Shruthika Varri

DATA ANALYST PROJECT

Customer Satisfaction Prediction

Data Set:

	A	B	C	D	E	F	G	H
1	Ticket ID	Customer Name	Customer Email	Customer Age	Customer Gender	Product Purchased	Date of Purchase	Ticket Type
2	1	Marisa Obrien	carrollallison@example.com	32	Other	GoPro Hero	22/03/2021	Technical issue
3	2	Jessica Rios	clarkeashley@example.com	42	Female	LG Smart TV	22/05/2021	Technical issue
4	3	Christopher Robbins	gonzalestracy@example.com	48	Other	Dell XPS	14/07/2020	Technical issue
5	4	Christina Dillon	bradleyolson@example.org	27	Female	Microsoft Office	13/11/2020	Billing inquiry
6	5	Alexander Carroll	bradleymark@example.com	67	Female	Autodesk AutoCAD	04/02/2020	Billing inquiry
7	6	Rebecca Fleming	sheenasmith@example.com	53	Male	Microsoft Office	28/07/2020	Cancellation request
8	7	Jacqueline Wright	donaldkeith@example.org	24	Other	Microsoft Surface	23/02/2020	Product inquiry
9	8	Denise Lee	joelwilliams@example.com	23	Male	Philips Hue Lights	09/08/2020	Refund request
10	9	Nicolas Wilson	joshua24@example.com	60	Other	Fitbit Versa Smartwatch	16/07/2020	Technical issue
11	10	William Dawson	clopez@example.com	27	Male	Dyson Vacuum Cleaner	06/03/2020	Refund request
12	11	Joseph Moreno	mbrown@example.org	48	Male	Nintendo Switch	19/01/2021	Cancellation request
13	12	Brandon Arnold	davisjohn@example.net	51	Male	Microsoft Xbox Controller	24/10/2021	Product inquiry
14	13	Tamara Hahn	jensenwilliam@example.net	27	Other	ntendo Switch Pro Controll	26/05/2021	Technical issue
15	14	Sandra Barnes	gwendolyn51@example.net	65	Other	Nest Thermostat	13/07/2020	Technical issue
16	15	Amy Hill	medinasteven@example.net	48	Female	Sony PlayStation	29/02/2020	Billing inquiry
17	16	Elizabeth Foley	amy41@example.net	18	Other	GoPro Action Camera	24/06/2021	Billing inquiry
18	17	Julia Salazar	watkinsbarbara@example.com	63	Other	Xbox	13/10/2021	Product inquiry
19	18	Joshua Castillo	mooredeborah@example.org	56	Female	Microsoft Xbox Controller	07/09/2020	Product inquiry
20	19	Wendy Davis	brenda20@example.net	19	Male	LG Washing Machine	23/09/2021	Product inquiry
21	20	Jeffrey Robertson	jameslopez@example.com	39	Female	Canon EOS	08/03/2021	Refund request
22	21	Suzanne Holmes	rogermcgrath@example.net	28	Other	HP Pavilion	31/05/2021	Refund request
23	22	Tanner Conley	zbond@example.net	66	Male	Fitbit Versa Smartwatch	04/02/2020	Cancellation request
24	23	Stephanie Nelson DVM	ljohnson@example.org	54	Female	Xbox	11/02/2020	Cancellation request
25	24	Antonio Scott	carlsonmatthew@example.org	63	Female	ntendo Switch Pro Controll	06/08/2021	Technical issue
26	25	John Cuevas	lbarron@example.org	19	Male	Amazon Kindle	25/05/2021	Product inquiry

1)Importing python libraries and setting the pandas options up to 17 columns:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')


#setting pandas options(maximum coloumns options)
pd.set_option('display.max_columns', 17)
```

2) Checking for null values:

df.isnull().sum()

```
Ticket ID                          0
Customer Name                      0
Customer Email                     0
Customer Age                       0
Customer Gender                    0
Product Purchased                  0
Date of Purchase                   0
Ticket Type                        0
Ticket Subject                     0
Ticket Description                 0
Ticket Status                      0
Resolution                      5700
Ticket Priority                    0
Ticket Channel                     0
First Response Time             2819
Time to Resolution              5700
Customer Satisfaction Rating       0
dtype: int64
```

This shows that Resolution, Ticket Response time and Time to Resolution has null values and need to be handled.

3) Checking duplicate values

df.duplicated().sum()

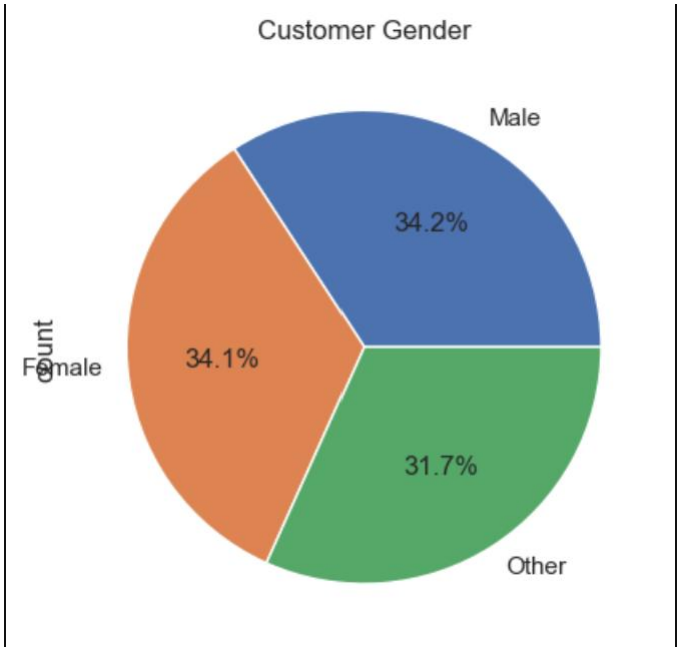
```
[584]: df.duplicated().sum()

[584]: 0
```

This shows that the data has no duplicate values in the data set.

4) Filling the null values with the forward fill method

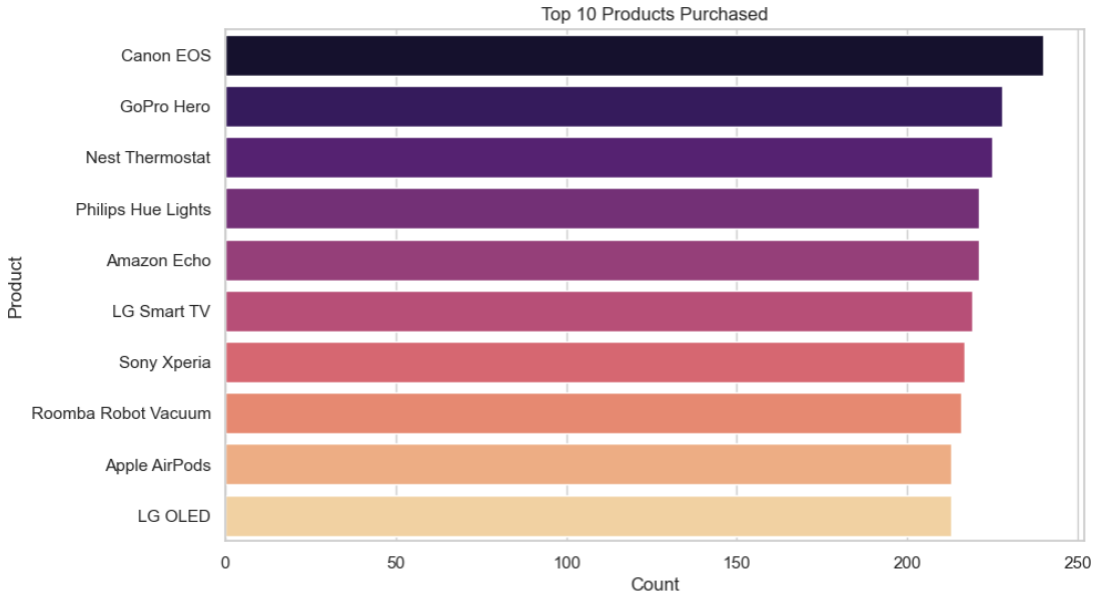
