**CEBD 1160: Branch performance analysis**

|  |  |
| --- | --- |
| **Name** | **Date** |
| **Arnab Roy** | **March 27, 2020** |

### **Resources**

* Python script for my analysis: project.py
* Results figure/saved file: figures/
* Dataset : branch\_data.xlsx
* Runtime-instructions in a file named RUNME.md

**Research Question**

Based on total account and total deposit amount of two conjugative months can we determine the performance of individual branch of the bank?

**Abstract**

In this project I have used a dataset of a bank named Southeast Bank Bangladesh Limited [1], which contains data of two months: April 2018 and May 2018.  Using these data, we may be able to visualize difference between performances of branches which will help top management of bank to take necessary measure to improve the performance of branches.

**Introduction**

The bank offered some new facilities (i.e.: offer more interest) to encourages its clients to do more deposit and also they introduced new features (i.e.: free debit cards, reduced account maintenance fee) for new clients. Bank management wanted to saw the response of client from year 2018 and wanted to decide whether they offer something like that in year 2020. To visualize the data I have used different libraries of python such as matplotlib, plotly [2] [3] etc. and I have also used Machine learning algorithm sklearn [4].

**Methods**

The data was organized into an excel file and I have uploaded it using pandas library and then save in to a data frame. After that I used this dataset to solve different problems. In following graphs I would like to show the growth (+/-) of Accounts and Total deposit amount for individual branches as well as overall bank in two months.

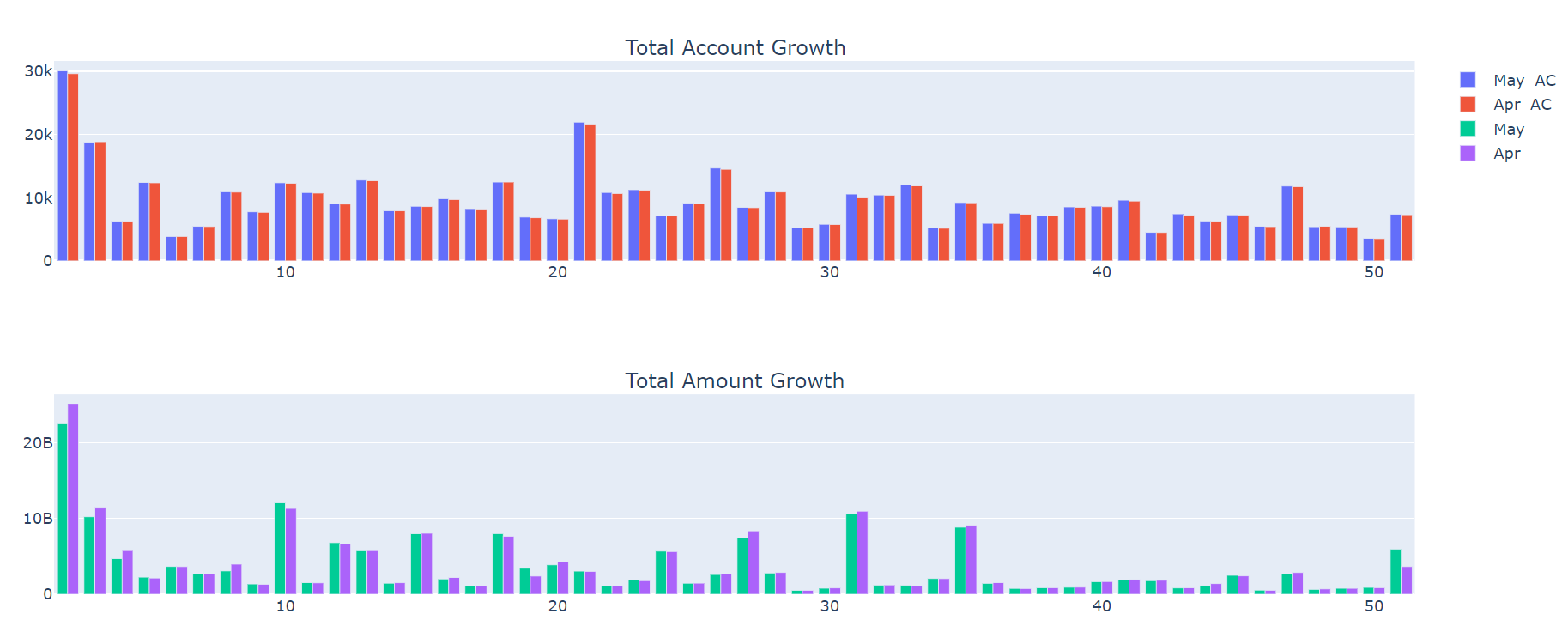


Figure 1: Branch wise Total Account and Total deposit

To show the consolidate growth of bank please see the file named: Total ac growth.html and Total amt growth.html.

Here I have also shown the overall plotting (using library seaborn) and heatmap (using library plotly) of three types of deposit called no cost deposit, low cost deposit and high cost deposit. No cost deposit accounts has no interest, such as chequing accounts, low cost deposit account has very low amount of deposit such as savings accounts, and high cost deposit accounts has high interest rates such as Fixed term deposit (FDR).

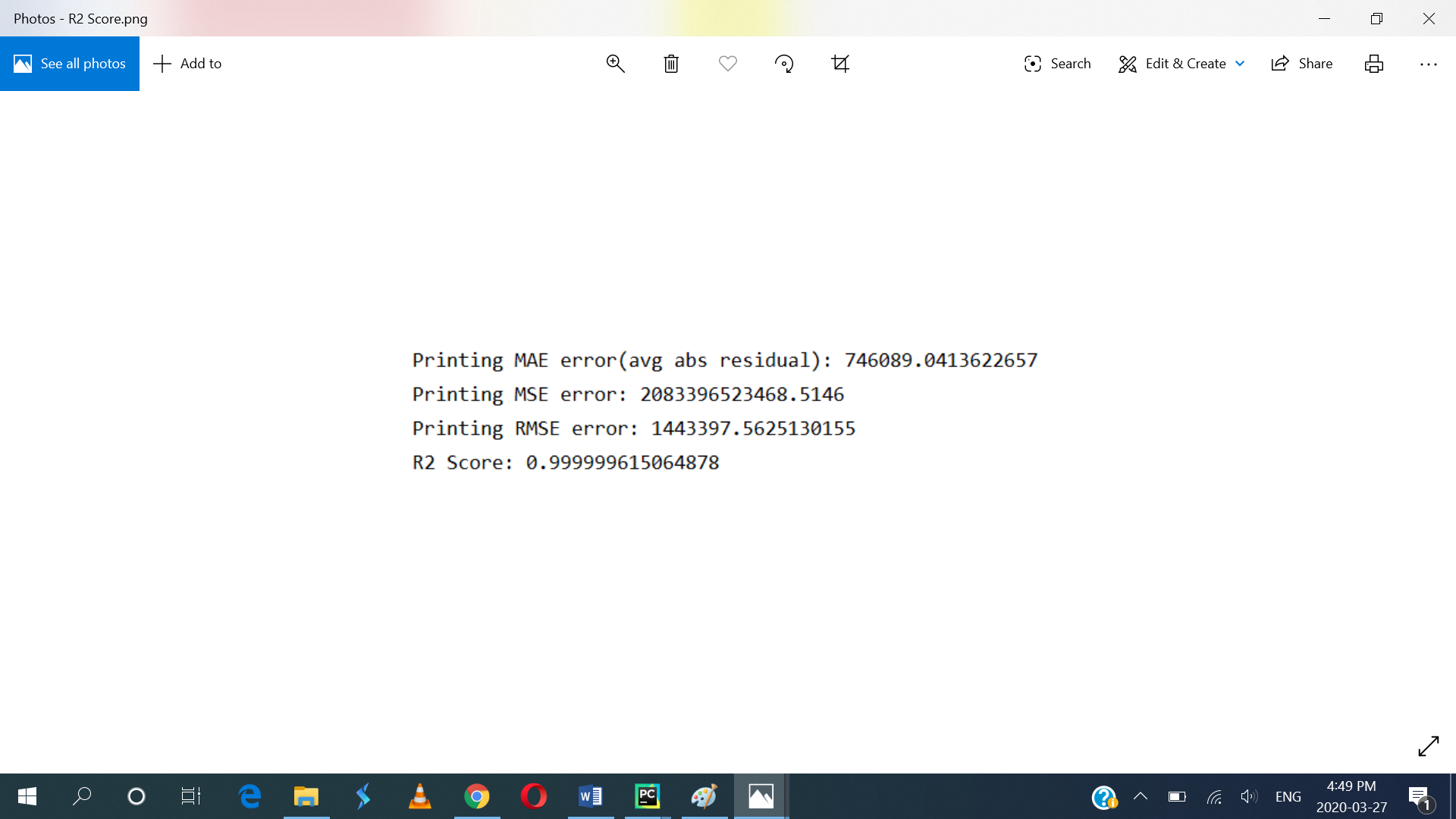


Figure 2: Overall plotting of different types of accounts



Figure 3: Heat map of different types of accounts

Finally I tried to use ML algorithm to predict branch wise total amount of May 2018. Following is different score of this prediction: For



For more machine learning plots please see figures: ML Fig 1, ML Fig 2, and ML Fig 3

**Discussion:**

Here I tried to give a clear idea about what happened in Apr 2018 and Mar 2018, but due to small amount of data I was unable to do the best. Both the account number and deposit amount increases in one month but not in significant amount. The reason behind this is, some client closes their old accounts and open new accounts to get better offers. As a result the growth was not significant.

**Reference:**

[1] <https://www.southeastbank.com.bd/>

[2] <https://github.com/cce-bigdataintro-1160/winter2020-site>

[3] <https://plotly.com/python/bar-charts/>

[3] <https://plotly.com/python/subplots/>

[4] <https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html>