Topic

Understanding and analysis of the behavior of the customer while purchasing a product on any e-commerce website

Done by – Team Vijay

Vijay Gupta (Team Lead)

Anubhav Anand

Uppari Ganesh

Deepika Yadav

Abhijeet Pradhan

Arindam Mondal

Chapter 1- Introduction

1.1 - Abstract: -

The new technology has radically changed the tradition way of doing online Shopping business. Online Shopping has become new type of retail Shopping. Online Shopping strengthening their sale base an financial resources by developing their own e-product and service to suit changing needs of the customer. The main objective of this study is to analyze buying behavior of customer towards online Shopping and to examine the problems of Consumer while doing with Online Shopping. The advancement in the technology paved new ways of delivering shopping goods facilities to the customer, such as Online Shopping. Online Shopping has become a popular shopping method over since the internet has declared to take over.

1.2 - Problem statement: -

This project aims to gain consumer insight of a multi-category online store through Exploratory Data Analysis (EDA). Various methods are utilised in order to obtain various angles and critically validate the analysis.

1.3 - Market / Customer/ Business Need Assessment-:

Online shopping is more and more driven by the ICT infrastructure development, online payment systems and the Internet penetration rate in World. Earlier studies showed that unlike brick and mortar shopping behavior, online shopping behavior is influenced by net connectivity, website esthetics (Constantinides, 2004), security, customers' experience, age and learning curve, etc. Studying these unique characteristics of online shopping and consumer behavior of online shoppers would benefit the tech-entrepreneurs and policymakers to craft their strategies properly for the market. This study empirically reveals the consumer behavior of online shoppers in India.

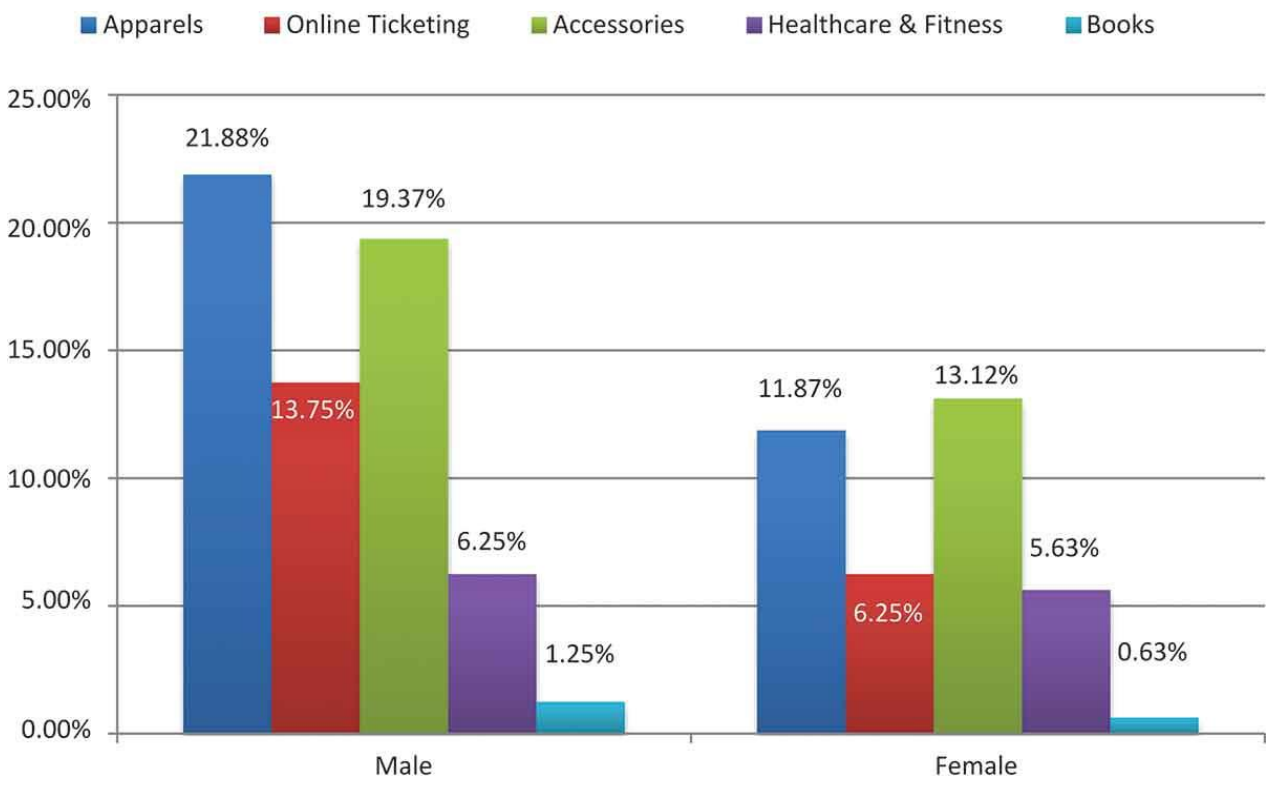
online shoppers are young (mostly below 40 years) similar to other parts of the world. They do online shopping because it saves time, offers home delivery, provides ease in shopping and offers more variety of products for apparels, accessories, and ticketing than that of brick and mortar stores. They mostly rely on price and their experience as the basis of the quality judgment of items in online shopping and for payment system they prefer cash on delivery option. Most of the shoppers get the information primarily from Facebook advertisements which is pursued by friends and family by following their "word of mouth" communication. However, privacy and inability to touch and feel are the most disliking factors for online shoppers. These findings of our study have both theoretical and practical implications.

1.4 -Target Specifications and Characterization:-

Both male and female respondents assured that there are specific reasons for choosing online shopping. 38.75% respondents (26.25% male, 12.5% female) mentioned saving time is their primary reason for choosing online shopping and about 29.38% of the interviewees prefer online shopping because of availability of the varieties of products. Nearly, 19% of the respondents prefer online shopping because product comparison is much easier for online shopping and 13.13% choose for a comfortable reason

