

Nick Gelstharp - Assignment 2

by Nicholas Gelstharp

Submission date: 08-Jan-2024 11:17AM (UTC+0000)

Submission ID: 220941331

File name: 40128_Nicholas_Gelstharp_Nick_Gelstharp_-_Assignment_2_547400_806143450.docx (1.21M)

Word count: 1877

Character count: 8339



Statistical Analysis of Retail Store Performance

Nicholas Gelstharp (2211042)
Data Analysis and Visualisation
1/8/24

Contents

Question 1:..... 2

Question 1a)..... 3

Question 1b) 5

Question 1c)..... 11

Question 2a)..... 13

Question 2b..... 14

Question 2c 14

Question 3a..... 15

Question 3b) 16

Question 3c)..... 16

Question 4a)..... 16

Question 4b) 16

Question 5a)..... 18

Question 5b) 19

Question 1:

In order to make the analysis of the information simpler. The first activity that I completed was to enter the information into a table within Excel. I have included a screenshot of this below

Store	Product Category	Daily Sales	Average Customer Rating	Location
1	Electronics	12	4.2	Urban
1	Electronics	15	4.2	Urban
1	Electronics	18	4.2	Urban
1	Electronics	20	4.2	Urban
1	Electronics	20	4.2	Urban
1	Electronics	25	4.2	Urban
1	Electronics	28	4.2	Urban
2	Apparel	10	3.8	Urban
2	Apparel	12	3.8	Urban
2	Apparel	15	3.8	Urban
2	Apparel	18	3.8	Urban
2	Apparel	20	3.8	Urban
3	Home Goods	8	4.1	Rural
3	Home Goods	10	4.1	Rural
3	Home Goods	12	4.1	Rural
3	Home Goods	14	4.1	Rural
3	Home Goods	16	4.1	Rural
3	Home Goods	18	4.1	Rural
4	Electronics	5	3.6	Rural
4	Electronics	8	3.6	Rural
4	Electronics	10	3.6	Rural
4	Electronics	12	3.6	Rural
4	Electronics	15	3.6	Rural
4	Electronics	20	3.6	Rural
4	Electronics	22	3.6	Rural
5	Home Goods	18	4.5	Urban
5	Home Goods	20	4.5	Urban
5	Home Goods	22	4.5	Urban
5	Home Goods	25	4.5	Urban
5	Home Goods	28	4.5	Urban
5	Home Goods	30	4.5	Urban

Question 1a)

The mean is the average of all of the daily sales. In order to calculate the mean I used the Average function in excel selecting the daily sales data as my range. This resulted in a mean figure of 16.94, meaning that across the 5 stores, the average daily sales figure could be rounded up to 17.

The screenshot shows an Excel spreadsheet with a table of daily sales data. The formula bar at the top displays the formula `=AVERAGE(Table1[Daily Sales])` and the result `16.96774`. The table has the following columns: Store, Product Category, Daily Sales, Average Customer Rating, and Location.

Store	Product Category	Daily Sales	Average Customer Rating	Location
1	Electronics	12	4.2	Urban
1	Electronics	15	4.2	Urban
1	Electronics	18	4.2	Urban
1	Electronics	20	4.2	Urban
1	Electronics	20	4.2	Urban
1	Electronics	25	4.2	Urban
1	Electronics	28	4.2	Urban
2	Apparel	10	3.8	Urban
2	Apparel	12	3.8	Urban
2	Apparel	15	3.8	Urban
2	Apparel	18	3.8	Urban
2	Apparel	20	3.8	Urban
3	Home Goods	8	4.1	Rural
3	Home Goods	10	4.1	Rural
3	Home Goods	12	4.1	Rural
3	Home Goods	14	4.1	Rural
3	Home Goods	16	4.1	Rural
3	Home Goods	18	4.1	Rural
4	Electronics	5	3.6	Rural
4	Electronics	8	3.6	Rural
4	Electronics	10	3.6	Rural
4	Electronics	12	3.6	Rural
4	Electronics	15	3.6	Rural
4	Electronics	20	3.6	Rural
4	Electronics	22	3.6	Rural
5	Home Goods	18	4.5	Urban
5	Home Goods	20	4.5	Urban
5	Home Goods	22	4.5	Urban
5	Home Goods	25	4.5	Urban
5	Home Goods	28	4.5	Urban
5	Home Goods	30	4.5	Urban

Looking at the data presented it is clear that store 2 and store 4 have many entries below the mean for the 5 stores across this period with store 4 in particular showing some very low figures in comparison. This indicates that they are likely to be the worst performing stores.

In order to calculate the median is the middle value of all of the data presented. In order to calculate this I used the MEDIAN function within Excel as shown in the screenshot below. This returned a figure of 18.

AutoSave Off Company sales Data.xlsx Nicholas Gelstarp (2211042) 2

File Home Insert Page Layout Formulas Data Review View Automate Help

Clipboard Font Alignment Styles Conditional Formatting Format as Table Cell Styles Cells Editing Add-ins Analyze Data

IS

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		Store	Product Category	Daily Sales	Average Customer Rating	Location							
3		1	Electronics	12	4.2	Urban			16.96774				
4		1	Electronics	15	4.2	Urban			18				
5		1	Electronics	18	4.2	Urban							
6		1	Electronics	20	4.2	Urban							
7		1	Electronics	20	4.2	Urban							
8		1	Electronics	25	4.2	Urban							
9		1	Electronics	28	4.2	Urban							
10		2	Apparel	10	3.8	Urban							
11		2	Apparel	12	3.8	Urban							
12		2	Apparel	15	3.8	Urban							
13		2	Apparel	18	3.8	Urban							
14		2	Apparel	20	3.8	Urban							
15		3	Home Goods	8	4.1	Rural							
16		3	Home Goods	10	4.1	Rural							
17		3	Home Goods	12	4.1	Rural							
18		3	Home Goods	14	4.1	Rural							
19		3	Home Goods	16	4.1	Rural							
20		3	Home Goods	18	4.1	Rural							
21		4	Electronics	5	3.6	Rural							
22		4	Electronics	8	3.6	Rural							
23		4	Electronics	10	3.6	Rural							
24		4	Electronics	12	3.6	Rural							
25		4	Electronics	15	3.6	Rural							
26		4	Electronics	20	3.6	Rural							
27		4	Electronics	22	3.6	Rural							
28		5	Home Goods	18	4.5	Urban							
29		5	Home Goods	20	4.5	Urban							
30		5	Home Goods	22	4.5	Urban							
31		5	Home Goods	25	4.5	Urban							
32		5	Home Goods	28	4.5	Urban							
33		5	Home Goods	30	4.5	Urban							
34													

Sheet1

Ready Accessibility: Good to go 100%

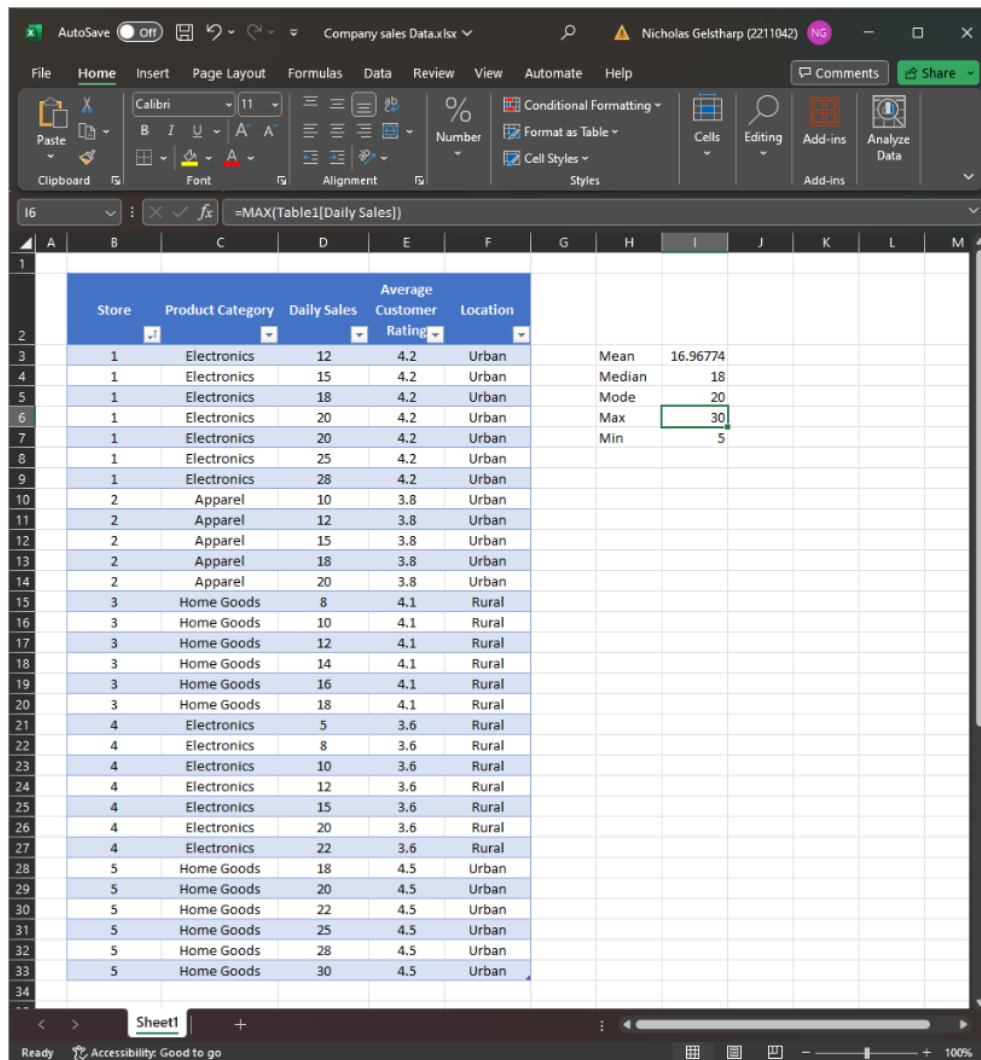
The mode is identifying the number that occurs most regularly in a dataset. In order to calculate the mode I used the MODE.SNGL function within Excel. The answer to this was 20.

