

Class 2, Task 2 CreditOne

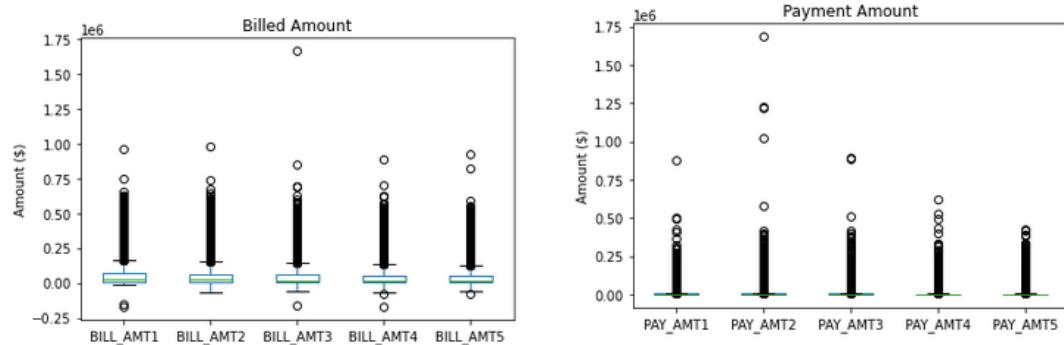
Population View

Client's average age is 35.49 years-old with 75% of clients being 41 years-old or younger. Most of the clients (82%) have, at least, a university degree. Also, 99% are married or single, and we have a 60/40 split for female to male.

75% of the clients have a credit limit of \$240K or below. The average credit limit is \$167K with a large Std. deviation (\$130K), the average monthly bill is \$63K and average monthly payment is \$5.2K.

Payment and Balances

Monthly balance distribution remains similar from month to month. Payments however, look more scatter as time elapses. It can be people trying to do catch up payments. Some balances are negative which means a client paid more than necessary (creating a credit)



Delinquency

- Bar plots show that delinquency has been increasing every month which is what CreditOne stated. September is the worst month with increased cases of one- and 3-month delays. Scatter plots also validate this month-by-month increase.
- Correlation (-0.32) showed that there is a tendency where the lower the credit limit, the higher the incidence of delinquency. Also, there is a small tendency (0.28) where the higher the bill to pay, the higher delinquency cases.
- Divorce clients seem to have the worst time paying on-time.
- Other values were evaluated vs. delinquency but not much relation was found and they followed the overall composition of the data sample. Ex: Visualization shows that females have a higher incidence of delinquency than males (circa 3:1 ratio), but this is expected as 60% of the clients are females.

Initial Recommendations

Data on the state of the economy (Ex: Employment rate, inflation, interest rates) and more information on client's income would be helpful to create an even better predictive model for default and credit limits.

Data Wrangling

We created a parallel dataframe that contains the fields in the table which some will be used in the model development phase. The goal was to use these variables as a reduction of the data set, but still meaningful while running thru correlation and visualization.

| | |
|----------------------|---|
| Bill_Avg, Pay_Avg | Average of monthly bill and payment |
| Bill_Amt, Pay_Amt | Sum of all 6 months of monthly bills and payments |
| Pay_Act | Sum of monthly payment status which reflects the tendency of payment by the client |
| Gender, Def, Edu_Num | Converted to numeric value, so we could get an idea of its correlation with other data. |
| AgeBin | Binning age for visualization. |