Spinny's Revenue Growth Through Demand and Pricing Optimisation

Introduction:

Spinny is an Indian company that provides a platform to buy and sell pre-owned cars online, bringing transparency and honesty to the pre-owned car market. The cumbersome process of buying and selling cars is replaced by eliminating unnecessary charges, reducing the paperwork, and eradicating all hassles.

It operates across the entire value chain of pre-owned cars, embedding superior technology and processes to deliver a premium experience to customers. Every car on the Spinny platform comes with a 200-point inspection checklist, 5-day no questions asked money-back guarantee, and 1-year after-sales warranty.

Objective:

We mainly aim to analyse Spinny's used car transaction data through this project to:

1. **Understand Demand:** Assess the demand for pre-owned cars in each region.
2. **Buying Trends:** Identify the demand for specific car models and their key attributes (e.g., condition, fuel type, mileage).
3. **Optimal Pricing:** Determine the pricing for each car, ensuring it is competitive and aligned with market trends.
4. **Increase Revenue:** Provide insights to increase revenue by boosting sales.

Business Impact:

This analysis will help Spinny in:

1. Provide better offerings to consumers, to boost sales.

2. Optimize pricing strategies based on different cars and location
3. Identify regions to run marketing campaigns to increase consumer awareness.
4. Provide detailed insights to build consumer trust and retention.
5. Offer certified pre-owned programs for cars with thorough inspections, providing added assurance to buyers and potentially justifying higher prices.

Dataset Information:

- Dataset: Spinny Analytics
- Number of Rows: 426880
- Number of Columns: 21

Each row in the dataset represents details of unique vehicles with specific attributes.

Explanation of Data Columns:

1. **Id (Transaction ID)**: A unique identifier for each car transaction in the dataset.
2. **Price:** The final selling price of the vehicle in Indian Rupees (INR).
3. **Year:** The year of manufacture of the vehicle, indicating its age.
4. **Manufacturer:** The brand or company that produced the vehicle.
5. **Model:** The specific model name of the vehicle produced by the manufacturer.
6. **Condition:** The physical and operational state of the vehicle at the time of sale.
7. **Cylinder:** The number of cylinders in the car's engine, influencing its power and performance.
8. **Fuel:** The type of fuel the vehicle uses, such as petrol, diesel, CNG, or electric.
9. **Odometer:** The total distance the vehicle has travelled, measured in kilometres.

10. **Title_status:** The condition of the car's, indicating if it is clean, or has missing parts or has liens.
11. **VIN (Vehicle Identification Number):** A unique serial number assigned to each vehicle for identification.
12. **Drive:** The drivetrain configuration of the vehicle (e.g., front-wheel drive, rear-wheel drive).
13. **Size:** The size classification of the vehicle, such as compact, midsize, or full-size.
14. **Type:** The type of vehicle, such as sedan, SUV, hatchback, or truck.
15. **Paint_color:** The exterior color of the vehicle.
16. **Image_url:** A link to an image of the vehicle for reference.
17. **State:** The Indian state where the car transaction took place.
18. **Posting_date:** The date when the car was listed for sale.
19. **Latitude:** The geographical latitude coordinate of the transaction location.
20. **Longitude:** The geographical longitude coordinate of the transaction location.

Data Cleaning and Preparation:

Data cleaning and preparation is an important step in the Data Analysis to get required insights. The Spinny dataset includes different attributes of vehicles and its pricing which has potential issues like missing values, inconsistent formats, and incorrect data types. This process will make the dataset ready for exploratory data analysis and to get further insights.

- **Data Import:** The CSV file containing Spinny pre-owned car data was imported from the drive into Google Colab for further processing.
- **Finding null Values:** The dataset contains 12,15,288 null values across different columns. We used **[.isna().sum().sum**() **]** to get null values count and used mean, mode, and word Unknown using **[.fillna]** to address null values.
- **Removing Duplicates:** The dataset contains no duplicate values. No modification is needed here.

- **Checking Datatypes:** Verifying the data types of each column is important to ensure they are correctly formatted, using **[.dtypes]**
- **Identifying the timeframe of the dataset:** Used .max() and .min() to find the timeframe of the dataset, and it shows that the data is for 122 years.
- **Outliers:** Outliers were calculated using Interquartile range with 25th and 75th percentiles, to check for any unusual values in the dataset that might affect calculations.
- **Normalization:** Normalization is important for comparisons, especially when data in different columns has varying ranges. It can be done using various methods, in this the min-max method was used to prepare the data for further analysis.

Data Overview:

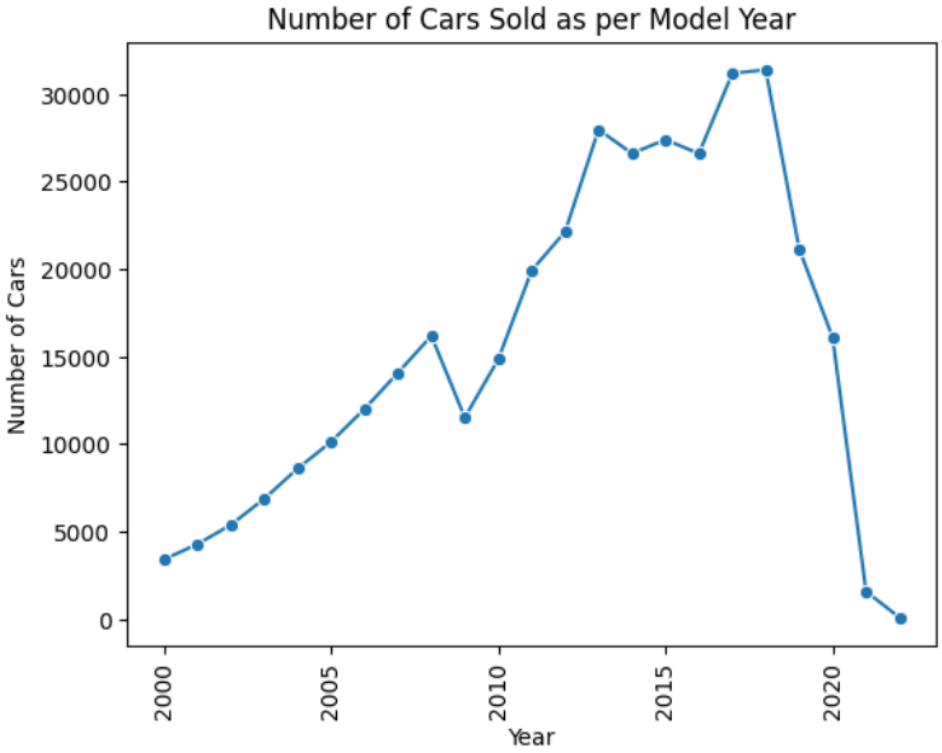
After cleaning and preparing the data, we reviewed everything to make sure the changes were correctly applied. Then, we analysed the following from the data to understand key points and to get a clear overview of the data.

1. Cars made by Ford have highest number of sale, while cars made by Morgan have lowest sale.
2. Number of Ford's car sold in Uttar Pradesh is at highest of 2013-2017 model.
3. BMW has made highest total revenue while being at 9th position in number of sales, which means BMW cars are most expensive.
4. Uttar Pradesh and Maharashtra has contributed to both, in terms of highest number of car sales, and highest revenue source.
5. Highest number of transaction happened on 2018 and 2019 model cars, which shows that buyer prefer buying cars which are recent ones.

Exploratory Data Analysis (EDA)

Hypothesis 1:

Car model greater than of certain year with lower number of sales, indicating consumer preference towards newer models.



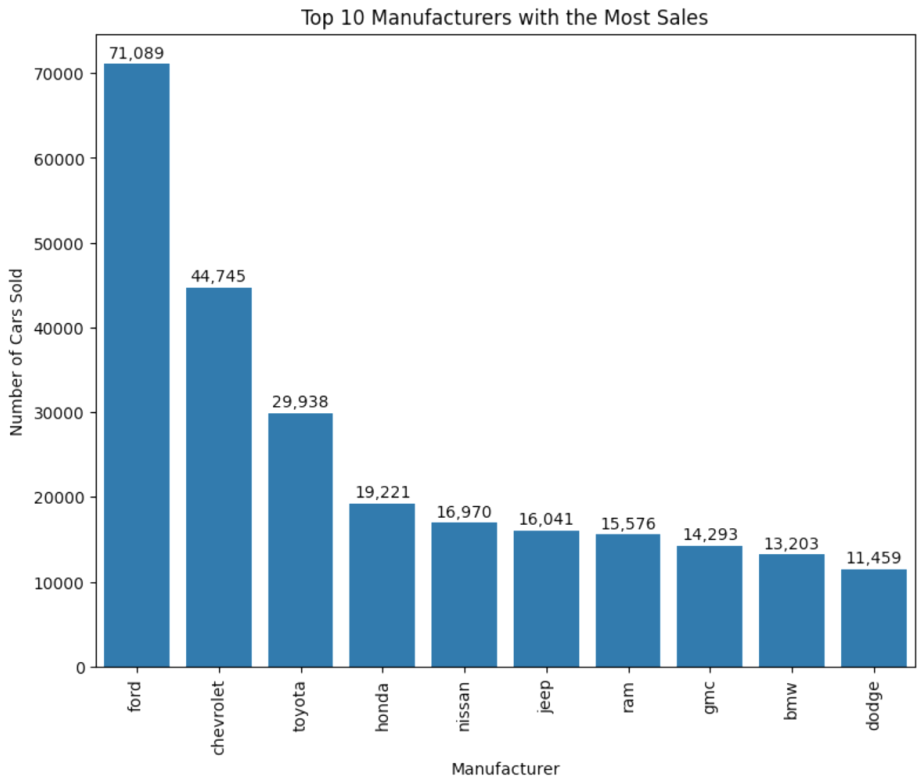
Analysis: From the analysis we can see that sales number has continuously increased till 2018 model and at sharper rate from 2013 model onwards, while numbers decreased after 2018 model. This suggests that buyers prefer cars which are not more than 10 years old and there is low availability of cars which are 3-4 years old.

Recommendation:

- Stock enough vehicles from high-demand years (<= 10 years) while minimizing excess stock for slower-selling models.
- Increase inventory of vehicles manufactured in last 3-4 years.
- Stay informed about changing customer preferences, technological advancements, and economic conditions to anticipate future sales trends and adapt strategies accordingly.

Hypothesis 2:

Manufacturers having high sales number with high average price results in high profits for company.

