

# Capstone Project Submission

## Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

### Team Member's Name, Email and Contribution

Team Member	Email id
Vivek Pawar	<a href="mailto:Vivek16pawar@gmail.com">Vivek16pawar@gmail.com</a>
Varsh Rani	<a href="mailto:e.varsharani@gmail.com">e.varsharani@gmail.com</a>
Tushar Gaikwad	<a href="mailto:Gaikwadtushar140@gmail.com">Gaikwadtushar140@gmail.com</a>
Rabista Parween	<a href="mailto:2001rabista@gmail.com">2001rabista@gmail.com</a>

Contribution	Responsibilities
Coding	Vivek Pawar
Presentation	Varsh Rani
Technical Documentation	Varsh Rani
Summary	Rabista Parween
Presentation Video	Individual

### Git Hub Repository

Github Link : [https://github.com/2001rabista/Unsupervised\\_ML\\_Capstone](https://github.com/2001rabista/Unsupervised_ML_Capstone)

## Summary

### **Problem Statement:**

Extraction/identification of major topics & themes discussed in news articles. Problem Description In this project, your task is to identify major customer segments on a transnational data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail. The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

### **Steps Involved:**

- Libraries Importing
- Data Loading And Data Inspection
- Data Manipulation
- EDA
- Feature Engineering & Data Cleaning
- RFM Analysis
- Model Preparation
- Outlier Removal
- Training Model
- K-Means Algorithm
- Selecting Optimal K value
- Silhouette score method
- Elbow Method
- Cluster 0 Analysis
- Cluster1 Analysis

## **Conclusion:**

We have got 2 clusters by applying k means algorithm. So the customers got segmented into 2 clusters. Online Retail Customer marketing team can now use different approaches to acquire the customers

### **Cluster 0**

#### Key Figures

- Frequency : 28.68
- Recency : 230
- Monetary : 3070
- RFM Score : 10.71

#### Top 5 Products

- WHITE HANGING HEART T-LIGHT HOLDER : 339
- REGENCY CAKESTAND 3 TIER : 268
- ASSORTED COLOUR BIRD ORNAMENT : 235
- PARTY BUNTING : 229
- REX CASH+CARRY JUMBO SHOPPER : 202

### **Cluster 1**

#### Key Figures

- Frequency : 37.67
- Recency : 134.64
- Monetary : 447.40
- RFM Score : 5.90

#### Top 5 Products

- WHITE HANGING HEART T-LIGHT HOLDER : 344
- REGENCY CAKESTAND 3 TIER : 271
- ASSORTED COLOUR BIRD ORNAMENT : 239
- PARTY BUNTING : 232
- REX CASH+CARRY JUMBO SHOPPER : 204