

## Project Two

Date

### Data and Descriptive Statistics

The collected data was as follows:

	Students	Internet Duration (hrs)	Missed lessons
1	Male	0	1
2	Male	0	0
3	Male	0	0
4	Male	1	0
5	Male	1	1
6	Male	2	1
7	Male	3	3
8	Male	3	3
9	Male	3	4
10	Male	3	3
11	Male	4	4
12	Male	4	4
13	Male	4	3
14	Male	4	3
15	Male	4	3
16	Male	5	4
17	Male	5	4
18	Male	5	4
19	Male	6	8
20	Male	6	5
21	Male	6	5
22	Male	6	5
23	Male	7	7
24	Male	8	7
25	Male	8	8
26	Male	10	8
27	Female	2	2
28	Female	3	3
29	Female	3	4
30	Female	4	4
31	Female	4	4
32	Female	4	4
33	Female	4	4
34	Female	5	4
35	Female	5	3
36	Female	5	5
37	Female	6	4

38	Female	6	5
39	Female	6	6
40	Female	6	9
41	Female	6	2
42	Female	6	8
43	Female	7	4
44	Female	7	6
45	Female	8	0
46	Female	8	7
47	Female	8	8
48	Female	8	8
49	Female	9	5
50	Female	9	6
51	Female	9	7
52	Female	9	7
53	Female	9	7

The information collected from the students was computed into Excel and analyzed to reveal the findings. The data was calculated based on the comparison of gender, the time spent on the internet and the number of lessons they missed. The results were categorized as:

- Gender-based distribution of the variables
- General distribution to compare the relationship between missing lessons and the time spent on the internet.

### Distribution of the Quantitative Data

- Gender-based distribution of the variables

Table 1.1.0a.

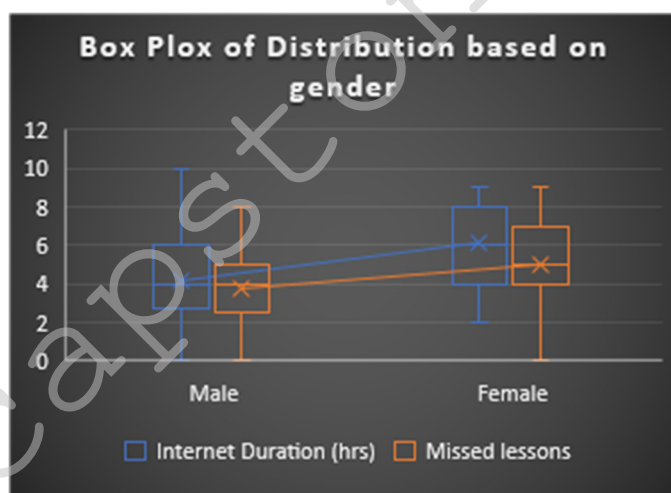


Table 1.1.1a

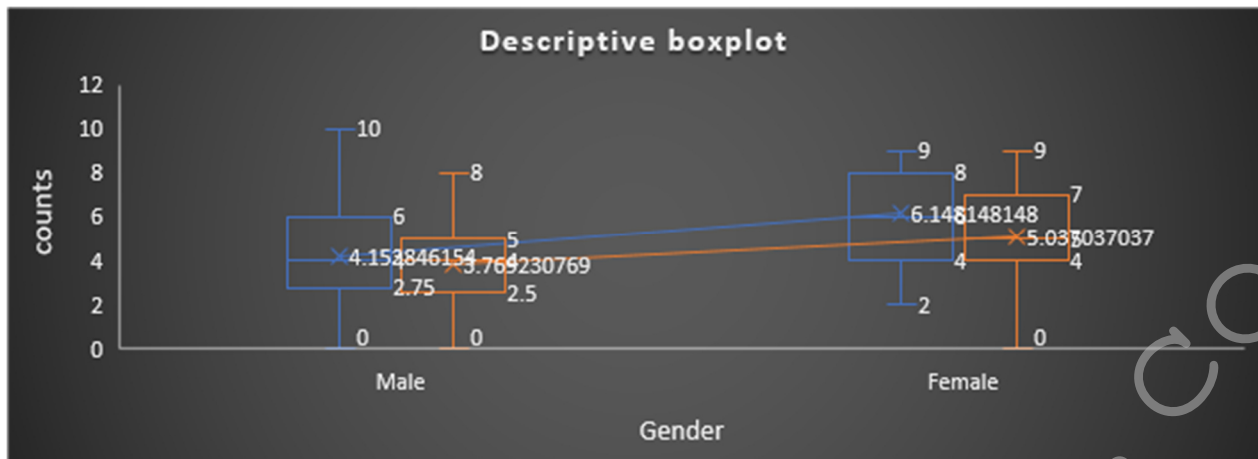


Table 1.2.0a:

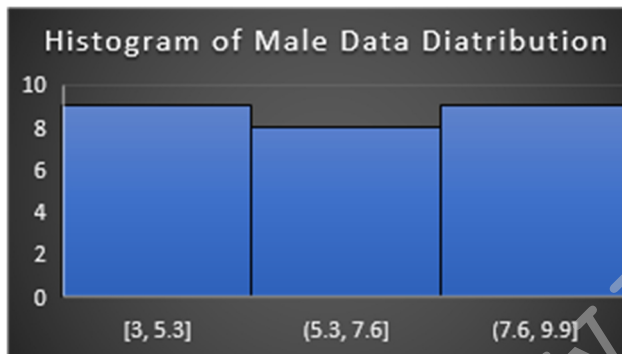
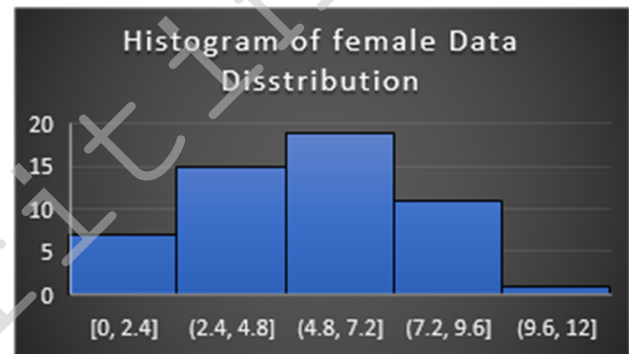


Table 1.2.1a



The findings of the boxplot diagram indicate the male respondents had a center of variation of approximately four missed lessons for every four hours of internet access. The males have a lower center of variation than the females. On the other hand, the female respondents had an average of 5 missed lessons for six hours on the internet. The findings suggest that the females spent more time on the internet than the males. The boxplot depicts outliers in both genders. The male had outliers of ten counts and eight counts for internet duration and missed lessons respectively. The distribution is positively skewed for the female gender.

b. The general descriptive statistics

Table 1.1.0b.