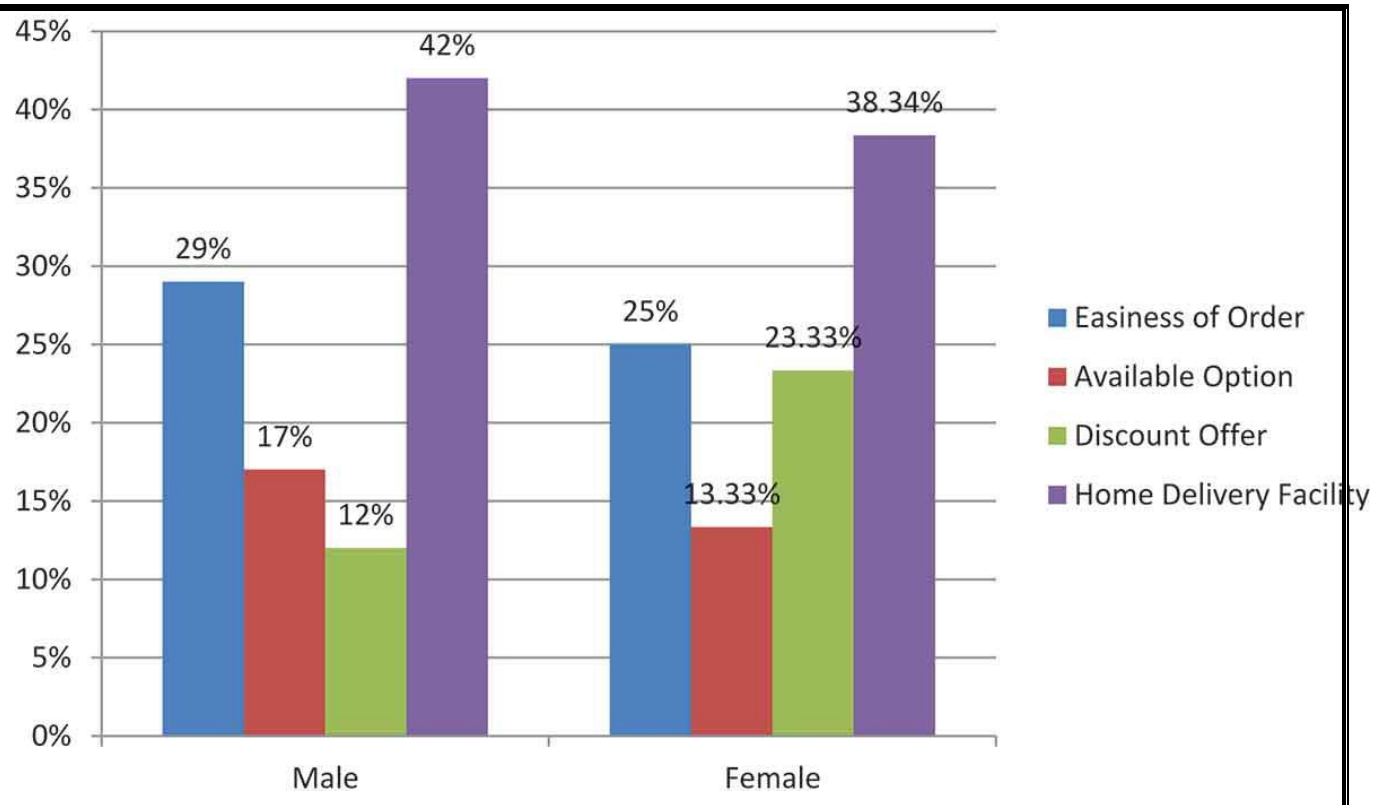
1.4.1 - Preference for product/service

While respondents were asked to know about the offered goods and services, 33.75% of the respondents preferred Apparels (21.88% male and 11.87% female) and accessories 32.49% (19.37% male and 13.12% female). 20% (16% male and 7% female) respondent preferred online ticketing. About 11.88% of the interviewees preferred healthcare and fitness products and only 1.88% respondents prefer books (see Figure)



1.4.2 - Factors for liking online shopping :-

Among the male respondents, 42% like online shopping because of the home delivery facility, while about 29% male respondents said that it is easier to order for which they would like to shop online. 12% of the male respondents like discount offer most for online shopping and 17% of the male respondents like available options for buying and payments gateways. On the other hand, 38.34% of the female respondents also do online shopping for a home delivery facility, 25% do for the easiness of ordering, rest of the 23.33% (13.33% female respondents) like for discount offer and available options respectively.


45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
29%
17%
12%
42%
25%
13.33%
23.33%
38.34%
Male
Female
Easiness of Order
Available Option
Discount Offer
Home Delivery Facility

Chapter 2 – Steps

Step 1: - Prototype Selection

Feasibility:

Performing a customer behavior analysis is becoming increasingly important for both B2C and B2B companies. This is because of a growing need for highly-personalized content that's unique to each customer. This project can be developed and deployed within a few years as SaaS(Software as a Service) for anyone to use.

However, you can't personalize content until you have a complete understanding of your customer's preferences and tendencies

Viability:

According to **Accenture**, 41% of customers switched companies last year due to poor **personalization**. As the retail industry grows in India and the world, there will always be small businesses existing which can use this service to improvise on their sales and data warehousing techniques. So, it is viable to survive in the long-term future as well but improvements are necessary as new technologies emerge.

Monetization:

We understand our customers' behaviors, it's time to capitalize on it. Use our analysis to determine what customers will think about these changes. Customers are habitual creatures, and some will push back on change even if it's for the better. These customers tend to be more loyal to your brand, so it's imperative you don't lose them as a result. Hence, this service is directly monetizable as it can be directly released as a service on completion which can be used by businesses.

Some services are also available which analysis the behavior and gave analysis report to companies:-

--Trifacta

--Vertica

--HubSpot

Step 2 : - Prototype Development

Githublink-
:https://github.com/vijay2020pc/FeynnProjects/tree/main/task%204

The behavior of the customer while purchasing a product on an e-commerce website

Understanding what are the key drivers of customers' buying behavior is always the holy grail for the eCommerce industry. The knowledge can be used to improve the shopping process and eventually result in higher sales and customer satisfaction.

Data Understanding

	event_time	event_type	product_id	category_id	category_code	brand	price	user_id	user_session
0	2019-11-01 00:00:00 UTC	view	1003461	2053013555631882655	electronics.smartphone	xiaomi	489.07	520088904	4d3b30da-a5e4-49df-b1a8-ba5943f1dd33
1	2019-11-01 00:00:00 UTC	view	5000088	2053013566100866035	appliances.sewing_machine	janome	293.65	530496790	8e5f4f83-366c-4f70-860e-ca7417414283
2	2019-11-01 00:00:01 UTC	view	17302664	2053013553853497655	NaN	creed	28.31	561587266	755422e7-9040-477b-9bd2-6a6e8fd97387
3	2019-11-01 00:00:01 UTC	view	3601530	2053013563810775923	appliances.kitchen.washer	lg	712.87	518085591	3bfb58cd-7892-48cc-8020-2f17e6de6e7f
4	2019-11-01 00:00:01 UTC	view	1004775	2053013555631882655	electronics.smartphone	xiaomi	183.27	558856683	313628f1-68b8-460d-84f6-cec7a8796ef2

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 67501979 entries, 0 to 67501978
Data columns (total 9 columns):
 #   Column         Dtype
---  ------         -----
 0   event_time     object
 1   event_type     object
 2   product_id     int64
 3   category_id    int64
 4   category_code  object
 5   brand          object
 6   price          float64
 7   user_id        int64
 8   user_session   object
dtypes: float64(1), int64(3), object(5)
memory usage: 4.5+ GB
```

Know your customers

How many customers visit the site?

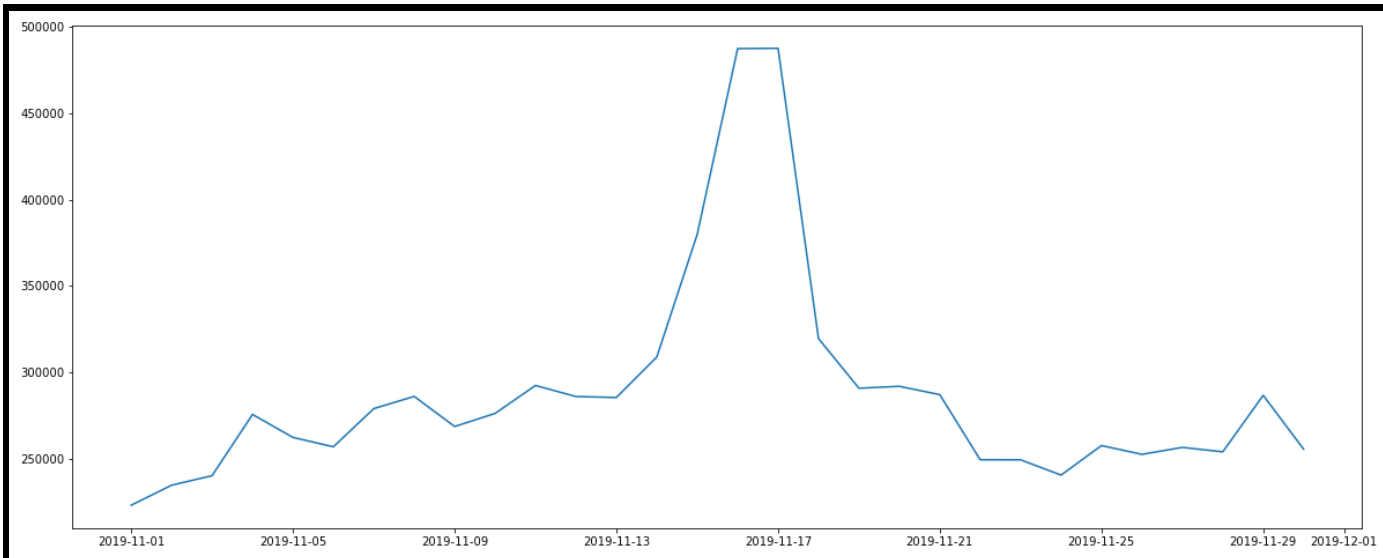
Know your Customers

