Project Title:
NEXT BEST ACTION

NAME	Roll No	Sap – ID
Adarsh Kandwal	R103218007	500068043
Ritik Kathuria	R103218117	500067286
Ritik Ranjan Tripathi	R103218118	500067646
Sarthak Saraiya	R103218129	500067734
Saubhagya Marwaha	R103218131	500068772

Under the guidance of

Dr. Hitesh Kumar Sharma

INDEX

Contents **Page no.**

1. Introduction 2
2. Objective 3
3. Target Audience 3
4. Technology Used(Language & Software) 3
5. ML problem mapping 4
6. Dataset description 4
7. Project Demonstration(Snapshots) 5 – 7
8. Future Score 8
9. Conclusion 8

INTRODUCTION

These days, people don't like to be overwhelmed by information. They don't notice or report online ads and posts, unsubscribe from newsletters, reject calls, and mute app notifications. Or simply forget to read bookmarked posts because…there is just too much content on the internet.

As described by marketing consultant Mark Schaefer, content shock develops when humans can no longer consume an ever-growing amount of content.

And he's not the only one talking about it. Other specialists are mentioning it too. Suneel Grover from SAS notes that it's "*important to objectively realize that your brand or product isn't the center of your consumer's world*." Most marketing-centric content is perceived as an interruption to a customer-oriented experience.

This changes the game for marketers. Businesses now must engage with customers on their terms: through preferred channels and at a preferred time, with offers that speak to their hearts. The shifting attention from a product to a consumer is the central idea of ***the next best action strategy.***

The **next best action (NBA)** marketing strategy aims at finding the optimal action a company must take during a customer interaction that will unobtrusively and smoothly lead a certain sale prospect to purchase. For instance, a user starts with the section showcasing sneakers in a mobile app, then reads reviews, bookmarks a few models, adds two pairs to a cart, and abandons it. The best next action may be to send a notification with a promo code for forgotten items. Another task is to define when to send the notification and how to not make people feel like they're being chased.

OBJECTIVE

- Given a dataset of transactions (Online Retail dataset from UCI Machine Learning repository) get the segments i.e clusters/segments. (Find common patterns and group them)
- Understand which marketing activities are most likely to move each individua l customer closer to purchase.
- Predict what to display to what group of users
- Using Machine Learning algorithms to find out the Next Best Action for a customer.
- Categorizing the customers in a particular segment based on their buying patterns.
- Predicting which kind of items they will buy in future based on their segmentation

TARGET AUDIENCE

We wish to target individuals who are keen on new innovation, assisting them with deciding. It tends to be a merchant for subcontracting a task or any individual who wishes to facilitate their work with the assistance of innovation.

TECHNOLOGY USED

In the given project, we have used the basics of machine learning predicting Customer behavior & Buying pattern, and the language we have entirely used is python.

The software we used while working on the project is Jupyter Notebook.

ML PROBLEM MAPPING

1. Given a dataset of transanctions (Online Retail dataset from UCI Machine Learning repository) get the segments i.e clusters/segments. (Find common patterns and group them)
2. Predict what to display to what group of users

Input: We will be using e-commerce data that contains the list of purchases in 1 year for 4,000 customers.

Output: The first goal is that we need to categorize our consumer base into appropriate customer segments. The second goal is we need to predict the purchases for the current year and the next year based on the customers' first purchase.

DATASET DESCRIPTION

This is 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.

- InvoiceNo: Invoice number. Nominal, a 6-digit integral number uniquely assigned to each transaction. If this code starts with letter 'c', it indicates a cancellation.
- StockCode: Product (item) code. Nominal, a 5-digit integral number uniquely assigned to each distinct product.
- Description: Product (item) name. Nominal.
- Quantity: The quantities of each product (item) per transaction. Numeric.
- InvoiceDate: Invice Date and time. Numeric, the day and time when each transaction was generated.
- UnitPrice: Unit price. Numeric, Product price per unit in sterling.
- CustomerID: Customer number. Nominal, a 5-digit integral number uniquely assigned to each customer.
- Country: Country name. Nominal, the name of the country where each customer resides.

PROJECT DEMONSTRATION (SNAPSHOTS)

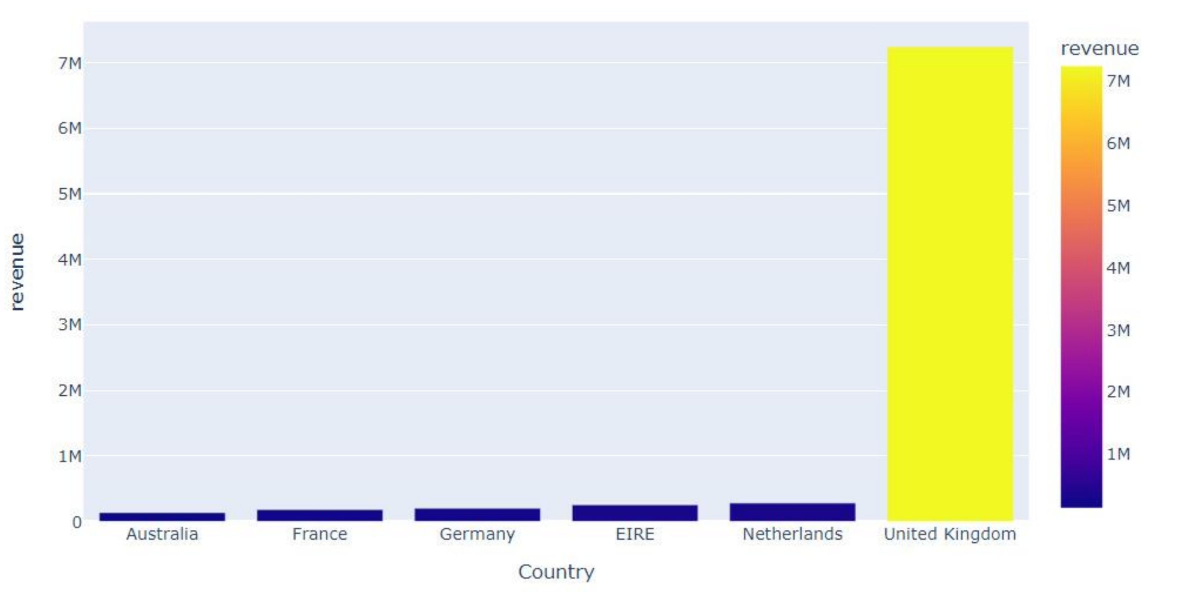
Now we sum the individual orders and group them on the basis of invoice number to remove the problem of duplicate rows for same order :

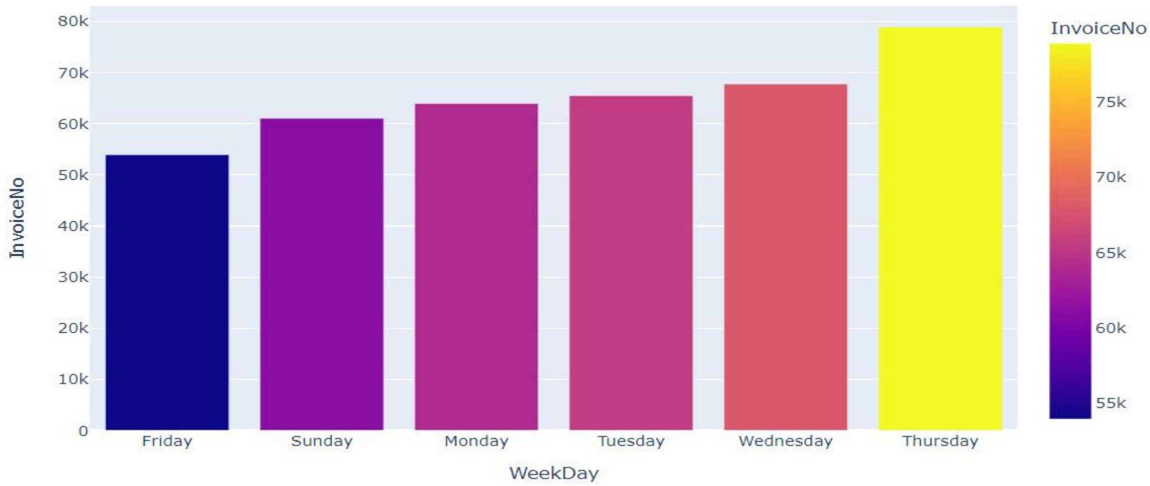
ı [28]:

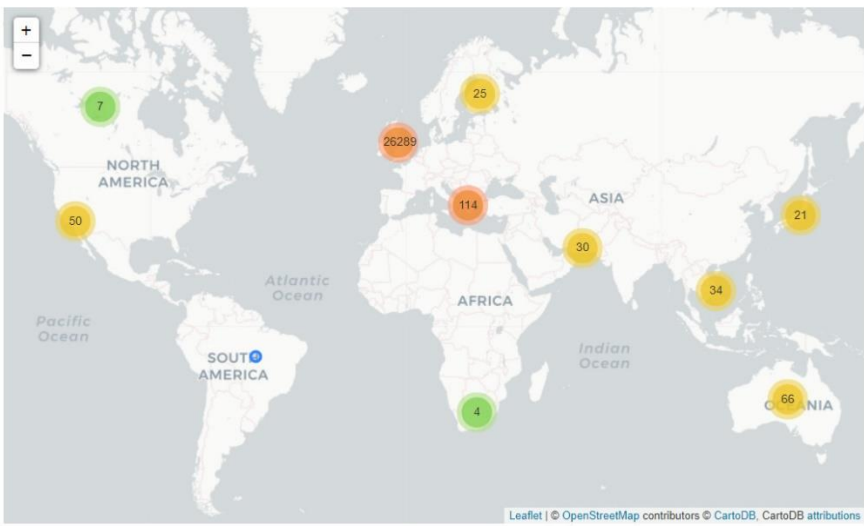
```
temp = df_cleaned.groupby(by=['CustomerID', 'InvoiceNo'], as_index=False)['TotalPrice'].sum()
basket_price = temp.rename(columns = {'TotalPrice': 'Basket Price'})

df_cleaned['InvoiceDate_int'] = df_cleaned['InvoiceDate'].astype('int64')
temp = df_cleaned.groupby(by=['CustomerID', 'InvoiceNo'], as_index=False)['InvoiceDate_int'].mean()
df_cleaned.drop('InvoiceDate_int', axis = 1, inplace=True)
basket_price.loc[:, 'InvoiceDate'] = pd.to_datetime(temp['InvoiceDate_int'])

basket_price = basket_price[basket_price['Basket Price'] > 0]
basket_price.sort_values('CustomerID')[:6]
```

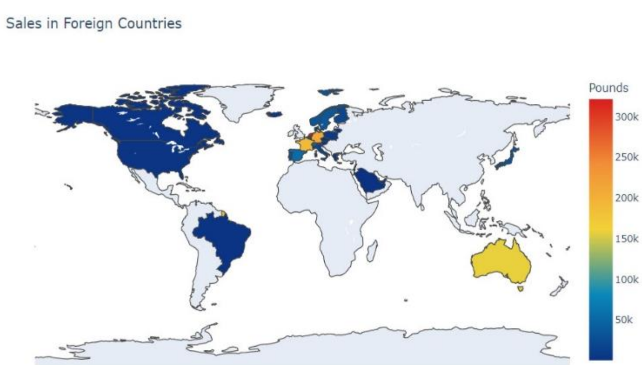
ıt[28]:

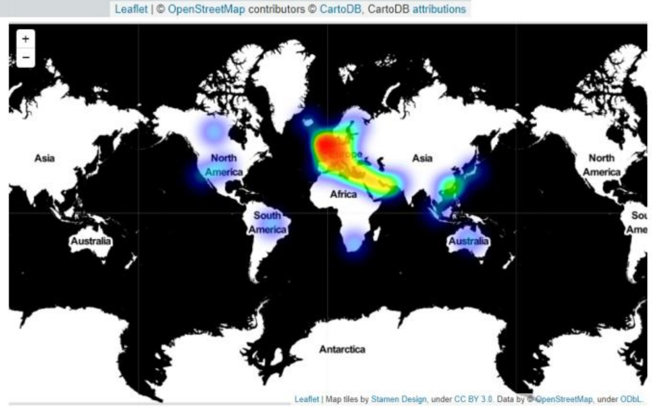
	CustomerID	InvoiceNo	Basket Price	InvoiceDate
1	12347.0	537626	711.79	2010-12-07 14:57:00.000001024
2	12347.0	542237	475.39	2011-01-26 14:29:59.999999744
3	12347.0	549222	636.25	2011-04-07 10:42:59.999999232
4	12347.0	556201	382.52	2011-06-09 13:01:00.000000256
5	12347.0	562032	584.91	2011-08-02 08:48:00.000000000
6	12347.0	573511	1294.32	2011-10-31 12:25:00.000001280



