

Shop Sales Data Report

1. Introduction:

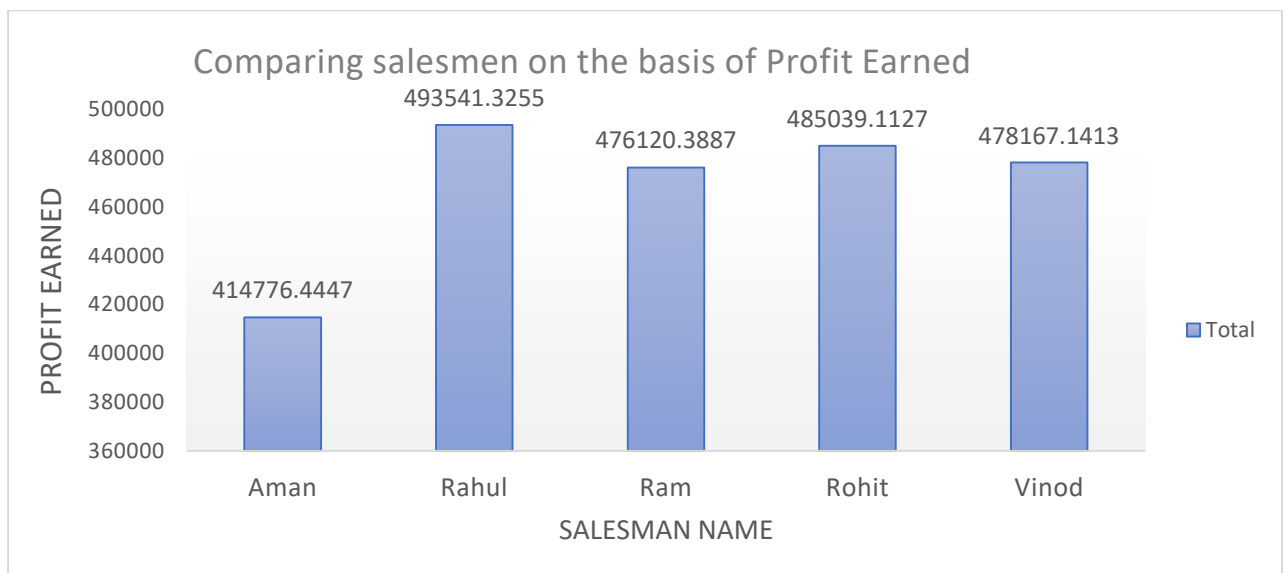
This dataset contains a plethora of information on sales transactions, offering valuable insights into the intricacies of retail operations. With meticulously crafted columns capturing crucial aspects of each transaction such as Date, Salesman, Item Name, Company, Quantity, and Amount, analysts and businesses have access to a wealth of actionable data. Whether it's identifying trends, optimizing inventory management, or refining sales strategies, this dataset serves as an invaluable resource for making informed decisions and exploring new paths for growth.

2. Questionnaire:

- Q1. Compare all the salesmen on the basis of profit earn.
- Q2. Find out most sold product over the period of May-September.
- Q3. Find out which of the two product sold the most over the year Computer or Laptop?
- Q4. Which item yield most average profit?
- Q5. Find out average sales of all the products and compare them.

3. Analytics:

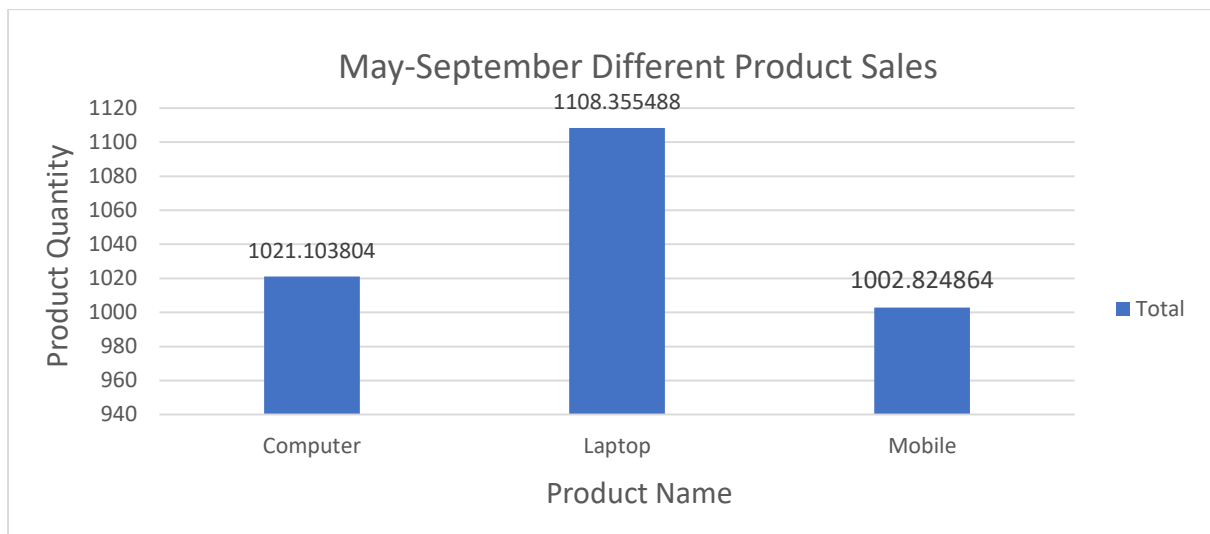
Q1. Compare all the salesmen on the basis of profit earn.