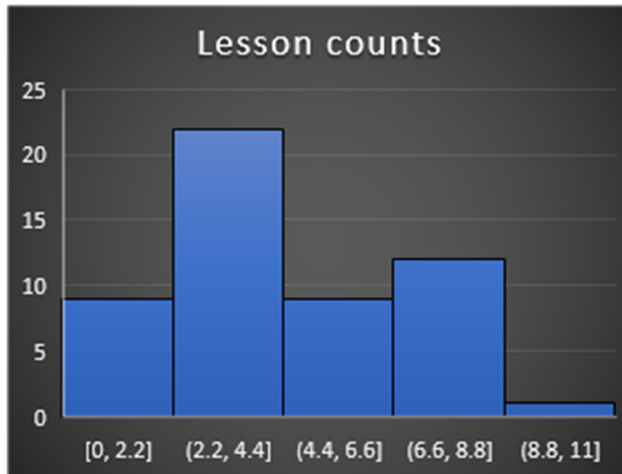


Table 1.1.1b

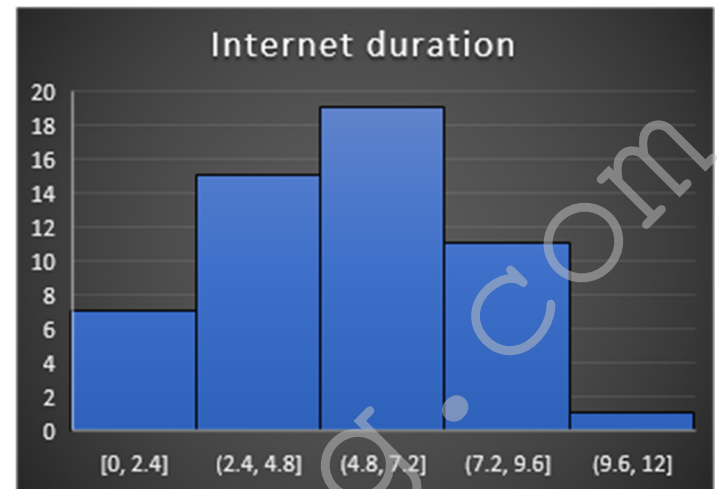


Table 1.2b

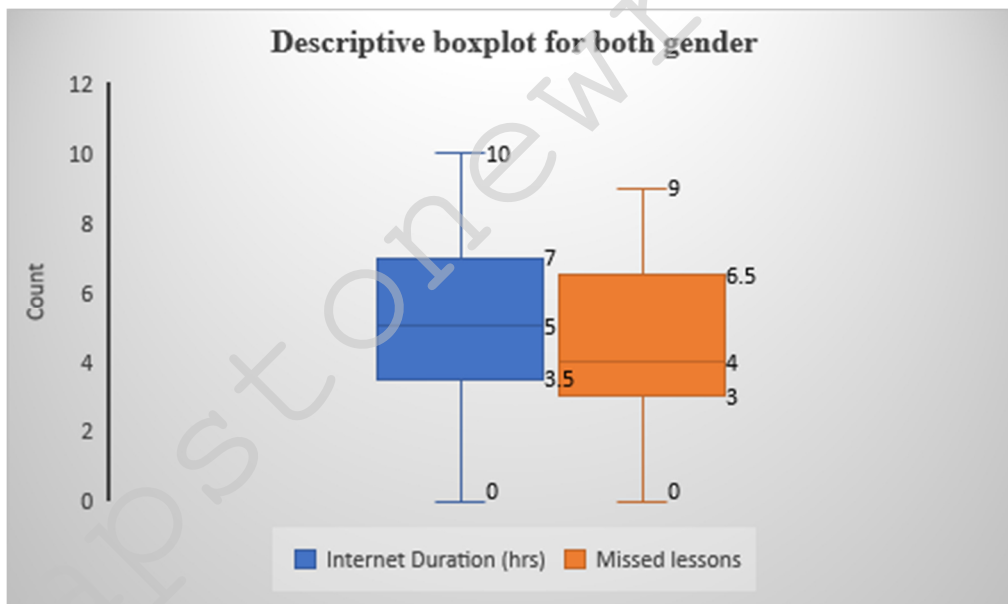


Table 1.1 b illustrates that the distribution across 4.8-7.2 had the highest incidences while the distributions of the extremities had the lowest frequencies. In the boxplot table 1.2b, the center of variation was 5 for the time spent on internet and 4 for the number of missed lessons. The distribution has outliers of ten and nine for the internet duration and the missed lessons

respectively. The mean of the data positively skews the distribution. The lesson counts' distribution has a positive skewness whereas the internet duration has a symmetrical skewness.

### **The z-score of the distribution**

Where  $z$  is the z score,  $x$  is the observations,  $\mu$  is the mean and  $\sigma$  is the standard deviation.

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

The z score of the lesson attendance will be

$$z = \frac{53 - 4.415}{2.37} = 20.5$$

On the other hand, the z score for the internet duration will be

$$z = \frac{53 - 5.17}{2.55} = 18.77$$

Therefore, the internet duration spent had a closer distribution about the mean data.

