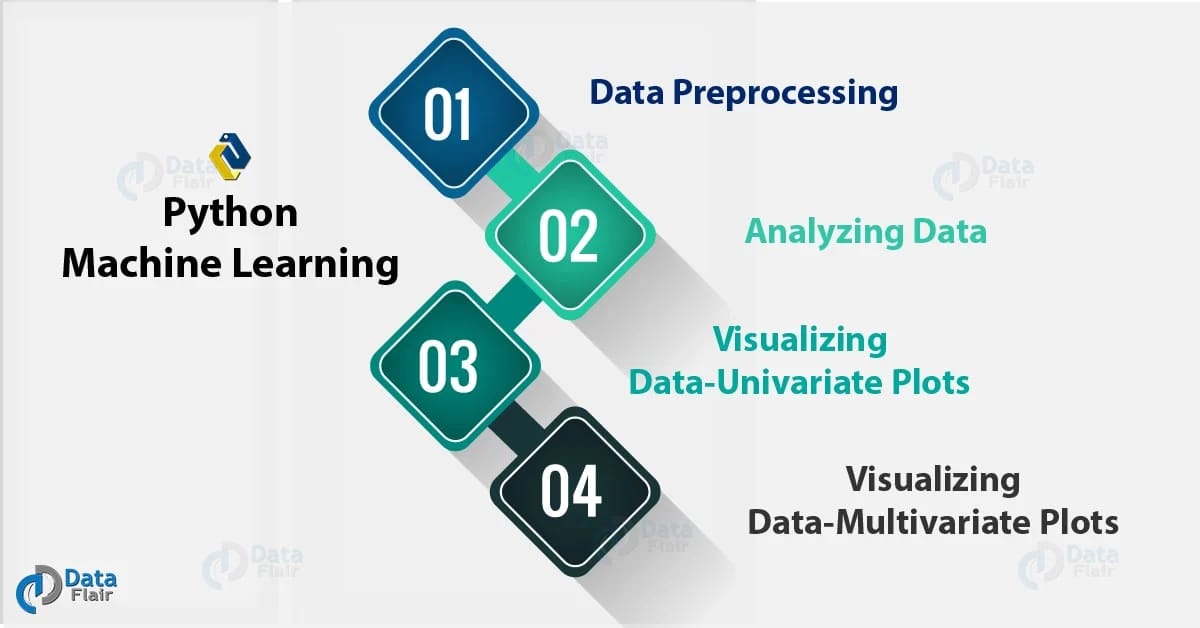
Play pivotal roles in this endeavor. Data pre-processing encompasses a range of activities, including data integration, analysis, cleaning, transformation, and dimension reduction.

Data pre-processing involves cleaning and preparing raw data to facilitate feature engineering. Meanwhile, feature engineering entails employing various techniques to manipulate the data. This may include adding or removing relevant features, handling missing data, encoding variables, and dealing with categorical variables, among other tasks.

Undoubtedly, feature engineering is a critical task that significantly influences the outcome of a model. It involves crafting new features based on existing data while pre-processing primarily focuses on cleaning and organizing the data.

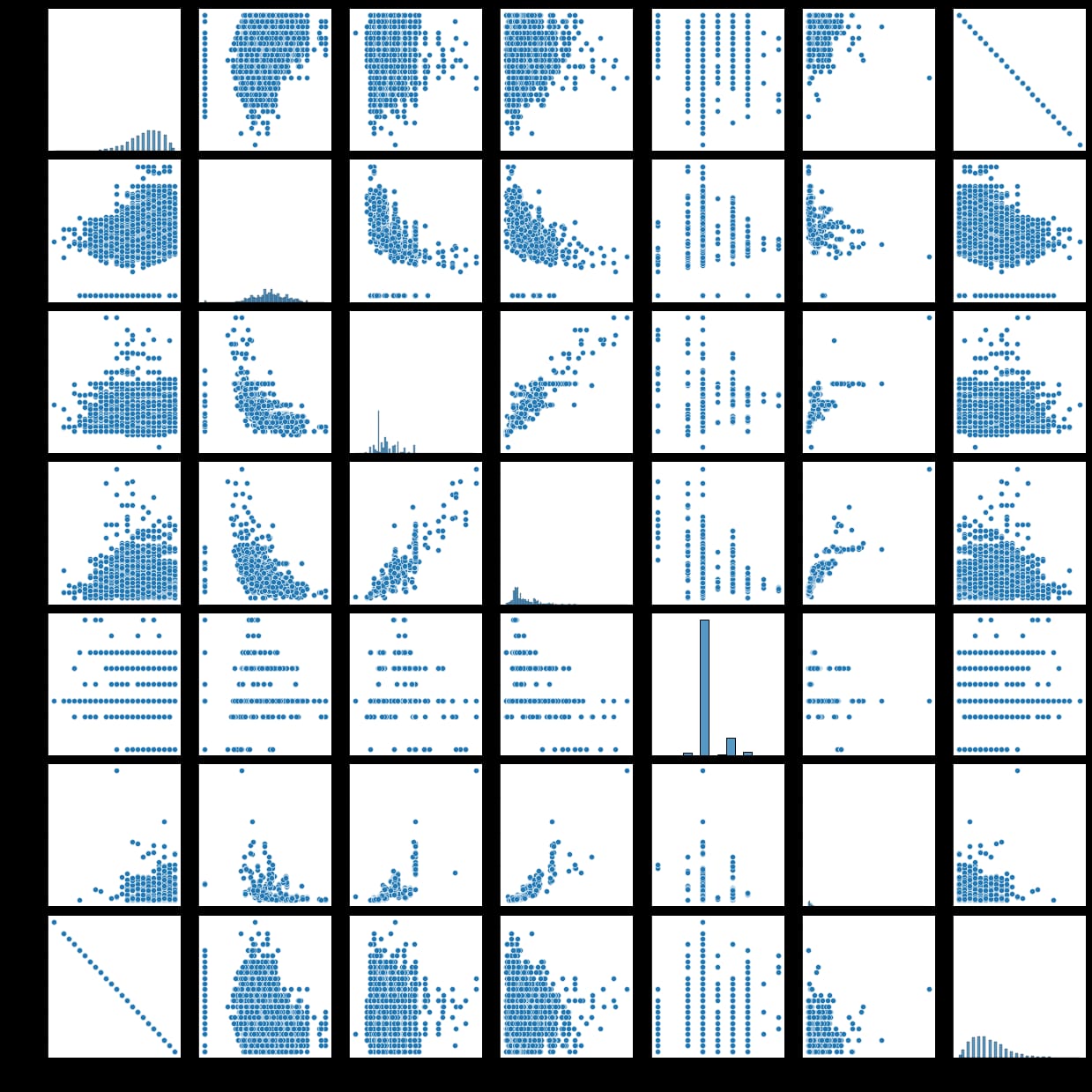
Python machine learning

Analyzing the Data

Before we make any inferences, we listen to our data by examining all variables in the data.

The main goal of data understanding is to gain general insights about the data, which covers the number of rows and columns, values in the data, datatypes, and Missing values in the dataset.

shape – shape will display the number of observations(rows) and features(columns) in the dataset

There are 7253 observations and 14 variables in our dataset

EDA Exploratory Data Analysis

Exploratory Data Analysis refers to the crucial process of performing initial investigations on data to discover patterns to check assumptions with the help of summary statistics and graphical representations.

EDA can be leveraged to check for outliers, patterns, and trends in the given data.

EDA helps to find meaningful patterns in data.

EDA provides in-depth insights into the data sets to solve our business problems.

EDA gives a clue to impute missing values in the dataset

**Feature Engineering**

engineering refers to the process of using domain knowledge to select and transform the most relevant variables from raw data when creating a predictive model using machine learning or statistical modeling. The main goal of Feature engineering is to create meaningful data from raw data.

**Conclusion**

this article, we tried to analyze the factors influencing the used car’s price.

Data Analysis helps to find the basic structure of the dataset.

Dropped columns that are not adding value to our analysis.

Performed Feature Engineering by adding some columns which contribute to our analysis.

Data Transformations have been used to normalize the columns.

We used different visualizations for EDA like Univariate, Bi-Variate, and Multivariate Analysis.

Through EDA, we got useful insights, and below are the factors influencing the price of the car and a few takeaways:

Most of the customers prefer 2 Seat cars hence the price of the 2-seat cars is higher than other cars.

The price of the car decreases as the Age of the car increases.

Customers prefer to purchase the First owner rather than the Second or Third.

Due to increased Fuel price, the customer prefers to purchase an Electric vehicle.

Automatic Transmission is easier than Manual.

This way, we perform EDA on the datasets to explore the data and extract all possible insights, which can help in model building and better decision making.

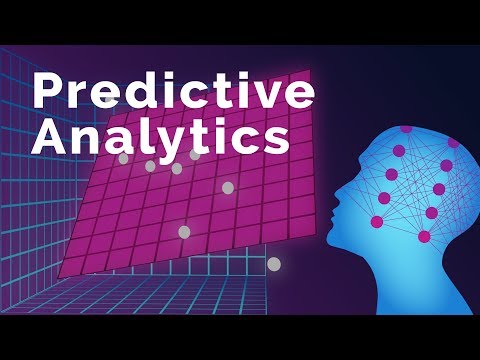
However, this was only an overview of how EDA works; you can go deeper into it and attempt the stages on larger datasets.

If the EDA process is clear and precise, our model will work better and gives higher accuracy!



**Predictive analytics in business**

analytics draws its power from a wide range of methods and technologies, including big data, data mining, statistical modeling, machine learning, and assorted mathematical processes. Organizations use predictive analytics to sift through current and historical data to detect trends and forecast events and conditions that should occur at a specific time, based on supplied parameters.

With predictive analytics, organizations can find and exploit patterns contained within data in order to detect risks and opportunities. Models can be designed, for instance, to discover relationships between various behavior factors. Such models enable the assessment of either the promise or risk presented by a particular set of conditions, guiding informed decision-making across various categories of supply chain and procurement events.