Problem Statement: Goal is to build a model to accurately predict (target column APPLICATION.STATUS) the status of a loan application. Please acclimatise yourself with the data first. You have to build -

1. Straight Through Process: Classify which cases are automatically (basis the set rules) being Approved or Declined by the system.
2. Underwriter Cases: How will you build a predictive model to automate the eventual decision being manually made by the Underwriters (Approved or Declined).

Summary:
Target Variable - Column "APPLICATION.STATUS"
Unique Loan Number - Column "APPLICATION.ID"
Straight Through - Column "QUEUE.ID", value is "Straight Through Process"
Underwriter - Column "QUEUE.ID", value is "Under.Writer"

Developer Notes:

1. Straight through processed is a scenario where the system was automatically able to decision the loan application basis the configured business rules engine. For e.g. a credit score of < 700 is a decline, etc.
2. Underwriter decisionsed are the ones where the system was unable to auto decision and a credit officer had to manually review and decision the loan application.
3. Ideally a good credit risk model for unsecured loans (https://www.investopedia.com/terms/u/unsecuredloan.asp) should have a STP rate of more than 95%.

Dataset: