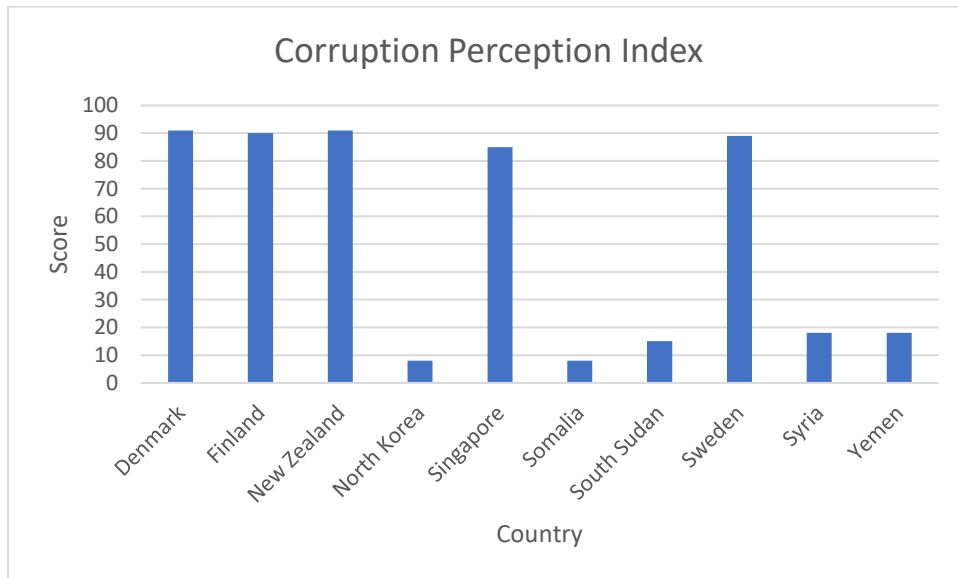**ETW1001 Assignment 1**

Name: Lim Zheng Haur

Student ID: 32023952

Part A: Data Visualization

i.

ii.

From the diagram in (i), we could see a column chart showing the Corruption Perception Index scores for Denmark, Finland, New Zealand, North Korea, Singapore, Somalia, South Sudan, Sweden, Syria and Yemen from 2015. It could be seen that Denmark, Finland, New Zealand, Singapore and Sweden have relatively high scores above 80. On the other hand, North Korea, Somalia, South Sudan, Syria and Yemen have very low scores that are lower than 20. This shows that Denmark, Finland, New Zealand, Singapore and Sweden have very less corruption within the country while North Korea, Somalia, South Sudan, Syria and Yemen are very corrupted. This highlights the vast difference of corruption between these countries.

iii.

Corruption Perception Index
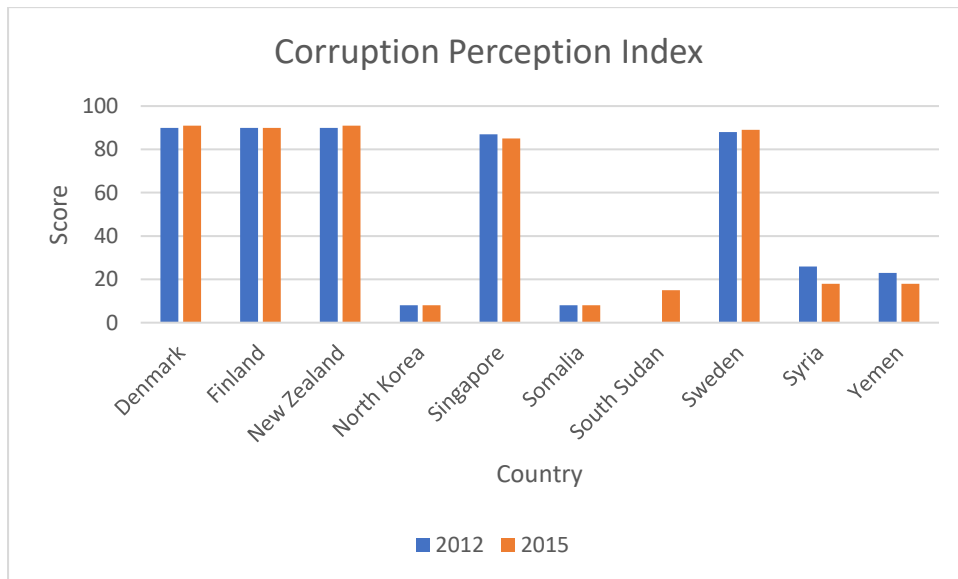
From the chart above, we could see that most of the countries which are Denmark, Finland, New Zealand, North Korea, Somalia and Sweden maintained their Corruption Perception Index scores from 2012 to 2015. However, there is a slight decrease for Singapore, Syria and Yemen from 2012 to 2015. This might be caused by many different factors but mainly due to the increase in corruption within the countries. As South Sudan did not have its data collected in 2012, a comparison could not be made.

