SECI2143-04 KEBARANGKALIAN STATISTIK & ANALISIS DATA (PROBABILITY & STATISTICAL DATA ANALYSIS) : PROJECT 1

SAIFUL HABIB DANIAL BIN SHIFUL ANUAR	A19EC0157
ARIF BIN SUHAIMI	A19EC0023
MUHAMAD ADLI HANIS BIN ABDUL SHUKOR	A19EC0091
MUHAMAD AMIN QAYYUM BIN MOHD RAZI	A19EC0093

LECTURER:

DR. SUHAILA MOHAMAD YUSUF

SCHOOL OF COMPUTING

FACULTY OF ENGINEERING

UNIVERSITI TEKNOLOGI MALAYSIA

2020

Table of content

INTRODUCTION

Probability & Statistical Data Analysis (SECI 2143) is planned to present a few measurable strategies as apparatuses to dissect the information. In the starting the understudies will be uncovered with different shapes of information. The information spoken to by the diverse sorts of factors are determined from distinctive sources; every day and mechanical activities. The investigation starts with the information representation outwardly.

The course will too investigate some methods of parameter estimation from diverse conveyances. Encourage information investigation is conducted by presenting the theory testing. A few models are utilized to fit bunches of data. At the conclusion of course the understudies ought to be able to apply a few measurable models in analyzing information utilizing accessible program. For our project 1, we were studying about *The number vehicles between block M22 and M23.* Block M22 and M23 are mostly occupied by first year students and a couple of second year students. Universiti Teknologi Malaysia (UTM) has written the law that only if you are a second year student and older can own vehicles (bicycles are legal for first year).

1.0 Objective

Our objective was to study the pattern behind the number of cars park between these two blocks. Other than that, we want to provide sufficient evidence to support the needs for students to bring their own vehicles. It is because; only these students can apply for a vehicle's sticker. However, the parking lot that provided between these two blocks were parked by tons of cars and under the porch there were many motorcycle. Sometimes there no empty parking spaces for parents or lecturers to park their vehicles if they have to come to these blocks. If there were any emergencies, for example one of these block was caught a fire, the fire and rescue team expected that there must be sufficient space to park their vehicles (based on the information that these two blocks for first year students) but the reality that there was wrong. That's the negative part.

CONTENT

2.1 Methodology (Scope of research)

The scope of our research is *Descriptive Statistics.* What is Descriptive Statistics? Descriptive Statistics is the term given to the examination of information that makes a difference portray, appear or summarize information in a significant way such that, for case, designs might develop from the information. Expressive insights don't , however, allow us to create conclusions past the information we have dissected or reach conclusions with respect to any speculations we might have made. They are essentially a way to depict our data. Descriptive Statistics are exceptionally critical since in the event that we basically presented our crude information it would be difficult to imagine what the information was appearing, particularly in case there was a part of it. Graphic measurements enhance us to display the information in a more important way, which permits less difficult elucidation of the information. The techniques are Graphical description, tabular description and parametric description.

2.2 Methodology (Type of data)

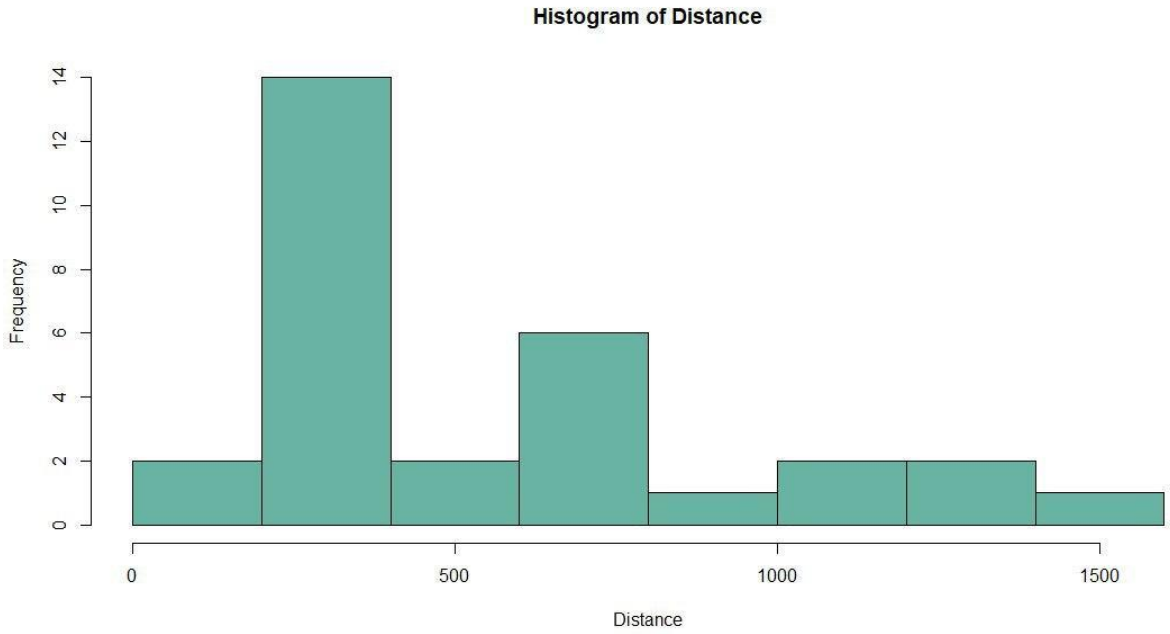
The type of data that we use was *Primary Data* (Google Form). Primary Data is a sort of information that's collected by analysts specifically from fundamental sources through interviews, overviews, tests, etc. Primary Data are ordinarily collected from the source where the data initially begins from and are respected as the most excellent kind of information in research. The sources of essential information are as a rule chosen and custom fitted particularly to meet the demands or necessities of a specific investigate. Moreover, before choosing an information collection source, things just like the point of the investigate and target populace have to be identified.

3.1 Results and Discussion *Table shows the respondent answers

No	Year Study	Block	School	Vehicle	Usage	Oil	Service	Distance	Park
1	FIRST	M22	COMPUTING	MOTOR CYCLE	SOMETIMES	15	ONCE IN A THREE MONTH	750	LESS THAN 12 HOURS
2	FIRST	M23	CIVIL	BICYCLE	ALWAYS	0	NEVER	250	MORE THAN 24 HOURS
3	FIRST	M23	COMPUTING	MOTOR CYCLE	RARELY	10	ONCE IN A THREE MONTH	100	BETWEEN 12 HOURS TO 24 HOURS
4	SECOND	M22	COMPUTING	CAR	ALWAYS	35	ONCE IN A THREE MONTH	250	LESS THAN 12 HOURS
5	FIRST	M22	COMPUTING	MOTOR CYCLE	SOMETIMES	17	ONCE IN A THREE MONTH	500	MORE THAN 24 HOURS
6	FIRST	M23	COMPUTING	BICYCLE	SOMETIMES	0	NEVER	250	MORE THAN 24 HOURS
7	FIRST	M22	COMPUTING	BICYCLE	ALWAYS	10	ONCE IN A THREE MONTH	750	LESS THAN 12 HOURS
8	FIRST	M23	COMPUTING	BICYCLE	RARELY	0	NEVER	250	MORE THAN 24 HOURS
9	FIRST	M22	COMPUTING	MOTOR CYCLE	SOMETIMES	7	ONCE IN A THREE MONTH	900	MORE THAN 24 HOURS
10	FIRST	M23	COMPUTING	BICYCLE	SOMETIMES	5	ONCE IN A THREE MONTH	250	BETWEEN 12 HOURS TO 24 HOURS
11	FIRST	M23	COMPUTING	BICYCLE	RARELY	0	NEVER	500	MORE THAN 24 HOURS
12	FIRST	M22	CIVIL	CAR	ALWAYS	18	ONCE IN A THREE MONTH	750	LESS THAN 12 HOURS
13	FIRST	M22	CIVIL	BICYCLE	SOMETIMES	0	NEVER	250	BETWEEN 12 HOURS TO 24 HOURS
14	FIRST	M22	CIVIL	CAR	RARELY	10	NEVER	250	MORE THAN 24 HOURS
15	THIRD	M23	COMPUTING	MOTOR CYCLE	ALWAYS	40	NEVER	1300	MORE THAN 24 HOURS
16	FIRST	M22	CIVIL	BICYCLE	RARELY	0	NEVER	250	BETWEEN 12 HOURS TO 24 HOURS
17	FIRST	M23	COMPUTING	CAR	RARELY	0	ONCE A MONTH	250	LESS THAN 12 HOURS
18	FIRST	M22	COMPUTING	CAR	SOMETIMES	19	ONCE IN A THREE MONTH	1200	BETWEEN 12 HOURS TO 24 HOURS
19	FIRST	M22	COMPUTING	CAR	SOMETIMES	30	SIX MONTH	250	BETWEEN 12 HOURS TO 24 HOURS
20	SECOND	M22	CIVIL	BICYCLE	ALWAYS	0	NEVER	250	MORE THAN 24 HOURS