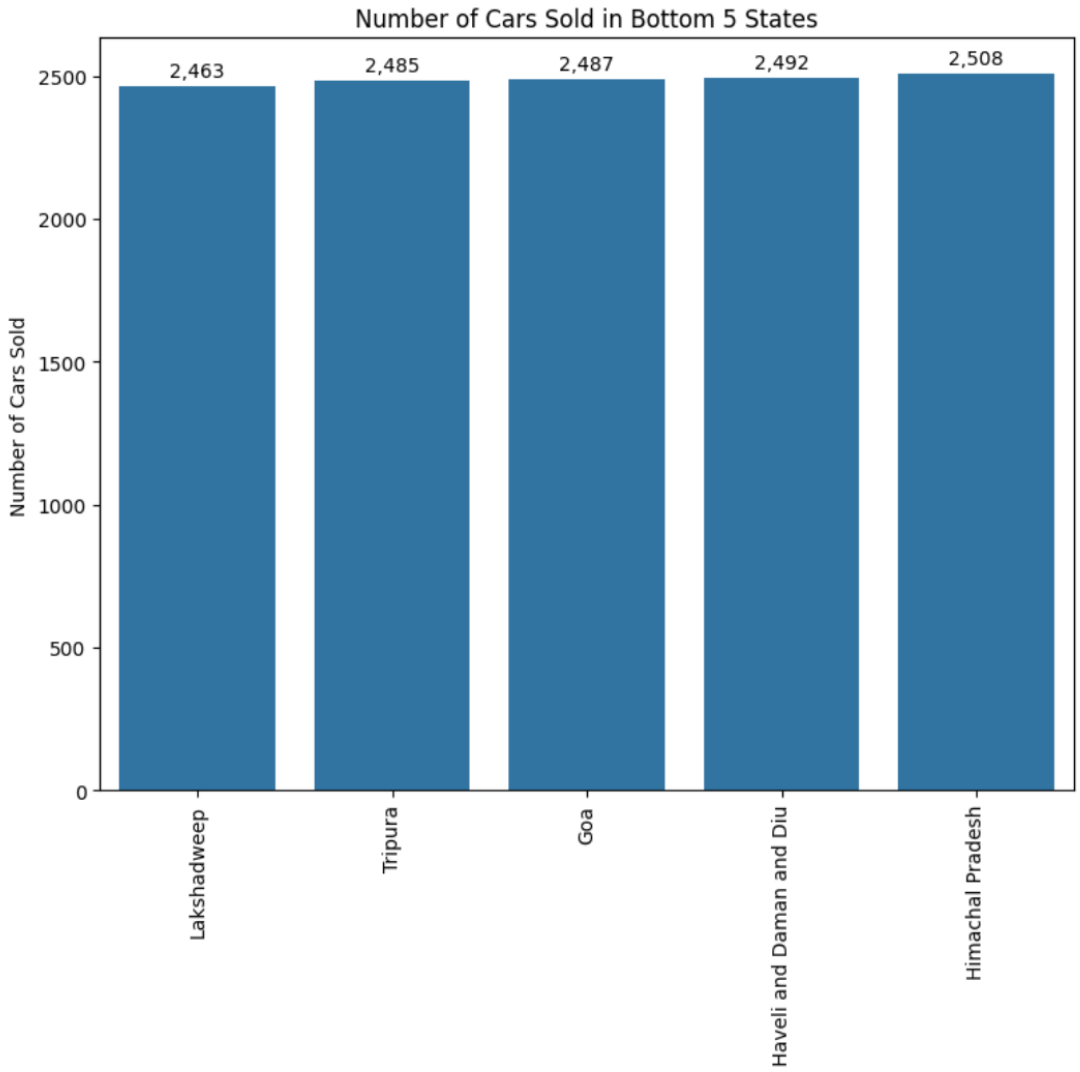


Analysis: High sales number or high total revenue results in high commission for brand resulting in increase in profits.

Recommendation: Provide timely sale offer to users for vehicles in-demand. Optimize pricing and features for brands which have low sales number but high average price.

Hypothesis 3:

States with lower number of transactions, leading to drop in market share and profits.



