```
In [1]: import pandas as pd
        import matplotlib
        import matplotlib.pyplot as plt
        import numpy as np
        import seaborn as sns
        %matplotlib inline
        from matplotlib import style
```

```
In [2]: kag = pd.read_csv("C://Users//kolli Mallikarjuna//OneDrive//Desktop//project data//KAG
        kag
```

Out[2]:

	ad_id	xyz_campaign_id	fb_campaign_id	age	gender	interest	Impressions	Clicks	Spe
0	708746	916	103916	30-34	M	15	7350	1	1.4300
1	708749	916	103917	30-34	M	16	17861	2	1.8200
2	708771	916	103920	30-34	M	20	693	0	0.0000
3	708815	916	103928	30-34	M	28	4259	1	1.2500
4	708818	916	103928	30-34	M	28	4133	1	1.2900
...	...	...	...	...	...	...	...	...	
1138	1314410	1178	179977	45-49	F	109	1129773	252	358.1899
1139	1314411	1178	179978	45-49	F	110	637549	120	173.8800
1140	1314412	1178	179979	45-49	F	111	151531	28	40.2899
1141	1314414	1178	179981	45-49	F	113	790253	135	198.7100
1142	1314415	1178	179982	45-49	F	114	513161	114	165.6099

1143 rows × 11 columns

```
In [3]: kag.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1143 entries, 0 to 1142
Data columns (total 11 columns):
 #   Column             Non-Null Count   Dtype
---  ------             --------------   -----
 0   ad_id              1143 non-null    int64
 1   xyz_campaign_id    1143 non-null    int64
 2   fb_campaign_id     1143 non-null    int64
 3   age                1143 non-null    object
 4   gender             1143 non-null    object
 5   interest           1143 non-null    int64
 6   Impressions        1143 non-null    int64
 7   Clicks             1143 non-null    int64
 8   Spent              1143 non-null    float64
 9   Total_Conversion   1143 non-null    int64
 10  Approved_Conversion 1143 non-null   int64
dtypes: float64(1), int64(8), object(2)
memory usage: 98.4+ KB
```

In [4]: `kag.describe()`

Out[4]:

	ad_id	xyz_campaign_id	fb_campaign_id	interest	Impressions	Clicks
count	1.143000e+03	1143.000000	1143.000000	1143.000000	1.143000e+03	1143.000000	1143.
mean	9.872611e+05	1067.382327	133783.989501	32.766404	1.867321e+05	33.390201	51.
std	1.939928e+05	121.629393	20500.308622	26.952131	3.127622e+05	56.892438	86.
min	7.087460e+05	916.000000	103916.000000	2.000000	8.700000e+01	0.000000	0.
25%	7.776325e+05	936.000000	115716.000000	16.000000	6.503500e+03	1.000000	1.
50%	1.121185e+06	1178.000000	144549.000000	25.000000	5.150900e+04	8.000000	12.
75%	1.121804e+06	1178.000000	144657.500000	31.000000	2.217690e+05	37.500000	60.
max	1.314415e+06	1178.000000	179982.000000	114.000000	3.052003e+06	421.000000	639.

In [5]: `kag.isna().sum()`

Out[5]:
```
ad_id                 0
xyz_campaign_id       0
fb_campaign_id        0
age                   0
gender                0
interest              0
Impressions           0
Clicks                0
Spent                 0
Total_Conversion      0
Approved_Conversion   0
dtype: int64
```

In [6]: `kag.shape`

Out[6]: `(1143, 11)`

In [7]: `kag.columns`

```
Out[7]:  Index(['ad_id', 'xyz_campaign_id', 'fb_campaign_id', 'age', 'gender',
                'interest', 'Impressions', 'Clicks', 'Spent', 'Total_Conversion',
                'Approved_Conversion'],
               dtype='object')
```
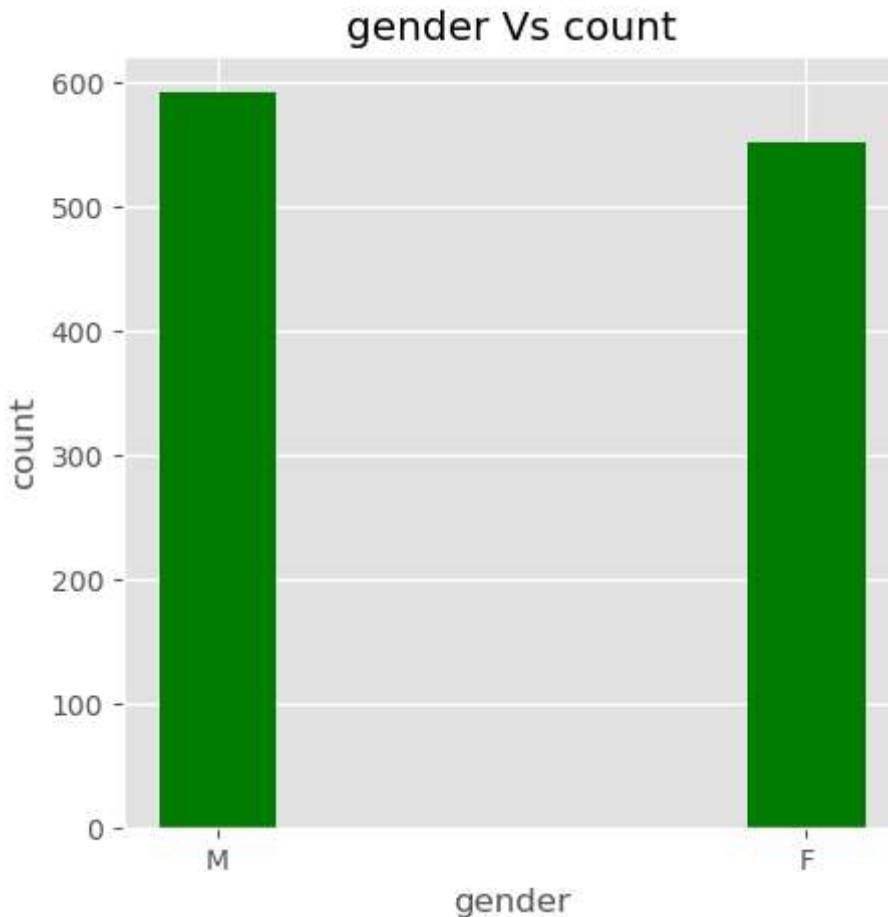
```
In [10]:  gender = kag.gender.value_counts().index
          gender
```

```
Out[10]:  Index(['M', 'F'], dtype='object', name='gender')
```

```
In [12]:  count =  kag.gender.value_counts().values
          count
```

```
Out[12]:  array([592, 551], dtype=int64)
```

```
In [25]:  style.use('ggplot')
          plt.figure(figsize=(5,5))
          plt.bar(gender,count,color = 'green', width = 0.2)
          plt.ylabel('count')
          plt.xlabel('gender')
          plt.title('gender Vs count');
```



```
In [33]:  xyz_id = kag.xyz_campaign_id.value_counts().index
          xyz_count = kag.xyz_campaign_id.value_counts().values
          xyz_id
```

```
Out[33]:  Index([1178, 936, 916], dtype='int64', name='xyz_campaign_id')
```
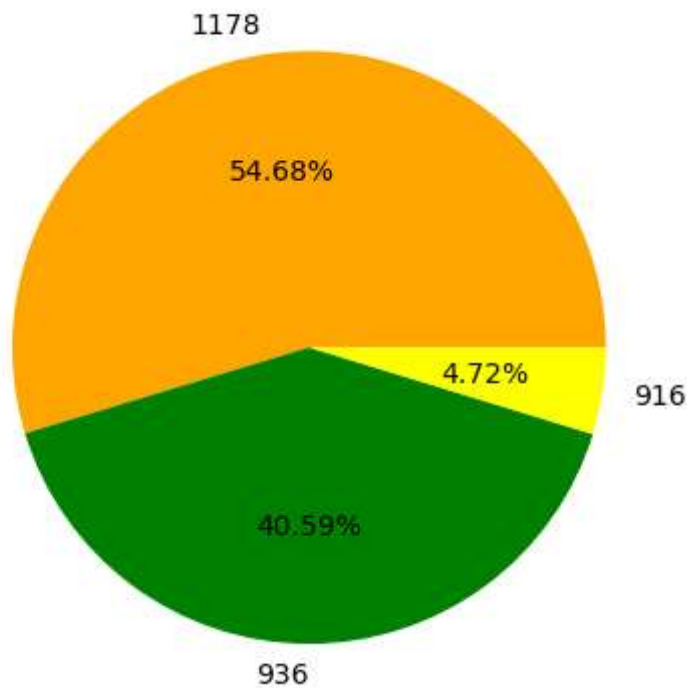
In [34]:
```python
xyz_count
```

Out[34]:
```
array([625, 464,  54], dtype=int64)
```

In [38]:
```python
plt.pie(xyz_count, labels = xyz_id, autopct = '%1.2f%%', colors = ['orange', 'green',
plt.title('xyz_Campaign_id vs count')
```

Out[38]:
```
Text(0.5, 1.0, 'xyz_Campaign_id vs count')
```

xyz_Campaign_id vs count

1178

54.68%

4.72%

916

40.59%

936

In [50]:
```python
agegrp = kag.groupby(['age','gender']).size().reset_index().rename(columns={0:'entries
agegrp
```

Out[50]:

	age	gender	entries
0	30-34	F	197
1	30-34	M	229
2	35-39	F	109
3	35-39	M	139
4	40-44	F	107
5	40-44	M	103
6	45-49	F	138
7	45-49	M	121