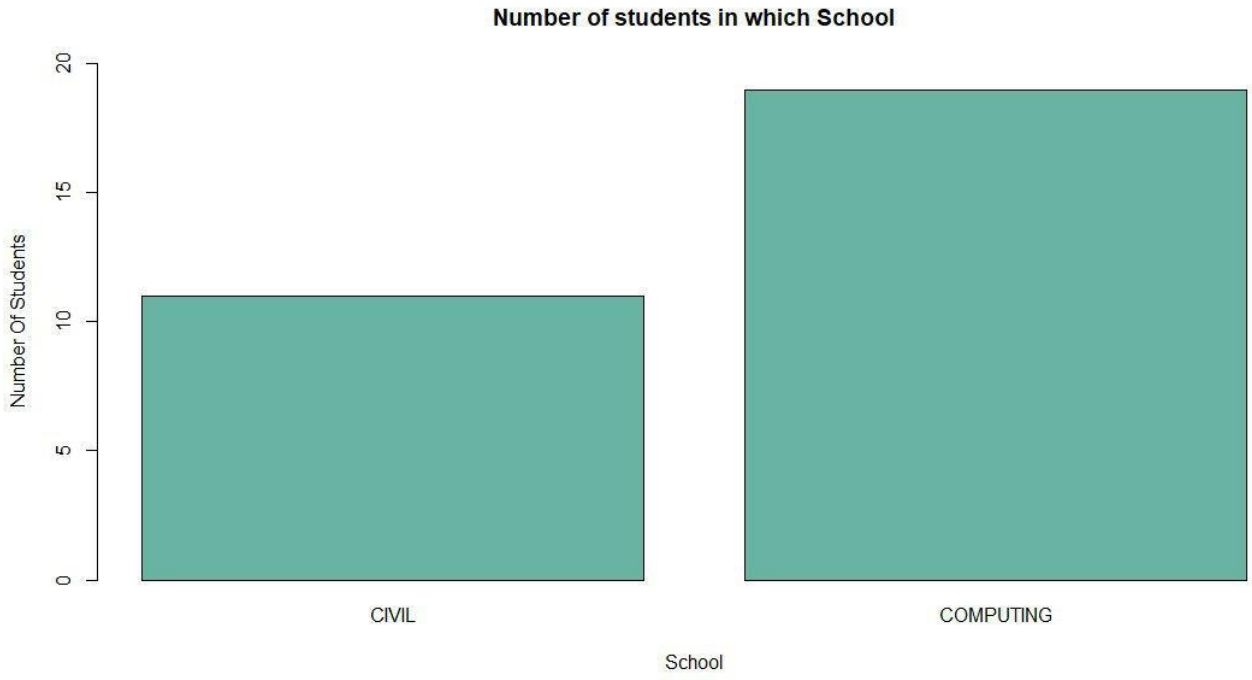
21	SECOND	M22	CIVIL	BICYCLE	ALWAYS	0	NEVER	1100	MORE THAN 24 HOURS
22	THIRD	M23	COMPUTING	MOTORCYCLE	SOMETIMES	6	ONCE IN A THREE MONTH	250	BETWEEN 12 HOURS TO 24 HOURS
23	THIRD	M22	COMPUTING	MOTORCYCLE	SOMETIMES	5	ONCE IN A THREE MONTH	250	BETWEEN 12 HOURS TO 24 HOURS
24	THIRD	M23	CIVIL	MOTORCYCLE	ALWAYS	8	ONCE IN A THREE MONTH	150	BETWEEN 12 HOURS TO 24 HOURS
25	SECOND	M23	CIVIL	CAR	SOMETIMES	35	ONCE A MONTH	250	LESS THAN 12 HOURS
26	FIRST	M22	COMPUTING	BICYCLE	RARELY	0	NEVER	750	MORE THAN 24 HOURS
27	FIRST	M22	COMPUTING	BICYCLE	RARELY	0	NEVER	750	MORE THAN 24 HOURS
28	SECOND	M23	CIVIL	CAR	ALWAYS	40	ONCE IN A THREE MONTH	750	LESS THAN 12 HOURS
29	SECOND	M23	CIVIL	MOTORCYCLE	SOMETIMES	10	ONCE IN A THREE MONTH	1500	BETWEEN 12 HOURS TO 24 HOURS
30	SECOND	M23	COMPUTING	MOTORCYCLE	SOMETIMES	12	ONCE IN A THREE MONTH	1400	BETWEEN 12 HOURS TO 24 HOURS

