

Feature Selection: Case Study

Feature selection is a crucial process in building machine learning models which involves selecting a subset of relevant features (variables, predictors) for use in model construction.

The given dataset consists of data related to ride-sharing service costs, encompassing various factors that could influence the dynamic pricing model of rides. It includes 1,000 records with the following key attributes:

- **Number_of_Riders**: The number of riders participating in the ride.
- **Number_of_Drivers**: The number of drivers available at the time of booking.
- **Location_Category**: The category of the location (Urban, Suburban, Rural).
- **Customer_Loyalty_Status**: The loyalty status of the customer (e.g., Silver, Regular).
- **Number_of_Past_Rides**: The number of past rides the customer has had.
- **Average_Ratings**: The average ratings given to the drivers by the customer.
- **Time_of_Booking**: The time of the day when the booking was made.
- **Vehicle_Type**: The type of vehicle booked for the ride.
- **Expected_Ride_Duration**: The expected duration of the ride in minutes.
- **Historical_Cost_of_Ride**: The cost of the ride, which serves as the target variable for predictive modeling.

Your task is to conduct an Exploratory Data Analysis (EDA) to identify and select the most important features that influence the historical cost of a ride.