iv.
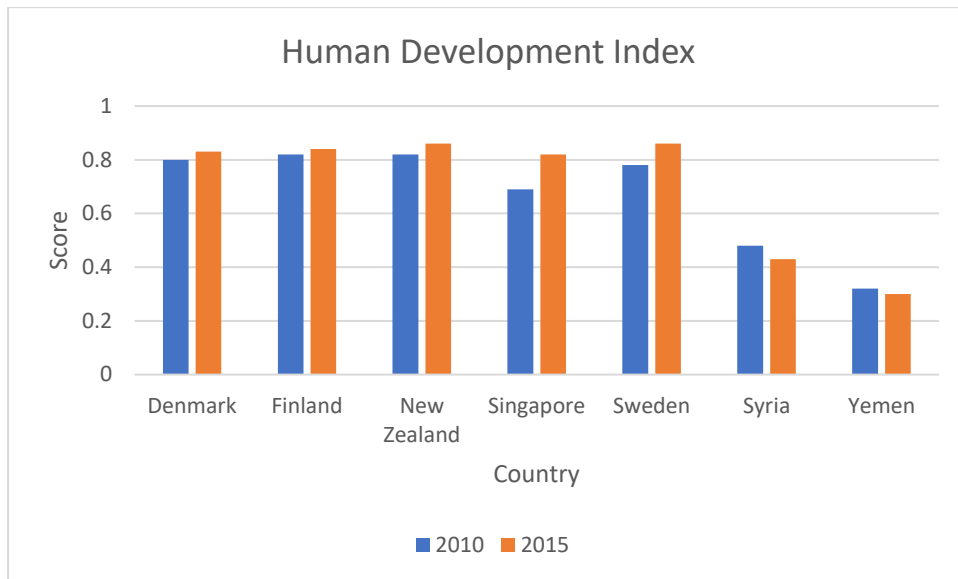
From this column chart, we could see the Human Development Index score for Denmark, Finland, New Zealand, Singapore, Sweden, Syria and Yemen. The data for Somalia, South Sudan and North Korea is not collected and could not be included in the chart. This chart shows that Denmark, Finland, New Zealand, Singapore and Sweden have a higher HDI score as compared to Syria and Yemen. We could also see a general trend that the higher scoring countries are able to increase their HDI score from 2010 to 2015 while the lower scoring countries, Syria and Yemen have decreasing HDI scores from 2010 to 2015.

