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
columns_to_drop = [0, 1]  # drop t
data = data.drop(data.columns[colu

print("\nDataset after dropping co
data.head()
```

```
Dataset after dropping columns by
index positions:
```

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 user 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 the
    '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 a mon
1	25 - 34	Female	Self-employed	Rarely
2	55 - 64	Female	Retired	Once a mon
3	55 - 64	Female	Unemployed	Rarely
4	65 and above	Male	Unemployed	Several times a 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['for_online_shopping'].value_
```

Out[9]:
```
4             42
2             38
3             36
1 (Never)     18
5 (Always)    17
Name: for_online_shopping, dtype:
int64
```

In [10]:
```
data['for_online_shopping'].map(ty
```

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: for_online_shopping, Length:
151, dtype: object
```

In [11]:
```
data['for_online_shopping'] = data
data['for_reading_articles'] = dat
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]:
```
# sorting percentage in each colum
(data.isnull().sum()/len(data)*100
```

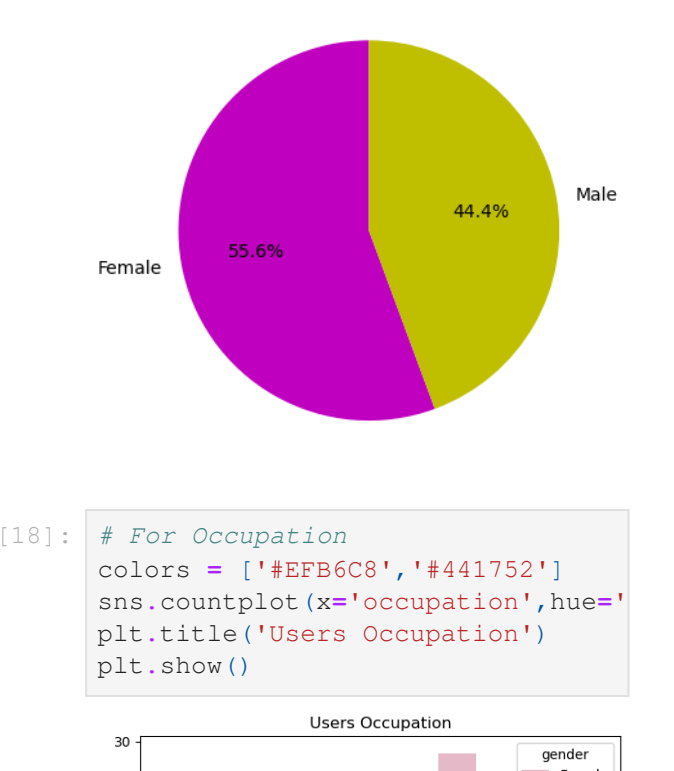
Out[13]:
```
additional_features         25.827
815
improvements                25.827
815
gender                       0.000
000
rating_for_finding_needs     0.000
000
accessiblity_for_disable     0.000
000
will_recommend               0.000
000
overall_experience           0.000
000
social_media_integration     0.000
000
for_online_discussions       0.000
000
age                          0.000
000
for_reading_articles         0.000
000
for_online_shopping          0.000
000
webplatform_types            0.000
000
reason                       0.000
000
device                       0.000
000
internet_usage               0.000
000
occupation                   0.000
000
for_watching_videos          0.000
000
dtype: float64
```

In [14]:
```
# let's grid the null value
data['additional_features'].fillna
data['improvements'].fillna(value=
```

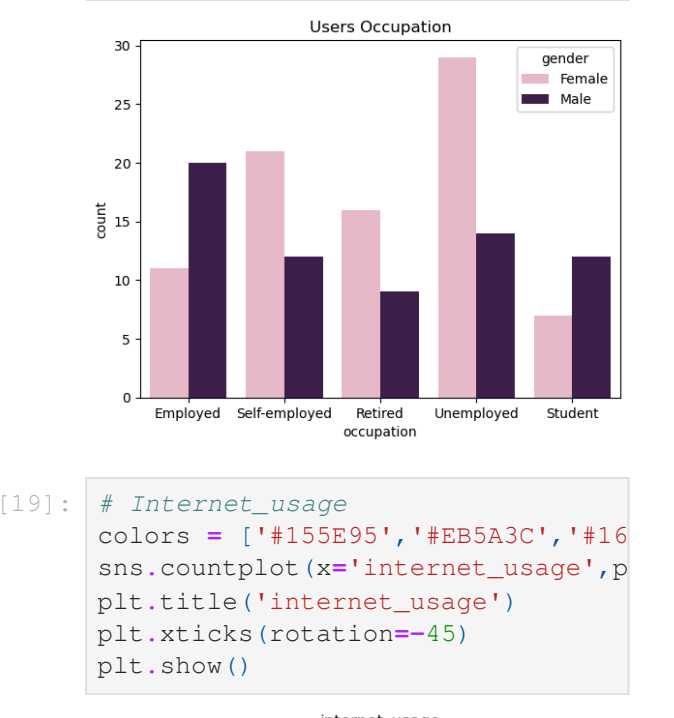
In [15]:
```
data.isnull().sum()
```

Out[15]:
```
age                         0
gender                      0
occupation                  0
internet_usage              0
device                      0
reason                      0
webplatform_types           0
for_online_shopping         0
for_reading_articles        0
for_watching_videos         0
for_online_discussions      0
social_media_integration    0
overall_experience          0
will_recommend              0
accessiblity_for_disable    0
rating_for_finding_needs    0
improvements                0
additional_features         0
dtype: int64
```

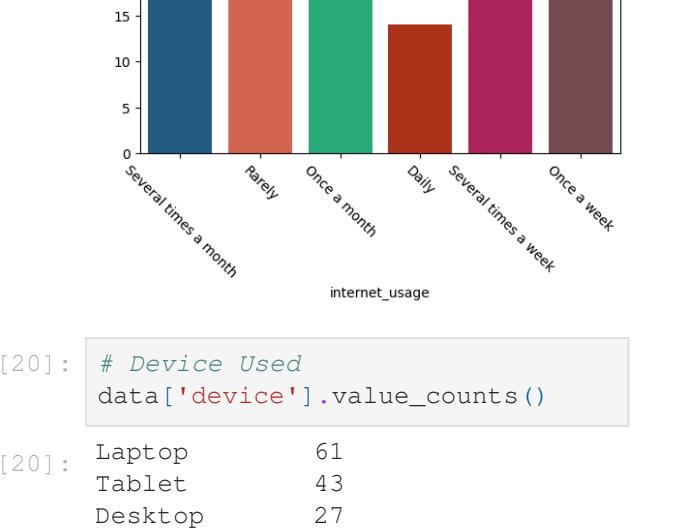
In [16]:
```
age = data['age'].value_counts(asc
ax = sns.barplot(x=age.index, y=ag
ax.set(xlabel='AGE', yticks=[], ti
ax.tick_params(length=0)
ax.bar_label(ax.containers[0])
plt.show()
```



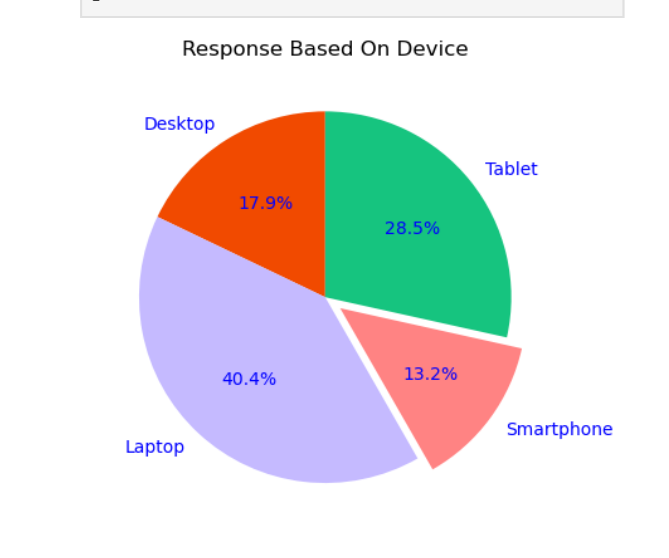
In [26]:
```
# Response Based on Gender
colors = ['m', 'y']
data.groupby('gender').size().plot
plt.title('Response Based On Gende
plt.show()
```



In [18]:
```
# For Occupation
colors = ['#EFB6C8','#441752']
sns.countplot(x='occupation',hue='
plt.title('Users Occupation')
plt.show()
```



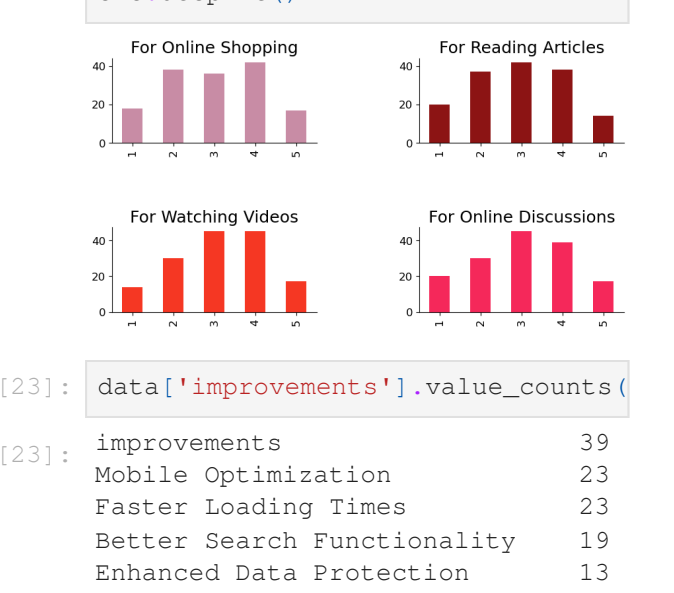
In [19]:
```
# Internet_usage
colors = ['#155E95','#EB5A3C','#16
sns.countplot(x='internet_usage',p
plt.title('internet_usage')
plt.xticks(rotation=-45)
plt.show()
```



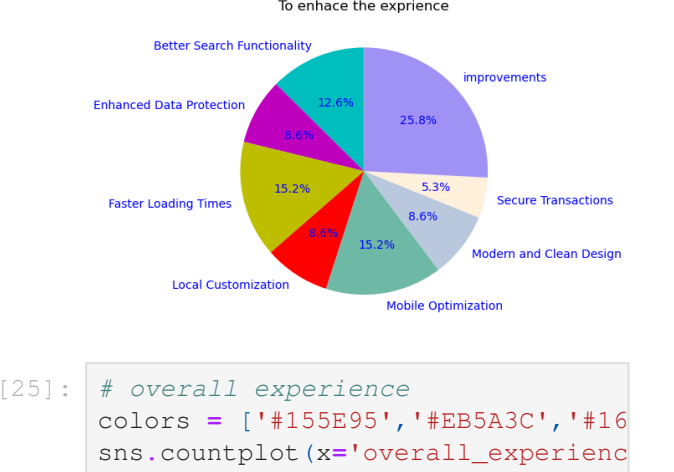
In [20]:
```
# Device Used
data['device'].value_counts()
```

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

In [21]:
```
colors = ['#F14A00','#C5BAFF','#FF
data.groupby('device').size().plot
plt.title('Response Based On Devic
plt.show()
```



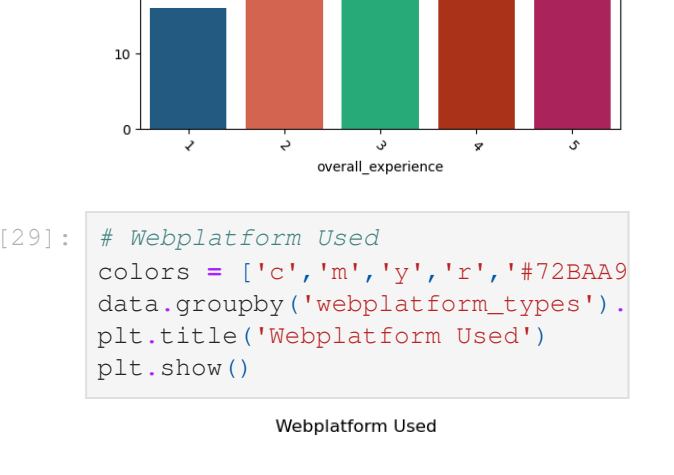
In [22]:
```
fig, axarr = plt.subplots(2, 2, fi
data['for_online_shopping'].value_
axarr[0][0].set_title("For Online
data['for_reading_articles'].value
axarr[0][1].set_title("For Reading
data['for_watching_videos'].value_
axarr[1][0].set_title("For Watchin
data['for_online_discussions'].val
axarr[1][1].set_title("For Online
plt.subplots_adjust(hspace=1.0)
plt.subplots_adjust(wspace=.5)
sns.despine()
```



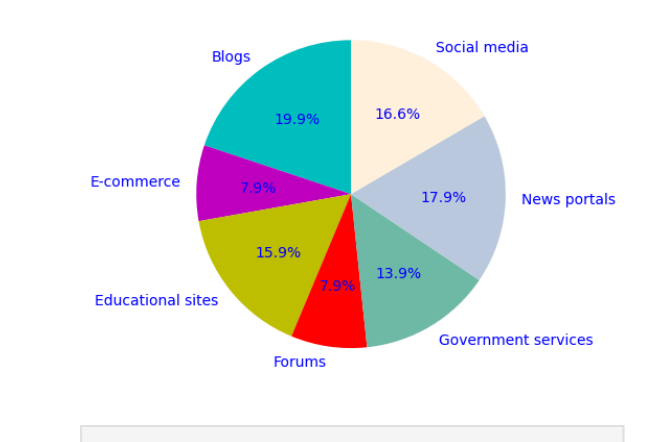
In [23]:
```
data['improvements'].value_counts(
```

Out[23]:
```
improvements                    39
Mobile Optimization             23
Faster Loading Times            23
Better Search Functionality     19
Enhanced Data Protection        13
Local Customization             13
Modern and Clean Design         13
Secure Transactions              8
Name: improvements, dtype: int64
```

In [24]:
```
# improvement
colors = ['c','m','y','r','#72BAA9
data.groupby('improvements').size(
plt.title('To enhace the exprience
plt.show()
```

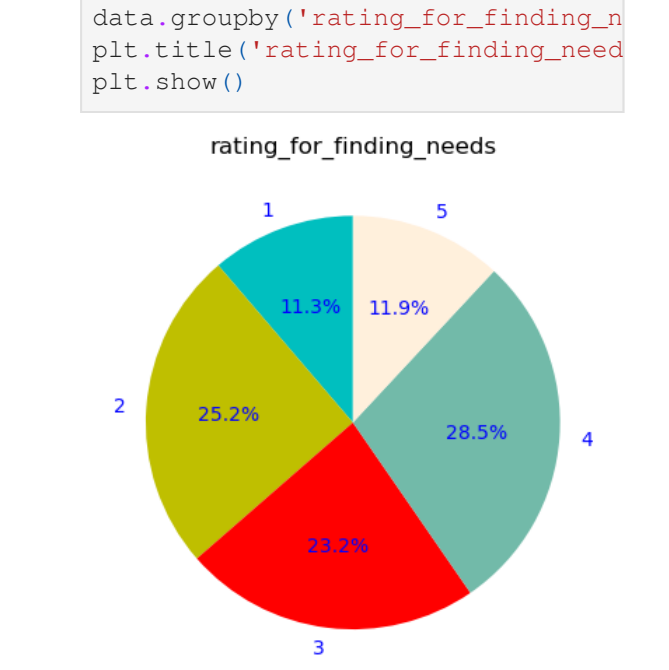


In [25]:
```
# overall experience
colors = ['#155E95','#EB5A3C','#16
sns.countplot(x='overall_experienc
plt.title('overall_experience')
plt.xticks(rotation=-45)
plt.show()
```

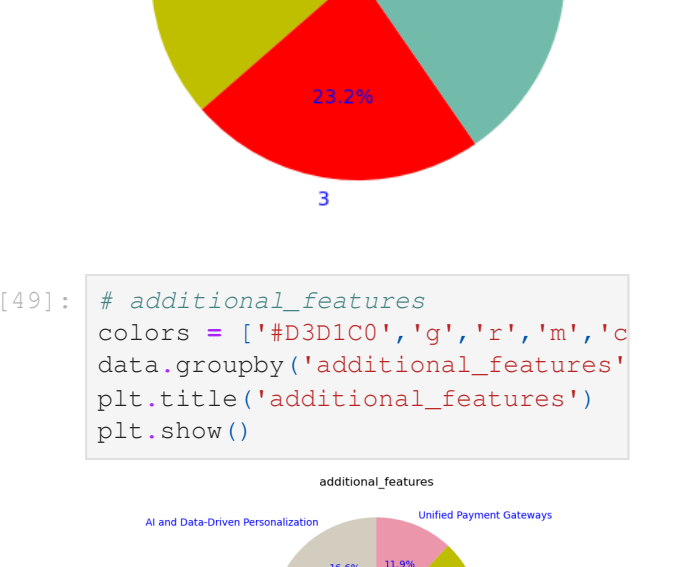


In [29]:
```
# Webplatform Used
colors = ['c','m','y','r','#72BAA9
data.groupby('webplatform_types').
plt.title('Webplatform Used')
plt.show()
```

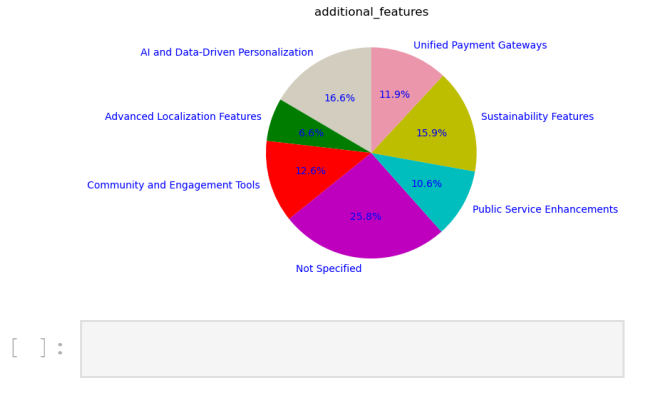
In [40]:
```
# webplatform likely recommended a
fig, axarr = plt.subplots(1, 2, fi
data['will_recommend'].value_count
axarr[0].set_title("will_recommend
data['accessiblity_for_disable'].v
axarr[1].set_title("accessiblity_f
plt.subplots_adjust(hspace=1.0)
plt.subplots_adjust(wspace=.5)
sns.despine()
```



In [45]:
```
# rating_for_finding_needs
colors = ['c','y','r','#72BAA9','#
data.groupby('rating_for_finding_n
plt.title('rating_for_finding_need
plt.show()
```



In [49]:
```
# additional_features
colors = ['#D3D1C0','g','r','m','c
data.groupby('additional_features'
plt.title('additional_features')
plt.show()
```



In [ ]: