

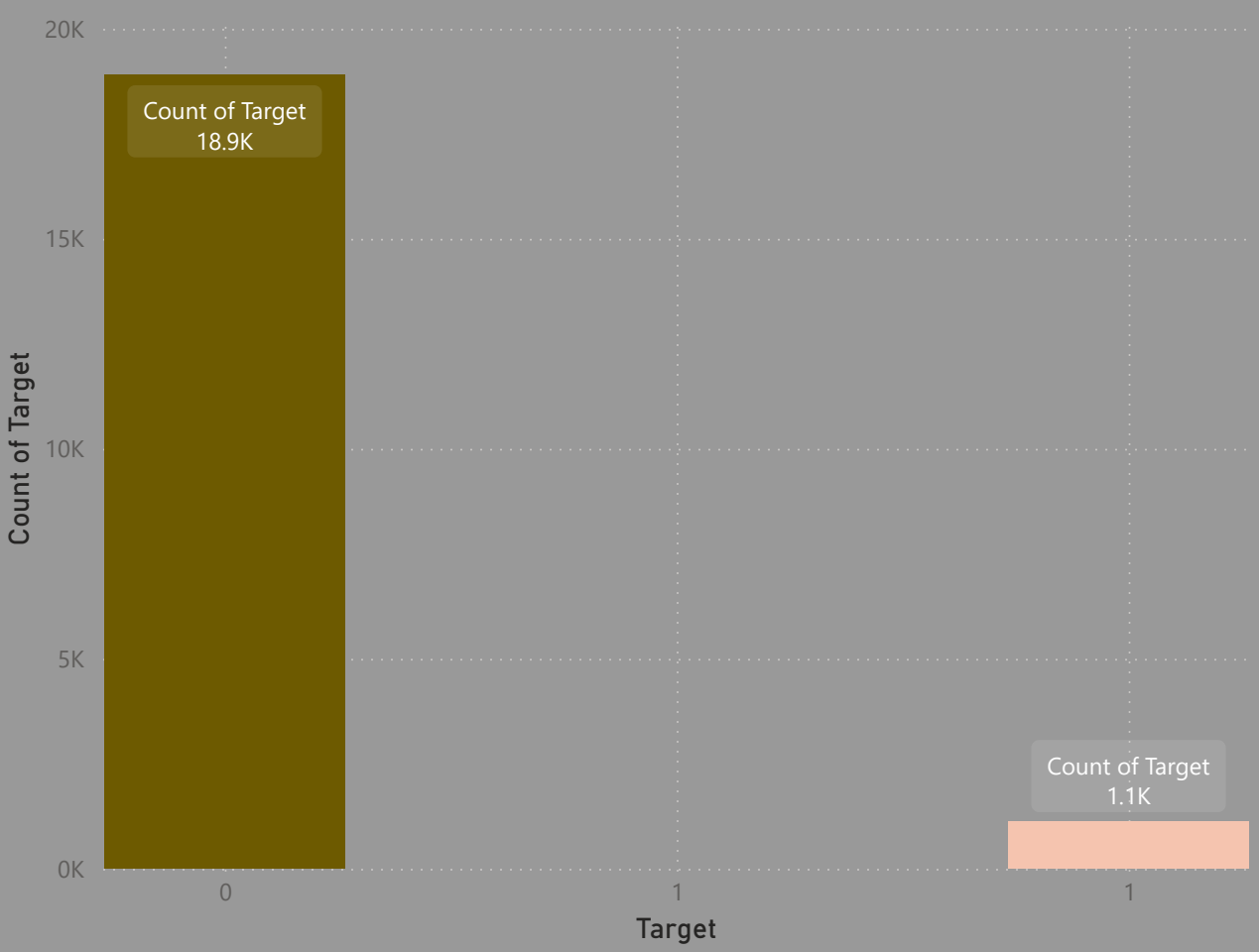
Overview of the report

- This report contains initial exploratory data analysis of Rene wind turbine maintenance project.
- Here I have tried to do an initial analysis on the available features. I have tried to find out whether they have linear or non-linear relationships among them and with the target variable.
- Also, I have tried to find out how strongly they are correlated to each other and to the target variable.
- I have also tried to gain information about the distribution of features and made dynamic box plots to detect outliers.
- I have also tried to make the best fit line for each feature based on scatter plots.

