

# Credit Card Fraud detection progress report

By Mohammed Shati

# The dataset columns

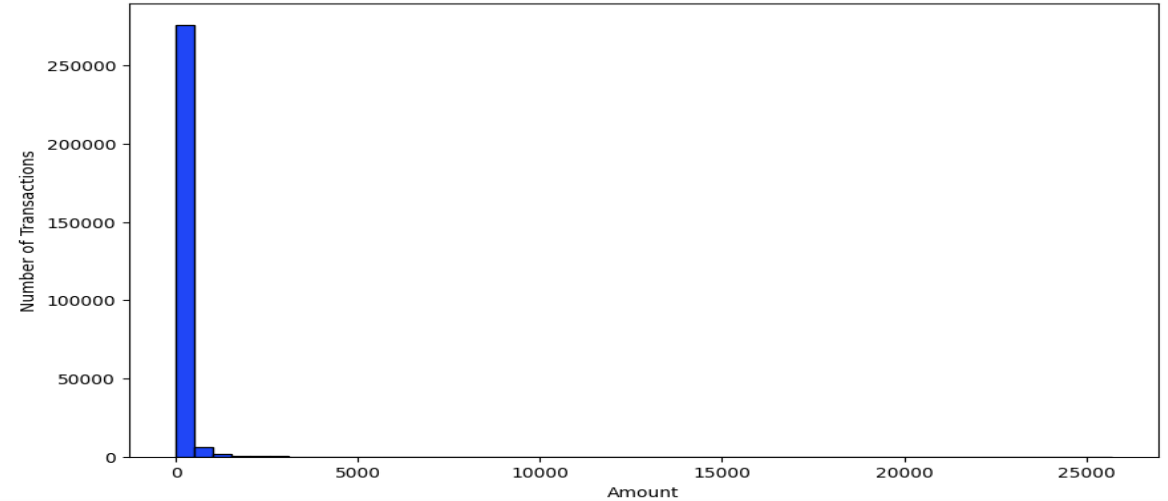
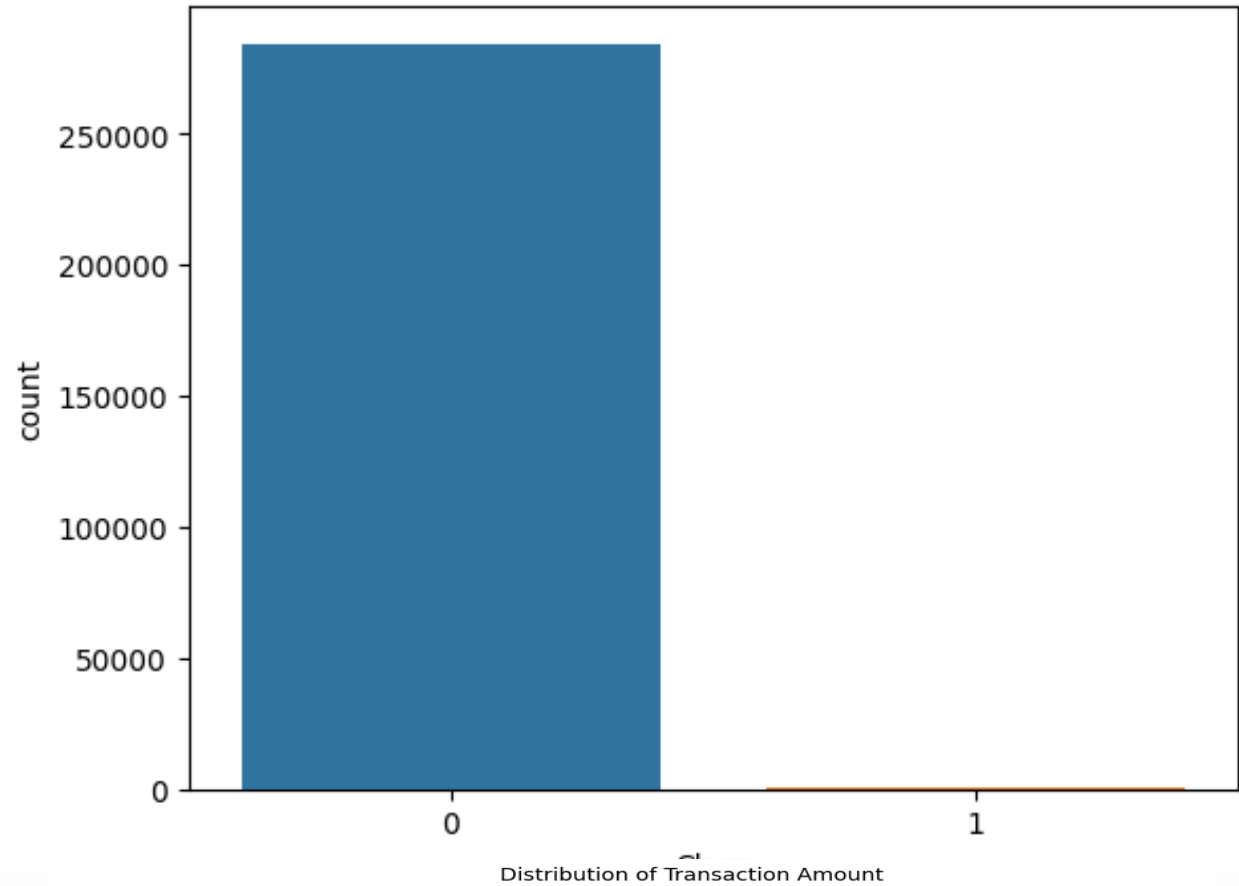
Time: The time (in seconds) elapsed between the current transaction and the first transaction in the dataset. V1 to V28: These are the anonymized features resulting from the PCA transformation. These are the primary components extracted from the original features, and their specific real-world interpretation is not publicly available. Amount: The transaction amount. Class: This is the target variable, where 0 indicates a legitimate transaction and 1 indicates a fraudulent transaction.

# The null and info

Inspecting the null values there is no null values and every column type is float

# EDA

Some EDA in here on the very top there is the class imbalance fraudulent vs legitimate. The dataset has very high legitimate transactions. In the very bottom on the right, the transaction amount is in hundreds. So I will deal with the imbalance dataset in the later phase of the project.



# Standard scaling and imputer

I built a pipeline and I added standard scaling and fitted them to the trainset