Flipkart Sales Product Analysis

In this raw data, A variety of products from different variety of brands have been registered with different prices and discounted prices in Year 2015 and 2016. We are going to analyze the data of Flipkart sales product from 2015 to 2016, we will also focus on other useful insights. This analysis is to leverage Flipkart sales product data to understand the market sales in year 2015 and 16 of products, Analyze total revenue and revenue each year, and identify highest discounted price of products and which product of categories highest sold and Find out which categories registered the minimum and maximum discounted prices and analyze how many users have responded to the rating.

Data collection methodology The data set has many data points such as uniq_id,crawl_timestamp,product_url,product_name,product_category_tree,pid,retail_price,discounted_price,image,is_FK_Advantage_product,description,product_rating,overall_rating,brand,product_specifications.

Statistical and Analytic Issues There are is null value in the data set columns named,retail_price,discounted_price,image,description,brand,product_specific-ations as it consist of the blank cells in the data sets for which data has not been provided. The retail price and discounted price columns have replace the null value with their median value but We have omitted the row of image,descripiton ,brand and product_specifications columns.

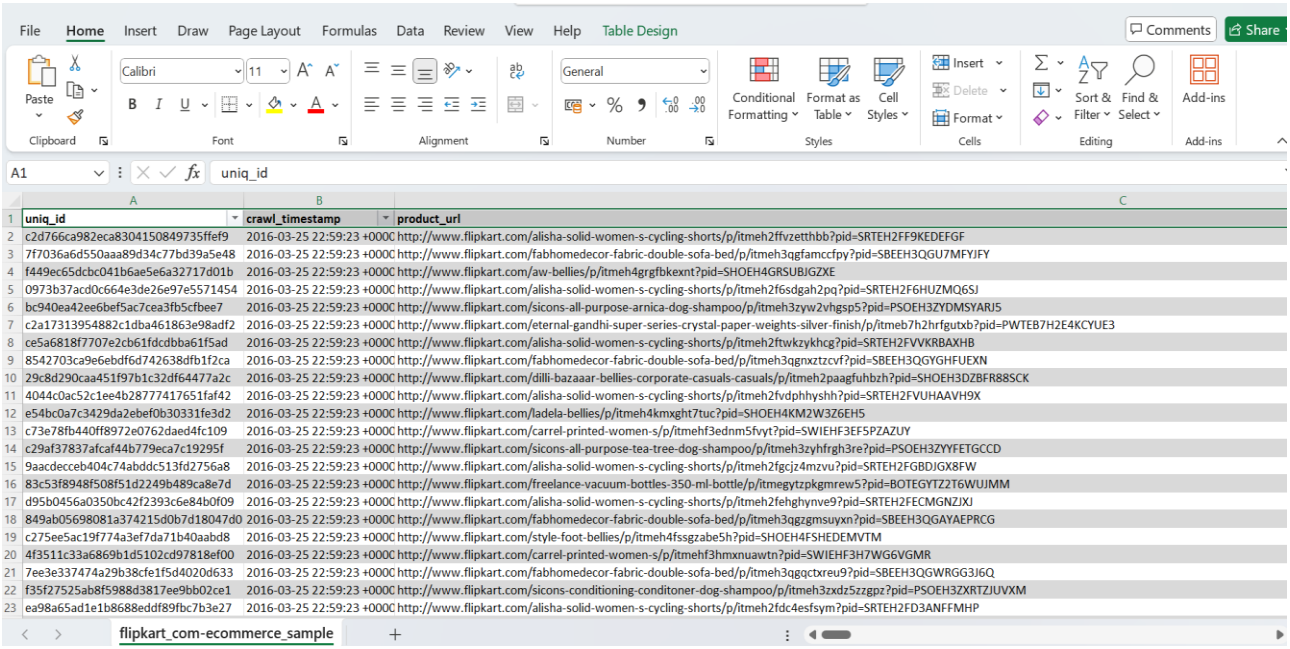
Data description of the data set of column

Data	Discription
Uniq _id	This column consist of unique id of product assigned to each record.
crawl_timestamp	This column contains the date and time of product registration
Product_url	This column consist of product url.
Product name	In this column, Name of the product which is registered
Product_category_tree	In this column,include different categories and Subcategories of products
pid	This column consist of product id assigned to each product
Retail_price	Retail price of product.
Discounted_price	Discounted price of product.
Image	This column consist of product image.
Is_FK_Advantage_product	In this column, a product is part of a flipkart advantage program such as fast shipping, Better customer service etc.
Description	Details of product.
Product_rating	Ratings of products.
Overall_rating	Overall rating of products.
Brand	This column contains a brand of product.
Product_specifications	Full details of product.

Table 1.1: Data description

DATA VISUALIZATION AND ANALYSIS

We will be analyzing the data with the help of some questions. Below is the figure of the data sheet in excel that will give you the hint that how the data is available to us.



uniq_id	crawl_timestamp	product_url
c2d766ca982eca8304150849735ffef9	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-cycling-shorts/p/itmeh2ffvzetthbb?pid=SRTEH2FF9KEDEFGF
7f7036a6d550aaa89d34c77bd39a5e48	2016-03-25 22:59:23 +0000	http://www.flipkart.com/fabhomedecor-fabric-double-sofa-bed/p/itmeh3qgfamccfpy?pid=SBEEH3QGU7MFYJFY
f449ec65dcbc041b6ae5e6a32717d01b	2016-03-25 22:59:23 +0000	http://www.flipkart.com/aw-bellies/p/itmeh4grgfbkexnt?pid=SHOEH4GRSUBJGZXE
0973b37acd0c664e3de26e97e5571454	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-cycling-shorts/p/itmeh2f6sdgah2pq?pid=SRTEH2F6HUZMQ6SJ
bc940ea42ee6bef5ac7cea3fb5cfbee7	2016-03-25 22:59:23 +0000	http://www.flipkart.com/sicons-all-purpose-arnica-dog-shampoo/p/itmeh3zyw2vhgsp5?pid=PSOEH3ZYDMSYARJ5
c2a17313954882c1dba461863e98adf2	2016-03-25 22:59:23 +0000	http://www.flipkart.com/eternal-gandhi-super-series-crystal-paper-weights-silver-finish/p/itmeb7h2hrfgutxb?pid=PWTEB7H2E4KCYUE3
ce5a6818f7707e2cb61fdcdbba61f5ad	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-cycling-shorts/p/itmeh2ftwkzykhcg?pid=SRTEH2FVVKRBAXHB
8542703ca9e6ebdf6d742638dfb1f2ca	2016-03-25 22:59:23 +0000	http://www.flipkart.com/fabhomedecor-fabric-double-sofa-bed/p/itmeh3qgnxztzcvf?pid=SBEEH3QGYGHFUEXN
29c8d290caa451f97b1c32df64477a2c	2016-03-25 22:59:23 +0000	http://www.flipkart.com/dilli-bazaaar-bellies-corporate-casuals-casuals/p/itmeh2paagfuhbzh?pid=SHOEH3DZBFR88SCK
4044c0ac52c1ee4b28777417651faf42	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-cycling-shorts/p/itmeh2fvdphhyshh?pid=SRTEH2FVUHAAVH9X
e54bc0a7c3429da2ebef0b30331fe3d2	2016-03-25 22:59:23 +0000	http://www.flipkart.com/ladela-bellies/p/itmeh4kmxght7tuc?pid=SHOEH4KM2W3Z6EH5
c73e78fb440ff8972e0762daed4fc109	2016-03-25 22:59:23 +0000	http://www.flipkart.com/carrel-printed-women-s/p/itmehf3ednm5fvyt?pid=SWIEHF3EF5PZAZUY
c29af37837afcaf44b779eca7c19295f	2016-03-25 22:59:23 +0000	http://www.flipkart.com/sicons-all-purpose-tea-tree-dog-shampoo/p/itmeh3zyhfrgh3re?pid=PSOEH3ZYYFETGCCD
9aacdecceb404c74abddc513fd2756a8	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-cycling-shorts/p/itmeh2fgcjz4mzvu?pid=SRTEH2FGBDJGX8FW
83c53f8948f508f51d2249b489ca8e7d	2016-03-25 22:59:23 +0000	http://www.flipkart.com/freelance-vacuum-bottles-350-ml-bottle/p/itmegytzpkgmrew5?pid=BOTEGYTZ2T6WUJMM
d95b0456a0350bc42f2393c6e84b0f09	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-cycling-shorts/p/itmeh2fehghynve9?pid=SRTEH2FECMGNZJXJ
849ab05698081a374215d0b7d18047d0	2016-03-25 22:59:23 +0000	http://www.flipkart.com/fabhomedecor-fabric-double-sofa-bed/p/itmeh3qgzgmsuyxn?pid=SBEEH3QGAYAEPRCG
c275ee5ac19f774a3ef7da71b40aabd8	2016-03-25 22:59:23 +0000	http://www.flipkart.com/style-foot-bellies/p/itmeh4fssgzabe5h?pid=SHOEH4FSHEDEMVTM
4f3511c33a6869b1d5102cd97818ef00	2016-03-25 22:59:23 +0000	http://www.flipkart.com/carrel-printed-women-s/p/itmehf3hmxnuawtn?pid=SWIEHF3H7WG6VGMR
7ee3e337474a29b38cfe1f5d4020d633	2016-03-25 22:59:23 +0000	http://www.flipkart.com/fabhomedecor-fabric-double-sofa-bed/p/itmeh3qgqctxreu9?pid=SBEEH3QGWRGG3J6Q
f35f27525ab8f5988d3817ee9bb02ce1	2016-03-25 22:59:23 +0000	http://www.flipkart.com/sicons-conditioning-conditoner-dog-shampoo/p/itmeh3zxdz5zzgpz?pid=PSOEH3ZXRTZJUVXM
ea98a65ad1e1b8688eddf89fbc7b3e27	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-cycling-shorts/p/itmeh2fdc4esfsym?pid=SRTEH2FD3ANFFMHP

Figure 2.1: Flipkart Sales Products Data set

Analysis will be easier by giving explanation to the following set of questions.

1. Analyze the highest sold product enteries and in year 2015 and 16 ? Which was the year where maximum product sold enteries ?
   Solution: The explanation of each and every line is provided inside the program itself.

```
[ ]: import plotly.express as px
     import numpy as np
     import pandas as pd
     import matplotlib.pyplot as plt
     import seaborn as sns
     import warnings as wr
     wr.filterwarnings('ignore')
```

Figure 2.2: Importing the required libraries and package

```
[8]: df=pd.read_csv('flipkart_com-ecommerce_sample.csv')
     df
```

[8]:

	uniq_id	crawl_timestamp	product_url	product_name	product_category_tree
0	c2d766ca982eca8304150849735ffef9	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-c...	Alisha Solid Women's Cycling Shorts	["Clothing >> Women's Clothing >> Lingerie, Sl...
1	7f7036a6d550aaa89d34c77bd39a5e48	2016-03-25 22:59:23 +0000	http://www.flipkart.com/fabhomedecor-fabric-do...	FabHomeDecor Fabric Double Sofa Bed	["Furniture >> Living Room Furniture >> Sofa B...
2	f449ec65dcbc041b6ae5e6a32717d01b	2016-03-25 22:59:23 +0000	http://www.flipkart.com/aw-bellies/p/itmeh4grg...	AW Bellies	["Footwear >> Women's Footwear >> Ballerinas >...
3	0973b37acd0c664e3de26e97e5571454	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-c...	Alisha Solid Women's Cycling Shorts	["Clothing >> Women's Clothing >> Lingerie, Sl...
4	bc940ea42ee6bef5ac7cea3fb5cfbee7	2016-03-25 22:59:23 +0000	http://www.flipkart.com/sicons-all-purpose-arn...	Sicons All Purpose Arnica Dog Shampoo	["Pet Supplies >> Grooming >> Skin & Coat Care...

Figure 2.3: Reading the csv file and show the data set

```
[4]: df.head()
```

[4]:

	uniq_id	crawl_timestamp	product_url	product_name	product_category_tree
0	c2d766ca982eca8304150849735ffef9	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-c...	Alisha Solid Women's Cycling Shorts	["Clothing >> Women's Clothing >> Lingerie, Sl...
1	7f7036a6d550aaa89d34c77bd39a5e48	2016-03-25 22:59:23 +0000	http://www.flipkart.com/fabhomedecor-fabric-do...	FabHomeDecor Fabric Double Sofa Bed	["Furniture >> Living Room Furniture >> Sofa B...
2	f449ec65dcbc041b6ae5e6a32717d01b	2016-03-25 22:59:23 +0000	http://www.flipkart.com/aw-bellies/p/itmeh4grg...	AW Bellies	["Footwear >> Women's Footwear >> Ballerinas >...
3	0973b37acd0c664e3de26e97e5571454	2016-03-25 22:59:23 +0000	http://www.flipkart.com/alisha-solid-women-s-c...	Alisha Solid Women's Cycling Shorts	["Clothing >> Women's Clothing >> Lingerie, Sl...
4	bc940ea42ee6bef5ac7cea3fb5cfbee7	2016-03-25 22:59:23 +0000	http://www.flipkart.com/sicons-all-purpose-arn...	Sicons All Purpose Arnica Dog Shampoo	["Pet Supplies >> Grooming >> Skin & Coat Care...

