BEN PLATT - Student

Northern Regional College | PLA50039683@nrc.ac.uk

Mathematics for IT Practitioners

MAths Assignment 2

Contents

[Introduction 2](#_Toc32570804)

[Hypothesis 2](#_Toc32570805)

[Questionnaire 3](#_Toc32570806)

[Sample 4](#_Toc32570807)

[Collecting the Data 4](#_Toc32570808)

[Mode, Mean, Median, Standard Deviation and Interquartile Range 5](#_Toc32570809)

[Data For staff 6](#_Toc32570810)

[Data for Students 9](#_Toc32570811)

[Conclusion 10](#_Toc32570812)

[Potential Factors that can Influence the Validity of Information 11](#_Toc32570813)

[Further Conclusions and Recommendations 12](#_Toc32570814)

[References 13](#_Toc32570815)

[Appendix A 13](#_Toc32570816)

# Introduction

I have been asked to make a survey about the Staff and Students commute to NRC. I also have to compile the data then analyse it. I am analysing it to identify trends and patterns within the data set. The survey will be an online survey and will take place at the start of a term. I will ask the staff and students to complete this and then work with that data set and examine its validity.

# Hypothesis

A hypothesis is a proposed explanation for a result made on the basis of limited provided evidence. It is usually the best starting point for any investigations.

Staff travel further to work than students.

The reason I think that staff travel further to work than students is that students tend to go to the college nearest to where they live whereas staff will have to go to the college that has jobs available that they can fulfil. This is due to different colleges having different courses that have available places that need filled.

Students will use the bus more often than staff.

The reason I think that Students use the bus more often than staff is that students can avail of a travel pass depending on the distance they have to travel from college. It is also more likely that Staff can drive more than students as they have had longer to practice.

Staff travel for longer than students.

The reason I think that staff travel longer than students to work is because they tend to travel from farther afield and along highly trafficked roads and lanes. Whereas students will travel less as they tend to attend a college closer to them. This is similar to Hypothesis 1.

Students will car share more than staff.

Students will end to Car share more than staff as they will live closer to one another and tend to get along more than staff would as they spend more time together outside of class. They also will maybe only have one or two members of their friend group that can drive. Whereas most staff will be able to drive.

Staff are more likely to travel by car than students.

The reason I think that staff are more likely to travel by is because they are more likely to have a full driver license rather than a provisional as most of the student base will just have come of age to learn to drive.

Staff are more likely to stay later than students.

The reason I think that Staff will stay behind is due to them having to plan and organize future lessons. Student however tend to want to go home to get their assignments done so they can relax and unwind.

# Questionnaire

**The Hypothesis I am trying to prove is staff are more likely to travel by car than students.**

1. Staff or Student?

Possible answers:

Staff

Student

The aim is to sample at least 20 students and 20 members of staff so that sufficient data can be collected from each group to make a valid comparison.

1. How do you travel to college?

Possible answers:

Car

Bus

Walk

Train

Taxi

Cycle

The aim is to collect data about the different methods of travelling to and from college so that we can see and compare how different groups of people travel.

3. If you chose car, Do you car share?

Yes

No

The aim is to see how many students/staff are helping to reduce CO2 emissions and also lets us know that more people that might not have taken the survey also travel by car.

4. How far do you travel to college?(km)

User Input

This is to allow us compare how the distance of the journey affects how the person travels.

5. How early do you arrive before class?(mins)

User Input

This is to allow us to see when the different groups (Staff and Student) arrive to college to see if it affects how they travel.

6. How long do you stay after your final class ends?(mins)

User Input

This is to allow us to see when the different groups (Staff and Student) stay after college ends to see if it affects how long after class finishes they stay.

7. How long does your average journey take?(mins)

User Input

This is to allow us to see if the length of the journey affects the users answers to the previous 2 questions.

8. What days do you attend college?(km)

Monday

Tuesday

Wednesday

Thursday

Friday

This is to allow us to prove that Staff spend longer in college than students.

