Acknowledgment:

|  |  |
| --- | --- |
| **A Statistical Analysis of the Indian Smartphone Users** | Submitted by  **ISHIKA JAISWAL(A22017) PULOKESHCHATTERJEE(A22026) SOUMYA ROY(A22034)**    ***PRAXIS BUSINESS SCHOOL PGPDS FALL BATCH’2022 24/10/2022*** |

I would like to thank Mrs. Sayantani Roychowdhury, our professor-in-charge for her support and guidance in completing our project on the topic (topic name). It was a great learning experience.

I would like to take this opportunity to express my gratitude to all my group members Ishika Jaiswal and Soumya RoyThe project would not have been successful without their cooperation and input.

Abstract:

This paper is an attempt to analyze Indian Smartphone users. Corresponding to the data collected, an attempt is made to draw statistical relationships between various attributes. Therefore, to achieve these objectives, the discussion in this paper mainly focuses on various factors that are important from the point of view of any new or existing smartphone users, factors that influence the acceptance and ownership of a new Smartphone.

The data on smartphone users have been collected primarily from various existing smartphone users by doing surveys through Google Forms.

From the data we have performed Graphical EDA and numerical EDA, hence trying to derive statistical significance between various features.

An attempt has been made to discuss the linear relationship or association between the continuous variables depending on the essential values of the measures like Covariance, Correlation, etc.

Finally, we have assumed one smartphone company that is planning to launch a new smartphone in the market can strategies its launch in a better way by utilizing our research findings.

Introduction:

The term „Smartphone‟, refers to a multimedia phone handset, which is a multifunctional electronic device that has features ranging from Camera, Audio-Video Playback, and Web browsing to a high-density screen display along with several other multimedia options. Smartphones have functions that are like those of computers. This is a single solution for all communication problems from voice calls to social media everything is accessible. A smartphone is a mini-computer in your hand that gives you access to media players, digital cameras, GPS, web browsing, videos, emails, audio-video playback, voice chatting, etc. These days most of the websites running a business online are launching their mobile versions. This application has contributed to the huge increase in online shopping and has been a potent reason for adopting Smartphones. In India, there are more than 30 core smartphone users, and the number is increasing rapidly. Even cities with a smaller population of less than 10 lakh also

have 6 percent of people with Smartphones. Nielsen Inform ate in one of its report states that 87% of Smartphone consumers use it for searching for things online whereas 80% use it for social networking. For chatting and emails, the percentage is 72%, and 59% use smartphones for video streaming and navigation. More than 30 percent of people use Smartphones for banking and shopping. Using a mobile phone for watching television online is also a new trend in urban India and 25% of consumers use Smartphones for this purpose. These numbers show that many Indian consumers have been using smartphones for various things, which are dominated by shopping, searching, social networking, taking selfies, entertainment, checking emails, doing business, etc. The major players in smartphone manufacturing in India are Apple, Samsung, Xiaomi, One Plus, and others.

The influx of Chinese smartphones seems to be growing steadily in India. As per International Data Corporation's (IDC) Quarterly Mobile Phone Tracker, 27 million smartphones were shipped in India in the first quarter of CY2017 with a 14.8% growth over the same period last year. In the first quarter, Chinese smartphones held a 51.4% share of smartphone shipments in India with a growth of a whopping 142.6% growth year-on-year. On the contrary, the share of homegrown smartphones dropped to 13.5% in the first quarter of this year from 40.5% last year. The single winner here was Chinese major Miami’s Redmi Note 4. It was the most shipped smartphone in India as against Samsung J2 in the quarter before that. The Quarterly Mobile Phone Tracker report has also spotted some trends in India, like the average selling price of smartphones, which has increased from $131 in Q1 2016 to $155 in Q1 2017. With most networks moving to 4G, 94.5% of smartphones shipped in the first quarter were 4Genabled, with Chinese smartphones leading the race here as well. Camera quality in most phones has also increased over time. As per the IDC report, five out of 10 smartphones shipped in Q1 had a rear camera with 13 megapixels or more. Indian companies such as Micromax, Lava, Reliance Jio, and Intext which together controlled 40 % market share in Q1 2016, have slumped further and now control just 13.5% of the smartphone market share. “Indian vendors are making attempts to recapture the lost ground with new launches in sub-$100 as well as in the mid-range segment. But intense competition from China-based vendors continues to be a major challenge and is expected to increase in the coming quarters.

