SUBMITTED BY: NIMERTA WADHWANI

CMS ID: 133-22-0009

ARTIFICIAL INTELLIGENCE &

MACHINE LEARNING

MINI PROJECT

ASSIGNED BY: SIR MUMTAZ KALOI

**Car Price Analysis Report**

**Introduction**

We were assigned a mini project by our Sir Mumtaz Kaloi, in which we had to use Python to apply basic statistical operations like mean, mode, standard deviation, and normal distribution to a real-world dataset. The goal is to analyze the data, interpret the statistical results, and draw meaningful insights that could guide decision-making in a real-world context.

And I chose the Car-Dataset available on Kaggle. The goal is to understand car prices through both statistical calculations and data visualizations. Multiple statistical evaluations were performed on car prices which appeared in the dataset for analyzing their distribution pattern.

Dataset Link: [Car-Dataset-Kaggle](https://www.kaggle.com/datasets/asinow/car-price-dataset)

**Statistical Analysis**

To analyze the car price dataset, the following statistical operations were performed:

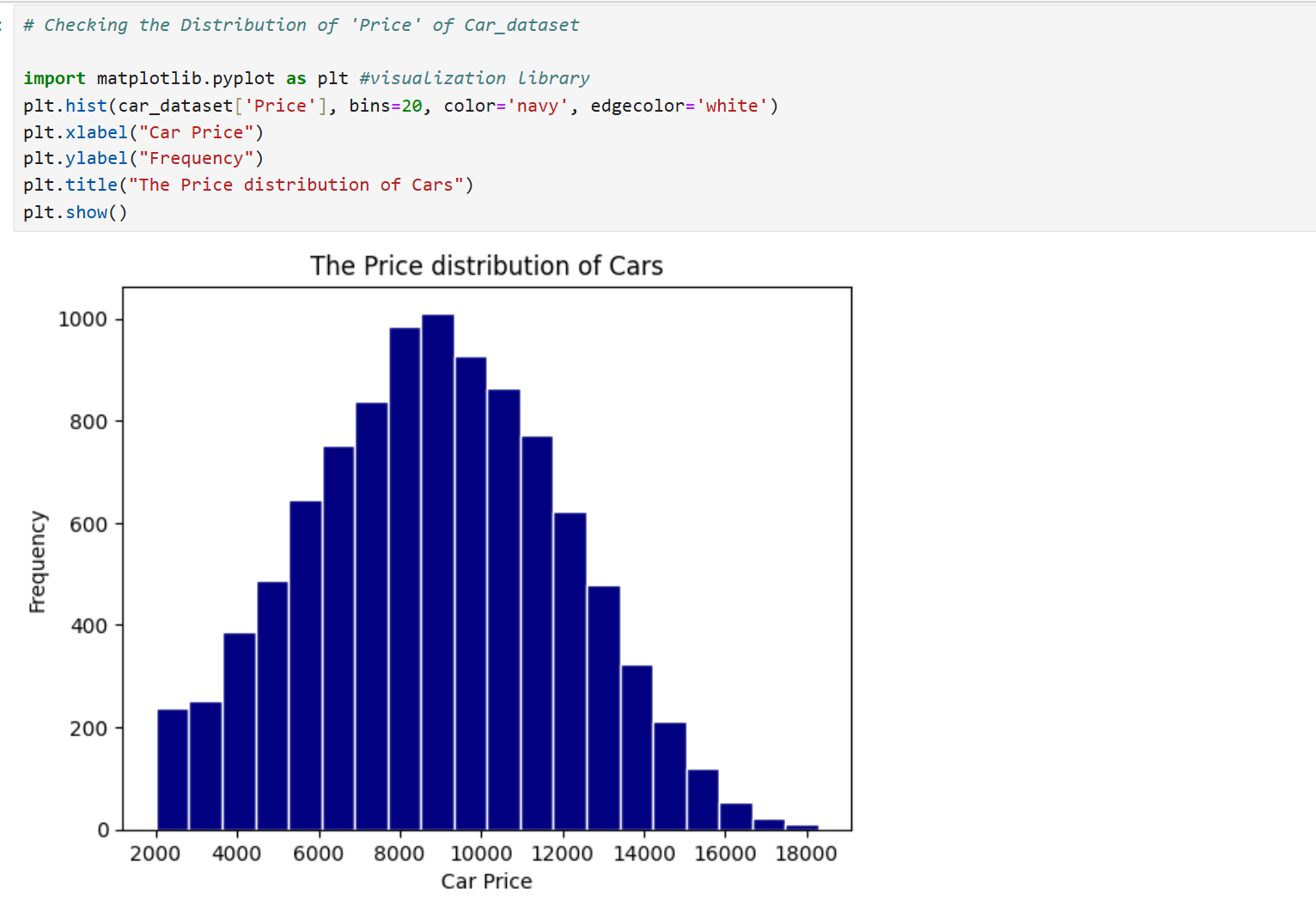


* **Mean (Average Price):** The mean price of cars in the dataset is 8852.96.
* **Median (Middle Value):** The median price is 8858.5, which indicates that half of the cars are priced below this value and half are above.
* **Mode (Most Frequent Value):** The most frequently occurring car price is 2000.
* **Standard Deviation:** The standard deviation of car prices is 3112.59, which shows the spread of prices around the mean.

**Data Distribution**

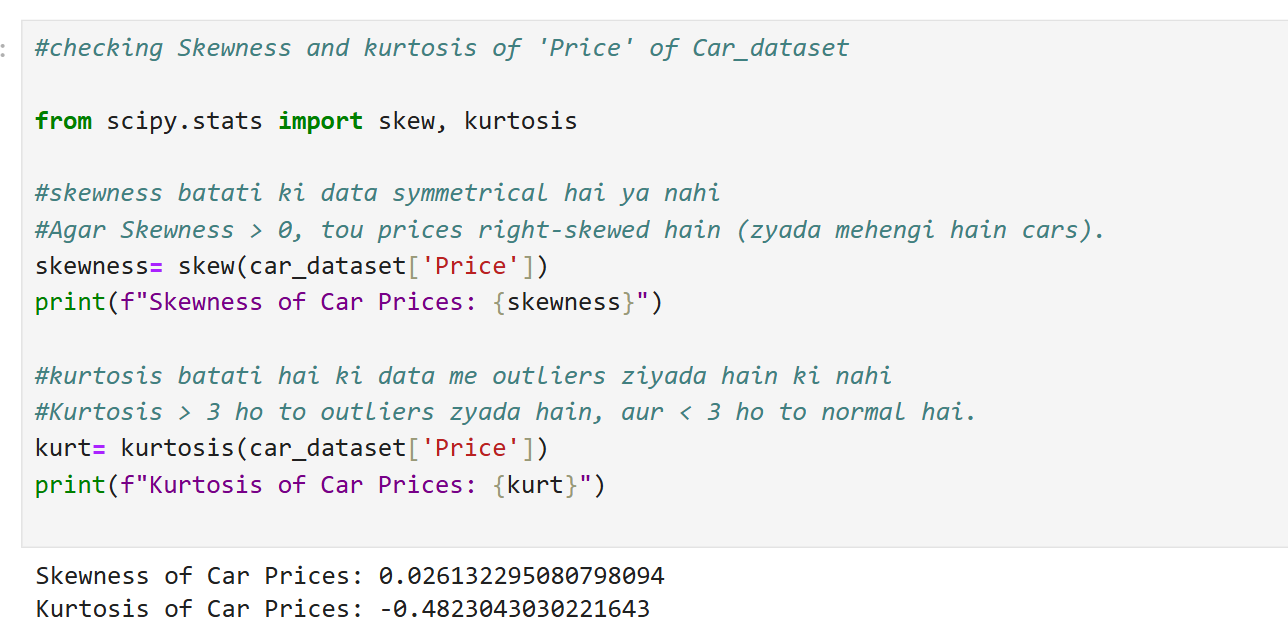
Histogram Analysis

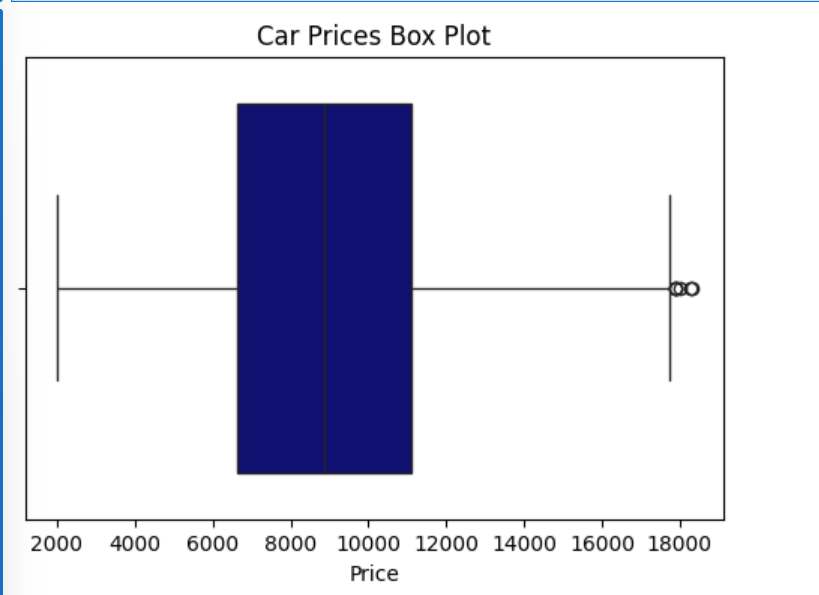
The car price distribution appeared through a histogram design that displayed frequency information. The graph depicts which price ranges occur how many times within the data. The histogram shape reveals how data distributes between normal distribution and skewed distributions.