3

	Store	Product Category	Daily Sales	Average Customer Rating	Location
1					
2					
3	1	Electronics	12	4.2	Urban
4	1	Electronics	15	4.2	Urban
5	1	Electronics	18	4.2	Urban
6	1	Electronics	20	4.2	Urban
7	1	Electronics	20	4.2	Urban
8	1	Electronics	25	4.2	Urban
9	1	Electronics	28	4.2	Urban
10	2	Apparel	10	3.8	Urban
11	2	Apparel	12	3.8	Urban
12	2	Apparel	15	3.8	Urban
13	2	Apparel	18	3.8	Urban
14	2	Apparel	20	3.8	Urban
15	3	Home Goods	8	4.1	Rural
16	3	Home Goods	10	4.1	Rural
17	3	Home Goods	12	4.1	Rural
18	3	Home Goods	14	4.1	Rural
19	3	Home Goods	16	4.1	Rural
20	3	Home Goods	18	4.1	Rural
21	4	Electronics	5	3.6	Rural
22	4	Electronics	8	3.6	Rural
23	4	Electronics	10	3.6	Rural
24	4	Electronics	12	3.6	Rural
25	4	Electronics	15	3.6	Rural
26	4	Electronics	20	3.6	Rural
27	4	Electronics	22	3.6	Rural
28	5	Home Goods	18	4.5	Urban
29	5	Home Goods	20	4.5	Urban
30	5	Home Goods	22	4.5	Urban
31	5	Home Goods	25	4.5	Urban
32	5	Home Goods	28	4.5	Urban
33	5	Home Goods	30	4.5	Urban
34					

Question 1b)

The range is the difference between the largest and the smallest value within a dataset. I calculated this by using the MAX function to find the highest value within the sales data



I then used the Min formula to find the smallest value within the sales data

AutoSave Off Company sales Data.xlsx Nicholas Gelstarp (2211042)

File Home Insert Page Layout Formulas Data Review View Automate Help

Clipboard Font Alignment Number Conditional Formatting Format as Table Cell Styles

Comments Share

I7: =MIN(Table1[Daily Sales])

	Store	Product Category	Daily Sales	Average Customer Rating	Location		
1							
2							
3	1	Electronics	12	4.2	Urban	Mean	16.96774
4	1	Electronics	15	4.2	Urban	Median	18
5	1	Electronics	18	4.2	Urban	Mode	20
6	1	Electronics	20	4.2	Urban	Max	30
7	1	Electronics	20	4.2	Urban	Min	5
8	1	Electronics	25	4.2	Urban		
9	1	Electronics	28	4.2	Urban		
10	2	Apparel	10	3.8	Urban		
11	2	Apparel	12	3.8	Urban		
12	2	Apparel	15	3.8	Urban		
13	2	Apparel	18	3.8	Urban		
14	2	Apparel	20	3.8	Urban		
15	3	Home Goods	8	4.1	Rural		
16	3	Home Goods	10	4.1	Rural		
17	3	Home Goods	12	4.1	Rural		
18	3	Home Goods	14	4.1	Rural		
19	3	Home Goods	16	4.1	Rural		
20	3	Home Goods	18	4.1	Rural		
21	4	Electronics	5	3.6	Rural		
22	4	Electronics	8	3.6	Rural		
23	4	Electronics	10	3.6	Rural		
24	4	Electronics	12	3.6	Rural		
25	4	Electronics	15	3.6	Rural		
26	4	Electronics	20	3.6	Rural		
27	4	Electronics	22	3.6	Rural		
28	5	Home Goods	18	4.5	Urban		
29	5	Home Goods	20	4.5	Urban		
30	5	Home Goods	22	4.5	Urban		
31	5	Home Goods	25	4.5	Urban		
32	5	Home Goods	28	4.5	Urban		
33	5	Home Goods	30	4.5	Urban		
34							

Sheet1

Ready Accessibility: Good to go

Once I had both of these value I then took the min away from the max to calculate the range which was 25.



AutoSave Off Company sales Data.xlsx Nicholas Gelstarp (2211042)

File Home Insert Page Layout Formulas Data Review View Automate Help

Clipboard Font Alignment Styles Conditional Formatting Format as Table Cell Styles