Figure 2.4: Use of head function

```
[11]: df.isnull().sum()
```

```
[11]: uniq_id                    0
      crawl_timestamp            0
      product_url                0
      product_name               0
      product_category_tree      0
      pid                        0
      retail_price              78
      discounted_price          78
      image                      3
      is_FK_Advantage_product    0
      description                2
      product_rating             0
      overall_rating             0
      brand                   5864
      product_specifications    14
      dtype: int64
```

Figure 2.5: Use of null.sum function

```
[12]: duplicated=df.duplicated()
      print('no.of duplicate instances:',duplicated.sum())
      df[duplicated]
```

```
no.of duplicate instances: 0
```

Figure 2.6: Use of duplicate function

```
[13]: df['retail_price'].fillna(df['retail_price'].median(),inplace=True)
      df['retail_price']
```

```
[13]: 0          999.0
      1        32157.0
      2          999.0
      3          699.0
      4          220.0
```

```
[15]: df['discounted_price'].fillna(df['discounted_price'].median(),inplace=True)
      df['discounted_price']
```

```
[15]: 0          379.0
      1        22646.0
      2          499.0
      3          267.0
      4          210.0
```

Figure 2.7: Use of fillna function to replace null value to median value

```
[17]: df=df.dropna()
      df
```

Figure 2.8: Use of dropna function to remove null value

```
[18]: df['timestamp'] = pd.to_datetime(df['crawl_timestamp'])
```

```
[19]: df['time']=df['timestamp'].dt.time
      df['year']=df['timestamp'].dt.year
      df['month']=df['timestamp'].dt.month
      df['day']=df['timestamp'].dt.day
      df['date']=df['timestamp'].dt.date
      df.drop(['crawl_timestamp'],axis=1,inplace=True)
```

```
[20]: df
```

[20]:

ge	is_FK_Advantage_product	description	product_rating	overall_rating	brand	product_specifications	timestamp	time	year	month	day	date
/...	False	Key Features of Alisha Solid Women's Cycling S...	No rating available	No rating available	Alisha	{"product_specification"=> [{"key"=>"Number of ...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25
/f...	False	FabHomeDecor Fabric Double Sofa Bed (Finish Co...	No rating available	No rating available	FabHomeDecor	{"product_specification"=> [{"key"=>"Installati...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25
		Key Features of										

Figure 2.9: Use of datetime function

```
[21]: df['discount_percentage']=(df['retail_price']-df['discounted_price'])/df['retail_price']*100
      df
```

[21]:

_Advantage_product	description	...	overall_rating	brand	product_specifications	timestamp	time	year	month	day	date	discount_percentage
False	Key Features of Alisha Solid Women's Cycling S...	...	No rating available	Alisha	{"product_specification"=> [{"key"=>"Number of ...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	62.062062
False	FabHomeDecor Fabric Double Sofa Bed (Finish Co...	...	No rating available	FabHomeDecor	{"product_specification"=> [{"key"=>"Installati...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	29.576764
False	Key Features of AW Bellies Sandals Wedges Heel...	...	No rating available	AW	{"product_specification"=> [{"key"=>"Ideal For"...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	50.050050

Figure 2.10: Added a discount percentage Column

```
[23]: df['main_category']=df['product_category_tree'].apply(lambda x:x.split(" >> ")[0].replace('["',' '))
      df
```

[23]:

Advantage_product	description	...	brand	product_specifications	timestamp	time	year	month	day	date	discount_percentage	main_category
False	Key Features of Alisha Solid Women's Cycling S...	...	Alisha	{"product_specification"=> [{"key"=>"Number of ...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	62.062062	Clothing
False	FabHomeDecor Fabric Double Sofa Bed (Finish Co...	...	FabHomeDecor	{"product_specification"=> [{"key"=>"Installati...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	29.576764	Furniture
False	Key Features of AW Bellies Sandals Wedges Heel...	...	AW	{"product_specification"=> [{"key"=>"Ideal For"...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	50.050050	Footwear
False	Key Features of Alisha Solid Women's Cycling S...	...	Alisha	{"product_specification"=> [{"key"=>"Number of ...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	61.802575	Clothing

Figure 2.11: Added a main_category Column

```
[79]: plt.figure(figsize=(8,6))
      sns.countplot(x='year', data=df,palette='BuPu')
      plt.title('flipkart: Higehst sold entries in year:',fontsize=20)
      plt.xlabel('Year',fontsize=12)
      plt.ylabel('counts',fontsize=12)
      plt.tight_layout()
      plt.show()
```

Figure 2.12: Plotting the graph

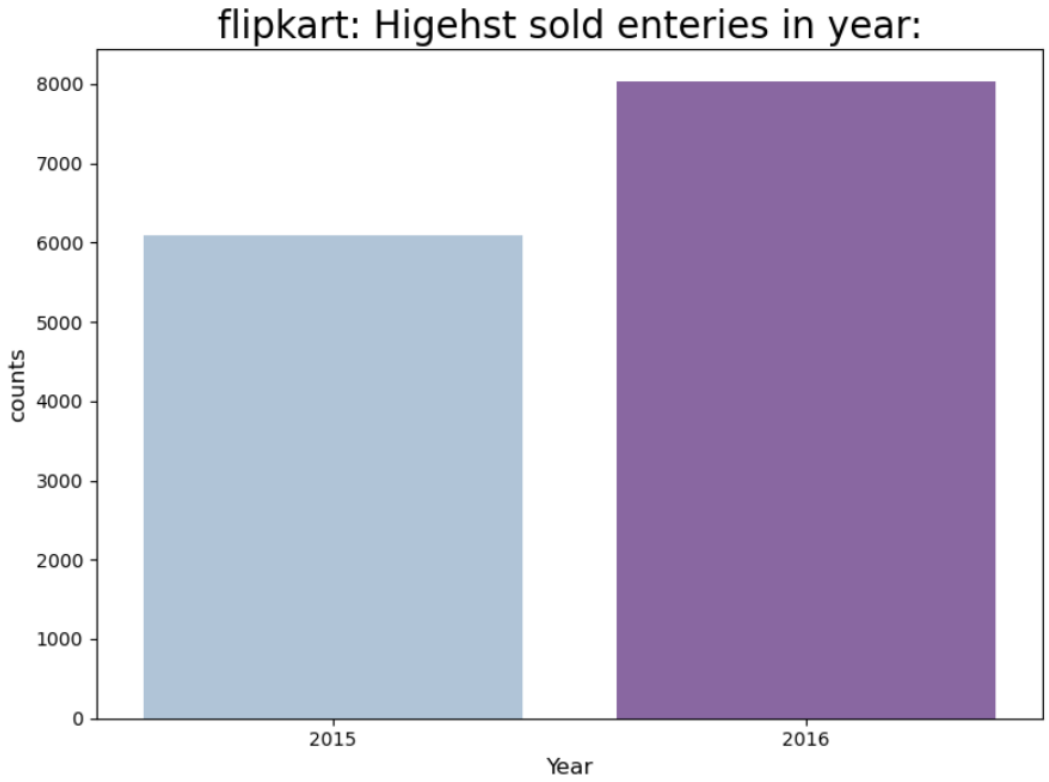


Figure 2.13: Highest sold enteries

```
[87]: plt.figure(figsize=(8, 6))
      sns.countplot(x='year', data=df[df['main_category'].isin(df['main_category'].head())],
                    palette='cubehelix', hue='main_category')
      plt.title('Flipkart: Category Wise Highest Products Sold Entries', fontsize=20)
      plt.xlabel('Year', fontsize=12)
      plt.ylabel('Counts', fontsize=12)
      plt.tight_layout()
      plt.show()
```

Figure 2.14: Plotting the count plot

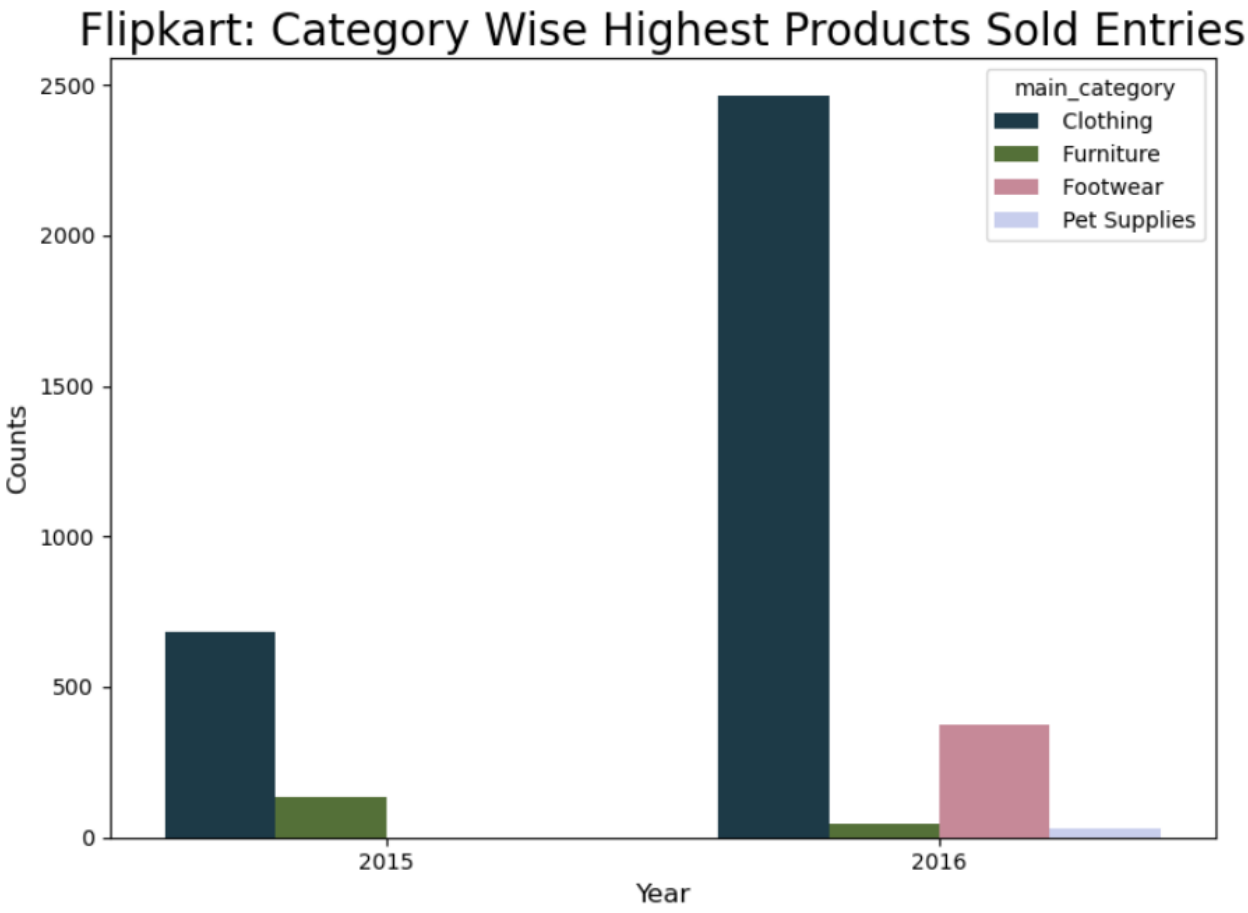


Figure 2.15: Category Wise Highest Products Sold Entries

Conclusion- From the above count plot, It is clear that the enteries of Flipkart's product is increasing over time and product enteries have increased more in 2016 than in 2015,This means that product entries have increased only when product sales have increased.In second plot, In one year, the clothing market in the product category has grown much more than other product categories but the furniture market has down in same year and some product categories as a footwear,pet supplies etc have grown gradually.

2. Flipkart sold which top 20 product categories in two years?
   Solution: The explanation of each and every line is provided inside the program itself.

```
[47]: main_category_counts=df['main_category'].value_counts()[:20]
      main_category_counts
```

```
[47]: main_category
      Jewellery                    3530
      Clothing                     3148
      Mobiles & Accessories        1098
      Automotive                   1012
      Home Decor & Festive Needs    861
      Home Furnishing               698
      Computers                     577
      Baby Care                     457
      Tools & Hardware              391
      Footwear                      375
      Kitchen & Dining              365
      Furniture                     180
      Pens & Stationery             174
      Beauty and Personal Care      155
      Bags, Wallets & Belts         151
      Sports & Fitness              108
      Toys & School Supplies        103
      Cameras & Accessories          82
      Home Improvement               81
      Watches                        48
      Name: count, dtype: int64
```

Figure 2.16: Use of count function

```
[89]: fig = px.bar(main_category_counts,x=main_category_counts.index,
                   y=main_category_counts.values,
                   color=main_category_counts)

      fig.update_layout(title='Top 20 Main Categories', width=1100, height=500)
      fig.show()
```

Figure 2.17: plotting the bar graph

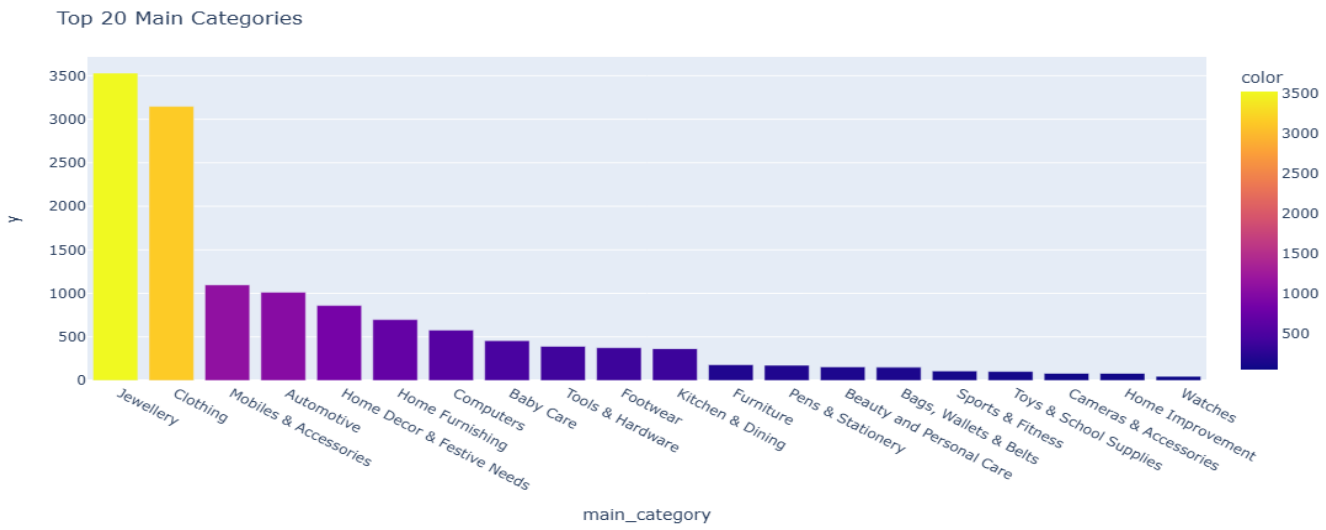


Figure 2.18: Top 20 main categories

```
[49]: df['secondary']=df['product_category_tree'].apply(lambda x:x.split(" >> ")[1] if len(x.split(" >> ")) >1 else '')
      df
```

[49]:

	_FK_Advantage_product	description	...	product_specifications	timestamp	time	year	month	day	date	discount_percentage	main_category	secondary
	False	Key Features of Alisha Solid Women's Cycling S...	...	{"product_specification"=> [{"key"=>"Number of ...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	62.062062	Clothing	Women's Clothing
	False	FabHomeDecor Fabric Double Sofa Bed (Finish Co...	...	{"product_specification"=> [{"key"=>"Installati...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	29.576764	Furniture	Living Room Furniture
	False	Key Features of AW Bellies Sandals Wedges Heel...	...	{"product_specification"=> [{"key"=>"Ideal For"...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	50.050050	Footwear	Women's Footwear
	False	Key Features of Alisha Solid Women's Cycling S...	...	{"product_specification"=> [{"key"=>"Number of ...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	61.802575	Clothing	Women's Clothing

Figure 2.19: using apply,lambda,and split functions

```
[51]: secondary_counts=df['secondary'].value_counts()[:20]
      secondary_counts
```

```
[51]: secondary
      Women's Clothing                2258
      Necklaces & Chains              1606
      Accessories & Spare parts        925
      Tablet Accessories               801
      Bangles, Bracelets & Armlets     724
      Rings"]                          710
      Men's Clothing                   542
      Tools                            403
      Kids' Clothing                   344
      Laptop Accessories               343
                                       323
      Showpieces                       303
      Mobile Accessories               303
      Bed Linen                        217
      Wall Decor & Clocks              212
      Network Components               202
      Baby & Kids Gifts                195
      Women's Footwear                 184
      Infant Wear                      173
      Curtains & Accessories           162
      Name: count, dtype: int64
```

Figure 2.20: using count function

```
[56]: fig = px.bar(y=secondary_counts.index,x=secondary_counts.values,
                   color=secondary_counts)

      fig.update_layout(title='Top 20 Secondary Counts by Count',
                        width=1100, height=500)
      fig.show()
```

Figure 2.21: Plotting bar graph

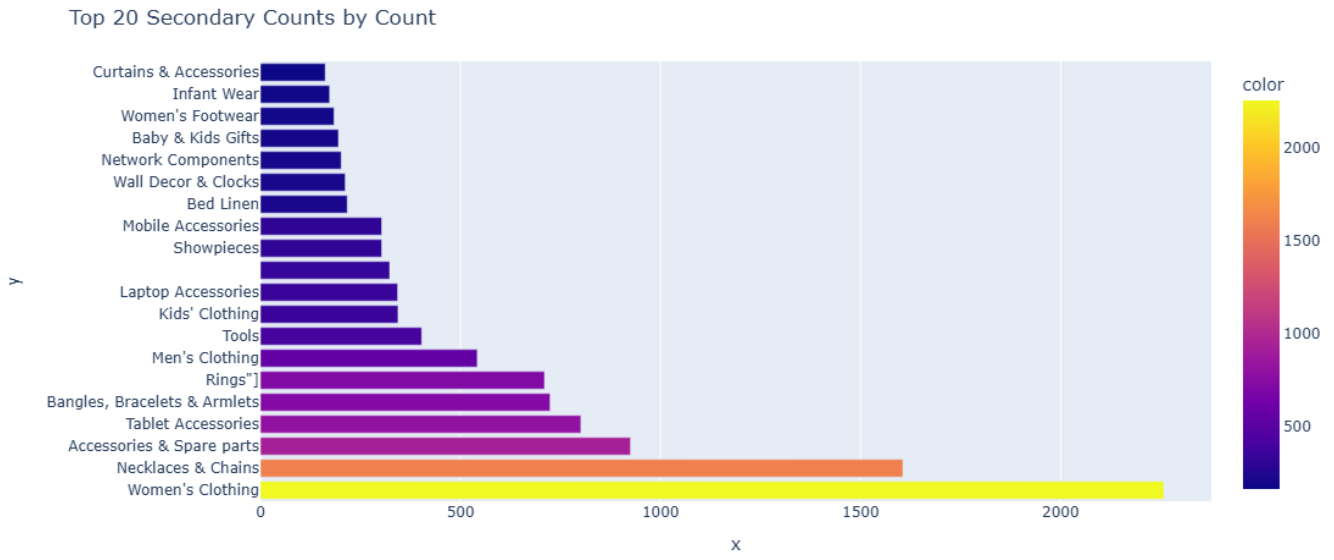


Figure 2.22: Top 20 secondary products categories

```
[57]: df['tertiary']=df['product_category_tree'].apply(lambda x:x.split(' >> ')[2] if len(x.split(" >> ")) >2 else '')
      df
```

[57]:

	image	is_FK_Advantage_product	description	...	timestamp	time	year	month	day	date	discount_percentage	main_category	secondary	tertiary
	hort/u/4/a/...	False	Key Features of Alisha Solid Women's Cycling S...	...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	62.062062	Clothing	Women's Clothing	Lingerie, Sleep & Swimwear
	ofa-bed/j/f...	False	FabHomeDecor Fabric Double Sofa Bed (Finish Co...	...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	29.576764	Furniture	Living Room Furniture	Sofa Beds & Futons
	hoe/7/z/z/r...	False	Key Features of AW Bellies Sandals Wedges Heel...	...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	50.050050	Footwear	Women's Footwear	Ballerinas
	hort/6/2/h/...	False	Key Features of Alisha Solid Women's Cycling S...	...	2016-03-25 22:59:23+00:00	22:59:23	2016	3	25	2016-03-25	61.802575	Clothing	Women's Clothing	Lingerie, Sleep & Swimwear

Figure 2.23: using apply,lambda and split functions

```
[58]: tertiary_counts=df['tertiary'].value_counts()[:20]
      tertiary_counts
```

```
[58]: tertiary
      Necklaces"]                  1567
                                   1452
      Lingerie, Sleep & Swimwear   1179
      Cases & Covers                796
      Western Wear                  736
      Car Interior & Exterior       677
      Bangles"]                     430
      Gardening Tools               343
      Bracelets"]                   251
      Shirts                        217
```

Figure 2.24: using count function

```
[62]: fig = px.bar(tertiary_counts,x=tertiary_counts.index,y=tertiary_counts.values,
                   color=tertiary_counts)

      fig.update_layout(title='Top 20 tertiary by Count', width=1100, height=500)
      fig.show()
```

Figure 2.25: plotting bar gragraph

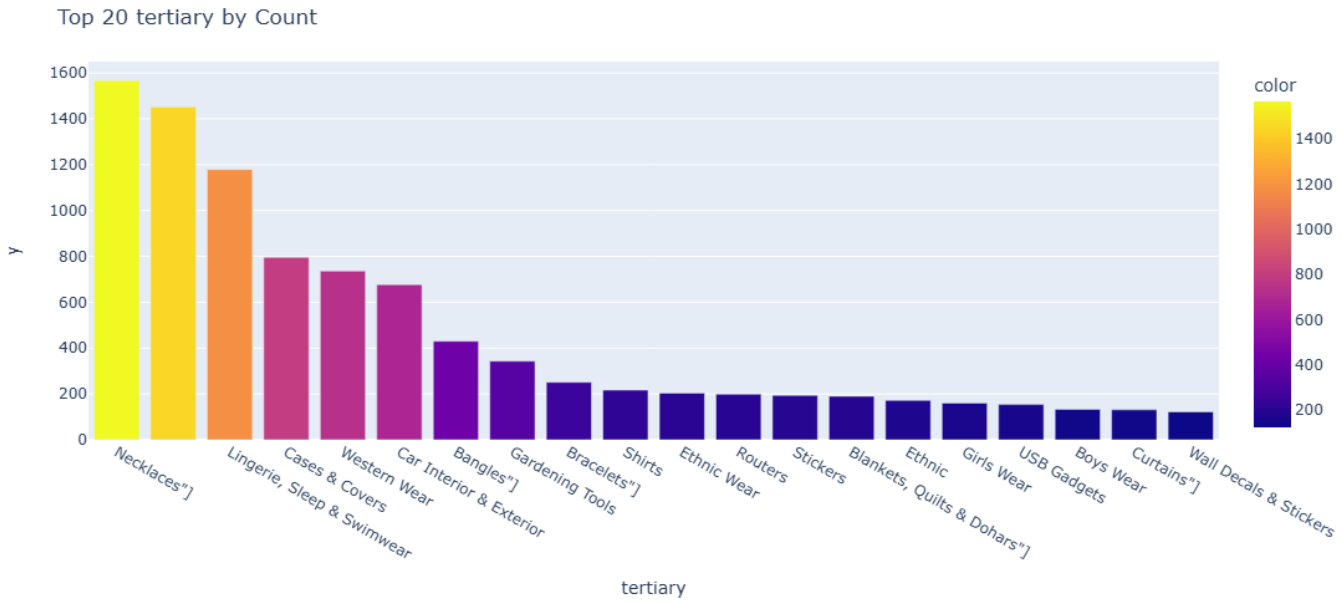


Figure 2.26: Top 20 tertiary products categories

```
[65]: df['quaternary']=df['product_category_tree'].apply(lambda x:x.split(' >> ')[3] if len(x.split(" >> ")) >3 else '')
      df
```

[65]:

	image	is_FK_Advantage_product	description	...	time	year	month	day	date	discount_percentage	main_category	secondary	tertiary	quaternary
	ıort/u/4/a/...	False	Key Features of Alisha Solid Women's Cycling S...	...	22:59:23	2016	3	25	2016-03-25	62.062062	Clothing	Women's Clothing	Lingerie, Sleep & Swimwear	Shorts
	ofa-bed/j/f...	False	FabHomeDecor Fabric Double Sofa Bed (Finish Co...	...	22:59:23	2016	3	25	2016-03-25	29.576764	Furniture	Living Room Furniture	Sofa Beds & Futons	FabHomeDecor Fabric Double Sofa Bed (Finish Co...
	ıoe/7/z/z/r...	False	Key Features of AW Bellies Sandals Wedges Heel...	...	22:59:23	2016	3	25	2016-03-25	50.050050	Footwear	Women's Footwear	Ballerinas	AW Bellies"]
	ıort/6/2/h/...	False	Key Features of Alisha Solid Women's Cycling S...	...	22:59:23	2016	3	25	2016-03-25	61.802575	Clothing	Women's Clothing	Lingerie, Sleep & Swimwear	Shorts
	/image/pet-shampoo/...	False	Specifications of Sicons All Purpose Arnica Do...	...	22:59:23	2016	3	25	2016-03-25	4.545455	Pet Supplies	Grooming	Skin & Coat Care	Shampoo

Figure 2.27: using apply,lambda and split functions

```
[67]: quaternary_counts=df['quaternary'].value_counts()[:20]
      quaternary_counts
```

```
[67]: quaternary
                                          4597
      Bras                                1036
      Car Interior                         659
      Shirts, Tops & Tunics                616
      Plant Containers & Sets              333
      TheLostPuppy Cases & Covers"]        229
      DailyObjects Cases & Covers"]        144
      Formal Shirts                        114
      Dresses & Skirts                     108
      Casual & Party Wear Shirts           103
      Enthopia Cases & Covers"]            101
      Woks & Kadhais                        89
      Car Spare Parts                       88
      Ethnic Wear                           84
      Religious Idols                       80
      Wallmantra Stickers"]                 75
      DeStudio Wall Decals & Stickers"]     71
      WallDesign Stickers"]                 65
      Pizza Cutters                         61
      Dungarees & Jumpsuits                 50
      Name: count, dtype: int64
```

Figure 2.28: count function

```
[69]: fig = px.bar(y=quaternary_counts.index,x=quaternary_counts.values,color=quaternary_counts)

      fig.update_layout(title='Top 20 quaternary Counts by Count', width=1100, height=500)
      fig.show()
```

