In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
# to ignore warnings
import warnings
warnings.filterwarnings('ignore')
```

In [2]:
```
# Reading Dataset
data = pd.read_excel('UserEngageme
data.head()
```

Out[2]:

	Timestamp	By clicking 'I agree' below, you are indicating that you have read and understood this consent form and agree to participate in this research study. If you are under 18 years old, you confirm that you have obtained parental or guardian consent to participate.	What is your age?	What is your gender?
0	2024/12/24 11:33:49 AM GMT+8	Agree	Under 18	Female
1	2024/12/24 11:33:49 AM GMT+8	Agree	25 - 34	Female
2	2024/12/24 11:33:49 AM GMT+8	Agree	55 - 64	Female
3	2024/12/24 11:33:49 AM GMT+8	Agree	55 - 64	Female
4	2024/12/24 11:33:49 AM GMT+8	Agree	65 and above	Male

In [3]:
```
# Data Cleaning And Preparation

# Dropping columns by index positi
columns_to_drop = [0, 1]  # Index
data = data.drop(data.columns[colu


data.head()
```

Out[3]:

	What is your age?	What is your gender?	What is your occupation?	How often do you use the internet?
0	Under 18	Female	Employed	Several times a month
1	25 - 34	Female	Self-employed	Rarely
2	55 - 64	Female	Retired	Once a month
3	55 - 64	Female	Unemployed	Rarely
4	65 and above	Male	Unemployed	Several times a month

In [4]:
```
# Retrieving column names
data.columns
```

Out[4]:
```
Index(['What is your age?', 'What
is your gender?', 'What is your oc
cupation?',
       'How often do you use the i
nternet?',
       'Which device do you primar
ily use to access the internet?',
       'What is your main reason f
or using the internet?',
       'What types of Malaysian we
b platforms do you frequently visi
t?',
       'How often do you interact
with online shopping on website
s?',
       'How often do you interact
with reading articles on website
s?',
       'How often do you interact
with watching videos on website
s?',
       'How often do you interact
with participating in online discu
ssions on websites?',
       'How often do you interact
with social media integration on w
ebsites?',
       'How satisfied are you with
the overall user experience of Mal
aysian web platforms? (Rate on a s
cale of 1-5, where 1 is 'Very Diss
atisfied' and 5 is 'Very Satisfie
d') ',
       'How likely are you to reco
mmend Malaysian web platforms to o
thers? (Rate on a scale of 1-5, wh
ere 1 is 'Very Unlikely' and 5 is
'Very Likely')',
       'How accessible do you find
Malaysian web platforms for people
with disabilities? (Rate on a scal
e of 1-5, where 1 is 'Not Accessib
le at All' and 5 is 'Very Accessib
le')',
       'How successful are you at
finding what you need on Malaysian
web platforms? (Rate on a scale of
1-5, where 1 is 'Never Successful'
and 5 is 'Always Successful')',
       'What improvements would yo
u suggest for Malaysian web platfo
rms to enhance your experience? ',
       'Are there any additional f
eatures you would like to see impl
emented on Malaysian web platform
s?'],
      dtype='object')
```

In [5]:
```
# Renaming columns
data.rename(columns={
    'What is your age?': 'age',
    'What is your gender?': 'gende
    'What is your occupation?': 'o
    'How often do you use the inte
    'Which device do you primarily
    'What is your main reason for
    'What types of Malaysian web p
    'How often do you interact wit
    'How often do you interact wit
    'How often do you interact wit
    'How often do you interact wit
    'How often do you interact wit
    "How satisfied are you with th
    'How likely are you to recomme
    'How accessible do you find Ma
    'How successful are you at fin
    "What improvements would you s
    'Are there any additional feat
}, inplace=True)

data.head()
```

Out[5]:

	age	gender	occupation	internet_usa
0	Under 18	Female	Employed	Several times mon
1	25 - 34	Female	Self-employed	Rar
2	55 - 64	Female	Retired	Once a mon
3	55 - 64	Female	Unemployed	Rar
4	65 and above	Male	Unemployed	Several times mon

In [6]:
```
# Replacing column values in a pan
data.age.value_counts()
```

Out[6]:
```
25 - 34         31
45 - 54         27
55 - 64         26
18 - 24         20
Under 18        18
35 - 44         18
65 and above    11
Name: age, dtype: int64
```

In [7]:
```
data['age'] = data['age'].map({'Un
                               '45
                               '55
                               '65
```

In [8]:
```
data.age.value_counts().sort_index
```

Out[8]:
```
0-18      18
18-24     20
25-34     31
35-44     18
45-54     27
55-64     26
65-100    11
Name: age, dtype: int64
```

In [9]:
```
data['online_shopping'].value_coun
```

Out[9]:
```
4             42
2             38
3             36
1 (Never)     18
5 (Always)    17
Name: online_shopping, dtype: int6
4
```

In [10]:
```
data['online_shopping'].map(type)
```

Out[10]:
```
0      <class 'int'>
1      <class 'int'>
2      <class 'int'>
3      <class 'int'>
4      <class 'int'>
           ...
