

PDAT 617 Assignment 3 Due on 06/29 11:59pm

Introduction: Customer Churn, also known as customer attrition, or customer turnover, is the loss of customers and it is an important and challenging problem for ecommerce and online businesses. The dataset used for this project was obtained from Kaggle. It contains data about customers who are withdrawing their account from a bank. Please finish the following questions based on the given data.

1. Load the data into a pandas dataframe (10 points)
2. Determine whether the missing values exist in dataframe (10 points)
3. Drop these 4 columns 'RowNumber', 'CustomerId', 'Surname', and 'CreditScore' in dataframe (10 points)
4. Select customers who live in France and have churned. Notice churned means Exited = 1 (15 points)
5. Select customers whose bank balance are between \$80,000 to \$100,000. (15 points)
6. Group the rows based on the geography-gender combinations and then give the average churn rate for each group. (20 points)
7. Set Geography column as the index. (10 points)
8. Check the number of unique Geography. (10 points)