Answer:

Aman earned a profit of \$43,476.4, Ram earned -\$476,120.38 (indicating a loss), Rohit earned \$485,039.11, and Vinod earned \$478,167.14. Rahul earned \$493,541.37, making him the salesman with the highest profit.

Q2. Find out most sold product over the period of May-September.

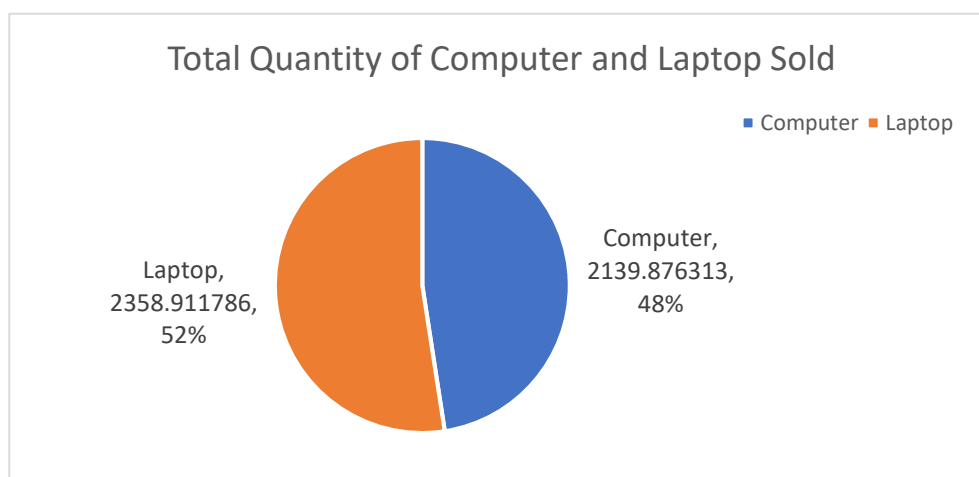


Date	Qty	Item Name
01-05-2019	14	Computer
02-05-2019	14	Laptop
03-05-2019	14	Mobile
04-05-2019	15	
05-05-2019	15	
06-05-2019	15	
07-05-2019	15	
08-05-2019	15	

Answer:

After filtering out the months from May to September, laptops recorded the highest sales.

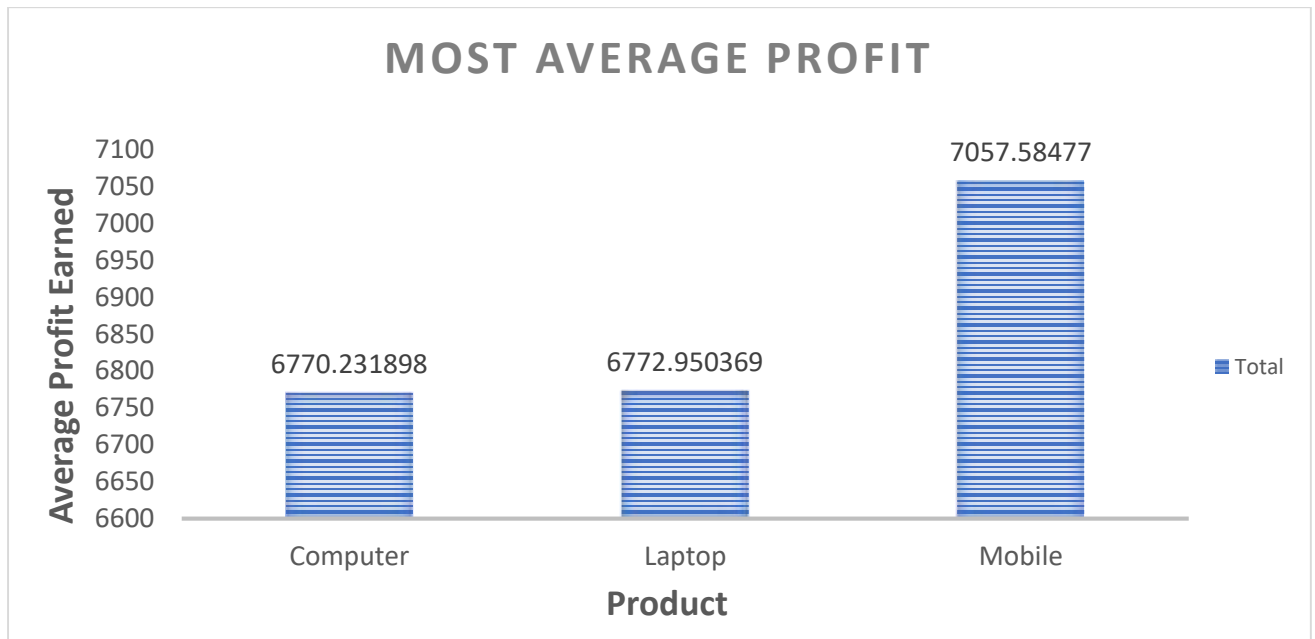
Q3. Find out which of the two product sold the most over the year Computer or Laptop?



Answer:

Laptop was sold more than Computer over the year.

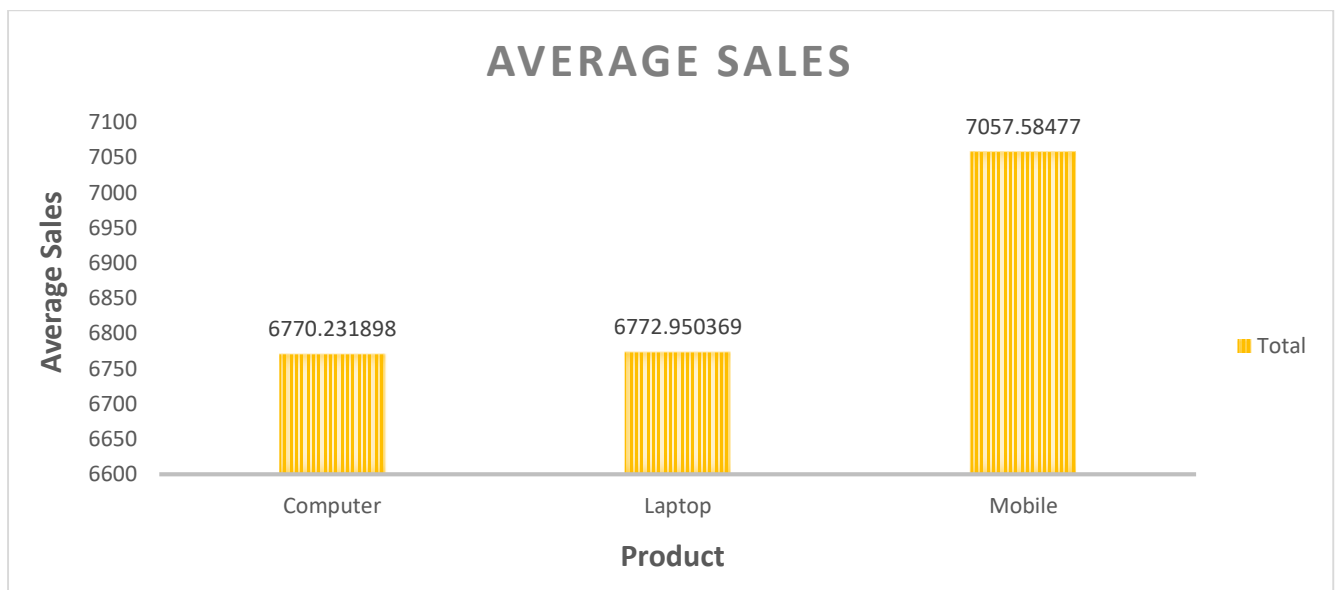
Q4. Which item yield most average profit?



Answer:

Mobiles have the highest average profit among all other products.

Q5. Find out average sales of all the products and compare them.



Answer:

The average sales for mobiles were \$7,057.584, for laptops were \$6,772.95, and for computers were \$6,770.23. Mobiles had the highest sales.

Regression:

SUMMARY OUTPUT						
<i>Regression Statistics</i>						
Multiple R	0.812617					
R Square	0.660347					
Adjusted R Square	0.629469					
Standard Error	1215.119					
Observations	13					
ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	31576697	31576697	21.38598	0.000753	
Residual	11	16241653	14776514			
Total	12	47818350				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	244.7062	754.0557	0.32452	0.751632	-1414.96	1904.372
X Variable	0.190729	0.041243	4.624498	0.000735	0.099954	0.281505

The regression model demonstrates a significant p-value, indicating a robust positive relationship between the amount and the profit earned, serving as the outcome variable. The model's predictive accuracy is substantiated by its high R-squared value of 0.660.

Correlation:

	<i>Qty</i>	<i>Amount</i>
Column 1	1	
Column 2	#DIV/0!	1

The correlation coefficient between units sold and revenue is 0.796, suggesting a strong positive correlation between the two variables.

Anova (Single Factor):

SUMMARY						
Groups	Count	Sum	Average	Variance		
Column 1	15	78.56643	5.237762	2.766871		
Column 2	15	50419.05	3361.27	3416099		
ANNOVA						
Source of Variance	SS	df	MS	F	P-Value	F crit
Between Group	84472135	1	84472135	49.45528	1.2E-07	4.195972
Without Group	47825420	28	170851			
Total	1.32E+08	29				

This ANOVA analysis assesses whether there is a significant difference between two groups (Column 1 and Column 2) in terms of the variable being measured. Here's what each part means:

Between Group Variance: This part compares the variance between the two groups. The high F-value (49.45528) and the very small p-value (1.2E-07) indicate that there is a significant difference between the groups. In simpler terms, there is strong evidence to suggest that the means of the two groups are different from each other.

Without Group Variance: This part examines the variance within each group. It provides information about the variability of scores within each group.

Overall, the results suggest that there is a significant difference between the groups in terms of the variable being measured. In other words, the factor represented by the groups has a significant effect on the outcome variable.

Anova two factor with Replication:

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	841600745	10	4160074	65535	#NUM!	#NUM!
Columns	0	0	65535	65535	#NUM!	#NUM!
Error	0	0	65535			
Total	41600745	10				

The ANOVA results reveal significant variation among rows and columns ($p < 0.001$), with degrees of freedom (df) values of 10 respectively. The error term has a degree of freedom of 0

Anova two factor without Replication:

Summary	Count	Sum	Average
4	1	7800	7800
5	1	3000	3000
4	1	2300	2300
3	1	7000	7000
3	1	1200	1200
4	1	2506.667	2506.667
5	1	2618.095	2618.095
6	1	2729.524	2729.524
7	1	2840.952	2840.952
6	1	4500	4500
7	1	3063.81	3063.81
1000		39559.05	3596.277

This output summarizes the count, sum, and average for various data groups. Each row corresponds to a group, with the first column indicating the group identifier. The second column shows the observation count, the third column displays the sum of values, and the fourth column presents the average value for each group. For instance, group 4 has one observation with a sum and average of \$7,800. Group 1000 comprises 13 observations with a sum of \$39,559.05 and an average of \$3,596.277.

Descriptive Statistics:

<i>Amount</i>	
Mean	6864.457
Standard Error	113.5651
Median	6984.647
Mode	1000
Standard Deviation	2100.186
Sample Variance	4410782
Kurtosis	-0.5078
Skewness	-0.36449
Range	9279.851
Minimum	1000
Maximum	10279.85
Sum	2347644
Count	342

4. Conclusion and Review:

The analysis of the shop sales dataset has uncovered valuable insights into retail operations, including comparisons of salesmen based on profit, identification of top-selling products, and assessment of item performance. Regression analysis reveals a strong positive relationship between sales volume and profitability, while correlation analysis highlights a robust positive correlation between units sold and revenue. ANOVA tests further illuminate variability within the dataset, aiding in trend identification and decision-making. Overall, the dataset serves as a valuable resource for optimizing sales strategies and driving growth in the retail sector. While comprehensive, additional variables could enhance its analytical capabilities.