```
visitor = data['user_id'].nunique()
print ("visitors: {}".format(visitor))
```

```
visitors: 3696117
```

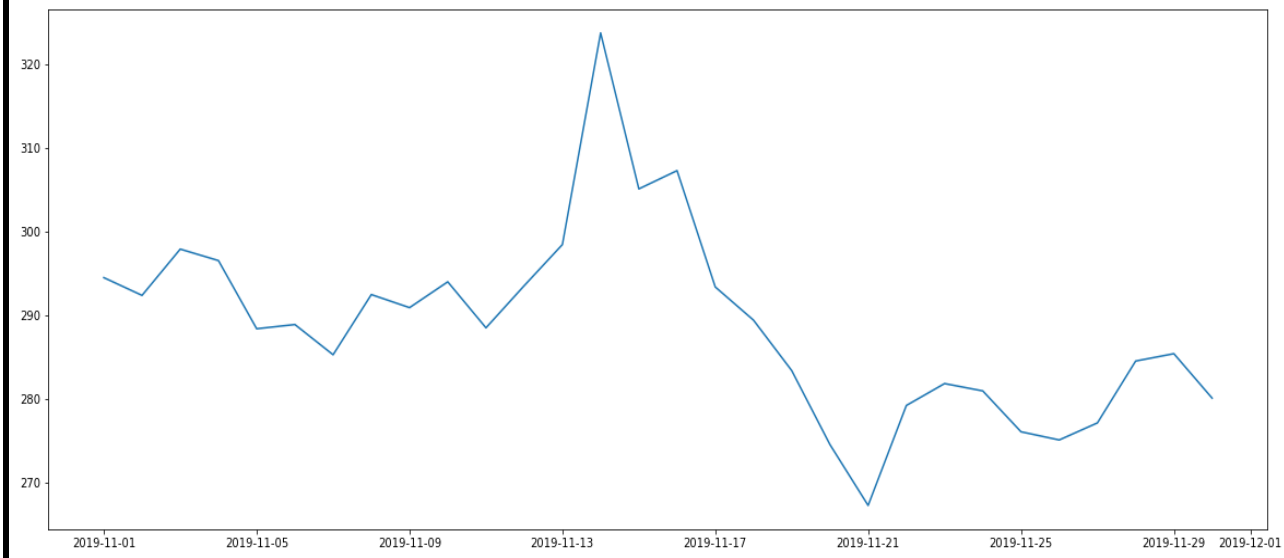
Visitors Daily Trend

```
x = pd.Series(visitor_by_date.index.values)
y = visitor_by_date['count']
plt.rcParams['figure.figsize'] = (20,8)

plt.plot(x,y)
plt.show()
```



Price Daily Trend

```
x = pd.Series(product_daily_price.index.values)
y = product_daily_price['price']
plt.rcParams['figure.figsize'] = (20,8)

plt.plot(x,y)
plt.show()
```

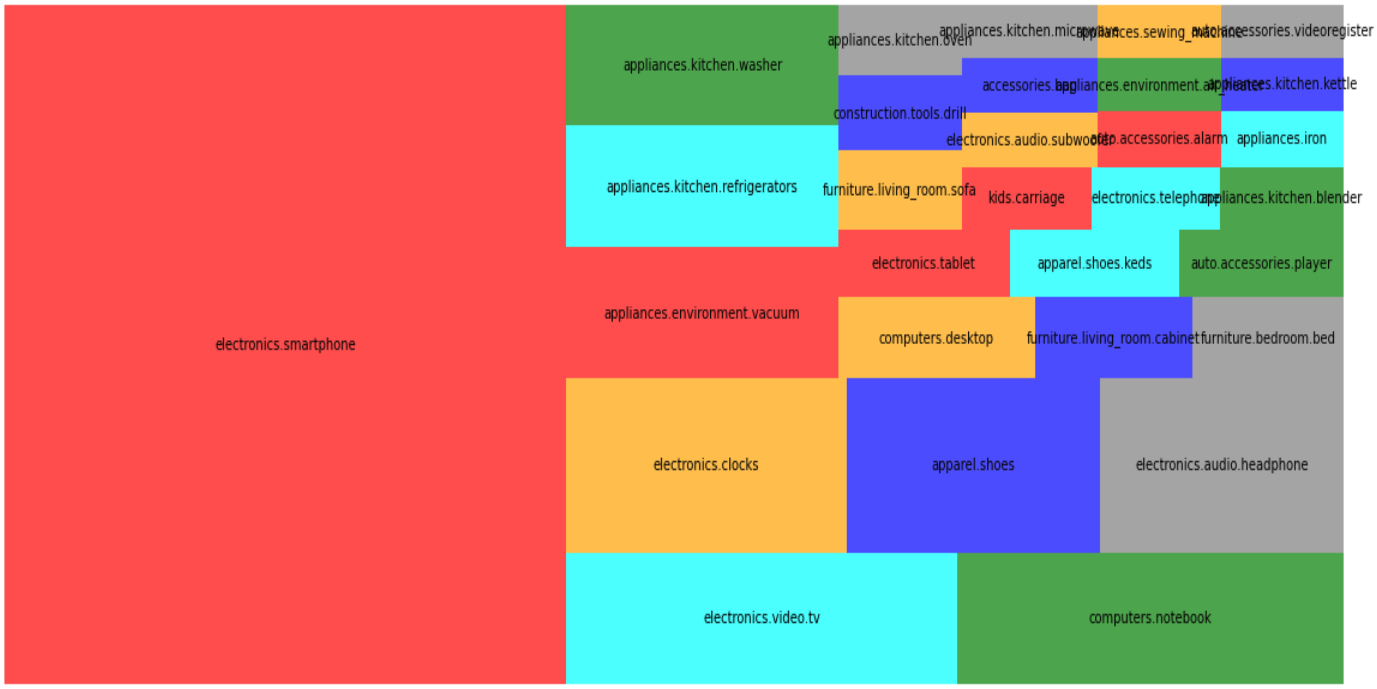


By Category and Product

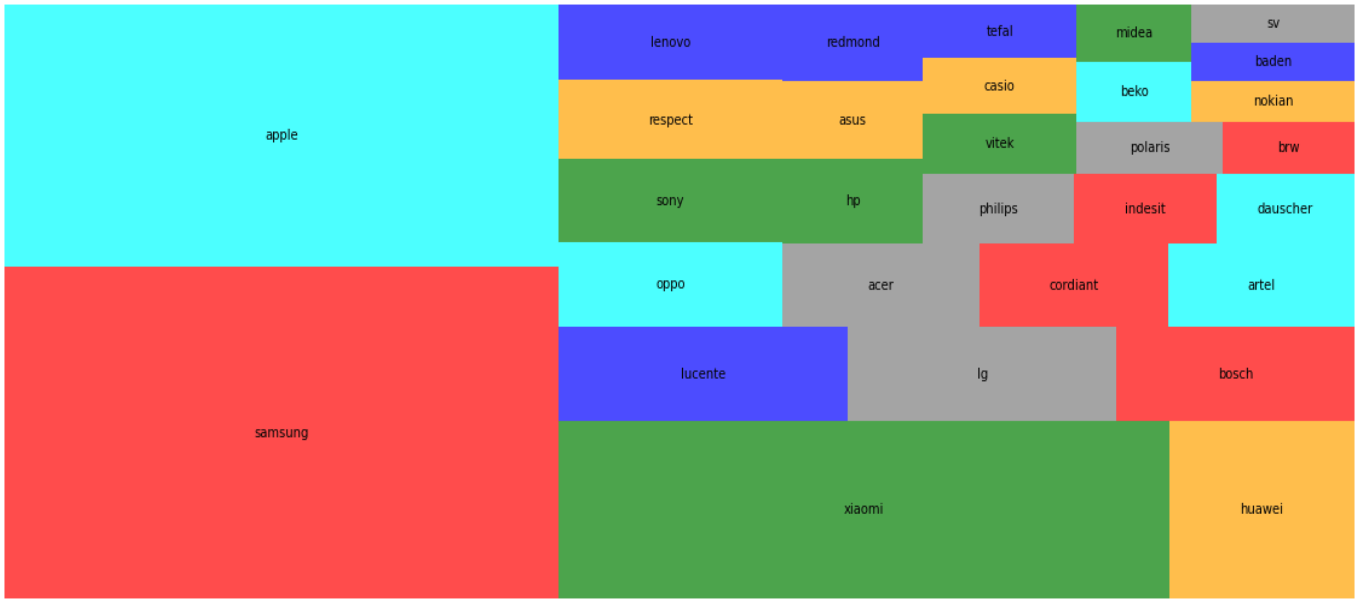
Which category do customers interact with the most and what brand do they view most.

By Category and Product

