# E-retail factors for customer activation and retention: A case study from Indian e-commerce customers

- *Problem statement and understanding*
- *EDA steps and visualizations*
- *Steps and assumptions used to complete the project*
- *analysis, and conclusion*

# Problem Statement

- ✓ With increasing competition in the online store space its very important that every action/strategies we make(we means the online stores) should be effective enough in bringing the desired change and increasing the chances of landing more customers that too in required period of time.

- ✓ Detecting the key factors/features which decides the market share of the stores .

- ✓ Getting a idea of what issues the customers face while shopping and what are the chances they may abandon the store leading to loss of potential customers

- ✓ Discussing the measures to be taken to overcome the shortcoming/lags if any and increasing the engagement of customers with the store leading to increased profitability.

# UNDERSTANDINGS:-

A thorough analysis of the features tells a lot about the patterns opted by the customers like

- Age groups more active in shopping online
- Area/regions contributing more in the customer base
- Preferences of customers
- What factors are helping certain online stores to attract more and how they are different from others who lags in the market size
- How good the stores is when it comes to connect with the values of customers (social, patriotic , gives them a sense of fulfilment , sense of proud) when customers engage and make payments
- How safe customers feel while making transactions and what mode/gateways are preffered when it comes to  make payments
- How building trust is a key feature which decides whether or not that store will see growth in number of customers
- Gender wise distribution of customers and how frequent are they in shopping online further classified using age groups
- What factors leads in abondaning the bag/cart
- Average  time in m
- inutes/spans a site or application can hold the customers and convert it into a purchase.

This is just a gist of what are the core points we are going to touch followed by the detailed analysis

# EXPLORATORY DATA ANALYSIS:
## WITH STEPS ,ASSUMPTIONS AND VISUALISATIONS

**Note:**-All the analysis is done in juypter notebook

## STEPS INVOLVED BEFORE ACTUAL ANALYSIS

- Step 1: importing all the necessary libraries which are going to be used in the notebook and eda

- Step2 :extracting the excel file with  sheetname as " datasheet" using pandas

- Step 3: observing the total number of columns/features check

- Step 3: checking the type of data in columns and separating all the categorical columns header in a separate list

- Step 4: checking for the nulls if any

- Step 5: creating a dataframe of column name, number of unique features, name of the unique features which helps us getting a clear idea of the categories

# ANALYSIS:

## *BASED ON THE GENDER OF THE CUSTOMERS FURTHER CLASSIFIED WITH THE AGE GROUPS*
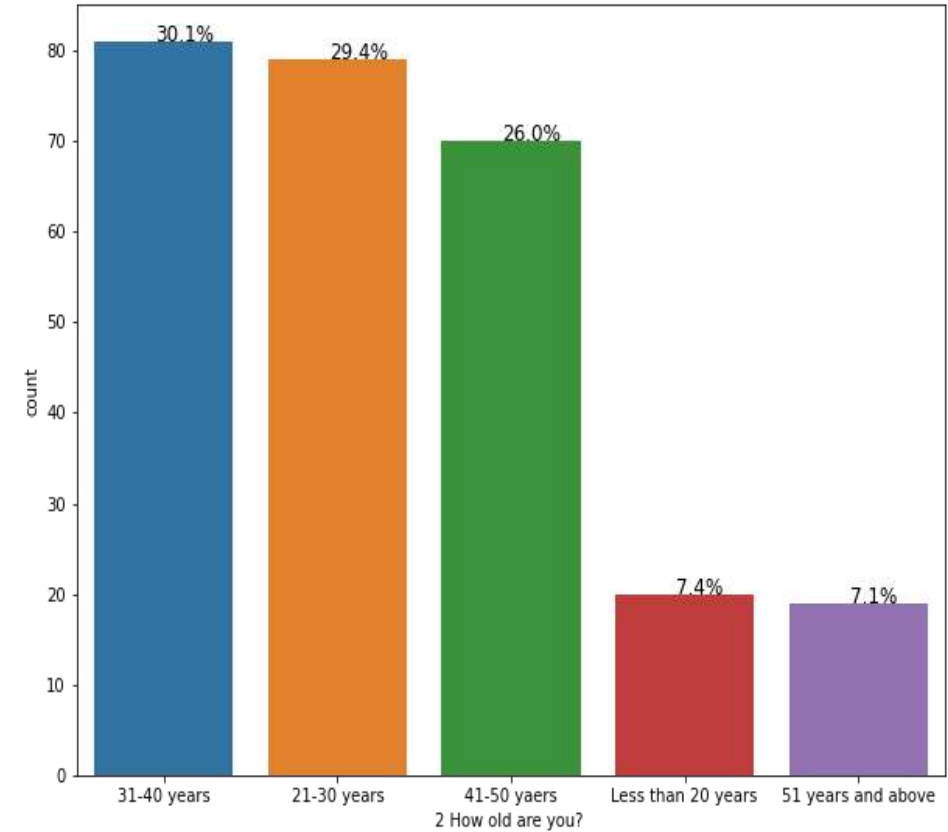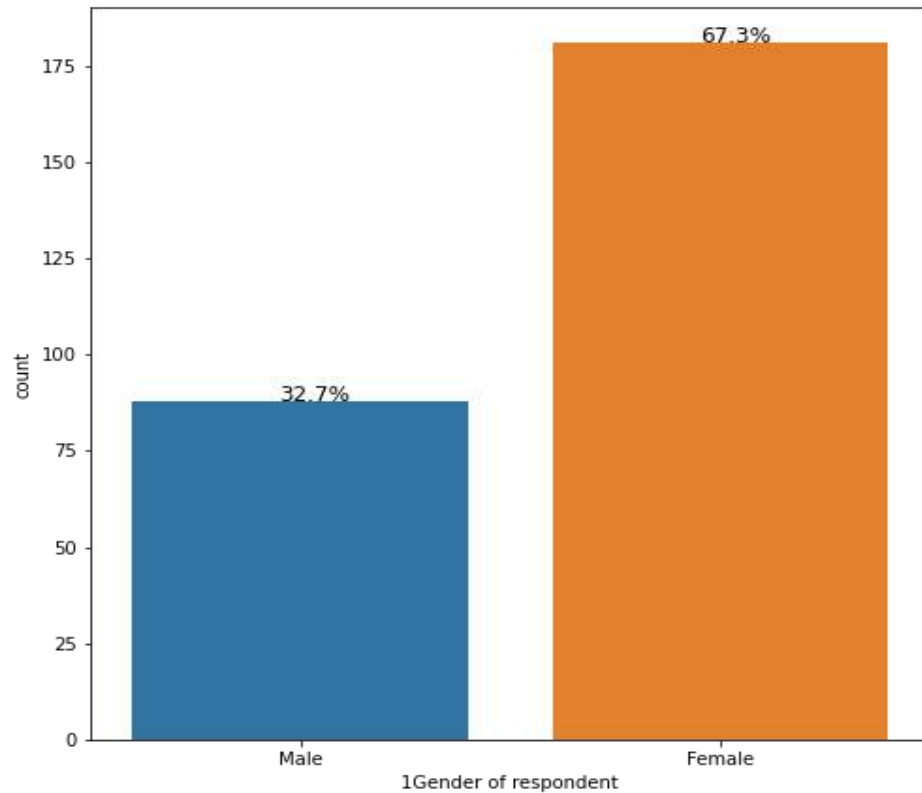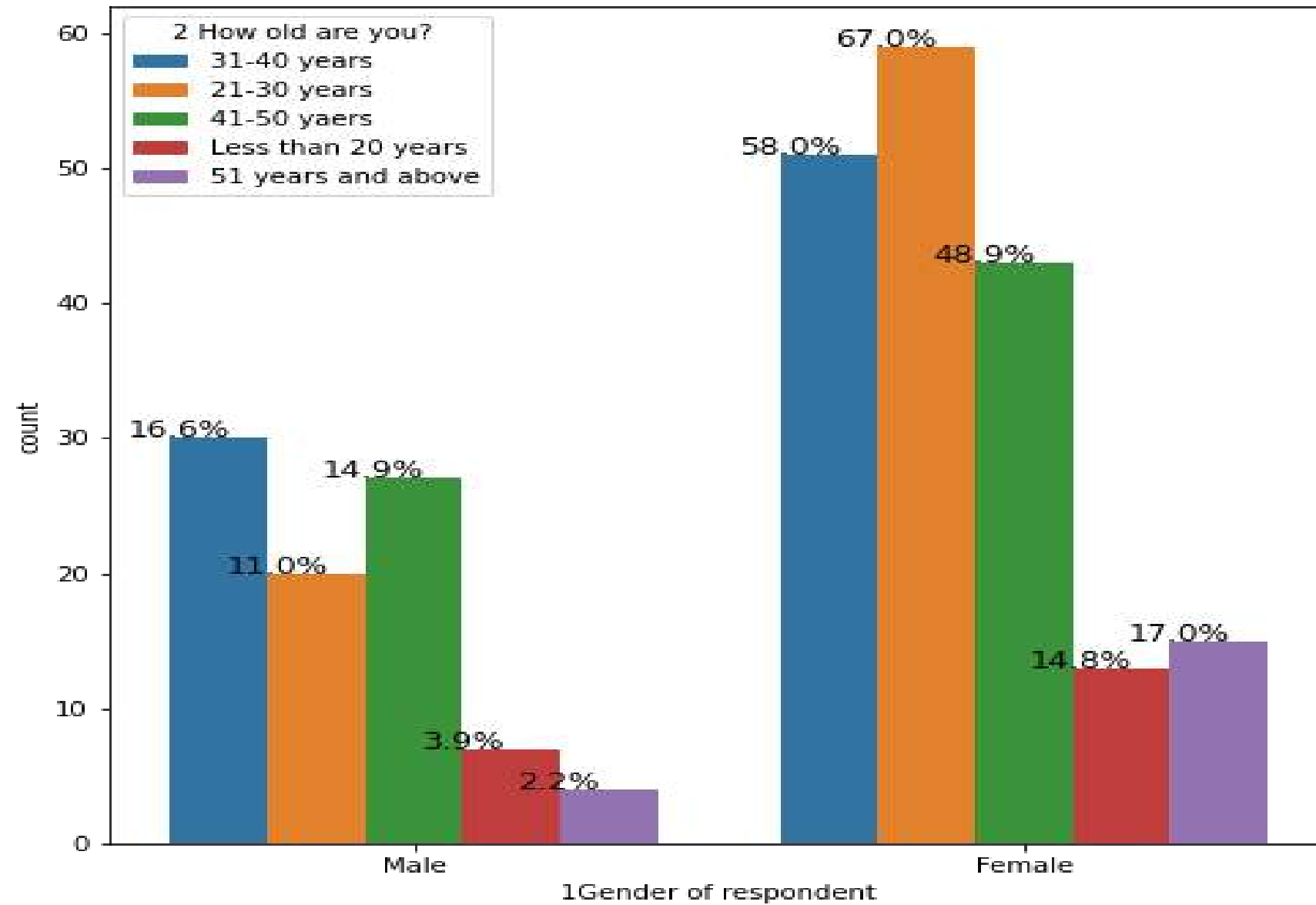
- **Its can be observed that females constitute a major portion of the customer base around 67% of the total and males constitute only 32.7%**

- **Customers in age groups 21-30,31-40,41-50 constitute almost same number of portions and a there is a huge difference to those in less that 20 years category and more than 51 years category.**

o There are various reasons for this huge gap like:

➤ For age group less than 20 years (in india) financial independence is a major issue which hinders that group to be a active category

➤ Not having access to smartphones/laptops or having parental control can also be the reason for such a low participation in our customerbase

▪ For age group above 51 years :

➤ Lack of interest from materialistics things after a certain age

➤ More focus on savings for the post retirement life

➤ Maybe the offerings from the is such that they are foused on engaging a certaing age group
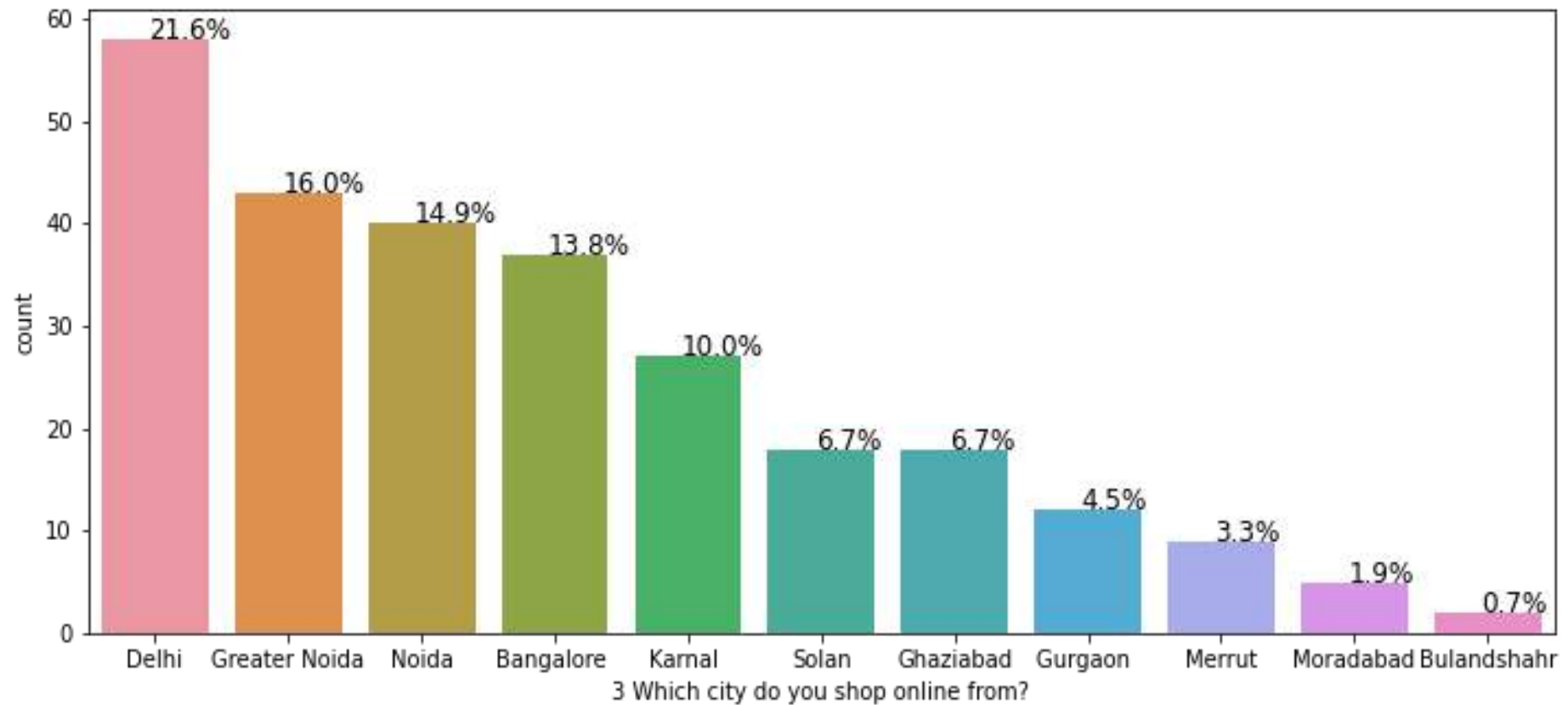
These might be the reasons for low percentages of post 51 years customers category

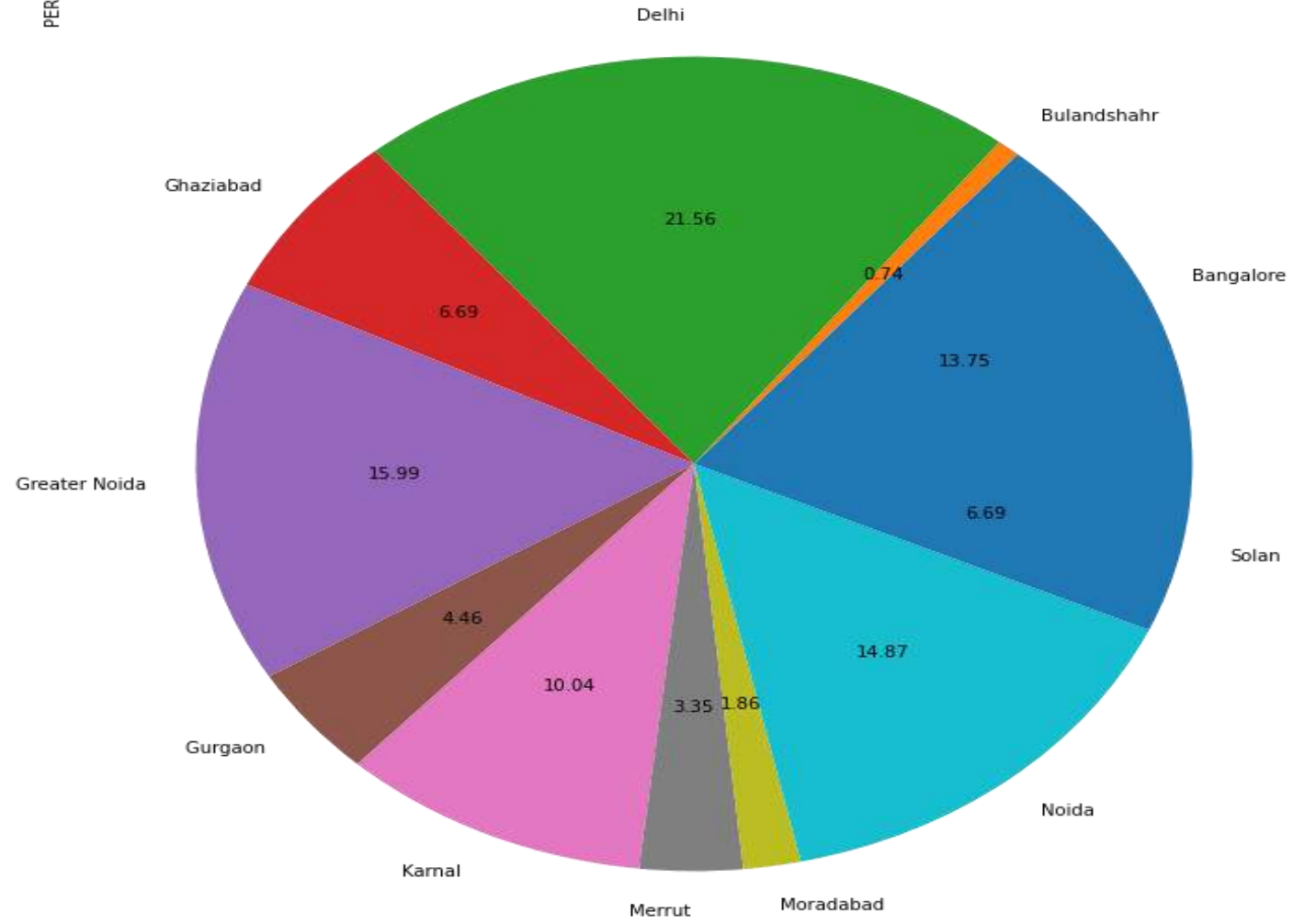<u>Measures that can be taken to make this distribution quiet even with low variance:</u>