Part B: Descriptive statistics

a. Table (a)

Category	Frequency	Percentage
Good	84	28.00%
Very Good	150	50.00%
Excellent	66	22.00%
Total	300	100.00%

b. Graph (b)
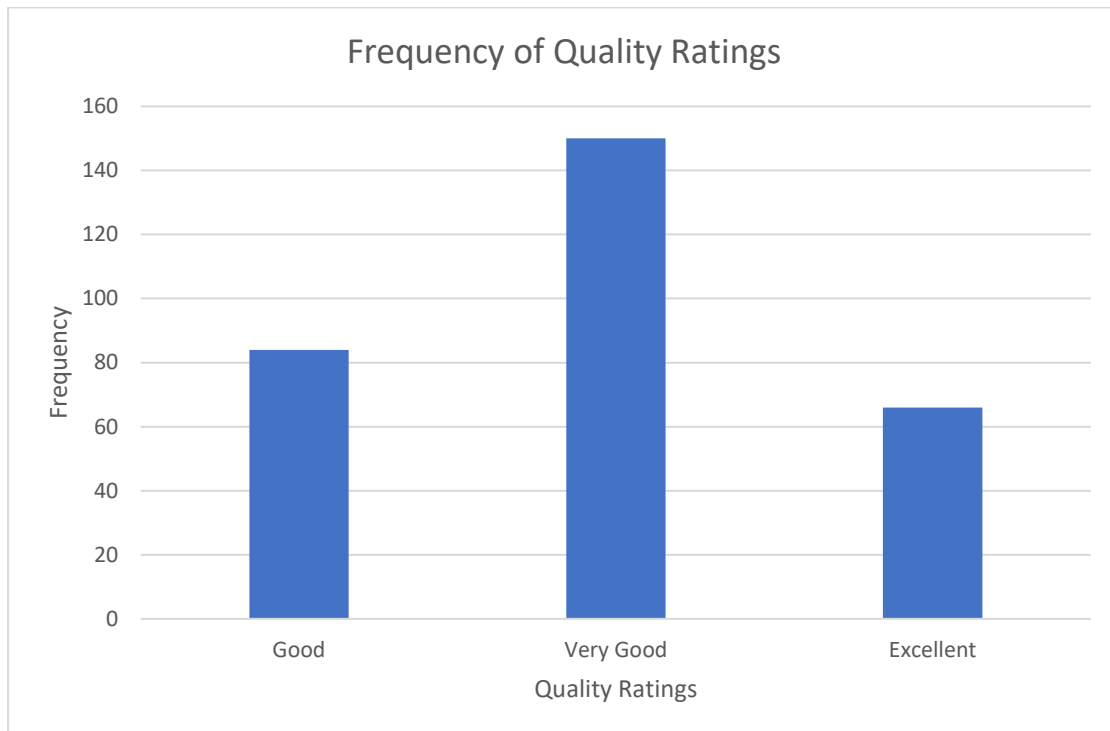
c. Table (c)

Summary Statistics	Good	Very Good	Excellent
Mean	19.82143	25.633333	35.01515
Standard Error	0.893522	0.6780955	1.297746
Median	19.5	25	34.5
Mode	20	28	32
Standard Deviation	8.189268	8.3049404	10.54294
Sample Variance	67.06411	68.972036	111.1536
Kurtosis	4.362796	-0.556534	8.695879
Skewness	0.978711	0.1448621	1.732834
Range	60	38	75
Minimum	-5	10	13
Maximum	55	48	88
Sum	1665	3845	2311
Count	84	150	66

d. Table (d)

	Count	300			
Lower bound	Upper bound	Midpoint	Excellent	Good	Very Good
-15	-5	-10	0	1	0
-5	5	0	0	0	0
5	15	10	1	23	20
15	25	20	9	40	56
25	35	30	26	17	56
35	45	40	22	2	17
45	55	50	7	1	1
55	65	60	0	0	0
65	75	70	0	0	0
75	85	80	0	0	0
85	95	90	1	0	0
			66	84	150

Percentage Frequency			Cumulative Percentage Frequency		
Excellent	Good	Very Good	Excellent	Good	Very Good
0.00%	1.19%	0.00%	0.00%	1.19%	0.00%
0.00%	0.00%	0.00%	0.00%	1.19%	0.00%
1.52%	27.38%	13.33%	1.52%	28.57%	13.33%
13.64%	47.62%	37.33%	15.15%	76.19%	50.67%
39.39%	20.24%	37.33%	54.55%	96.43%	88.00%
33.33%	2.38%	11.33%	87.88%	98.81%	99.33%
10.61%	1.19%	0.67%	98.48%	100.00%	100.00%
0.00%	0.00%	0.00%	98.48%	100.00%	100.00%
0.00%	0.00%	0.00%	98.48%	100.00%	100.00%
0.00%	0.00%	0.00%	98.48%	100.00%	100.00%
1.52%	0.00%	0.00%	100.00%	100.00%	100.00%
100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

e.

