LOGISTICS REGRESSION

BY: HONEY NEIL SHAHU

Data Understanding

To perform data understanding, preparation, and exploratory data analysis (EDA) on the provided dataset, we can follow these steps:

- **Load the data** and understand its structure.
- **Summarize the data** to get basic statistics.
- **Identify missing values** and handle them.
- **Analyze categorical and numerical features**.
- **Visualize the data** to identify patterns and relationships.

- For briefing the data we followed following steps:

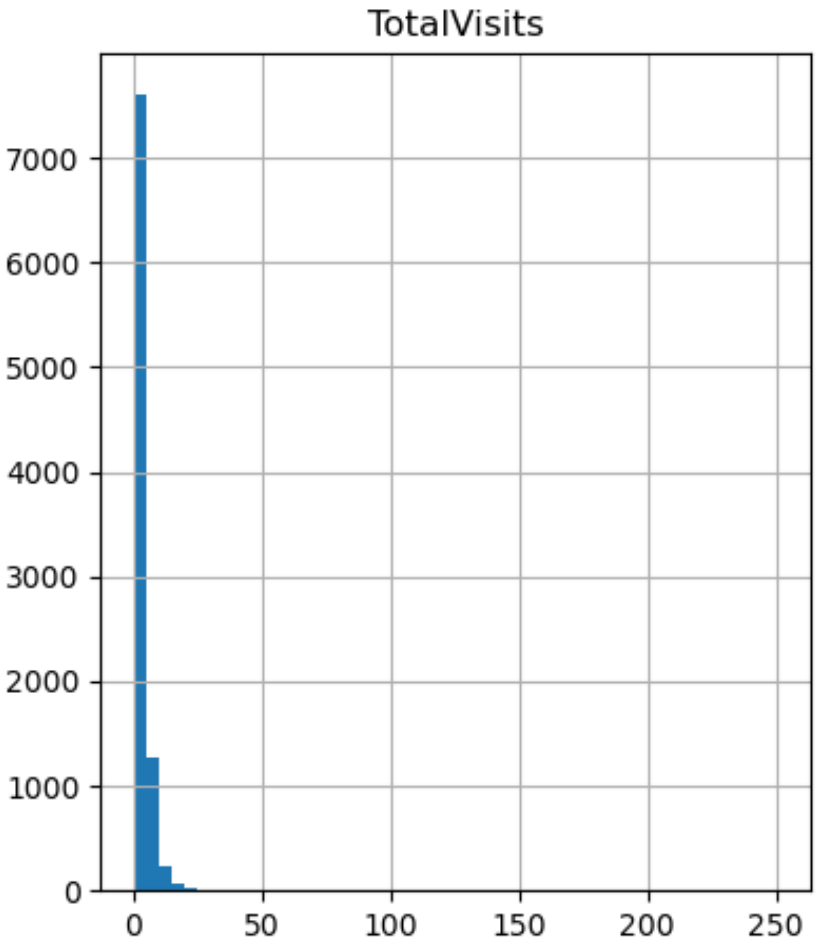
1. Loading the data
2. Summarizing the data
3. Identifying the missing values
4. Analysing categorical and numerical features
5. Visualizing the data

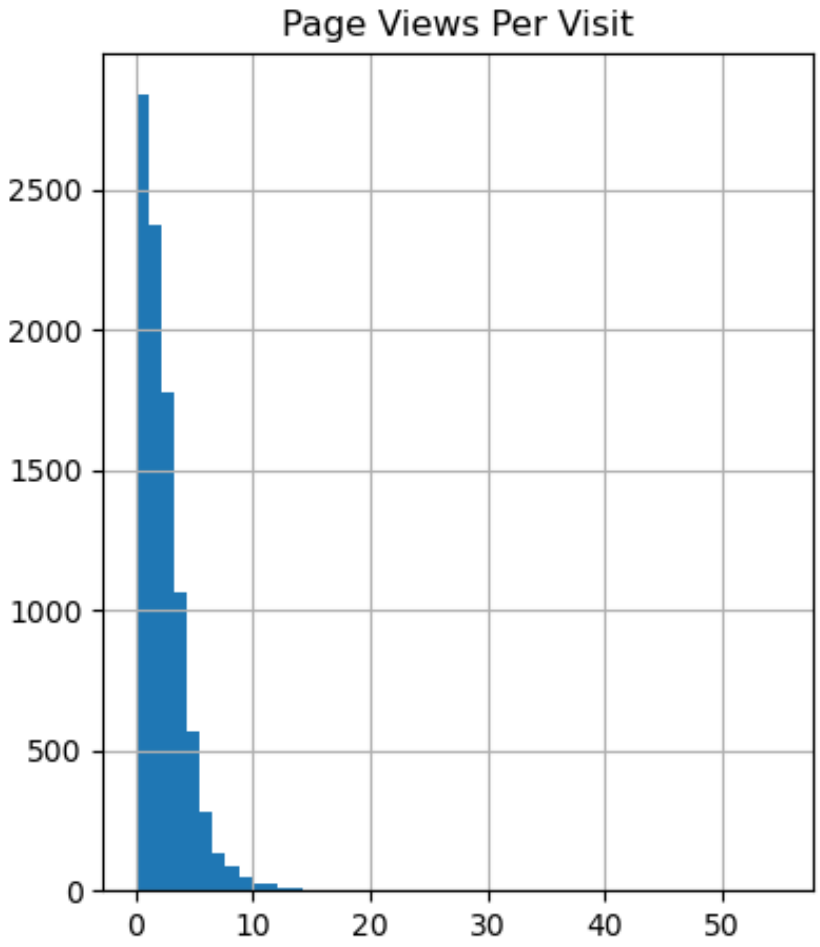
Model Building and Evaluation

For model building and evaluation using the same dataset, we'll follow these steps:

- **Preprocess the data**: Handle missing values and encode categorical variables.
- **Split the data**: Divide the data into training and test sets.
- **Build the model**: Train a logistic regression model.
- **Evaluate the model**: Assess the model's performance using appropriate metrics.

EDA


TotalVisits
7000
6000
5000
4000
3000
2000
1000
0
0
50
100
150
200
250


Page Views Per Visit
2500
2000
1500
1000
500
0
0
10
20
30
40
50

1. Preprocess and Training the Model

- Define the preprocessing for numeric and categorical features
- Combine preprocessing steps
- Define the target variable
- Separate features and target variable
- Apply preprocessing
- Split the data into training and test sets
- Initialize and train the logistic regression model

2. Retrieve and analyze the coefficients

- Get feature names after one-hot encoding
- Get the coefficients of the logistic regression model
- Create a DataFrame to view feature names with their corresponding coefficients
- Sort the DataFrame by the absolute value of the coefficients
- Display the top 3 features

Thank You