In [1]:

```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt # visualizing data
import seaborn as sns
from collections import Counter
%matplotlib inline
import os
print(os.listdir("../workspace"))
```

['.ipynb_checkpoints', '.metadata', '.recommenders', 'BlackFriday.csv', 'Untitled.ipynb', 'Untitled1.ipynb']

In [2]:

```python
df = pd.read_csv('../workspace/BlackFriday.csv')
df.shape
```

Out[2]:

(537577, 12)

In [3]:

```python
df.describe()
```

Out[3]:

|	User_ID	Occupation	Marital_Status	Product_Category_1	Product_Category_2
count	5.375770e+05	537577.00000	537577.000000	537577.000000	370591.000000
mean	1.002992e+06	8.08271	0.408797	5.295546	9.842144
std	1.714393e+03	6.52412	0.491612	3.750701	5.087259
min	1.000001e+06	0.00000	0.000000	1.000000	2.000000
25%	1.001495e+06	2.00000	0.000000	1.000000	5.000000
50%	1.003031e+06	7.00000	0.000000	5.000000	9.000000
75%	1.004417e+06	14.00000	1.000000	8.000000	15.000000
max	1.006040e+06	20.00000	1.000000	18.000000	18.000000
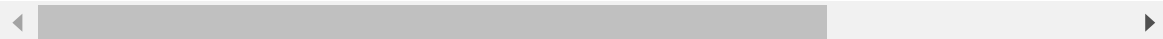
In [4]:

```
df.describe()
df.head()
```

Out[4]:

|	User_ID	Product_ID	Gender	Age	Occupation	City_Category	Stay_In_Current_City_Years
0	1000001	P00069042	F	0-17	10	A	2
1	1000001	P00248942	F	0-17	10	A	2
2	1000001	P00087842	F	0-17	10	A	2
3	1000001	P00085442	F	0-17	10	A	2
4	1000002	P00285442	M	55+	16	C	4+

In [5]:

```
# Gender

explode = (0.1,0)
fig1, ax1 = plt.subplots(figsize=(12,7))
ax1.pie(df['Gender'].value_counts(), explode=explode,labels=['Male','Female'], autopct=
'%1.1f%%',
        shadow=True, startangle=90)
# Equal aspect ratio ensures that pie is drawn as a circle
ax1.axis('equal')
plt.tight_layout()
plt.legend()
plt.show()
```

