Take Home Assignment 1
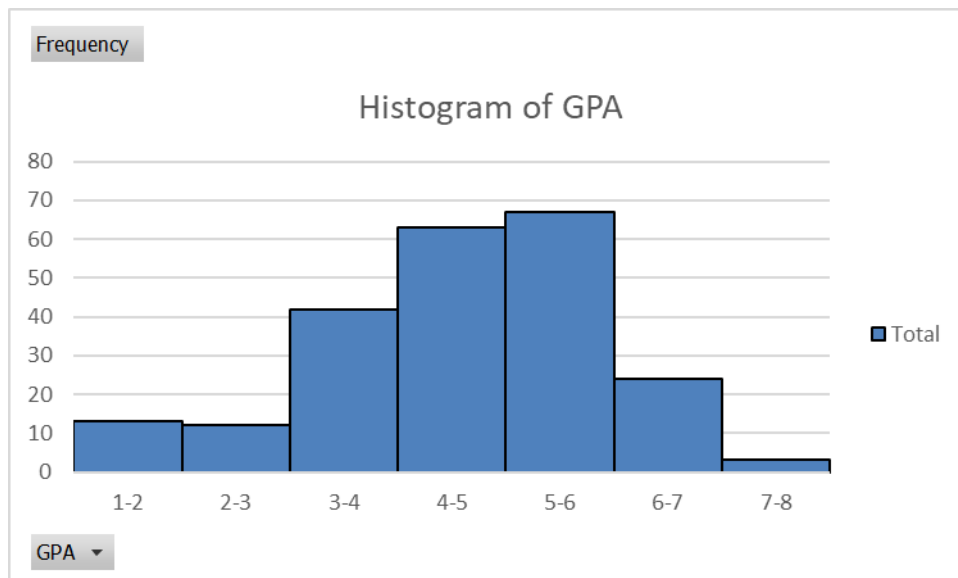
Numerical Descriptive Measures

Brendan Doncaster – N9819363

(a) Using the **Pivot Table** in **Excel**, construct a **frequency distribution table** of GPA and a **histogram** displaying the table.

Frequency Distribution Table:

GPA	Frequency
1-2	13
2-3	12
3-4	42
4-5	63
5-6	67
6-7	24
7-8	3
Grand Total	**224**

Histogram:

(b) Using **Excel,** compute the **sample mean ($\overline{X}$)** and **sample standard deviation (s)**. No working is required. *(Round your answers to 3 significant figures.)*

Sample Mean: 4.60

Sample Standard Deviation: 1.332

(c) Applying the _common sense method_ locate the **Q1**, **Q2** and **Q3** of the test scores. Show your working.

Starting bracket = [1 to 224]

224/2 = 112, therefore the two next brackets hold 112 each, which are [1 to 112] and [113 to 224]

112/2 = 56, therefore the next brackets hole 56 each, which are [1 to 56], [57 to 112], [113 to 168] and [169 to 224]

The quartiles Q1, Q2 and Q3 exist in between those 4 brackets, and since the starting amount was even, the quartiles must be found by averaging the edges of the brackets.

Q1 = (56 + 57) / 2
Q2 = (112 + 113) / 2
Q3 = (168 + 169) / 2

Then after sorting the data and inserting the GPA from these positions we get

Q1 = (3.78 + 3.79) / 2 = 3.785

Q2 = (4.76 + 4.77) / 2 = 4.765

Q3 = (5.61 + 5.61) / 2 = 5.61

(d) By comparing Q2 and the sample mean comment on the skewness of the data. Explain your answer briefly.

The data can be seen to be skewed to the left, this is because the sample mean (4.60) is lower than Q2, the median (4.765).

(e) An observation is a suspected outlier if it is more than 1.5 times the inter-quartile range (1.5 × IQR) below Q1 or above Q3. Use this rule to locate any outliers in the dataset.

IQR = Q3 − Q1 = 5.61 − 3.785 = 1.825

1.5 * IQR = 2.7375

Lower outliers = Q1 − 1.5 * IQR = 3.785 − 2.7375 = 1.0475

Upper outliers = Q3 + 1.5 * IQR = 5.61 + 2.7375 = 8.3475

Therefore the outliers are any GPA lower than 1.0475 or higher than 8.3475 (which is impossible)

Outlying data set:

1.00
1.00
1.00