Figure 2.29: plotting bar graph

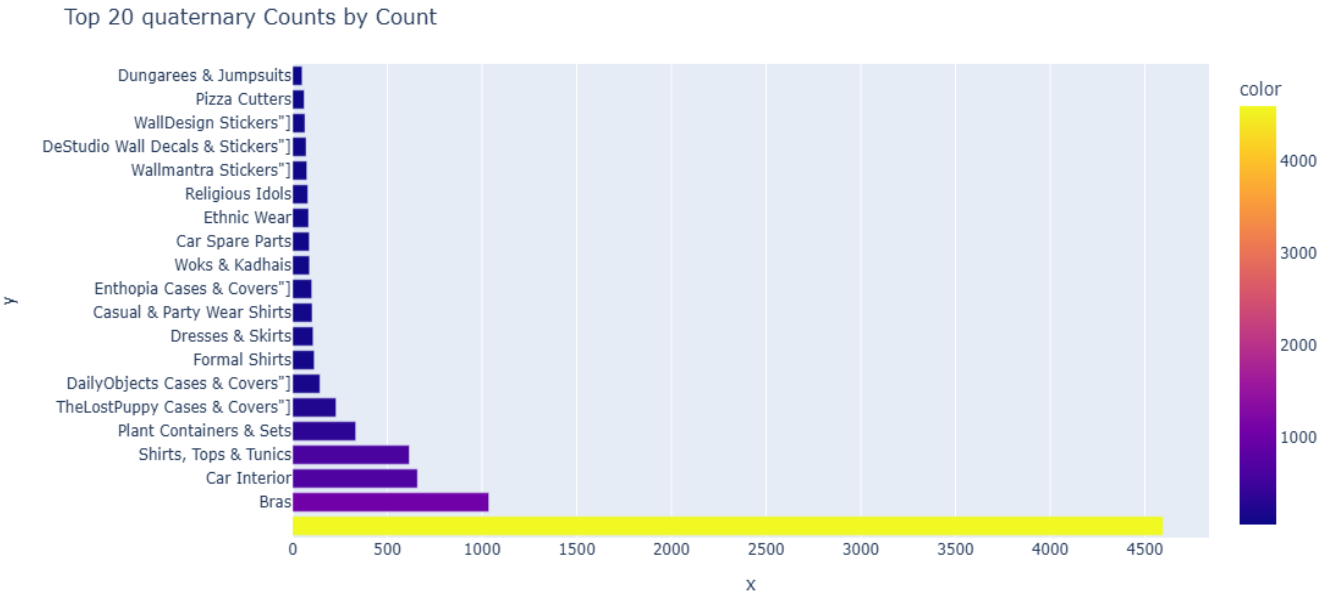


Figure 2.30: Top 20 quaternary products categories

Conclusion- From all the charts given above, we get that the products categories are Clothing, Jewellery,Footwear,Mobile Accessories etc and sub categories are Women's Clothing,Men's Clothing,Necklaces & Chains,Accessories & Spare parts,Tablet Accessories etc.

- In two years, the highest sales have been of jewellery and the highest sales in jewellery have also been of necklaces and chains.
- In two years, the second highest sales have been of clothing and the second highest sales in clothing have also been of women's clothing ,men's clothing and kid's clothing.
- Women's clothing items are Bras,tops,tunics,dresses,skirts,dungarees & Jumpsuits and its sales have been more.
- men's clothing items are shirts,formal shirts,casual and party wear shirts and its sales have been more.
- Apart from this, Mobiles Accessories & Automotive, Home Decor & Festive Needs, HomeFurnishing, Computer, Baby Care, Tools & Hardware, Footwear, Kitchen & Dining, Furniture, Pens & Stationery, Beauty & Personal Care, Bags, Wallets & Belts, Sports & Fitness, Toys & School Supplies, Cameras & Accessories, Home Improvement, Watches have been sold.

3. Analyze the total revenue of flipkart sale's product?
   Solution: The explanation of each and every line is provided inside the program itself.

```
[112]: df_2015=df[df['year']==2015]
       total_revenue_2015=df_2015['discounted_price'].sum()
       total_revenue_result_2015_formated=f"₹{total_revenue_2015:,.2f}"
       total_revenue_result_2015=pd.DataFrame({'year':[2015],'total_revenue':[total_revenue_result_2015_formated]})
       total_revenue_result_2015
```

[112]:

	year	total_revenue
0	2015	₹8,243,208.00

```
[114]: df_2016=df[df['year']==2016]
       total_revenue_2016=df_2016['discounted_price'].sum()
       total_revenue_2016_formated=f"₹{total_revenue_2016:,.2f}"
       total_revenue_result_2016=pd.DataFrame({'year':[2016],'total_revenue':[total_revenue_2016_formated]})
       total_revenue_result_2016
```

[114]:

	year	total_revenue
0	2016	₹25,805,718.00

Figure 2.31: using sum function and format method

```
[116]: merged= pd.concat([total_revenue_result_2015, total_revenue_result_2016])
       merged
```

[116]:

	year	total_revenue
0	2015	₹8,243,208.00
0	2016	₹25,805,718.00

```
[81]: merged['total_revenue'] =merged['total_revenue'].replace({'₹': '', ',': ''},regex=True).astype(float)
      merged['total_revenue']
```

```
[81]: 0     8243208.0
      0    25805718.0
      Name: total_revenue, dtype: float64
```

Figure 2.32: using concat and astype method

```
[126]: import matplotlib.pyplot as plt
       merged = {'total_revenue': [500, 700]}
       plt.figure(figsize=(5, 6))
       labels = [2015, 2016]
       plt.pie(merged['total_revenue'], labels=labels, autopct='%1.2f%%', colors=['Tomato', 'SteelBlue'])
       plt.title('Flipkart: Year wise Total revenue', fontsize=20)
       plt.legend()
       plt.tight_layout()
       plt.show()
```

Figure 2.33: plotting pie chart

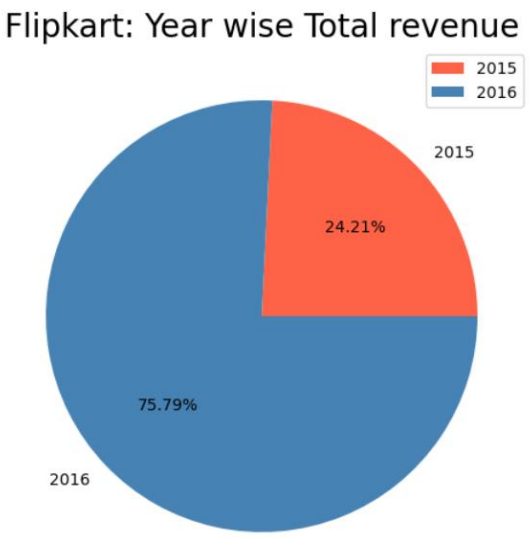


Figure 2.34: Year Wise total revenue

```
[85]: total_revenue=((total_revenue_2015+total_revenue_2016))
      total_revenue
```

```
[85]: 34048926.0
```

Figure 2.35:using adittion operator

Conclusion- Flipkart achieved revenue of INR 8,243,208.00 in 2015 and INR 25,805,718.00 in 2016.

- Flipkart has imporoved its selling perfomance by INR 17,562,510 (51.58%) in one year and it is a significant improvement from previous year.
- Flipkart achieved total revenue of INR 34,048,926 in 2015 & 16.

4. What is the maximum and minimum discount percentage with product category and maximum and minimum discount price with product category?
   Solution: The explanation of each and every line is provided inside the program itself.

```
[70]: max_row = df.loc[df['discount_percentage'].idxmax()]
      print(f"Maximum Discount Percentage: {max_row['discount_percentage']}")
      print(f"Category with Maximum Discount: {max_row['main_category']}")

      Maximum Discount Percentage: 96.53333333333333
      Category with Maximum Discount:  Home Furnishing

[72]: min_row=df.loc[df['discount_percentage'].idxmin()]
      print(f"minimun discount percentage:{min_row['discount_percentage']}")
      print(f"category with minimum discount:{min_row['main_category']}")

      minimun discount percentage:0.0
      category with minimum discount: Eternal Gandhi Super Series Crystal Paper Weight..."]
```

Figure 2.36:using loc and f-string

```
[73]: max_row_discount=df.loc[df['discounted_price'].idxmax()]
      print(f"maximum discount price:{max_row_discount['discounted_price']}")
      print(f"category with maximum discount:{max_row_discount['main_category']}")

      maximum discount price:162825.0
      category with maximum discount: Furniture

[74]: min_row_discount=df.loc[df['discounted_price'].idxmin()]
      print(f"minimun discount price:{min_row_discount['discounted_price']}")
      print(f"category with minimun discount:{min_row_discount['main_category']}")

      minimun discount price:35.0
      category with minimun discount: Kitchen & Dining
```

Figure 2.37:using loc and f-string

Conclusion- The highest discount percentage of 96.55% was received on home furnishing product category and the lowest discount percentage of 0% was received on watches,This means that the watches are sold at retail prices without discounts.

The highest discount price of INR 571230.0 was received on watches and the lowest discount price of 35.0 was received on kitchen and dining.

5. Analyze the Product Ratings ?
   Solution: The explanation of each and every line is provided inside the program itself.

```
[79]: product_rating_counts=df['product_rating'].value_counts()
      product_rating_counts
```

```
[79]: product_rating
      No rating available    13198
      5                        341
      4                        114
      1                         93
      3                         71
      2                         31
      4.2                       31
      4.5                       25
      3.7                       23
      4.3                       21
      3.5                       18
      2.5                       17
      3.6                       13
      3.8                       11
      4.8                       11
```

Figure 2.38:using count method

```
[104]: plt.figure(figsize=(12,8))
       sns.countplot(x='product_rating',data=df,palette='dark')
       plt.title('flipkart:-count of ratings:',fontsize=20)
       plt.xlabel('ratings',fontsize=12)
       plt.ylabel('counts',fontsize=12)
       plt.xticks(rotation=90)
       plt.tight_layout()
       plt.show()
```

Figure 2.39: plotting count plot

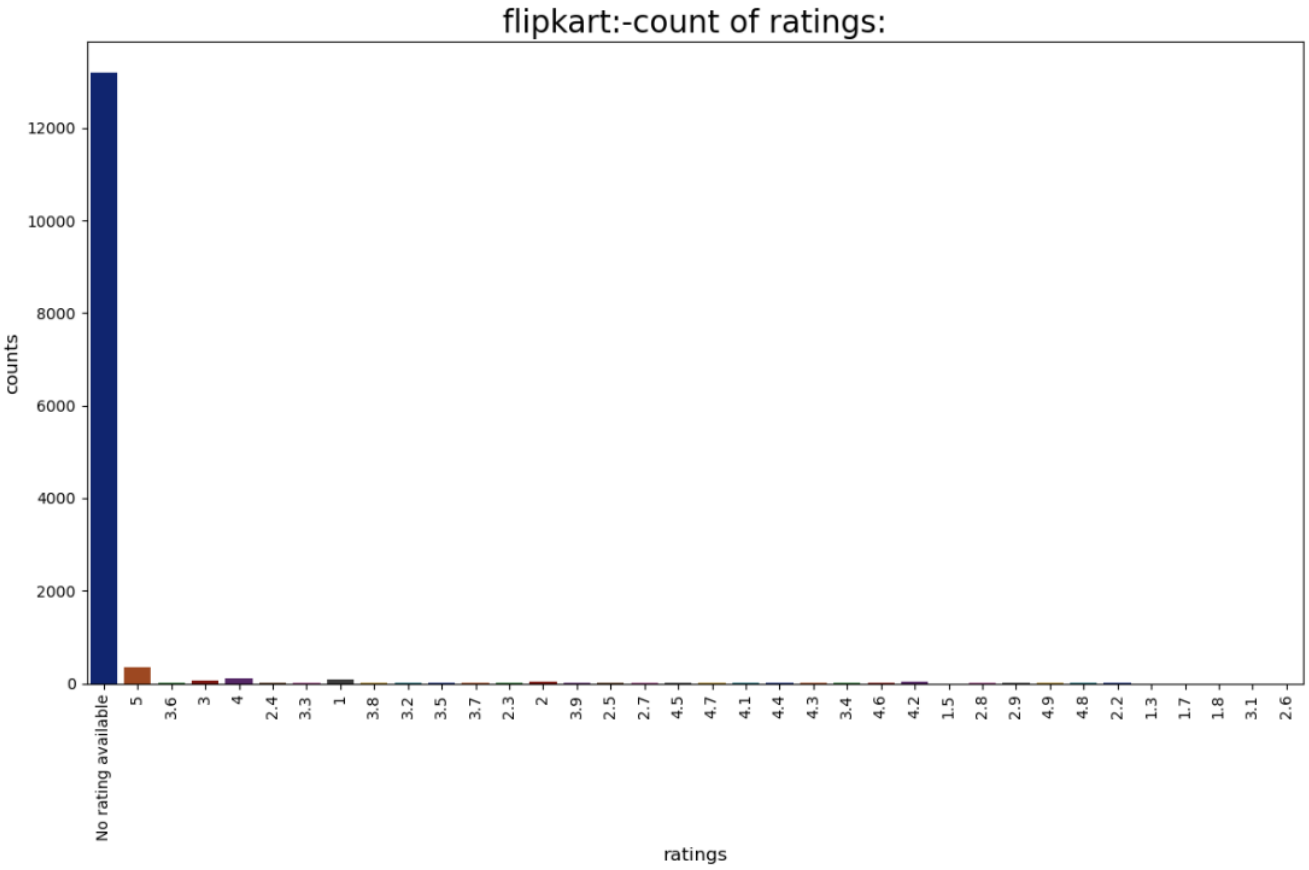


Figure 2.40: count of Ratings

Conclusion- We know that the market should grow more and more, so product rating is the most important factor because according to the rating, some important changes can be made in the service of that product or the product that the product grows in the market.

- The graph above shows that customers have not rated most of the products and we know that customers do not rate or respond to the product, so it should be considered to create a policy that allows the customer to give maximum ratings and feedback.
- Some customers have rated 5. This means that the product which has got 5 ratings is the best product.