- Launching new ad campaigns focusing specially on the portions which contribute low (in this case Male and those lying in age group (less than 20 and above 50)
- For those who are financially dependent (less than 20 years) making pocket friendly carts which is affordable with the pocketmoney/savings

(for school /college going students)

- For the senior ceitizens improvising the readability on site or making a separate mode on application/site (which includes large fontsize of texts with suitable headings and colors) as they may not be able to parse through because of vision problems
- Focusing more on medical products category like(pulse/bp monitoring euiqpments , pain relieif  equipments,belts etc) which is oftenly used by the senior citizens
- These are some of the measures that can be taken to overcome uneven distribution In the customer base

# CITYWISE DISTRIBUTION OF CUSTOMERS

PER CITY

Delhi 21.56
Bulandshahr 0.74
Bangalore 13.75
Solan 6.69
Noida 14.87
Moradabad 1.86
Merrut 3.35
Karnal 10.04
Gurgaon 4.46
Greater Noida 15.99
Ghaziabad 6.69

Its very clear from the plot that major metropolitan cities like Delhi ,Greater Noida ,Bangalore contributes a lot in total percent of online purchase being made.

There are various factors which may be leading to low percentage of customers from tier-2 ,tier-3 cities like:

- Campaigns run by these stores are not able to influence customers from these cities as they do in case of top tier cities

- People having issues in building trust as rumours spread very easily and unavailability of resources to cross check these rumours make people that they are true and affecting the sales in online purchase

- Longer Delivery periods and unavailabitiy of features like replacements policies  in some regions even make it harder to aquire new customers

- Economical weaker cities

Measures that can be taken:

➢ Working on the logistics and making it more penetrable so it can reach

➢ To backward cities and regions

➢ Running ads and campaigns that target people from these cities helping them to build trust so new customers can be aquired and retaining the old ones

# Trends of customers based on years they been purchasing from these online stores
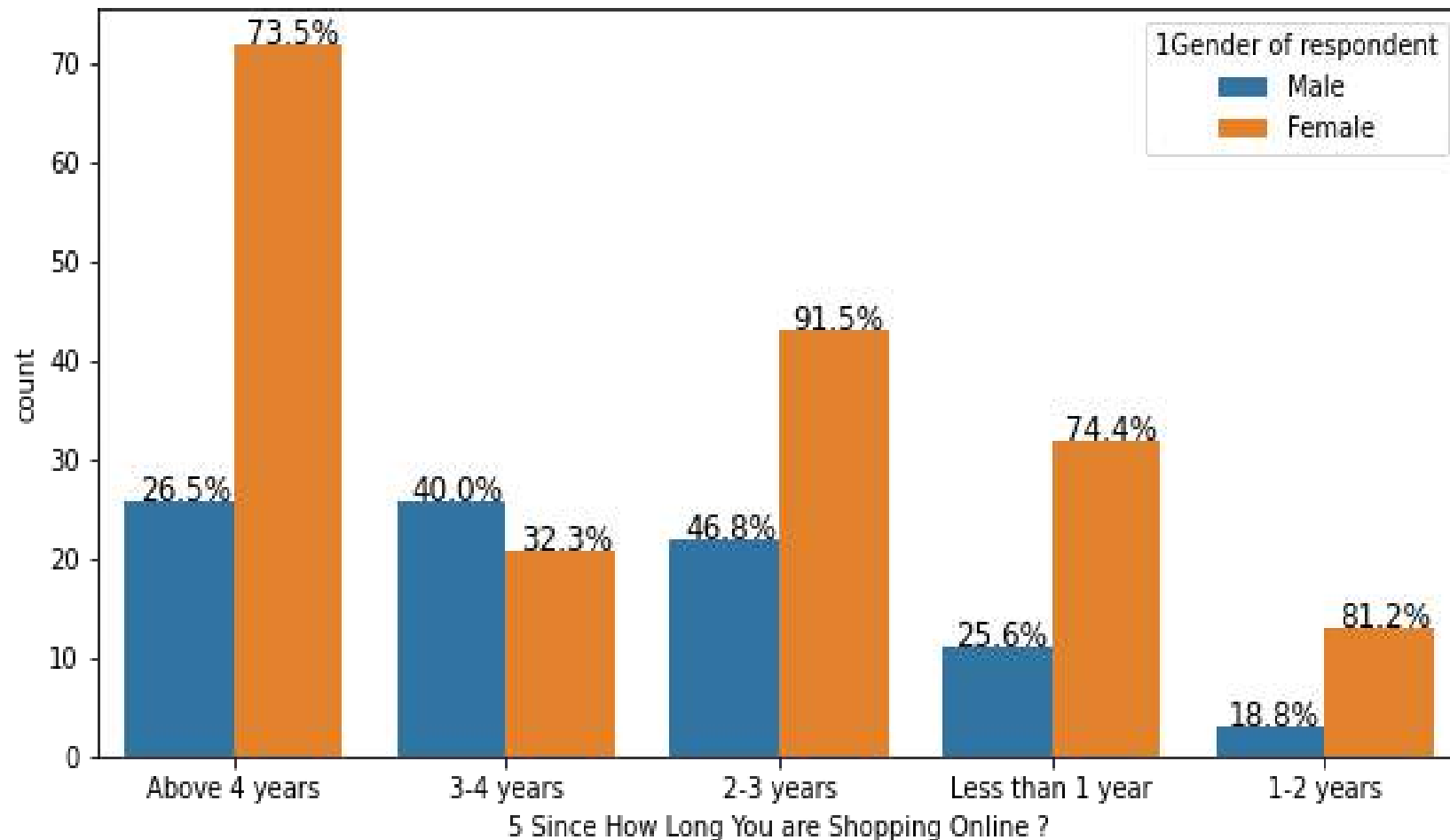
It can be observed that there is a good amount of retention of customers as number of year increases.

Main reason maybe the trust a customer develops ,smoothened processes ,flexibility they enjoy while shopping online
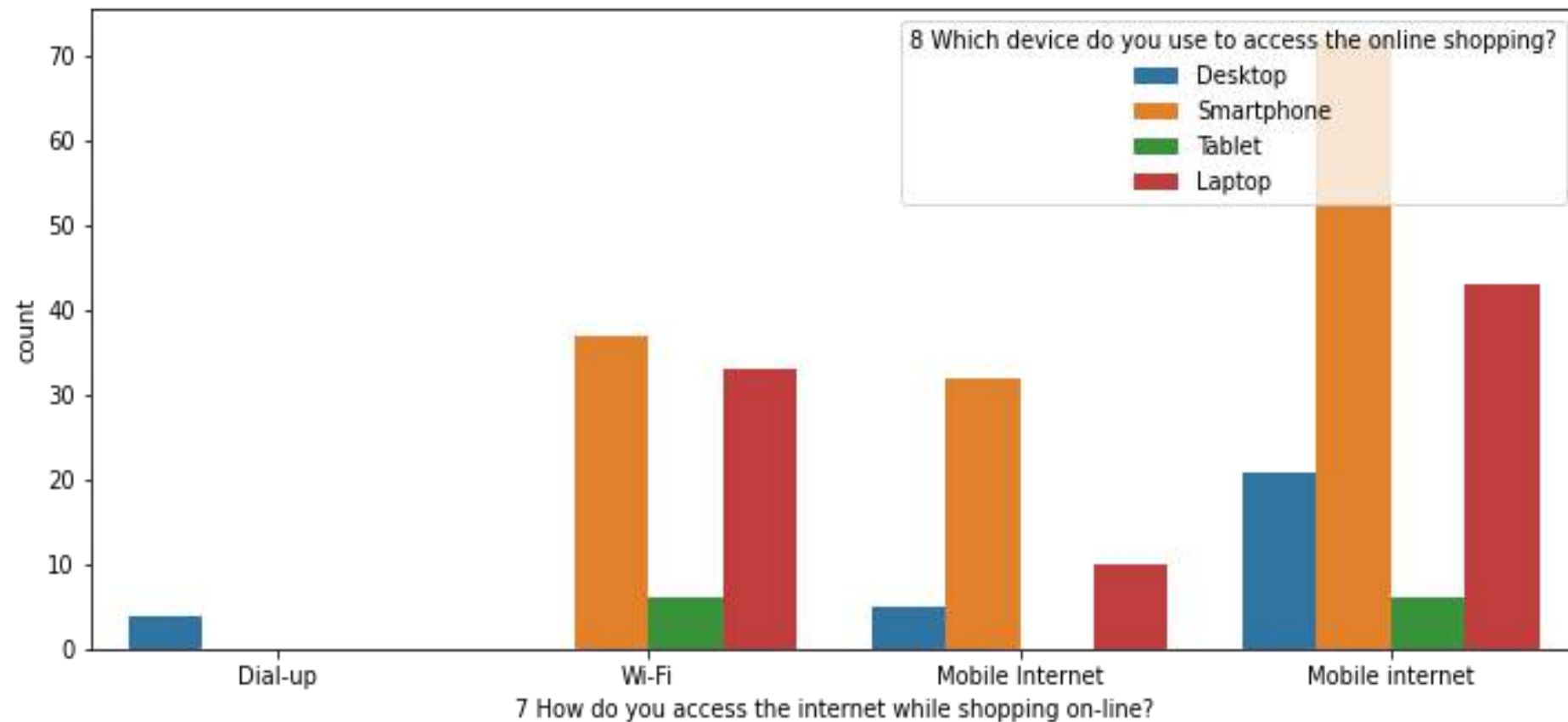
Wide variety to choose from,more options available

Easy transactions as technology is getting better

Measures to be taken:

Retention seems pretty fine gotta work more on acquisition of new customers

# Based on how customers access the internet while shopping online and further categorized on device they use

Mobile internet remains top choice for accessing the internet and that too on smart phones followed by the laptops

Quiet obvious that smart phones as being quiet handy in day to day activities and most of the customers have easy access to it is top choice for online shopping followed by laptops

Measures that can be taken:

➤ As we know maximum customers have easy to smartphones and mobile internet more focus should be on improving the mobile application interfacability and quickability.

➤ Making the apps more appealing with improved graphics which creates some positive vibes when customers use it

➤ Making the application lite versions which can be easily used on low edge phones giving access to economical weaker sections too giving them a sense of fulfillness and pride.

# Plot showing the distribution when it comes to browsers
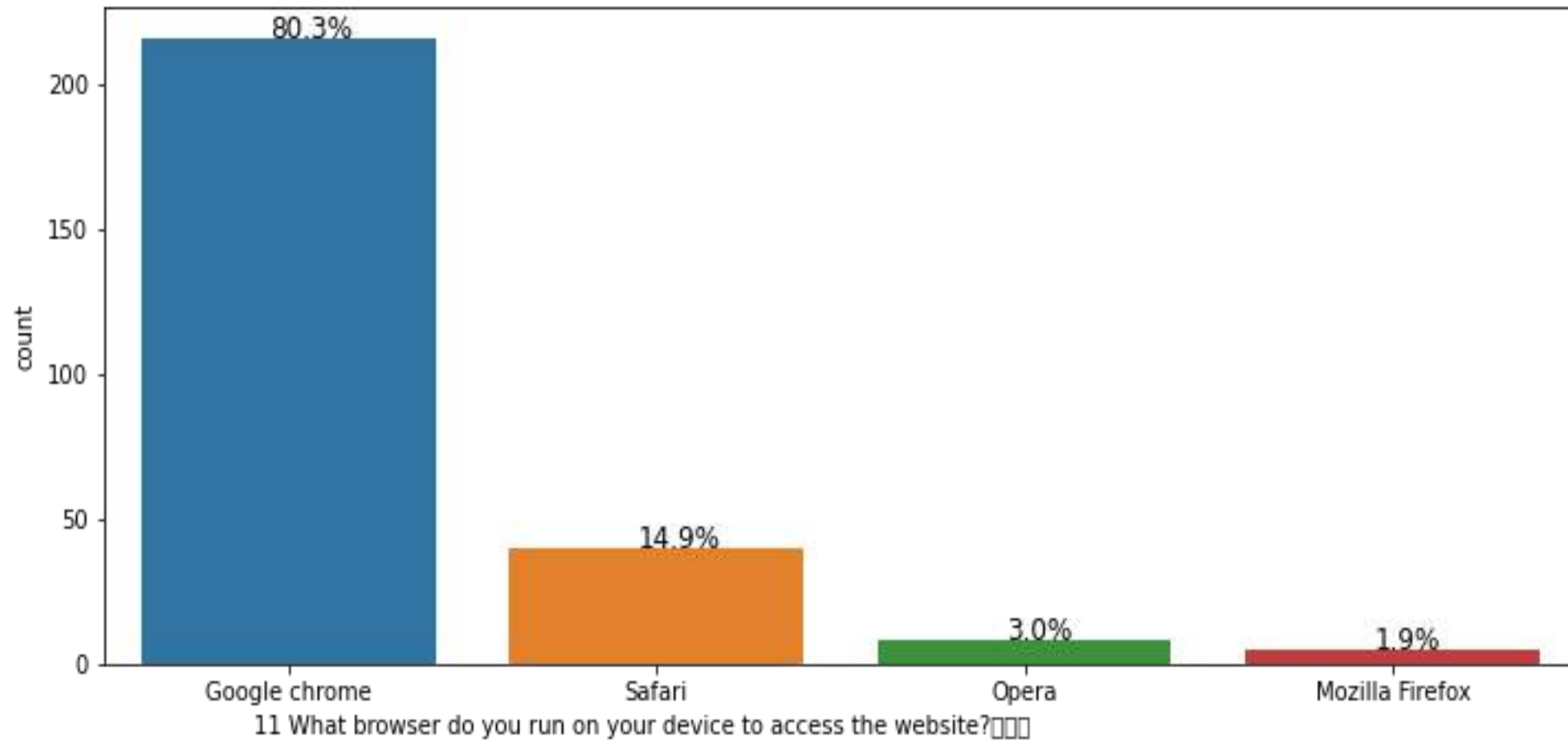
Google chrome is the most  widely used browser for online shopping

Reasons are obvious :

- Better performance when it comes to loading large sites
- Per-downloaded browser in the phone which gives it a edge above as most people start using it on a daily basis and starts getting used to it.
- Opens almost all sites with minor or low crashes
- Privacy is intact and transactions can be done easy without any hassle

Measures to be taken:

- Making the sites more readable
- Removig unnecessary graphics and loads addons from the pages
- Improving the interface

# Fast loading websites and applications countplot
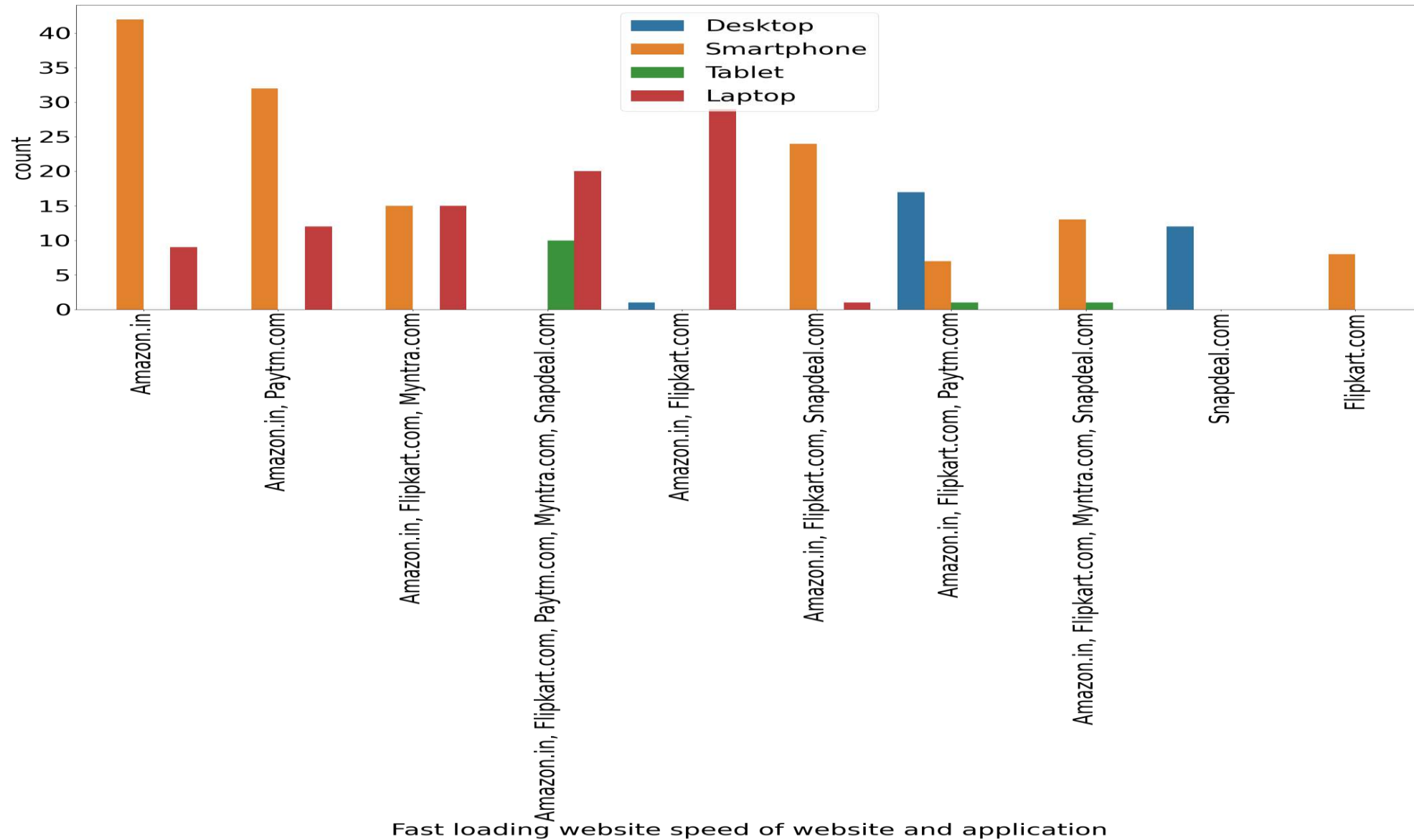
- It can be observed that most of the people are using either mobilephones/smartphones or laptops to make a purchase or surf through the store
- In this category its easily visible that amazon did a great job with their mobile application and site as it takes less time to load and takes advantage over the competitors
- Snapdeal ,flipkart have to work on their applications and sites as they are not very responsive as compared with the competitors

***Benefits :***

- Taking a edge over the competition as fast loading and fast responses can make the surfing smooth for the customers and can convert them into a sale
- Saves times of customers as they don't have to wait much
- Prevents churning of customers to the competitors

# Relationship between loading/processing speed and device used for online shopping
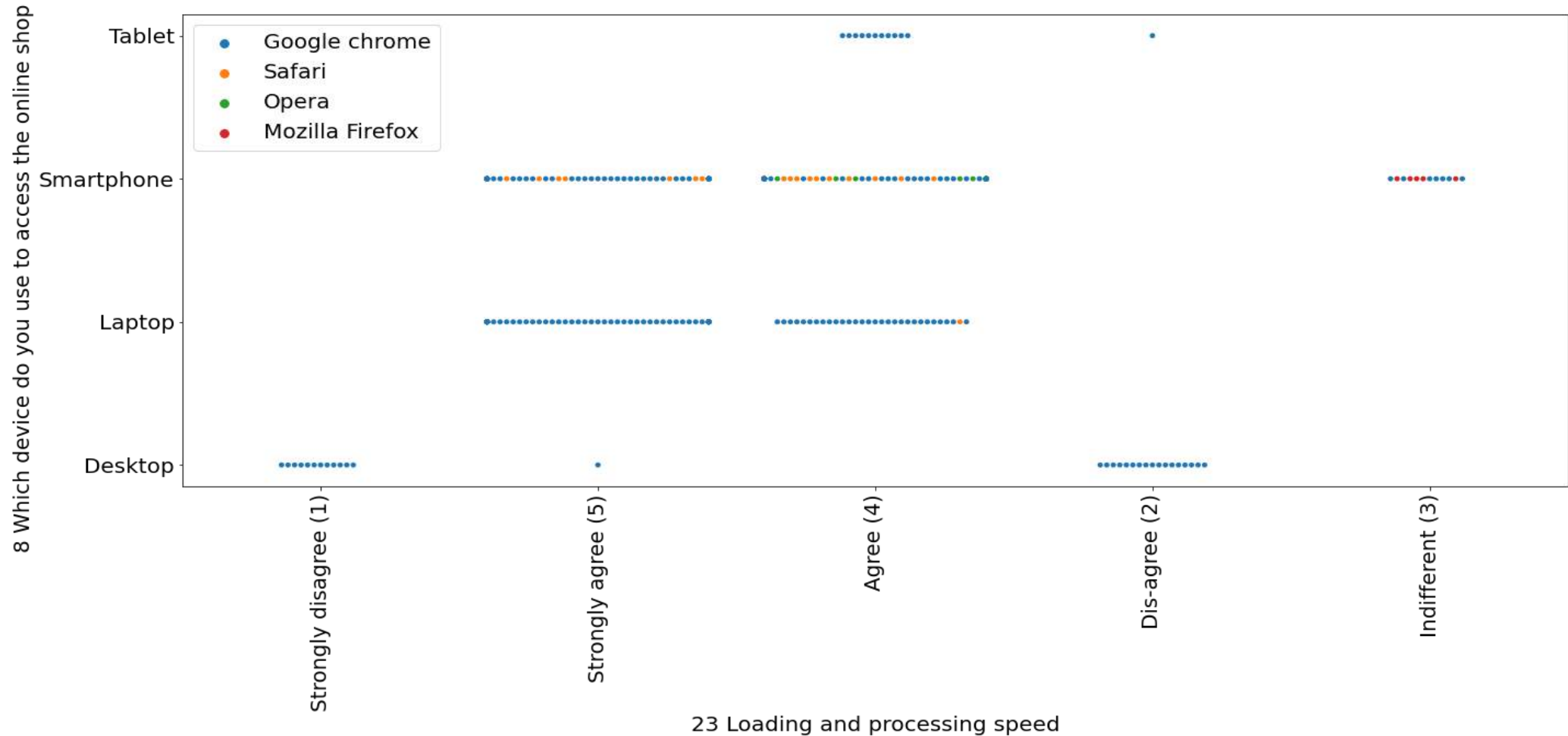
- This swamplot gives a better understanding about the relationship between loading/processing speed and devices being used

- A large proportion of people strongly agree to the fact that using smart phones /laptops for browsing the  e-stores are better when it comes to

- Loading and processing speed which we further classified using hue

- Few people opts for tablets and few customers thinks its indifferent to take into account the loadingness/processing speeds factor when it comes to online shopping

Reasons for maximum customers browsing on smartphones /laptops maybe:

- Easily accessible

- E stores also have very interactive application for smartphones which fulfils the needs of customers

- Saves time of the customers

# Plot describing how the customers landed on the store first time ever
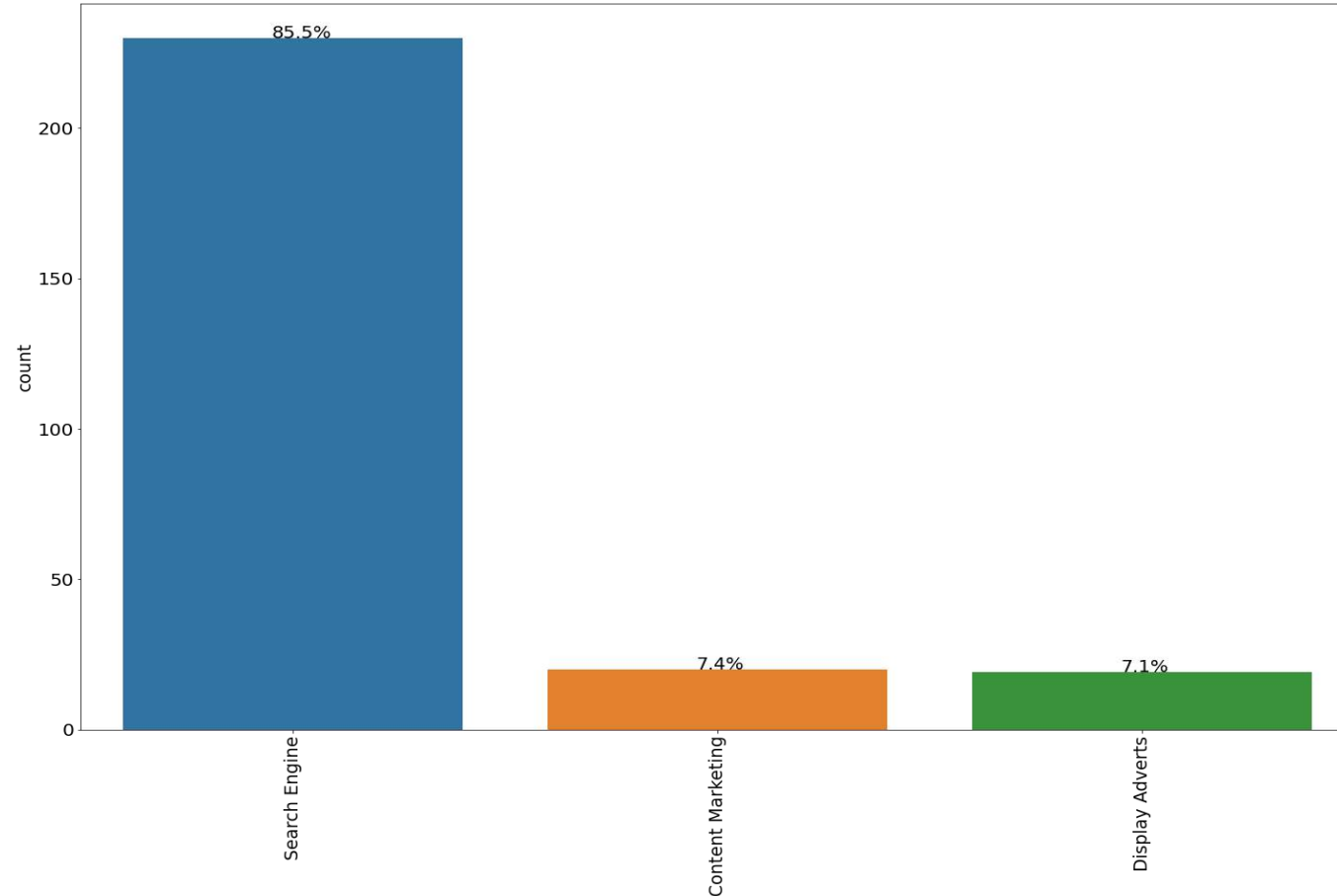


12 Which channel did you follow to arrive at your favorite online store for the first time?

- Its very clear from the analysis that majority of the customers landed on the store directly through the search engine
- But there is no denial fact that content marketing and display adverts had played a good role in impacting the rise in customers

Reasons maybe:

❖Though they came to know it through the ads/content marketing only but the first thing anyone will do is a google search so this maybe the reason where there is exceptional percentage in search engine category (85.5%) and 7.4 ,7.1 for content marketing and advert displays respectively

❖Or poor connect with the customers through the adverts and campaigns might also be the reason

❖Not that impactful in creating a sense of connect from customers perspective

# Measures that can be taken:

- Improvising the approach of marketing and building new strategies whose sole purpose is to create a sense of necessity in the customers mind

- Using SEOs ,SMOs for improving the rank of the store on search engines
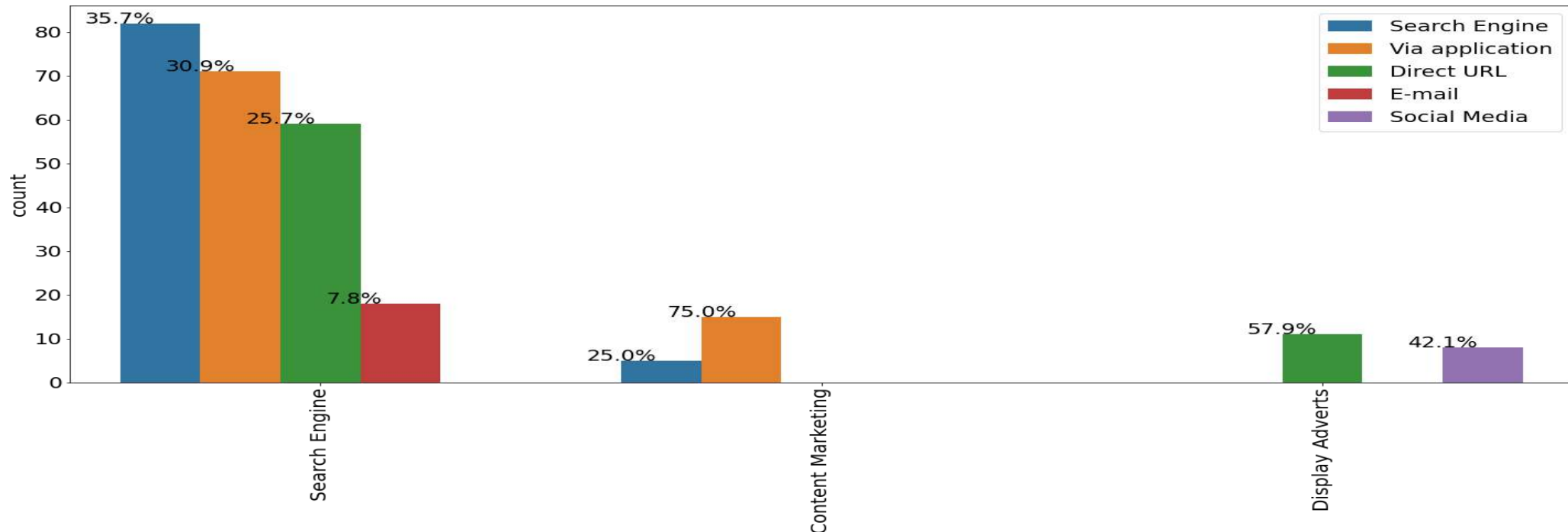


12 Which channel did you follow to arrive at your favorite online store for the first time?

Still after the first visit its easily noticeable from the further analysis that:

Search engine still remains the top choice in visiting the stores

Around 31% customers download the app from playstore/Appstore and continue browsing from there

Content marketing is still better than display adverts

Measure that can be taken:

As we know there is good amount of customers shifting to mobile applictions after the first search keeping up and updating ,optimizing the apps constantly so to retain them there

Simple adverts  containing the links such that there is an urge among the customers to browse/shop from the store

# Plotting a countplot to know about the average time required before a purchase is made
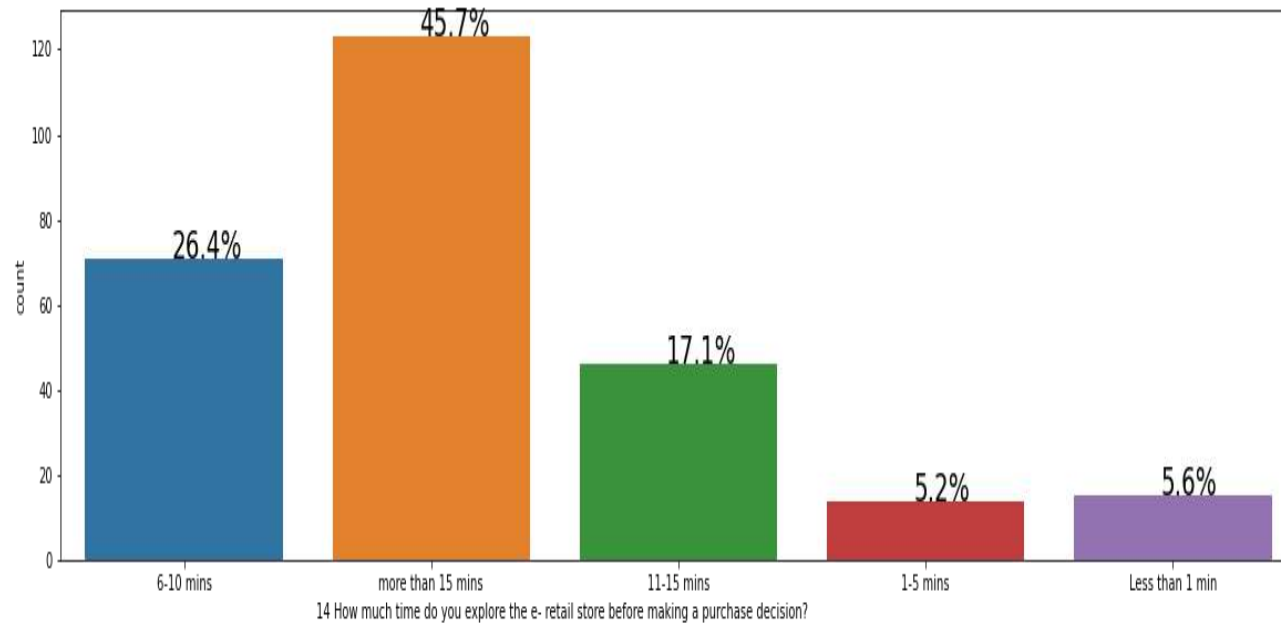
Around 45.7 percent of customers take more than 15 mins to make a purchase decision

Around 26.4 % of customers lie in category of 6-10 mins

Its interesting to note that around 5.6 percent of the customers don't even take a minute before reaching at a point of decision of purchase

Our main aim as  a e-store should be to drop those in more than 15 mins category to the 6-10 min category as its seems achievable and quiet logical

Dropping these decisions timings will help in preventing customers from abandoning the cart/bags as their may be shift in the mood of the customers or they might find better deal at competitors store

Providing best deals , using optimised timers which gives a sense of time limit in customers mind

Creating different modes Depending on the attention spans of various categories of customers

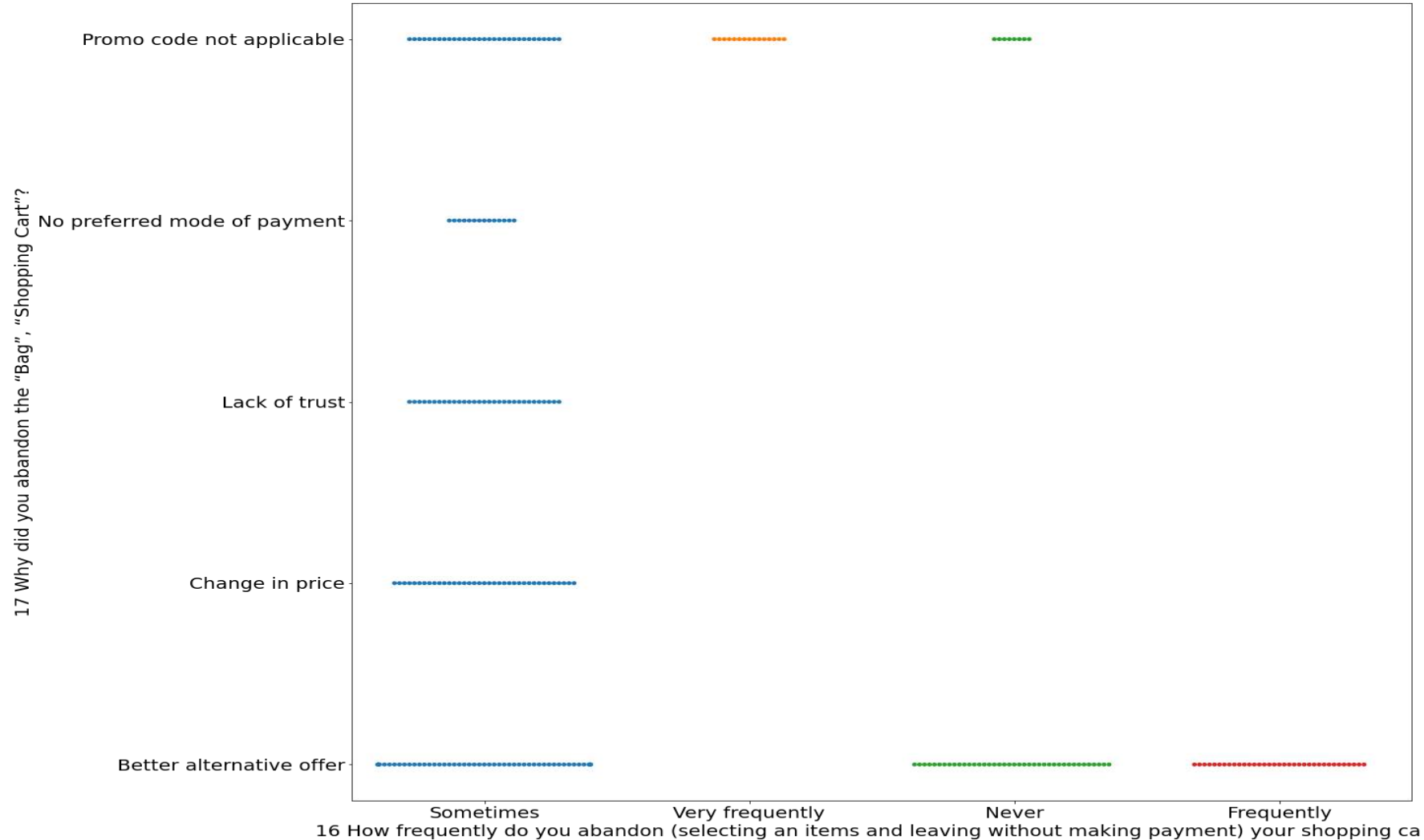# Further categorized on gender basis
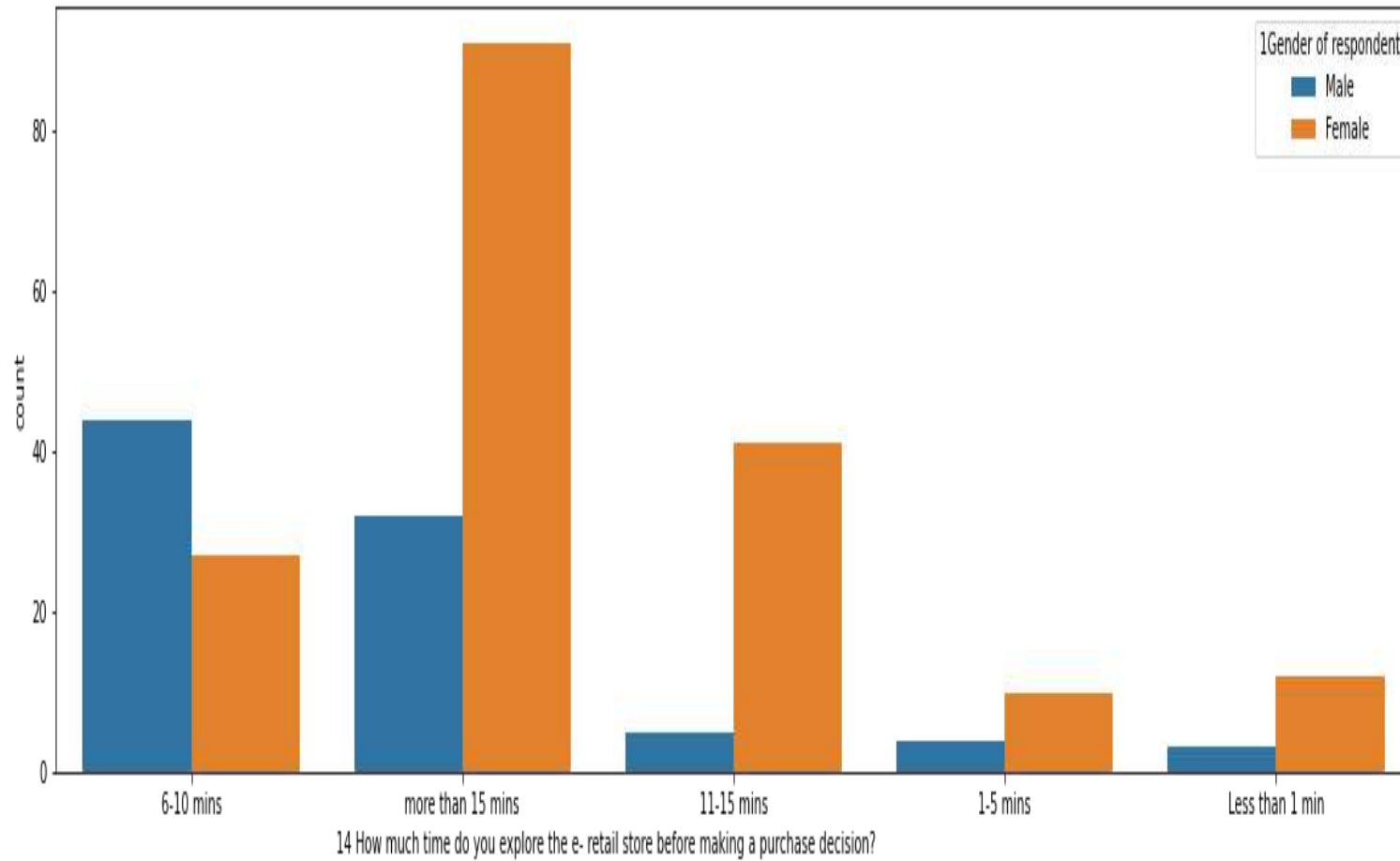
Women take comparatively more time to reach to the the decision as compared to men

Another plot describing decision time based on the age categories
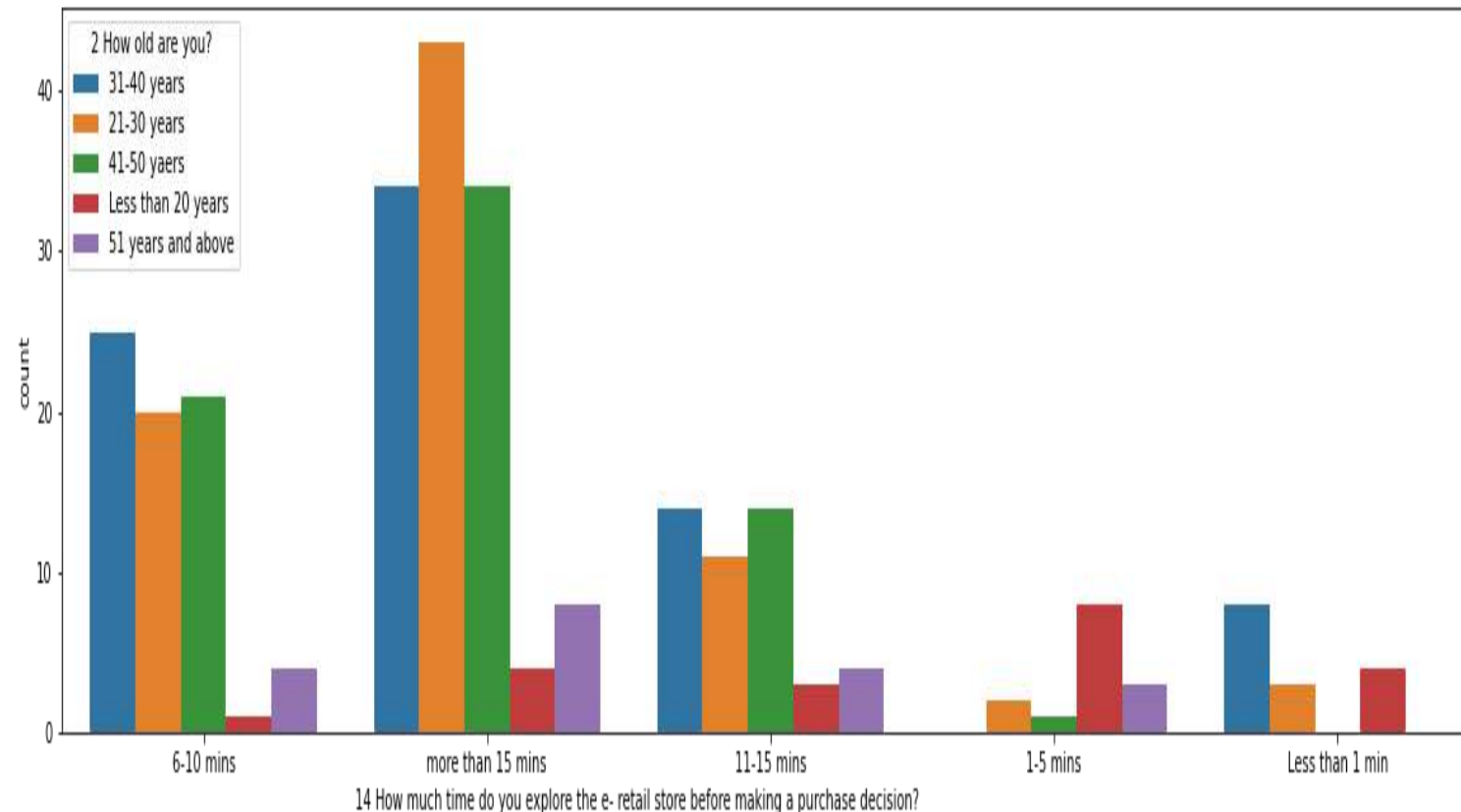
# Plot describing relationships between how frequently someone abandon the bag and the reason they abandon the bag

- It can be observed very easily that customers who frequently abandon the bag/cart is because the promo codes are not  applicable

- There is no unique pattern for those who lie in "sometime" category

  They are likely to abandon the bag in all cases

Measures that can be taken:

- Informing about the schemes/policies in a very clear way so that customers should not feel like being clickbated which may destroy the reputation of store

- Regular feedbacks ,followups for those who have items in the cart and they aint making the purchase

- Rewarding those who very rarely or never abandons the bag/cart giving them a sense of importance which further will help in retaining those customers

# Importance of complete detailed information of listed seller and about the product in a purchase decision



20 Complete information on listed seller and product being offered is important for purchase decision.
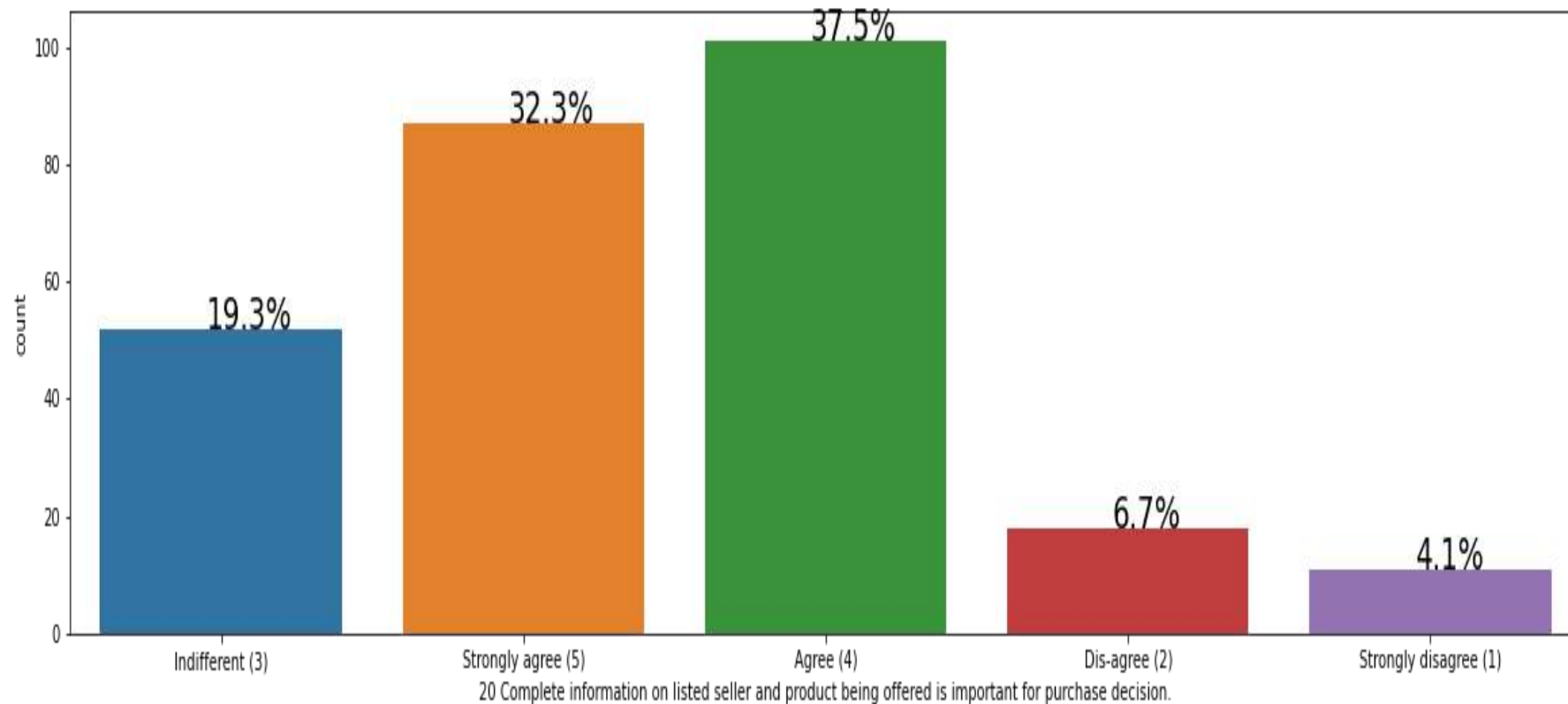
- It can be seen very clearly that around 37.5 % strongly agree and around 32.3% agree to the fact that yes detailed information about the seller and product is a necessary condition influencing purchase decision
- Its quiet logical too that giving complete information helps in building trust with the customers
- Overcoming the fear of getting scammed /online frauds
- A healthy relationship with customers means easy referrals to the new customers a they are also taken into trust by the satisfied customers