**Data Collection:**

As our main intention is to know all the important features that can be considered while buying a new smartphone and in turn, the new company which is expected to come up with a new smartphone, can utilize our research findings for their design, so we have only collected data from existing smartphone users through a survey as per response recorded through a Google Form.

*So, we have worked with primary data only.*

In total, we have collected 286 responses from people who fall into different income brackets belonging to various socio-economic cultures. From that collected data we have performed exploratory data analysis intending to draw some statistical relation between various important features.

**Sampling Technique:**

Here we have used the Snowball sampling technique while collecting the data.

Overview and Visualization of Data as received through the Google form:

Before going for the Exploratory data analysis of the data let’s have look at the visualization of the data as collected from the primary survey:

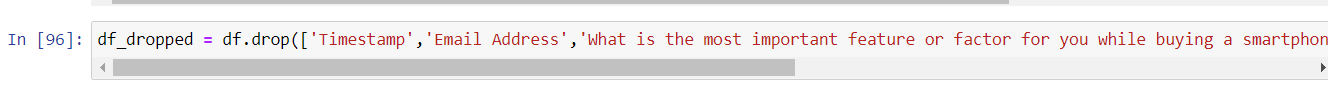
So, seeing at the data, we have figured out that there are many entries from people aged below 30, making the plot completely skewed, to counter the same we have subdivided the entire data set into two parts, which will be discussed below.

**Understanding and analyzing the data:**

To begin with, we have a sort of skewed data as there are influences from the point of age group; we have subdivided the data into two parts:

Age bracket less than and equal to 30 and age above 30.

Now, we have selected some of the features:

****

We have selected a few columns which are not relevant and those which have multiple options selected by the users, and those columns are separated from the main dataset and have been taken into consideration at a later stage.

*Not relevant columns are as follows:*

|  |  |
| --- | --- |
| Timestamp | Email Address |

Columns storing multiple responses from the users are also separated and analysed at a later stage, these are as follows:

|  |  |  |
| --- | --- | --- |
| 'What is the most important feature or factor for you while buying a smartphone? | 'Which brand will you prefer? | For what purpose do you use your smartphone? |

Also, from the total of 286 data, we removed two records because we have found these twos are ambiguous. These two records are taken out from the following column: “what is your preferable build type”

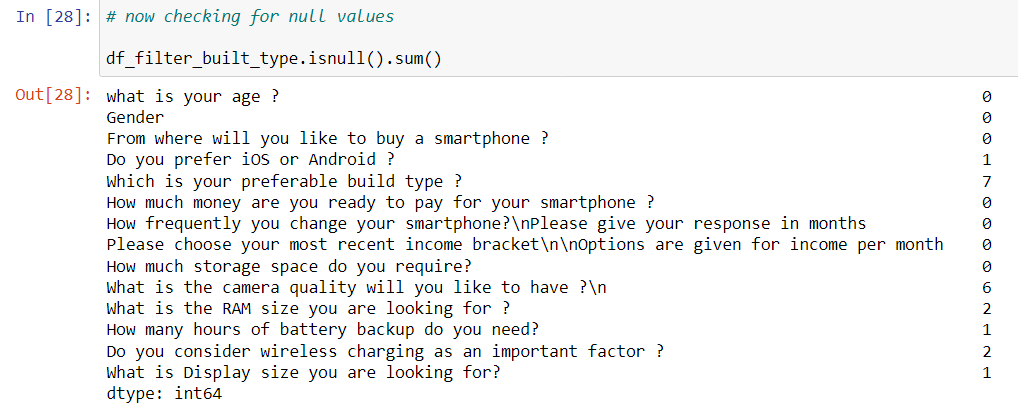
A picture containing scatter chart

Description automatically generated

So, now we are left with 284 records.

Replacing Null values:

Since all the null values in the categorical columns are replaced by the modal value of the already available data with us.



Categorization according to data types:

Text

Description automatically generated

Here we have written some python codes to differentiate categorical and continuous variables.

After Replacing the Null value: Here we can see that the null values are successfully replaced with the moral values, shown below:

Text

Description automatically generated

Removing Outliers from the dataset:

Text

Description automatically generated

Where have we landed after the removal of Outliers?

Text

Description automatically generated

|  |  |  |
| --- | --- | --- |
| Features | Before the Removal of Outliers | After the Removal of Outliers |
| What is your age? | 284 | 278 |
| Money ready to be paid | 278 | 255 |
| Frequency of changing the handset | 255 | 249 |

**Analysis of the continuous variables of age below 30**

Table

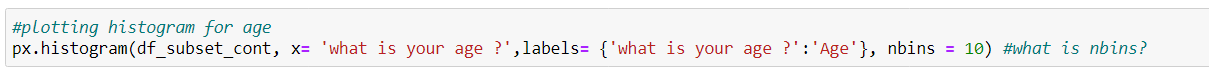
Description automatically generated with medium confidence

Now what we are getting is shown below:

|  |  |  |  |
| --- | --- | --- | --- |
| Measures | Age | Money Spent | Frequency of changing |
| Mean | 20.18 | 24227.83 | 30.09 |
| Standard Deviation | 4.30 | 12266.03 | 17.72 |
|  |  |  |  |

Univariate Analysis on the Continuous Variables:

* Now the data that we have left after doing necessary data cleaning, we are trying to do a Univariate analysis for the same and we have done a graphical EDA here:

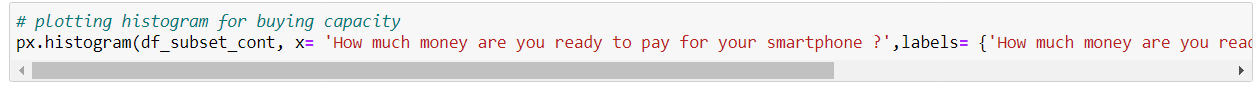


Chart, histogram

Description automatically generated

Here we have created one Histogram on Ages, and from here we can say that our data is positively or rightly skewed

* Here we are trying to plot the Histogram on the feature: Buying Capacity

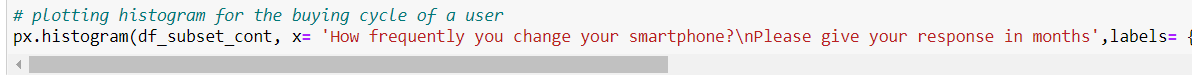


Chart, histogram

Description automatically generated

It shows the distribution is mostly positively skewed.

* Here we are plotting the Histogram on Buying Cycle



The below plot shows that it is rightly or positively skewed.

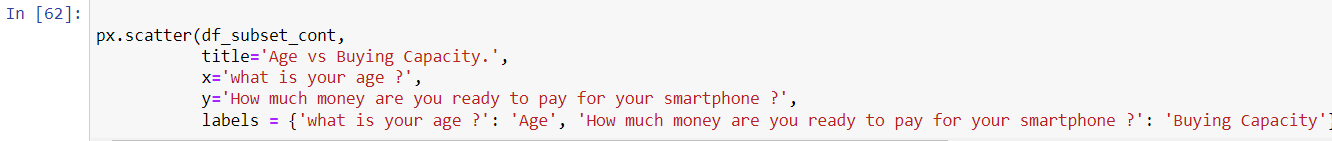
Chart, histogram

Description automatically generated

**Bivariate Analysis:**

Here We will be performing Bivariate analysis and trying to plot a Scatter plot to represent our data and try to derive some meaningful association between the variables taken.

