

Sales Data Samples Report

Introduction:

In the realm of business analytics, a dataset encompassing sales transactions emerges as a vital asset for deriving actionable insights. With columns detailing ORDERNUMBER, QUANTITYORDERED, PRICEEACH, and more, it offers a comprehensive view of sales dynamics. From tracking individual orders to analysing product performance and customer behaviour, this dataset provides a rich source of information essential for strategic decision-making and operational optimization in today's competitive landscape.

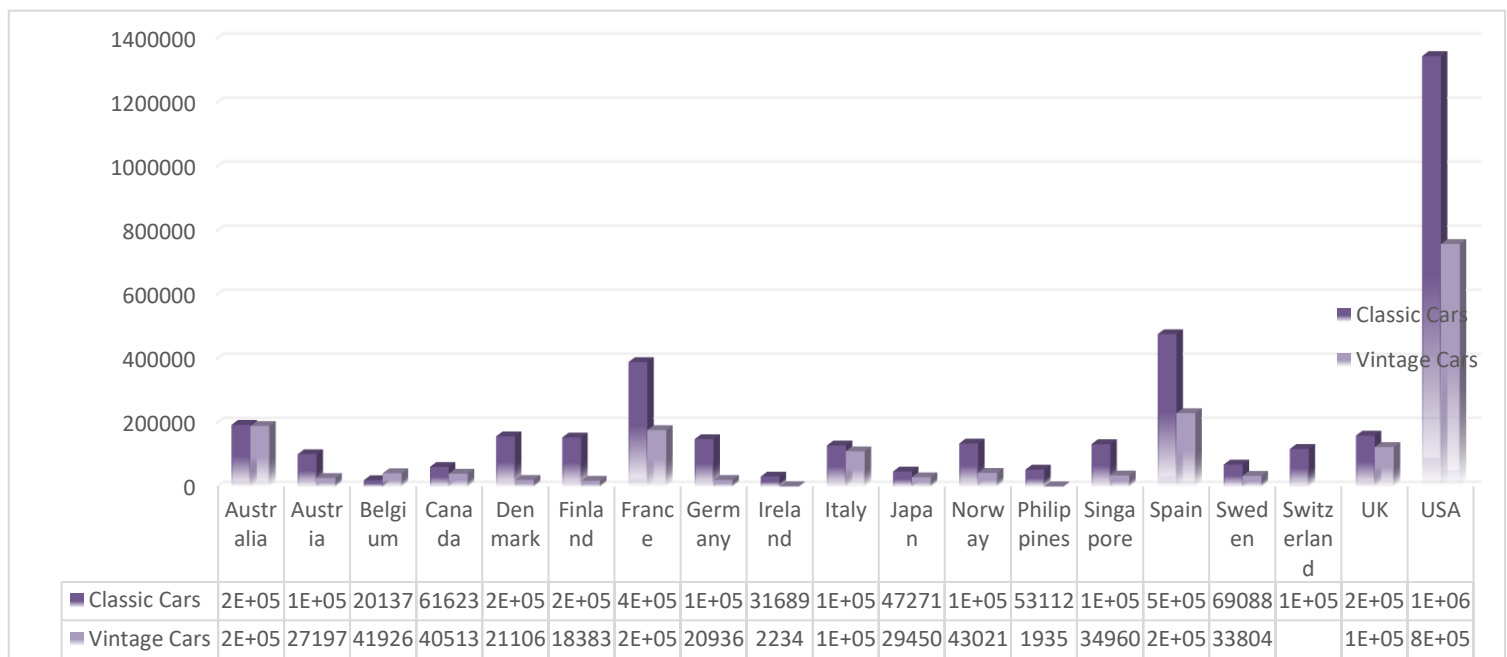
Questionnaire:

1. Compare the sale of Vintage cars and Classic cars for all the countries.
2. Find out average sales of all the products? which product yield most sale?
3. Which country yields most of the profit for Motorcycles, Trucks and buses?
4. Compare sales of all the items for the years of 2004, 2005.
5. Compare all the countries based on deal size.

Analytics:

1. Compare the sale of Vintage cars and Classic cars for all the countries.

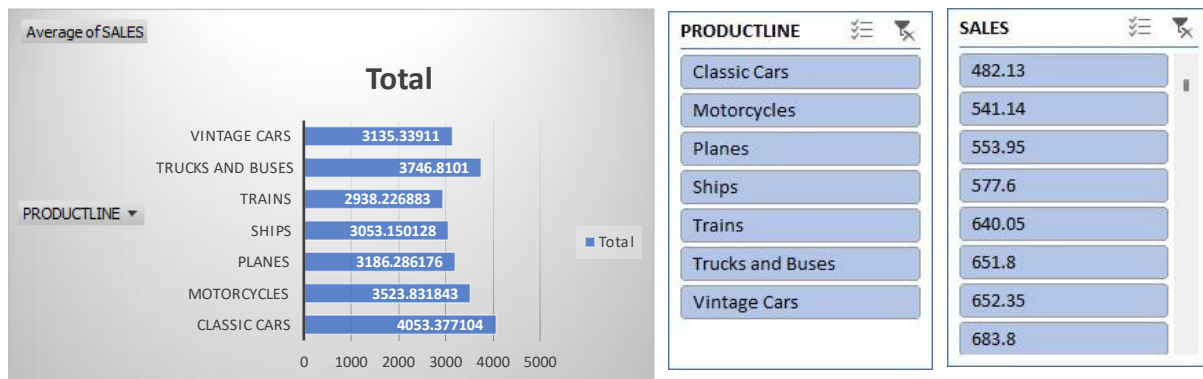
Ans:-The comparsion of sale of Vintage cars and Classic cars for all the countries is given below:-



COUNTRY	PRODUCTLINE	SALES
Australia	Classic Cars	541.14
Austria	Motorcycles	553.95
Belgium	Planes	577.6
Canada	Ships	640.05
Denmark	Trains	652.35
Finland	Trucks and Buses	683.8
France	Vintage Cars	694.6
Germany		702.6

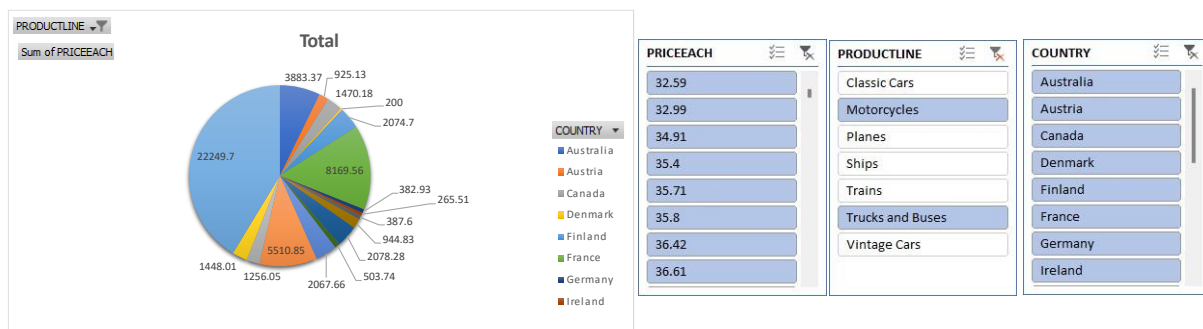
2. Find out average sales of all the products? which product yield most sale?

Ans:



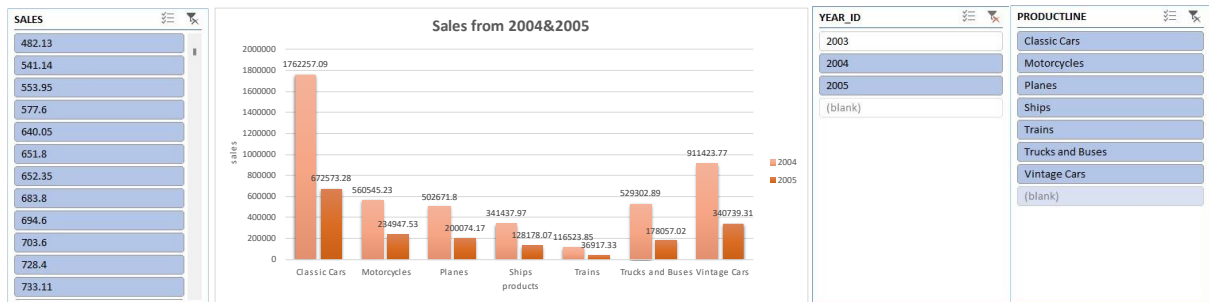
3. Which country yields most of the profit for Motorcycles, Trucks and buses?

Ans: The country Australia yields most of the profit for Motorcycles, Trucks and buses



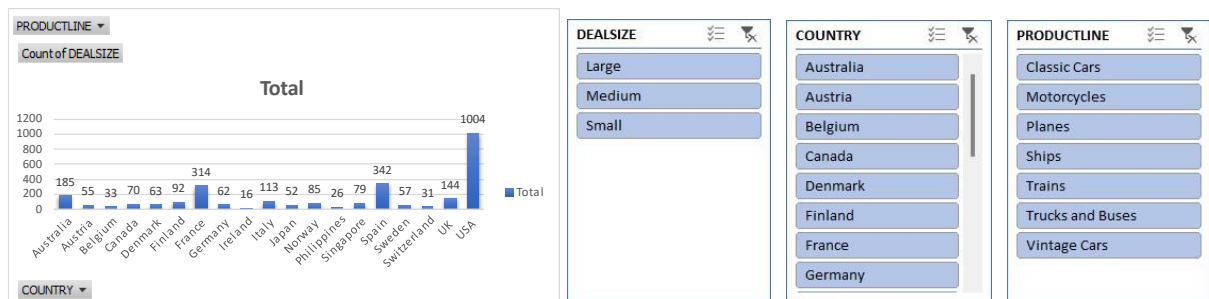
4. Compare sales of all the items for the years of 2004, 2005.

Ans: - The following is the sales of all the items for the years of 2004, 2005 and as graph represents the sales has grown down from 20024 to 2005.



5. Compare all the countries based on deal size.

Ans. The comparison of all the countries based on deal size are:



Regression and Anova

SUMMARY OUTPUT						
<i>Regression Statistics</i>						
Multiple R	0.657840928					
R Square	0.432754687					
Adjusted R Square	0.432553607					
Standard Error	1387.45926					
Observations	2823					
<i>ANOVA</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	4142995200	4142995200	2152.157001	0	
Residual	2821	5430546866	1925043.199			
Total	2822	9573542065				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	1470.590019	111.4099971	13.19980305	1.20143E-38	1689.043329	-1252.13671
PRICE EACH	60.05936566	1.294624334	46.39134619	0	57.52085944	62.59787188

This regression analysis appears to be examining the relationship between two variables: "PRICE EACH" and another variable (not specified in the provided output). Here are the results:

1. **Regression Equation:** The regression equation can be written as: $Y = -1470.59 (\text{PRICE EACH}) + 60.06$ where:
 - Y represents the dependent variable Quantity.
 - X represents the independent variable "PRICE EACH".
2. **Interpretation of Coefficients:**
 - The intercept coefficient (-1470.59) suggests that when the "PRICE EACH" variable is zero, the estimated value of the dependent variable is -1470.59. However, depending on the context, this interpretation might not make sense practically.
 - The coefficient for "PRICE EACH" (60.06) suggests that for every one-unit increase in "PRICE EACH", the estimated value of the dependent variable increases by 60.06 units.
3. **Statistical Significance:**
 - The p-value associated with the coefficient for "PRICE EACH" is 0.00, indicating that the coefficient is statistically significant at conventional levels of significance (typically $\alpha = 0.05$).
 - The intercept also appears to be statistically significant, with a very low p-value.
4. **Goodness of Fit:**
 - The R-squared value (0.433) indicates that approximately 43.3% of the variance in the dependent variable is explained by the independent variable "PRICE EACH".
 - The adjusted R-squared value (0.433) adjusts the R-squared value for the number of predictors in the model.
5. **ANOVA:**
 - The ANOVA table indicates that the regression model as a whole is statistically significant, as the p-value associated with the F-statistic is 0.00.
6. **Standard Error:**
 - The standard error (1387.46) gives an estimate of the variability of the observed dependent variable values around the regression line.
7. **Observations:**
 - The analysis is based on a sample of 2823 observations.

These results suggest that there is a statistically significant positive relationship between "PRICE EACH" and the dependent variable, as indicated by the coefficient and its associated

p-value. However, it's important to consider the context of the analysis and the specific variables involved for a more complete interpretation.

CORRELATION:

The correlation coefficient you calculated (0.657840928) represents the strength. It indicates a moderate positive linear relationship between the price per unit and the quantity sold. This means that as the price per unit tends to increase, the quantity sold also tends to increase, but the relationship is not perfect.

Descriptive Statistics:

<i>SALES</i>	
Mean	3553.889072
Standard Error	34.66589212
Median	3184.8
Mode	3003
Standard Deviation	1841.865106
Sample Variance	3392467.068
Kurtosis	1.792676469
Skewness	1.161076001
Range	13600.67
Minimum	482.13
Maximum	14082.8
Sum	10032628.85
Count	2823

Conclusion and Review:

In conclusion, the analysis of the provided sales dataset offers a window into the intricacies of business operations, shedding light on customer preferences, product performance, and market trends. By leveraging the insights gleaned from this dataset, businesses can make informed decisions, streamline processes, and drive growth. As the landscape of data analytics continues to evolve, harnessing the power of such datasets remains instrumental in staying competitive and responsive to the ever-changing demands of the market.