146    <class 'int'>
147    <class 'str'>
148    <class 'int'>
149    <class 'str'>
150    <class 'int'>
Name: online_shopping, Length: 15
1, dtype: object
```

In [11]:
```
data['online_shopping'] = data['on
data['reading_articles'] = data['r
data['for_watching_videos'] = data
data['social_media_integration'] =
```

In [12]:
```
data.head()
```

Out[12]:

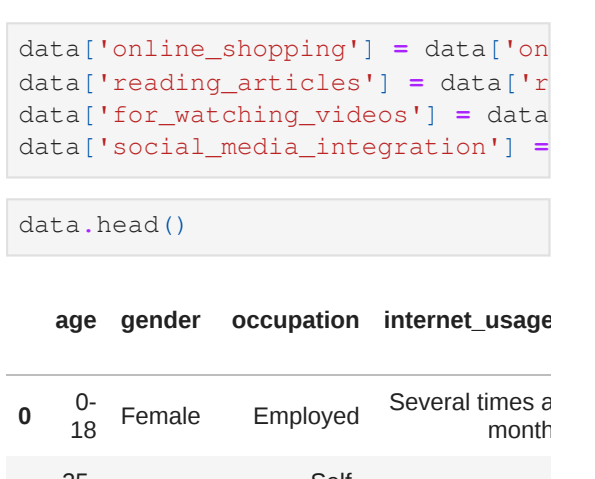
	age	gender	occupation	internet_usage
0	0-18	Female	Employed	Several times a month
1	25-34	Female	Self-employed	Rarely
2	55-64	Female	Retired	Once a month
3	55-64	Female	Unemployed	Rarely
4	65-100	Male	Unemployed	Several times a month

In [13]:
```
# Dealing With Missing Data
# Sorting Percentage Of Null Value
(data.isnull().sum()/len(data)*100
```

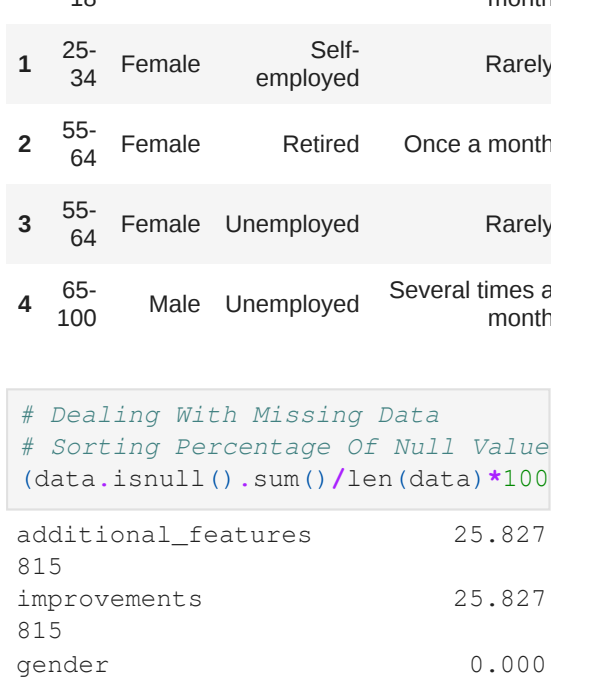
Out[13]:
```
additional_features           25.827
815
improvements                  25.827
815
gender                         0.000
000
rating_for_finding_needs       0.000
000
accessiblity_for_disable       0.000
000
will_recommend                 0.000
000
overall_experience             0.000
000
social_media_integration       0.000
000
for_online_discussions         0.000
000
age                            0.000
000
reading_articles               0.000
000
online_shopping                0.000
000
Malaysian web platforms        0.000
000
reason                         0.000
000
device used                    0.000
000
internet_usage                 0.000
000
occupation                     0.000
000
for_watching_videos            0.000
000
dtype: float64
```

In [14]:
```
# Data analysis using Data Visuali

# Responses based on age
age = data['age'].value_counts(asc
ax = sns.barplot(x=age.index, y=ag
ax.set(xlabel='AGE', yticks=[], ti
ax.tick_params(length=0)
ax.bar_label(ax.containers[0])
plt.show()
```



In [15]:
```
#Response based on gender
colors = ['#81BFDA', '#F72C5B']
data.groupby('gender').size().plot
plt.title('Responses Based On Gend
plt.show()
```

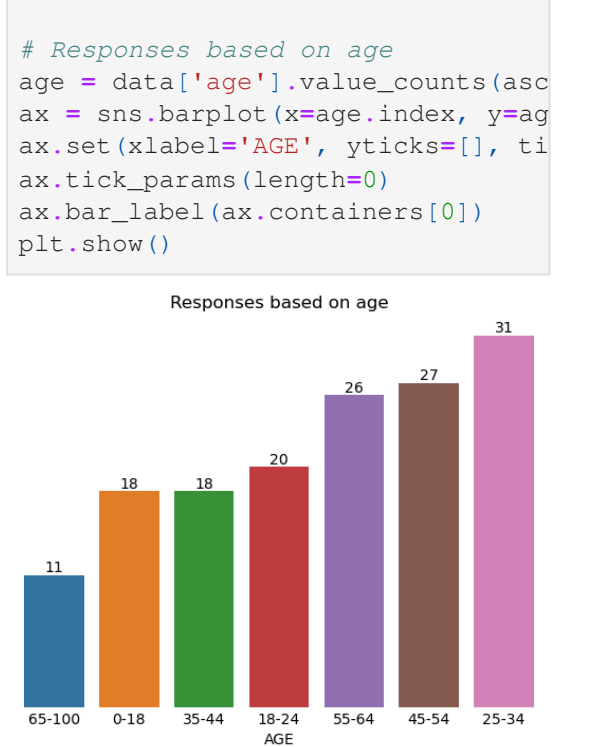


In [16]:
```
#User Occupation
colors = ['#A888B5','#6A669D']
sns.countplot(x='occupation',hue =
plt.title('User Occupation')
sns.color_palette()
plt.show()
```

In [17]:
```
# How often user's use the interne
data['internet_usage'].value_count
```

Out[17]:
```
Several times a month    37
Once a month             35
Once a week              29
Several times a week     19
Rarely                   17
Daily                    14
Name: internet_usage, dtype: int64
```

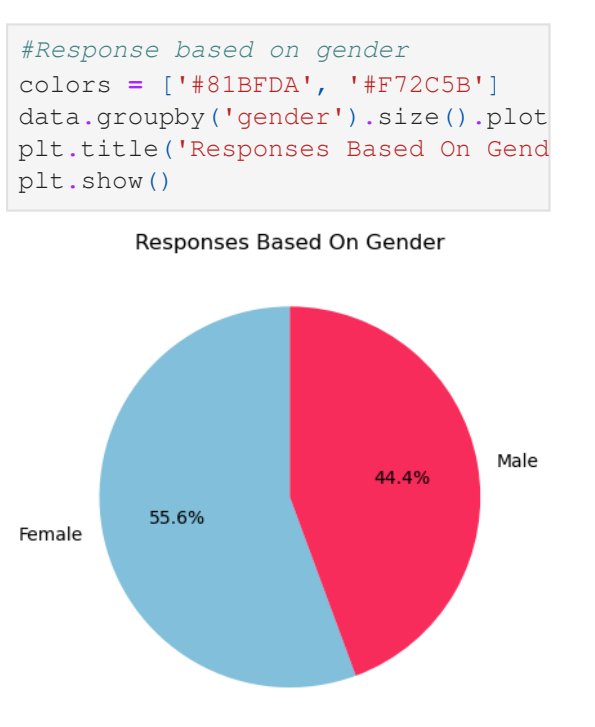
In [18]:
```
sns.countplot(x= 'internet_usage'
plt.title('Internet usage')
plt.xticks(rotation=-45)
plt.show()
```



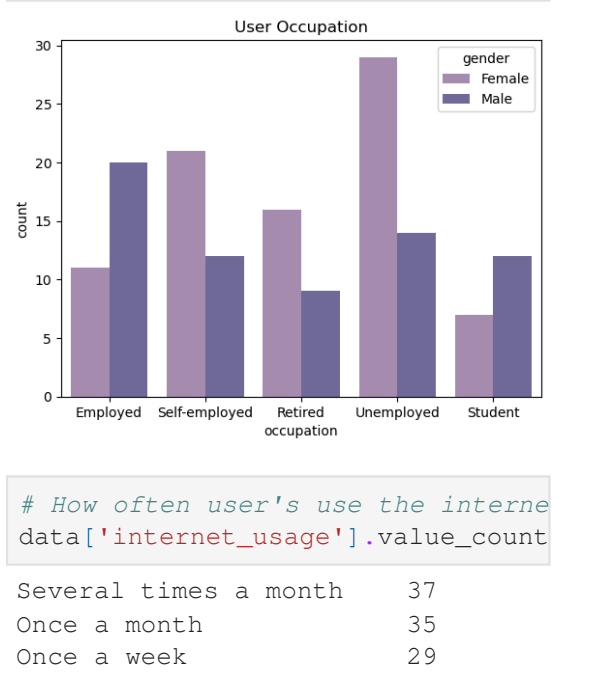
In [19]:
```
data['device used'].value_counts()
```

Out[19]:
```
Laptop        61
Tablet        43
Desktop       27
Smartphone    20
Name: device used, dtype: int64
```

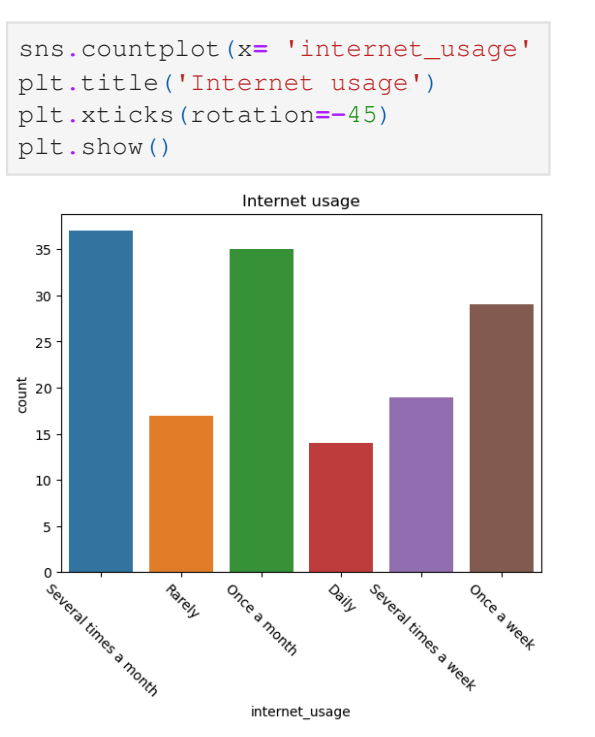
In [20]:
```
# Devices used while using interne
Slices = ['Desktop','Tablet','Smar
colors = ['#A35C7A','#155E95','#44
data.groupby('device used').size()
plt.title('Devices used while usin
plt.show()
```



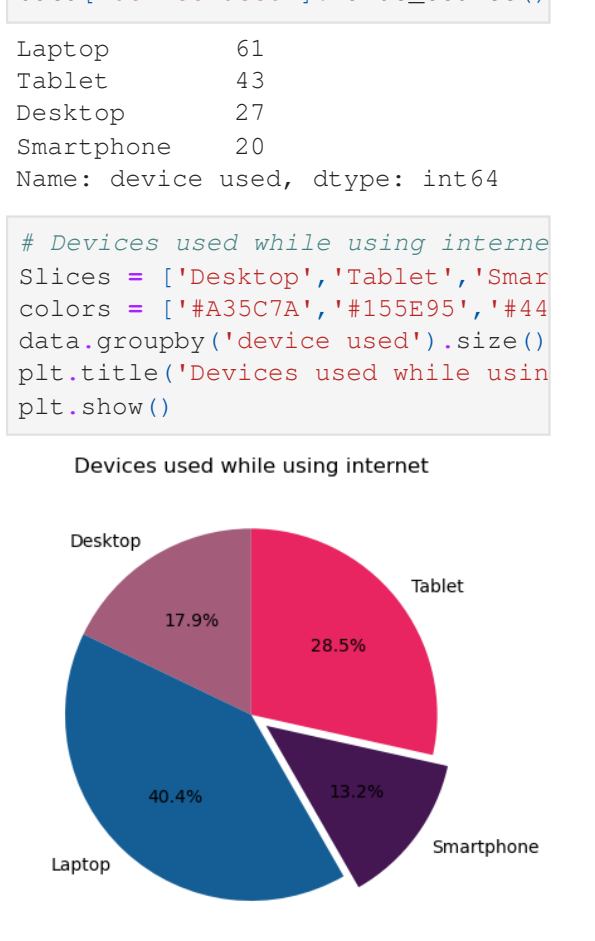
In [21]:
```
# User Responses
fig, axarr = plt.subplots(2,2, fig
data['online_shopping'].value_coun
axarr[0][0].set_title(" Online Sho
data['reading_articles'].value_cou
axarr[0][1].set_title(" Reading Ar
data['for_watching_videos'].value_
axarr[1][0].set_title("For Watchin
data['for_online_discussions'].val
axarr[1][1].set_title("For Online
plt.subplots_adjust(hspace=1.0)
plt.subplots_adjust(wspace=.5)
sns.despine()
```



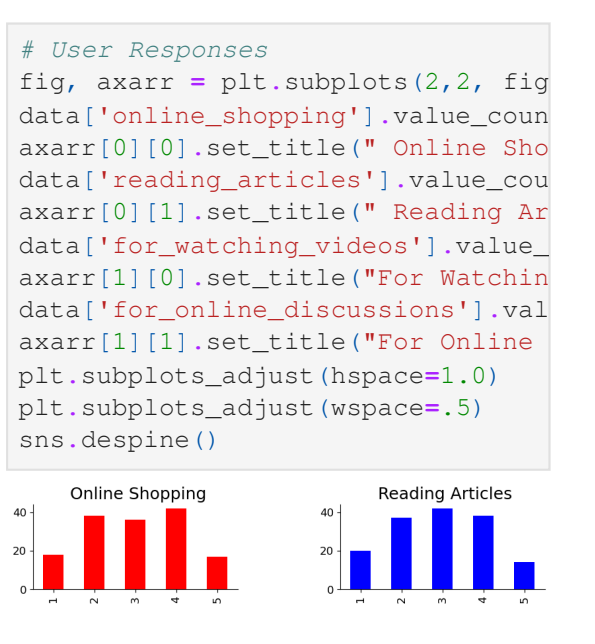
In [22]:
```
# user response for social media
sns.countplot(x= 'social_media_int
plt.title('social_media_integratio
plt.xticks(rotation=-45)
plt.show()
```



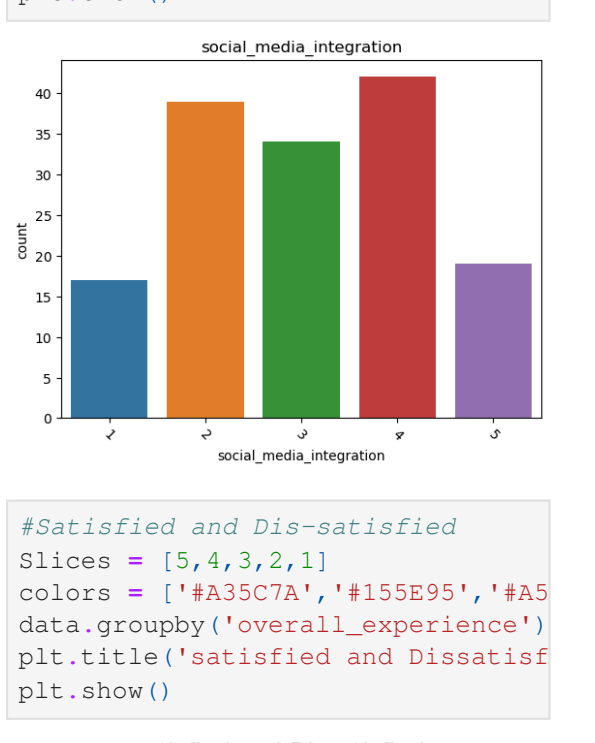
In [31]:
```
#Satisfied and Dis-satisfied