Cells Editing Add-ins Analyze Data

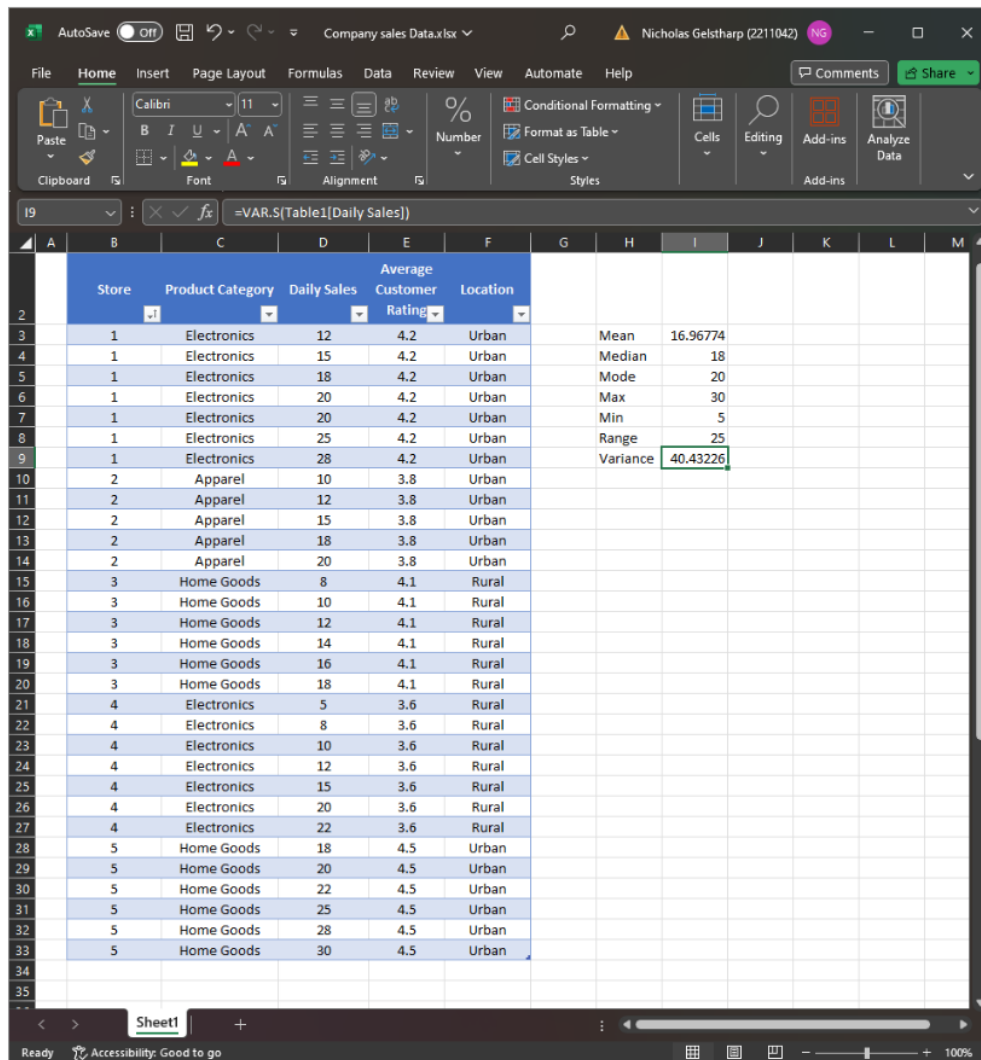
18 =I6-I7

	Store	Product Category	Daily Sales	Average Customer Rating	Location		
1							
2							
3	1	Electronics	12	4.2	Urban	Mean	16.96774
4	1	Electronics	15	4.2	Urban	Median	18
5	1	Electronics	18	4.2	Urban	Mode	20
6	1	Electronics	20	4.2	Urban	Max	30
7	1	Electronics	20	4.2	Urban	Min	5
8	1	Electronics	25	4.2	Urban	Range	25
9	1	Electronics	28	4.2	Urban		
10	2	Apparel	10	3.8	Urban		
11	2	Apparel	12	3.8	Urban		
12	2	Apparel	15	3.8	Urban		
13	2	Apparel	18	3.8	Urban		
14	2	Apparel	20	3.8	Urban		
15	3	Home Goods	8	4.1	Rural		
16	3	Home Goods	10	4.1	Rural		
17	3	Home Goods	12	4.1	Rural		
18	3	Home Goods	14	4.1	Rural		
19	3	Home Goods	16	4.1	Rural		
20	3	Home Goods	18	4.1	Rural		
21	4	Electronics	5	3.6	Rural		
22	4	Electronics	8	3.6	Rural		
23	4	Electronics	10	3.6	Rural		
24	4	Electronics	12	3.6	Rural		
25	4	Electronics	15	3.6	Rural		
26	4	Electronics	20	3.6	Rural		
27	4	Electronics	22	3.6	Rural		
28	5	Home Goods	18	4.5	Urban		
29	5	Home Goods	20	4.5	Urban		
30	5	Home Goods	22	4.5	Urban		
31	5	Home Goods	25	4.5	Urban		
32	5	Home Goods	28	4.5	Urban		
33	5	Home Goods	30	4.5	Urban		
34							

Sheet1

Ready Accessibility: Good to go

Variance is a calculation that can demonstrate how much variance there is from the mean value. This is helpful in identifying how spread out the numbers may be. In order to perform this calculation I used the VAR.S function within Excel as the data represents a sample dataset from a week. The answer was 40.43



Standard Deviation can tell us how dispersed a certain range of data may be. The lower the value, the less dispersed the data within the range is. In order to calculate this I first of all highlighted the sales data and then selected the data tab in Excel and chose Data Analysis

Store	Product Category	Daily Sales	Average Customer Rating	Location		
1	Electronics	12	4.2	Urban	Mean	16.96774
1	Electronics	15	4.2	Urban	Median	18
1	Electronics	18	4.2	Urban	Mode	20
1	Electronics	20	4.2	Urban	Max	30
1	Electronics	20	4.2	Urban	Min	5
1	Electronics	25	4.2	Urban	Range	25
1	Electronics	28	4.2	Urban	Variance	40.43226
2	Apparel	10	3.8	Urban		
2	Apparel	12	3.8	Urban		
2	Apparel	15	3.8	Urban		
2	Apparel	18	3.8	Urban		
2	Apparel	20	3.8	Urban		
3	Home Goods	8	4.1	Rural		
3	Home Goods	10	4.1	Rural		
3	Home Goods	12	4.1	Rural		
3	Home Goods	14	4.1	Rural		
3	Home Goods	16	4.1	Rural		
3	Home Goods	18	4.1	Rural		
4	Electronics	5	3.6	Rural		
4	Electronics	8	3.6	Rural		
4	Electronics	10	3.6	Rural		
4	Electronics	12	3.6	Rural		
4	Electronics	15	3.6	Rural		
4	Electronics	20	3.6	Rural		
4	Electronics	22	3.6	Rural		
5	Home Goods	18	4.5	Urban		
5	Home Goods	20	4.5	Urban		
5	Home Goods	22	4.5	Urban		
5	Home Goods	25	4.5	Urban		
5	Home Goods	28	4.5	Urban		
5	Home Goods	30	4.5	Urban		

Descriptive Statistics ? X

Input

Input Range:

Grouped By: ☒ Columns ☐ Rows

☐ Labels in First Row

Output options

☐ Output Range:

☒ New Worksheet Ply:

☐ New Workbook

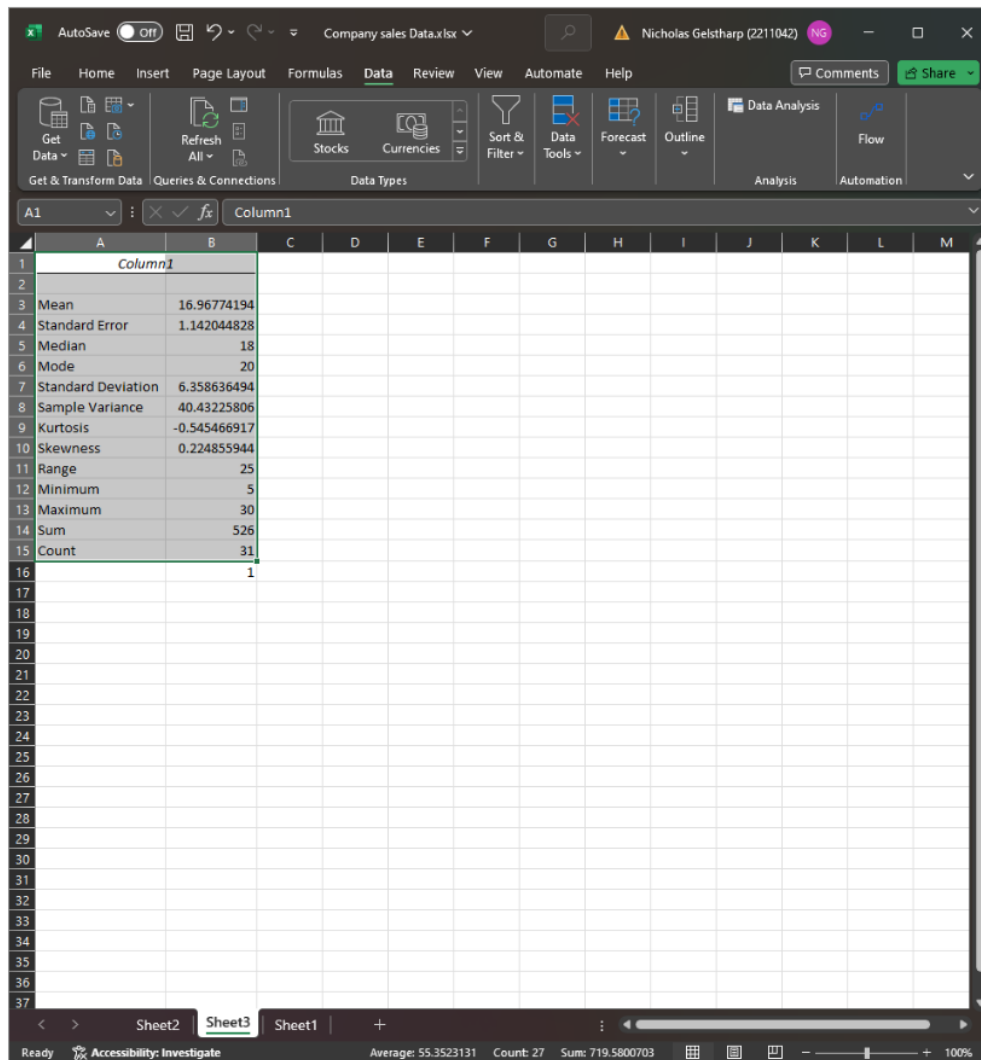
☒ Summary statistics

☐ Confidence Level for Mean: %

☐ Kth Largest:

☐ Kth Smallest:

This then calculated a variety of different descriptive statistics that confirmed the standard deviation is 6.36. Through using this analysis I was also able to check that the previous methods of calculation that I had used were also accurate.



Question 1c)

Using these techniques I am able to see clearly that there are some significant outliers within the data. For example the Range of 25 show in comparison to the Max of 30 and the min of 5 shows that there is a wide range of data. The sale of 5 for store 4 is significantly under the mean and median. Of the 6 worst performing days store 4 appears 3 times, as do electronics in the product category. 5 out of the 6 stores for the worst performing days were also based in rural locations.

Store	Product Category	Daily Sales (in thousands of pounds)	Average Customer Rating (on a scale of 1 to 5)	Location
4	Electronics	5	3.6	Rural
3	Home Goods	8	4.1	Rural
4	Electronics	8	3.6	Rural
2	Apparel	10	3.8	Urban
3	Home Goods	10	4.1	Rural
4	Electronics	10	3.6	Rural

Looking at the other end of the scale store 5 appears in the top 6 for daily sales metrics 4 times and all of the stores in this section are based in Urban locations which strongly indicates that the urban stores are by far the best performing for daily sales.

5	Home Goods	22	4.5	Urban
1	Electronics	25	4.2	Urban
5	Home Goods	25	4.5	Urban
1	Electronics	28	4.2	Urban
5	Home Goods	28	4.5	Urban
5	Home Goods	30	4.5	Urban

You can also see that Electronics and home goods appears at both ends of this scale, which means that we can likely conclude that the relevance of location is playing a significant role in the results being achieved by the stores.

Question 2a)

In order to calculate the 75th percentile of the daily sales for each store you must first organise the data in order from smallest to largest. I have done this in the table below

Store 1	Store 2	Store 3	Store 4	Store 5
12	10	8	5	18
15	12	10	8	20
18	15	12	10	22
20	18	14	12	25
20	20	16	15	28
25		18	20	30
28			22	

Now that I have the data ordered I can perform the calculation . In order to do this I must calculate the index which is $(k / 100) * (n - 1) + 1$



K represents the percentile I am looking for

N is the number of entries within the dataset

So the calculation for store 1 is $(75/100) * (7-1) + 1$ the result of which is 5.5. This means that the 75th percentile for store 1 is entry 6 in the dataset (25) for store 1. I have performed this calculation for each of the other store below.

$$\text{Store 2} = (75/100) * (5-1) + 1 = 4$$

Store 2 is anything beyond and including the 4th entry in the dataset which is 18

$$\text{Store 3} = (75/100) * (6-1) + 1 = 4.75$$

Store 3 is anything beyond and including the 5th entry in the dataset which is 16

$$\text{Store 4} = (75/100) * (7-1) + 1 = 5.5$$

Store 4 is anything beyond and including the 6th entry in the dataset which is 20

$$\text{Store 5} = (75/100) * (6-1) + 1 = 4.75$$

Store 5 is anything beyond and including the 5th entry in the dataset which is 28

$$\text{Doing this for all stores} = (75/100) * (38-1) + 1 = 28.25$$

All stores is anything beyond and including the 29th entry in the dataset which is 20

This figure of 20 is very useful in analysing the performance of the other stores and even in comparing the 75th percentile for each as I can now already see that stores 1,5 and 4 have a larger 75th percentile than the stores as a whole which means that they are likely to be the best performing stores without having to look through the data in detail. This number of 20 is useful to indicate on a daily basis what a good day would be in comparison to the stores as a whole, whereas the individual one may be better for saying when an individual store has done better. This sort of calculation can sometimes be used for bonuses and things of this nature.

Question 2b



Using the all stores 75th percentile figure of 20, I can now see the store that stores 1, 4 and 5 have higher 75th percentile figures for the period that is being analysed. This means that they are beating the overall 75th and are the higher performing stores. Store 2 and 3 therefore are the worst performing stores.

Question 2c

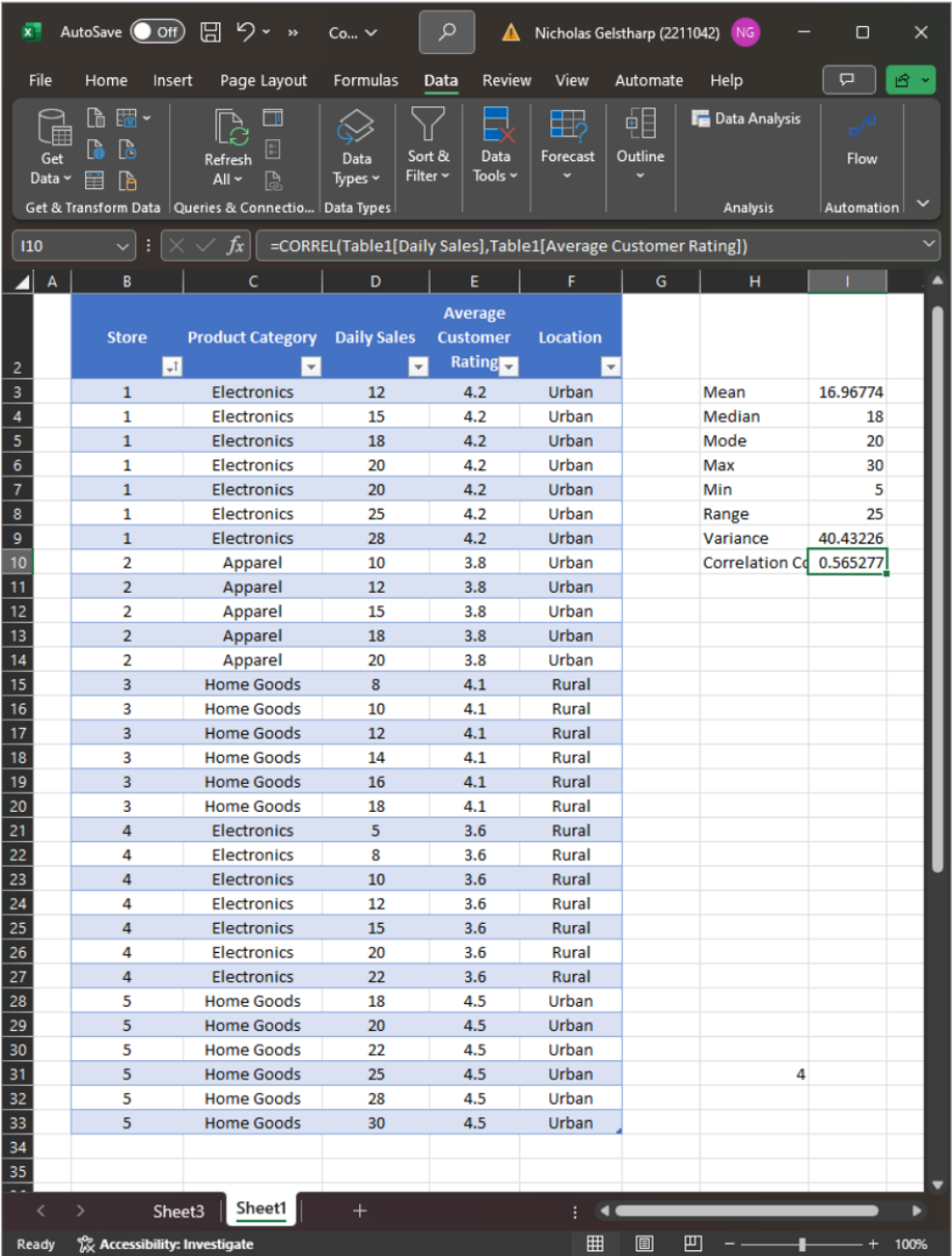
The implication of finding the 75th percentile can be useful in the sales strategy. As I have discussed previously this can be useful for knowing when a daily sales figure has beaten this percentile figure compared to all stores, or just for its own figures. This can be very useful in setting targets as for example for store 3 a daily sales target of 16 would be appropriate as this would be a strong daily sales target compared to historic figures. However, this figure would represent a very poor daily sales for store 5, where a target of 28 sales would be more appropriate. In this way individual targets can be set for performance for each of the branches that are achievable within their settings and target markets.

This is also useful in analysing the sales performance based on product and location as each will have separate percentile figures which can be further manipulate to generate sales target data. Also this does indicate that urban stores are performing much better, however, this is not likely to be the full picture as with urban settings there is likely to be a higher cost in renting stores etc and you would likely require that the sales for these stores to be significantly higher than their cheaper to run rural counterparts. The overhead costs would be needed to determine whether or not it would be to invest in rural or urban settings should the business be looking to expand.



Question 3a

The correlation coefficient is a measurement of the relationship between two datasets, essentially explaining how one dataset is likely to affect another. In order to calculate the correlation coefficient between the average customer rating and the sales for each store I used the CORREL function within Excel as seen in the screenshot below



Question 3b)

This means that overall there is a moderate positive relationship between the two datasets across the dataset as a whole with an overall result of 0.56. The result of the correlation coefficient will always be between 0 and 1, with 0 indicating that there is no relationship between the two datasets and a result of 1 being a perfect correlation between the two dataset. A result of 0.56 is in the middle and indicates that although there is a positive result, it is not strong enough to be able to determine that the results are definitely related and one will affect the other. This means that a more positive customer rating is likely to result in higher daily sales figures but it is very much not guaranteed.

Question 3c)

I would recommend that in order to improve Sales performance that each store is measured on it's own statistics and that consumer feedback and sales targets are generated from that information. As I have discussed, setting targets for each of the stores based on an overall figure seems unreasonable due to the large variance in the figures depending on the location. This recommendation is based on the assumption that each store has different running costs and that the organisation is able to sustain lower figures for certain stores, especially those based in rural areas. I think that the data for consumer ratings should also be analysed on a daily basis instead of taking an average so that the impact of higher sales can be sought on customer service and any potential outliers can be identified so that we know the figures aren't skewed in anyway.

Question 4a)

The null hypothesis for this would be that there is no significant difference in sales performance between two groups of stores.: those located in urban areas and those in rural areas. An alternative hypothesis would be that home goods sell better in urban areas.

Question 4b)

In order to perform a t-test on the data I used the data analysis tab within excel

D16																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										</
-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

t-Test: Paired Two Sample for Means

Input

Variable 1 Range:

Variable 2 Range:

Hypothesized Mean Difference:

☐ Labels

Alpha:

Output options

☐ Output Range:

☒ New Worksheet Ply:

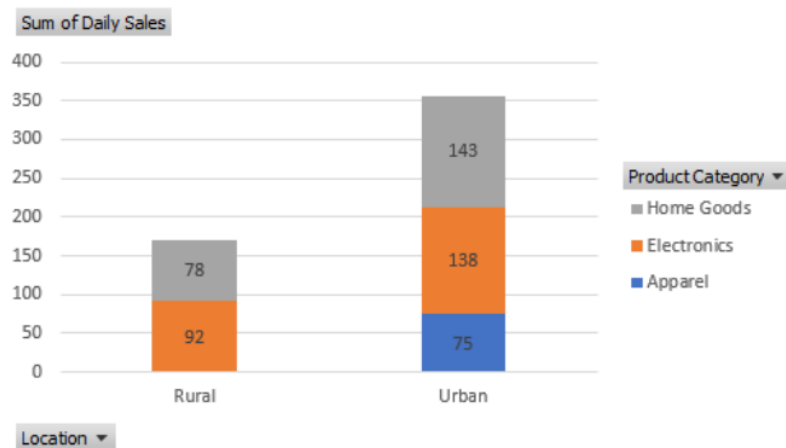
☐ New Workbook

OK Cancel Help

This gave the following results which include the test results, tstat, degrees of freedom (df) and P stats as required.

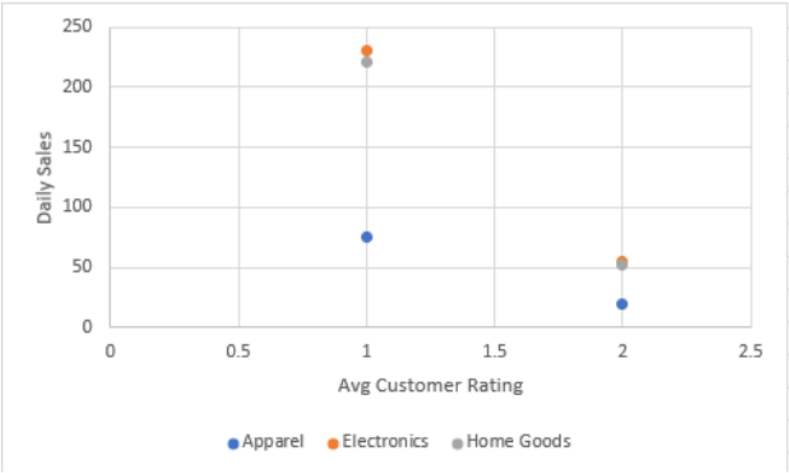
	A	B	C
1	t-Test: Paired Two Sample for Means		
2			
3		Variable 1	Variable 2
4	Mean	13.07692308	17.76923077
5	Variance	25.24358974	26.19230769
6	Observations	13	13
7	Pearson Correlation	0.250291831	
8	Hypothesized Mean Difference	0	
9	df	12	
10	t Stat	-2.724372854	
11	P(T<=t) one-tail	0.009228894	
12	t Critical one-tail	1.782287556	
13	P(T<=t) two-tail	0.018457788	
14	t Critical two-tail	2.17881283	
15			

Question 5a)



The stacked bar chart above shows the split of sales depending on rural or urban locations. From this information it is clear that no apparel sales are made in the rural settings and that in all categories the urban stores outperform their rural counterparts by significant daily sales amounts.

Question 5b)



Nick Gelstharp - Assignment 2

GRADEMARK REPORT

FINAL GRADE

GENERAL COMMENTS

58/100

Question 1	7
Question 2	14
Question 3	12
Question 4	12
Question 5	7
Writing Skills	6

PAGE 1

PAGE 2

PAGE 3

PAGE 4



Comment 1

the values are to be calculated and presented here

PAGE 5



Comment 2

screenshots to be cropped to highlight the relevant sections



Comment 3

the mean, median, and mode collectively provide insights into the distribution of daily sales for each store, helping to understand the central tendency and variability in the data.

PAGE 6

PAGE 7



Comment 4

these measures should be calculated for each store



Comment 5

Likewise outliers should be identified using IQR or z score method



Comment 6

no comment



Comment 7

Q3 values for daily sales in each store provide actionable insights for tailoring the company's sales strategy.



Comment 8

Q3 of 20 indicates
that 75% of the daily sales in Store 1 are below 20.



Comment 9

in addition, strategies to boost sales on days with historically lower performance could be explored. also

Consider
diversifying product offerings or implementing targeted promotions



Comment 10

more specifics should be included

PAGE 17



Comment 11

well analysed



Comment 12

strategies should be proposed. example: Explore collaborations or partnerships with influencers or other businesses to increase visibility and credibility. or Establish a feedback mechanism to continuously monitor and address customer concerns.



Comment 13

hypothesis should be shown in corresponding notations



Comment 14

There is a significant difference in sales performance between the two groups of stores-alternate hypothesis

PAGE 18



Comment 15

a brief explanation is expected

PAGE 19

PAGE 20
