IMPACT OF CAR FEATURES

Project Description

The automotive industry has been rapidly evolving over the past few decades, with a growing focus on fuel efficiency, environmental sustainability, and technological innovation. With increasing competition among manufacturers and a changing consumer landscape, it has become more important than ever to understand the factors that drive consumer demand for cars.

In recent years, there has been a growing trend towards electric and hybrid vehicles and increased interest in alternative fuel sources such as hydrogen and natural gas. At the same time, traditional gasoline-powered cars remain dominant in the market, with varying fuel types and grades available to consumers.

For the given dataset, as a Data Analyst, the client has asked How can a car manufacturer optimize pricing and product development decisions to maximize profitability while meeting consumer demand?

This problem could be approached by analyzing the relationship between a car's features, market category, and pricing, and identifying which features and categories are most popular among consumers and most profitable for the manufacturer. By using data analysis techniques such as regression analysis and market segmentation, the manufacturer could develop a pricing strategy that balances consumer demand with profitability, and identify which product features to focus on in future product development efforts. This could help the manufacturer improve its competitiveness in the market and increase its profitability over time.

**Data Cleaning**

The percentage of missing data is calculated for all columns using the formula

=(COUNTBLANK(A2:A11915)/COUNTA($A$2:$A$11915))\*100

Only 4 columns has missing data. They are Engine Fuel type, Engine HP, Engine cylinders and Number of doors.

Engine fuel type is a categorical column and three records which are having missing data under this column are removed. Other three columns are having numerical data and hence median value is filled in the blank cells.

**TASK 1 : How does the popularity of a car model vary across different market categories?**

* **Task 1.A**  A pivot table is created to show the number of car models in each market category and their corresponding sum of their popularity scores.

|  |  |  |
| --- | --- | --- |
| **Row Labels** | **Count of Model** | **Sum of Popularity** |
| Crossover | 1110 | 1715242 |
| Crossover,Diesel | 7 | 6111 |
| Crossover,Exotic,Luxury,High-Performance | 1 | 238 |
| Crossover,Exotic,Luxury,Performance | 1 | 238 |
| Crossover,Factory Tuner,Luxury,High-Performance | 26 | 47410 |
| Crossover,Factory Tuner,Luxury,Performance | 5 | 13037 |
| Crossover,Factory Tuner,Performance | 4 | 840 |
| Crossover,Flex Fuel | 64 | 132720 |
| Crossover,Flex Fuel,Luxury | 10 | 11732 |
| Crossover,Flex Fuel,Luxury,Performance | 6 | 9744 |
| Crossover,Flex Fuel,Performance | 6 | 33942 |
| Crossover,Hatchback | 72 | 120650 |
| Crossover,Hatchback,Factory Tuner,Performance | 6 | 12054 |
| Crossover,Hatchback,Luxury | 7 | 1428 |
| Crossover,Hatchback,Performance | 6 | 12054 |
| Crossover,Hybrid | 42 | 107662 |
| Crossover,Luxury | 410 | 362665 |
| Crossover,Luxury,Diesel | 34 | 73080 |
| Crossover,Luxury,High-Performance | 9 | 9335 |
| Crossover,Luxury,Hybrid | 24 | 15142 |
| Crossover,Luxury,Performance | 113 | 151968 |
| Crossover,Luxury,Performance,Hybrid | 2 | 7832 |
| Crossover,Performance | 69 | 178431 |
| Diesel | 84 | 145396 |
| Diesel,Luxury | 51 | 116025 |
| Exotic,Factory Tuner,High-Performance | 21 | 21974 |
| Exotic,Factory Tuner,Luxury,High-Performance | 52 | 26912 |
| Exotic,Factory Tuner,Luxury,Performance | 3 | 1560 |
| Exotic,Flex Fuel,Factory Tuner,Luxury,High-Performance | 13 | 6760 |
| Exotic,Flex Fuel,Luxury,High-Performance | 11 | 5720 |
| Exotic,High-Performance | 261 | 331818 |
| Exotic,Luxury | 12 | 1352 |
| Exotic,Luxury,High-Performance | 79 | 36899 |
| Exotic,Luxury,High-Performance,Hybrid | 1 | 204 |
| Exotic,Luxury,Performance | 36 | 7813 |
| Exotic,Performance | 10 | 13910 |
| Factory Tuner,High-Performance | 106 | 205790 |
| Factory Tuner,Luxury | 2 | 1234 |
| Factory Tuner,Luxury,High-Performance | 215 | 458674 |
| Factory Tuner,Luxury,Performance | 31 | 43816 |
| Factory Tuner,Performance | 92 | 156004 |
| Flex Fuel | 872 | 1933488 |
| Flex Fuel,Diesel | 16 | 90512 |
| Flex Fuel,Factory Tuner,Luxury,High-Performance | 1 | 258 |
| Flex Fuel,Hybrid | 2 | 310 |
| Flex Fuel,Luxury | 39 | 29115 |
| Flex Fuel,Luxury,High-Performance | 33 | 29004 |
| Flex Fuel,Luxury,Performance | 28 | 38642 |
| Flex Fuel,Performance | 87 | 146201 |
| Flex Fuel,Performance,Hybrid | 2 | 310 |
| Hatchback | 641 | 845393 |
| Hatchback,Diesel | 14 | 12222 |
| Hatchback,Factory Tuner,High-Performance | 13 | 15667 |
| Hatchback,Factory Tuner,Luxury,Performance | 9 | 7982 |
| Hatchback,Factory Tuner,Performance | 22 | 47499 |
| Hatchback,Flex Fuel | 7 | 39599 |
| Hatchback,Hybrid | 72 | 152730 |
| Hatchback,Luxury | 46 | 63457 |
| Hatchback,Luxury,Hybrid | 3 | 1362 |
| Hatchback,Luxury,Performance | 38 | 59513 |
| Hatchback,Performance | 252 | 261991 |
| High-Performance | 199 | 362468 |
| Hybrid | 123 | 258985 |
| Luxury | 855 | 942772 |
| Luxury,High-Performance | 334 | 557118 |
| Luxury,High-Performance,Hybrid | 12 | 6826 |
| Luxury,Hybrid | 52 | 35029 |
| Luxury,Performance | 673 | 869930 |
| Luxury,Performance,Hybrid | 11 | 25665 |
| N/A | 3739 | 6273477 |
| Performance | 601 | 810673 |
| Performance,Hybrid | 1 | 155 |
| (blank) |  |  |
| **Grand Total** | **11911** | **18523769** |