## Sample

We took a sample of 40 people (20 staff and 20 students) this allows us to have an even mix of both groups of people. They are two distinct groups of people that both go to the same place but for different reasons. There are about 1000 students and 100 staff members in the Newtownabbey Campus alone. This means I have about a 2% reading of students and a 20% of the staff.

## Collecting the Data

The google form was sent out via email and each user will answer all 8 questions and click the submit button. It will then be put into a .csv file and edited into a neater format to be read from and analysed.

# Mode, Mean, Median, Standard Deviation and Interquartile Range

Mean

The mean is the sum off all the numbers divided by the amount of data entries.

Mode

This is the number that appears most often within a set of numbers.

Median

This is the number that falls in the middle of a data set, in order to find this, you will have to list out the data in chronological order then find the value that is in the exact middle. If there are 2 numbers, then you must find the middle of those.

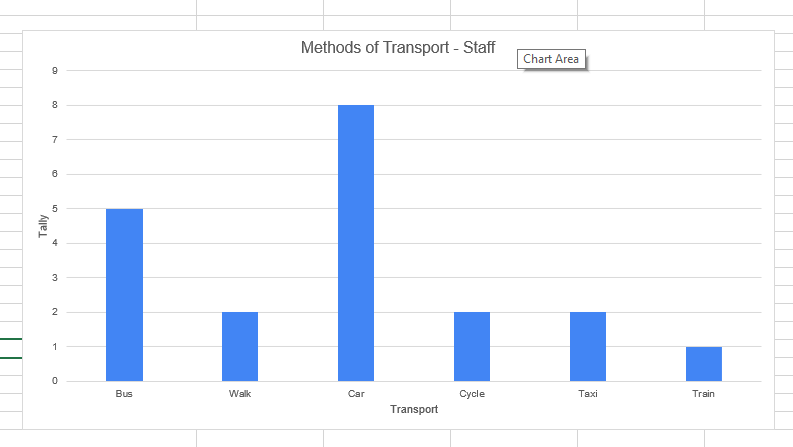
Standard Deviation

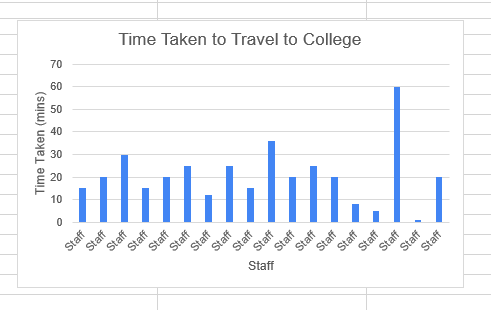
‘In statistics, the standard deviation is a measure of the amount of variation or dispersion of a set of values. A low standard deviation indicates that the values tend to be close to the mean of the set, while a high standard deviation indicates that the values are spread out over a wider range’ (En.wikipedia.org, 2020)