```
[ ]  top_category_n = 30
     top_category = data['category_code'].value_counts()[:top_category_n].sort_values(ascending=False)
     df = pd.DataFrame({'count':top_category, 'top_category':top_category })

     squarify.plot(sizes=top_category, label=top_category.index.array, color=["red","cyan","green","orange","blue","grey"], alpha=.7  )
     plt.axis('off')
     plt.show()
```




```
top_brand_n = 30
top_brand = data['brand'].value_counts()[:top_brand_n].sort_values(ascending=False)
df = pd.DataFrame({'count':top_brand, 'top_category':top_brand.index.array })

squarify.plot(sizes=top_brand, label=top_brand.index.array, color=["red","cyan","green","orange","blue","grey"], alpha=.7  )
plt.axis('off')
plt.show()
```

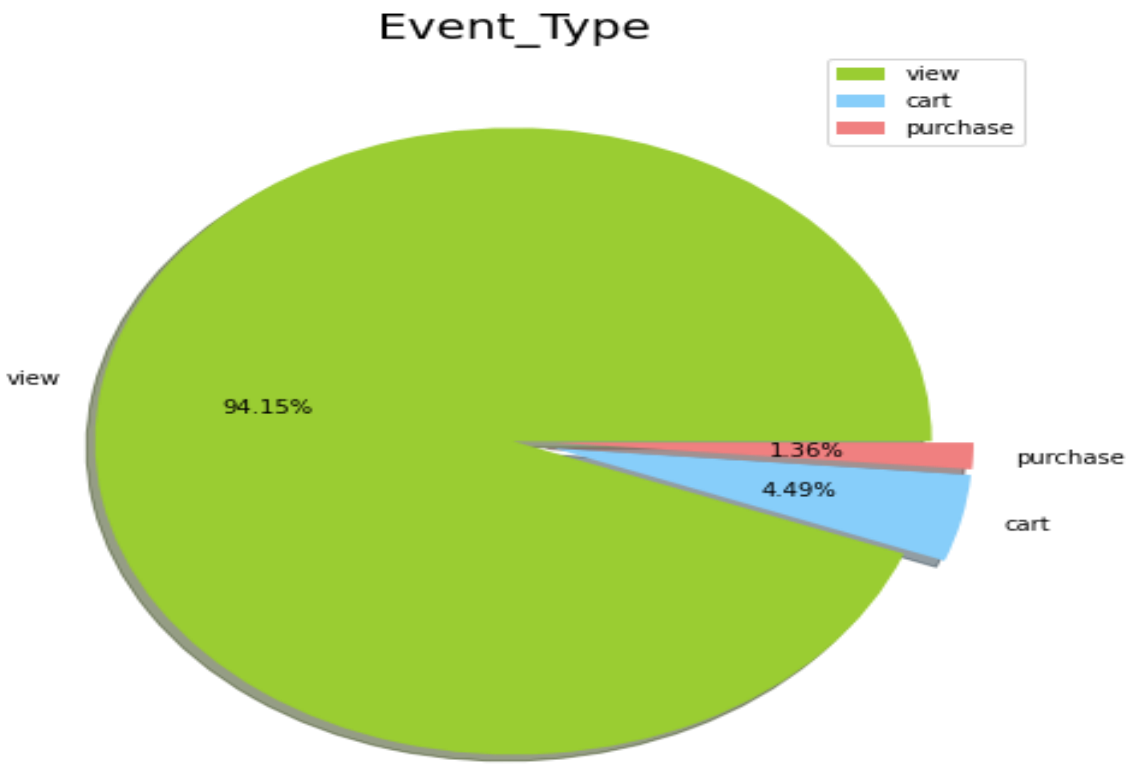


Purchase Behavior

What is the distribution of eventype?

```
[ ] data['event_type'].value_counts()

    view        63556110
    cart         3028930
    purchase      916939
    Name: event_type, dtype: int64
```

```
labels = ['view', 'cart','purchase']
size = data['event_type'].value_counts()
colors = ['yellowgreen', 'lightskyblue','lightcoral']
explode = [0, 0.1,0.1]

plt.rcParams['figure.figsize'] = (8, 8)
plt.pie(size, colors = colors, explode = explode, labels = labels, shadow = True, autopct = '%.2f%%')
plt.title('Event_Type', fontsize = 20)
plt.axis('off')
plt.legend()
plt.show()
```



What items do customers buy?

event_type = "purchase", what item do customers buy?

```
[ ]  purchase = data.loc[data['event_type'] == 'purchase']
     purchase = purchase.dropna(axis='rows')
     purchase
```

	event_time	event_type	product_id	category_id	category_code	brand	price	user_id	user_session	event_date
168	2019-11-01 00:01:04 UTC	purchase	1005161	2053013555631882655	electronics.smartphone	xiaomi	211.92	513351129	e6b7ce9b-1938-4e20-976c-8b4163aea11d	2019-11-01
707	2019-11-01 00:04:51 UTC	purchase	1004856	2053013555631882655	electronics.smartphone	samsung	128.42	562958505	0f039697-fedc-40fa-8830-39c1a024351d	2019-11-01
939	2019-11-01 00:06:33 UTC	purchase	1801881	2053013554415534427	electronics.video.tv	samsung	488.80	557746614	4d76d6d3-fff5-4880-8327-e9e57b618e0e	2019-11-01
942	2019-11-01 00:06:34 UTC	purchase	5800823	2053013553945772349	electronics.audio.subwoofer	nakamichi	123.56	514166940	8ef5214a-86ad-4d0b-8df3-4280dd411b47	2019-11-01
1107	2019-11-01 00:07:38 UTC	purchase	30000218	2127425436764865054	construction.tools.welding	magnetta	254.78	515240495	0253151d-5c84-4809-ba02-38ac405494e1	2019-11-01
...	...	...	...	...	...	...	...	...	...	...
67501716	2019-11-30 23:58:08 UTC	purchase	1004767	2053013555631882655	electronics.smartphone	samsung	235.60	574868869	878a1538-ebe3-4d7f-b773-1b057b1971eb	2019-11-30
67501725	2019-11-30 23:58:14 UTC	purchase	1004874	2053013555631882655	electronics.smartphone	samsung	346.70	547804983	717566cf-ef93-4078-ba8f-169a3ac9f1a0	2019-11-30
67501745	2019-11-30 23:58:22 UTC	purchase	1005130	2053013555631882655	electronics.smartphone	apple	1437.02	515582054	829c20b5-696e-4a8a-8a9f-171014a3ecbe	2019-11-30
67501824	2019-11-30 23:58:57 UTC	purchase	1004767	2053013555631882655	electronics.smartphone	samsung	235.60	579876821	ca50e291-43f3-4ca2-9e13-20ee6b8b25f0	2019-11-30
67501863	2019-11-30 23:59:15 UTC	purchase	3701309	2053013565983425517	appliances.environment.vacuum	polaris	89.32	543733099	a65116f4-ac53-4a41-ad68-6606788e674c	2019-11-30

659256 rows × 10 columns

What brands do the customers buy?

What brands the customers buy?