* Scatterplot between Age and Buying Capacity:



Chart, scatter chart

Description automatically generated

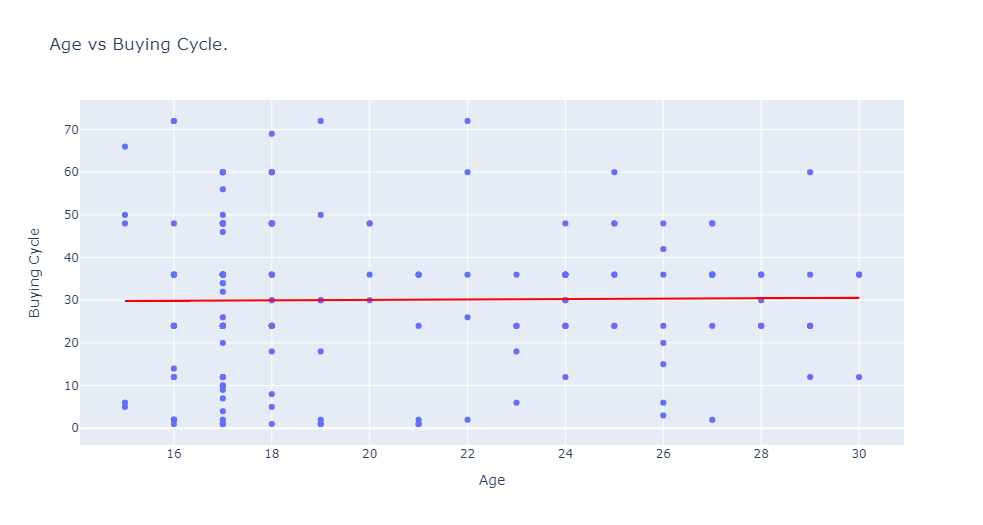
Here we have got OLS(Ordinary Least Square) as below :

Age – 23; Buying Capacity – 33.84; R2 = 0.124

* Scatterplot between Age and Buying Cycle:

*A picture containing scatter chart

Description automatically generated*



Calculation of Correlation between the variables plotted:

* Here we can say that the linear association between the variables is weak except between Age and Buying capacity which has moderate association between the two.

Graphical user interface, text, application

Description automatically generated

Plotting of Stacked Bar Chart:

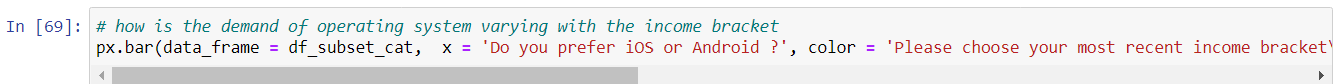
1. Gender and Preferred Buying Window: It shows from the plot that if we consider Gender, then both Males and Females are preferring to buy their smartphones from Physical Stores.
2. Also after Physical stores if we go by the preference of buying windows we can arrange them in the following orders: Flipkart, Amazon.Others



Chart, bar chart

Description automatically generated

1. IOS or Android and Income Bracket: It shows from the plot that if we consider IOS or Android, then 67.10% of the people in the income bracket of 15000 to 30000 per month, are the ones who are likely to buy Android smartphones and people preferring IOS are of 36%.
2. Also, it is seen that these people with the said income bracket are the most contributor while buying a new smartphone, in this age bracket.



Chart, bar chart

Description automatically generated

4.plotting the battery backup and ram, ram is kept in legends:

Most of the people are preferring a battery backup between 12-16 hrs. and amongst these people most of them are also preferring a RAM of 8GB. So, we can conclude that while buying a smartphone the combination of 12-16 hrs. of battery backup with a RAM of 8 GB, are preferred for the age bracket less than equal to 30.

Among the other categories of battery backup, 8GB RAM is the most preferred one.

Chart, bar chart

Description automatically generated

Analysis for Age Bracket greater than 30:

Now, let’s have a look at the analysis with other portion of the data which is greater than 30 years.

Measures of central tendency are as follows:

Table

Description automatically generated

|  |  |  |  |
| --- | --- | --- | --- |
| Measures | Age | Money Spent | Frequency of changing |
| Mean | 46.43 | 2101.46 | 33.58 |
| Standard Deviation | 7.06 | 10977.43 | 17.72 |
|  |  |  |  |

Univariate Analysis on the Continuous Variables:

**Age**

As we can see in the below representation the distribution of age for the responders above 30 are almost normal.

Chart, histogram

Description automatically generated

BUYING CAPACITY

The data distribution for buying capacity is positively skewed.

Chart, histogram

Description automatically generated

Buying cycle

The data distribution for buying capacity is almost normally distributed.

Chart, histogram

Description automatically generated

Bivariate analysis

**Age vs Buying capacity**

Below is a scatter plot drawn between age and buying capacity. As we can see the trend fitted shows negative association.

The regression coefficients are as follows:

Buying capacity: -398.129

Age: 39588.5

R- squared value – 0.0655

Chart, scatter chart

Description automatically generated

**Age vs Buying cycle**

Below is a scatter plot drawn between age and buying cycle.

The regression coefficients are as follows:

Buying cycle: 0.473

Age: 11.53

R- squared value – 0.035

Chart, scatter chart

Description automatically generated

**As we can see from the below snippet below variables have negative and weak association between them:**

**Age and buying capacity**

**Buying capacity and buying cycle**

**Variables having positive association between them:**

**Age and buying cycle**

**Graphical user interface, text, application, email

Description automatically generated**

**Analysis of categorical variable for age above 30**

From the below bar chart, we can conclude that most of the people are preferring to buy a smartphone from a physical store. Among these people the count of male is on the higher side.

If we arrange the shopping preferences in a descending order, the order will be as follows:

Physical Store > Amazon> Flipkart> others

Chart, bar chart

Description automatically generated

From the below chart we can conclude that among the people who are falling under the 15000 to 30000 income bracket takes up around 36 percent of share among the android users.

For the people who are preferring IOS the number is equally divided among 3 income group which are as follows:

30000 to 45000, 15000 to 30000, above 75000

Chart, bar chart, treemap chart

Description automatically generated

**plotting the battery backup and ram, ram is kept in legends.**

From the below graph we can conclude that most of the people are preferring a battery backup of more than 20 hours. If we arrange the count of battery backup in a descending order, then we will get it as follows:

Above 20 > 16-20> 12-16> 8-12> 4-8

Also, we can see that the most preferred combination of ram and battery backup are as follows in a descending order:

Above 20 hours and 8GB of ram

16-20 hours of battery back up and 8 GB ram

12-16 hours of battery backup and 12 gab rams.

Chart, bar chart

Description automatically generated

**Analysis on the multiple option variables.**

This analysis is done on all the sample together, this data is not divided between age <= 30 and age > 30.

Plotting of bar chart on Buying factor:

From the below chart, we can see that processor stands out to be the most important factor for buying a smartphone. If we arrange the buying factors in descending order the order will be as follows:

Processor > Storage space > Price> Camera> RAM> Looks

Chart, bar chart

Description automatically generated

Plotting a bar chart on Brands:

SAMSUNG stands out to be the most preferred brand. If we arrange the brand preferences in descending order the order will be as follows:

SAMSUNG > ONEPLUS > XIAOMI > APPLE = OTHERS

Chart, bar chart

Description automatically generated

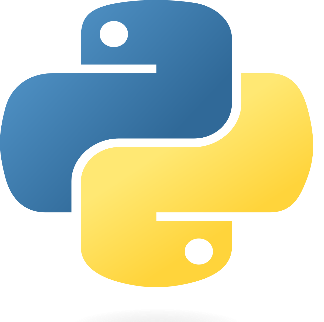
Plotting a bar chart on the Purpose of buying smartphones

Personal use and Photography come out to be the most important purpose for buying a smartphone.

Chart, bar chart

Description automatically generated

**TOOLS USED**



[This Photo](https://en.wikiversity.org/wiki/Python_Concepts) by Unknown Author is licensed under [CC BY-SA](https://creativecommons.org/licenses/by-sa/3.0/)



[This Photo](http://ectdtips.com/2016/07/making-text-hyperlink-blue-in-microsoft-word/) by Unknown Author is licensed under [CC BY-SA](https://creativecommons.org/licenses/by-sa/3.0/)