3.2 Distance



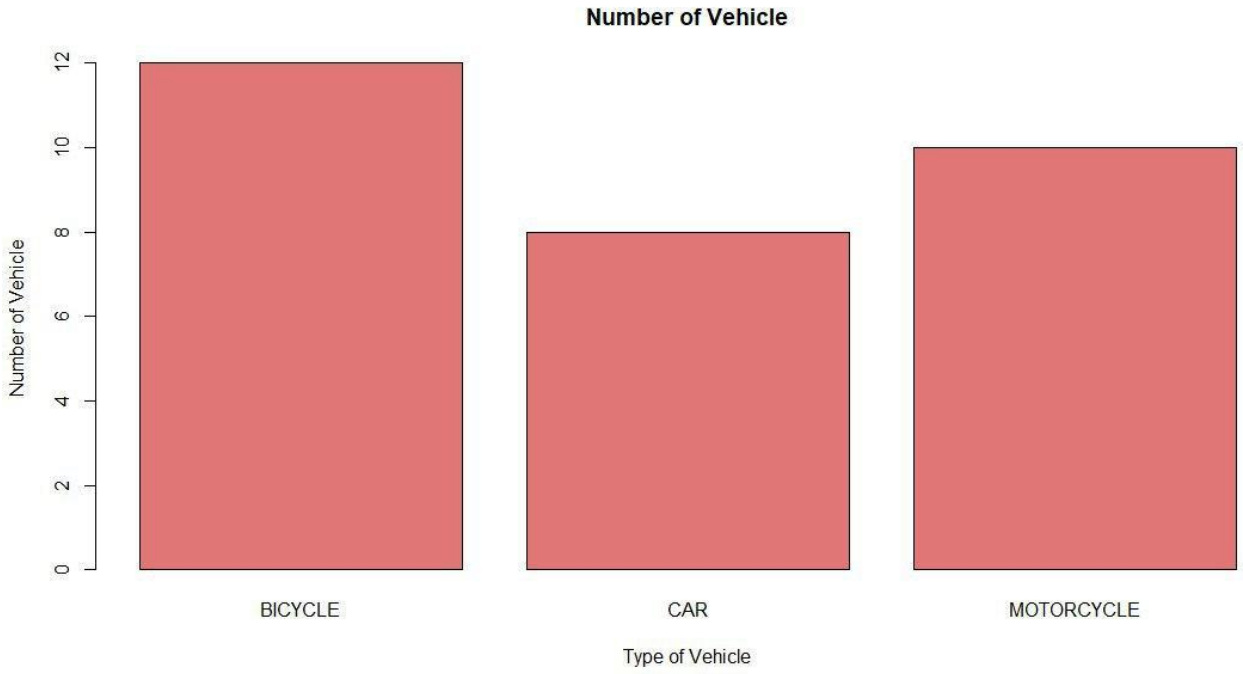
We can say, first of all this data does not distribute symmetrically. This is because, if we cut this graph in the middle we did not get the same mirror image of both half. Other than that, we can see that the mode of this graph is in the range 0 to 500. Note that the distance is in meter (m). This is because, all of our respondent are from M22 and M23. These respondent mostly study in Computer Science and their faculty is just 500m from their hostels. The rest which have the distance more than 500m are from same blocks but their faculty is at Civil Engineering and others. Based on the graph, there is a small number of respondent who have the distance more than 1500m. We can assume with only this data that the respondent's faculty is far from KTDI.