```
top_sellers = purchase.groupby('brand')['brand'].agg(['count']).sort_values('count', ascending=False)
top_sellers.head(20)
```

brand	count
samsung	198670
apple	165681
xiaomi	57909
huawei	23466
oppo	15080
lg	11828
artel	7269
lenovo	6546
acer	6402
bosch	5718
indesit	5187
respect	4557
hp	4002
midea	3984
elenberg	3944
haier	3826
beko	3813
casio	3477
tefal	3343
vitek	3095

Data Modeling

In this project,To use XGB classifier. XGBoost is an implementation of gradient boosted decision trees with good performance.

Another reason to choose xgboost is because of its great performance, considering the amount of data in this project. If we need to create dummy variables, the computation complexity increases dramatically, which might not favorable for quick exploration

Data Modeling

```
[ ] X_train, X_test, y_train, y_test = train_test_split(features,
                                                        is_purchased,
                                                        test_size = 0.3,
                                                        random_state = 0)
```

```
[ ] from xgboost import XGBClassifier
    model = XGBClassifier(learning_rate=0.1)
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)
```

Evaluate the Results and Measure Performance

Evaluate the Results and Measure Performance

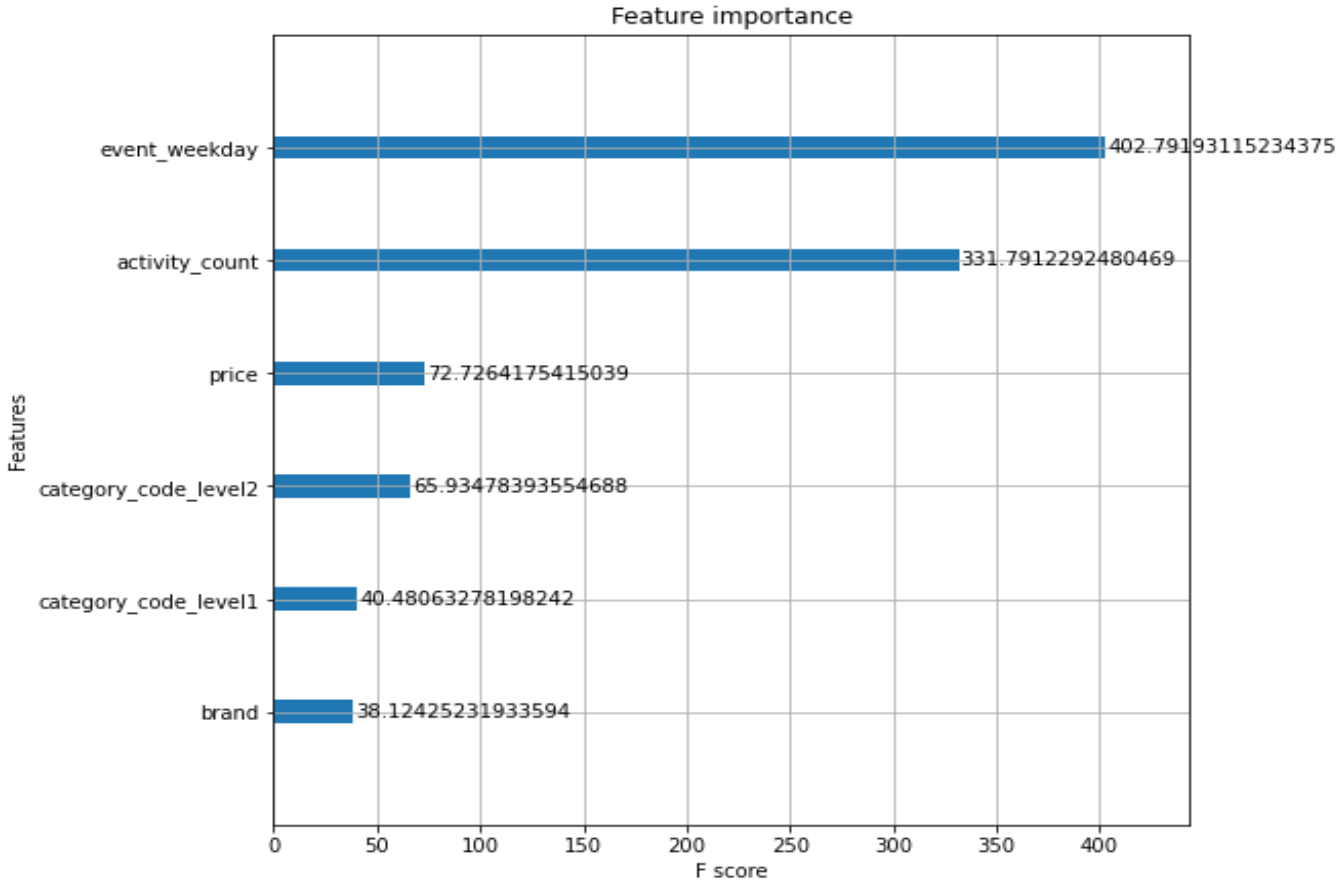
```
[ ] print("Accuracy:",metrics.accuracy_score(y_test, y_pred))
    print("Precision:",metrics.precision_score(y_test, y_pred))
    print("Recall:",metrics.recall_score(y_test, y_pred))
    print("fbeta:",metrics.fbeta_score(y_test, y_pred, average='weighted', beta=0.5))
```

```
Accuracy: 0.6983033333333334
Precision: 0.674964120179752
Recall: 0.7650082332784448
fbeta: 0.6991476840030184
```

Feature importance

Feature importance

