# Scripts Order and their Description

1. **Script1(Language R)\_Convert\_complete\_data\_to\_Fraud\_NotFraus\_Classes(0,1)**

* It is written in R (one can use R Studio to run it)
* It is about changing the data in terms of a Machine Learning problem
* Two datasets used:
* Retailer Locations\_Examples.csv (F1)
* ASU\_D11\_Retailer\_Daily\_Sales\_200206.csv (F2)
* Used **‘(Additional) Date Reference(s) for Cashings’** and **‘Retailer ID(s)’** columns of **F1** to mark the transaction as **fraud** and merged these with **‘Cashings Date’** and **‘retailer’** columnsof **F2** respectively.
* Final file created for Machine Learning Models: **‘total\_data.csv’**

1. **Script2(Python)\_Experiment 1**

* This file is in Python(Anaconda Jupyter can be used to run)
* Complete data set is used
* It consists of different Models and their evaluation

1. **Script3(Python)\_Experiment 2**

* This file is in Python(Anaconda Jupyter can be used to run)
* Data of retailers used who have committed frauds at least once
* It consists of different Models and their evaluation(Threshold Analysis)
* Transactions can be differentiated into different buckets(red, green, blue) as per the threshold value set

1. **Script4(Python)\_Experiment 3**

* This file is in Python(Anaconda Jupyter can be used to run)
* Retailers specific models created(started with retailers with max fraud count)
* It consists of different Models and their evaluation
* Transactions can be differentiated into different buckets(red, green, blue) as per the threshold value set

1. **Script5(Python)\_Address to Latitude & Longitude**

* Used Python for this part(Spyder or Jupyter can be used to run)
* It converts Retailer address to Latitude and Longitude
* After modelling, if any retailer’s transaction is marked to be in red bucket, this script can be used to trace the **nearest 5 retailers**
* Their data can be further analyzed to identify any illicit pattern