### **The linear Regression Analysis**

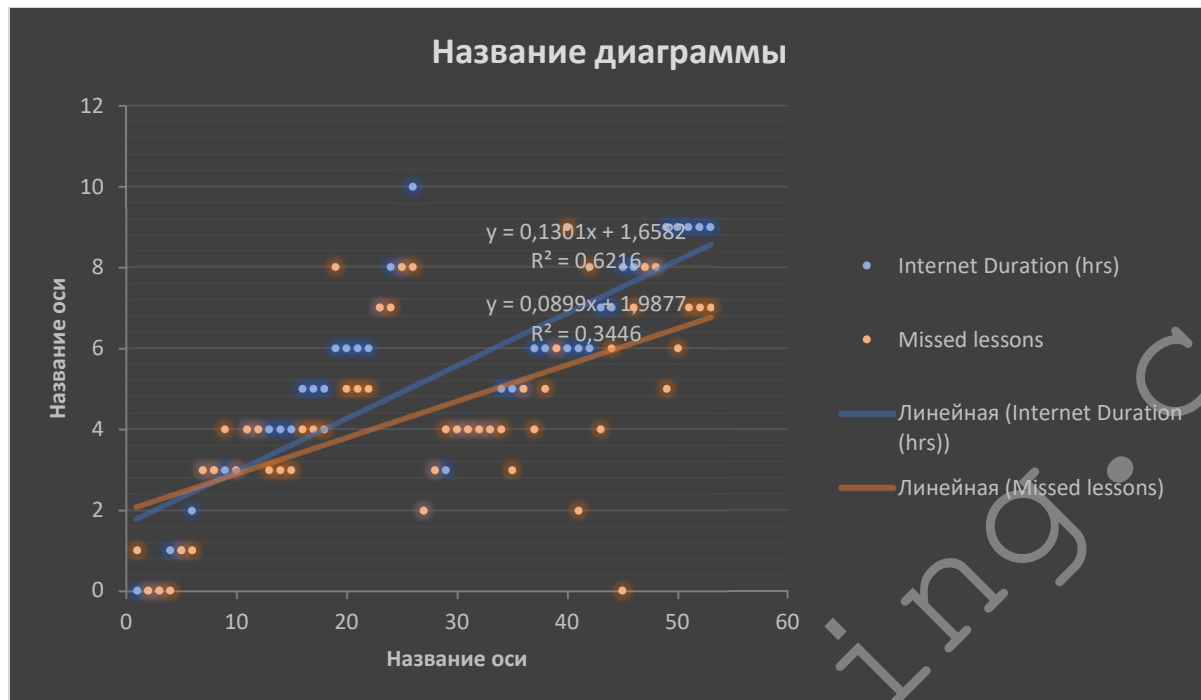
In the regression analysis, the independent variable is the internet duration whereas the dependent variable is the count of the lessons missed. The reason is that the According to previous findings, the number of lessons missed depends on the duration spent on the internet.

**Table 2.1: Regression Analysis**

<b>Regression Statistics</b>								
<b>Multiple R</b>	0.773254684							
<b>R Square</b>	0.597922807							
<b>Adjusted R Square</b>	0.590038941							
<b>Standard Error</b>	1.631129003							
<b>Observations</b>	53							
<b>ANOVA</b>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
<b>Regression</b>	1	201.782025	201.782	75.84132	1.15E-11			
<b>Residual</b>	51	135.6896731	2.660582					
<b>Total</b>	52	337.4716981						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
<b>Intercept</b>	1.49247535	0.478020423	3.1222	0.002956	0.532809	2.452141	0.532809	2.452141
<b>Missed lessons</b>	0.832900882	0.095640182	8.708692	1.15E-11	0.640895	1.024907	0.640895	1.024907

The scatter points

Table 3.1



The data has a uniform distribution across both variables, with an even rise in every variable. The clusters associate along the distribution such that they have a consistent separation along the y-intercepts. However, some outliers are persistent with the number of missed lessons on both sides of its y-intercept. The pattern of the lost lessons was more randomly distributed about the intercept than the duration spent on the internet.

### Correlation Coefficient and Linear Regression:

The correlation was calculated through the Excel data analysis tool. The correlation coefficient of the intercept lies at 1.49247535 compared to the missing lessons at 0.832900882. The coefficient has a relationship with the scatter points because the data is distributed across the y-intercept with an elevation in the missing lesson count for every rise in the internet duration. However, the missing lessons record slightly lower values because of the extended intercept position.

The results of the regression confirm the earlier perspectives as follows:

- For every increase in the duration spent on the internet, there is a correspondent increase in the number of lessons missed. However, there are insignificant scatter points from this finding (Al-Rahmi & Othman, 2013).
- The women spend more time on the internet than the male respondents, leading to higher losses in the number of lessons among the women than the men.

capstonewriting.com



## References

- Al-Rahmi, W., & Othman, M. (2013). The impact of social media use on academic performance among university students: A pilot study. *Journal of information systems research and innovation*, 4(12), 1-10.

capstonewriting.com