In [59]:
```python
matplotlib.rcParams['figure.figsize']=(8,5)
sns.barplot(x='age',y='entries',hue = 'gender', data = agegrp, errwidth = 0, width = 6
```
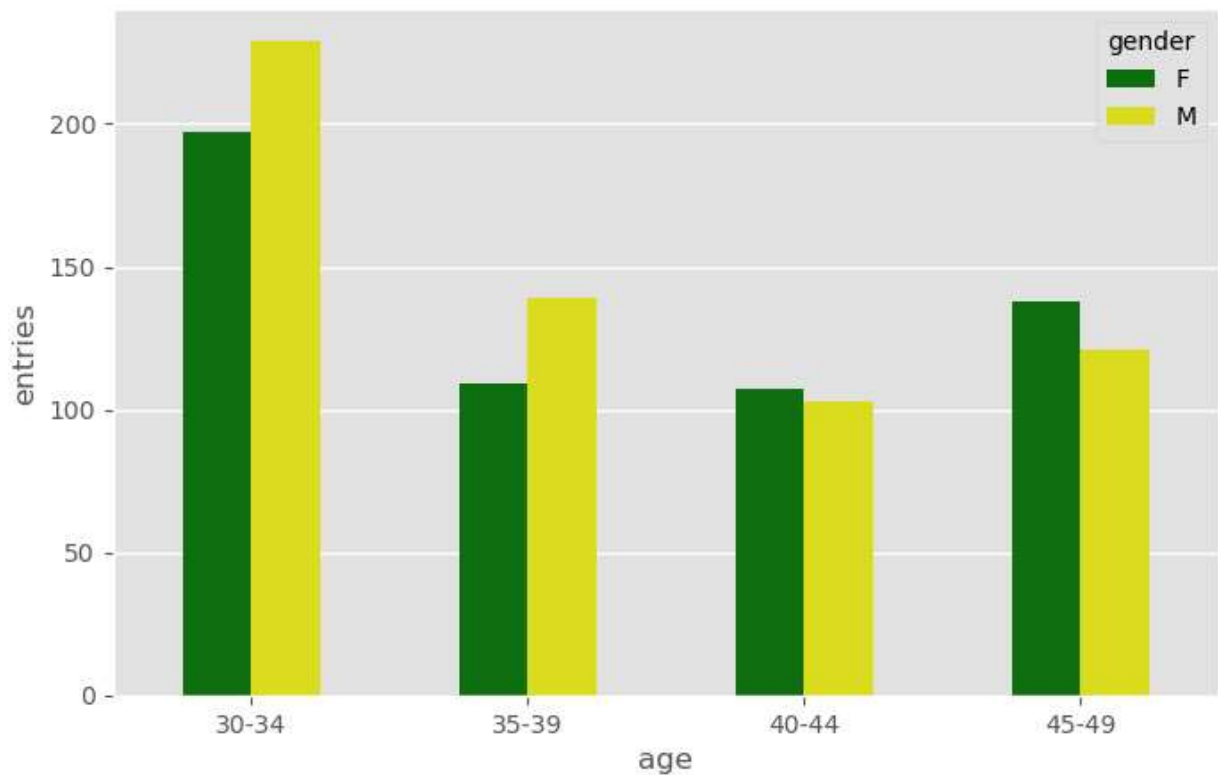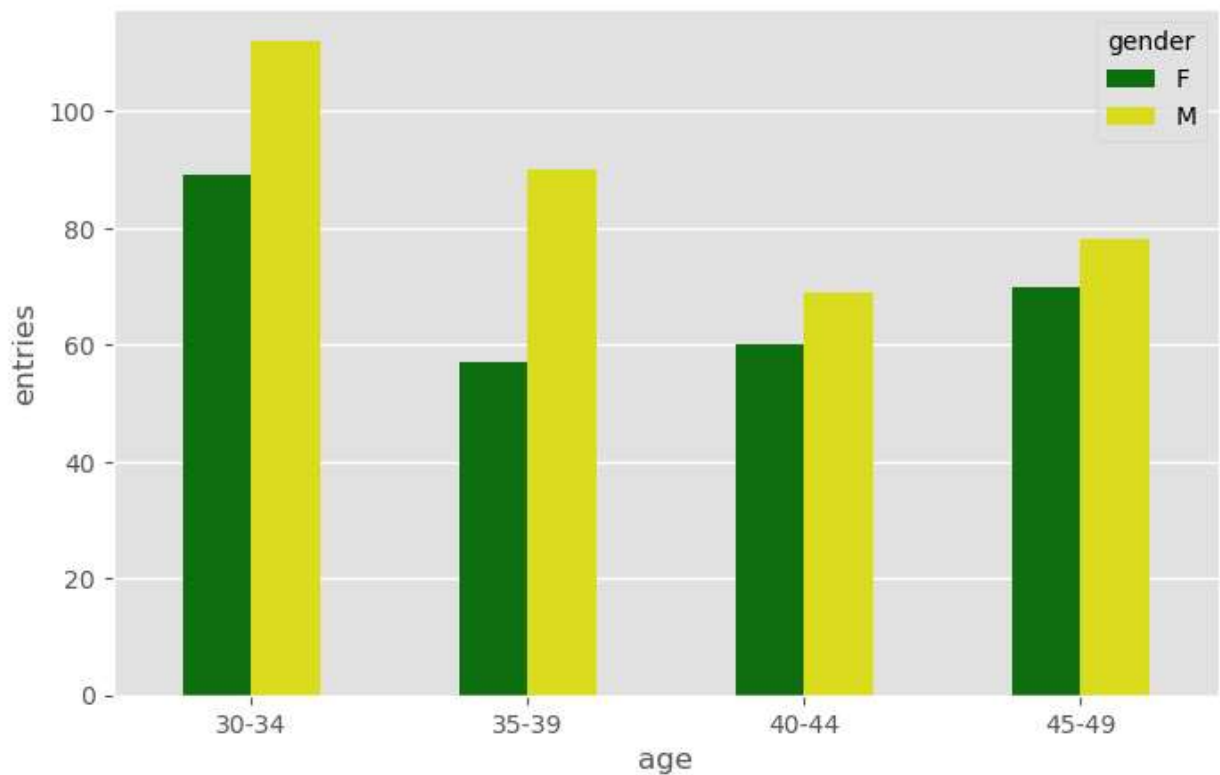
These is the data of the overall male and female participating in the campaigns with age group seggregation.

```
In [67]: xyz1178 = kag[kag['xyz_campaign_id']==1178].groupby(['gender','age']).size().reset_ind
         xyz1178
```

Out[67]:

	gender	age	entries
0	F	30-34	89
1	F	35-39	57
2	F	40-44	60
3	F	45-49	70
4	M	30-34	112
5	M	35-39	90
6	M	40-44	69
7	M	45-49	78

```
In [68]: matplotlib.rcParams['figure.figsize']=(8,5)
         sns.barplot(x='age',y='entries',hue = 'gender', data = xyz1178, errwidth = 0, width =
```

This is the data of xyz_campaign_id 1178., male and female participating in the campaigns with age group seggregation.

```
In [69]: xyz936 = kag[kag['xyz_campaign_id']==936].groupby(['gender','age']).size().reset_index
         xyz936
```

Out[69]:

	gender	age	entries
0	F	30-34	97
1	F	35-39	49
2	F	40-44	46
3	F	45-49	64
4	M	30-34	99
5	M	35-39	40
6	M	40-44	29
7	M	45-49	40

```
In [70]: matplotlib.rcParams['figure.figsize']=(8,5)
         sns.barplot(x='age',y='entries',hue = 'gender', data = xyz936, errwidth = 0, width = 0
```

This is the data of xyz_campaign_id 936., male and female participating in the campaigns with age group seggregation.

```
In [71]: xyz916 = kag[kag['xyz_campaign_id']==916].groupby(['gender','age']).size().reset_index
         xyz916
```

Out[71]:

	gender	age	entries
0	F	30-34	11
1	F	35-39	3
2	F	40-44	1
3	F	45-49	4
4	M	30-34	18
5	M	35-39	9
6	M	40-44	5
7	M	45-49	3

```
In [72]: matplotlib.rcParams['figure.figsize']=(8,5)
         sns.barplot(x='age',y='entries',hue = 'gender', data = xyz916, errwidth = 0, width = 6
```
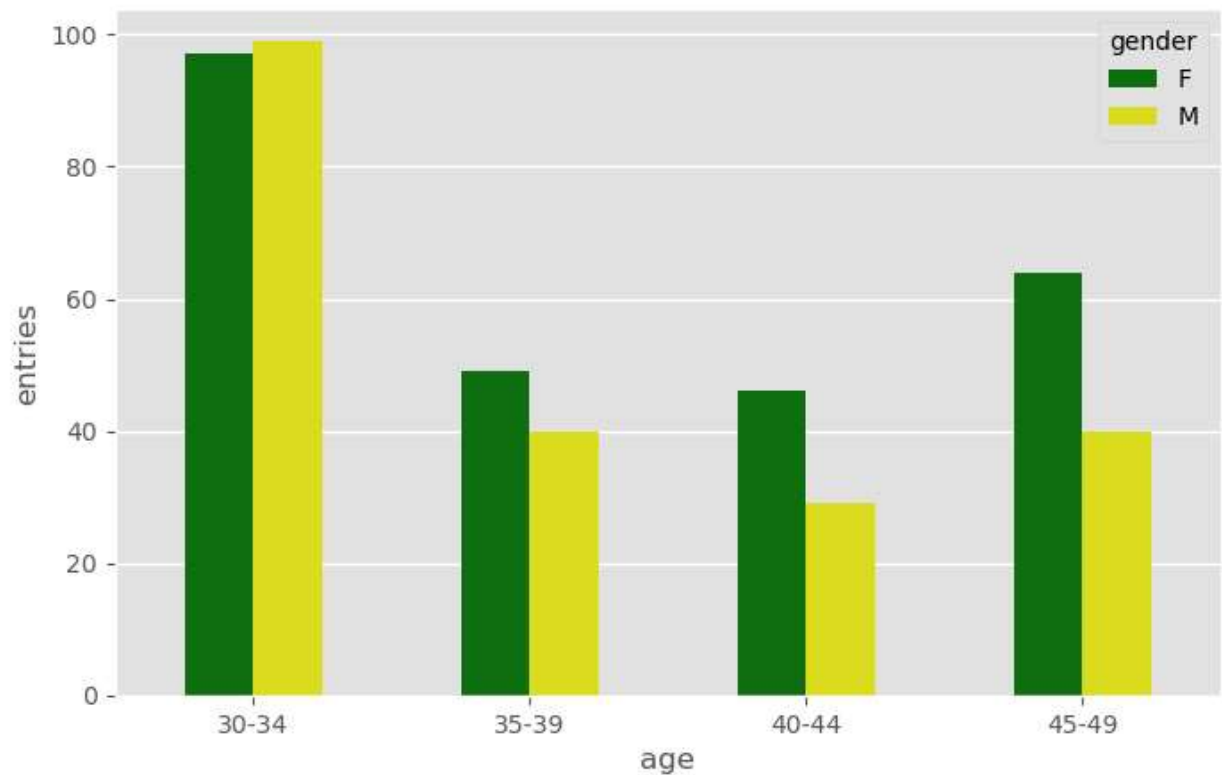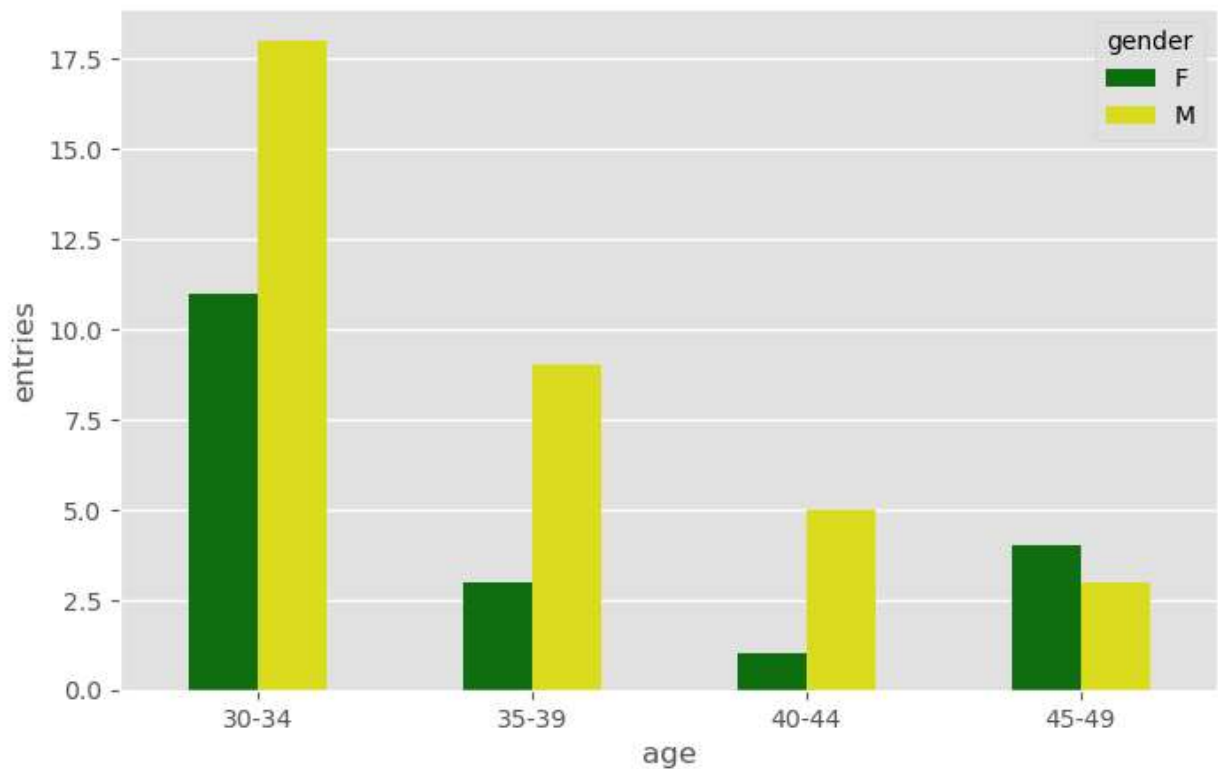
This is the data of xyz_campaign_id 916., male and female participating in the campaigns with age group seggregation.

```
In [77]:   fbid = kag.fb_campaign_id.value_counts().index
           fbid
```

```
Out[77]:   Index([144536, 144562, 144599, 144636, 144611, 144674, 144572, 144537, 144586,
                  144585,
                  ...
                  115802, 115812, 115854, 115856, 115858, 115866, 115870, 115876, 115878,
                  179982],
                 dtype='int64', name='fb_campaign_id', length=691)
```

```
In [97]:   fbcount =  kag.fb_campaign_id.value_counts().values
           fbcount
```

Out[97]:
```
array([6, 6, 6, 6, 6, 6, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5,
       5, 5, 5, 5, 5, 5, 5, 5, 5, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4,
       4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4,
       4, 4, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3,
       3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3,
       3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
       2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
       2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
       2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
       2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
       1, 1, 1, 1, 1, 1, 1], dtype=int64)
```

In [95]:
```python
plt.pie(fbcount[:20], labels = fbid[:20], autopct = '%1.2f%%', colors = ['orange', 'gr
plt.title('fb_Campaign_id vs count')
```

Out[95]:   Text(0.5, 1.0, 'fb_Campaign_id vs count')

fb_Campaign_id vs count
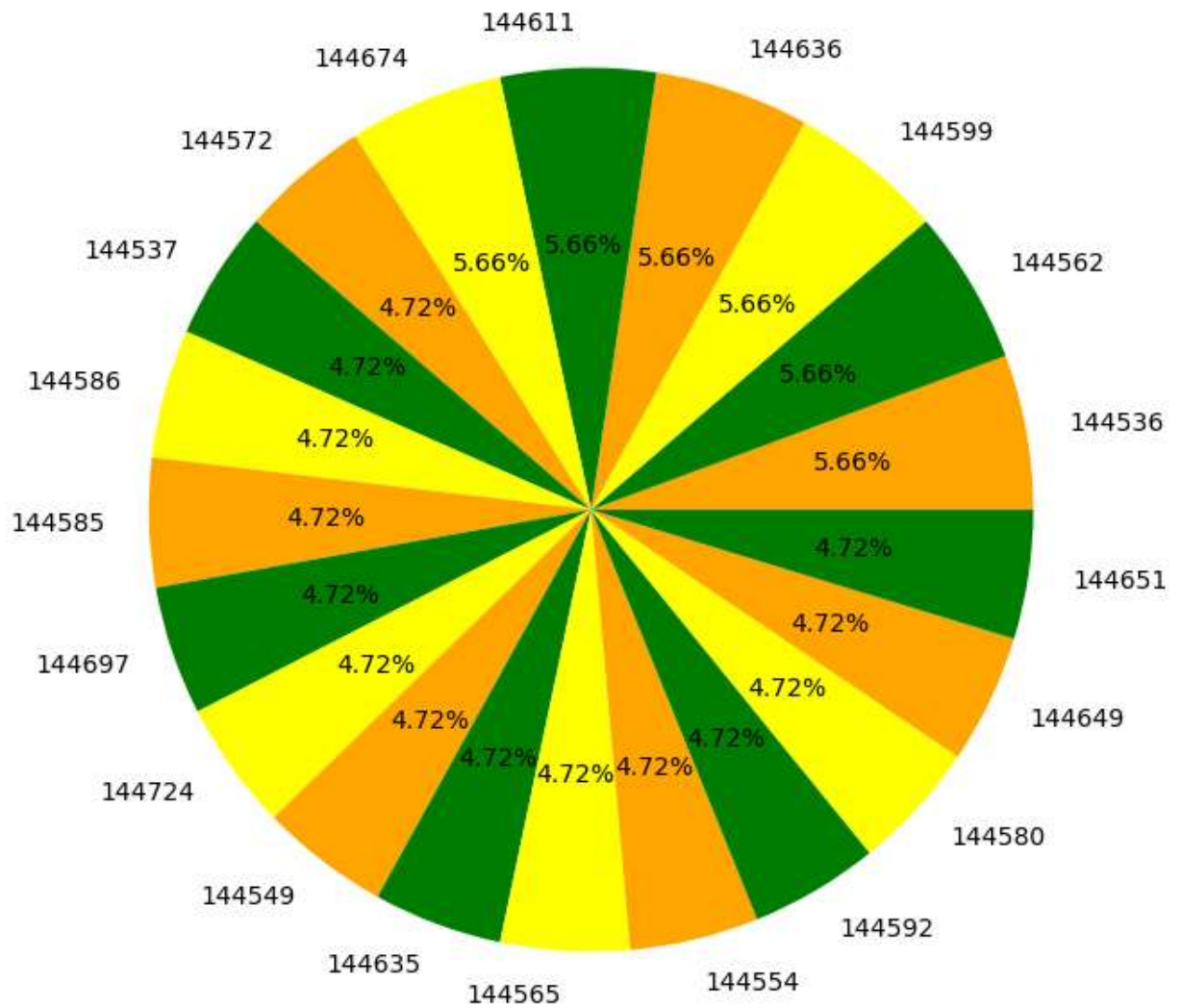

```
imprsn = kag[kag['Impressions']>100000].groupby(['gender','interest']).size().reset_in
imprsn
```

Out[115]:

|	gender	interest	entries
0	F	2	1
1	F	7	2
2	F	10	16
3	F	15	7
4	F	16	28
...	...	...	...
70	M	109	2
71	M	110	4
72	M	112	2
73	M	113	2
74	M	114	1

75 rows × 3 columns

In [111...
```python
imprsn_gender = imprsn.gender.value_counts().index
imprsn_gender
```

Out[111]:
```
Index(['F', 'M'], dtype='object', name='gender')
```

In [116...
```python
imprsn_count = imprsn.gender.value_counts().values
imprsn_count
```

Out[116]:
```
array([38, 37], dtype=int64)
```

In [118...
```python
style.use('ggplot')
plt.figure(figsize=(4,4))
plt.bar(imprsn_gender,imprsn_count,color = 'green', width = 0.2)
plt.ylabel('imprsn_count')
plt.xlabel('imprsn_gender')
plt.title('gender Vs count');
```

gender Vs count
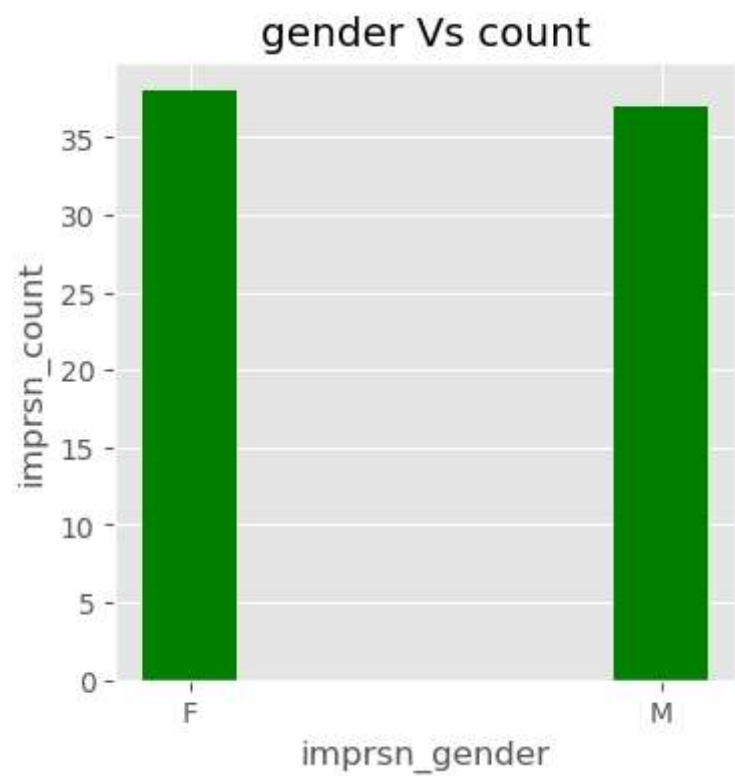

```
In [138…   Acnvsn = kag[kag['age']=='30-34'].groupby(['gender','xyz_campaign_id','Total_Conversic
           Acnvsn
```

Out[138]:

	gender	xyz_campaign_id	Total_Conversion	Approved_Conversion	0
0	F	916	1	0	7
1	F	916	1	1	4
2	F	936	0	0	1
3	F	936	1	0	58
4	F	936	1	1	27
...	...	...	...	...	...
96	M	1178	26	14	1
97	M	1178	28	14	1
98	M	1178	31	7	1
99	M	1178	40	21	1
100	M	1178	60	17	1

101 rows × 5 columns

```
In [143…   matplotlib.rcParams['figure.figsize']=(8,5)
           sns.barplot(x='Total_Conversion',y='Approved_Conversion',hue = 'xyz_campaign_id', data
```
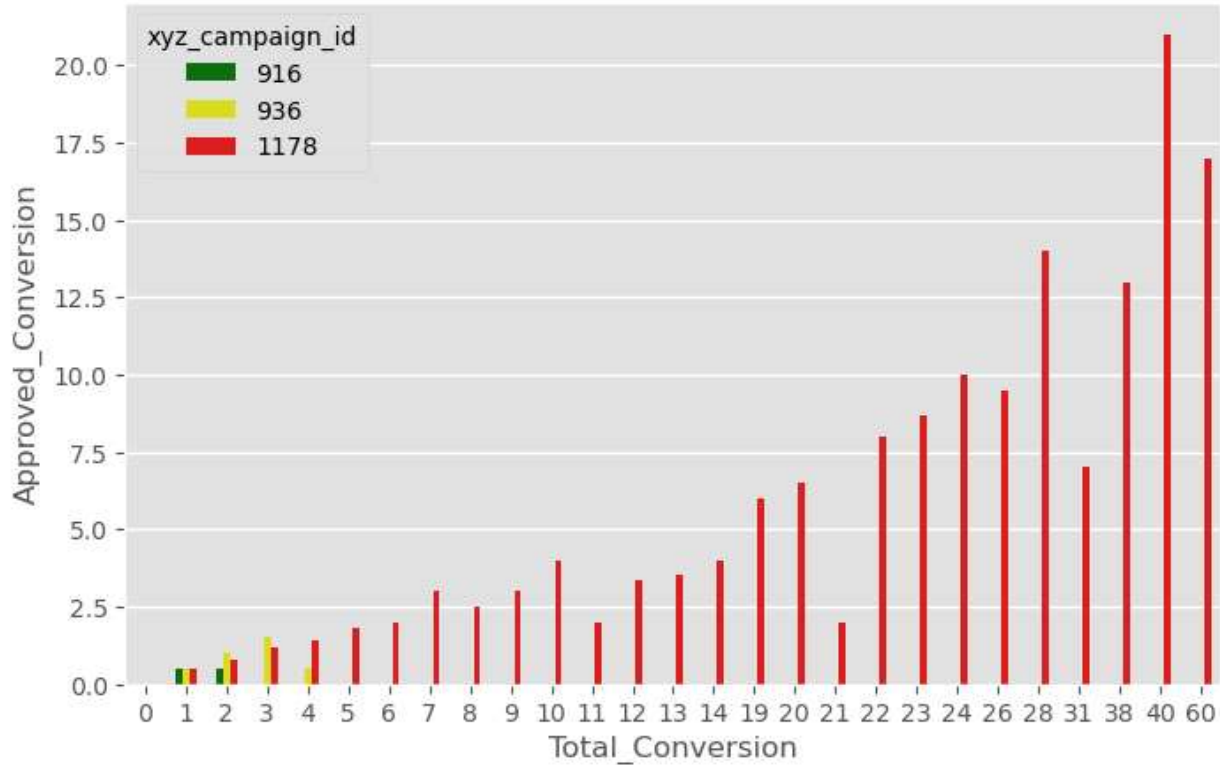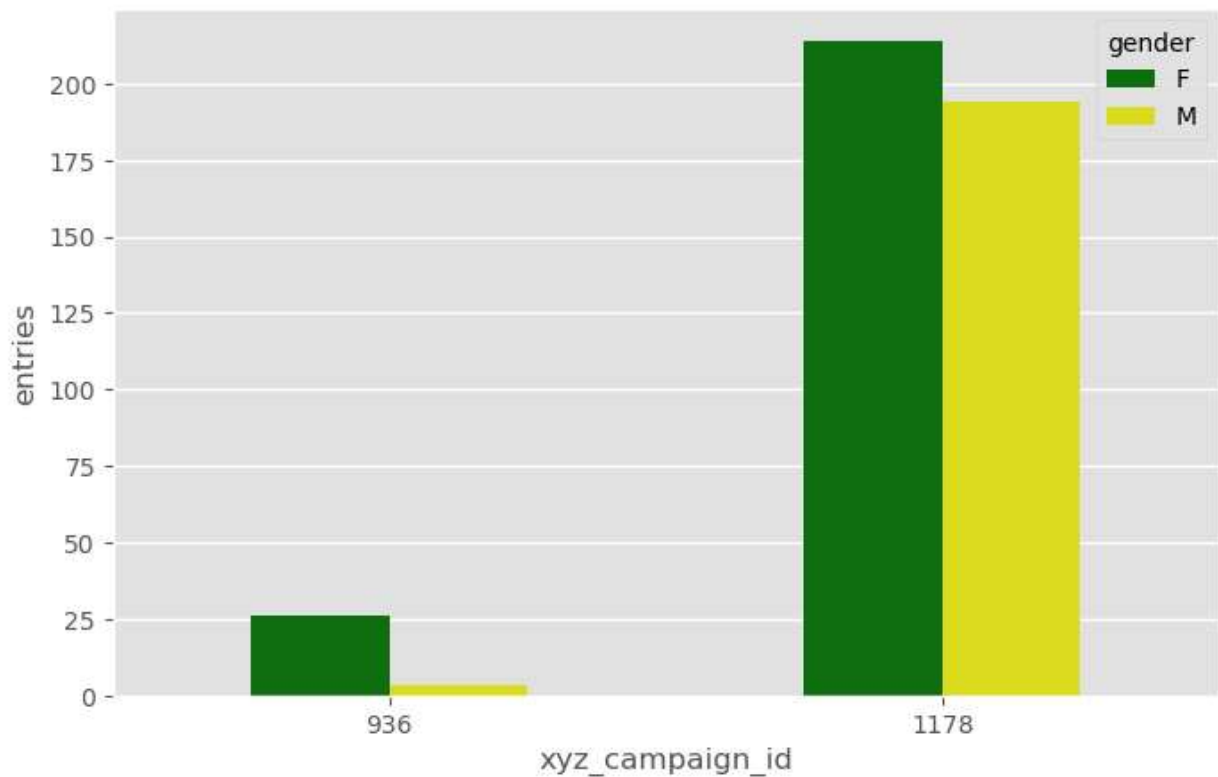
In [135…
```
clicks = kag[kag['Clicks']>=20].groupby(['gender','xyz_campaign_id']).size().reset_ind
clicks
```

Out[135]:

	gender	xyz_campaign_id	entries
0	F	936	26
1	F	1178	214
2	M	936	3
3	M	1178	194

In [137…
```
matplotlib.rcParams['figure.figsize']=(8,5)
sns.barplot(x='xyz_campaign_id',y='entries',hue = 'gender', data =clicks, errwidth = 6
```
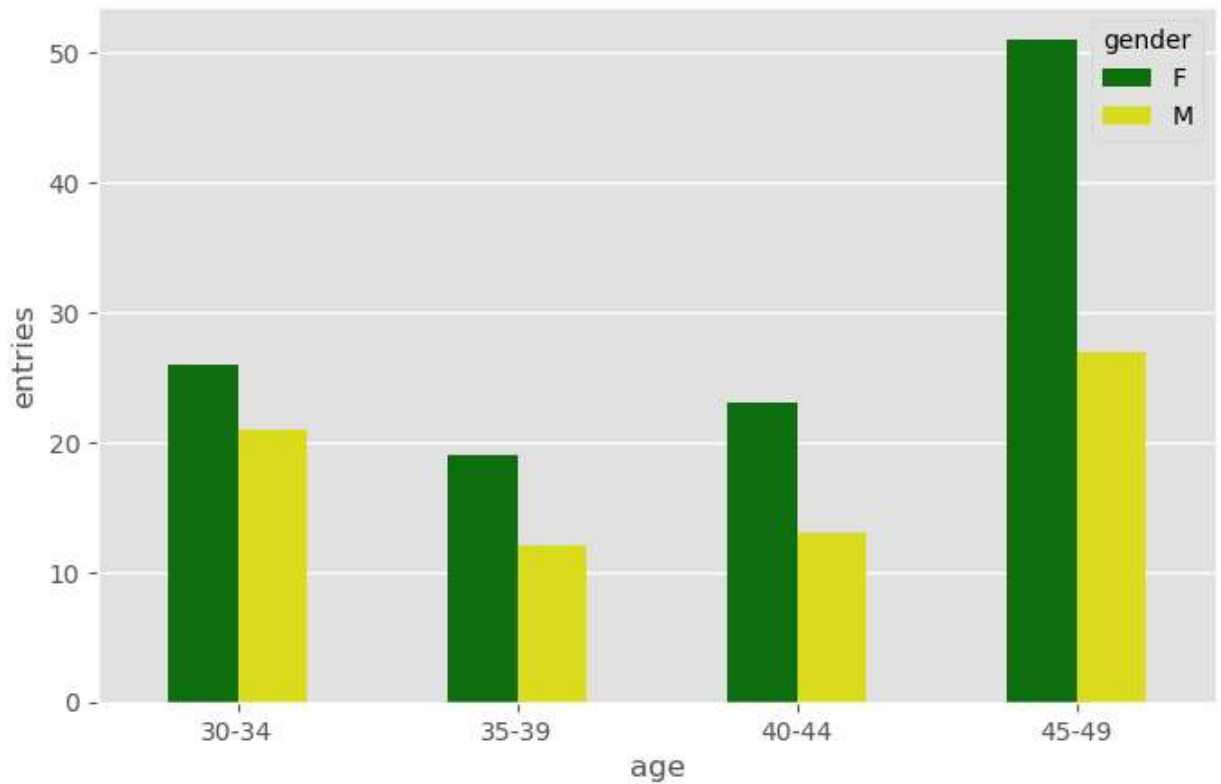
Out[137]:
```
<Axes: xlabel='xyz_campaign_id', ylabel='entries'>
```

In [144…
```python
spent = kag[kag['Spent']>100].groupby(['gender','age']).size().reset_index().rename(co
spent
```

Out[144]:

	gender	age	entries
0	F	30-34	26
1	F	35-39	19
2	F	40-44	23
3	F	45-49	51
4	M	30-34	21
5	M	35-39	12
6	M	40-44	13
7	M	45-49	27

In [148…
```python
matplotlib.rcParams['figure.figsize']=(8,5)
sns.barplot(x='age',y='entries',hue = 'gender', data =spent, errwidth = 0, width = 0.5
```

Out[148]:   `<Axes: xlabel='age', ylabel='entries'>`

In [119…   `kag.columns`

Out[119]:   Index(['ad_id', 'xyz_campaign_id', 'fb_campaign_id', 'age', 'gender',
                   'interest', 'Impressions', 'Clicks', 'Spent', 'Total_Conversion',
                   'Approved_Conversion'],
                  dtype='object')