# Financial: Analysis on Customer data


## Rahul Gupta - DSC 680 - Winter 2020


https://github.com/rahulgupta271/DSC680_Project_2_Financial_Analysis_Customer_data

## Which Domain?

This project will concentrate on the patterns in financial and the correlations that influence shifts. I do have background in the banking industry, I'm fascinated in how financial companies decide to rollout the new promotions based on the customer behavior.

**Reference** for this project are mentioned below:


1. Data source link - https://www.openml.org/d/1461

2. Chris Albon, April 2018, “Python Machine Learning Cookbook”, ISBN 9781491989388.

3. Jake Huneycutt, May 18 2018, "Implementing a Random Forest Classification Model in Python",

https://medium.com/@hjhuney/implementing-a-random-forest-classification-model-in-python- 583891c99652

4. Will Koehrsen, Jan 9 2018, "Hyperparameter Tuning the Random Forest in Python",

https://towardsdatascience.com/hyperparameter-tuning-the-random-forest-in-python-using- scikit-learn-28d2aa77dd74

5. Mohammed Sunasra, Nov 11, 2017, "Performance Metrics for Classification problems in Machine Learning", https://medium.com/thalus-ai/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b

## Which Data?

The data collection consists of a Portuguese bank's direct marketing strategy outcomes.

Data set is available at the link - https://www.openml.org/d/1461

## Research Questions? Benefits? Why analyzes these data?

My key analysis concern is What drives Customer in banking?" "Drilling deep, I'm going to try to find clear explanations for high and low customers between different products and between different divisions within those banks..

I can visually review the results, see some apparent patterns and see where each bank product has flourished. Then I'm going to try to classify the associations that influence revenue. Most definitely, I'm going to try heatmaps and uncertainty matrixes and see how the features apply to each other.

My motivation in this area is to chat to a friend who works for an analytics company focused specifically on financial. analytics

## What Method?

At first, I'm going to use R and/or Python for visualization. I'm more likely to attempt to build visualizations of Power BI, when many businesses are using Power BI and searching for analysts who are fluent in it. I'm more inclined to use R to discover correlations, as I've always found it easier to use than Python.

## Potential Issues?

For me, one of the greatest problems I see would be that I could not identify any real meaningful associations, as my intuition would be that there are actually a variety of correlating characteristics. This is going to go off track if I concentrate more on one thing than the others and lose sight of time control.

## Concluding Remarks

I'm interested in getting a career in analytics, and as financial analytics appears to be a growing focus in the world of data science, I decided to discuss this field in order to gain more insight in learning what banks should do to influence customers to open account in their bank.