```
plot_importance(model, max_num_features=10, importance_type ='gain')
plt.rcParams['figure.figsize'] = (40,10)
plt.show()
```



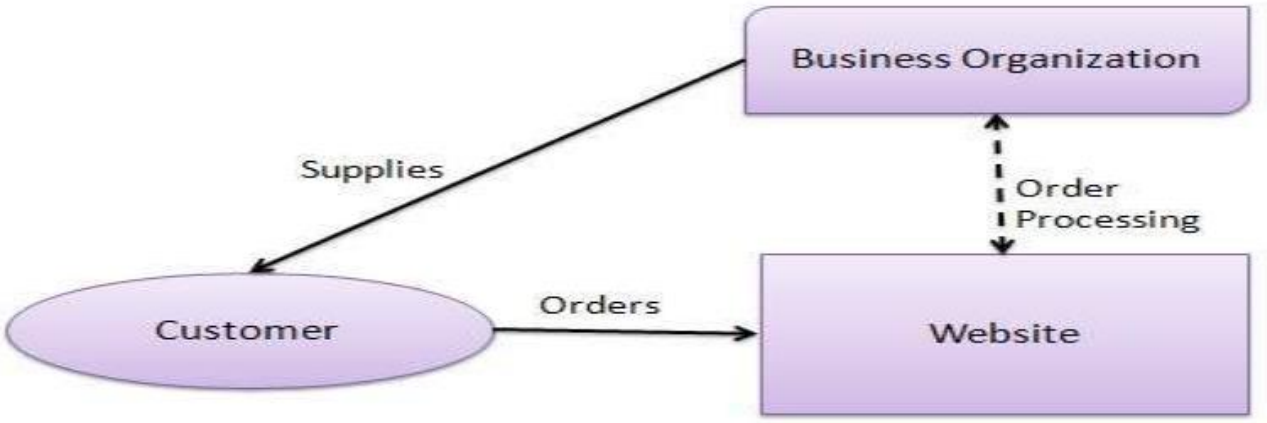
Step 3 :- Business Modelling

Before we choose the e-commerce business model that will enable us to sell our products to customers, we'll need to figure out who our customer is. For instance, are we selling directly to consumers, or are we supplying something to another business or a government entity?

Business-to-consumer (B2C)

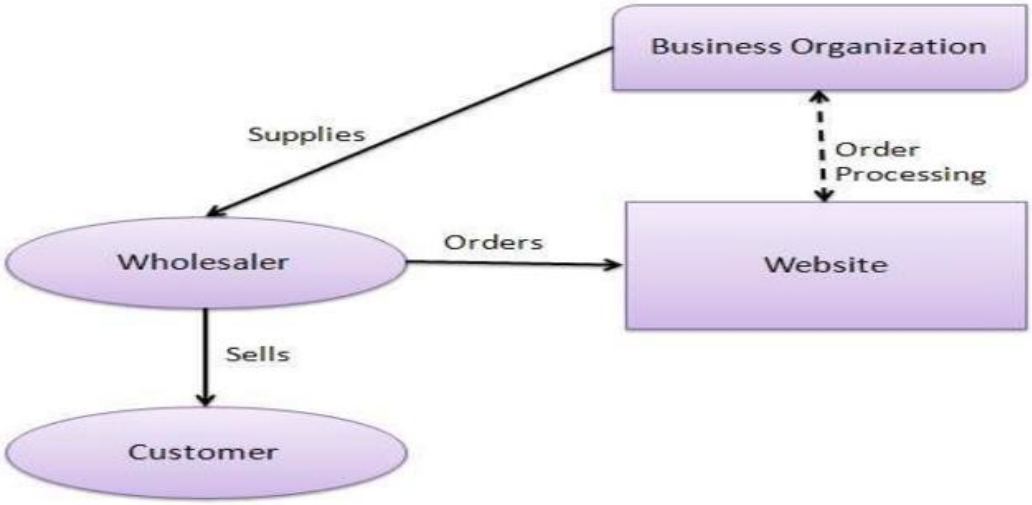
A website following the B2C business model sells its products directly to a customer. A customer can view the products shown on the website. The customer can choose a product and order the same. The website will then

send a notification to the business organization via email and the organization will dispatch the product/goods to the customer.



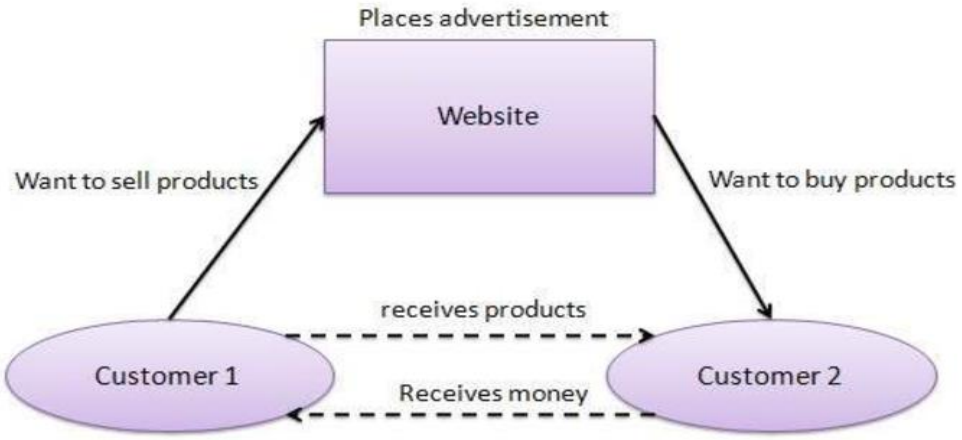
Business-to-business (B2B)

A website following the B2B business model sells its products to an intermediate buyer who then sells the product to the final customer. As an example, a wholesaler places an order from a company's website and after receiving the consignment, sells the endproduct to the final customer who comes to buy the product at one of its retail outlets.



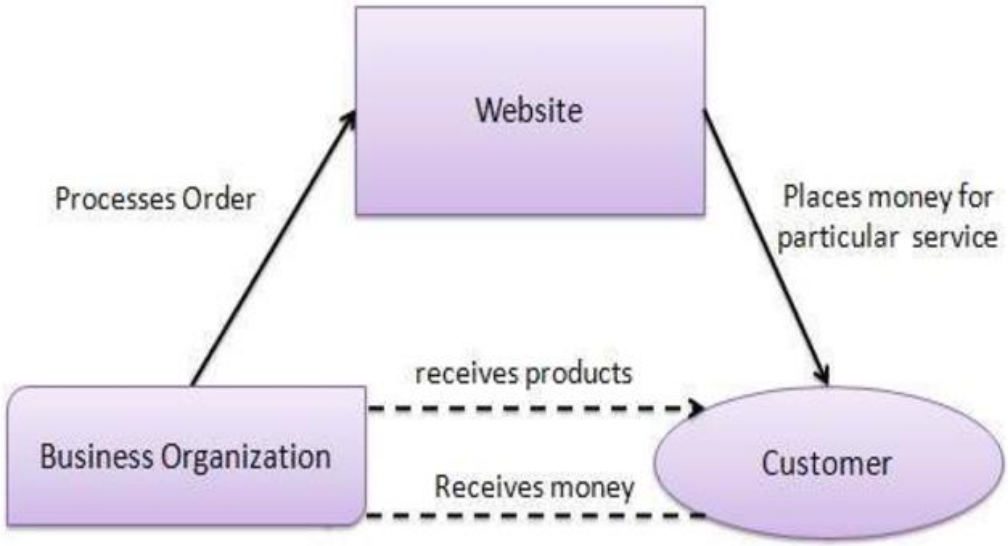
Consumer-to-consumer (C2C) or peer-to-peer (P2P)

A website following the C2C business model helps consumers to sell their assets like residential property, cars, motorcycles, etc., or rent a room by publishing their information on the website. Website may or may not charge the consumer for its services. Another consumer may opt to buy the product of the first customer by viewing the post/advertisement on the website.



Consumer-to-business (C2B)

In this model, a consumer approaches a website showing multiple business organizations for a particular service. The consumer places an estimate of amount he/she wants to spend for a particular service. For example, the comparison of interest rates of personal loan/car loan provided by various banks via websites. A business organization who fulfills the consumer's requirement within the specified budget, approaches the customer and provides its services.



Business-to-government (B2G)

In this model, a consumer approaches a website showing multiple business organizations for a particular service. The consumer places an estimate of amount he/she wants to spend for a particular service. For example, the comparison of interest rates of personal loan/car loan provided by various banks via websites. A business organization who fulfills the consumer's requirement within the specified budget, approaches the customer and provides its services.

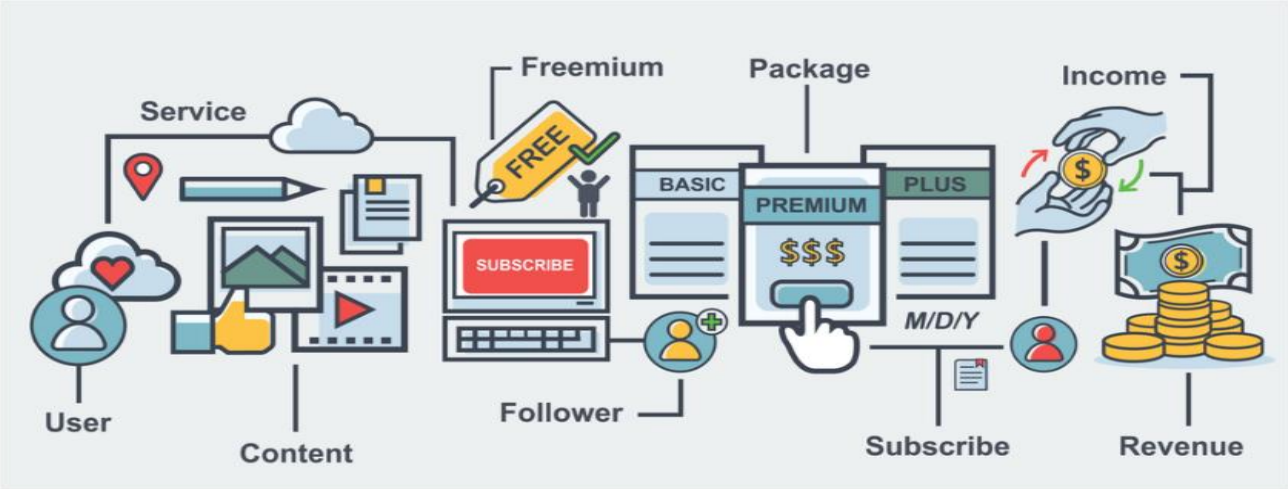
E-commerce business models

There are a few core types of e-commerce business models that online businesses use to reach customers and drive sales and we will use them according to aur customer analysis and provide to companies.

Subscriptions

Subscription e-commerce business models have risen dramatically in popularity. With a subscription, our customer signs up to receive products in regular intervals. Subscriptions enable businesses to develop strong recurring revenue streams and maintain customer relationships for longer than many other businesses that rely on one-off purchases.

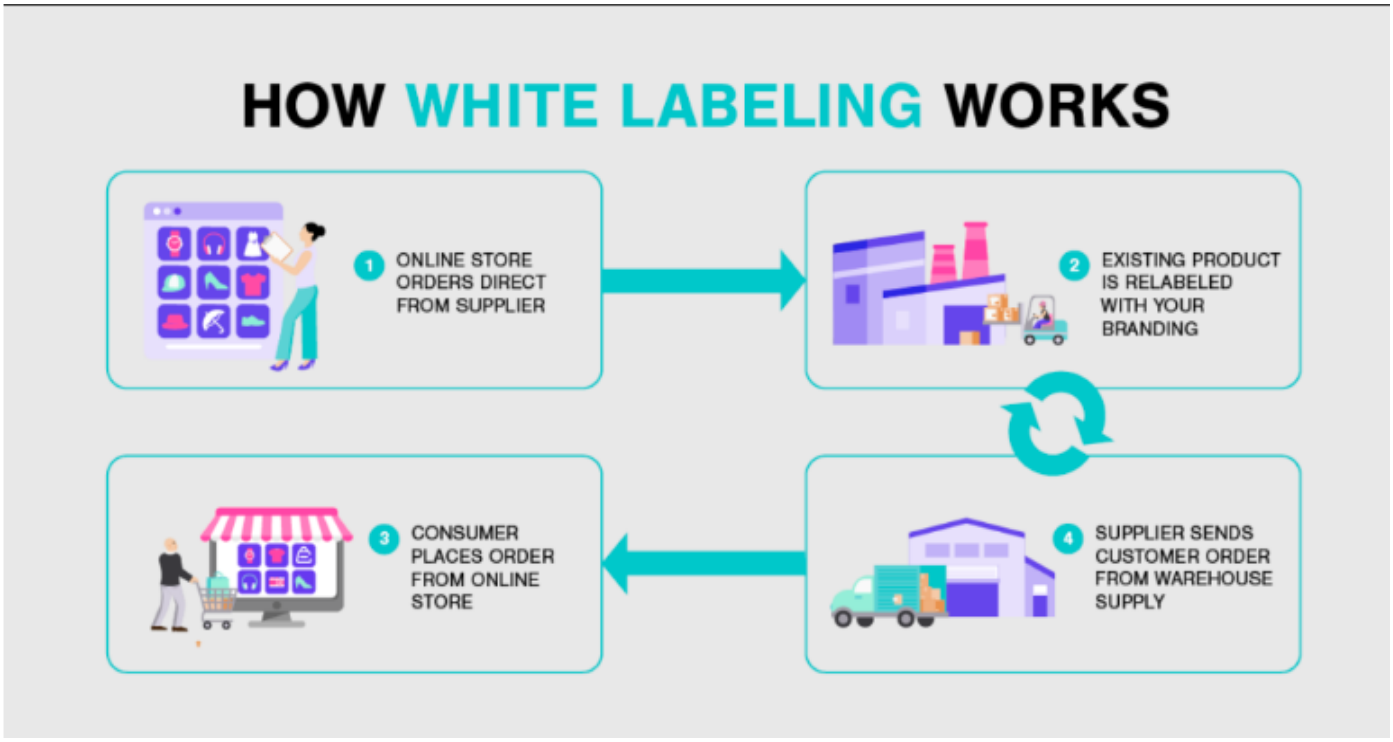
Subscriptions have worked for many different types of businesses — from toothbrushes and razors to cleaning products, meal kits and clothing. If your product is something that customers will use frequently and either need to replace, reuse or may run out of, a subscription e-commerce business model could work for you.