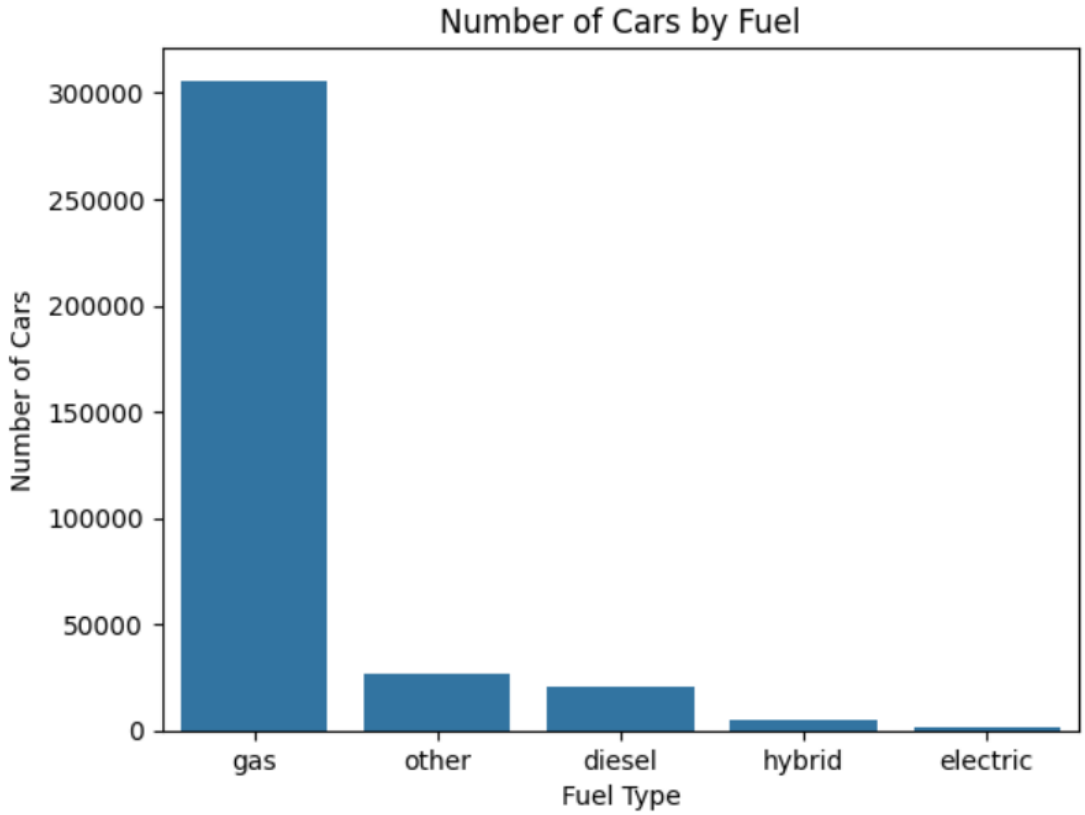
Analysis: Some of the regions have comparatively lower sales number of many manufacturers. This indicates less availability of in-demand cars and low consumer awareness.

Recommendation: Increase in-demand car listings in those regions at right price point. Improve consumer awareness by targeted marketing campaigns for in-demand brands at right price point.

Hypothesis 4:

Fuel type with most transactions, tend to have most sales number and higher average price. While certain newer fuel type showing increasing share, indicating consumer preference shifts.

```
       fuel    % Share
0       gas  85.051857
1     other   7.449457
2    diesel   5.725880
3    hybrid   1.345319
4  electric   0.427487
Text(0.5, 1.0, 'Number of Cars by Fuel')
```