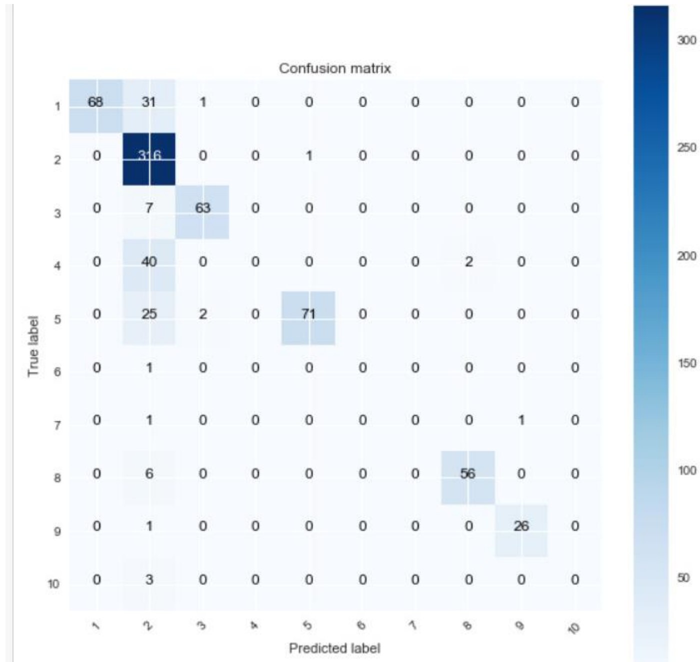

InvoiceNo
80k
70k
60k
50k
40k
30k
20k
10k
0
Friday
Sunday
Monday
Tuesday
Wednesday
Thursday
WeekDay
InvoiceNo
75k
70k
65k
60k
55k

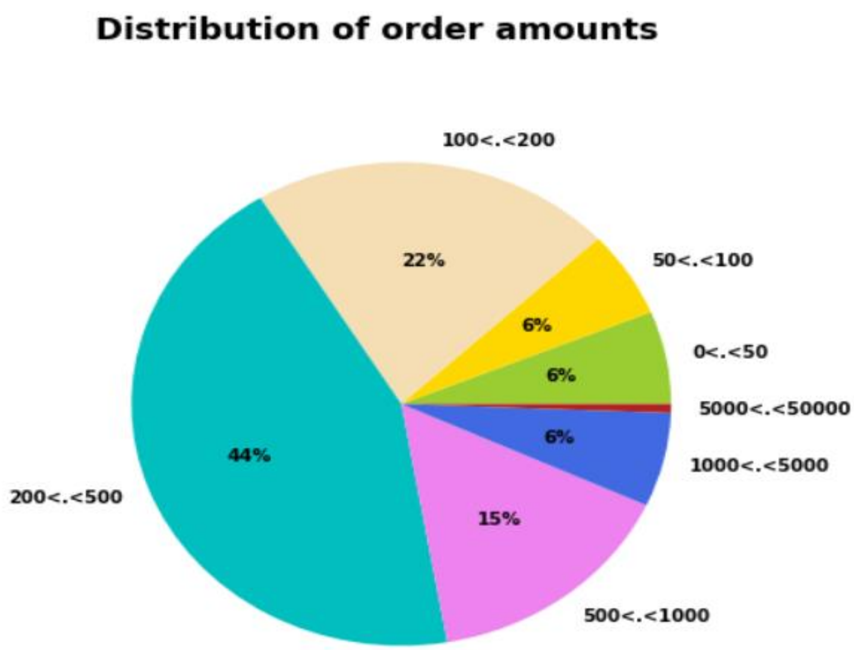

7
25
26289
114
50
21
30
34
66
4
NORTH AMERICA
ASIA
Atlantic Ocean
AFRICA
Pacific Ocean
Indian Ocean
AMERICA
Leaflet | © OpenStreetMap contributors © CartoDB, CartoDB attributions


Sales in Foreign Countries
Pounds
300k
250k
200k
150k
100k
50k


Asia
North America
Asia
North America
Africa
South America
Australia
Australia
Antarctica
Leaflet | Map tiles by Stamen Design, under CC BY 3.0. Data by © OpenStreetMap, under ODbL.

PROJECT DEMONSTRATION (SNAPSHOTS)






```
Precision : 89.22%
```

> Accuracy on test dataset is good considering we use 10 months old data to predict on new data.

FUTURE SCOPE

- We wish to smooth out the results with more accuracy than we have already.
- We hope to work with many other algorithms and try-test the accuracy of each algorithm in terms of time complexity and accuracy of results.

CONCLUSION

- While working on the project, we came across many difficulties. However, we tried our level best to tackle them in the best way possible.
- We are able to categorize the customers in a particular segment based on their buying patterns.
- The precision of the model is about 89.2%.