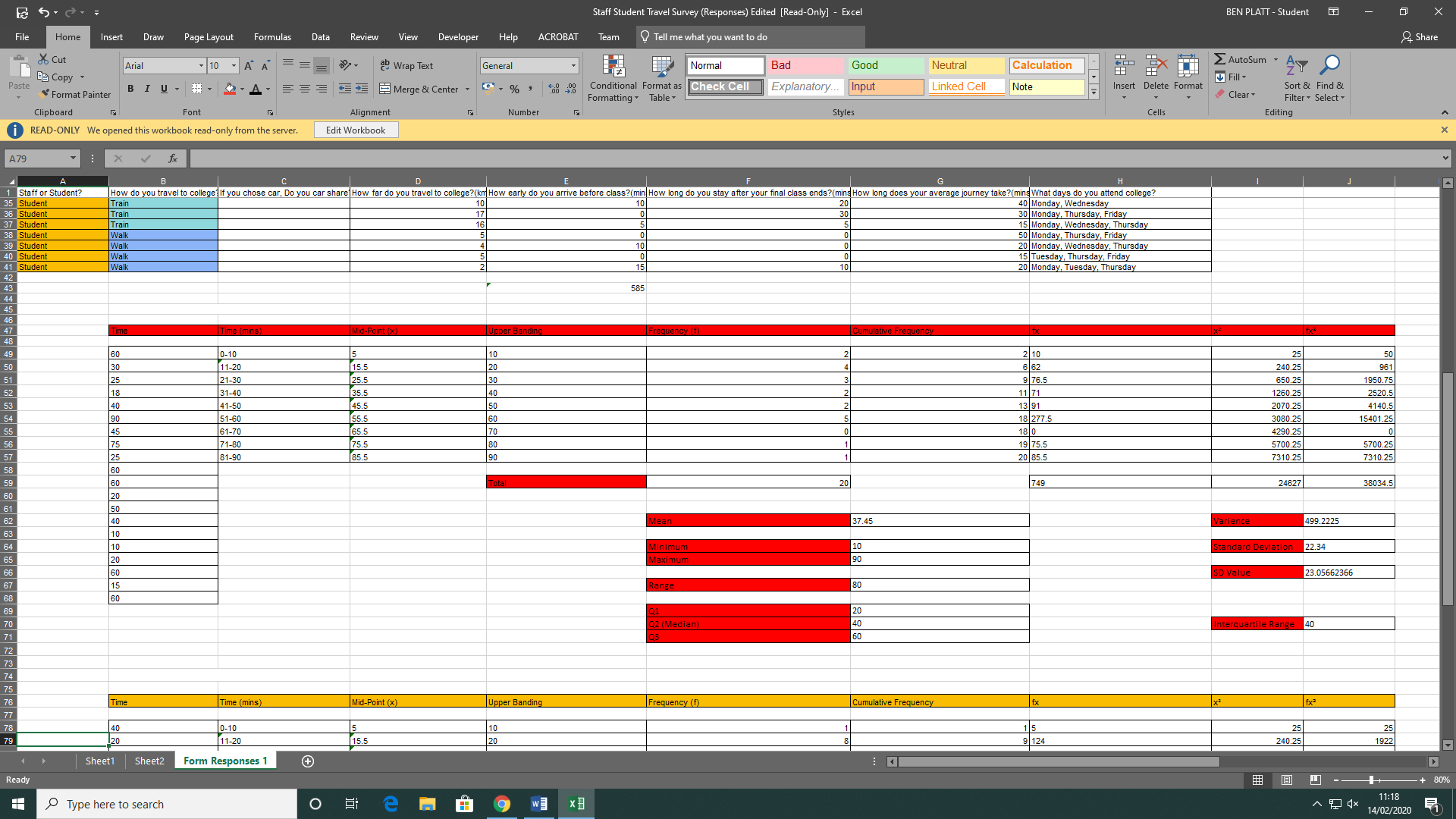
Interquartile Range

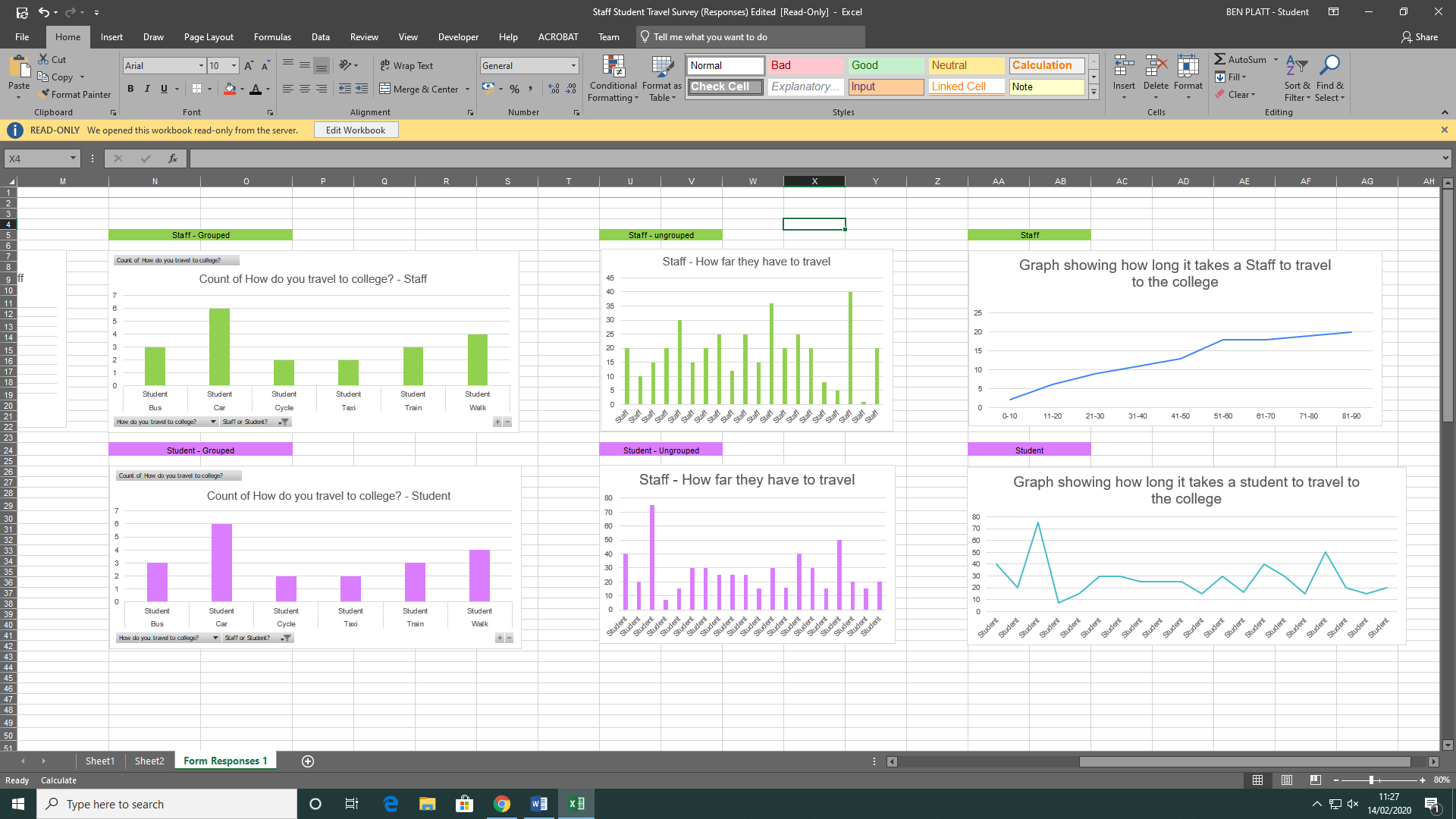
The value for the 1st quartile and the 3rd quartile of the set of data must be calculated first in order to then calculate the interquartile range. Once those values are calculated, you subtract the value of Q1 from Q3

## Data For staff

The bar chart above shows the different methods of transport the staff members take when traveling to the college. The graph shows that majority of the staff members take their car to work. This shows that the staff live at least a significant distance away to have to drive rather than walk or cycle.

