**JPMC Case # 1 Description**

*In this case, you will use Tableau (*[*www.tableau.com*](http://www.tableau.com)*) to perform Exploratory Data Analysis of the two data sets provided.*

You are a data analyst working for a very large financial institution and for the next several weeks you have been asked to review and examine two large data sets that have been provided for you. The data sets are:

1. **Modified ATM Withdrawals (Tableau hyper file, n= 11,094,637).**

This data set has all withdrawals from various ATM devices across the US. The data includes date and time of the withdrawals, and the amount of $20 bills withdrawn. These amounts can be anywhere from 1 – 25, so $20 to $500 (USD). There is no data on deposits, and no location is provided for any of the ATMs. That data can be provided in a future release of the data set, or can be merge at a later time, but at this moment, all we have is date, time, ATM number and withdrawals.

1. **Modified Household Data (Tableau hyper file, n= 2,125,591).**

This data set is a snippet of customers and their interactions with the financial institution for a 12 month period. The data includes information about each client, their approximate age (grouped), their occupation (grouped), if they are a checking account holder, a user of the mobile application and how often they use an ATM or at Teller at a physical location. The groupings for age and occupation will limit our analysis, but that is the data we currently have.

So, let’s get a better understanding of these two data sets and begin to explore any insights that the data can provide us. We will approach the data with questions that are of interest to us, and our management team, and determine what additional business questions we might be able to answer and suggestions for additional data that we could utilize.

**Modified Household Data**

Question #1 -- What is the distribution of clients by age range (age grouping)? Are there any insights here? How could this data be improved?

Question #2 – Taking the age range as a starting point, use the other dimensions (so, primary checking, mobile app, ATM withdrawals, etc) to describe these customers.

**Modified ATM Withdrawal Data**

Question #1 – Based on all the ATMs, how do withdrawals vary over the calendar year? Is there are any trend in the data?

Question #2 – How do withdrawals vary over the day (use a 24 hour clock)? What trends do you see?

**Data Dictionary**

Modified ATM Withdrawals Data Set

|  |  |
| --- | --- |
| ***Variable*** | ***Description*** |
| ATM | ATM ID. Unique identifier for each ATM. |
| Date | Calendar Date of Transaction. Includes Month, Day, Year |
| Time | Military Time (24 hour clock) of transaction. Tableau has a default year, which will not be used, don’t be confused by it. |
| Trans AMT | The number of $20 bills removed in the withdrawal transaction |

Modified Household Extract Data Set

|  |  |
| --- | --- |
| ***Variable*** | ***Description*** |
| ID | Unique Identifier for each customer |
| Period | Date of Transaction |
| Checking | Does the customer have a checking account? True or False |
| Debit | Does the customer have a debit card? True or False |
| Debit Active | Has the customer used the debit card? True or False |
| Onl Mob Enrolled | Is the customer enrolled in online/mobile application? True or False |
| Onl Mob Act | Is the online/mobile application active? True or False |
| Prim Chk | Does the customer have a primary checking account with Chase? True or False |
| Check Dep LT200 | The number of deposits less than $200 during the time period |
| Check Dep GT200 | The number of deposits greater than $200 during the time period |
| Teller Cash Deposit | The number of deposits (any amount) using the teller at a Chase physical location |
| Teller WDLT500 | The number of withdrawals (less than $500)during the time period using the teller at a Chase physical location |
| Teller WDGT500 | The number of withdrawals (greater than $500) during the time period using the teller at a Chase physical location |
| ATM Deposits | The number of deposits made by the customer using an ATM |
| ATM Withdrawals | The number of withdrawals made by the customer using an ATM |
| ATM Transfers | The number of transfers made by the customer using an ATM |
| QK Pay Received | The number of transactions during the period in which the customer received any amount via QK Pay |
| QK Pay Sent | The number of transactions during the period in which the customer used QK Pay to send money to another individual |
| Mobile Deposit | The number of deposits made by the customer using their mobile device during the period |
| Phone Transfer | The number of transactions made by the customer during the time period in which they transferred money from one account to another on the phone |
| Online Transfer | The number of transactions made by the customer during the time period in which they transferred money from one account to another using the web app |
| Teller Transactions | The number of transactions made by the customer during the time period in which they walked into the branch and worked with a teller |
| Teller Deposits | The number of transactions made by the customer during the time period in which they walked into a branch and made a deposit with a teller |
| Age Range | The approximate age of the customer |
| Occupation | Dropdown choice of occupation for the customer. Very limited. |