Target
1
0

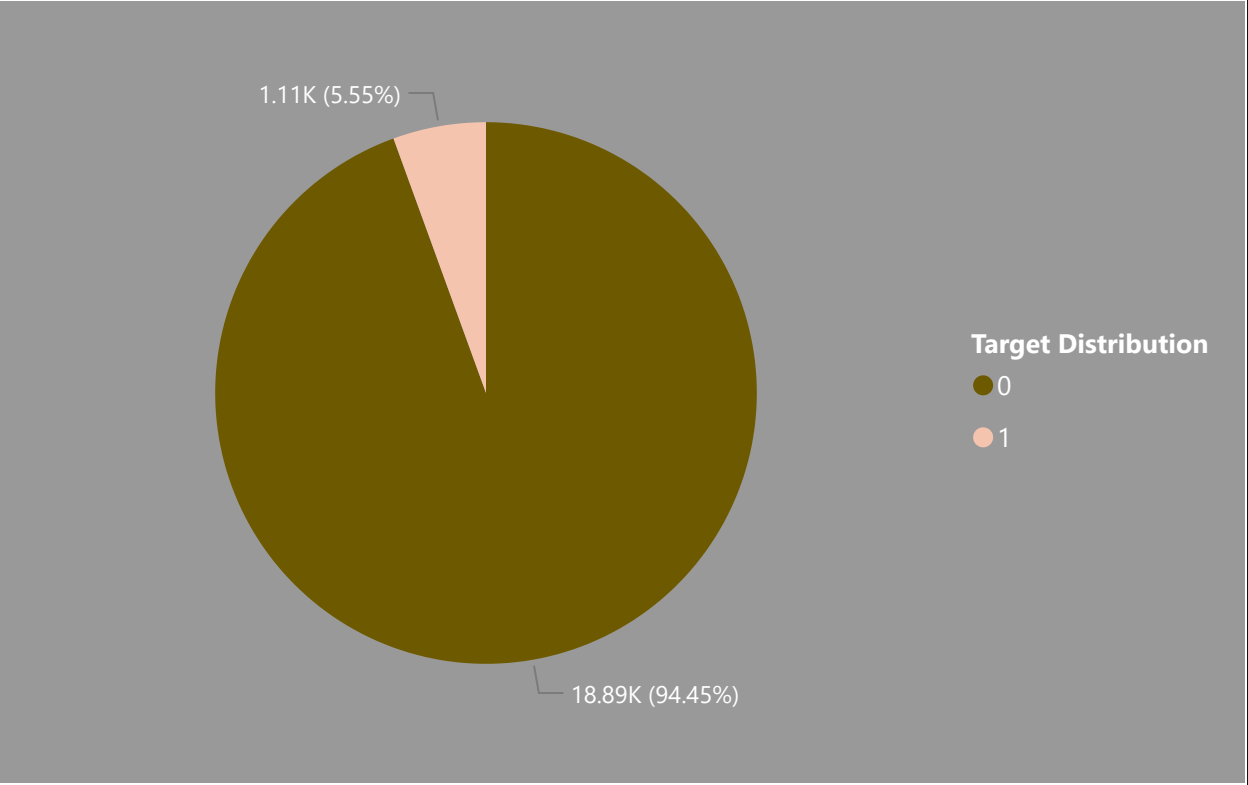
Count of Target

Count of Target by Target



Count of Target

Count of Target by Target



Features

☒ V1

☐ V10

☐ V11

☐ V12

☐ V13

☐ V14

☐ V15

☐ V16

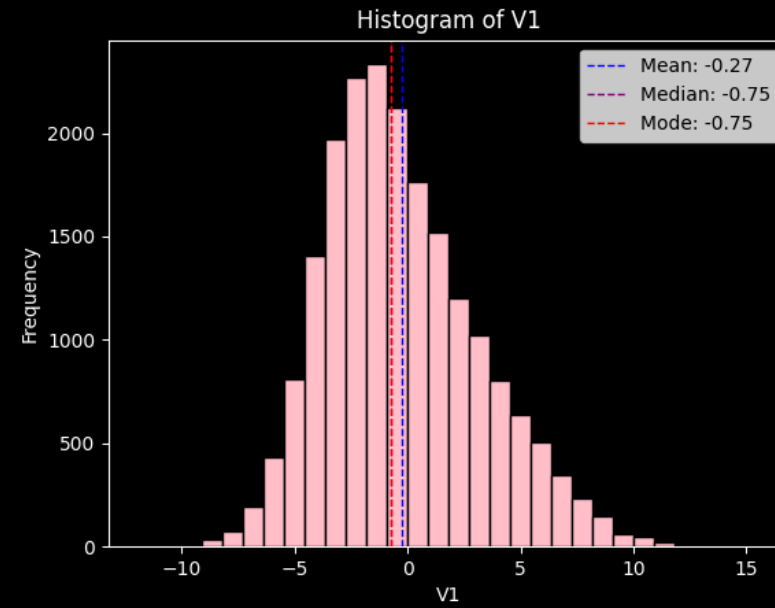
☐ V17

☐ V18

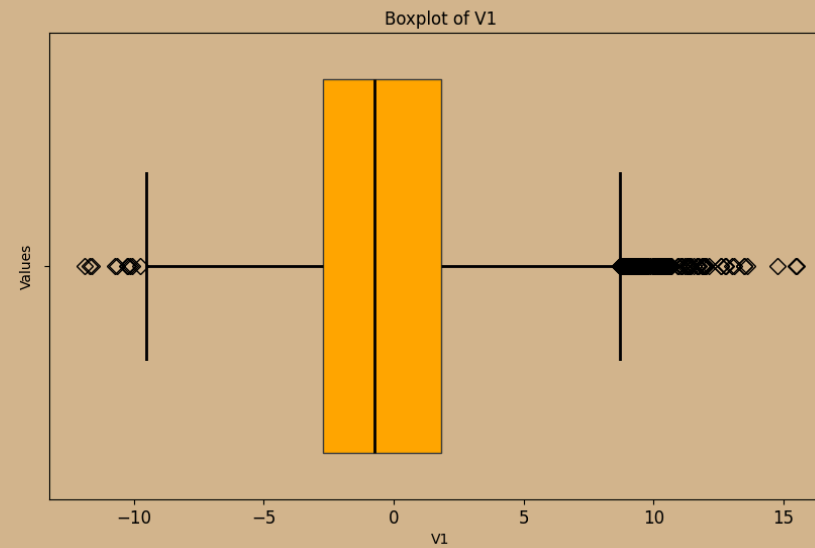
☐ V19

☐ V2

Distribution of different features