In [7]:
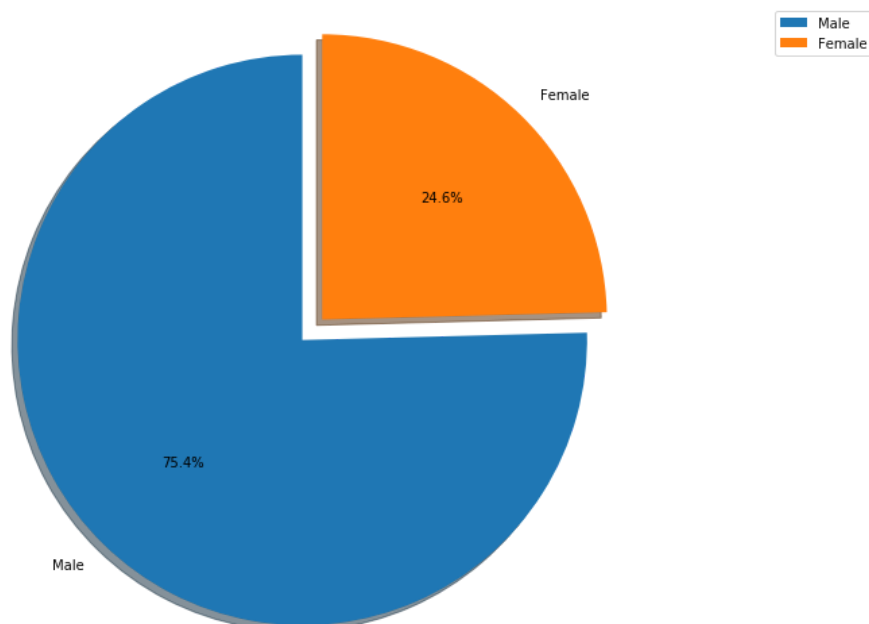
```python
# Gender wise purchase
def plot(group,column,plot):
    ax=plt.figure(figsize=(12,6))
    df.groupby(group)[column].sum().sort_values().plot(kind=plot)

plot('Gender','Purchase','bar')
```
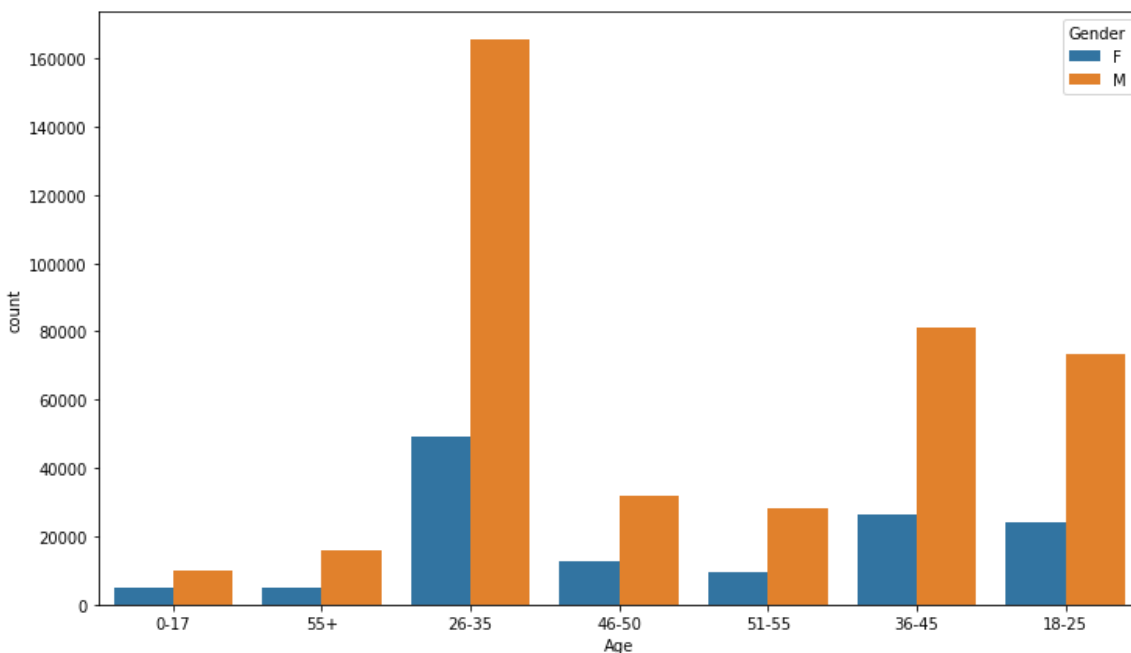

