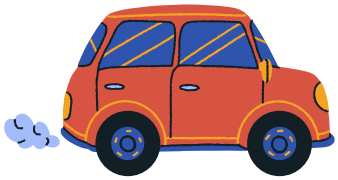
SkillCircle
THINK BEYOND NINE TO FIVE

MINI PROJECT

Exploratory Data Analysis (EDA) on Vehicle Insurance Dataset



Objective:
The primary objective of this project is to conduct an in-depth Exploratory Data Analysis (EDA) on a dataset related to vehicle insurance. Through this analysis, students will gain valuable insights into the patterns, trends, and factors influencing insurance claims. The project encompasses various aspects of data preprocessing, visualization, and statistical analysis.

Dataset Overview:
The dataset contains information related to vehicle insurance, including details about insured individuals, their vehicles, and insurance claims. Students will explore columns such as age, gender, region, insurance premiums, policy types, and more. The ultimate goal is to derive meaningful insights that can inform decision-making processes within the insurance domain.

DATASET

Key Components:

Data Loading and Inspection:
- Understand the structure of the dataset.
- Identify the types of information available.

Data Cleaning:
- Handle missing values and outliers appropriately.

Data Visualization:

- Utilize various visualization techniques to explore the distribution of key variables.

Feature Analysis:
- Examine the relationship between features and the target variable (insurance claims).

Age Distribution:
- Analyze the age distribution within the dataset and its impact on insurance claims.

Premium Analysis:
- Investigate the distribution of insurance premiums and their correlation with claim frequencies.

Claim Frequencies:
- Explore factors contributing to higher claim frequencies.

Gender Analysis:
- Investigate the role of gender in insurance claims.

Vehicle Age and Claims:
- Examine the impact of vehicle age on the likelihood of a claim.

Region-wise Analysis:

- Analyze regional patterns in insurance claims.

Policy Analysis:

- Explore the distribution and impact of different insurance policy types.

Claim Frequency by Vehicle Damage:

- Investigate the relationship between vehicle damage and claim frequencies.

Customer Loyalty:

- Analyze if the number of policies held by a customer influences claim likelihood.

Time Analysis:

- If applicable, explore temporal patterns in insurance claims.

The project will conclude with a summary of key findings and insights derived from the EDA. Students will gain practical experience in data analysis, hypothesis generation, and interpretation of results, providing a solid foundation for more advanced studies in data science.

Note: This project serves as a valuable exercise for students to apply EDA techniques in a real-world scenario and enhance their skills in data exploration and analysis.