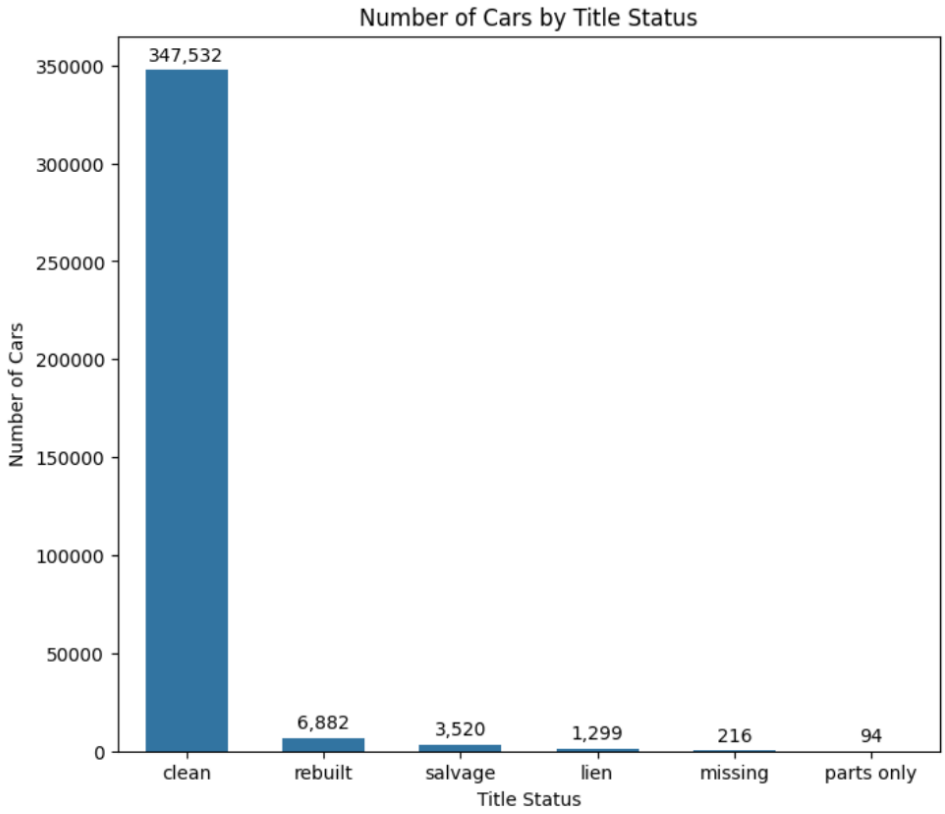


Analysis: Most transaction is for gasoline fuel type for 85% vehicles while consumer are shifting towards other low-cost fuel options such as electric and hybrid type. Diesel vehicles, while present, cater to niche markets such as trucks and commercial vehicles.

Recommendation: Provide users with quality listings of gas type vehicles to maintain current revenue. Encourage users to purchase electric and hybrid type by sending personalized offers, to adapt to changing preference.

Hypothesis 5:

A certain titles status car with high number of transactions, showing high consumer preference.



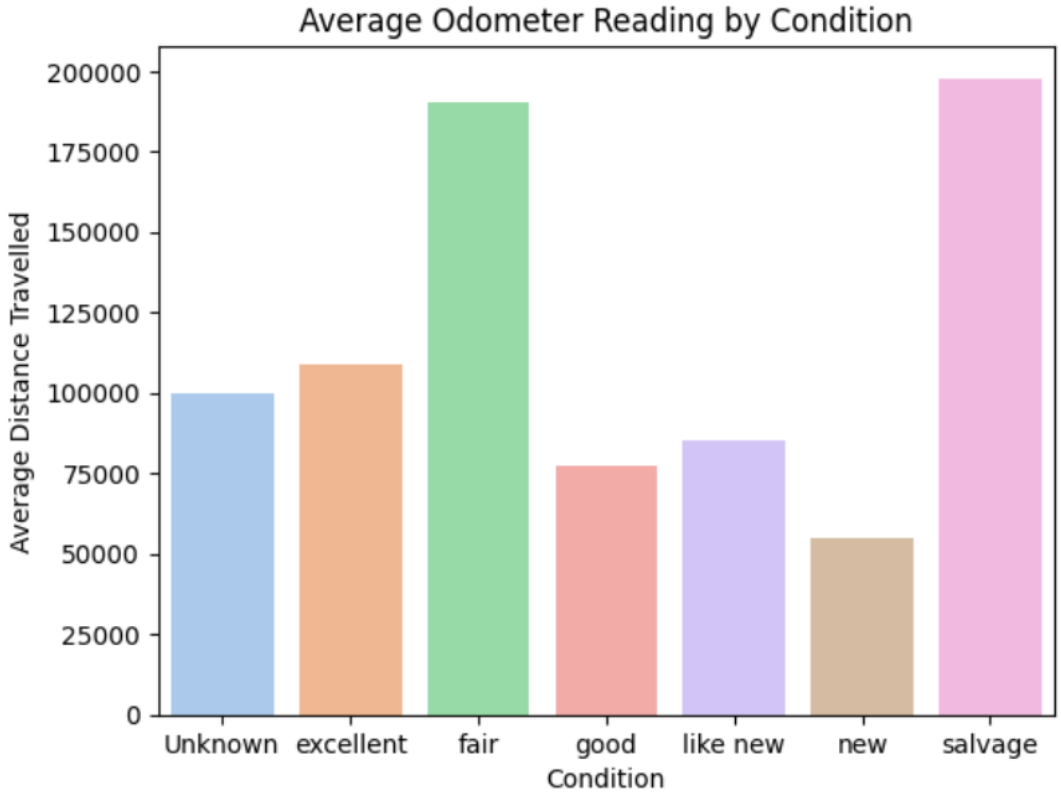
Analysis: It indicates that title status is a significant factor influencing its demand. Cars with 'clean' titles, implying a clear ownership history and no major damage, are perceived as more desirable. While, titles like 'salvage', 'rebuilt', or 'lien' suggest potential risks or complications, leading to lower sales number.

Recommendation:

- **Pricing Strategy**: Cars with 'clean' titles can be priced at a premium, to increase profits. However, cars with other title statuses might require more competitive pricing to attract buyers.
- **Risk Assessment**: For cars with 'salvage', 'rebuilt', or 'lien' titles, conduct thorough inspections and due diligence to assess the extent of any damage or potential issues. This helps mitigate risks and ensures accurate pricing.
- **Targeted Marketing**: Develop targeted marketing campaigns for cars with different title statuses. Emphasize the value proposition of 'clean' title cars and highlight any unique features or benefits of cars with other title statuses.

Hypothesis 6:

Vehicles which have travelled lesser distance tend to be of good condition, indicating higher demand and average price.



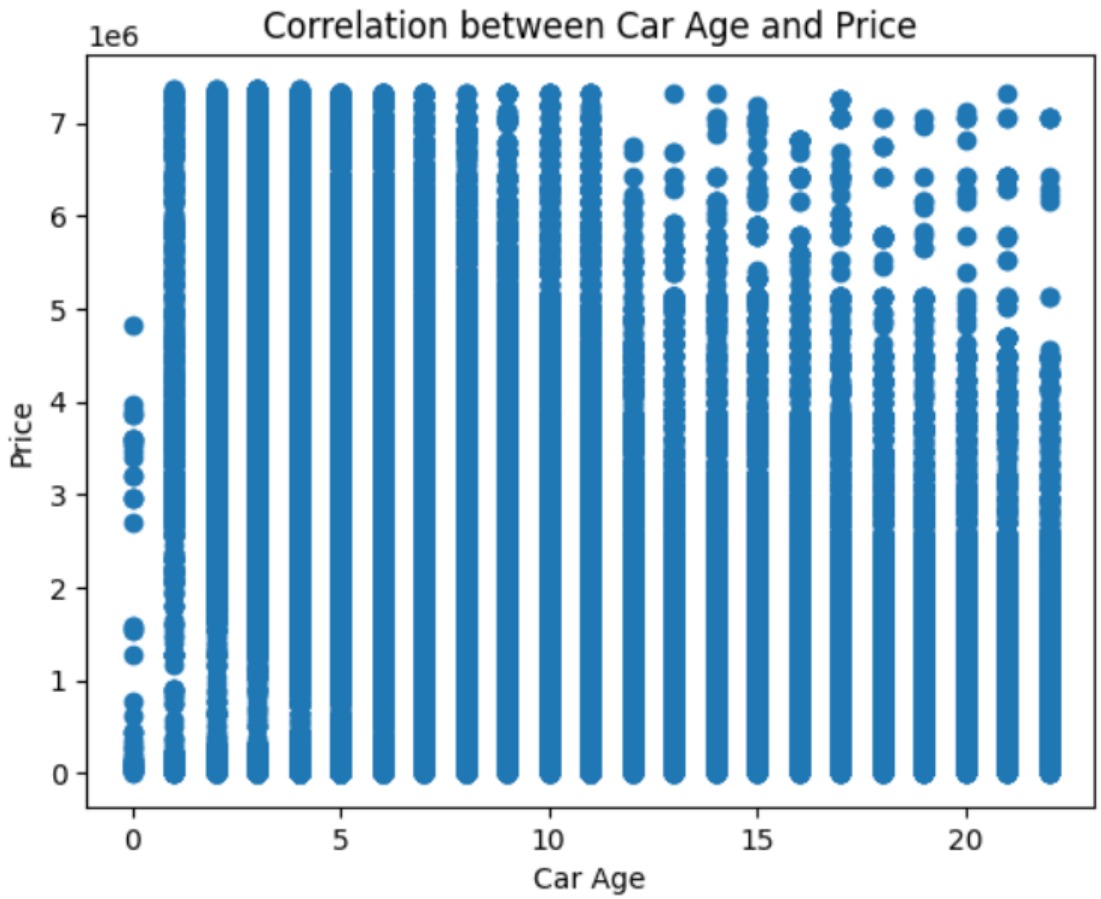
Analysis: From the analysis, we can see that lesser travelled cars are in better condition, tending to command higher price and more sales.