df['First Response Time'].fillna(method='ffill', inplace=True)

5) Pie chart on distribution of Customers based on their gender



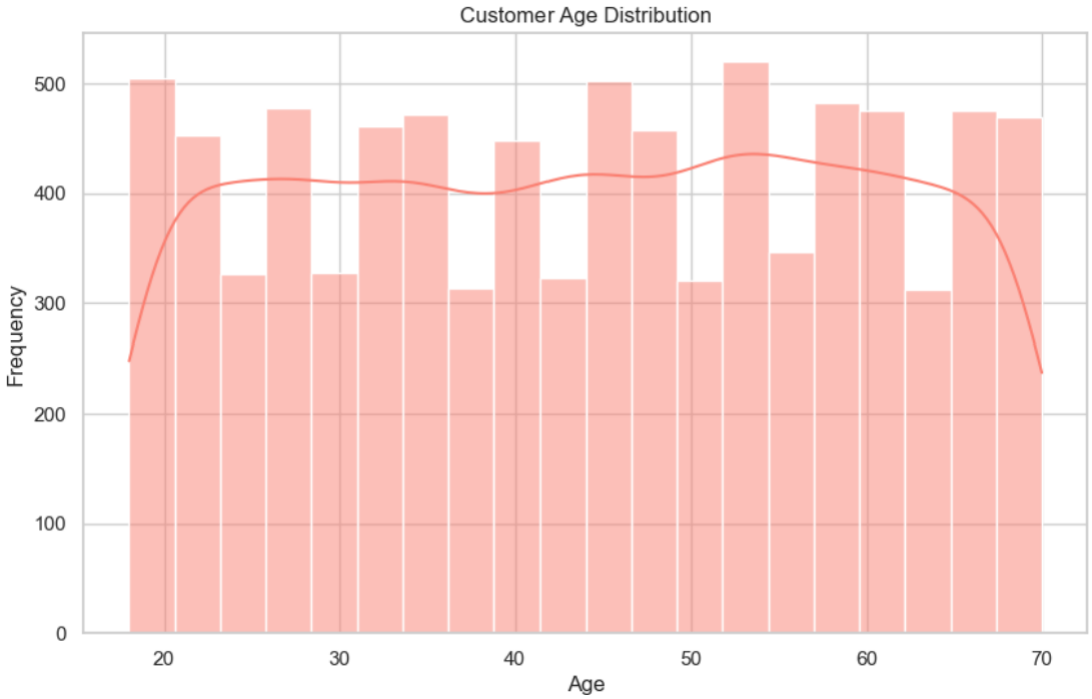
We have equal number of females and males as shown in the diagram apart from the others who haven't mentioned