# Relationship between user friendly interface and ease of navigation in website
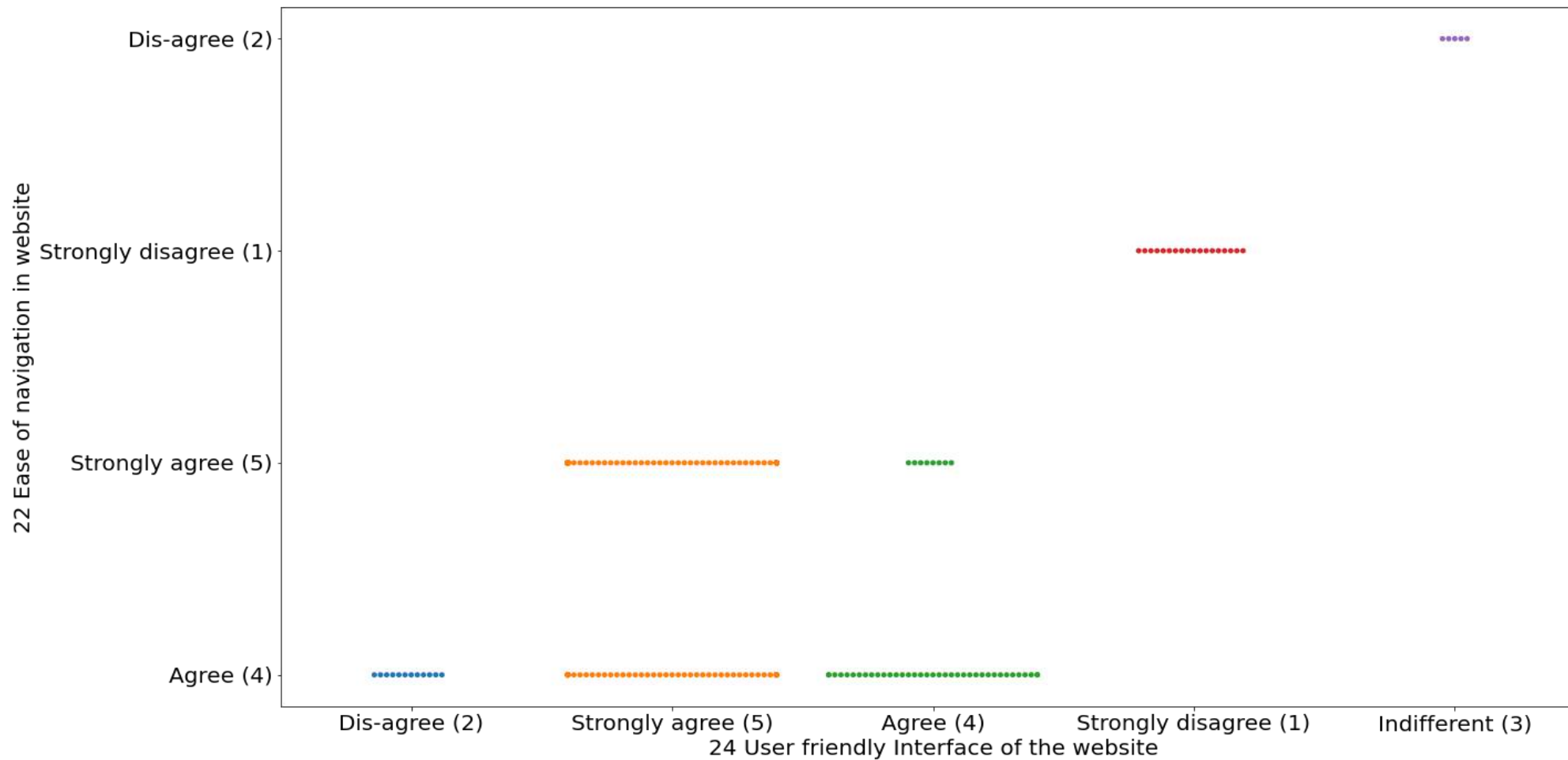
- Similar kind of patterns exist between the the 2 features and can be depicted by the plot in the previous slide
- Wide strips in the plots in agree and strongly agree column conclude the fact that yes as the user friendliness increases there is a ease of navigation
- Improving the experience of customers giving us the better results
- It can be also viewed from the machine learning modelling perspective  that as there is high correlation between the two we can drop one of the columns.

# Plot to check the dependency of " trust that online store will make it delivery on time" on   convenient payments methods
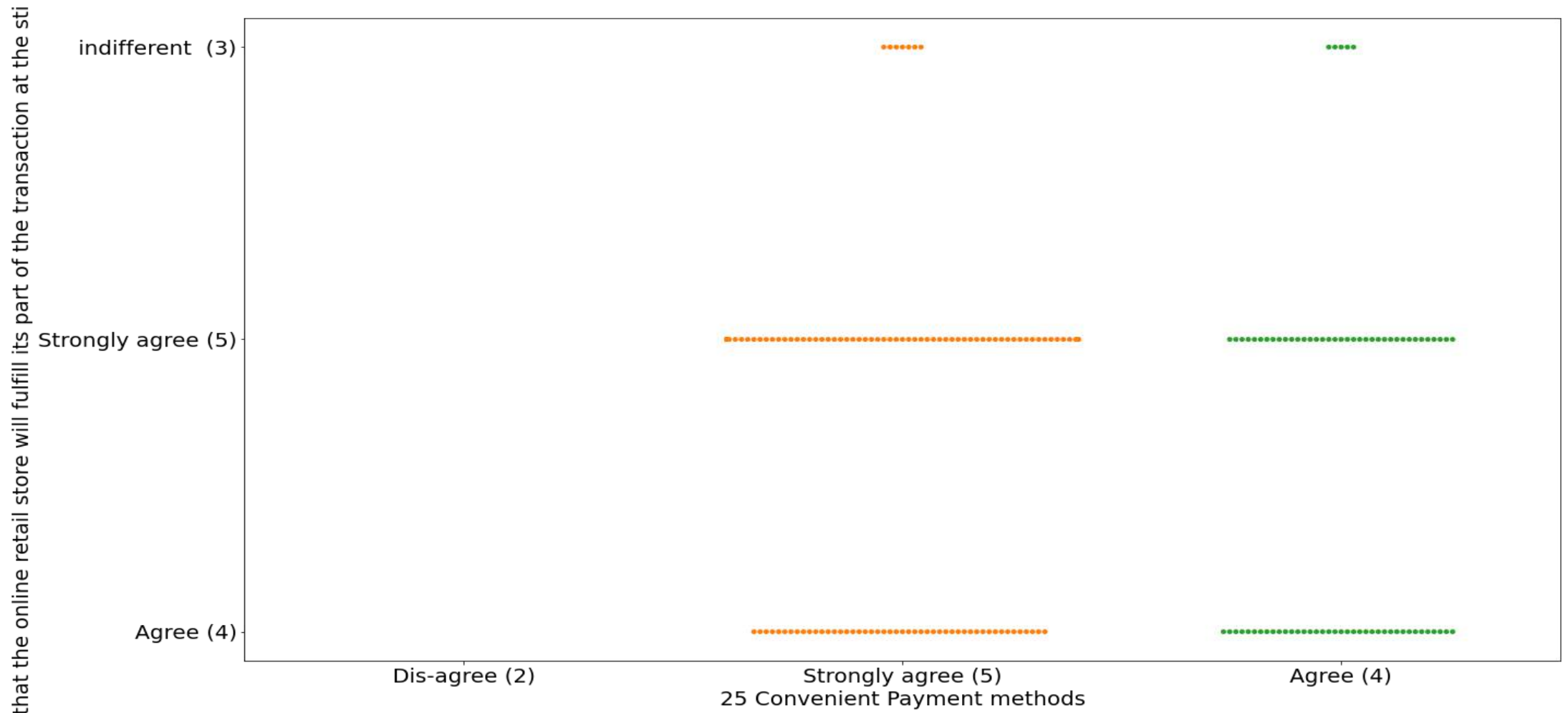
- We can see there is a genuine and very logical kind of response between both the features
- Building trust and convenient mode of payments goes hands in hands
- Wide strip on both agree and strongly agree columns confirms it that yes both are equally important features for purchase decision
- If the payment is smooth without much hassle and that too with proper feedback one wont doubt the authenticity of the store and automatically there is going to be the trust with the store
- Trust is very important feature when it comes to e-stores so taking that in consideration will be our top most priority

# Assistance /support as a important  feature hued with multiple channels for assistance /support
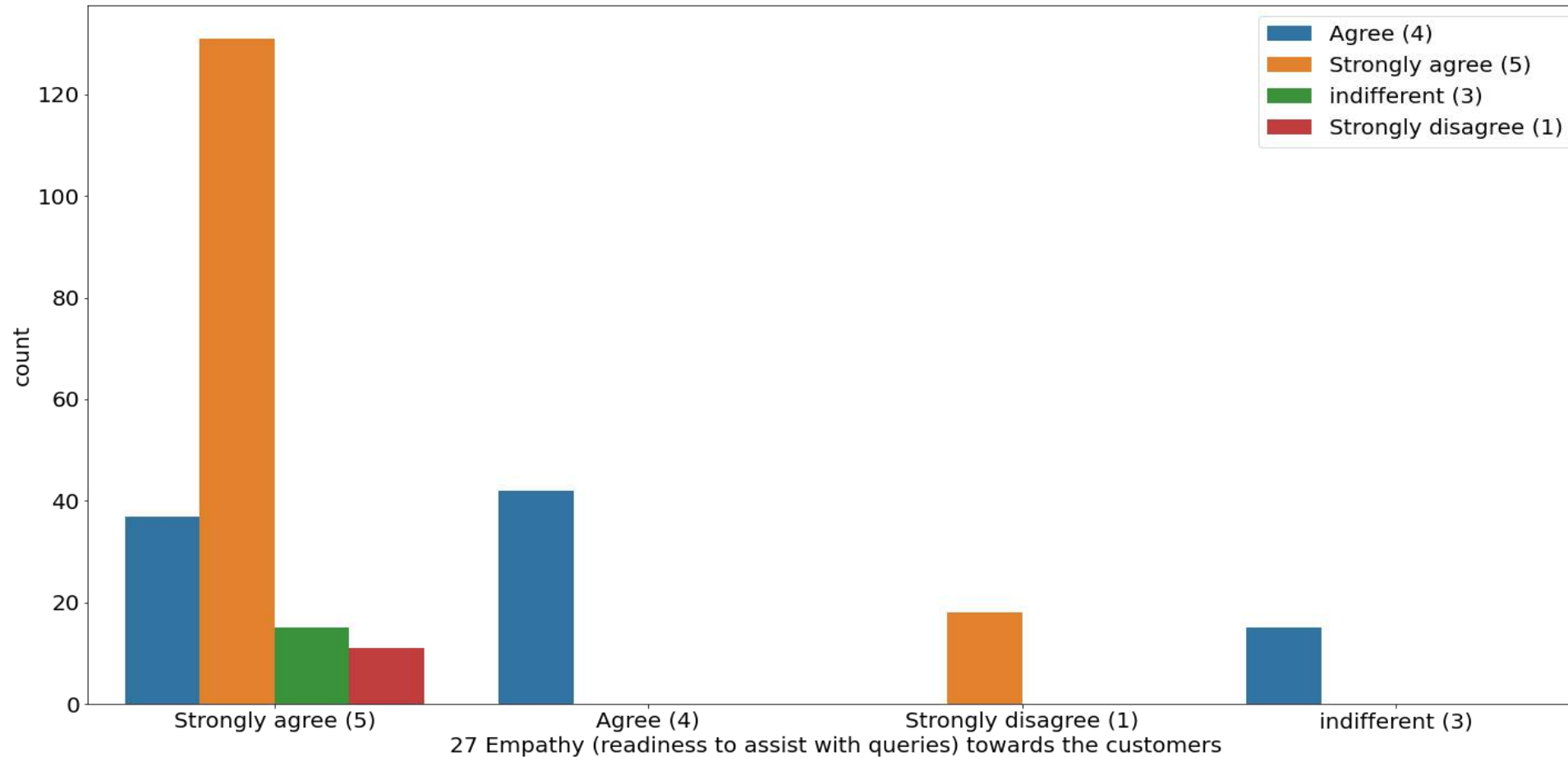
- Majority of the customers are in the favour of empathy towards the customers and having multiple channels to support them with
- Makes it easier for the customer to interact with the assistance team giving a sense of security and satisfraction among the customers
- Help with building trust
- Better support will also help in retaining the customers
- Improves the experience

# Take on importance of privacy on influencing the decisions
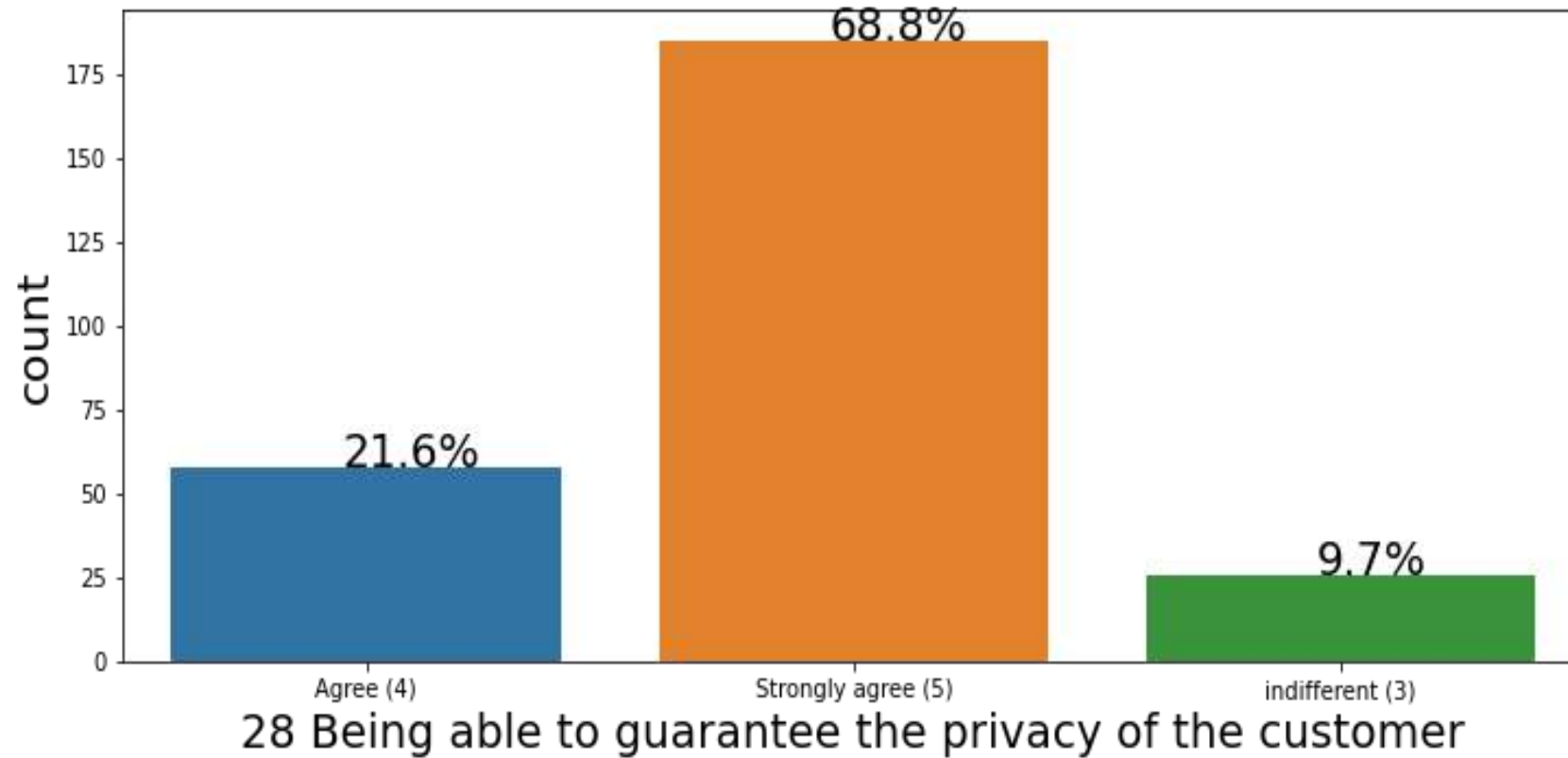
- There is no denying fact and also concluded by the plot that privacy of the customer data(personal and bank details) should be among one of the top priorities when features are considered
- Around 70% of the customers agreed to this in the feedback
- Again it helps in building a sense of safety and protection from the frauds /scams
- Data breach is a serious issue and can lead to loss of potential as well as retained customers and is also a threat to the reputation of the company

# Relationship between "having monetary benefits and discounts" with enjoyment derived from shopping online
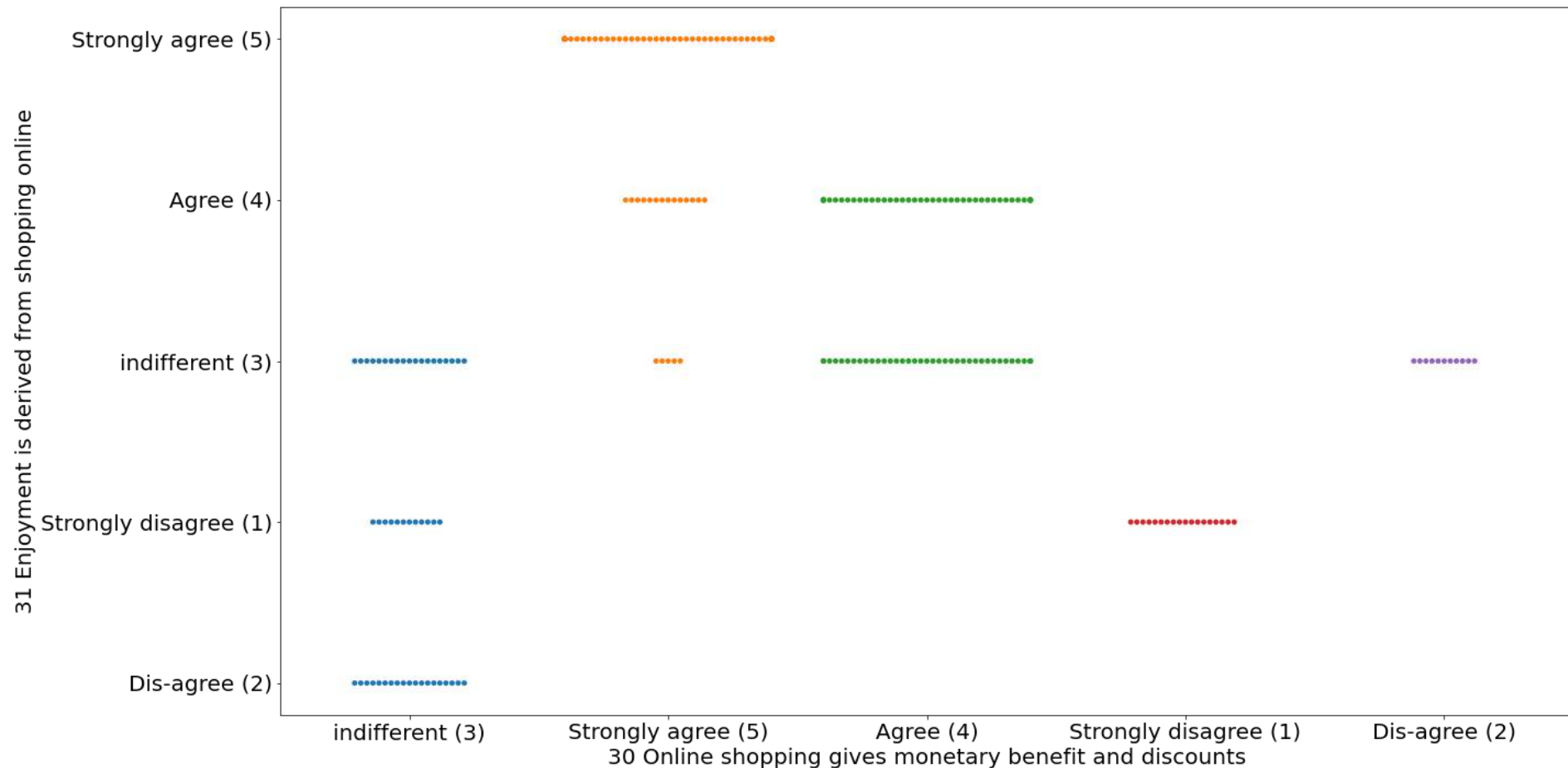
- Kind of decrease in number of customers can be seen from strongly agree to disagree which also states that a quiet good number of customers agree or strongly agree to the fact that there are monetary savings and discounts and enjoyment is derived from shopping online
- Enjoyment is highly correlated with the monetary savings
- Follows a similar trend In both cases

