

Amazon Customer Behavior Survey

In []:

```
In [1]: #import Libraries  
import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
In [3]: #Loading the dataset from csv
df=pd.read_csv("Amazon Customer Behavior Survey.csv")
df
```

Out[3]:

	Timestamp	age	Gender	Purchase_Frequency	Purchase_Categories	Personalized_Recomm
0	2023/06/04 1:28:19 PM GMT+5:30	23	Female	Few times a month	Beauty and Personal Care	
1	2023/06/04 2:30:44 PM GMT+5:30	23	Female	Once a month	Clothing and Fashion	
2	2023/06/04 5:04:56 PM GMT+5:30	24	Prefer not to say	Few times a month	Groceries and Gourmet Food;Clothing and Fashion	
3	2023/06/04 5:13:00 PM GMT+5:30	24	Female	Once a month	Beauty and Personal Care;Clothing and Fashion;...	
4	2023/06/04 5:28:06 PM GMT+5:30	22	Female	Less than once a month	Beauty and Personal Care;Clothing and Fashion	
...	
597	2023/06/12 4:02:02 PM GMT+5:30	23	Female	Once a week	Beauty and Personal Care	
598	2023/06/12 4:02:53 PM GMT+5:30	23	Female	Once a week	Clothing and Fashion	
599	2023/06/12 4:03:59 PM GMT+5:30	23	Female	Once a month	Beauty and Personal Care	
600	2023/06/12 9:57:20 PM GMT+5:30	23	Female	Few times a month	Beauty and Personal Care;Clothing and Fashion;...	
601	2023/06/16 9:16:05 AM GMT+5:30	23	Female	Once a week	Clothing and Fashion	

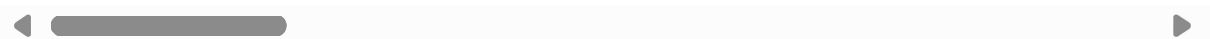
602 rows × 23 columns

In [4]: `df.head(10)`

Out[4]:

	Timestamp	age	Gender	Purchase_Frequency	Purchase_Categories	Personalized_Recommen
0	2023/06/04 1:28:19 PM GMT+5:30	23	Female	Few times a month	Beauty and Personal Care	
1	2023/06/04 2:30:44 PM GMT+5:30	23	Female	Once a month	Clothing and Fashion	
2	2023/06/04 5:04:56 PM GMT+5:30	24	Prefer not to say	Few times a month	Groceries and Gourmet Food;Clothing and Fashion	
3	2023/06/04 5:13:00 PM GMT+5:30	24	Female	Once a month	Beauty and Personal Care;Clothing and Fashion;...	
4	2023/06/04 5:28:06 PM GMT+5:30	22	Female	Less than once a month	Beauty and Personal Care;Clothing and Fashion	
5	2023/06/04 6:01:59 PM GMT+5:30	21	Female	Less than once a month	Clothing and Fashion	
6	2023/06/04 6:31:41 PM GMT+5:30	22	Female	Less than once a month	Clothing and Fashion	
7	2023/06/04 7:13:12 PM GMT+5:30	21	Female	Few times a month	Beauty and Personal Care;Clothing and Fashion	
8	2023/06/04 7:23:21 PM GMT+5:30	20	Female	Less than once a month	Beauty and Personal Care;Clothing and Fashion	
9	2023/06/04 7:33:12 PM GMT+5:30	23	Female	Less than once a month	Beauty and Personal Care;Clothing and Fashion	

10 rows × 23 columns



In [5]: `df.shape`

Out[5]: (602, 23)

In [6]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 602 entries, 0 to 601
Data columns (total 23 columns):
#   Column                                          Non-Null Count  Dtype
---  -
0   Timestamp                                     602 non-null    object
1   age                                             602 non-null    int64
2   Gender                                         602 non-null    object
3   Purchase_Frequency                           602 non-null    object
4   Purchase_Categories                           602 non-null    object
5   Personalized_Recommendation_Frequency        602 non-null    object
6   Browsing_Frequency                           602 non-null    object
7   Product_Search_Method                        600 non-null    object
8   Search_Result_Exploration                    602 non-null    object
9   Customer_Reviews_Importance                  602 non-null    int64
10  Add_to_Cart_Browsing                         602 non-null    object
11  Cart_Completion_Frequency                    602 non-null    object
12  Cart_Abandonment_Factors                     602 non-null    object
13  Saveforlater_Frequency                       602 non-null    object
14  Review_Left                                   602 non-null    object
15  Review_Reliability                           602 non-null    object
16  Review_Helpfulness                           602 non-null    object
17  Personalized_Recommendation_Frequency        602 non-null    int64
18  Recommendation_Helpfulness                    602 non-null    object
19  Rating_Accuracy                             602 non-null    int64
20  Shopping_Satisfaction                         602 non-null    int64
21  Service_Appreciation                         602 non-null    object
22  Improvement_Areas                            602 non-null    object
dtypes: int64(5), object(18)
memory usage: 108.3+ KB
```

In [7]: df.columns

```
Out[7]: Index(['Timestamp', 'age', 'Gender', 'Purchase_Frequency',
               'Purchase_Categories', 'Personalized_Recommendation_Frequency',
               'Browsing_Frequency', 'Product_Search_Method',
               'Search_Result_Exploration', 'Customer_Reviews_Importance',
               'Add_to_Cart_Browsing', 'Cart_Completion_Frequency',
               'Cart_Abandonment_Factors', 'Saveforlater_Frequency', 'Review_Left',
               'Review_Reliability', 'Review_Helpfulness',
               'Personalized_Recommendation_Frequency ', 'Recommendation_Helpfulness',
               'Rating_Accuracy ', 'Shopping_Satisfaction', 'Service_Appreciation',
               'Improvement_Areas'],
              dtype='object')
```

```
In [8]: df.isnull()
```

```
Out[8]:
```

	Timestamp	age	Gender	Purchase_Frequency	Purchase_Categories	Personalized_Recomr
0	False	False	False	False	False	
1	False	False	False	False	False	
2	False	False	False	False	False	
3	False	False	False	False	False	
4	False	False	False	False	False	
...
597	False	False	False	False	False	
598	False	False	False	False	False	
599	False	False	False	False	False	
600	False	False	False	False	False	
601	False	False	False	False	False	

602 rows × 23 columns



```
In [9]: df.isnull().sum()
```

```
Out[9]: Timestamp      0
age      0
Gender    0
Purchase_Frequency    0
Purchase_Categories    0
Personalized_Recommendation_Frequency    0
Browsing_Frequency    0
Product_Search_Method    2
Search_Result_Exploration    0
Customer_Reviews_Importance    0
Add_to_Cart_Browsing    0
Cart_Completion_Frequency    0
Cart_Abandonment_Factors    0
Saveforlater_Frequency    0
Review_Left    0
Review_Reliability    0
Review_Helpfulness    0
Personalized_Recommendation_Frequency    0
Recommendation_Helpfulness    0
Rating_Accuracy    0
Shopping_Satisfaction    0
Service_Appreciation    0
Improvement_Areas    0
dtype: int64
```

In [10]: `df.describe()`

Out[10]:

	age	Customer_Reviews_Importance	Personalized_Recommendation_Frequency	Rati
count	602.000000	602.000000	602.000000	
mean	30.790698	2.480066	2.699336	
std	10.193276	1.185226	1.042028	
min	3.000000	1.000000	1.000000	
25%	23.000000	1.000000	2.000000	
50%	26.000000	3.000000	3.000000	
75%	36.000000	3.000000	3.000000	
max	67.000000	5.000000	5.000000	

In [12]: `df.Service_Appreciation.value_counts()`

Out[12]:

Service_Appreciation	
Product recommendations	185
Competitive prices	182
Wide product selection	150
User-friendly website/app interface	80
.	1
Customer service	1
Customer service	1
Quick delivery	1
All the above	1

Name: count, dtype: int64

```
In [13]: df.Purchase_Categories.value_counts()
```

```
Out[13]: Purchase_Categories
Beauty and Personal Care
106
Clothing and Fashion
106
others
48
Beauty and Personal Care;Clothing and Fashion
46
Beauty and Personal Care;Clothing and Fashion;Home and Kitchen
42
Groceries and Gourmet Food;Beauty and Personal Care;Clothing and Fashion;Home
and Kitchen;others      32
Clothing and Fashion;Home and Kitchen
27
Home and Kitchen
24
Beauty and Personal Care;Home and Kitchen
21
Clothing and Fashion;Home and Kitchen;others
16
Clothing and Fashion;others
14
Groceries and Gourmet Food
14
Groceries and Gourmet Food;Beauty and Personal Care;Clothing and Fashion;Home
and Kitchen      14
Beauty and Personal Care;Clothing and Fashion;others
12
Groceries and Gourmet Food;Beauty and Personal Care;Clothing and Fashion
10
Home and Kitchen;others
9
Beauty and Personal Care;Clothing and Fashion;Home and Kitchen;others
8
Beauty and Personal Care;others
7
Groceries and Gourmet Food;Beauty and Personal Care
7
Groceries and Gourmet Food;Home and Kitchen;others
6
Groceries and Gourmet Food;Clothing and Fashion
6
Groceries and Gourmet Food;Home and Kitchen
5
Beauty and Personal Care;Home and Kitchen;others
5
Groceries and Gourmet Food;Beauty and Personal Care;Home and Kitchen
4
Groceries and Gourmet Food;Clothing and Fashion;Home and Kitchen
4
Groceries and Gourmet Food;Clothing and Fashion;Home and Kitchen;others
3
Groceries and Gourmet Food;Beauty and Personal Care;others
3
Groceries and Gourmet Food;Clothing and Fashion;others
2
```


Groceries and Gourmet Food;Beauty and Personal Care;Clothing and Fashion;other
rs 1
Name: count, dtype: int64

```
In [14]: df.Purchase_Frequency.value_counts()
```

```
Out[14]: Purchase_Frequency
Few times a month      203
Less than once a month 124
Once a week           112
Once a month          107
Multiple times a week   56
Name: count, dtype: int64
```

```
In [18]: df["Shopping_Satisfaction"].max()
```

```
Out[18]: 5
```

```
In [19]: df["Shopping_Satisfaction"].min()
```

```
Out[19]: 1
```

```
In [23]: df[["age", "Gender"]]
```

```
Out[23]:
```

	age	Gender
0	23	Female
1	23	Female
2	24	Prefer not to say
3	24	Female
4	22	Female
...
597	23	Female
598	23	Female
599	23	Female
600	23	Female
601	23	Female

602 rows × 2 columns

```
In [24]: df[df["Purchase_Categories"]=="Beauty and Personal Care"]
```

```
Out[24]:
```

	Timestamp	age	Gender	Purchase_Frequency	Purchase_Categories	Personalized_Recomm
0	2023/06/04 1:28:19 PM GMT+5:30	23	Female	Few times a month	Beauty and Personal Care	
14	2023/06/04 8:02:28 PM GMT+5:30	21	Female	Less than once a month	Beauty and Personal Care	
19	2023/06/04 8:48:59 PM GMT+5:30	25	Female	Less than once a month	Beauty and Personal Care	
27	2023/06/04 9:29:59 PM GMT+5:30	23	Female	Less than once a month	Beauty and Personal Care	
59	2023/06/05 1:32:57 PM GMT+5:30	25	Female	Once a month	Beauty and Personal Care	
...	
594	2023/06/12 3:59:10 PM GMT+5:30	23	Female	Once a month	Beauty and Personal Care	
595	2023/06/12 3:59:59 PM GMT+5:30	23	Female	Few times a month	Beauty and Personal Care	
596	2023/06/12 4:00:56 PM GMT+5:30	25	Female	Once a week	Beauty and Personal Care	
597	2023/06/12 4:02:02 PM GMT+5:30	23	Female	Once a week	Beauty and Personal Care	
599	2023/06/12 4:03:59 PM GMT+5:30	23	Female	Once a month	Beauty and Personal Care	

106 rows × 23 columns



```
In [25]: g=df.groupby('Purchase_Categories')
g
```

```
Out[25]: <pandas.core.groupby.generic.DataFrameGroupBy object at 0x0000028B4E4BC250>
```

```
In [26]: for Purchase_Categories, Purchase_Categories_df in g:
          print(Purchase_Categories)
          print(Purchase_Categories_df)
```

```
595          Competitive prices      Customer service responsiveness
596          Product recommendations      Customer service responsiveness
597          Competitive prices      Customer service responsiveness
599          Wide product selection      Product quality and accuracy
```

```
[106 rows x 23 columns]
```

```
Beauty and Personal Care;Clothing and Fashion
```

	Timestamp	age	Gender
4	2023/06/04 5:28:06 PM GMT+5:30	22	Female
7	2023/06/04 7:13:12 PM GMT+5:30	21	Female
8	2023/06/04 7:23:21 PM GMT+5:30	20	Female
9	2023/06/04 7:33:12 PM GMT+5:30	23	Female
12	2023/06/04 8:00:11 PM GMT+5:30	23	Female
26	2023/06/04 9:22:58 PM GMT+5:30	22	Female
44	2023/06/05 12:57:05 AM GMT+5:30	26	Female
45	2023/06/05 2:56:10 AM GMT+5:30	24	Female
52	2023/06/05 10:37:36 AM GMT+5:30	24	Female
57	2023/06/05 1:27:08 PM GMT+5:30	23	Female
64	2023/06/05 1:45:30 PM GMT+5:30	24	Female
70	2023/06/05 2:24:22 PM GMT+5:30	24	Female

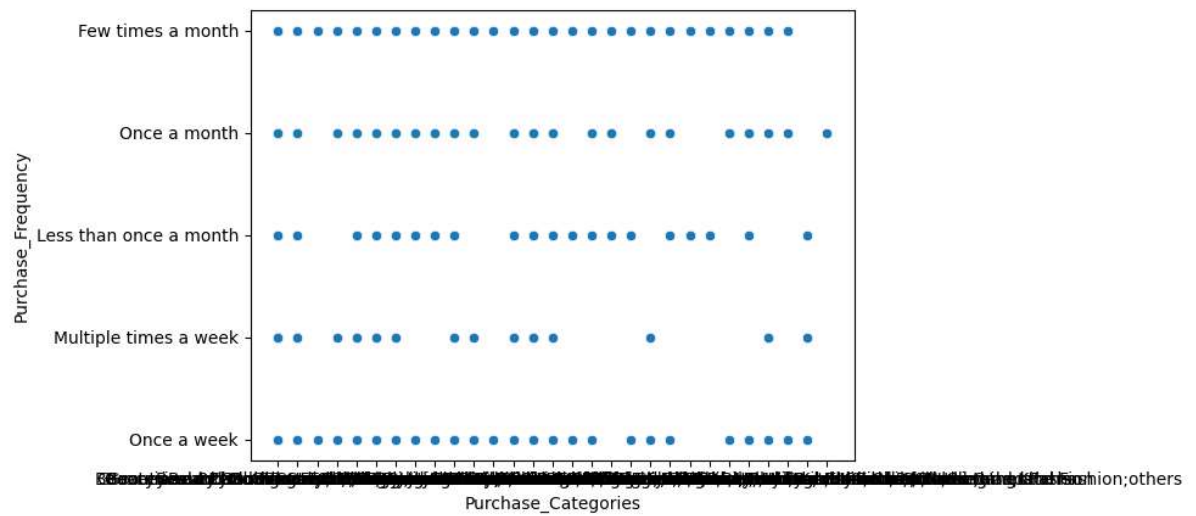
```
In [29]: %matplotlib inline
          g.plot()
```

C:\Users\sampr\AppData\Local\Programs\Python\Python311\Lib\site-packages\pandas\plotting_matplotlib\core.py:509: RuntimeWarning: More than 20 figures have been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`). Consider using `matplotlib.pyplot.close()`.

```
fig = self.plt.figure(figsize=self.figsize)
```

```
Out[29]: Purchase_Categories
Beauty and Personal Care
Axes(0.125,0.11;0.775x0.77)
Beauty and Personal Care;Clothing and Fashion
Axes(0.125,0.11;0.775x0.77)
Beauty and Personal Care;Clothing and Fashion;Home and Kitchen
Axes(0.125,0.11;0.775x0.77)
Beauty and Personal Care;Clothing and Fashion;Home and Kitchen;others
Axes(0.125,0.11;0.775x0.77)
Beauty and Personal Care;Clothing and Fashion;others
Axes(0.125,0.11;0.775x0.77)
Beauty and Personal Care;Home and Kitchen
```