Slices = [5,4,3,2,1]
colors = ['#A35C7A','#155E95','#A5
data.groupby('overall_experience').
plt.title('satisfied and Dissatisf
plt.show()
```



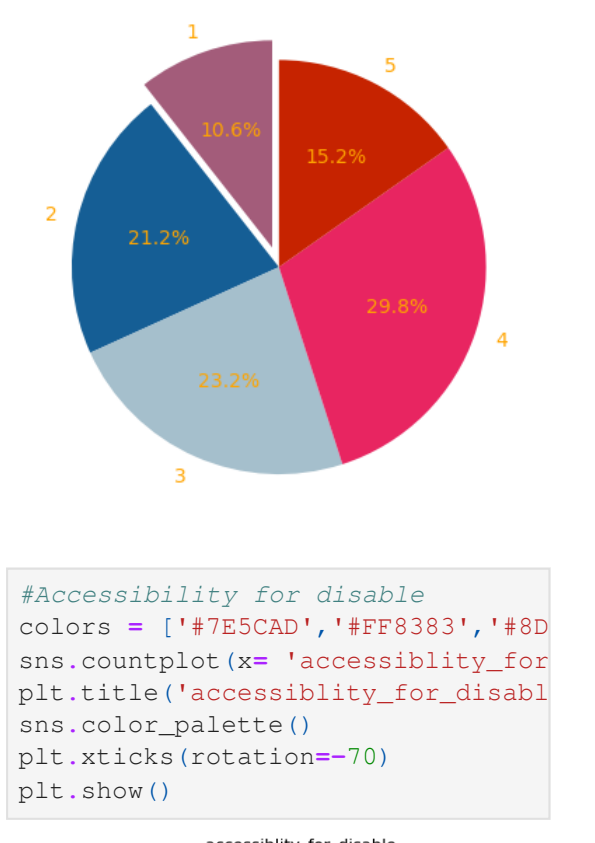
In [24]:
```
#Accessibility for disable
colors = ['#7E5CAD','#F8383','#8D
sns.countplot(x = 'accessiblity_for
plt.title('accessiblity_for_disabl
sns.color_palette()
plt.xticks(rotation=-70)
plt.show()
```



In [25]:
```
# Improvements
sns.countplot(x= 'improvements' ,
plt.title('improvements')
plt.xticks(rotation=-45)
plt.show()
```



In [30]:
```
#Additional Features
colors = ['c','m','y','r','#72BAA9
data.groupby('additional_features'
plt.title('Additional features')
plt.show()
```



In [42]:
```
fig, axarr = plt.subplots(1,2, fig
data['will_recommend'].value_coun
axarr[0].set_title(" Will recommend
data['rating_for_finding_needs'].v
axarr[1].set_title("Rating for fin
plt.subplots_adjust(hspace=1.0)
plt.subplots_adjust(wspace=.5)
sns.despine()
```

In [ ]: