CUSTOMER SEGMENTATION

DETAILED PROJECT REPORT

AKASH

iNeuron

PROJECT DETAIL

Project Title	Customer Segmentation
Technology	Business Intelligence
Domain	E Commerce
Project Difficulty level	Advanced
Programming Language Used	Python
Tools Used	Jupyter Notebook, MS-Excel, Pycharm

OBJECTIVE

- The goal of this project is to analyse Customer Behavior, based on a combination of features and divide them into different group of clusters.

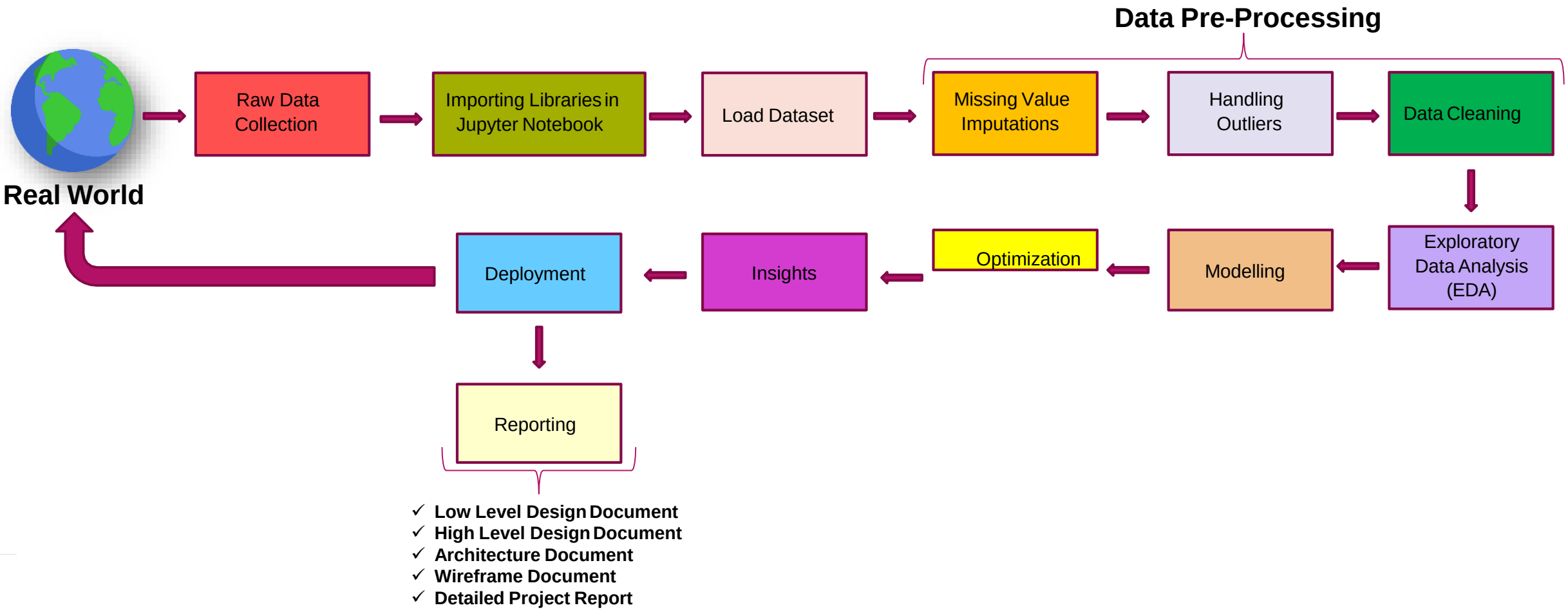
PROBLEM STATEMENT

- Not all customers are same. To know which group is your customer and their preferences is a big part for success in your business. Unsupervised machine learning can help marketers to know their audience globally and engage them with their products accordingly.

.

- Here we can classify millions of people's interests through their activity and also through other surveys online & offline and cluster them in specific group of their interest .

