

Mark awobokun 2311796

by Mark Awobokun

Submission date: 12-Jan-2024 08:35AM (UTC+0000)

Submission ID: 221683436

File name: 51099_Mark_Awobokun_Mark_awobokun_2311796_547400_767867913.docx (239.09K)

Word count: 954

Character count: 4873

Data analysis and Visualisation

Descriptive statistics

Mark Awobokun (2311796)

Contents

Mark Awobokun (2311796) 0

Learning in a Digital Era..... **Error! Bookmark not defined.**

E-portfolio **Error! Bookmark not defined.**

Task 1.....  1 **Error! Bookmark not defined.**

Task 2..... **Error! Bookmark not defined.**

Task 3..... **Error! Bookmark not defined.**

Task 4..... **Error! Bookmark not defined.**

Part A

Question 1



- For store 1 the median is 20000, the mode is 20000 and the mean is 19714.29. For store 2 the median is 15000, the mean is 15000 and there is no mode as no number in daily sales comes up more than once. For store 3 the median is 13000, the mean is 13000 and there is also no median for store 3. For store 4 the median is 12000, the mean is 13142.86 and there is no mode for store 4. For store 5 the median is 23500, the mean is 23833.33 and there is no mode for store 5. The mean is a method used to calculate a value of a set of numbers which is done by adding all numbers involved and dividing by the number of numbers involved, it is also a way to gauge performance over time, showing that store 5 has the highest mean meaning it is the best performing store for home goods. The median is the middlemost number among values arranged by numerical order and if 2 numbers are in the middle they are added together and divided by 2. The mode is the highest occurring number in a set of data for example in the list given only store 1 has a mode.
- The range is the highest value in the data set minus the lowest value in the data set, the range for store 1 is $28000 - 12000 = 16000$, the variance is 30238095.24 and the standard deviation is 5498.9176. for store 2 the range is 10000, the variance is 17000000 and the SD is 4123.1056. for store 3 the range is 10000, the Var is 14000000 and the SD is 3741.6574. for store 4 the range is 17000, the Var is 38809523.81 and the SD is 6229.7290. for store 5 the range is 12000, the Var is 21766666.67 and the SD is 4665.4760. the variance is used to find the spread of data against the mean value as well as the SD. If the Sd is larger than the mean the data set is spread far from the mean and more variability, whilst when the SD is lower than the mean the data set is tighter to the mean and less variability occurs.
- Stores 1,2,3 and 4 have no outliers. Store 5 has outliers on 18000 daily sales and 20000 daily sales. The outliers could lead to misleading interpretations on the relationship with the variables, also to identify the outliers I found the first quartile, the third quartile, the interquartile, the upper bound, lower bound and then used the OR function to find any outliers.
- Note all decimal places are to 4 points for degree of accuracy for the SD while all others are 2 points of degree.

Question 2



- The third quartile for store 1 is 22500, for store 2 is 18000, for store 3 is 15500, for store 4 is 17500 and for store 5 is 27250. Store 5 has the highest third quartile. The third quartile means that 75 percent of the data points are found when arranged in increasing order which also means most data points are below or equal to this and since store 3 has the lowest quartile for home goods it indicates it as the least performing store for home goods and store 4 has a lower third quartile than store 1 means it performs or sells less electronics than store 1 making stores 1,2 and 5 the highest performing store in each of their products.

- The 75th percentile for electronics is 20000 making store 1 the only performing store for electronics, the 75th percentile for home goods is 22750 making store 5 the only performing store for home goods.
- The company can decide to supply more product to the stores performing well and reduce the prices for the underperforming stores to urge them to sell more.

Section B

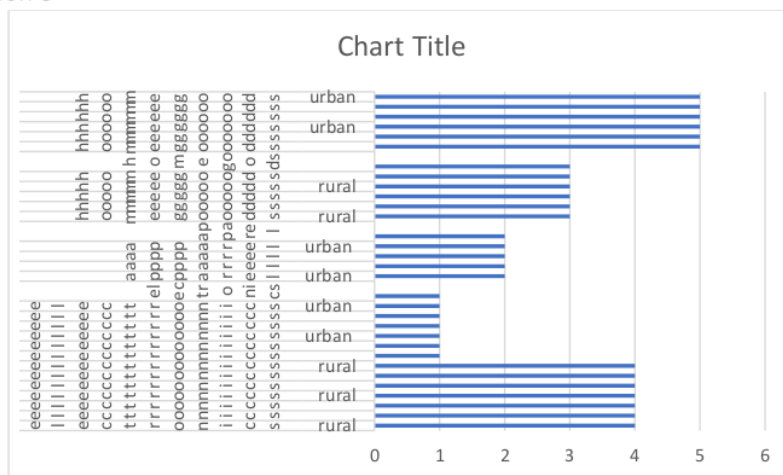
Question 3

- The correlation coefficient is 0.5653.
- If the correlation coefficient cannot be 1 or be less than -1. So, if the correlation is positive, it shows that it has a positive relationship with the data and since it is mid-way of 1 it is a strong relationship.
- To improve customer rating, we offer incentives to them, a referral program can be established, leveraging moments of customer happiness and to improve sales performance we need to have the right mindset to work, improving business acumen and setting concrete goals.

Question 4

- The null hypothesis would be if the mean of the 2 populations is the same while the alternative hypothesis is when the difference between the 2 populations means is not zero. Thus, for null $\mu_1 = \mu_2$ and for alternative hypothesis $\mu_1 \neq \mu_2$
- My t-test value was 0.003527

Question 5



References



- [1] H. Team. "15 Proven Strategies to Improve Sales Performance." Highspot.
<https://www.highspot.com/en-gb/blog/15-proven-strategies-to-improve-sales-performance/>
(accessed Jan 12,).
- [2] "Customer experience (CX) | Teradata." <https://www.teradata.com/solutions/customer-experience> (accessed Jan 12,).
- [3] "11. Correlation and regression | The BMJ." The BMJ | The BMJ: leading general medical journal. Research. Education. Comment. <https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression> (accessed Jan 12,).
- [4] "Interpret all statistics for Store Descriptive Statistics." <https://support.minitab.com/en-us/minitab/20/help-and-how-to/statistics/basic-statistics/how-to/store-descriptive-statistics/interpret-the-statistics/all-statistics/#q3> (accessed Jan 12,).
- [5] S. Hirol, "Navigating Outliers for Accurate Data Analysis & Decisions",
<https://statusneo.com/navigating-outliers-for-accurate-data-analysis-decisions/> (accessed Jan 12, 2024).
- [6] "Mode: What It Is in Statistics and How to Calculate It." Investopedia.
<https://www.investopedia.com/terms/m/mode.asp> (accessed Jan 12,).
- [7] "Median: What It Is and How to Calculate It, With Examples." Investopedia.
<https://www.investopedia.com/terms/m/median.asp> (accessed Jan 12,).
- [8] "What Is a Mean? Definition in Math and Formula for Calculation." Investopedia.
<https://www.investopedia.com/terms/m/mean.asp> (accessed Jan 12,).

FINAL GRADE

GENERAL COMMENTS

22/100

Question 1	4
Question 2	5
Question 3	7
Question 4	2
Question 5	2
Writing Skills	2

PAGE 1

PAGE 2



Comment 1

wrong module contents. proof reading before uploading is recommended

PAGE 3



Comment 2

clear labelling of subsections would help in easy navigation



Comment 3

mean, median,
and mode collectively provide insights into the distribution of daily sales for
each store, helping to understand the central please refer to lecture notes available in Moodle
for more clarification

tendency and variability in the
data.



Comment 4

the range, variance, and standard deviation capture the variability and spread of the daily sales data.



Comment 5

well analysed



Comment 6

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

PAGE 4



Comment 7

good finding, however, Focus could be on customer engagement and satisfaction to sustain consistent sales



Comment 8

more specifics of how the value is achieved should have been included



Comment 9

strength of relationship is wrongly judged



Comment 10

well thought



Comment 11

not well explained



Comment 12

Interpret the test results, including the t-statistic, degrees of freedom, and p-value

PAGE 5



Comment 13

IEEE referencing should be followed. creditable sources should have been included