```
In [33]: df=pd.read_csv("Amazon Customer Behavior Survey.csv")
sns.scatterplot(x="Purchase_Categories",y="Purchase_Frequency",data=df)
plt.show()
```



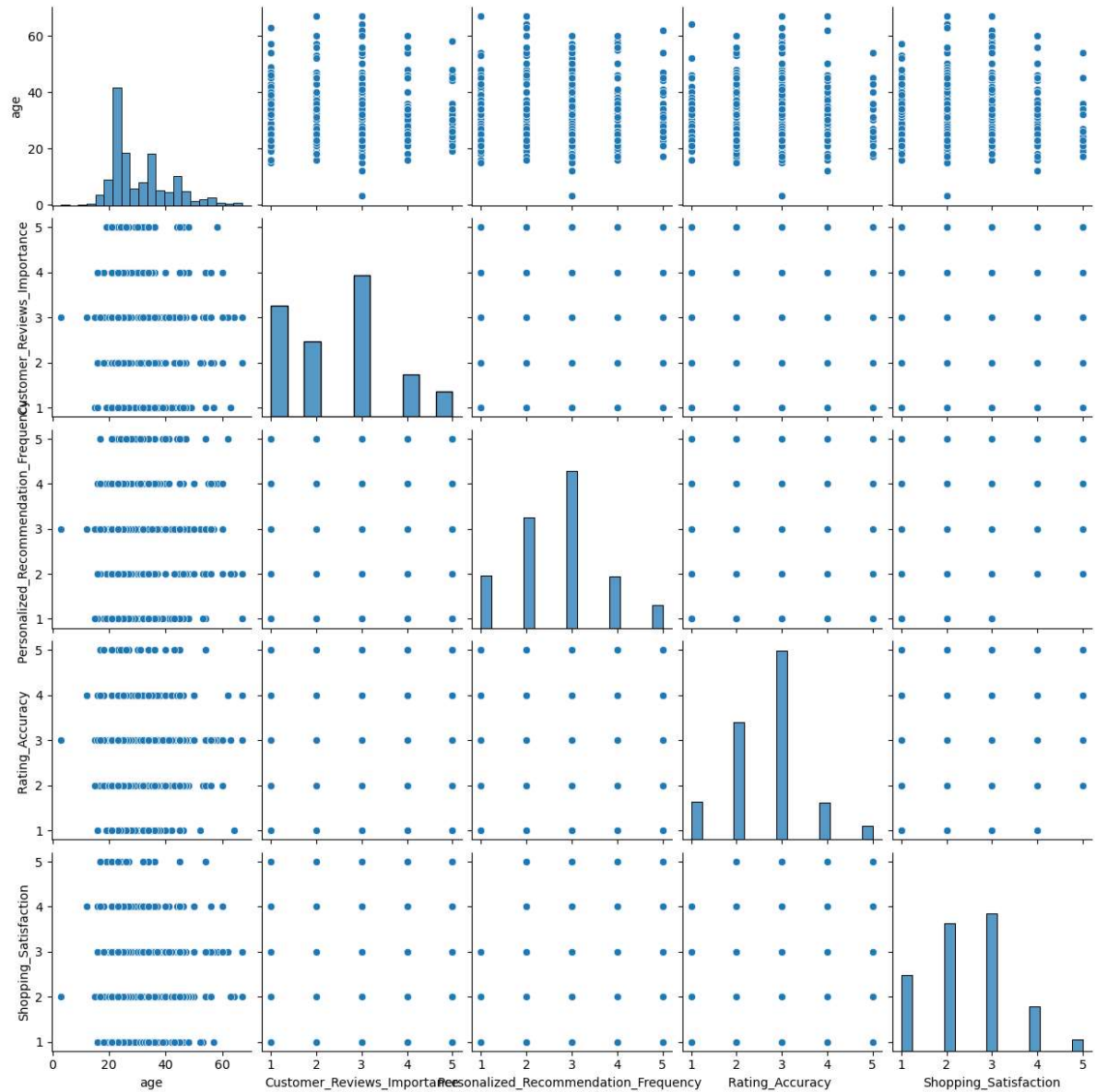
```
In [37]: var=pd.read_csv("Amazon Customer Behavior Survey.csv").head(10)
var
```

Out[37]:

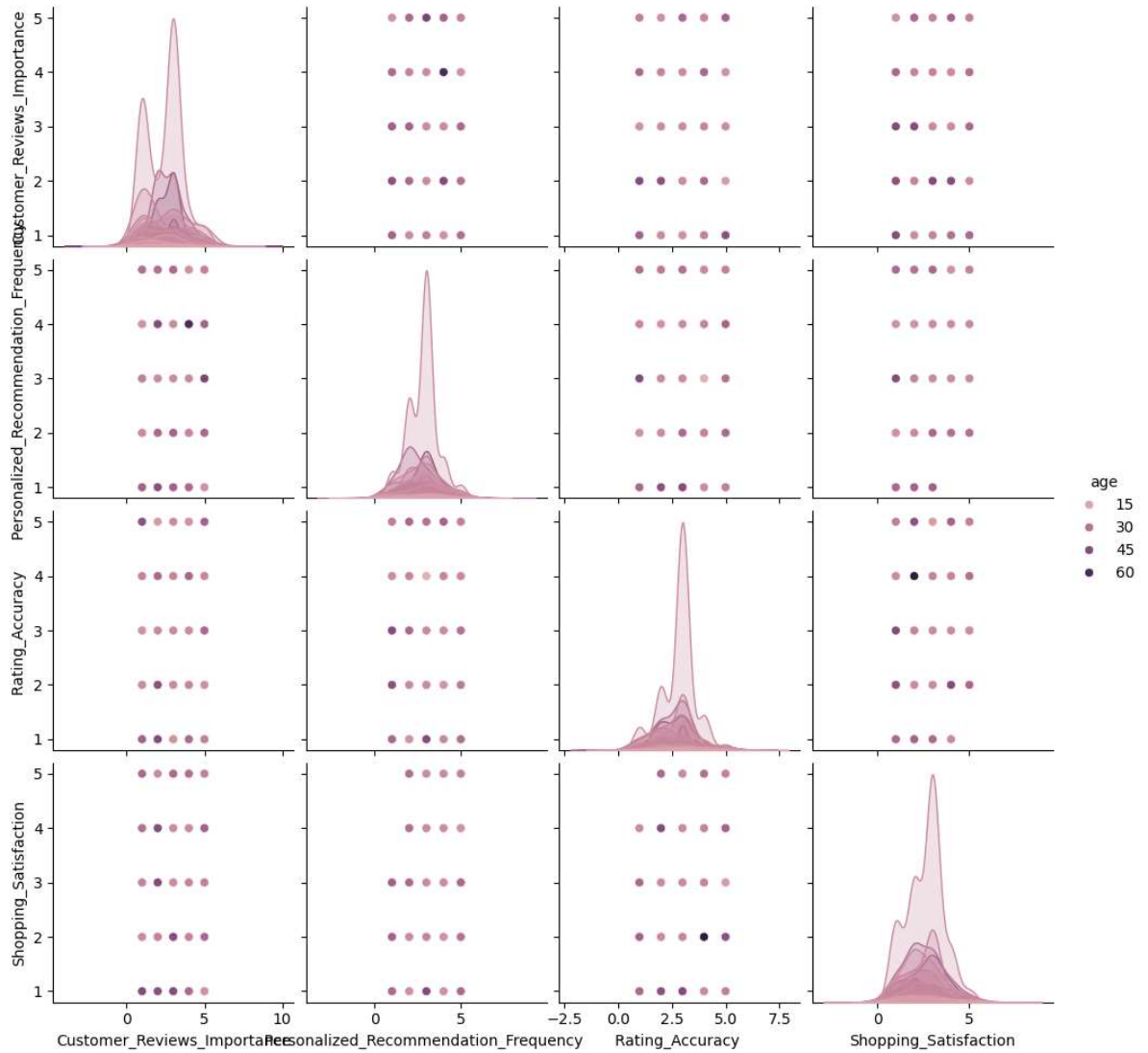
	Timestamp	age	Gender	Purchase_Frequency	Purchase_Categories	Personalized_Recommen
0	2023/06/04 1:28:19 PM GMT+5:30	23	Female	Few times a month	Beauty and Personal Care	
1	2023/06/04 2:30:44 PM GMT+5:30	23	Female	Once a month	Clothing and Fashion	
2	2023/06/04 5:04:56 PM GMT+5:30	24	Prefer not to say	Few times a month	Groceries and Gourmet Food;Clothing and Fashion	
3	2023/06/04 5:13:00 PM GMT+5:30	24	Female	Once a month	Beauty and Personal Care;Clothing and Fashion;...	
4	2023/06/04 5:28:06 PM GMT+5:30	22	Female	Less than once a month	Beauty and Personal Care;Clothing and Fashion	
5	2023/06/04 6:01:59 PM GMT+5:30	21	Female	Less than once a month	Clothing and Fashion	
6	2023/06/04 6:31:41 PM GMT+5:30	22	Female	Less than once a month	Clothing and Fashion	
7	2023/06/04 7:13:12 PM GMT+5:30	21	Female	Few times a month	Beauty and Personal Care;Clothing and Fashion	
8	2023/06/04 7:23:21 PM GMT+5:30	20	Female	Less than once a month	Beauty and Personal Care;Clothing and Fashion	
9	2023/06/04 7:33:12 PM GMT+5:30	23	Female	Less than once a month	Beauty and Personal Care;Clothing and Fashion	

10 rows × 23 columns

```
In [46]: sns.pairplot(df)
plt.show()
```



```
In [48]: sns.pairplot(df, hue="age")  
plt.show()
```



```
In [ ]:
```