* **Task 1.B:** A combo chart is created to visualizes the relationship between market category and popularity.

**Insights :** Flex fuel, Cross over and Luxury are top three market categories based on sum of popularity.

Hatchback with flex fuel, Flex fuel with diesel and Cross over with flexfuel and Performance are the top 3 market categories based on average of popularity.

**TASK 2**

Insight Required : What is the relationship between the cars’ engine power and its price?

A scatter plot is created between price on Y axis and engine power on X-axis. A trendline is added to visualise the relationship between these variables.

The plot exhibits a positive trendline suggesting the there is a direct correlation between the engine power and price of the cars.

The price of the car increases with its engine power.

**Task 3:** Which are the car features more important in determining the car price?

Using regression analysis, the variables which have strong relationship with car price are identified. Then a bar chart is created to understand their relative importance based on the coefficient values.

Based on the regression analysis conducted on Engine power, Engine cylinders, number of doors, highway MPG, city MPG and popularity, the most affecting features are engine cylinders and number of doors.

Engine cylinders increases with car price.

Two doored cars are costlier than four doored cars.

**Task 4: How does the average price of a car vary across different manufacturers?**

**Task 4.A**

A pivot table is created to determine the average prices of cars for each manufacturer.

**Task 4B:** A vertical bar chart to visualize the variation of average price among all manufacturers.

**Insight : Bugati has the highest average price amongst all brands.**

+

**Task 5: What is the relationship between the fuel efficiency and the number of cylinders in a car engine?**

Task 5.A : Create a scatter plot with no. of cylinders over X-axis and Highway MPG over Y-axis. Then create a trendline to visually estimate the slope of the relationship and assess its significance.

Task5.B: Calculate the correlation coefficient between the number of cylinders and highway MPG to quantify the strength and direction of the relationship.

|  |  |  |
| --- | --- | --- |
|  | *Engine Cylinders* | *highway MPG* |
| Engine Cylinders | 1 |  |
| highway MPG | -0.60094 | 1 |

The negative trendline shows that the performance of the car reduces with increased number of cylinders in engine.

## **Building the Dashboard:**

Now for the Next portion of the Project is to create the Interactive Dashboard.

Filters and slicers are used to make the chart interactive.

**Task 1:** How does the distribution of car prices vary by brand and body style?

* Stacked column chart to show the distribution of car prices by brand and body style based on filters and slicers to make the chart interactive. The total MSRP for each brand and body style is calculated using Pivot Tables.

**Insights :**

* Chevrolet, Mercedez Benz and Ford are the top 3 brands based on total MSRP.
* Sedan, 4dr SUV and Coupe are the top 3 body styles based on total MSRP.

**Task 2:** Which car brands have the highest and lowest average MSRPs, and how does this vary by body style?

Bugatti has highest average MSRP and Plymouth has lowest Average MSRP.

Coupe, Sedan and 4dr SUV are the top 3 designs amongst all brands based on average MSRP.

**Task 3:** How do the different feature such as transmission type affect the MSRP, and how does this vary by body style?

* Scatter plot chart to visualize the relationship between MSRP and transmission type, with different symbols for each body style.
* The average MSRP is calculated for each combination of transmission type and body style using Pivot Tables.

Convertible style has higher MSRP across all transmission types. 2Dr SUV has least MSRP across its transmission style.

**Task 4:** How does the fuel efficiency of cars vary across different body styles and model years?

* Line chart is used to show the trend of fuel efficiency (MPG) over time for each body style. The average MPG for each combination of body style and model year using Pivot Tables. Body styles like 4Dr hatchback, 2Dr hatchback and sedan performs very well with higher overall fuel efficiency. All brands can focus on these 3 styles to increase its sales.

**Task 5:** How does the car's horsepower, MPG, and price vary across different Brands?

* Bubble chart is created to visualize the relationship between horsepower, MPG, and price across different car brands.
* Bolt EV, i3, spark EV are the top high performance models based on their mileage. Performance of brands reduces when they are designed with Higher Engine HP. Price is independent of both horsepower and mileage.

**RESULT :**  The impact of various car features affecting its sales are thoroughly visualized by plotting bar charts, scatter plots, bubble charts, combo charts to derive as many as useful insights for the manufacturers as well as buyers.

**Link for Excel File:** [**Click here**](https://docs.google.com/spreadsheets/d/1XzFNs7-PLs801TnxXnVZy4FQ-h6l6dqB/edit?usp=sharing&ouid=110267340843358264811&rtpof=true&sd=true)