***Capstone Project – Online Retail Customer Segmentation***

**Table of Contents**

1. Problem Statement

2. Project Objective

3. Data Description

4. Data Pre-processing Steps and Inspiration

5. Choosing the Algorithm for the Project

6. Motivation and Reasons for Choosing the Algorithm

7. Assumptions

8. Model Evaluation and Techniques

9. Inferences from the Same

10. Future Possibilities of the Project

11. Conclusion

12. References

**Problem Statement**

An online retail store is trying to understand the various customer purchase patterns for their firm, you are required to give enough evidence based insights to provide the same.

**Project Objective**

The objective of this project is to find: -

1. Using the above data, find useful insights about the customer purchasing history that can be an added advantage for the online retailer.

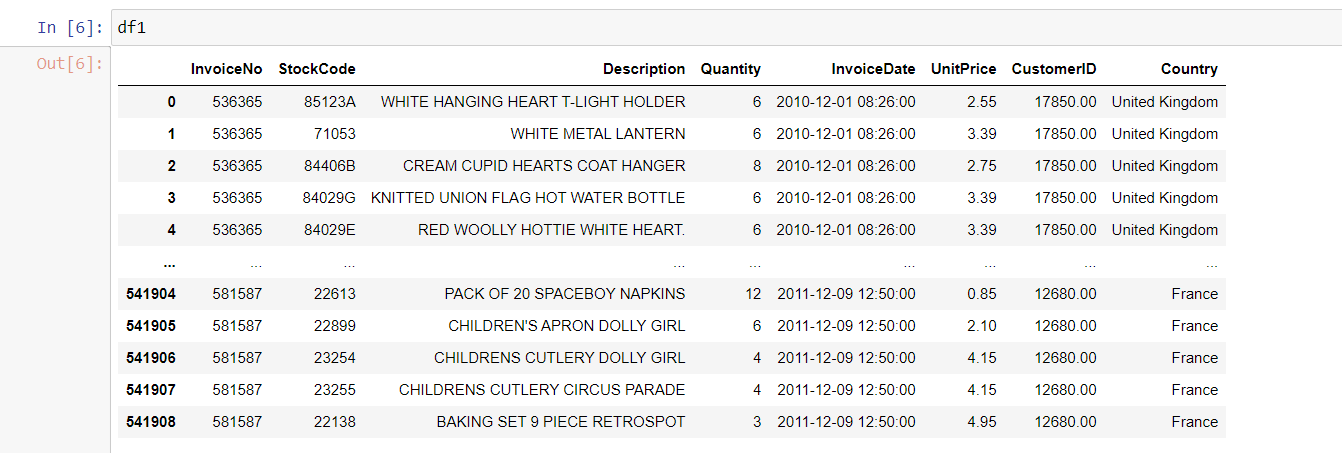
2. Segment the customers based on their purchasing behaviour

**Data Description**

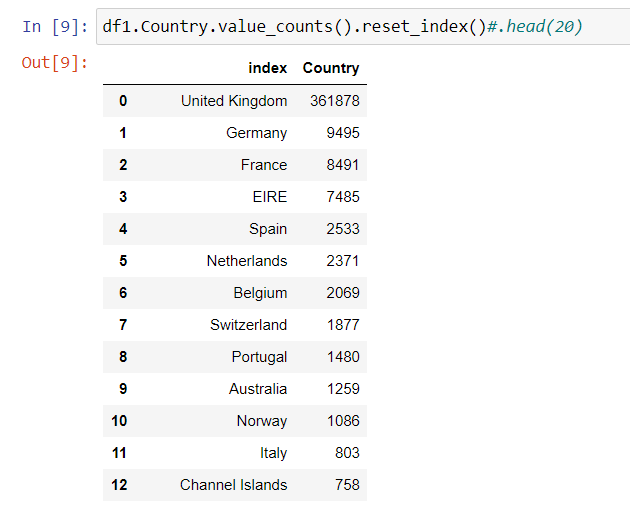
The dataset available is of online retail list which consists of details of customer purchase