ARCHITECTURE



DATASET INFORMATION

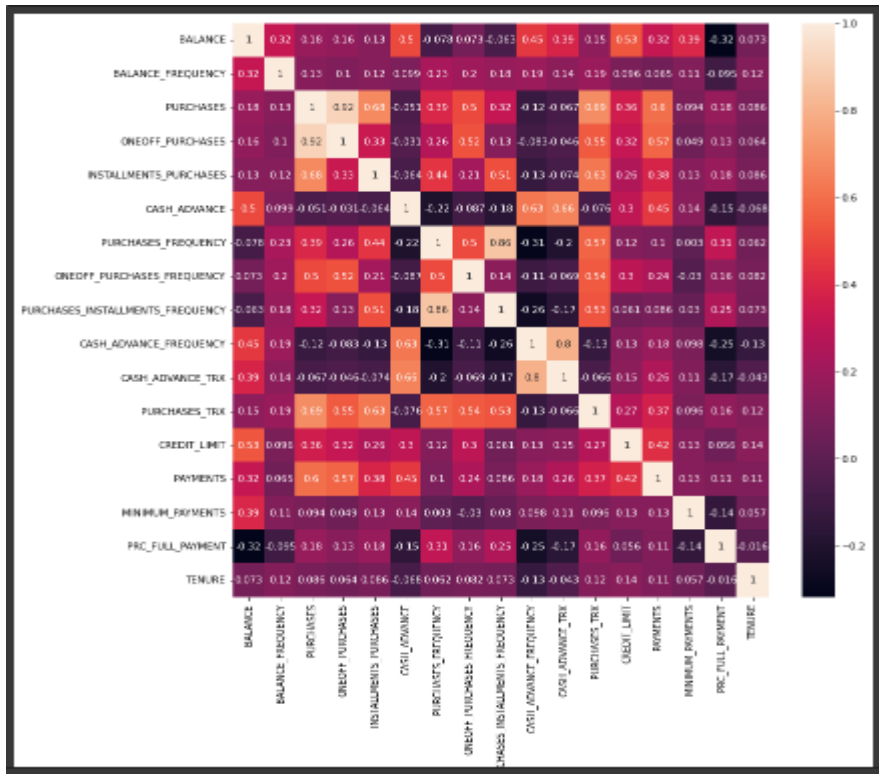
The sample Dataset summarizes the usage behavior of about 9000 active credit card holders during the last 6 months. The file is at a customer level with 18 behavioral variables.

Dataset

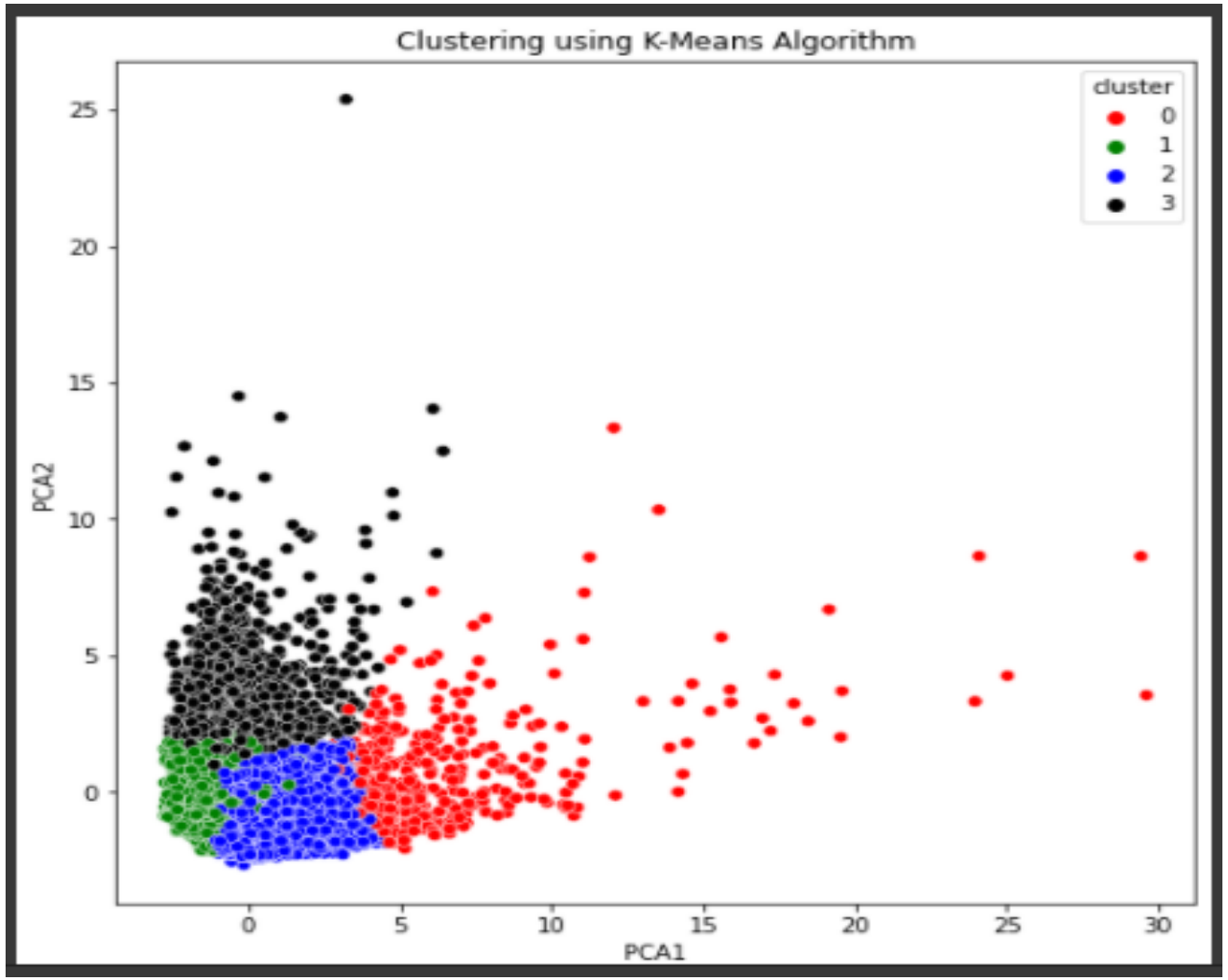
- Balance Frequency
- Purchases
- One-off Purchases
- Installment Purchases
- Cash Advance
- Purchases Frequency
- One-off Purchases Frequency
- Purchases Installments Frequency
- Cluster
- Balance
- Cash Advance Frequency
- Cash Advance TRX
- Purchases TRX
- Credit Limit
- Payments
- Minimum Payments
- PRC Full payment
- Tenure

