

WQD 7005 Data Mining

Milestone 4 Interpretation of data

Name: Ng Kang Wei

Student ID: WQD170068

Video presentation: <https://youtu.be/a0alaGTWEfY> (please turn on subtitle)

In the previous milestone, the focus is on a plantation stock, Sime Darby Plantation Berhad (5285). It is found that Kawan Food Bhd (7216) has the highest positive correlation with Sime Darby and Hock Heng Stone Industries Berhad (5165) has the highest negative correlation with Sime Darby.

In this milestone, the data undergo quantitative analysis and qualitative analysis.

Quantitative analysis

In continuance with the previous milestone, the correlation analysis has been performed on Sime Darby and the 2 stocks that have the highest positive correlation and negative correlation are identified. Correlation analysis is part of quantitative analysis as well. The quantitative analysis in this milestone is on the variables of the identified stocks. The mean, standard deviation, minimum value, the percentiles and the maximum values of all the variables.

The figure 1 below shows the stock data collected for Sime Darby Plantation Berhad. Figure 2 shows the quantitative analysis on Sime Darby data. The last line in Figure 2 show the total changes of the closing price of Sime Darby in this period. In the period of 28th Feb 2019 to 3rd Apr 2019, the price of Sime Darby has decrease 1.17%

Sime Darby		date	stock_code	open	close	high	low	vol	change_percentage
896		2019-02-28	5285	5.16	5.11	5.16	5.03	7237	-0.968992
1916		2019-02-28	5285	5.16	5.10	5.16	5.03	37831	-1.162791
2932		2019-03-04	5285	5.05	5.10	5.17	5.05	4273	0.990099
3948		2019-03-04	5285	5.05	5.11	5.17	5.05	22020	1.188119
4964		2019-03-05	5285	5.11	5.09	5.16	5.06	4036	-0.391389
5980		2019-03-06	5285	5.09	5.11	5.12	5.09	1971	0.392927
6641		2019-03-06	5285	5.09	5.13	5.15	5.09	9399	0.785855
7660		2019-03-07	5285	5.11	5.15	5.18	5.11	2369	0.782779
8679		2019-03-07	5285	5.11	5.14	5.18	5.11	7747	0.587084
9698		2019-03-08	5285	5.10	5.15	5.15	5.08	9358	0.980392
10717		2019-03-08	5285	5.10	5.10	5.16	5.08	30586	0.000000
11736		2019-03-11	5285	5.05	5.10	5.14	5.05	20298	0.990099
12754		2019-03-12	5285	5.06	5.08	5.09	5.03	19532	0.395257
13769		2019-03-13	5285	5.06	5.09	5.10	5.06	18418	0.592885
14784		2019-03-13	5285	5.06	5.09	5.10	5.06	20156	0.592885
15799		2019-03-15	5285	5.08	5.08	5.09	5.07	11292	0.000000
16814		2019-03-15	5285	5.08	5.05	5.09	5.05	101692	-0.590551
17825		2019-03-18	5285	5.05	5.10	5.16	5.05	21302	0.990099
18836		2019-03-19	5285	5.09	5.09	5.10	5.06	9966	0.000000
19847		2019-03-20	5285	5.05	5.15	5.16	5.05	9380	1.980198
20857		2019-03-21	5285	5.17	5.06	5.18	5.05	34091	-2.127660
21867		2019-03-22	5285	5.04	5.07	5.07	5.01	13669	0.595238
22877		2019-03-25	5285	5.03	5.06	5.07	4.99	24110	0.596421
23887		2019-03-26	5285	5.05	5.08	5.12	5.00	5923	0.594059
24897		2019-03-27	5285	5.06	5.09	5.13	5.06	14265	0.592885
25906		2019-03-28	5285	5.06	5.02	5.09	4.99	22864	-0.790514
26915		2019-04-02	5285	5.15	5.02	5.18	5.01	29465	-2.524272
27924		2019-04-03	5285	5.05	5.05	5.11	5.04	18957	0.000000

Figure 1: Sime Darby data

	open	close	high	low	vol	change_percentage
count	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000
mean	5.082857	5.091786	5.133571	5.050357	19007.392857	0.181111
std	0.039333	0.034432	0.036537	0.031796	18891.616442	0.993241
min	5.030000	5.020000	5.070000	4.990000	1971.000000	-2.524272
25%	5.050000	5.077500	5.100000	5.030000	8955.250000	-0.097847
50%	5.070000	5.090000	5.145000	5.050000	16341.500000	0.589985
75%	5.102500	5.110000	5.160000	5.062500	22231.000000	0.783548
max	5.170000	5.150000	5.180000	5.110000	101692.000000	1.980198
The percentage of change during this period: -1.174168%						

Figure 2: Quantitative Analysis on Sime Darby data

Kawan									
	date	stock_code	open	close	high	low	vol	change_percentage	
410	2019-02-28	7216	1.65	1.65	1.65	1.65	400	0.000000	
1430	2019-02-28	7216	1.65	1.65	1.65	1.65	940	0.000000	
2449	2019-03-04	7216	1.78	1.82	1.85	1.74	1118	2.247191	
3465	2019-03-04	7216	1.78	1.75	1.85	1.74	1612	-1.685393	
4482	2019-03-05	7216	1.78	1.78	1.78	1.78	10	0.000000	
5497	2019-03-06	7216	1.75	1.75	1.75	1.75	50	0.000000	
6876	2019-03-06	7216	1.75	1.75	1.75	1.75	120	0.000000	
7895	2019-03-07	7216	1.78	1.78	1.78	1.78	327	0.000000	
8914	2019-03-07	7216	1.78	1.80	1.81	1.78	935	1.123596	
9933	2019-03-08	7216	1.78	1.78	1.80	1.78	108	0.000000	
10952	2019-03-08	7216	1.78	1.77	1.80	1.77	978	-0.561798	
11970	2019-03-11	7216	0.00	1.77	0.00	0.00	0	inf	
12988	2019-03-12	7216	1.77	1.75	1.77	1.75	99	-1.129944	
14003	2019-03-13	7216	1.75	1.75	1.75	1.74	236	0.000000	
15018	2019-03-13	7216	1.75	1.75	1.75	1.74	236	0.000000	
16033	2019-03-15	7216	1.75	1.75	1.75	1.75	204	0.000000	
17048	2019-03-15	7216	1.75	1.76	1.77	1.75	481	0.571429	
18059	2019-03-18	7216	1.76	1.75	1.76	1.75	170	-0.568182	
19070	2019-03-19	7216	1.79	1.76	1.79	1.76	314	-1.675978	
20080	2019-03-20	7216	1.77	1.77	1.78	1.77	363	0.000000	
21090	2019-03-21	7216	1.79	1.71	1.79	1.70	2106	-4.469274	
22100	2019-03-22	7216	1.70	1.70	1.70	1.69	584	0.000000	
23110	2019-03-25	7216	1.70	1.67	1.70	1.67	383	-1.764706	
24120	2019-03-26	7216	1.68	1.72	1.72	1.67	79	2.380952	
25130	2019-03-27	7216	0.00	1.72	0.00	0.00	0	inf	
26139	2019-03-28	7216	1.66	1.68	1.68	1.65	569	1.204819	
27148	2019-04-02	7216	1.78	1.68	1.78	1.68	501	-5.617978	
28157	2019-04-03	7216	1.70	1.68	1.77	1.68	320	-1.176471	

Figure 3: Kawan data

	open	close	high	low	vol	change_percentage
count	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000
mean	1.620000	1.737500	1.633214	1.604286	472.964286	inf
std	0.459613	0.044773	0.463905	0.455179	503.863700	NaN
min	0.000000	1.650000	0.000000	0.000000	0.000000	-5.617978
25%	1.700000	1.707500	1.715000	1.677500	117.000000	-0.708622
50%	1.750000	1.750000	1.765000	1.740000	323.500000	0.000000
75%	1.780000	1.770000	1.782500	1.752500	572.750000	0.142857
max	1.790000	1.820000	1.850000	1.780000	2106.000000	inf
The percentage of change during this period: 1.818182%						

Figure 4: Quantitative Analysis of Kawan

Figure 3 shows the data collected on Kawan Food Bhd, figure 4 shows the quantitative analysis on Kawan. The closing price of Kawan has increase by 1.18% during this period.