Skewness & Kurtosis

* Skewness: The skewness of car prices is 0.026, indicating that the data is approximately symmetric.
* Kurtosis: The kurtosis value is -0.48, which suggests that the dataset has a slightly flatter distribution than a normal distribution, meaning fewer extreme outliers.

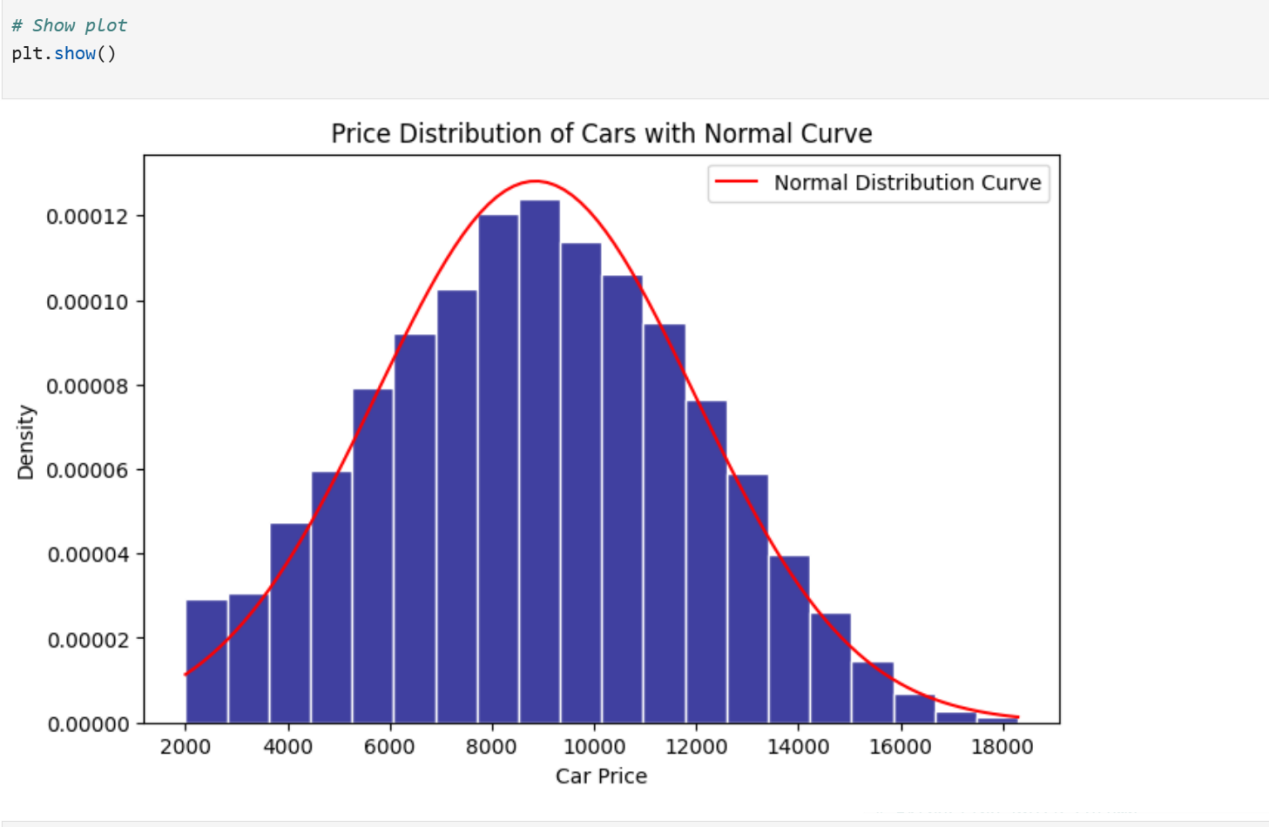


****

**Normal Distribution Curve**

The distribution of car prices was verified with a bell-shaped curve added to the histogram. Statistical computation of mean and standard deviation allowed the placement of the normal distribution curve. The price data follows a normal distribution pattern although it shows slight deviations from typical normal distributions.





**Interpretation of Results:**

*Mean:*

The mean car price functions as an average calculation created from the complete set of cars within the dataset. The variable shows what normal pricing levels appear in the car information collection.

*Mode:*

Most frequently observed price value appears as the mode. Market-based pricing strategies might be influenced by frequent occurrences of particular prices which also reveals prominent price segments.

*Standard Deviation:*

Standard deviation expresses numerical variation between prices and their mean value. A high standard deviation reveals there exists wide price differences between cars but a low standard deviation shows most vehicles maintain average price levels.

*Normal Distribution Analysis:*

The gathered data underwent evaluation for following normal distribution patterns. A normal distribution of prices in the dataset allows more accurate prediction through models that assume normal distribution.

**Conclusion**

The project delivered fundamental information about vehicle price tendencies by applying statistical calculations and visualization elements. Using this approach researchers gained their first exposure to statistical concepts that included mean and median and mode alongside standard deviation and skewness and kurtosis. Python libraries Pandas and Matplotlib and Seaborn enabled the generation of effective visualization of data.