INSIGHTS

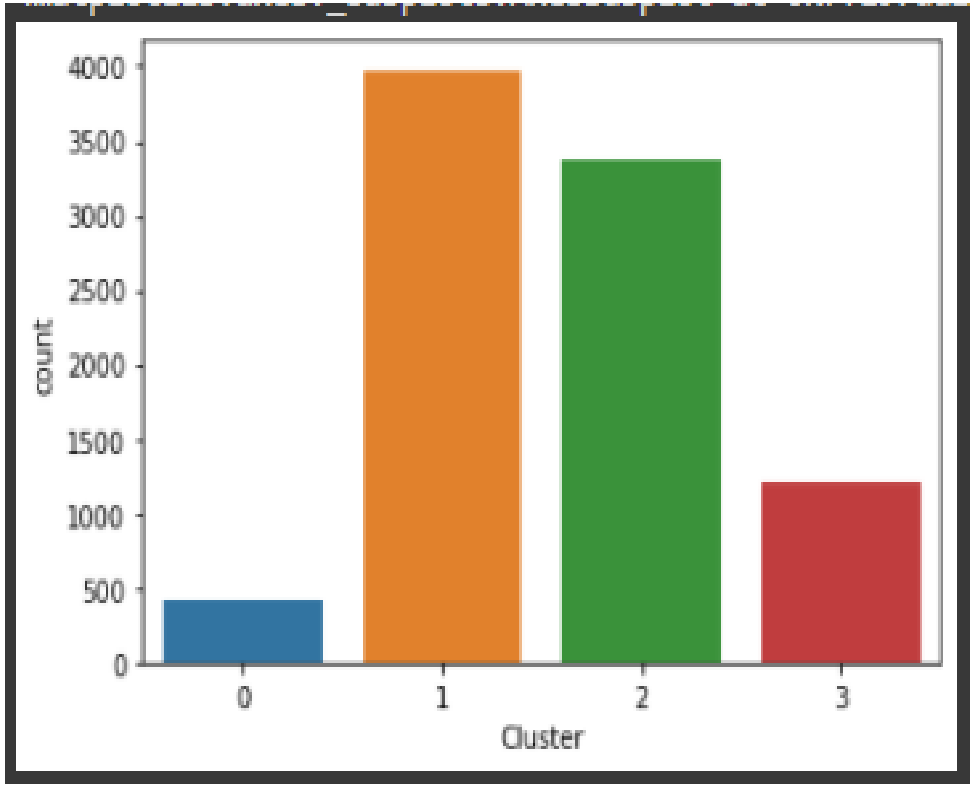
Correlation between variable



Dividing customers in different Clusters.



Distribution of customer groups



KEY PERFORMANCE INDICATOR (KPI)

1. Spending capacity
2. Amount of available cash
3. **Age Distrubution**
4. Confusion Matrics
5. Geographical location

Q & A

Q1) What was the type of data?

Ans) The data was the combination of numerical and Categorical values.

Q 3) What's the complete flow you followed in this Project?

Ans) We converted unsupervised learning problem into supervised learning problem

Q4) What techniques were you using for modeling?

Ans) –K-means to create clusters and decision tree for training.

Q 2) What were the libraries that you used in Python?

Ans) I used Pandas, NumPy and Matplotlib ,sklearn and Seaborn libraries .

-Visualizing relation of independent variables with each other and output variables

-Removing outliers

-Cleaning data and imputing if null values are present.

-Converting Numerical data into Categorical values.

THANK YOU