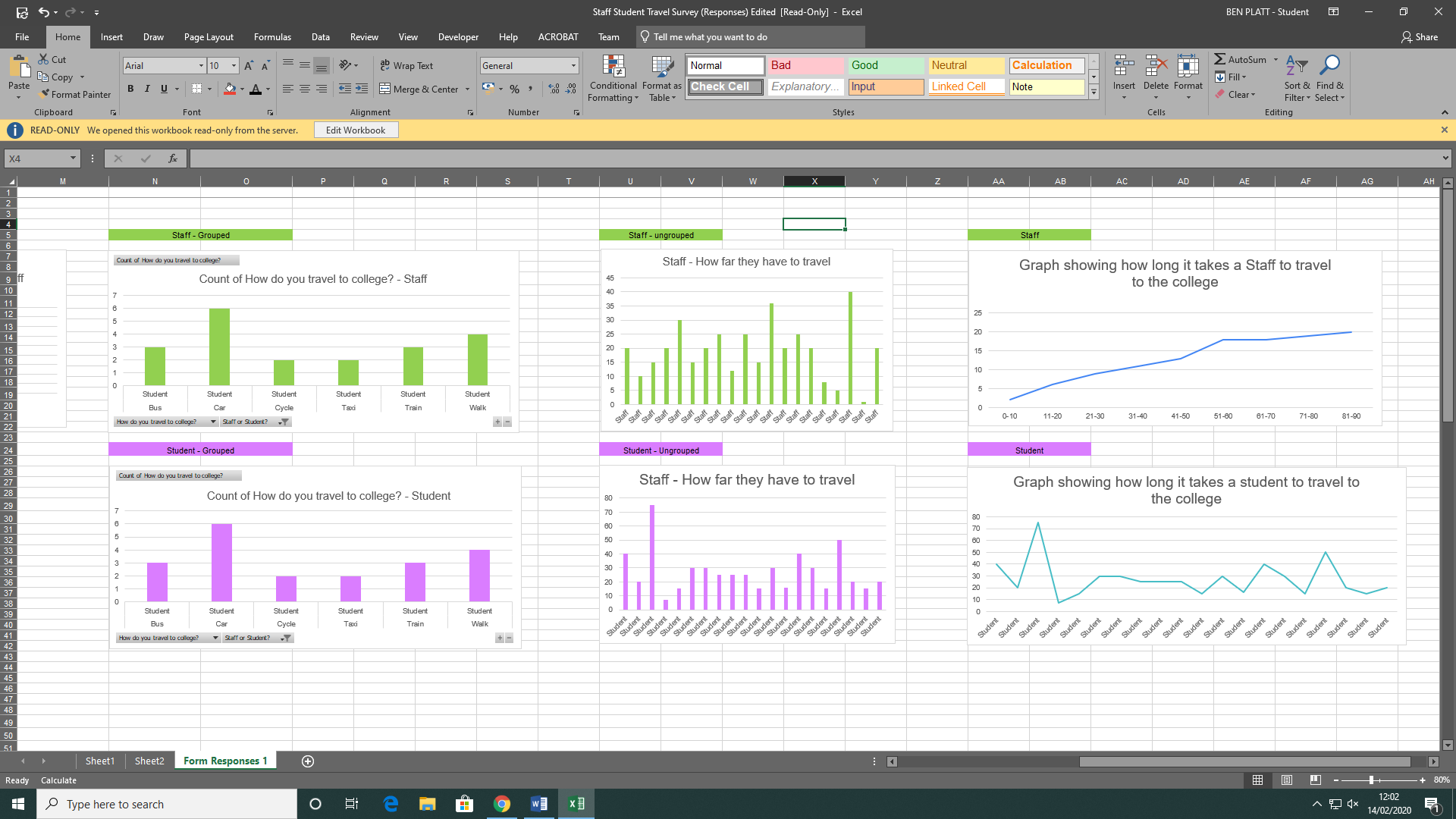
This is an example of ungrouped data. It helps show us that staff tend to take longer than 20 minutes travelling to college.

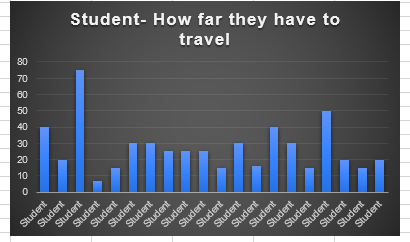


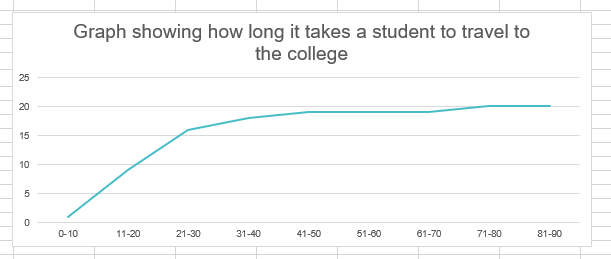
0

*Cumulative Frequency Graph showing Time Taken to Travel to NRC (Staff)*

## Data for Students

The bar chart above shows the different methods of transport the Students take when traveling to the college. The graph shows that majority of the students take their car to college. This shows that the staff live at least a significant distance away to have to drive rather than walk or cycle.