White-labeling

With white-labeling, you choose a product that's already manufactured and sold. However, you'd brand the merchandise with your company's name, design a label and packaging and then sell the product. When you white-label goods, you may have to buy in large quantities. There's a higher risk for dead stock this way, and you'll also need to consider warehousing this inventory.

White-labeling isn't the best fit for every business, especially those that create specialty products with proprietary formulations. However, it can be good for those who are looking to build a brand. White-labeling is common in the beauty industry where some formulas are created at a central lab and then rebranded by various companies.



On-demand manufacturing

If inventory space is a concern for you, you might want to consider on-demand manufacturing as an e-commerce business model. In this type of model, you create the good when the customer orders it. This way, you're not stuck with stock that you don't know whether you can move. You can also potentially offer lots of customization for your customers

On-demand manufacturing (sometimes called print-on-demand or POD) can sometimes mean orders have a longer lead time to get to customers because they'll have to wait for their product to be made versus plucked off a shelf. It's important to know whether your customer is willing to wait and be sure to be upfront about your timeline.

Dropshipping

An extremely popular e-commerce business model is dropshipping. With dropshipping, you don't have to worry about housing inventory, keeping track of it or even shipping it to your customers. You'll work with a warehouse partner who will keep your inventory and manage all of the operations once the orders have been placed. That can free you up to think about the other elements of your business, including merchandising, branding and marketing.

There is a fee to working with a warehouse for a dropshipping business. But, if you're selling at a volume that you simply can't sustain in your office alone, dropshipping is an option. It also allows you to experiment with what you sell without having to pay upfront for stock that may not end up selling.

Wholesaling and warehousing

With wholesaling, you often keep a significant amount of inventory — a big upfront investment — and require buyers to purchase from you in bulk. You'll need a lot of space to keep your products, but if you have the real estate, you can offer many types of goods, which can give you a competitive advantage. Volume is the name of the game with a wholesaling business.

Wholesaling used to be exclusively for B2B businesses, but it's expanded into allowing customers to purchase directly from wholesalers. This generally happens by providing an option for customers to buy in smaller quantities. These transactions often bear higher profit margins, since you're selling goods individually or in smaller packs.

Step 4 : - Financial Modelling

The model can provide service to e-commerce websites, where behavioural data is generated by, or in response to, a customer's engagement with a business.

The Indian e-commerce sector has received more investment than any other sector in 2021

The Indian online grocery market is estimated to reach US$ 26.93billion in 2027 from US$ 3.95 billion in FY21, expanding at a CAGR of 33%

31% business in India are focused on implementing new analytical methods and building AI models to improve their services.

99% business are implementing strategies related to recognising customers their customers across various platforms.

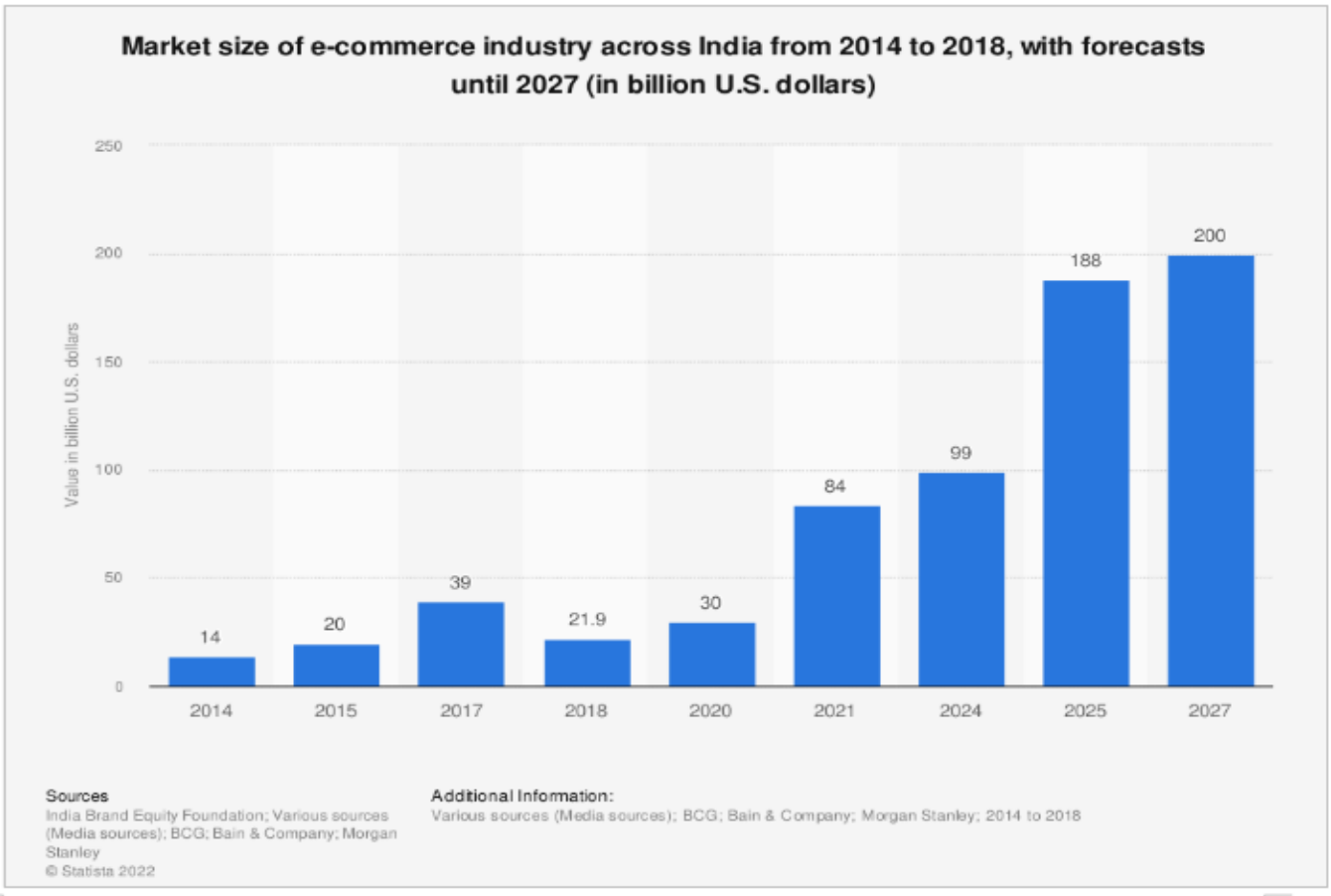
So, from above reports it would be appropriate to charge our service for 20,000 every three months. After increasing business, we would increase our charges and increase the features in the model.

To develop a model, it would take 2 to 3 weeks for which we need two data scientist, one machine learning engineer to perform algorithm and a software developer to build software.

If “s” is the salary of employees and “q” is the maintenance of the company, “y” is the profit of company and “m” is the pricing of product then a simple financial equation will be

y = m*x(t) – (s+q)

where x(t) is the market growth



Chapter 3 – Conclusion and References

Conclusion: -

Though online shopping is very common outside India, its growth in Indian market, which is a large customer market, is still not in line with the global market. According to India B2C E-Commerce Report 2013, e-tailing accounts for less than 1 percent of the overall retail market in India in 2012. While it accounts for over 5 percent if the total retail market in China and 10% in the UK and the US. This shows that only a small fraction of internet users in India are currently online shoppers. The reason could be that it is not the technology but the way customers feel about high-tech purchasing that is holding back the development of the industry (Reda, 1995). Hence this is imperative that marketers understand the depth of customer intentions for this medium of retailing. It can be concluded on the basis of study that online shopping is gaining popularity among people of young generation. Higher income groups and educated people are purchasing more via e-retailing websites. People have hesitations in doing online shopping due to security concerns. At the same time people are resistant to change because of. All the impact factors, flawless adoption in consumer's life, acceptance of e-commerce business style, responses to e-tailing and benefits of e-commerce are discussed in details backed by available literature. At the end of the literature review the importance trust and security factor and the future need is also discussed. Due to adoption of e-commerce in life, consumers have started expecting more which can impact on their buying behaviour. The factors of the e-commerce that influenced consumer behaviours need to be carefully analysed by the online retailers, who can utilize the appropriate marketing communications to respond the change in customer's purchase decision making process and improve their performance. Hence in terms of promising future research directions, two avenues deserve special mention, in researcher's opinion, first is, there is need to study in detail the impact of

e-commerce on consumer buying behaviour and second is to find out to what extent the behaviour can be changed.

References : -

1. https://towardsdatascience.com/data-analytics-in-e-commerce-retail-7ea42b561c2f
2. https://towardsdatascience.com/exploratory-data-analysis-on-e-commerce-data-be24c72b32b2
3. https://github.com/suciaulyaputri/EDA-on-E-Commerce-Behaviour-Data
4. https://www.investopedia.com/terms/b/businessmodel.asp#:~:text=For%20instance%2C%20direct%20sales%2C%20franchising,sporting%20organizations%20like%20the%20NBA
5. https://alcorfund.com/insight/18-business-model-example-explained/
6. https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-to-time-series-analysis/
7. https://www.analyticsvidhya.com/blog/2021/10/machine-learning-for-stock-market-prediction-with-step-by-step-implementation/
8. https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store