**Insights from the data:**

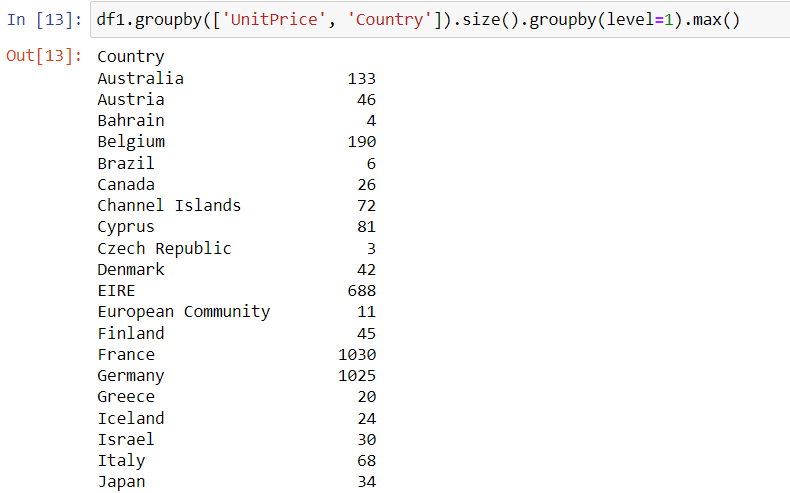
1. Data contents

****

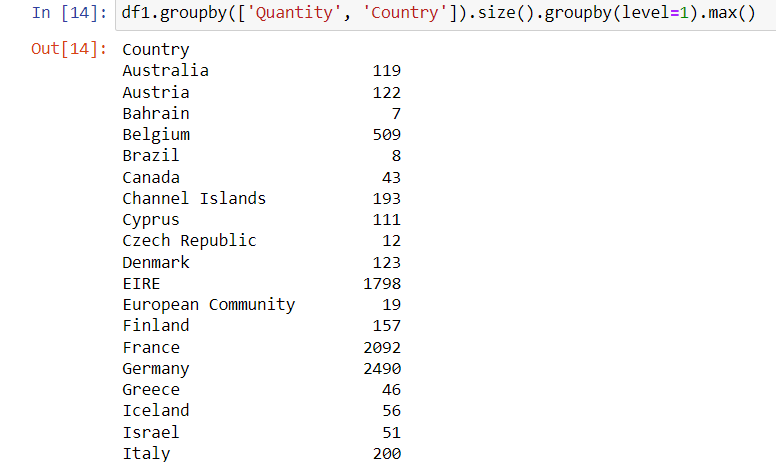
1. Highest number of products brought by a country



1. Unit prices spend my customers from each country



1. Total Quantity bought by customers from each country



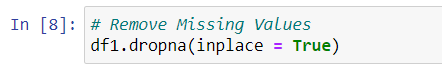
**Data Preprocessing Steps And Inspiration**

The preprocessing of the data included the following steps:

1. First step: - Convert Invoice Date to datetime format



1. Second step: - Drop all null values



**Choosing the RFM segmentation For the Project.**

I have chosen the RFM segmentation for this project for the following reasons:

1. RFM analysis allows marketers to target specific clusters of customers with communications that are much more relevant for their particular behavior
2. It utilizes objective, numerical scales that yield a concise and informative high-level depiction of customers.
3. It is simple – marketers can use it effectively without the need for data scientists or sophisticated software.
4. It is intuitive – the output of this segmentation method is easy to understand and interpret.

**Assumptions**

The following assumptions were made in order to create RFM segmentation.

1. Customers who purchased recently are more likely to purchase again than the customers who purchased a long time ago.
2. Customers who have purchased more in the past are more likely to respond than the customers who have made fewer purchases.
3. Customers who have spent more (in total for all purchases) in the past are more likely to respond than the customer who have spent less.

**Model Evaluation and Technique**

The following techniques and steps were involved in the evaluation of the model.

1. Find out the Recency, Frequency and Monetary values of the given data.
2. Convert them to RFM score values
3. Make Segmentations based on RFM scores.
4. Plot the Data

**Inferences from the Project**

The idea is to segment customers based on when their last purchase was, how often they’ve purchased in the past, and how much they’ve spent overall. All three of these measures have proven to be effective predictors of a customer's willingness to engage in marketing messages and offers.

**Limitations**

1. **RFM analysis of the customer database does not take into account the context of purchases (such as seasonality) and customer characteristics (such as demographics).** Such data is also important for offering products.
2. **RFM segmentation will not bring results when there is a small amount of initial data.** The database for analysis should be at least ten thousand people.
3. **It is not suitable for companies that sell large and expensive products.** For example, real estate is usually purchased once or twice in a lifetime, so customers are unlikely to return for a new apartment after a year.
4. **Data has to be updated and the analysis repeated, as customers can move from one segment to another.** The standard period is once every six months, but if there are rapid changes in sales, it is worth doing it once every two months.

**Conclusion**

1. RFM analysis can help brands identify their most valuable customers and develop targeted marketing strategies. However, analysing customer data can be challenging, particularly for organizations that don’t have a strong data foundation.
2. The RFM model enables you to get a holistic view of your new and existing customers. Hence, creating strategies to maximize the lifetime value of your customers.

**References**

1. <https://altcraft.com/blog/rfm-analysis-what-it-is-and-how-to-manage-it>
2. <https://github.com/Gaelim/Mall-Customer-Segmentation/blob/main/Mall%20Clustering%20Project.ipynb>
3. https://www.kaggle.com/code/ekrembayar/rfm-analysis-online-retail-ii