# Considering loyalty program as a benefit of online shopping
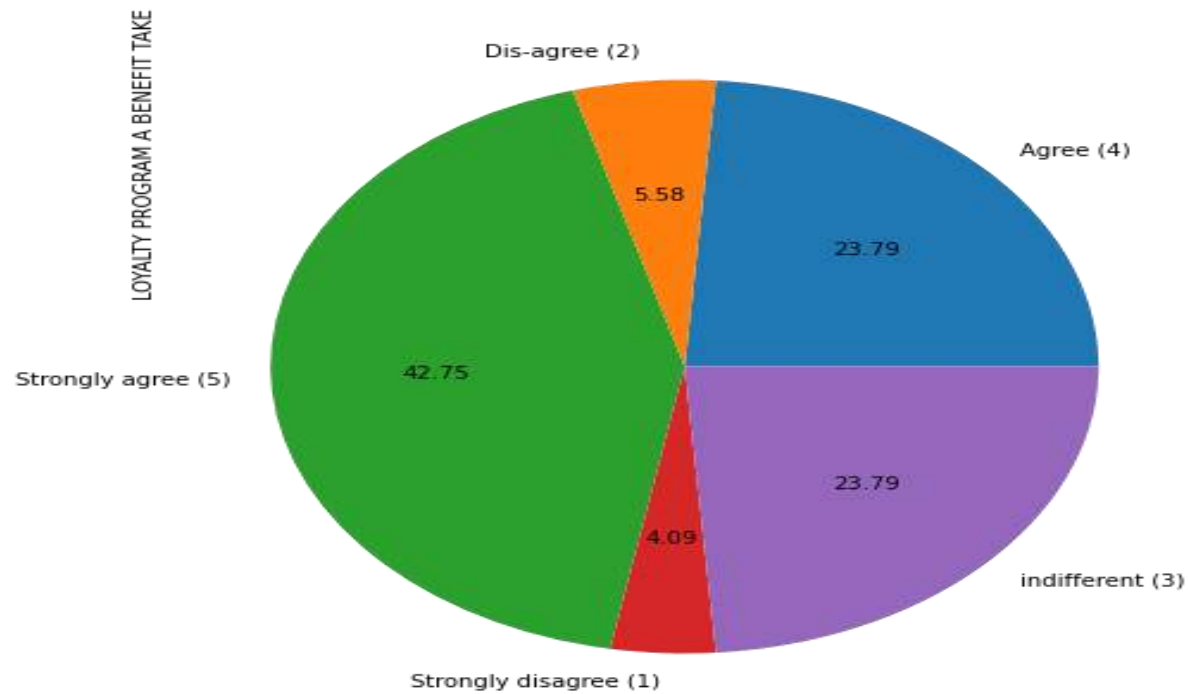
- Around 42.75% and 23.75% strongly agree and agree respectively with the statement that Loyalty program is a benefit of online shopping

- Its also interesting to note that around 23.79 % of customers feel that's its indifferent and have nothing to do with considering about loyalty programs when it comes to online shopping

- One of the reason for such a high percentage of people opting for "strongly agree " and "agree " is that there is a sense of inclusion and importance developed among the customers which is very good sign

- More inputs should be taken from those who feel its indifferent as chances are high that 23.79% of customers might be a having a strong perspective which might help us in understanding the behaviour of the customers

# Satisfaction and trust relationship
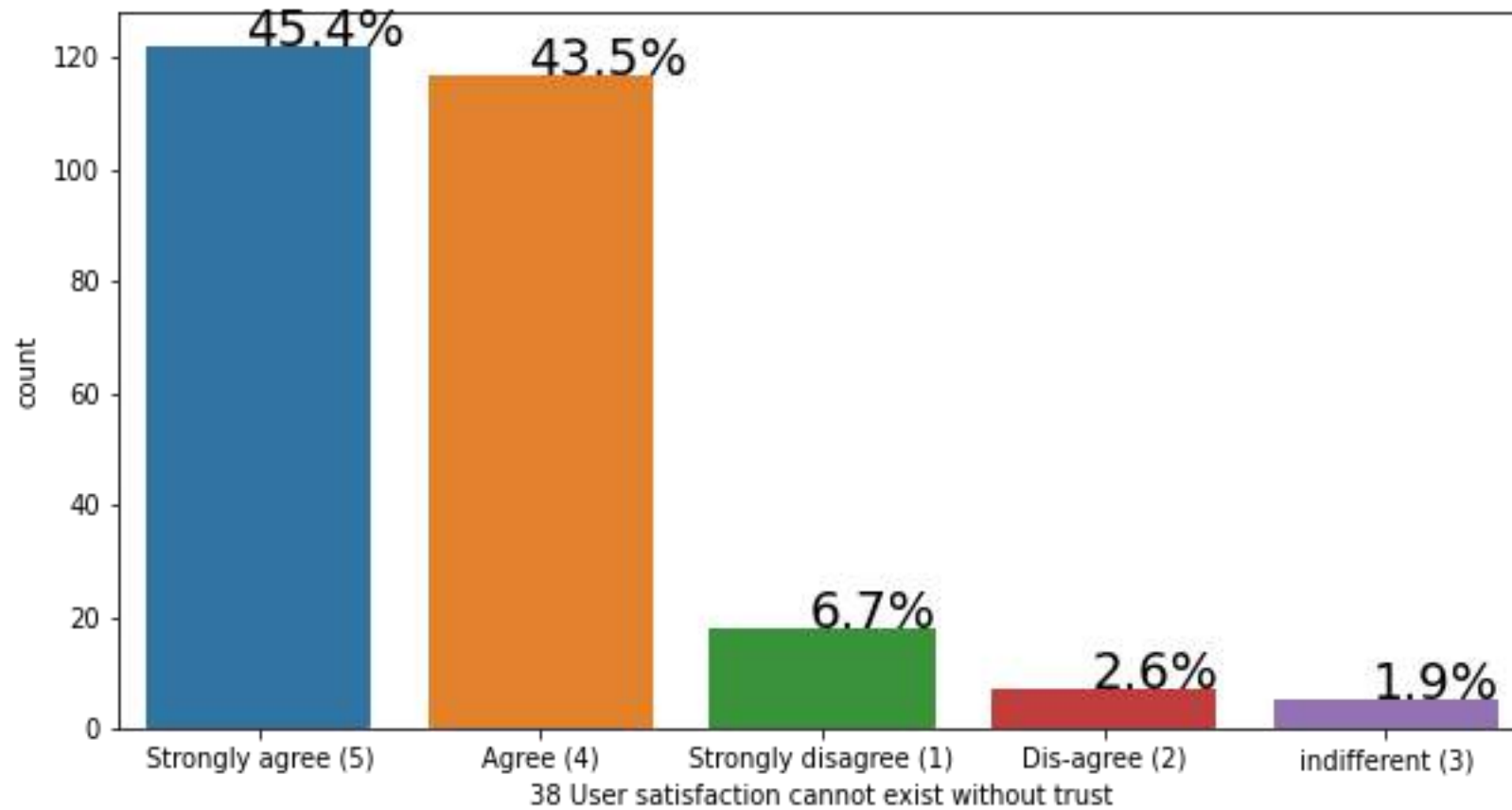
- This might be one of the most obvious as well as one of the most important relationship to be taken into consideration when it comes to selecting the features

- Huge amount of positive responses in this category tells that trust is what most of the customers seeks online

- As businesses work on the principle of customer satisfaction this relationship and analysis cant be ignored

# Patronizing the online stores take
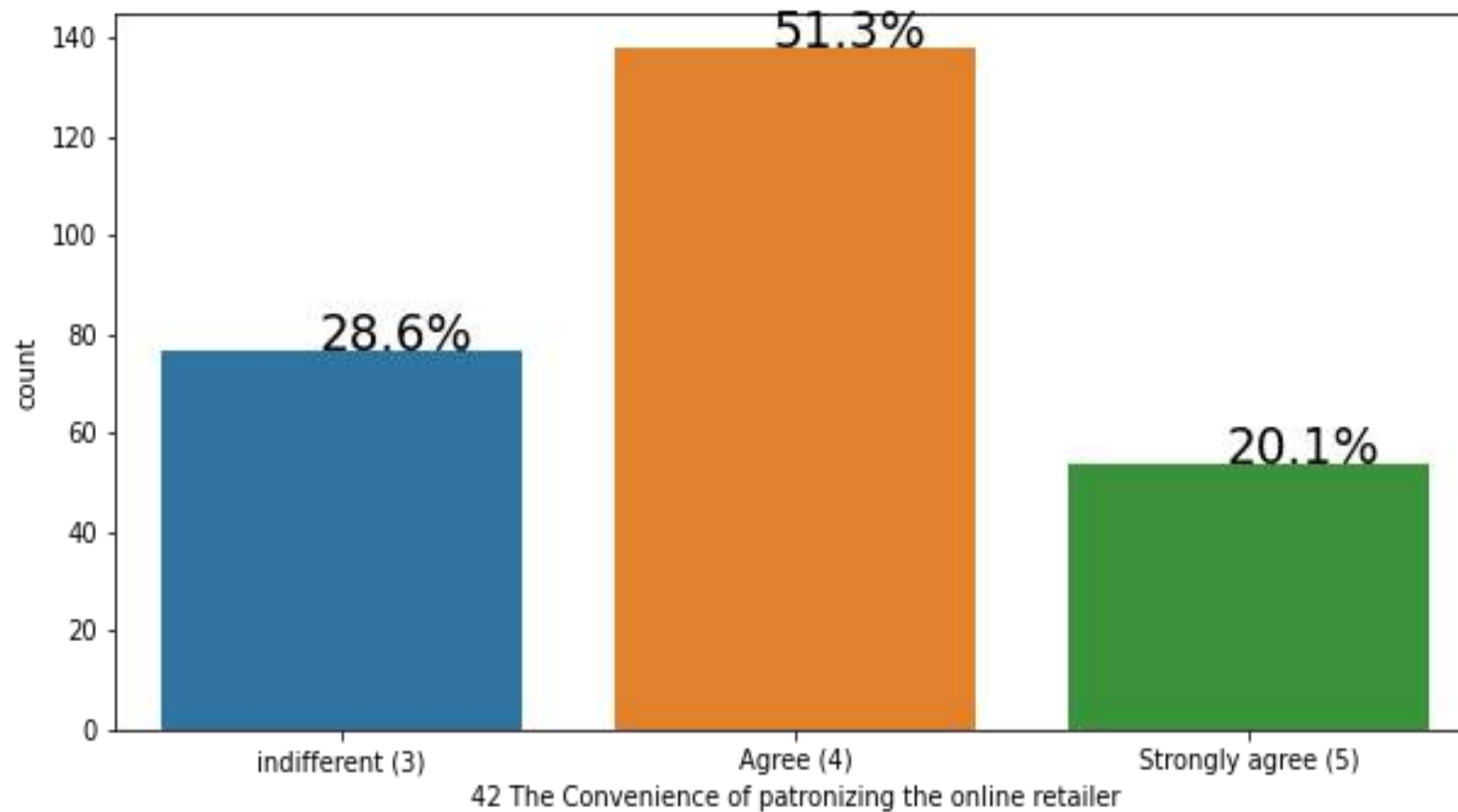
- Around 51.3 % of the customers agree to the fact that yes there is a sense of patriotism related with the e-store
- Around 28.6 % finds it unrelated and to the surprise 20.1% of the customers strongly agree to this
- From the trend its quiet obvious that e-stores can easily take the advantage of nationalism and patriotic feeling attached with the customers
- That's why we also observer that (in case of india) all the e-stores come up with some kind of offers/discounts/special offers to engage the customers
- Alibaba a huge success in china got the advantage of patriotism though it was backed by strong government policies too but still cant deny the fact that people intentionally support Alibaba because it gives them the sense of connect and feels like supporting what we say home built rather than any outsider

# Online shopping can be adventurous too
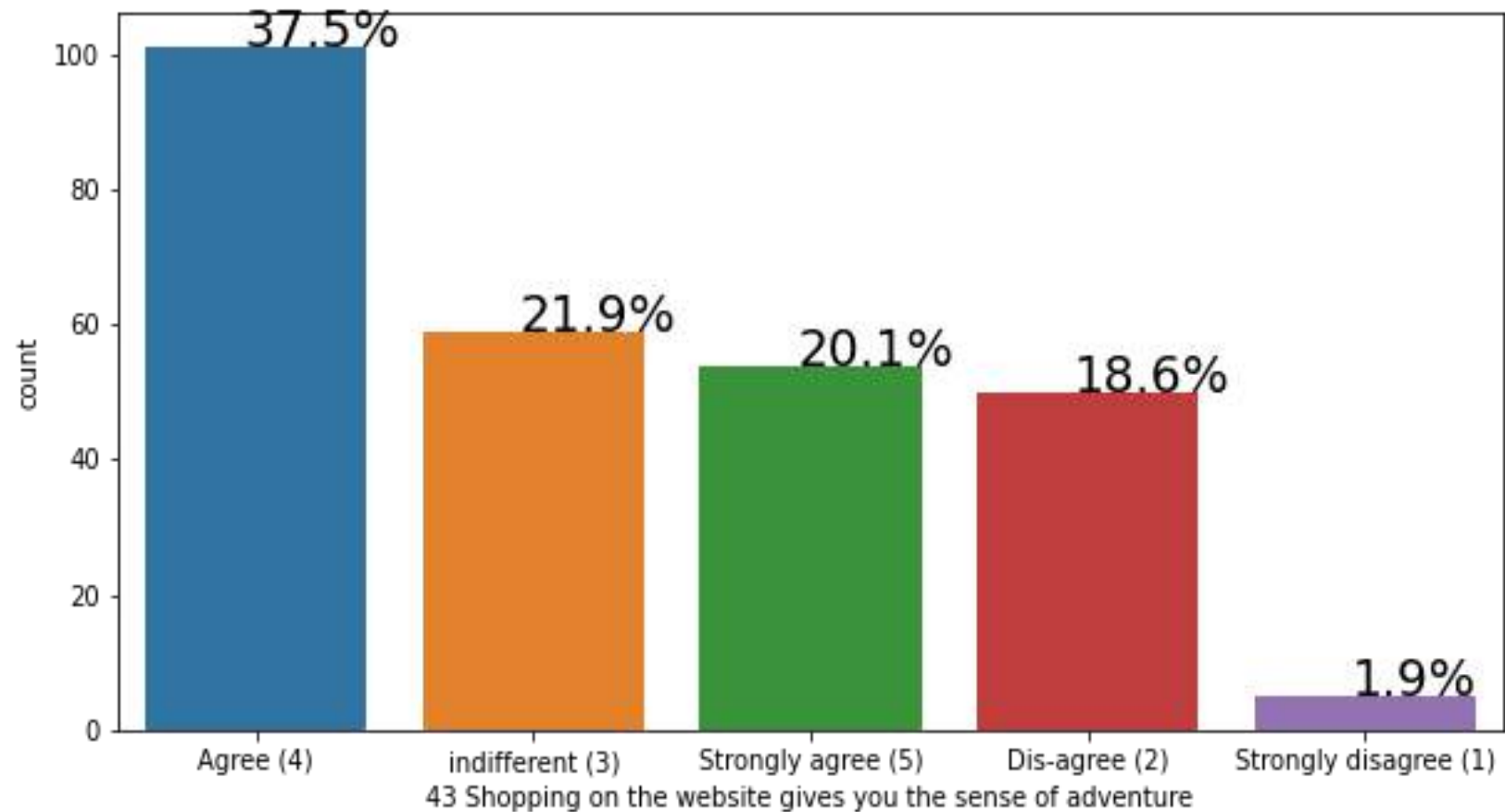
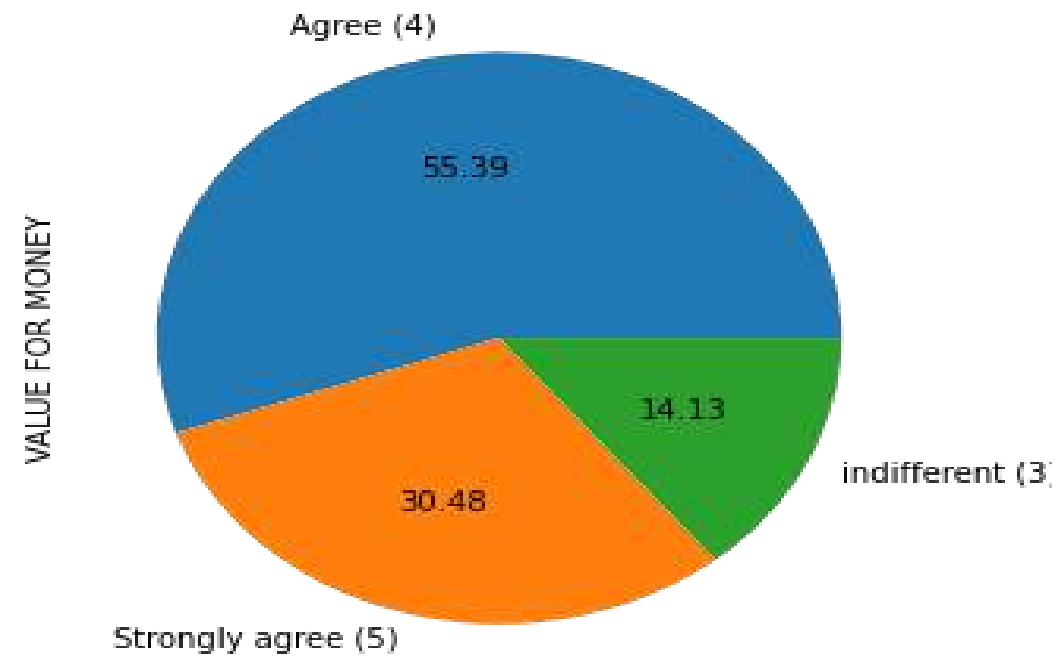# Plot describing sense of fulfilment of certain roles
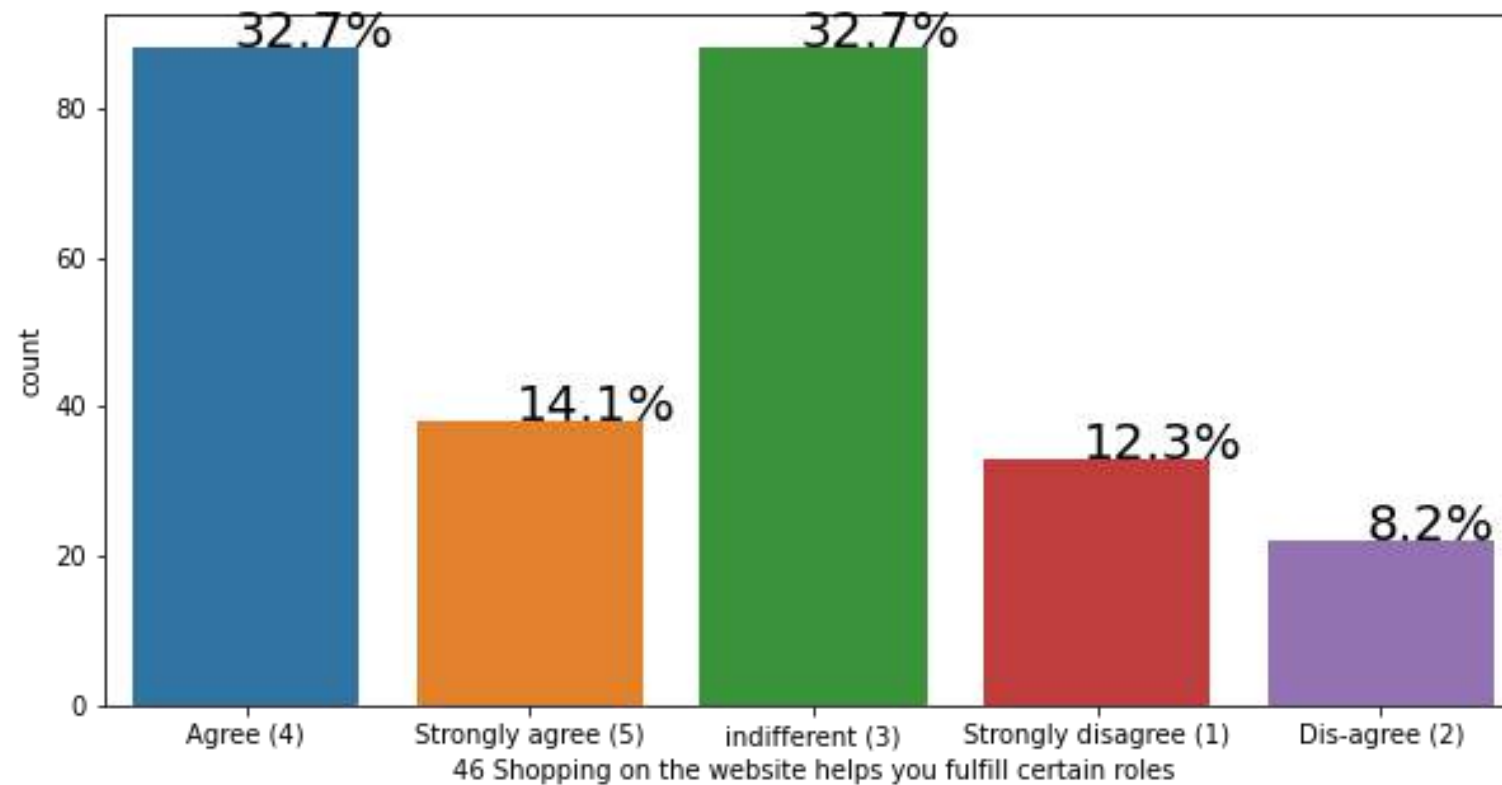
Around 32.7% of customers believe that there is kind of fulfilment of roles when it comes to online shopping
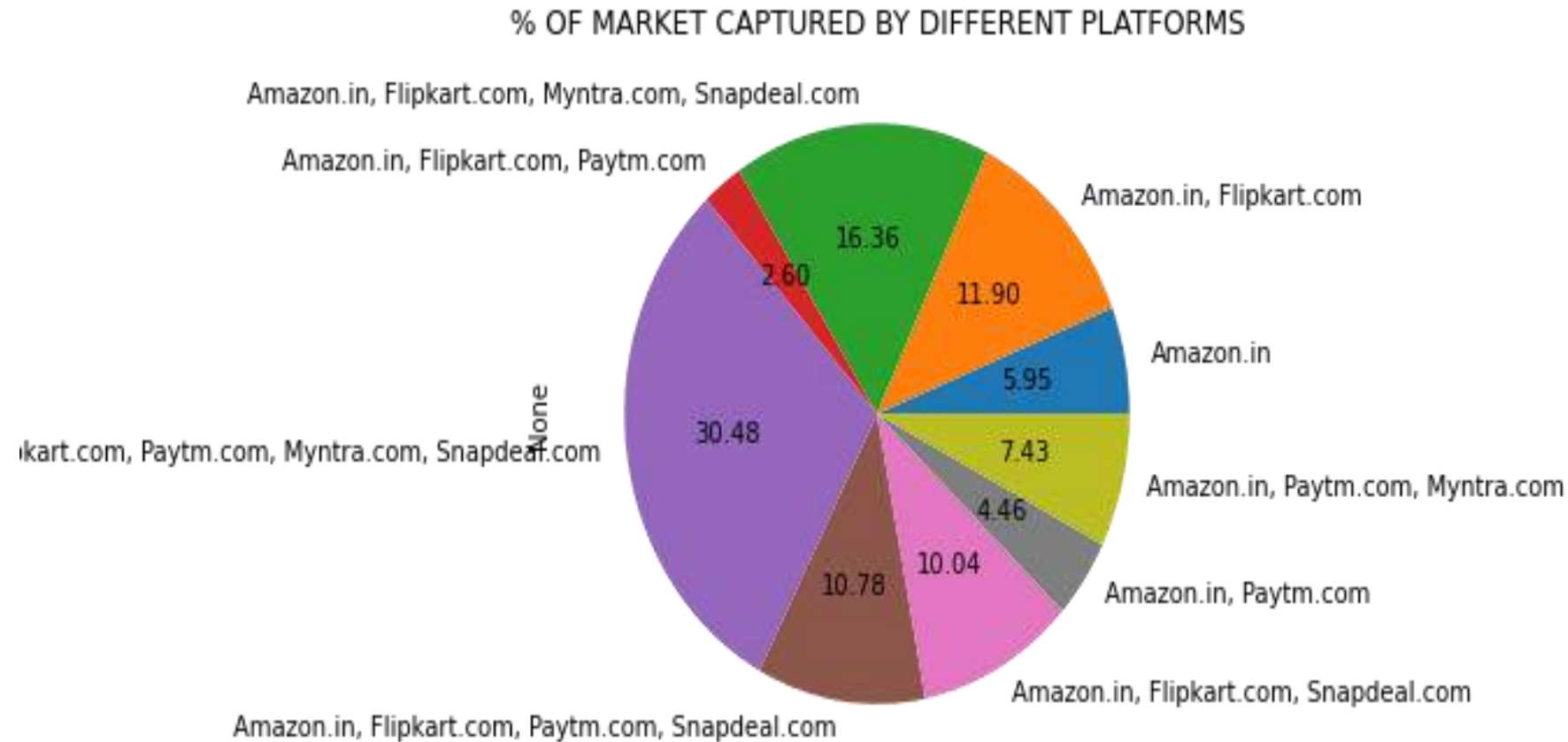
14 % strongly agree to the point

Its also surprising to see that 32.7 % don't feel any relationship with the leading point
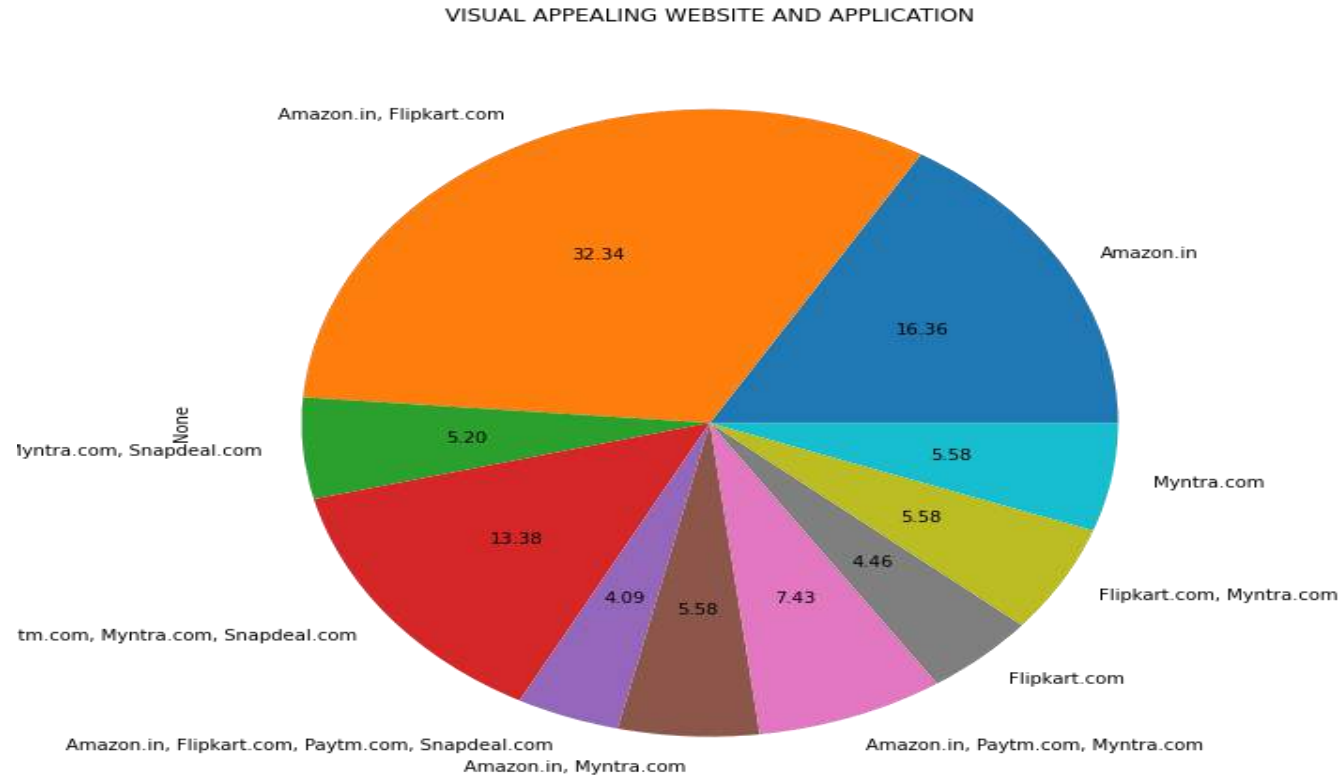
# Value for money as a feature

- Quiet a large portion agree or strongly agree that there is a value for money when online shopping is considered

- With even around 85% of customers voting in favour of value for money there is no doubt that this is one of the important features of online shopping and should be considered when we take this into consideration

- Value for money is a key factor in deciding which store holds the maximum cap in market

- Almost all the customers will look for this and compare with the other stores if they are getting a better deal then arrive at a decision

# Market share of different e-stores as reviewed by the customers



% OF MARKET CAPTURED BY DIFFERENT PLATFORMS

# Visual appealing website and application



VISUAL APPEALING WEBSITE AND APPLICATION

Amazon.in, Flipkart.com — 32.34
Amazon.in — 16.36
Myntra.com — 5.58
Flipkart.com, Myntra.com — 5.58
Flipkart.com — 4.46
Amazon.in, Paytm.com, Myntra.com — 7.43
Amazon.in, Myntra.com — 5.58
Amazon.in, Flipkart.com, Paytm.com, Snapdeal.com — 4.09
tm.com, Myntra.com, Snapdeal.com — 13.38
1yntra.com, Snapdeal.com — 5.20

AROUND 16% OF PEOPLE THINKS THAT AMAZON DOES IT BETTER WHEN IT COMES TO VISAUL APPEALING WEBSITE THEN COMES MYNTRA WHO IS FURTHER FOLLOWED BY FLIPKART

AND ALMOST 32.34% OF POPULATION BELIEIVES THAT AMAZON AND FLIPAKRT ARE BOTH GOOD WHEN IT COMES TO VISYAL APPEALING WEBSITE AND APPLICATION

There are high instances that if the website is visually appealing it tends to attract more customers

Plays a high role in the asthetics region

Specially the age groups less than 20 and 21-30 agrees are more influenced by the asthetics

# Quickiness to complete purchase



Quickness to complete purchase
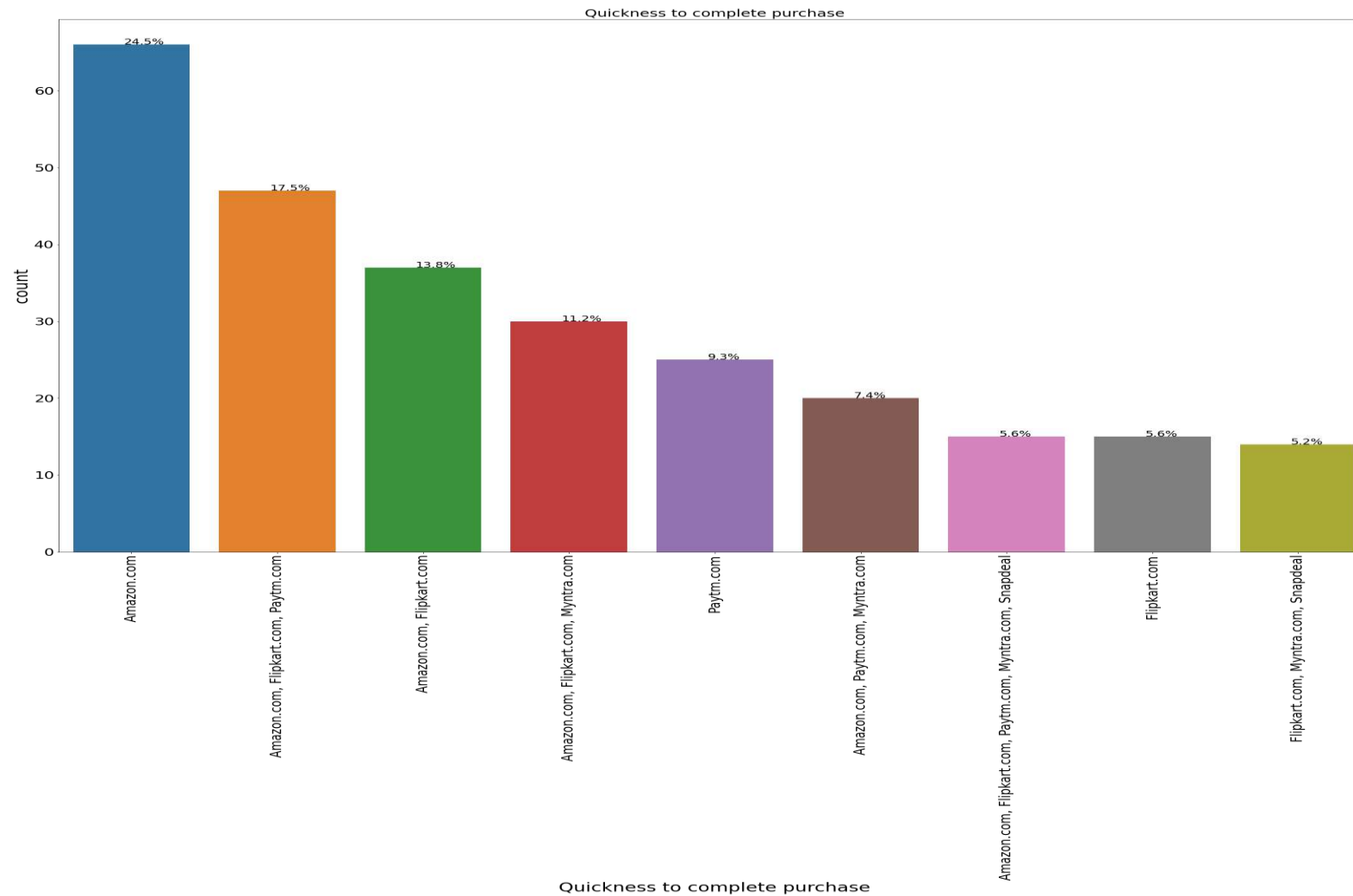
- This is very important feature which decides whether the customer will make an purchase or abandon the bag/cart