3.3 Number of Students



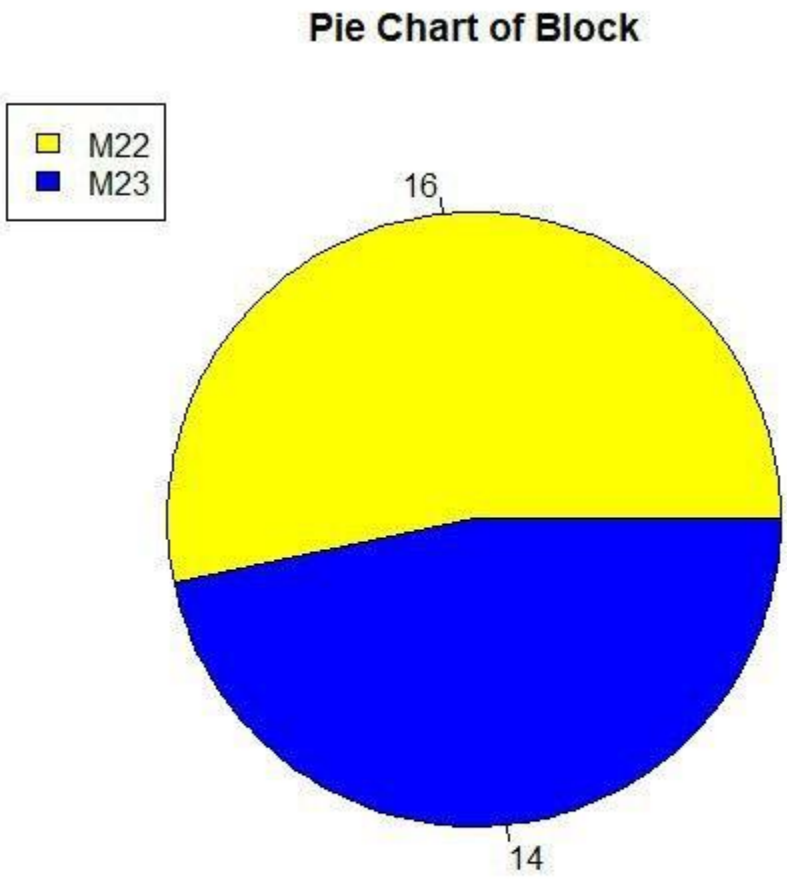
This graph shows the number respondents who chose between civil or computing. Based on this graph, we can see that most of our respondents were from School of Computing. It is because most of student who stay in KTDI were studying at School of Computing which is just 500m away from KTDI. However, this data is not a good evidence for student to bring their vehicles because students can simply walk to the faculty. Does not need any transportation other than UTM Fleet.

