Presenter: Lex

Data Science DS-SG-04– General Assembly – 8 July 2017

3 Project Proposals

Project Titles

Presenter: Lex

Predict the S&P500 Index using the Stock prices of the 4 Tech Giants Google, Facebook, Amazon and Apple

Project Proposal 1: by Lex

Predict the S&P500 Index

	A	B	C	D	E	F
1	date	AAPL	AMZN	FB	GOOG	SP500
2	01/08/2012	86.69	232.09	28.00	315.16	1,375.14
3	02/08/2012	86.83	230.81	27.77	313.20	1,365.00
4	03/08/2012	87.96	234.97	28.76	319.47	1,390.99
5	06/08/2012	88.94	233.99	29.42	320.21	1,394.23
6	07/08/2012	88.70	236.56	29.06	319.07	1,401.35
7	08/08/2012	88.55	234.38	30.59	319.92	1,402.22
8	09/08/2012	88.68	234.06	31.30	319.98	1,402.80
9	10/08/2012	88.81	232.75	31.72	319.80	1,405.87
10	13/08/2012	90.00	232.44	30.95	328.77	1,404.11
11	14/08/2012	90.24	233.19	30.10	333.08	1,403.93
12	15/08/2012	90.12	237.42	29.85	332.52	1,405.53
13	16/08/2012	90.91	241.55	30.14	335.18	1,415.51
14	17/08/2012	92.59	241.17	29.66	337.31	1,418.16
15	20/08/2012	95.02	240.35	30.73	336.51	1,418.13
16	21/08/2012	93.72	239.45	30.82	333.51	1,413.17
17	22/08/2012	95.55	243.10	31.08	337.33	1,413.49
18	23/08/2012	94.66	241.20	31.54	337.14	1,402.08
19	24/08/2012	94.75	245.74	32.47	338.05	1,411.13

<u>The Problem: What's the background and scope of the project idea?</u>

- Can the overall market (S&P500 Index Price) be predicted by the Stock prices of 4 technology giant's stocks (Facebook, Apple, Google, Amazon) from 1 Aug 2012 to 1 Aug 2016 (4 years)

<u>What problem are you attempting to address or solve?</u>

- Solve the problem of having to pick stocks. Investors just need to buy the tech giants if there is strong correlation

<u>Who may it matter to?</u>

Investors who buy the 4 tech giants stocks

Predict the S&P500 Index

	A	B	C	D	E	F
1	date	AAPL	AMZN	FB	GOOG	SP500
991	08/07/2016	96.68	745.81	117.43	705.63	2,129.90
992	11/07/2016	96.98	753.78	117.31	715.09	2,137.16
993	12/07/2016	97.42	748.21	117.95	720.64	2,152.14
994	13/07/2016	96.87	742.63	118.91	716.98	2,152.43
995	14/07/2016	98.79	741.20	119.68	720.95	2,163.75
996	15/07/2016	98.78	735.44	117.77	719.85	2,161.74
997	18/07/2016	99.83	736.07	120.31	733.78	2,166.89
998	19/07/2016	99.87	739.95	120.21	736.96	2,163.78
999	20/07/2016	99.96	745.72	120.57	741.19	2,173.02
1000	21/07/2016	99.43	744.43	119.87	738.63	2,165.17
1001	22/07/2016	98.66	744.86	119.24	742.74	2,175.03
1002	25/07/2016	97.34	739.61	119.09	739.77	2,168.48
1003	26/07/2016	96.67	735.59	119.04	738.42	2,169.18
1004	27/07/2016	102.95	736.67	117.40	741.77	2,166.58
1005	28/07/2016	104.34	752.61	117.27	745.91	2,170.06
1006	29/07/2016	104.21	758.81	118.01	768.79	2,173.60
1007	01/08/2016	106.05	767.74	116.92	772.88	2,170.84
1008						

Data: What data exists to help solve this problem?
- S&P Index Price, Stock Prices (close at end of the day) of Facebook, Apple, Google, Amazon

Where is it coming from?
- S&P Official website, Kaggle

What does the data look like?
- 1,007 days from 1/08/2012 to 01/08/2016, with total 6 columns (Date, 4 companies, 1 Index). Total 6,036 data points

What is the observation?
Positive correlation at first glance

Organize with Sections

Lightning Presentation

Hypotheses: Given the problem and data you're aware of, what do you believe is the solution?

- The S&P500 Index is widely regarded as the best single gauge of large US company stocks, hence it approximates the “overall market”. It includes the 500 leading companies and captures approximately 80% coverage of available market capitalization.
- The hypothesis is that the 4 tech giants prices move in the same direction as the overall market.
- The solution is that using a combination of all 4 tech giants is the best predictor of the market.

What does success look like?

- Success will show when all 4 tech giants prices, when put into the model **together**, is the best predictor of the S&P500 market. If individual stocks, such as FB, are even better than all 4 stocks as a predictor of the S&P500, then we also have a new discovery, i.e. that FB alone can better predict the S&P500 than 4 tech giant stocks.

Best month or day to invest in US Stock Market

Project Proposal 2: by Lex

The Problem: What's the background and scope of the project idea?

Which day or month will stocks have the highest prices?

	A	B	C	D
1	date	day	month	SP500
2	01/08/2012	1	8	1,375.14
3	02/08/2012	2	8	1,365.00
4	03/08/2012	3	8	1,390.99
5	06/08/2012	6	8	1,394.23
6	07/08/2012	7	8	1,401.35
7	08/08/2012	8	8	1,402.22
8	09/08/2012	9	8	1,402.80
9	10/08/2012	10	8	1,405.87
10	13/08/2012	13	8	1,404.11
11	14/08/2012	14	8	1,403.93
12	15/08/2012	15	8	1,405.53
13	16/08/2012	16	8	1,415.51
14	17/08/2012	17	8	1,418.16
15	20/08/2012	20	8	1,418.13
16	21/08/2012	21	8	1,413.17
17	22/08/2012	22	8	1,413.49
18	23/08/2012	23	8	1,402.08
19	24/08/2012	24	8	1,411.13

S&P500 Historical Data

What problem are you attempting to address or solve?

- Investors know which day or which month to log in to their trading accounts and buy stocks.
- Is it true there is phenonmenon such as “sell in May and go away”: meaning low prices from May and June

Who may it matter to?

Investors who like to invest in the overall market (S&P500 Index), purely based on the date (e.g. 23rd of the month), or the month itself.

Data: What data exists to help solve this problem?

- S&P Index Price

Where is it coming from?

- S&P Official website

What does the data look like?

- 4 years of data from 1/08/2012 to 01/08/2016, with total 2 columns (Date, 1 Index).

What is the observation?

- As the price increases over the 4 years, the data set has to be split into 4 sets, 1 year each (i.e. 2012, 2013, 2014, 2015 onwards)

Video Perfect

	A	B	C	D
1	date	day	month	SP500
2	01/08/2012	1	8	1,375.14
3	02/08/2012	2	8	1,365.00
4	03/08/2012	3	8	1,390.99
5	06/08/2012	6	8	1,394.23
6	07/08/2012	7	8	1,401.35
7	08/08/2012	8	8	1,402.22
8	09/08/2012	9	8	1,402.80
9	10/08/2012	10	8	1,405.87
10	13/08/2012	13	8	1,404.11
11	14/08/2012	14	8	1,403.93
12	15/08/2012	15	8	1,405.53
13	16/08/2012	16	8	1,415.51
14	17/08/2012	17	8	1,418.16
15	20/08/2012	20	8	1,418.13
16	21/08/2012	21	8	1,413.17
17	22/08/2012	22	8	1,413.49
18	23/08/2012	23	8	1,402.08
19	24/08/2012	24	8	1,411.13

Lightning Presentation
Hypotheses: Given the problem and data you're aware of, what do you believe is the solution?

- The hypothesis is that there are actually no months or days where stock prices are higher than other months.
- The solution is it does not matter which month you fit into the model, the price is still not having a large difference

What does success look like?
- Success will show no correlation between month vs. price, or day vs. price.

Can Apple's stock price predict Google's stock price?

My eyes are hazy! I need more sleep!

Apple VS Google

» The Problem: What's the background and scope of the project idea?

- By looking at the Apple's stock price, can we predict Google's stock price?

What problem are you attempting to address or solve?

- An intellectual curiosity: does the rise of Apple mean a fall in Google? Or do they instead have a symbiotic relationship where when Apple does well, Google also does well.

» Who may it matter to?

- People who are curious about business. Google's businesses such as Adwords, are they helped by Apple's businesses, such as the iPhone?

	A	B	E
1	date	AAPL	GOOG
2	01/08/2012	86.69	315.16
3	02/08/2012	86.83	313.20
4	03/08/2012	87.96	319.47
5	06/08/2012	88.94	320.21
6	07/08/2012	88.70	319.07
7	08/08/2012	88.55	319.92
8	09/08/2012	88.68	319.98
9	10/08/2012	88.81	319.80
10	13/08/2012	90.00	328.77
11	14/08/2012	90.24	333.08
12	15/08/2012	90.12	332.52
13	16/08/2012	90.91	335.18
14	17/08/2012	92.59	337.31
15	20/08/2012	95.02	336.51
16	21/08/2012	93.72	333.51
17	22/08/2012	95.55	337.33
18	23/08/2012	94.66	337.14
19	24/08/2012	94.75	338.05

Price of Apple (AAPL) and Google (GOOGL) across 4 years

Data analysis

Data: What data exists to help solve this problem?

- Stock Prices (close at end of the day) of Apple, Google

Where is it coming from?

- Kaggle

What does the data look like?

- 1,007 days from 1/08/2012 to 01/08/2016, with total 3 columns (Date, 2 companies)

What is the observation?

- Wildly varying but it seems Google is performing better than Apple and Apple is not a good predictor of Apple's performance

Lightning Presentation

Hypotheses: Given the problem and data you're aware of, what do you believe is the solution?

- The hypothesis is that when Apple rises or falls, Google is not affected.

The solution is that Apple prices do not affect Google stock prices.

What does success look like?

- Success will show a low or non-existent correlation between Apple stock prices and Google stock prices.

I hope you enjoyed the presentation