This is an example of ungrouped data. It helps show us that students tend to take longer than 20 minutes but rarely take less than 10 minutes travelling to college.



*Cumulative Frequency Graph showing Time Taken to Travel to NRC (Student)*

# Conclusion

The mean cumulative frequency for the staff was 37.45 minutes and was 24.98 minutes for students. The mean time for staff to travel to work was 12.47 minutes longer than the students. This proves that staff had a longer travel time to the college compared to the students.

The mode cumulative frequency for the staff was 60 minutes and the mode for students was 15. The most common time for students was 15 compared to the 60 minutes of staff members. This shows that the staff members tend to have a longer time travelling than students.

The median cumulative frequency for staff was 40 minutes and 25 minutes for students

The standard deviation for staff was 22.34 and the standard deviation for students was 14.69.

The interquartile range for staff was 40 and the interquartile range for students was 14.25. This tell us that the data given by the staff was more wide spread.

From all the information collected and calculated, it proves that the hypothesis is correct, “Staff travel longer to college than students”.

# Potential Factors that can Influence the Validity of Information

|  |  |
| --- | --- |
| **Factors** | **How they influence validity of information?** |
| Source of data | Primary data is data retrieved by the author themselves and secondary data is  collected by another person. The data I used is Primary data as I made the questionnaire to collect the data for me. |
| Appropriateness | In the questionnaire there was a question at the end, ‘What days do you attend college?’. This question isn’t relevant as it doesn’t have anything to do with the length of time take to travel to college. |
| Discrete or continuous data? | Discreet data is data that can only have certain values, an example would be the question ‘Staff or Student?’. This is because you are one or the other. Continuous Data is a constantly changing data type, an example of this is the question that ask about how long it takes the user to travel to College. |
| Number of variables | A variable is a data value that isn’t fixed or is liable to change. In this the variable was the distance the student/staff member lived from the college. |
| Bias | If a person was biasedly answering their questions it could make the whole data set unreliable. |
| Control groups | A control group is a group of items used to benchmark the other values. An example would be a high-end PC being the benchmark for how well software’s run. |
| Sample Size | If the sample size is too small then it can lead to inaccurate information as you may only have surveyed a group of individuals that walk when the majority drive. |
| Methods to gather data | Poorly designed and worded questions can lead to the wrong answers being received and the whole data set being off topic. |
| Calculations | When results are calculated the numbers tend to contain decimal points. For the most accurate data the numbers should not be rounded up and should remain to 2 decimal points. This can lead toanswers being slightly off. |

# Further Conclusions and Recommendations

If I was to do the questionnaire again one question, I would replace is How many days do you attend college? This is because the question doesn’t help prove the hypothesis either correct or incorrect and is irrelevant. All the other questions are justified in the Questionnaire section as to why it is used within the survey.

Overall, the outcome of this questionnaire proves all 6 pat hypotheses correct. This is due to there being an evident difference in the average travel time of Students and Staff Members. The overall difference was 12.47 minutes. This was calculated by taking one average away from the other.

The day which they completed the survey can affect the validity of my results. One day they could’ve been in extra traffic and thus would’ve had a longer journey time. Another factor could be that they came earlier than usual, and they would’ve had to go through less traffic. This in turn means that their journey time is shorter than usual.

# References

* En.wikipedia.org. (2020). *Standard deviation*. [online] Available at: https://en.wikipedia.org/wiki/Standard\_deviation [Accessed 14 Feb. 2020].

# Appendix A