3.4 Types of vehicles



This graph shows the number of vehicles that park at the parking lot between block M22 and M23. We can see the highest number is for bicycles, then motorcycles and lastly cars. The number of bicycles can give us the overview that there are many first year students that use bicycle as their transportation.

3.5 Number of Residents (respondents)



This pie chart show the number of respondent who stayed at M22 or M23. From this pie chart, we can say that most of the vehicles park belongs to the students who stayed at M22. But based on the percentage, M22 has 53.33% while M23 has 46.67%. It is just a small difference in number of respondent from M22 and M23 which is M22 with 16 respondents while M23 with 14 respondents.

3.6 Frequency Distribution for oil

Oil spent,RM	Frequency,f	Cumulative frequency	Middle point, x	x^2	fx	fx^2
0-10	20	20	5	25	100	500
11-20	5	25	15.5	240.25	77.5	1201.25
21-30	1	26	25.5	650.25	25.5	650.25
31-40	4	30	35.5	1260.25	142	5041
				TOTAL	345	7392.5

Mean = 11.5

Median = 7.5

Mode = 5

Varience = 118.1034

Std deviation = 10.8675

Kurtosis = 0.945

From these several calculations :

- The average money spent on oil is RM 11.5
- Kurtosis 0.945 with excess -2.055 < 0. It is **PLATYKURTIC**
- The mode that we obtain from the distribution is 5 (RM 5.00). this show us that the students tend to spend RM 5 on oil.

3.7 Frequency Distribution for hours park

Hours	Frequency,f	Cumulative frequency	Middle point,x	x^2	fx	fx^2
0-12	7	7	6	36	42	252
13-24	11	18	18.5	342.25	203.5	3764.75
25-36	12	30	30.5	930.25	366	11163
				TOTAL	611.5	15173.75

Mean = 20.3833

Median = 21

Mode = 30.5

Variance = 93.6325

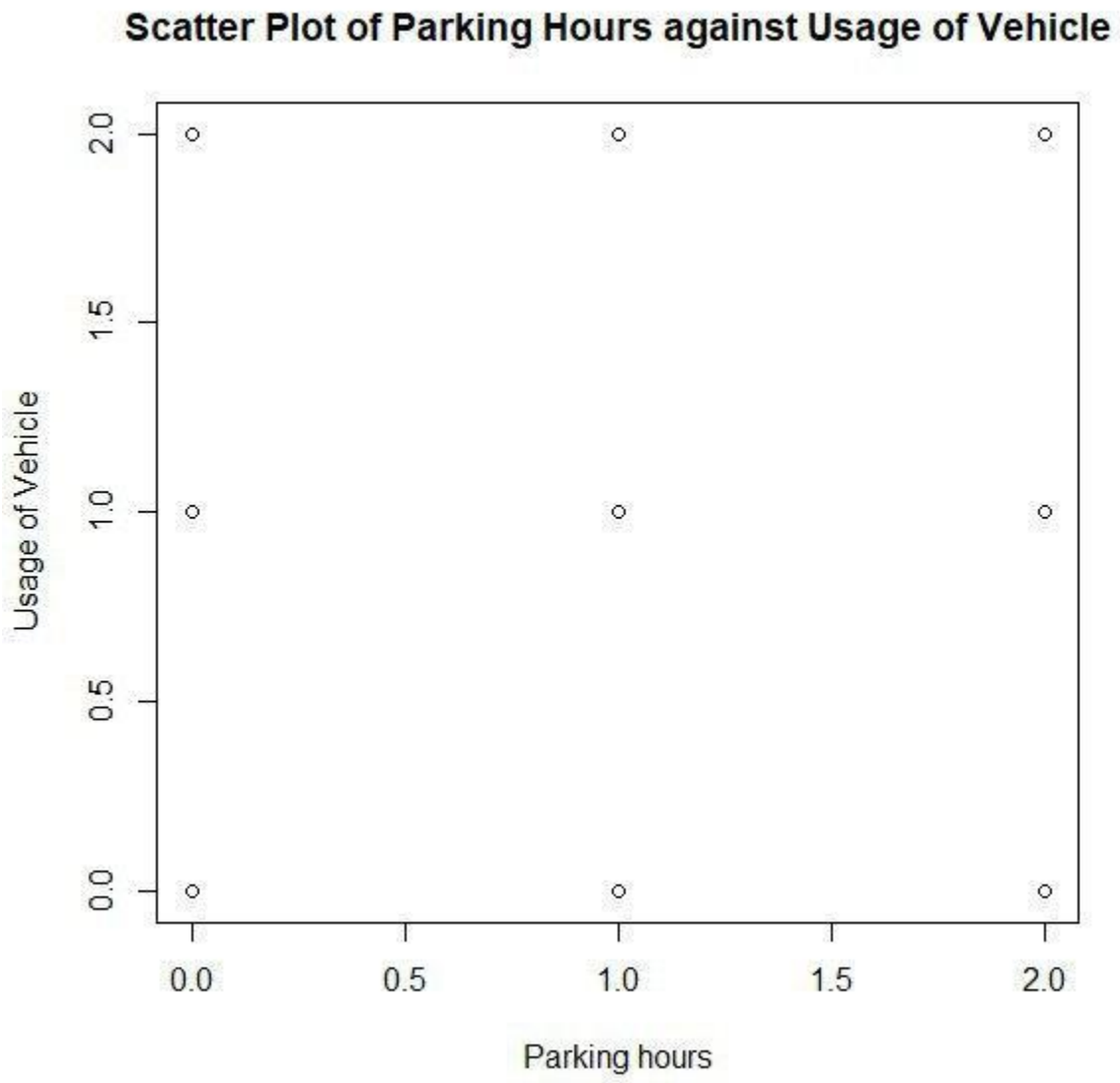
Std deviation = 9.6764

Kurtosis = 1.2936

From these calculations :

- The average hours park at the parking lot is 20.3833 hours.
- The kurtosis 1.2936 with the excess -1.7064. It is **PLATYKURTIC**
- The mode that we obtain from the distribution is 30.5. This show us that there are many respondent who park their vehicles more than more than 30 hours. Maybe that use their vehicles for emergency for for any urgent purpose.

3.8 Scatter Plot



Parking Hours: 0.0 is less than 12 hours ; 1.0 is between 12 hours to 24 hours ; 2.0 is more than 24 hours

Usage of Vehicle: 0.0 is rarely ; 1.0 is sometimes ; 2.0 is always

CONCLUSION

Based on the survey that we did, the data that we receive and many calculations, we can give some conclusion based on the data. Firstly, based on **3.3,** we can see that students from School of Computing more likely to use vehicle then students from School of Civil Engineering. Other than that, based on the graph at **3.4** students from M22 and M23 use bicycle more than car and motorcycle. This shows that the students still follows the law set by the UTM. Finally, we successfully achieve the objectives for this project. We successfully interpret the data that we obtain, and we can see the pattern behind the numbers of vehicles. Other than that, we successfully obtain the information about the first year who bring their own vehicles. Then, we successfully understand the monthly expenses for their vehicles (refer 3.6).

REFERENCE

- *https://www.dummies.com/education/math/statistics/how-to-identify-skew-and-symmetry-in-a-statistical-histogram/*
- *https://www.formpl.us/blog/primary-data*
- *https://conjointly.com/kb/descriptive-statistics/*