Hock Heng								
	date	stock_code	open	close	high	low	vol	change_percentage
683	2019-02-28	5165	0.550	0.560	0.560	0.550	200	1.818182
1703	2019-02-28	5165	0.550	0.550	0.560	0.550	250	0.000000
2722	2019-03-04	5165	0.550	0.550	0.550	0.550	150	0.000000
3738	2019-03-04	5165	0.550	0.550	0.550	0.550	150	0.000000
4754	2019-03-05	5165	0.000	0.550	0.000	0.000	0	inf
5770	2019-03-06	5165	0.550	0.550	0.550	0.550	130	0.000000
6431	2019-03-06	5165	0.550	0.550	0.550	0.550	130	0.000000
7450	2019-03-07	5165	0.000	0.550	0.000	0.000	0	inf
8469	2019-03-07	5165	0.545	0.540	0.545	0.540	150	-0.917431
9488	2019-03-08	5165	0.000	0.540	0.000	0.000	0	inf
10507	2019-03-08	5165	0.550	0.560	0.560	0.550	170	1.818182
11526	2019-03-11	5165	0.550	0.535	0.550	0.535	400	-2.727273
12544	2019-03-12	5165	0.000	0.535	0.000	0.000	0	inf
13562	2019-03-13	5165	0.550	0.545	0.550	0.545	300	-0.909091
14577	2019-03-13	5165	0.550	0.545	0.550	0.545	300	-0.909091
15592	2019-03-15	5165	0.580	0.580	0.580	0.580	317	0.000000
16607	2019-03-15	5165	0.580	0.580	0.580	0.575	954	0.000000
17618	2019-03-18	5165	0.560	0.580	0.580	0.560	610	3.571429
18629	2019-03-19	5165	0.570	0.580	0.580	0.570	290	1.754386
19640	2019-03-20	5165	0.590	0.590	0.595	0.590	440	0.000000
20650	2019-03-21	5165	0.600	0.640	0.665	0.600	4506	6.666667
21660	2019-03-22	5165	0.650	0.675	0.680	0.650	6597	3.846154
22670	2019-03-25	5165	0.670	0.650	0.670	0.650	1944	-2.985075
23680	2019-03-26	5165	0.650	0.640	0.660	0.640	600	-1.538462
24690	2019-03-27	5165	0.630	0.630	0.645	0.615	1976	0.000000
25699	2019-03-28	5165	0.630	0.665	0.670	0.630	1905	5.555556
26708	2019-04-02	5165	0.650	0.705	0.705	0.650	7392	8.461538
27717	2019-04-03	5165	0.705	0.705	0.705	0.680	2446	0.000000

Figure 5: Hock Heng data

	open	close	high	low	vol	change_percentage
count	28.000000	28.000000	28.000000	28.000000	28.000000	28.000000
mean	0.503929	0.586786	0.513929	0.500179	1153.821429	inf
std	0.214149	0.053649	0.220069	0.211925	1934.849579	NaN
min	0.000000	0.535000	0.000000	0.000000	0.000000	-2.985075
25%	0.550000	0.550000	0.550000	0.545000	150.000000	0.000000
50%	0.550000	0.560000	0.560000	0.550000	300.000000	0.000000
75%	0.607500	0.632500	0.648750	0.603750	1191.750000	4.273504
max	0.705000	0.705000	0.705000	0.680000	7392.000000	inf
The percentage of change during this period: 25.892857%						

Figure 6: Quantitative Analysis on Hock Heng

Figure 5 shows the data collected on Hock Heng Stone Industries, figure 6 shows the quantitative analysis on Hock Heng. During this period, Hock Heng prices has increase by 25.89%.

Qualitative Analysis

Qualitative analysis is content analysis that analyze text. Find information from text or news to get an insight in the stock movement. Twitter posts and news article are collected previously. The information within are sift through with SQL wildcard query.

The focus is still on Sime Darby Plantation Stock. Keywords such as palm or oil are used.

The SQL query used is

```
select * from stock_twitter where tweet like '%palm%';
```

The result from the database is shown below:

twitter_user	date_posted	tweet	favorite_count
mpochHQ	Sun Mar 17 05:34:23 +0000 2019	RT @teresakok: Teresa Kok : EU's 'high risk' ruling on palm oil shows double standard KLSE Screener https://t.co/UxaYcvCIEH	0
AliceChoo5	Sat Mar 16 09:40:39 +0000 2019	RT @teresakok: Teresa Kok : EU's 'high risk' ruling on palm oil shows double standard KLSE Screener https://t.co/UxaYcvCIEH	0
xineyeo	Sat Mar 16 05:47:11 +0000 2019	RT @teresakok: Teresa Kok : EU's 'high risk' ruling on palm oil shows double standard KLSE Screener https://t.co/UxaYcvCIEH	0
alanling2tweet	Fri Mar 15 10:21:22 +0000 2019	RT @teresakok: Teresa Kok : EU's 'high risk' ruling on palm oil shows double standard KLSE Screener https://t.co/UxaYcvCIEH	0
DDannfelt	Fri Mar 15 07:22:31 +0000 2019	RT @teresakok: Teresa Kok : EU's 'high risk' ruling on palm oil shows double standard KLSE Screener https://t.co/UxaYcvCIEH	0
teresakok	Fri Mar 15 06:43:14 +0000 2019	Teresa Kok : EU's 'high risk' ruling on palm oil shows double standard KLSE Screener https://t.co/UxaYcvCIEH	19

News articles are also searched with SQL query:

```
select * from stock_news where title LIKE '%palm%';
```

The result from the database is shown below:

title	link	description	published_date
M'sia threatens WTO challenge to EU's move to drop palm biofuel	https://www.malaysiakini.com/news/468266	Foreign Ministry calls EU's move a 'calculated political act'	Sat, 16 Mar 2019 13:58:56 GMT

Both of the results from twitter and news article are referring to the issue where European Union is planning to drop palm biofuel. Such a news would surely have an impact on the price and movement of the Sime Darby stock which is heavily reliant on palm oil.