Box plots of different features



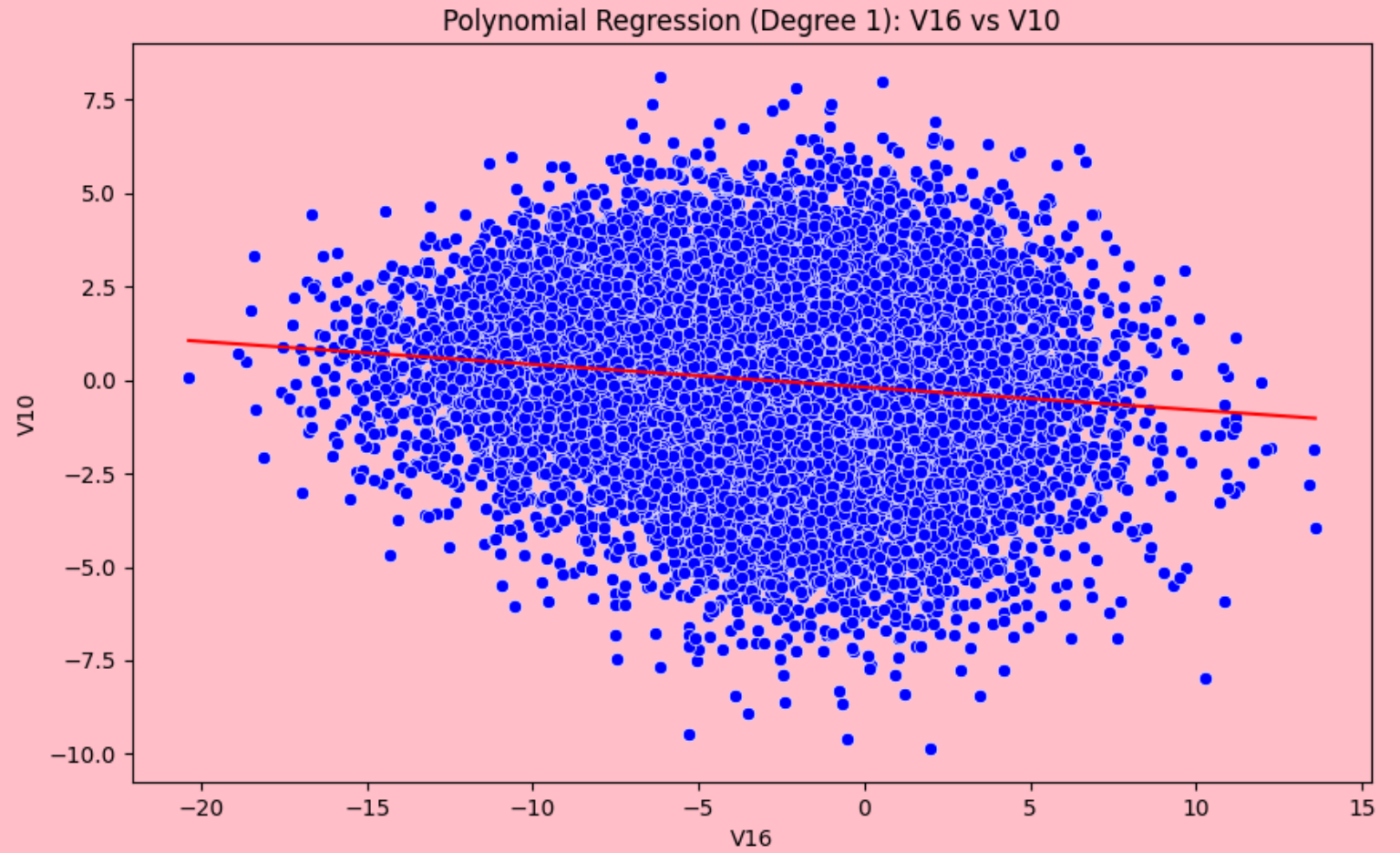
X Variables

- ☐ V1
- ☐ V10
- ☐ V11
- ☐ V12
- ☐ V13
- ☐ V14
- ☐ V15
- ☒ V16
- ☐ V17
- ☐ V18

Y Variables

- ☐ V1
- ☒ V10
- ☐ V11
- ☐ V12
- ☐ V13
- ☐ V14
- ☐ V15
- ☐ V16
- ☐ V17
- ☐ V18

Best fit line plot



Feature - Target point-biserial correlation after SMOTE