Yes, the level of quality rating is different for the three group of restaurants. There is an unusual figure of RM -5 for the minimum of "Good" ratings. There should not be negative values for meal price. There is also a usual outlier for the maximum of "Excellent" ratings of RM 88. These extreme values or outliers will cause the mean of the data to affect greatly. Therefore, a more accurate measure of centrality would be to use the median or mode.

f. Graph (f)

Percentage Frequency Polygon for Quality Ratings
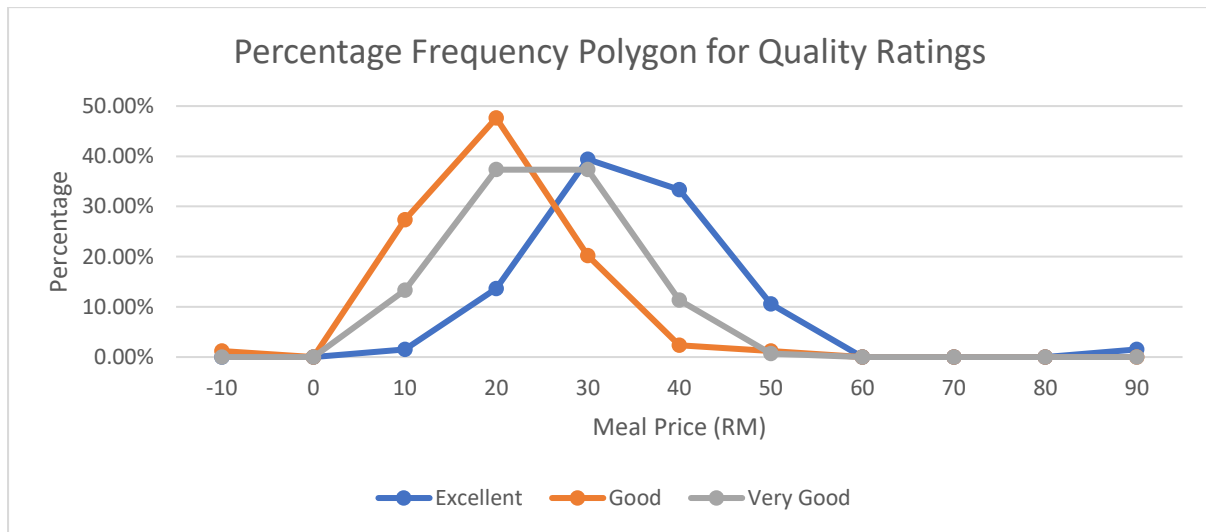
g.

The frequency polygon for "Good" quality rating is slightly skewed positively and have a mode of 48% at the food price class of RM15 - RM25. Meanwhile, the frequency polygon for "Very Good" quality rating is almost symmetrical with two modes at RM15 - RM25 and RM25 - RM35. Lastly, the frequency polygon for "Excellent" quality rating has a positive skewness and a mode of 39% at RM25 - RM35.

h.

The four measures of variability are the standard deviation, sample variance, range and interquartile range. However, in Table(c), only the standard deviation, sample variance and range is given while the interquartile range is not stated. The standard deviation of "Good" ratings is 8.19, "Very Good" ratings is8.30 and "Excellent" ratings is 10.54. The standard deviation shows that the "Excellent" ratings is the most variable, followed by "Very Good" and "Good" ratings. As the sample variance is relative to the standard deviation, the results also show that "Excellent" ratings is the most variable followed by "Very Good" and "Good". The range of the data shows that "Excellent" ratings has the highest variability but surprisingly "Good" has a higher range than "Very Good". This might be caused by a few outliers in "Good" ratings that cause it to have a greater range.

Part C: Estimation

i.
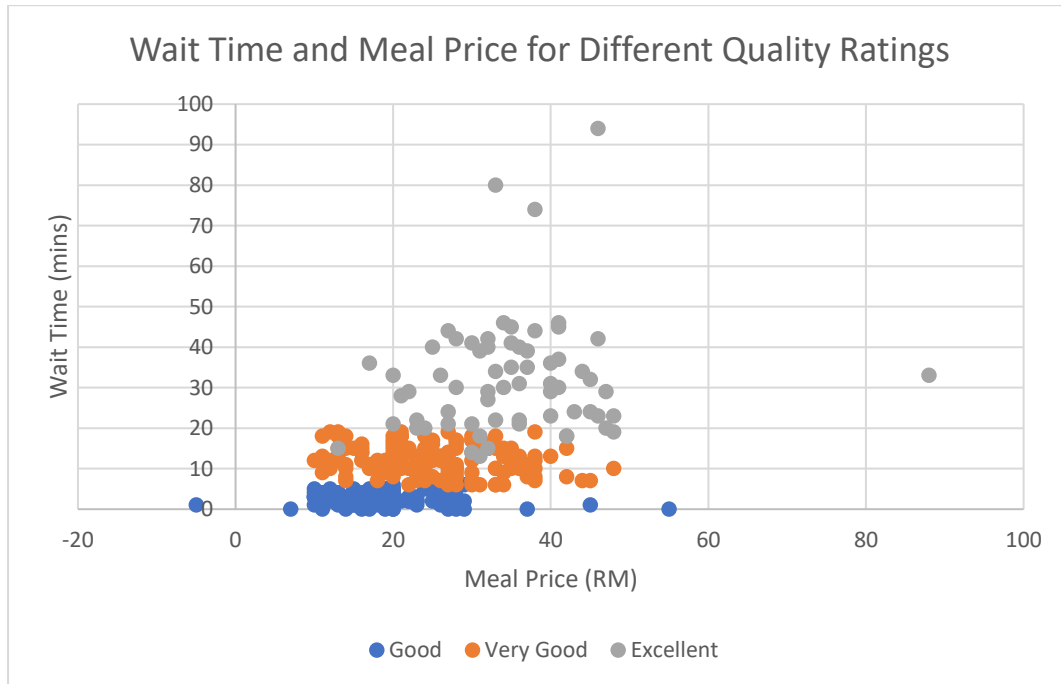
The diagram is a scatter plot which shows the wait time and meal price for each quality ratings. There is an observation where the better the quality rating, the higher the waiting time and meal price. This is an unusual phenomenon and might be caused by many possible reasons. One of the possible reasons for this is that the low waiting time meal is prepared long before consumed and the quality of the food is lacking while the longer waiting time allows for better preparation of the food by the restaurant. This also explains the higher price for the food that requires longer waiting time. There might be many more different causes for this unusual phenomenon other than this reason. It might also simply due to the personality of the customer providing the feedback as they are more generous to pay for more expensive meals and provide better quality ratings for the restaurant.

ii.

	Good	Very Good	Excellent
Mean	19.82142857	25.63333333	35.01515152
Sample std	8.189268197	8.304940445	10.54294139
Sample size	84	150	66
CV (t score)	1.98895978	1.976013178	1.997137908
Standard error	0.893522414	0.678095548	1.297746374
ME	1.777180145	1.339925738	2.59177848
Lower limit	18.04424843	24.2934076	32.42337304
Upper limit	21.59860872	26.97325907	37.60692999

It is 95% that the mean meal price for "Good" quality ratings population is between RM18.04 and RM21.60.

It is 95% that the mean meal price for "Very Good" quality ratings population is between RM24.29 and RM26.97.

It is 95% that the mean meal price for "Excellent" quality ratings population is between RM32.42 and RM37.61.