6) Representation of Products purchased



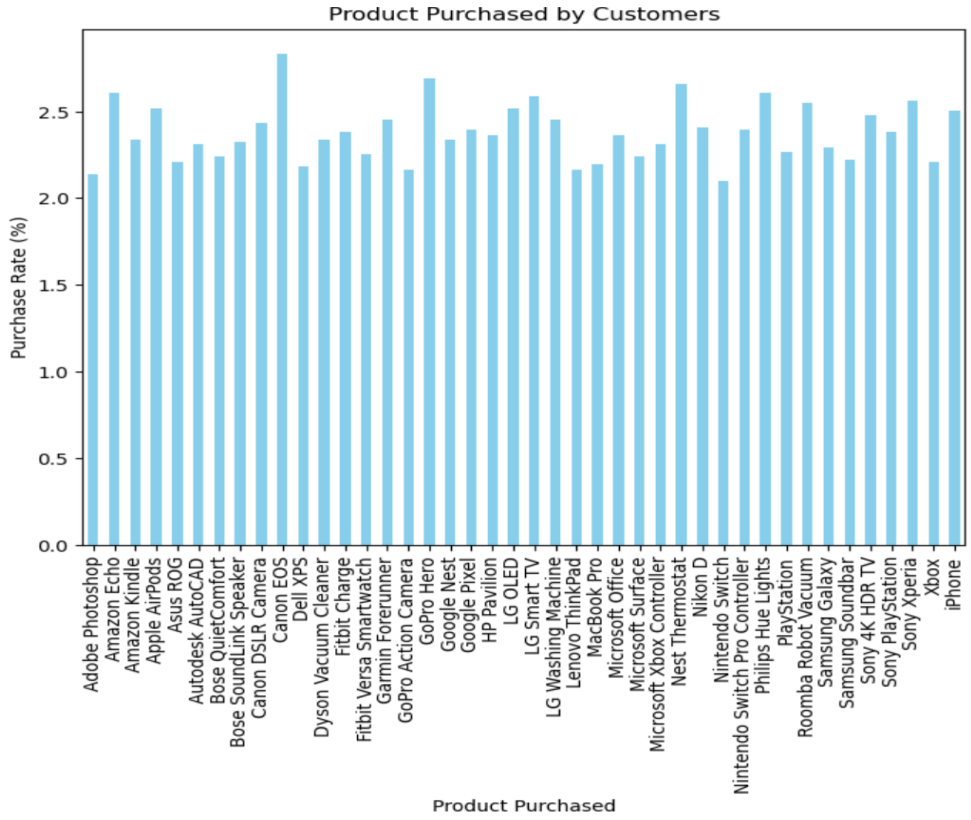
This shows the distribution of each product purchased with the most purchased being the Canon EOS and the least is LG OLED

7) Distribution of Customer age

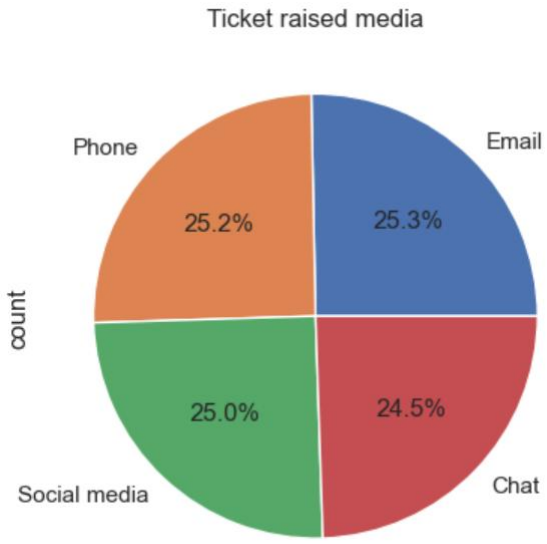


As we see the ages are distributed more or less from the ages of 20 to 70

8) Percentage of products purchased by customers

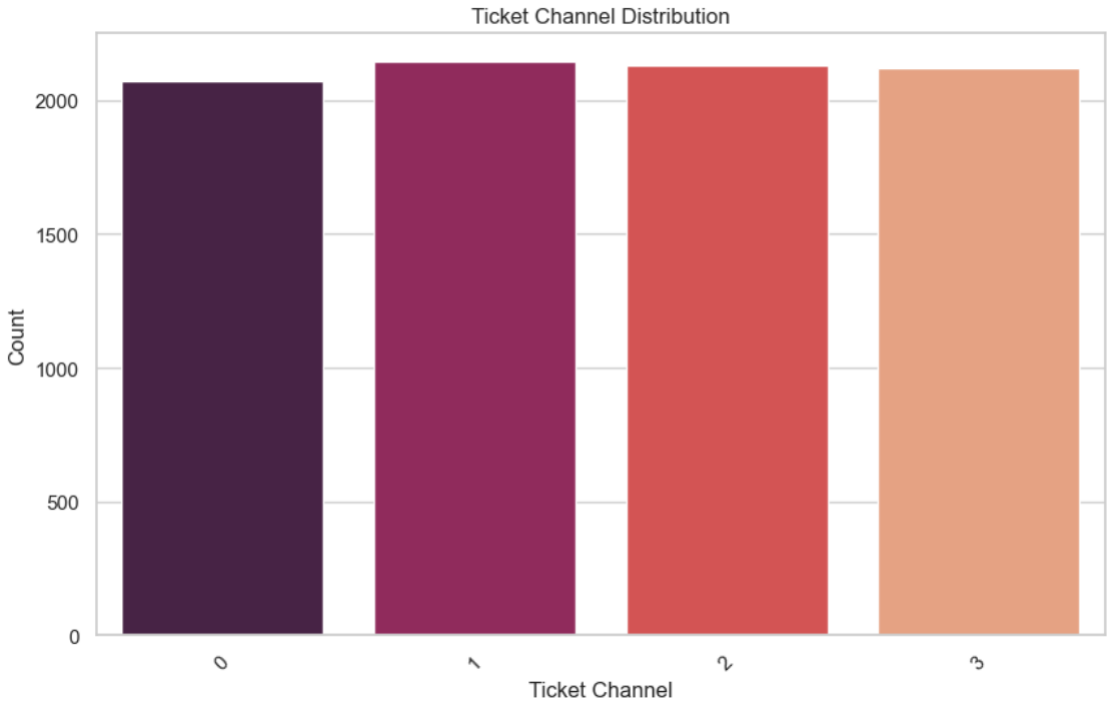


9) Distribution of Tickets raised by media

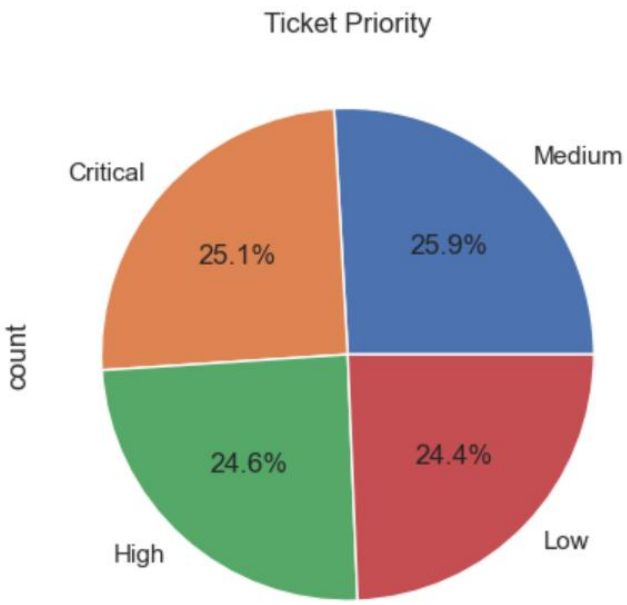


We have a equal distribution of tickets raised by the customers with 25% form each media.

10) Ticket Channel distribution

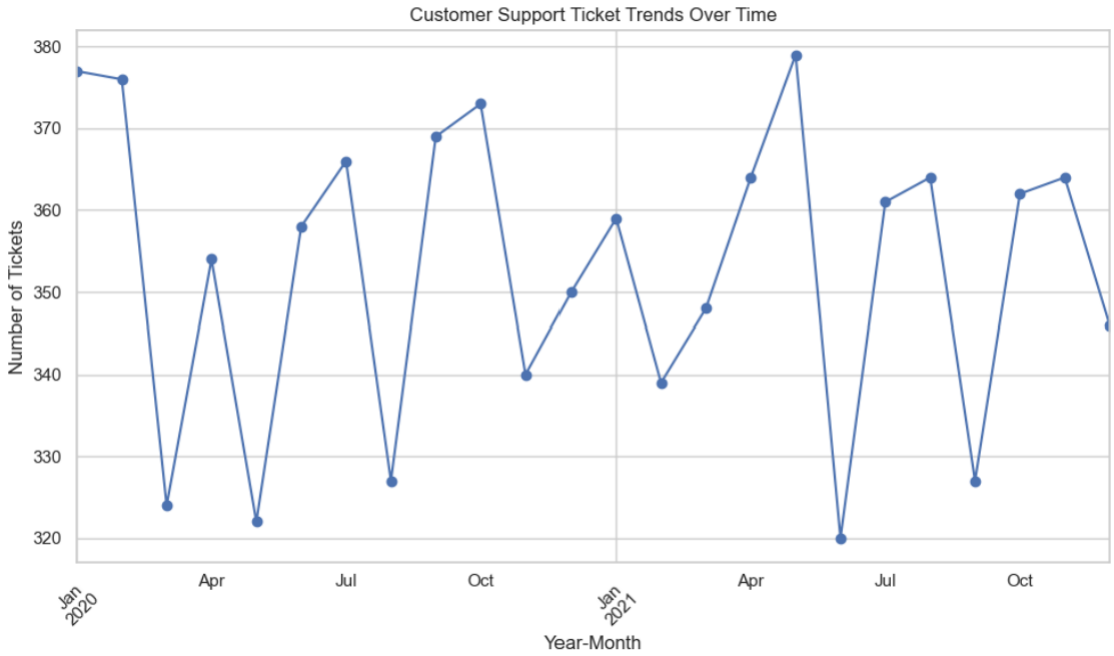


11) Ticket priority



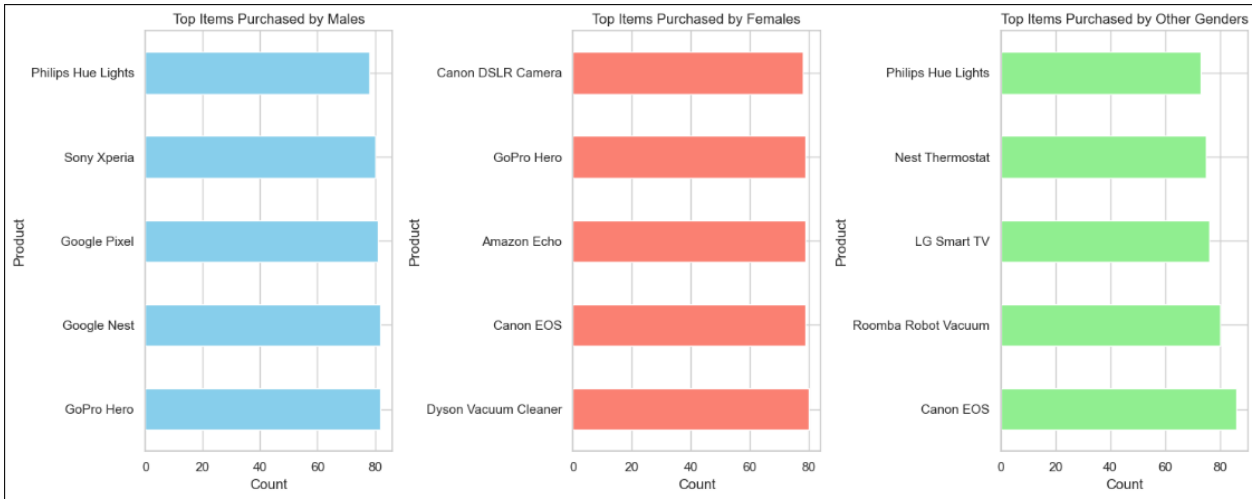
Tickets are either critical or medium priority mostly

12) Customer support ticket trends over time

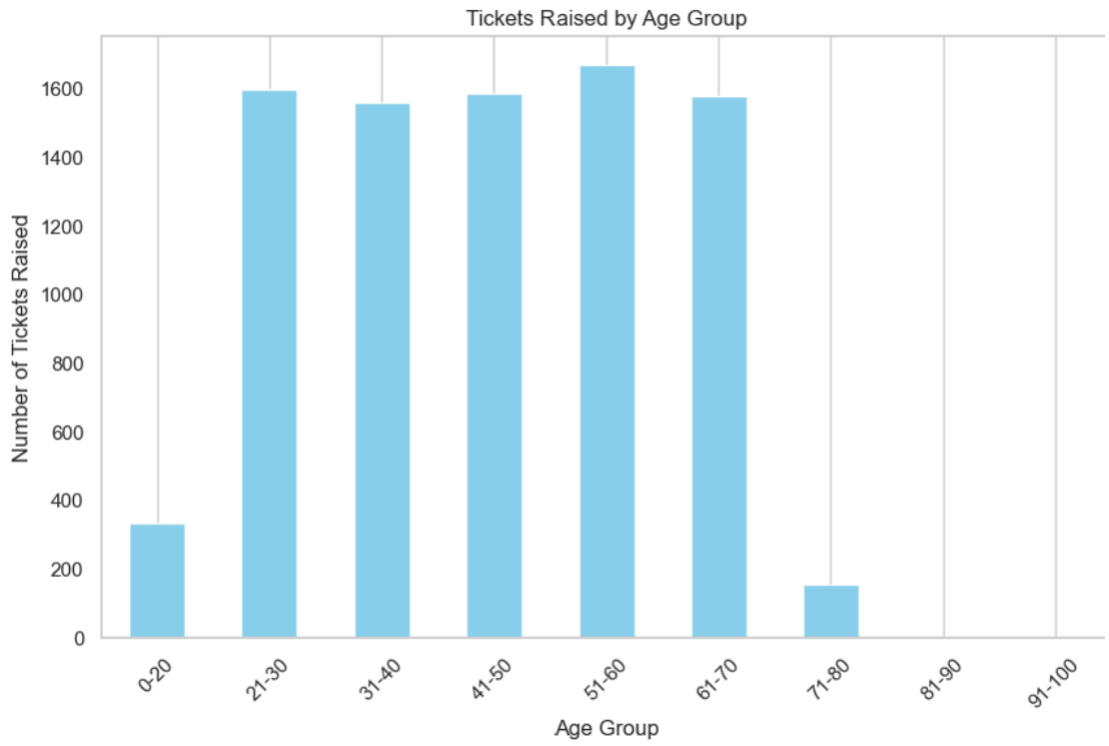


13) Rating of Customer satisfaction trend

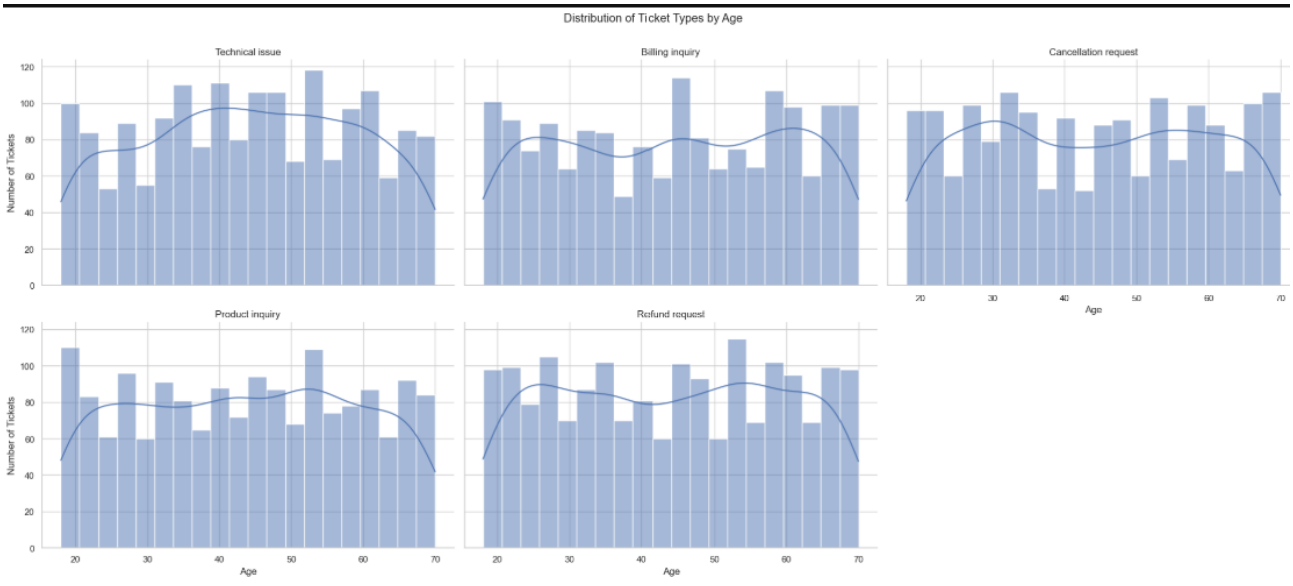
14) Top items purchased based on gender



15) Tickets raised by different age groups



16) Distribution of tickets by age



17) Top 10 Feature Importances