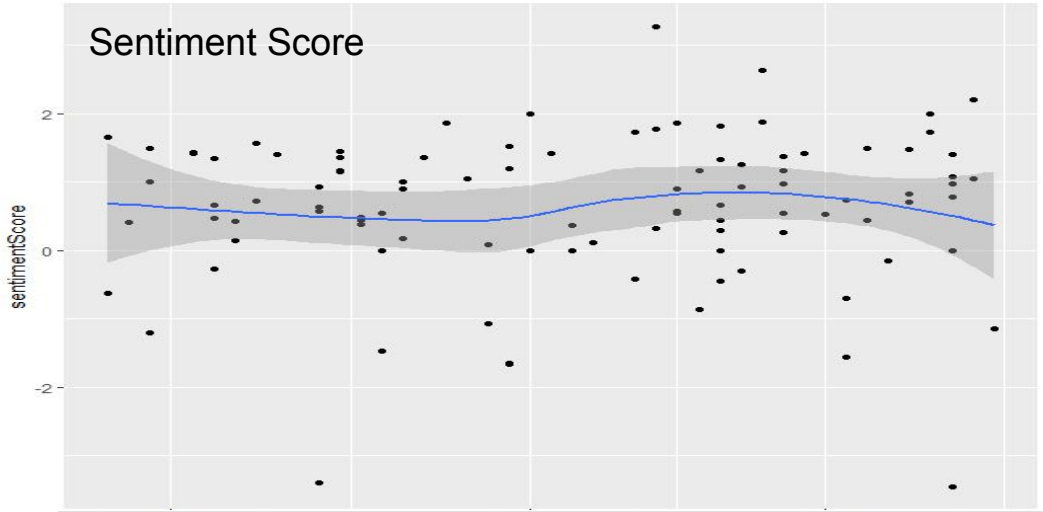
❑ News Article Web Crawling - Process

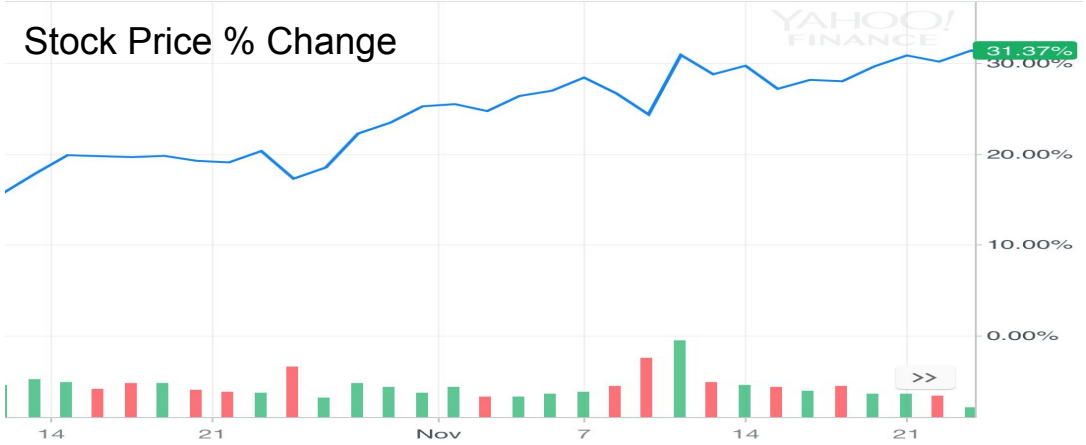


❑ Webcrawler - InnoDB

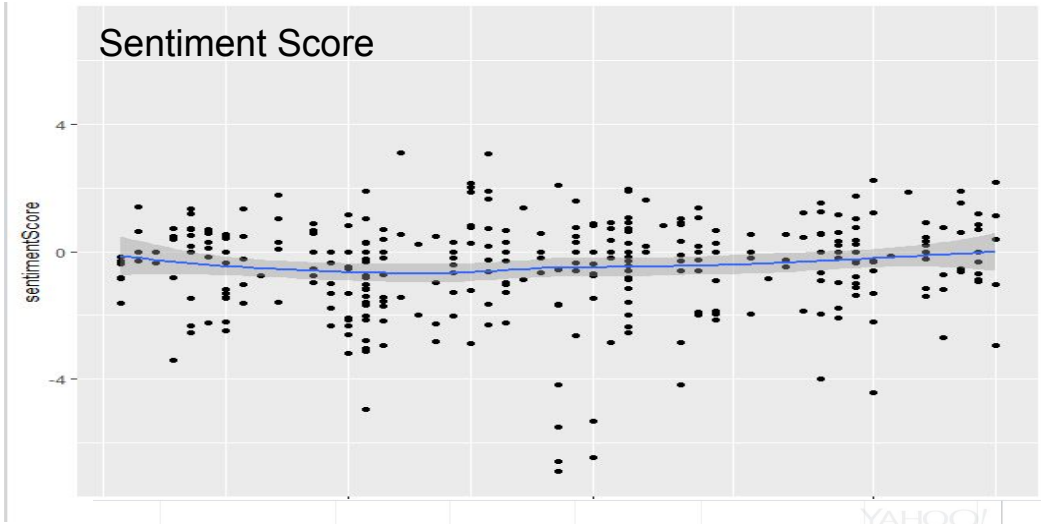
	id	articleURL	company	domain	date	sentScore
Edit Copy Delete	1	http://www.fool.com/investing/2018/02/28/2-best-dr...	Nvidia	fool.com	2018-02-28	0.77
Edit Copy Delete	2	http://www.fool.com/investing/2018/02/28/is-advanc...	Nvidia	fool.com	2018-02-28	0.64
Edit Copy Delete	3	http://www.fool.com/investing/2018/02/27/3-stocks-...	Nvidia	fool.com	2018-02-27	2.24
Edit Copy Delete	4	http://www.fool.com/investing/2018/02/27/wall-stre...	Nvidia	fool.com	2018-02-27	0.28
Edit Copy Delete	5	http://www.fool.com/investing/2018/02/26/why-is-th...	Nvidia	fool.com	2018-02-26	0.13
Edit Copy Delete	6	http://www.fool.com/investing/2018/02/23/amazon-an...	Nvidia	fool.com	2018-02-23	-0.32
Edit Copy Delete	7	http://www.fool.com/investing/2018/02/23/dont-wast...	Nvidia	fool.com	2018-02-23	0.32
Edit Copy Delete	8	http://www.fool.com/investing/2018/02/22/3-stocks-...	Nvidia	fool.com	2018-02-22	3.17
Edit Copy Delete	9	http://www.fool.com/investing/2018/02/22/forget-cr...	Nvidia	fool.com	2018-02-22	0.50