- There are high chances that more complexity in payment processes will give time to customers to rethink about the purchase decision which migh sometimes lead to loss of customers

- As we can see from the plot amazon does extremely well when its comes to readiness/Quickiness of  complete purchase channel with around 24.5% of customers voting in favour of amazon

# Plot describing relationship between "perceived trustworthiness" and "availability of several payment option"
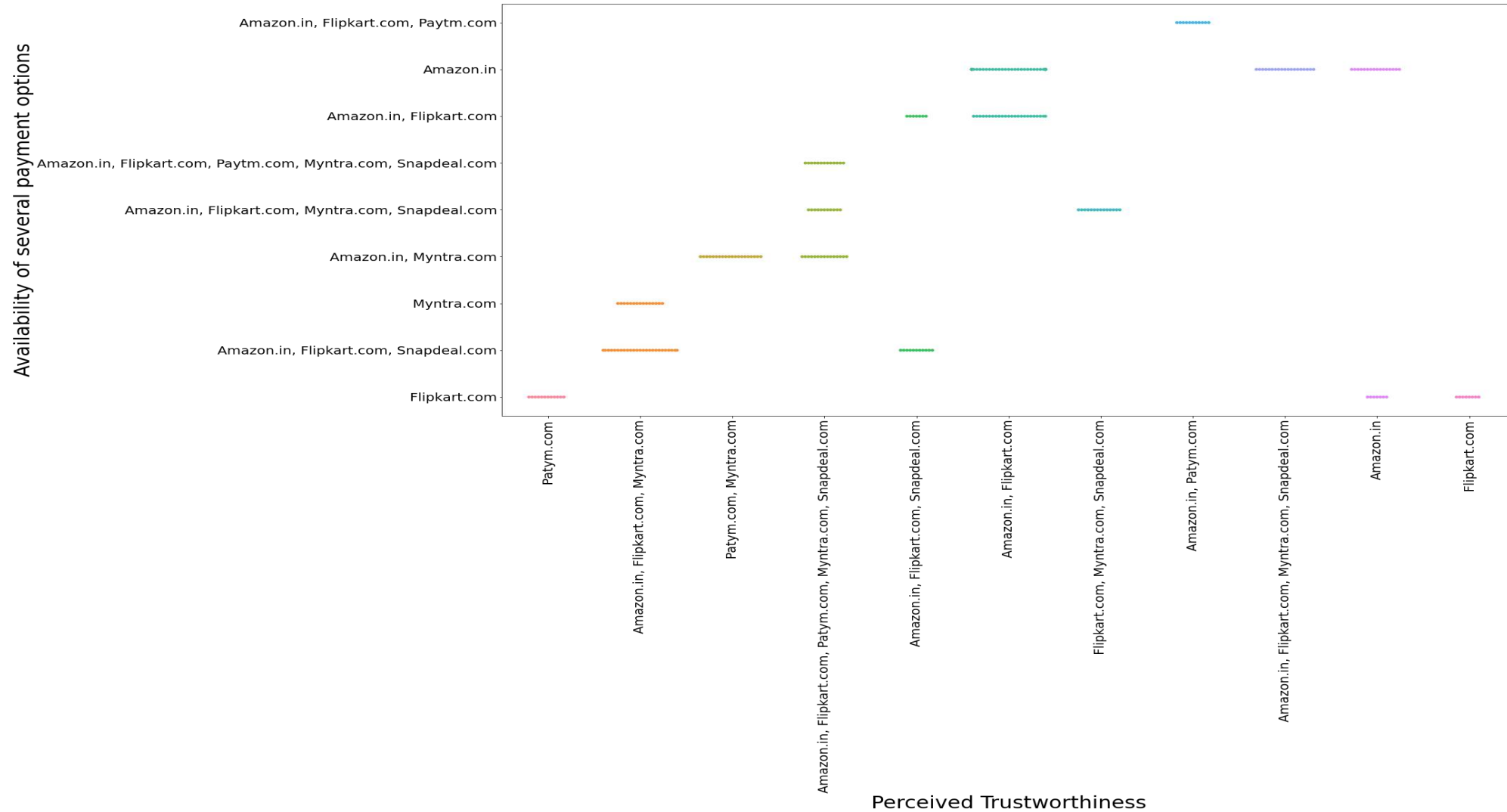
- AMAZON ALONE AND AMAZON AND FLIPKART BOTH COMBINED TOGETHER HAVE ALMOST SAME LEVELS OF PERCIEVED TRUSTORTHINESS WITH INCREASE OF AVAILABILITY OF PAYMENT OPTION

# Longer time in displaying graphics and photos
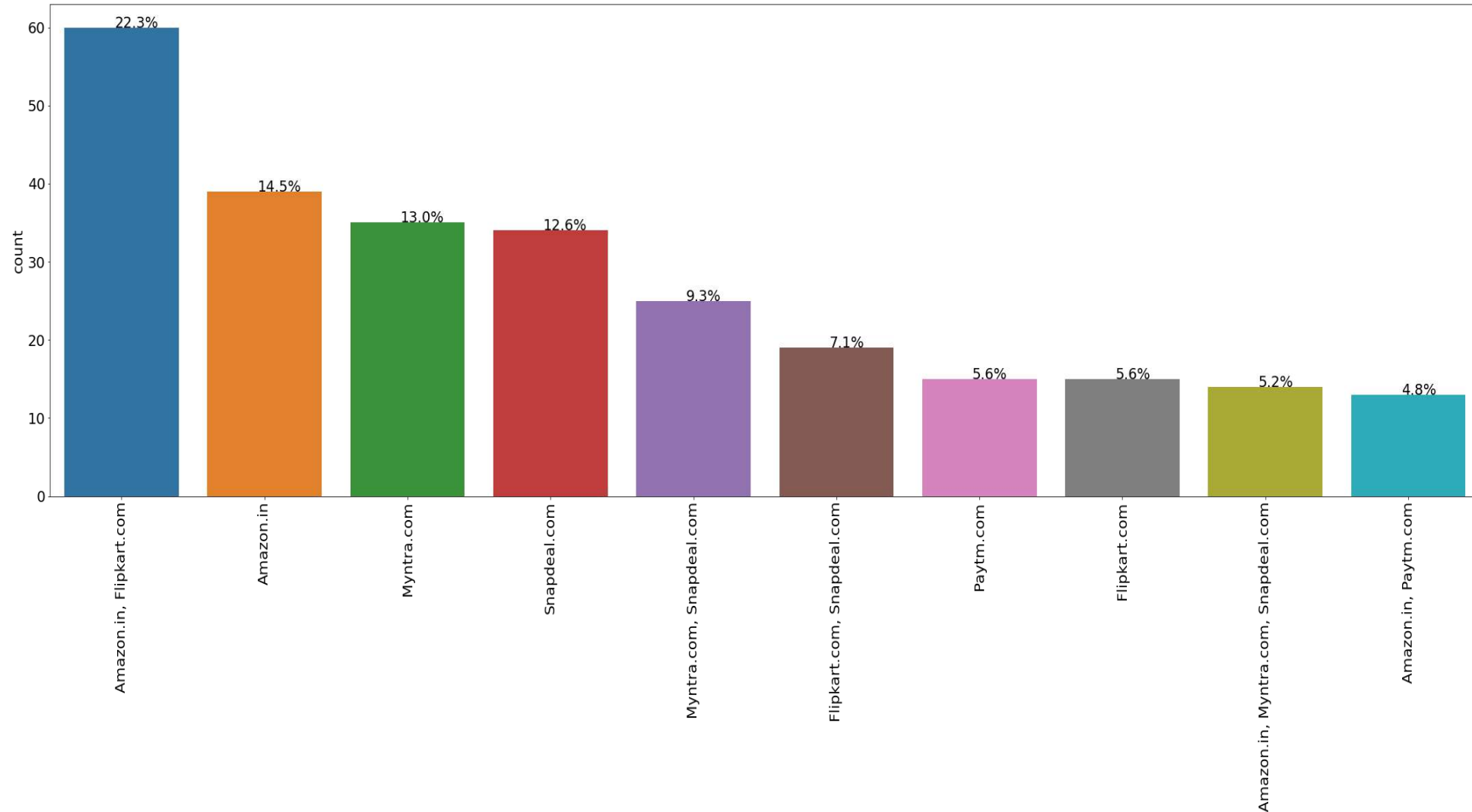


Longer time in displaying graphics and photos (promotion, sales period)

- This feature is a point of concern for most e-stores
- They might have the best deals ,better products ,wide range of products but if the graphics and photos does not load chances are high customers are going to scroll it without even checking it
- Can also lead to negative opinions about the store if this problem persists
- From the plot its quiet clear that around 14.5% ,13.0%  and 12.6% customers respectively are voting for amazon , snapdeal and  myntra

  which is a significant amount of customers and cant be ignored

# Longer delivery periods plot
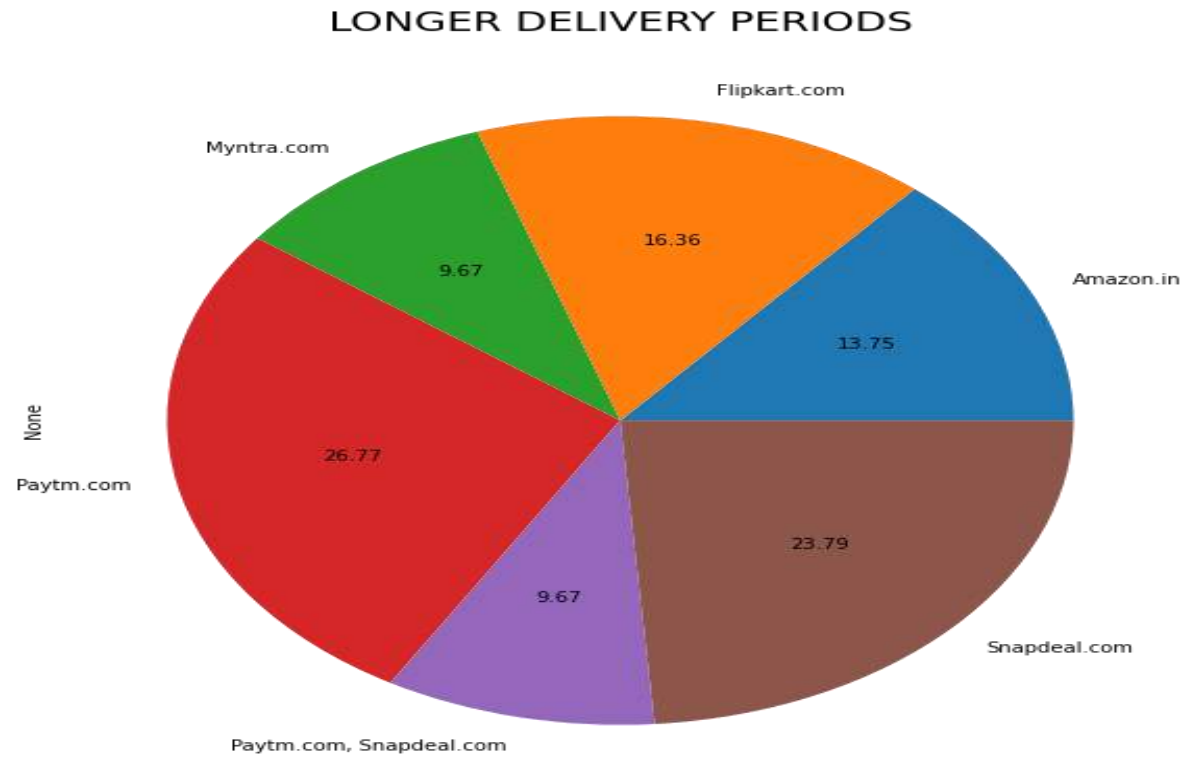
- Its one of the most prominent feature which almost all the e-store try to achieve and a deliver at the same time
- Customer satisfraction is highly related to time taken in delivery of the products i.e shorter delivery periods means higher satisfraction for customers
- Every e-store optimizes its logistics in a way such that they can achieve the best possible delivery periods gaining the trust of customers and increasing the credibility of store
- Though its clear from the plot that 26.7 % voted for paytm and 23.79% voted for snapdeal which is a considerable amount of customers and cant be ignored
- Affecting the sales of these stores

# Most recommended Indian e-stores



Which of the Indian online retailer would you recommend to a friend?
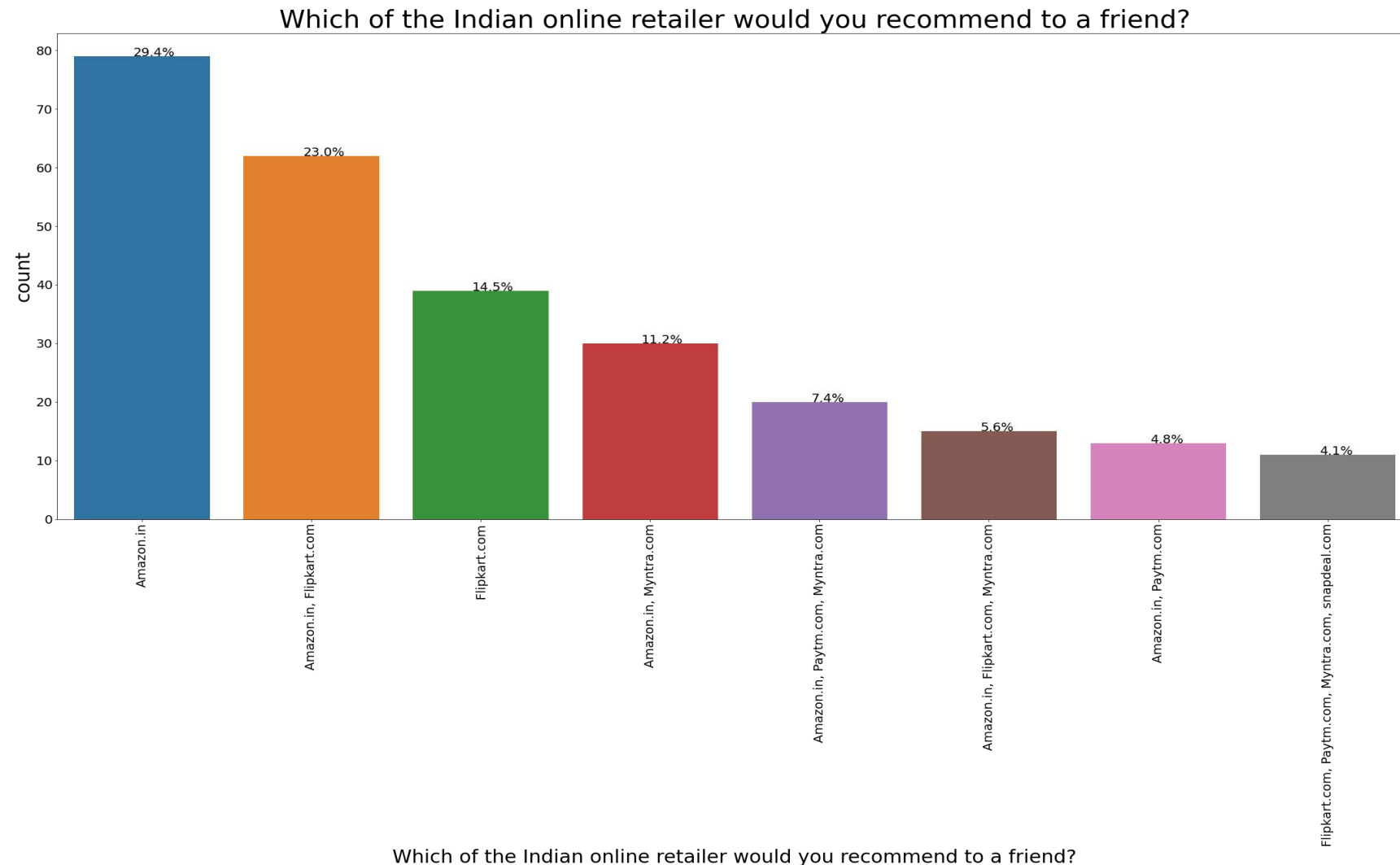
- We are assuming this as our target feature for modelling a ML model

 Reason for choosing this:

- A customer will only recommend/suggest the store which he/she finds best , credible, trustworthy and value for money

- The customer is retained as well as he/she is activating new customers which is the best thing for any store

- WOM(word of mouth) marketing is one of the most powerful forms of advertising as 88% of consumers trust their friends' recommendations over traditional media.

- Companies can encourage WOM marketing by exceeding expectations on a product, providing good customer service, and giving exclusive information to consumers