The news article in question is:

Teresa Kok : EU's 'high risk' ruling on palm oil shows double standard

TheEdge

Fri, Mar 15, 2019 - 1 month ago

KUALA LUMPUR (March 15): The EU's decision to classify palm oil as "high risk" in proposed legislation on biofuel is based on inaccurate and discriminatory factors, Malaysia's Primary Industry Minister Teresa Kok says in statement.

* Decision of the European Commission to phase out palm oil as biofuel in Europe on the basis that palm oil causes deforestation is totally without foundation: Kok

* Malaysia has made clear that if the so-called Delegated Act on biofuel is adopted, that it will investigate retaliatory action: statement

** "Entire Delegated Act process has been based on the politics of protectionism:" Kok

* Malaysia notes palm oil produces eight times more oil than U.S. soybean oil per hectare but the European Commission classifies soybean as "low risk" for political reasons

* Decision to lower the definition of smallholders to 2 hectares compared with the previously proposed 2-5 hectares in the draft is discriminatory towards smallholders in palm-producing countries

Sime Darby									
	date	stock_code	open	close	high	low	vol	change_percentage	
896	2019-02-28	5285	5.16	5.11	5.16	5.03	7237	-0.968992	
1916	2019-02-28	5285	5.16	5.10	5.16	5.03	37831	-1.162791	
2932	2019-03-04	5285	5.05	5.10	5.17	5.05	4273	0.990099	
3948	2019-03-04	5285	5.05	5.11	5.17	5.05	22020	1.188119	
4964	2019-03-05	5285	5.11	5.09	5.16	5.06	4036	-0.391389	
5980	2019-03-06	5285	5.09	5.11	5.12	5.09	1971	0.392927	
6641	2019-03-06	5285	5.09	5.13	5.15	5.09	9399	0.785855	
7660	2019-03-07	5285	5.11	5.15	5.18	5.11	2369	0.782779	
8679	2019-03-07	5285	5.11	5.14	5.18	5.11	7747	0.587084	
9698	2019-03-08	5285	5.10	5.15	5.15	5.08	9358	0.980392	
10717	2019-03-08	5285	5.10	5.10	5.16	5.08	30586	0.000000	
11736	2019-03-11	5285	5.05	5.10	5.14	5.05	20298	0.990099	
12754	2019-03-12	5285	5.06	5.08	5.09	5.03	19532	0.395257	
13769	2019-03-13	5285	5.06	5.09	5.10	5.06	18418	0.592885	
14784	2019-03-13	5285	5.06	5.09	5.10	5.06	20156	0.592885	
15799	2019-03-15	5285	5.08	5.08	5.09	5.07	11292	0.000000	
16814	2019-03-15	5285	5.08	5.05	5.09	5.05	101692	-0.590551	
17825	2019-03-18	5285	5.05	5.10	5.16	5.05	21302	0.990099	
18836	2019-03-19	5285	5.09	5.09	5.10	5.06	9966	0.000000	
19847	2019-03-20	5285	5.05	5.15	5.16	5.05	9380	1.980198	
20857	2019-03-21	5285	5.17	5.06	5.18	5.05	34091	-2.127660	
21867	2019-03-22	5285	5.04	5.07	5.07	5.01	13669	0.595238	
22877	2019-03-25	5285	5.03	5.06	5.07	4.99	24110	0.596421	
23887	2019-03-26	5285	5.05	5.08	5.12	5.00	5923	0.594059	
24897	2019-03-27	5285	5.06	5.09	5.13	5.06	14265	0.592885	
25906	2019-03-28	5285	5.06	5.02	5.09	4.99	22864	-0.790514	
26915	2019-04-02	5285	5.15	5.02	5.18	5.01	29465	-2.524272	
27924	2019-04-03	5285	5.05	5.05	5.11	5.04	18957	0.000000	

The figure above highlighted the day which the news of EU is planning to drop palm biofuel is released. The price of Sime Darby stock fell, it close with a lower price than its opening price. It register the highest volume on that day, as investors heard the news and try to sell off the stock.