Competitive Automobile Market Analysis

Project – Part 1

Project Objective

- The objective of the project is to analyze a car dataset to identify key factors for the potential car models to launch.
- Extract actionable insights from the given data, addressing key areas such as new product development/innovation and competitive analysis/benchmarking, etc.
- You are expected to apply the analytical skills gathered during previous all sprints for this project. Excel should be used based on requirement of the given analysis.
- The objective of this sprint is to clean the dataset (if required), perform project tasks mentioned in following slides, and summarize your key findings.

In automotive industry, a car manufacturer, faces a shifting market, drivers are becoming increasingly eco-conscious, prioritizing fuel efficiency and safety alongside performance. Data analytics is now applied at every stage of production. Smith, a data analyst with a popular automobile giant, needs to refer to a car dataset and identify the potential models to launch. To arrive at a conclusion, the upcoming tasks need to be performed.

Note: The output needs to be presented visually using appropriate graphs/charts, wherever applicable.

To implement the whole project, you need to perform the following steps:

Part 1:

- Data pre-processing (if required).
- Solving business questions according to the given tasks.
- Write interpretation for each task individually.

Part 2:

- Solving business questions according to the given tasks.
- Write **the** interpretation for each task individually.

Part 3:

- Solving business questions according to the given tasks.
- Write **the** interpretation for each task individually.
- Provide final business conclusions to support your analysis for entire project.

- Summarization of all the interpretations obtained from all the tasks .
- Provide final business conclusions to support your analysis for entire project.
- Prepare presentation ppt with project requirements, findings from analysis, summary and business conclusion (Use a maximum 15 slides).

Note: Write an interpretation from the output produced for each task in this project. Summarize all the interpretations at the end of all tasks.

Data Preprocessing and Performing Tasks

Part -1

Task 1:

Identify the top 10 cars with the highest miles per gallon. Also, identify the names of the car manufacturers.

Note: You may consider either city mileage or highway mileage.

Task 2:

Identify the top 10 cars with the highest miles per gallon based on displacement. Also, identify the names of their car manufacturers.

Note: You may consider either city mileage or highway mileage.

Task 3:

Identify the top 3 manufacturers with the highest number of cars of the hatchback body type.

Task 4:

Compare the mileages of cars produced by all manufacturers and perform outlier analysis.

Note: You may consider either city mileage or highway mileage.

Task 5:

Present the relationship between:

(a) displacement and mileage

(b) mileage and fuel tank capacity

Note: You may consider either city mileage or highway mileage.

Task 6:

Identify a car that would be best for adventurous drives.

Note: Consider the following areas for analysis: car size, engine power, safety features like traction control, and ground clearance. Based on these areas, infer the correct columns in the data for analysis.

Task 7:

A person choosing a car for family usage may consider the following areas for analysis.

1. Car size

2. Safety features for adults and children

3. Mileage

4. Comfort during long distance travel

5. Built-in entertainment options

Using the columns above, identify a car suitable

for family usage.

Task 8:

The car manufacturer decides to launch two types of cars.

1. Compact city car (primarily for city usage)

2. Family car (primarily for long distance travel)

Identify the potential competitor cars for the two types of cars stated above.

Note: Write an interpretation from the output produced for each task in this project. Summarize all the interpretations at the end of all tasks.

Competitive Automobile Market Analysis

Project – Part 2

In continuation of the car dataset analysis through finding key factors for the potential launching of car models, present the most diverse car manufacturers in the given country. Perform project tasks mentioned in following slides and summarize your key findings.

In project – part 1, we have considered car dataset to identify the potential models to launch Let's consider few more business questions to present insights about the most diverse car manufacturer in the given country. In this sprint, use the pre-processed data which has been obtained after completion of project – part 1.

Present the insights by performing the following tasks.

Note: The output needs to be presented visually using appropriate graphs/charts, wherever applicable.

Write interpretation from output produced for each task in this project. Summarize all the interpretations at the end of the all tasks.

Task 1:

Identify outliers from following columns using Histogram method.

1. Displacement(Bin Size=980)

2.City_Mileage(miles/gallon)(Bin Size=0.88)

3.Fuel_Tank_Capacity(Bin Size=3)

Remove the identified outliers from the dataset.

Task 2:

- Identify the top three car manufacturers who have the highest number of variants in the hatchback category.
- Identify the top three manufacturers who have the highest number of variants in the sedan category.
- Identify the top three manufacturers who have the

highest number of variants in the SUV category.

Task 3:

Find the most popular car body type, by count, from the analysis of Task 2.

Car body types can fall under categories such as SUV, sedan, and hatchback**.**

Task 4:

Identify the top two manufacturers who offer the widest range of cars variants. Use appropriate charts/graphs to present the output, capturing the number of car variants under each body type for the two manufacturers identified above.

Task 5:

Calculate and compare variations between City_Mileage data using different measures of dispersion of the following three manufacturers.

1.Tata

2. Maruti Suzuki

3. Hyundai

Which manufacturer offers the wide variation of

city mileage?

Task 6:

a. Display and check how City_Mileage depends on Fuel_Tank_Capacity using scatter plot along with trend line and Pearson's product-moment coefficient(R).

b. Display and check how Displacement depends on Fuel_Tank_Capacity using scatter plot along with

trend line and Pearson's product-moment coefficient(R).

Note: Write an interpretation from the output produced for each task in this project. Summarize all the interpretations at the end of all tasks.

Competitive Automobile Market Analysis

Project - Part- 3

In continuation of the car dataset analysis through finding key factors for the potential launching of car models, analyze fuel efficiency data and driving patterns to identify opportunities for reducing carbon emissions. Perform project tasks mentioned in following slides and summarize your key findings.

In project – part 1, we have considered car dataset to identify the potential models to launch

In project – part 2, few more business questions have been considered to present insights about the most diverse car manufacturer in the given country.

In this sprint, let's narrow down the analysis to observe fuel efficiency data and driving patterns to identify opportunities for reducing carbon emissions by following one more round of data pre-processing and performing few more adverse tasks.

Note: The output needs to be presented visually using appropriate graphs/charts, wherever applicable.

Write interpretation from output produced for each task in this project. Summarize all the interpretations at the end of the all tasks.

Task 1: Data Preprocessing

A) Import the car dataset into excel.

B) Clean and format data

a) Check for missing values and inconsistencies.

b) Convert categorical variables to numerical codes (e.g., 1. CNG, 2. Petrol for fuel types etc).

c) Standardize units (e.g., convert inches to meters).

C) Calculate additional fuel efficiency metrics:

a) Combine city and highway mileage into a weighted average based on typical driving patterns.

[Formula: Avg. Combined Mileage = (City_Mileage + Highway_Mileage)/2]

b) Calculate distances driven from the data of combined city_highway mileage and fuel tank capacity.

[Formula: Distance = (Fuel Tank Capacity * Avg. Combined Mileage)]

Task 2 – Data Analysis

A) **Visualize fuel efficiency distribution:** Compare fuel efficiency across car makes, models, fuel types, and other relevant features using Bar chart.

B) **Analyse driving patterns:**

a) Plot average speeds and distances driven against fuel efficiency to identify trends.

b) Use scatterplots and correlation matrices to explore relationships between fuel efficiency and other variables.

C) Identifying Opportunities for Reduced Emissions: Group cars by fuel efficiency:

a) Segment cars into categories like "high efficiency," "moderate efficiency," and "low efficiency."

b) Analyze the distribution of features within each category.

c) Identify common characteristics of cars with high fuel efficiency.

- Provide final business conclusions to support your analysis for entire project.
- Summarization of all the interpretations obtained from all the tasks
- Prepare presentation ppt with project requirements, findings from analysis, summary and business conclusion (Use maximum 15 slides).
- **Note:** Write an interpretation from the output produced for each task in this project. Summarize all the interpretations at the end of all tasks.