Recommendation:

- Prioritize the listing and sale of cars with lower mileage, due to higher demand and sell at better prices.
- Provide vehicle history reports, including detailed mileage information and any major repairs, addressing concerns about mileage accuracy and potential hidden issues.
- Target marketing efforts based on customer preferences for mileage. Some customers might be highly sensitive to mileage, while others might be more flexible and prioritize other factors like features or condition.
- Consider offering certified pre-owned programs for cars with lower mileage and thorough inspections, providing added assurance to buyers and potentially justifying higher prices.

Correlation Analysis:

1. **A negative correlation between car age and price shows that sales price goes down as car model becomes older.**



Analysis: The magnitude of the correlation value (-0.6) indicates the strength of the relationship. A stronger negative correlation suggests a more pronounced decrease in price with increasing car age.

Recommendation: To improve sales and profitability, optimizing pricing strategies by adjusting prices based on the age can be considered. It'll also helps potential buyers understand the expected price range for cars of different ages.

2. **There is a negative correlation between price and odometer and a positive correlation between car age and odometer which indicates that more used car commands less price and is older.**

Analysis: The above correlation indicates that cars which have travelled more and are older model tends to have lower price.

Recommendation: Target customers based on preferences for mileage and prioritize listing of cars which are newer model and have travelled lesser distance. Some customers might be highly sensitive to mileage, while others might be more flexible and prioritize other factors like features or condition.

Tools and Libraries Used:

- **Libraries Used:**
  - NumPy for numerical computations.
  - Pandas for data cleaning and manipulation.
  - Matplotlib and Seaborn for data visualization.

- **Environment:** Google Colab for data processing and analysis.

Recommendations Summary:

- Stock enough vehicles from high-demand years (<= 10 years), especially of last 3-4 years, while minimizing excess stock for slower-selling models.
- Improve consumer awareness at low sales region by targeted marketing campaigns for in-demand brands at right price point.
- Provide users with quality listings of gas type vehicles to maintain current revenue. Encourage users to purchase electric and hybrid type by sending personalized offers, to adapt to changing preference.
- Cars with 'clean' titles, less 'odometer' value, 'car age' less than 5 years and in close to newer 'condition' can be priced at a premium, to increase profits.
- Provide vehicle history reports, including detailed key attributes information and any major accidents and repairs to address concerns about mileage accuracy and potential hidden issues if any, which will build trust and consumer retention.
- Consider offering certified pre-owned programs for cars with lower mileage and thorough inspections, providing added assurance to buyers and potentially justifying higher prices.

Limitations of Analysis:

Data Scope: The dataset is of 122 years, but analysis has been done only for last 22 years to get latest purchase preference and trends.

Outliers: It has been identified and removed to maintain data integrity. Also, cars having price less than INR 5000 has been removed after removing outliers to get right pricing.