Wolt Customer Analysis

Sijan Shrestha

Customer Segmentation

- Customer segmentation is a crucial for a company to make useful strategies like marketing strategies, advertising strategies, target customers , target market etc.
- Customer segmentation can be done on following basis:

  -Demography (age, sex, gender, marital status etc)

  -Geography (location)

  -behaviour (how customer purchase)

  -dormant customers (registered but not engaged)

- In this project , I have done segmentation on the basis of Geography, behaviour and dormant customer

Steps followed:

- Data extraction
- Data cleaning
- Data exploration
- Data segmentation (Geography, behaviour and dormant)
- Analysis on the basis of outcomes.

Data extraction

We can use pandas function read_csv() for data extraction. It can be done as follows:

df=pd.read_csv('dataset_for_analyst_assignment_20201120.csv'). There are 21983 rows and 30 columns in this dataset.

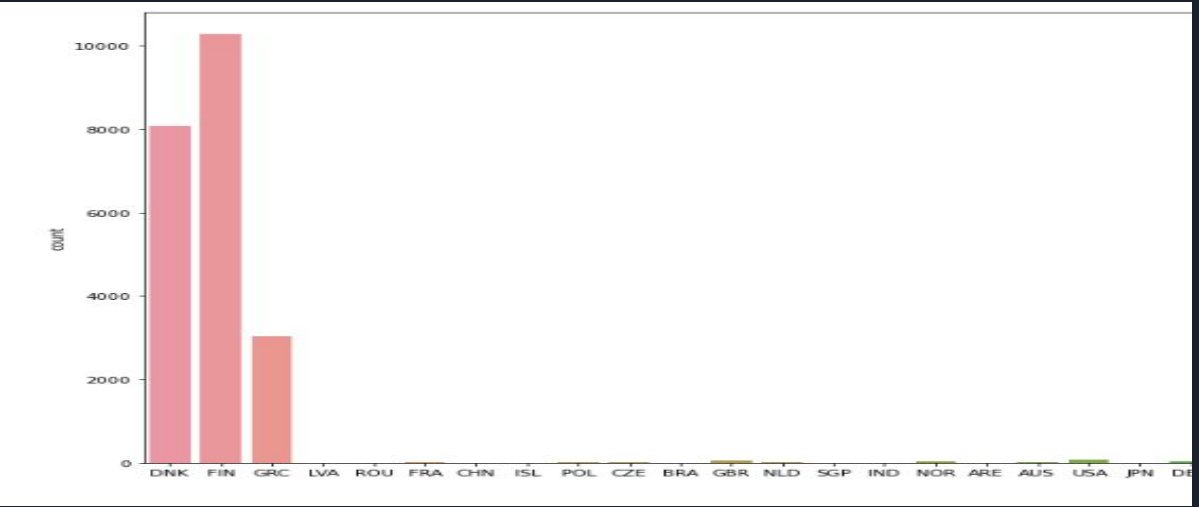
	REGISTRATION_DATE	REGISTRATION_COUNTRY	PURCHASE_COUNT	PURCHASE_COUNT_DELIVERY	PURCHASE_COUNT_TAKEAWAY	FIRST_PURCHASE_DAY	LAST_PU
0	2019-09-01 00:00:00.000	DNK	0	NaN	NaN	NaN	
1	2019-09-01 00:00:00.000	FIN	1	1.0	0.0	2020-09-02 00:00:00.000	2020-09-0
2	2019-09-01 00:00:00.000	DNK	19	19.0	0.0	2019-12-10 00:00:00.000	2020-05-2
3	2019-09-01 00:00:00.000	FIN	0	NaN	NaN	NaN	
4	2019-09-01 00:00:00.000	GRC	0	NaN	NaN	NaN	
...	...	...	...	...	...	...	
21978	2019-09-30 00:00:00.000	GRC	1	1.0	0.0	2020-01-23 00:00:00.000	2020-01-2

Data Cleaning

- Data cleaning and preparing data for further analysis is crucial part of data analysis.
- The raw dataset can have unwanted, duplicate , incorrect or incomplete data which should be corrected.
- In this project, null values has be removed , duplicate data has been checked and unwanted column has been removed .
- dropna() function helped to eliminate the null value. df.duplicated() is used to check for duplicate values.

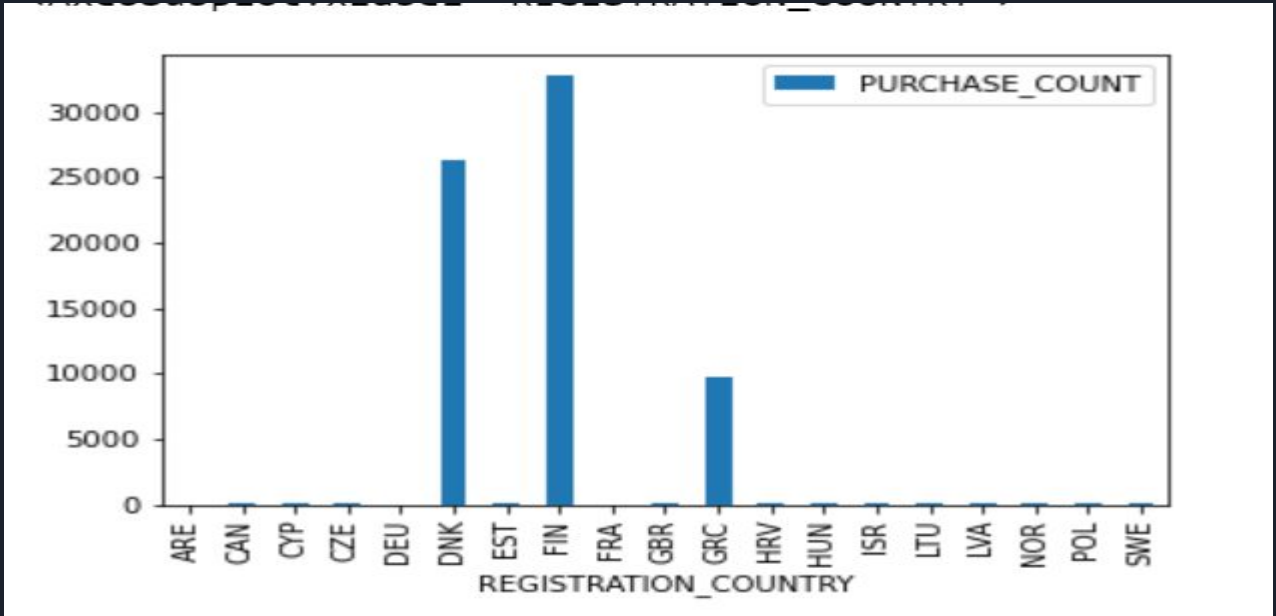
Data exploration

- While examining the dataset, there are 21983 rows and 30 columns.
- df['REGISTRATION_COUNTRY'].nunique(): provides information that there are 59 different countries in the dataset.
- Finland, Denmark and Greece are the most prominent market so far.



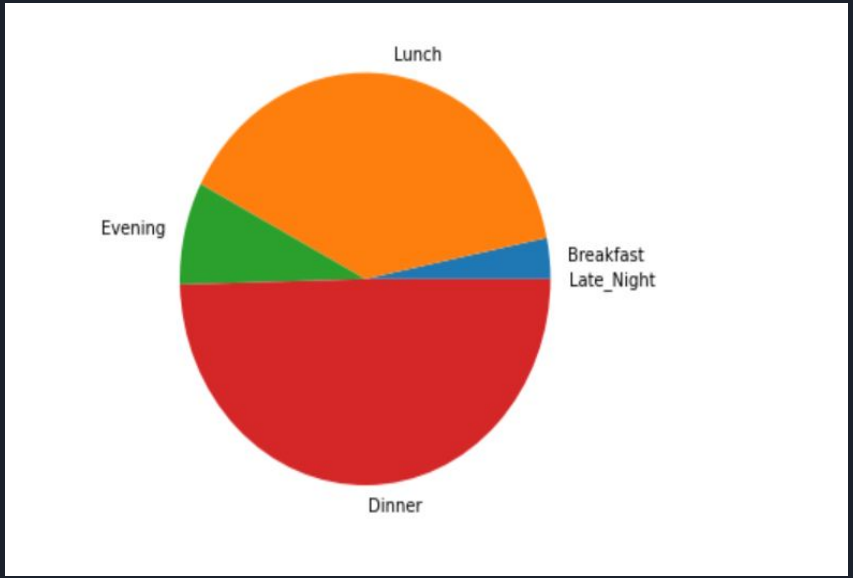
Purchase count

Among the registered countries, Finland and Denmark has highest number of purchases.



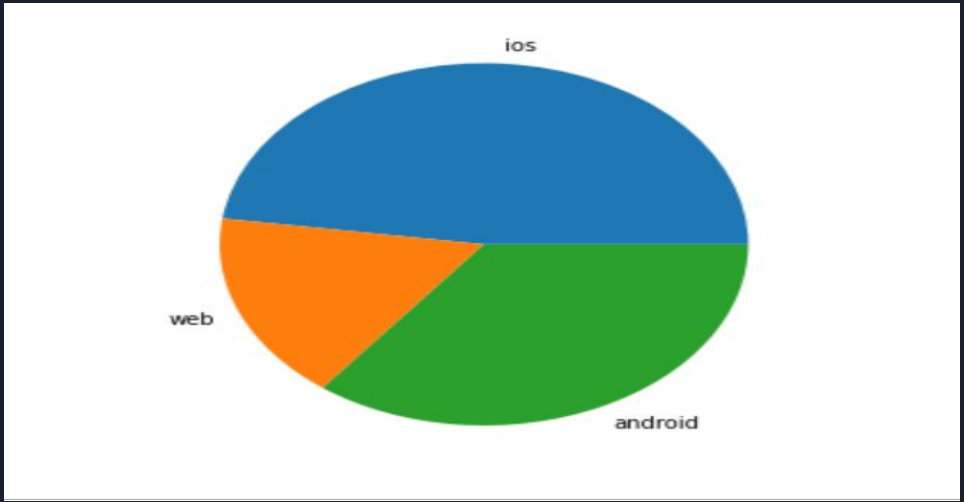
Distribution of purchases.

- Analysing purchasing behavior is important to make business strategies.
- In this dataset, there are 4 different time that customers is purchasing.
- Among them, dinner is the winner and lunch comes in the second place.



Preferred device used to purchase:

- In this dataset, the customers of Wolt has used one from ios, android or web purchases.
- From the data, we can segment the customers on the basis of this behaviour.



In Finland , ios is used most.

```
df_Fin_device=df_Fin['PREFERRED_DEVICE'].value_counts()
df_Fin_device
```

```
android    1662
ios        1346
web         566
Name: PREFERRED_DEVICE, dtype: int64
```

RFM

RFM(Receny, Frequency and Monetary value) is best used to understand the customer's behaviour)

- How recent customer has purchased, how frequently and how much money they have spent are key to understand the customer's value.
- There are still huge Dormant customers who has registered but are not active.
- The lowest recency, highest frequency and monetary amounts are our best customers.
- According to the RFM analysis, there are only three countries that shows lowest recent: Finland, Denmark and Greece.

```
df_recent_buyer['country'].value_counts()
```

```
FIN    127
DNK    113
GRC     23
Name: country, dtype: int64
```

RFM...

- Following outcome shows some of the customers with low recency and and high frequency and significant amount of monetary value.

USER_ID	country	recency	frequency	monetary_value	cost_per_order
3361	FIN	842	125	2459.160	19.673280
18316	DNK	842	111	3487.352	31.417586
1049	FIN	842	102	2143.416	21.013882
15528	FIN	842	96	2166.692	22.569708
5504	DNK	842	85	1875.236	22.061600

Dormant customers

- The analysis show that there are huge number of registered users who are idle.
- About 9955 customers has not purchase a single order.
- Wolt has to market their products so that they can reach these customers.

```
df_dormant['REGISTRATION_COUNTRY'].value_counts()
```

```
FIN    4826
DNK    3096
GRC    1510
USA      70
GBR      52
SWE      41
DEU      38
EST      31
ITA      24
NOR      21
```

Summary

- Customer segmentation is a crucial part in business analysis.
- It helps the company to build different business strategies.
- In this analysis, I have done basic exploratory analysis and done customer segmentation on the basis of geography and customer behavior.
- I have used RFM (Recency, Frequency and Monetary Value) table to analyse the customer.

Thank you !