Edit Copy Delete	10	http://www.fool.com/investing/2018/02/21/3-stocks-...	Nvidia	fool.com	2018-02-21	1.38
Edit Copy Delete	11	http://www.fool.com/investing/2018/02/28/why-i-wil...	Apple	fool.com	2018-02-28	-3.09
Edit Copy Delete	12	http://www.fool.com/investing/2018/02/28/apple-inc...	Apple	fool.com	2018-02-28	2.39
Edit Copy Delete	13	http://www.fool.com/investing/2018/02/28/4-reasons...	Apple	fool.com	2018-02-28	1.30
Edit Copy Delete	14	http://www.fool.com/investing/2018/02/28/you-gotta...	Apple	fool.com	2018-02-28	0.74

❑ NVDA


Sentiment Score
sentimentScore
2
0
-2


Stock Price % Change
YAHOO! FINANCE
31.37%
30.00%
20.00%
10.00%
0.00%
>>
14
21
Nov
7
14
21

❑ TSLA


Sentiment Score
sentimentScore
4
0
-4


Stock Price % Change
YAHOO! FINANCE
5.00%
0.00%
-5.00%
-10.00%
-14.42%
-20.00%
21
Nov
7
14
21
Dec
7

Front End - Home Page

FinanceTeam

New Member? Existing Member?

Let's Get Started

Enter Desired Username:

Login

Enter Desired Password:

Password

Confirm Password:

Confirm Password

Sign Me Up!

Welcome Back!

Enter Username:

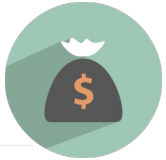
Username

Enter Password:

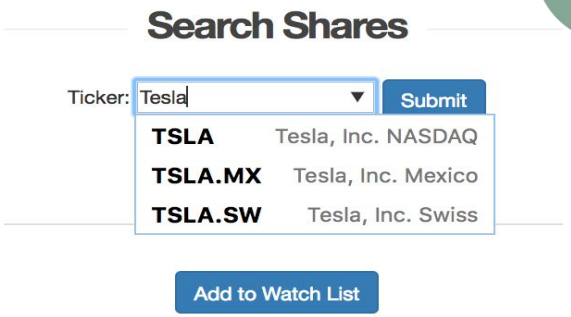
Password

Let's Go!

Front End - Stock Search Page



FinanceTeam



Search Shares

Ticker: TSLA Submit

Date	Symbol	Company Name	Price	Change	Change Percent
Fri Mar 02 2018 14:22:54 GMT-0600 (CST)	TSLA	Tesla Inc.	$331.99	$1.06	0.0032%

Add to Watch List

Front End - Profile & Navigation Page

☰NAVIGATION

FinanceTeam

Welcome, 7

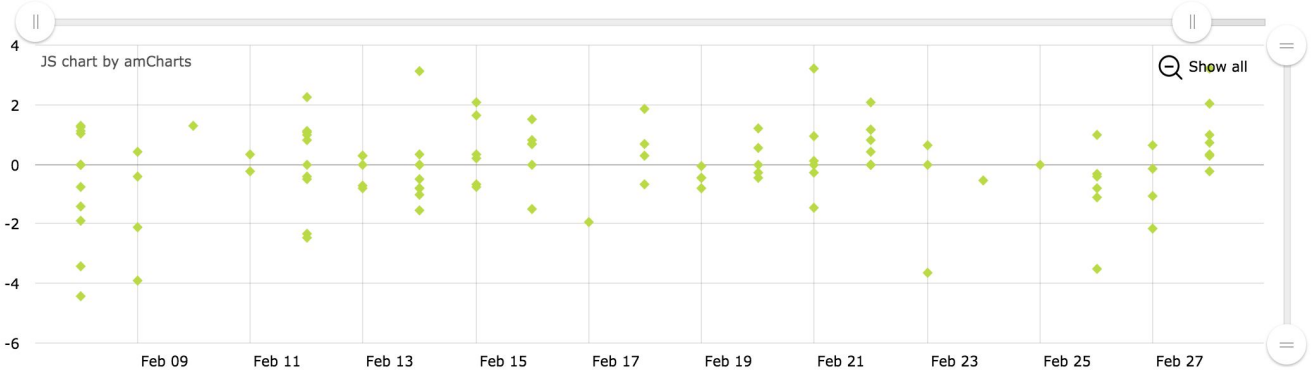
Watch List

Trans ID	Company Name	Symbol	Sentiment Analysis

Front End - Stock Portfolio Tool

Watch List

Trans ID	Company Name	Symbol	Sentiment Analysis		
6	tesla	TSLA	0.002833	Remove	View Details
8	apple	AAPL	-0.186917	Remove	View Details
9	nvidia	NVDA	0.476333	Remove	View Details



Log Out

❑ Interpretation

- Depending on trend
  - Positive linear slope[1, >]
    - High risk
    - High chance for a short term growth (Short term long buys)
  - Near zero risk sentiment score (low risk) [-1, 1]
    - Stable in growth
    - Long term investment opportunity?
  - Negative linear slope [<, -1]
    - High risk
    - High chance stock continue decreases (good for shorting)

❑ Results & Accuracy

- ❑ Qualitative Analysis
  - ❑ Data correlates with interpretation
  - ❑ Sentiment score relates to risk by absolute displacement from 0

Portfolio Risk Management

❑ Improvement Opportunities

- Crawler and Sentiment Scoring
  - Removing unrelated articles
  - More sources
  - Newer finance dictionary
- Big Data Management (Hadoop) and Cluster Computation(Apache Spark)
- Monte Carlo
  - Adapting the Monte Carlo model
    - Studying other versions
    - Neural Networks
- Merging Sentiment Scoring and Monte Carlo

❑ Overview

The back-end and front-end currently stands strong with minor needs of improvement

ALPHA:

- Functionality of Back-End: Monte Carlo (Long Term Analysis) & WebCrawler/Sentiment Scoring (Immediate Analysis), Spark Testing, Database
- Functionality of Front-End: Login, Register, Personal Portfolio Stock Watch List, Search & Add Stock to Portfolio, Interactive Chart Analysis of Sentiment Scoring

BETA Goals:

- Merging Sentiment Scoring and Monte Carlo
- Scale with Apache Spark & Hadoop
- Base and run the project with AWS