LEADS SCORING PROJECT

PROBLEM STATEMENT

- X Education sells online courses to industry professionals.

❏ X Education gets a lot of leads, its lead conversion rate is very poor. For example, if,   say, they acquire 100 leads in a day, only about 30 of them are converted.

❏ To make this process more efficient, the company wishes to identify the most   potential leads, also knownas 'Hot Leads'.

❏ If they successfully identify this set of leads, the lead conversion rate should go upas the sales team will now be focusing more on communicating with the potential leads   rather than making calls to everyone.
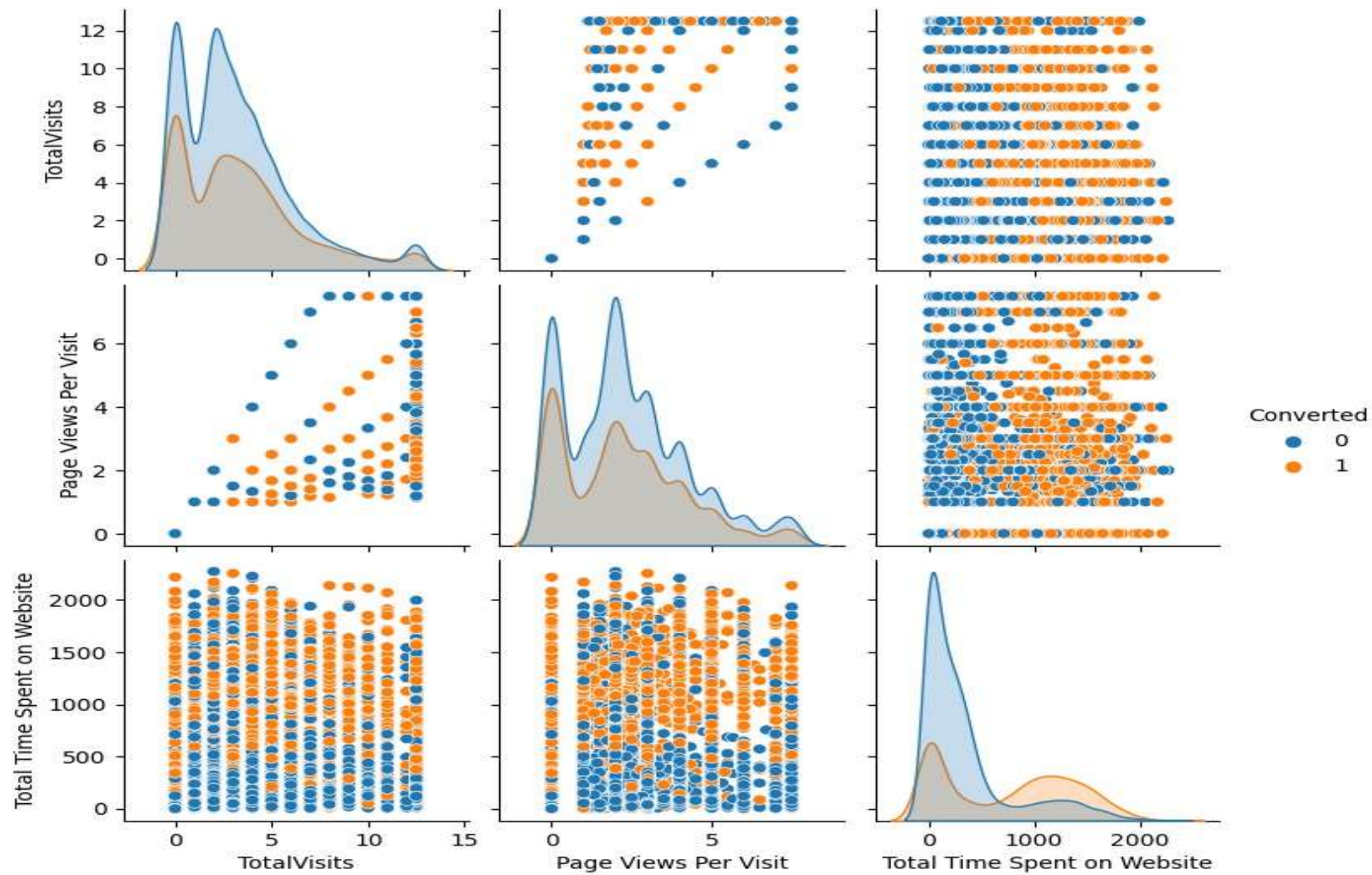
PROPOSED METHODOLOGY

Data cleaning and data manipulation

1. Check and handle duplicate data.

2. Check and handle NA values and missing values.

3. Drop columns, if it contains a large number of missing values and are not useful for the analysis.

4. Imputation of the values, if necessary.
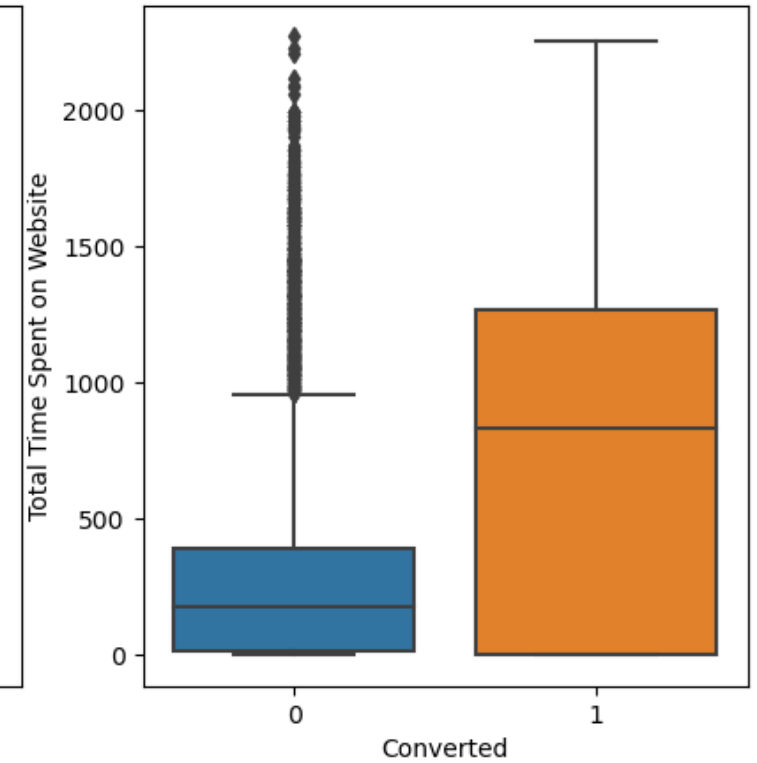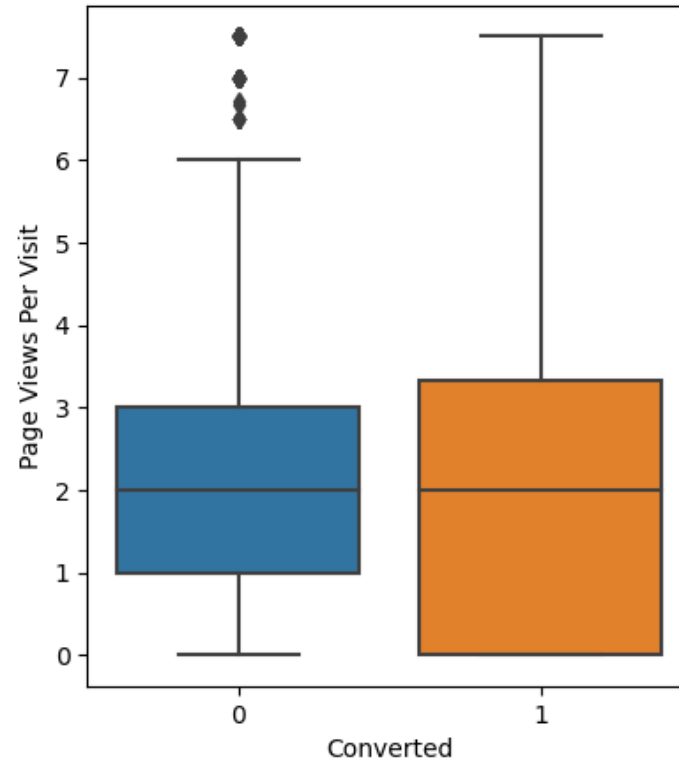
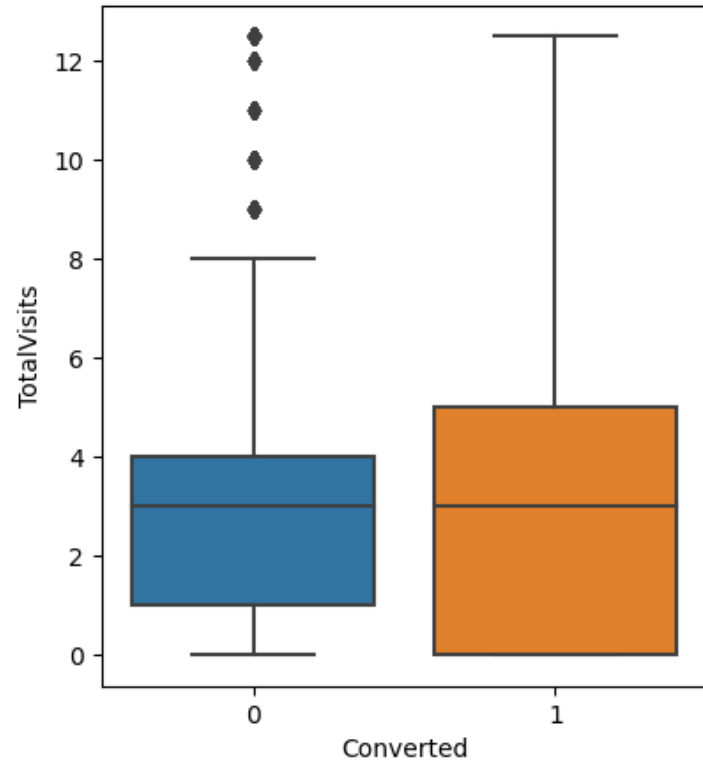5. Check and handle outliers in data.

Exploratory Data Analysis (EDA)

1. Univariate data analysis: value count, distribution of variables, etc.

2. Bivariate data analysis: correlation coefficients and pattern between the variables etc.

3. Feature Scaling & Dummy variables and encoding of the data.

4. Classification technique: logistic regression is used for model making and prediction.

5. Validation of the model.

6. Model presentation

EDA

Box Plot

Heat Map