In [8]:
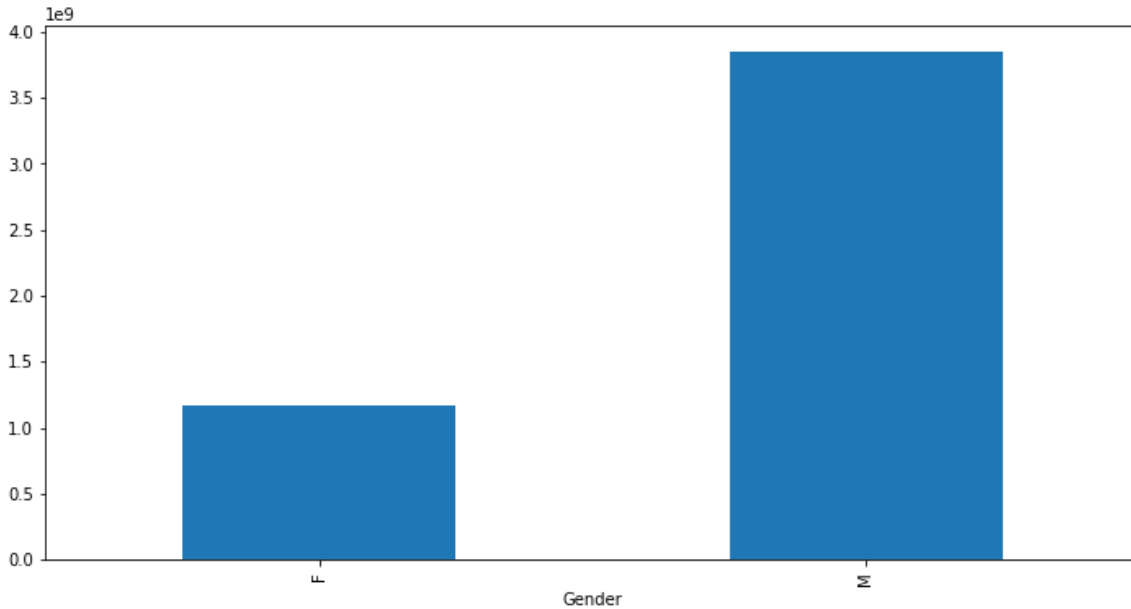
```python
# Age wise data
fig1, ax1 = plt.subplots(figsize=(12,7))
sns.countplot(df['Age'],hue=df['Gender'])
```
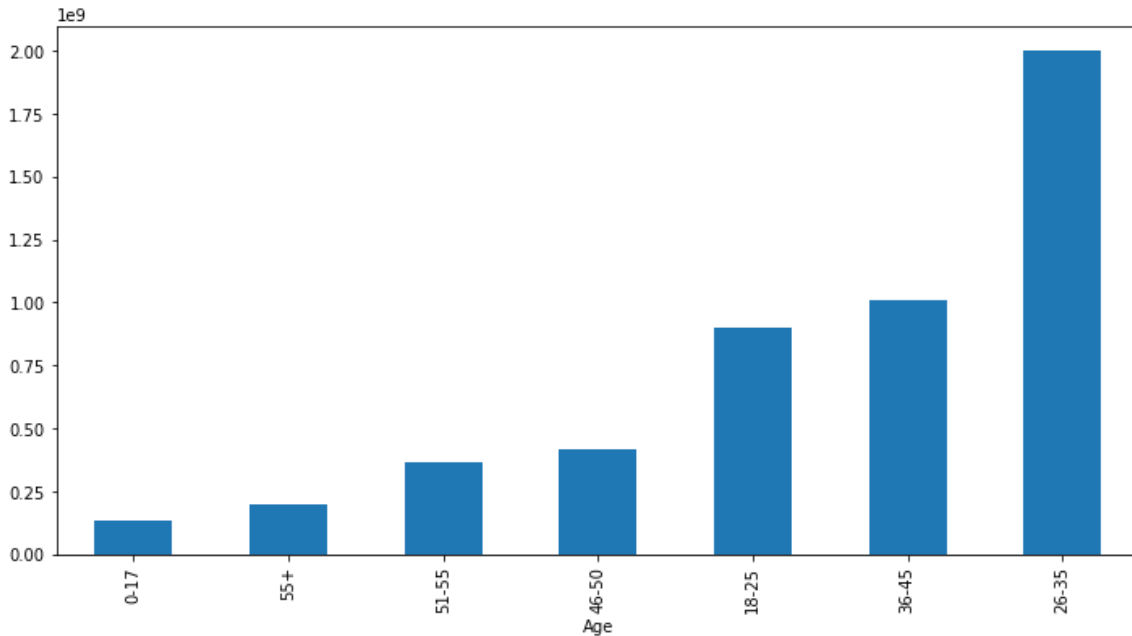
Out[8]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x2164e2a0cc8>
```
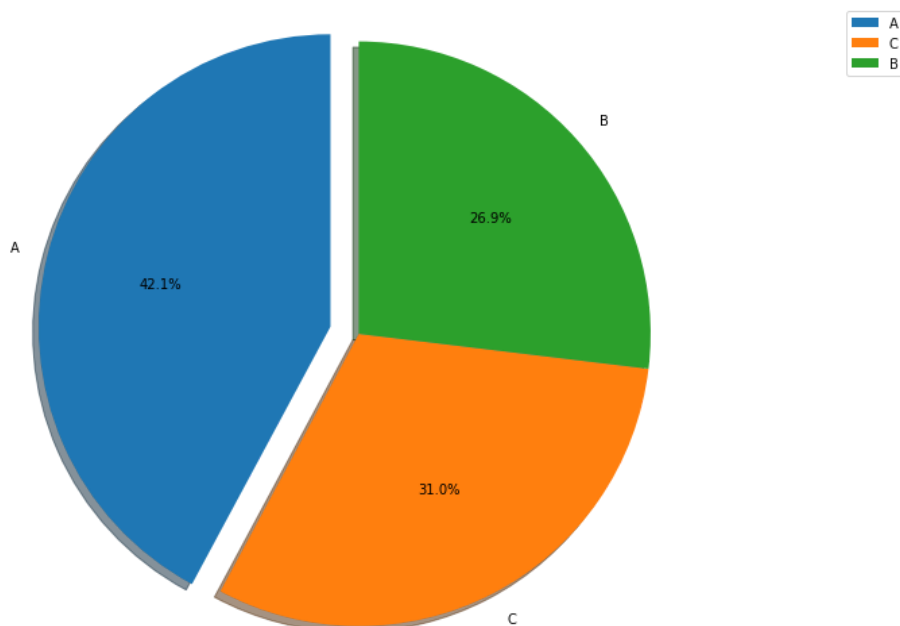
In [9]:

```python
# Age wise purchase
plot('Age','Purchase','bar')
```
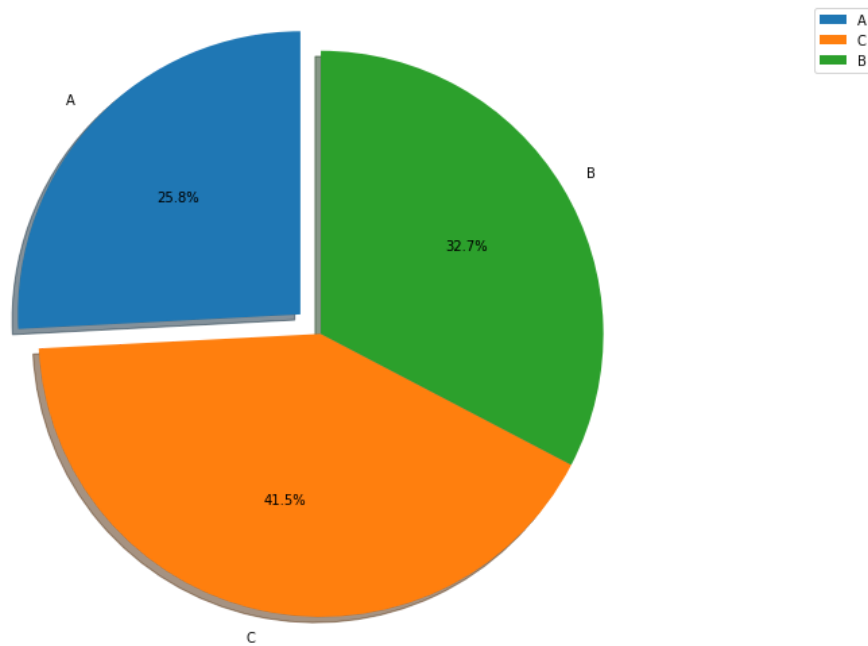

In [10]:

```python
# City
explode = (0.1, 0, 0)
fig1, ax1 = plt.subplots(figsize=(12,7))
ax1.pie(df['City_Category'].value_counts(),explode=explode, labels=df['City_Category'].
unique(), autopct='%1.1f%%',
        shadow=True, startangle=90)
# Equal aspect ratio ensures that pie is drawn as a circle
ax1.axis('equal')
plt.tight_layout()
plt.legend()
plt.show()
```
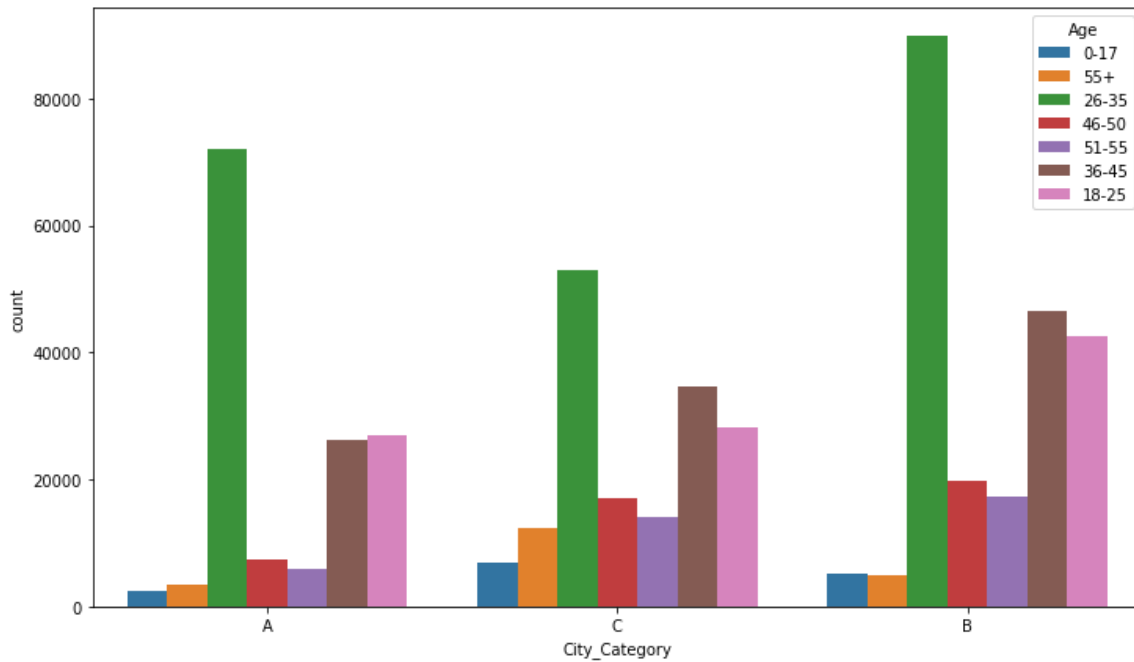
In [11]:

```python
# City
explode = (0.1, 0, 0)
fig1, ax1 = plt.subplots(figsize=(12,7))
ax1.pie(df.groupby('City_Category')['Purchase'].sum(),explode=explode, labels=df['City_
Category'].unique(), autopct='%1.1f%%',
        shadow=True, startangle=90)
# Equal aspect ratio ensures that pie is drawn as a circle
ax1.axis('equal')
plt.tight_layout()
plt.legend()
plt.show()
```

In [12]:

```
fig1, ax1 = plt.subplots(figsize=(12,7))
sns.countplot(df['City_Category'],hue=df['Age'])
```
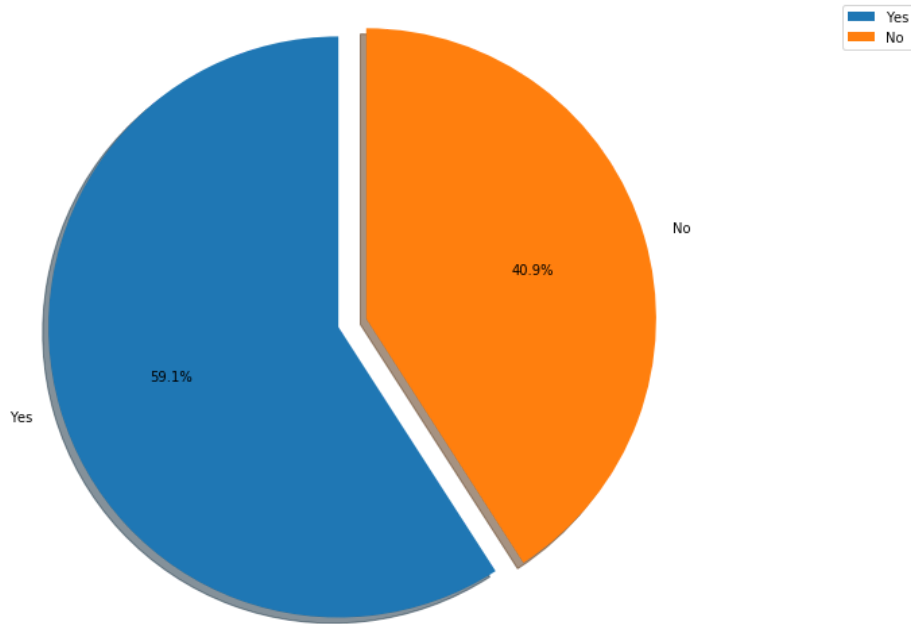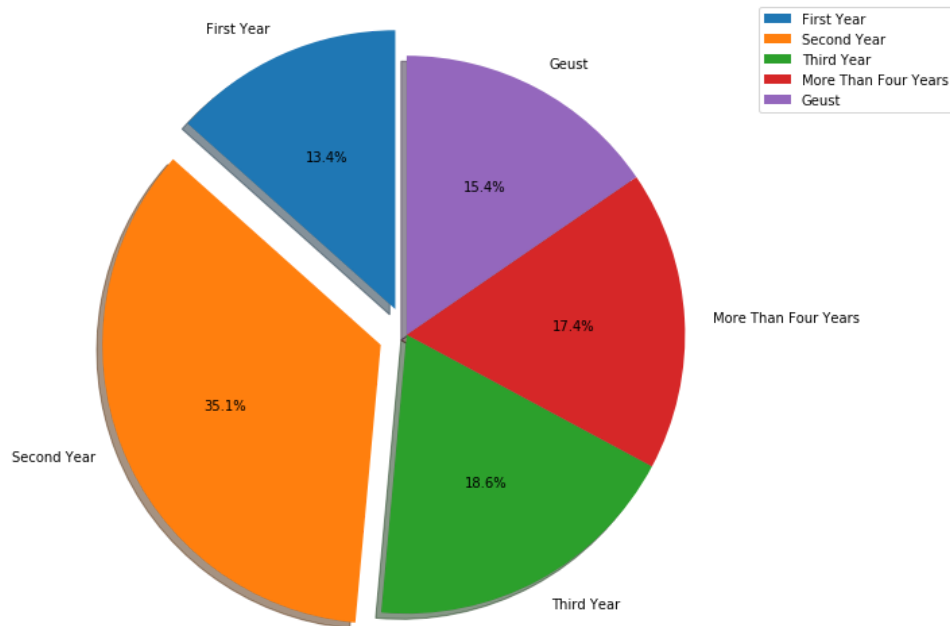
Out[12]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x2164fc19548>
```

In [13]:

```python
#label=['Underage 0-17','Retired +55','Middleage 26-35','46-50 y/o','Oldman 51-55','Mid
dleage+ 36-45','Youth']
explode = (0.1, 0)
fig1, ax1 = plt.subplots(figsize=(12,7))
ax1.pie(df['Marital_Status'].value_counts(),explode=explode, labels=['Yes','No'], autop
ct='%1.1f%%',
        shadow=True, startangle=90)
# Equal aspect ratio ensures that pie is drawn as a circle
ax1.axis('equal')
plt.tight_layout()
plt.legend()
plt.show()
```

In [14]:

```python
# Stability
labels=['First Year','Second Year','Third Year','More Than Four Years','Geust']
explode = (0.1, 0.1,0,0,0)
fig1, ax1 = plt.subplots(figsize=(12,7))
ax1.pie(df.groupby('Stay_In_Current_City_Years')['Purchase'].sum(),explode=explode, lab
els=labels, autopct='%1.1f%%',
        shadow=True, startangle=90)
# Equal aspect ratio ensures that pie is drawn as a circle
ax1.axis('equal')
plt.tight_layout()
plt.legend()
plt.show()
```

In [15]:

```python
# Stability
labels=['First Year','Second Year','Third Year','More Than Four Years','Geust']
#label=['Underage 0-17','Retired +55','Middleage 26-35','46-50 y/o','Oldman 51-55','Mid
dleage+ 36-45','Youth']
explode = (0.1, 0.1,0,0,0)
fig1, ax1 = plt.subplots(figsize=(12,7))
ax1.pie(df['Stay_In_Current_City_Years'].value_counts(),explode=explode, labels=labels,
autopct='%1.1f%%',
        shadow=True, startangle=90)
# Equal aspect ratio ensures that pie is drawn as a circle
ax1.axis('equal')
plt.tight_layout()
plt.legend()
plt.show()
```
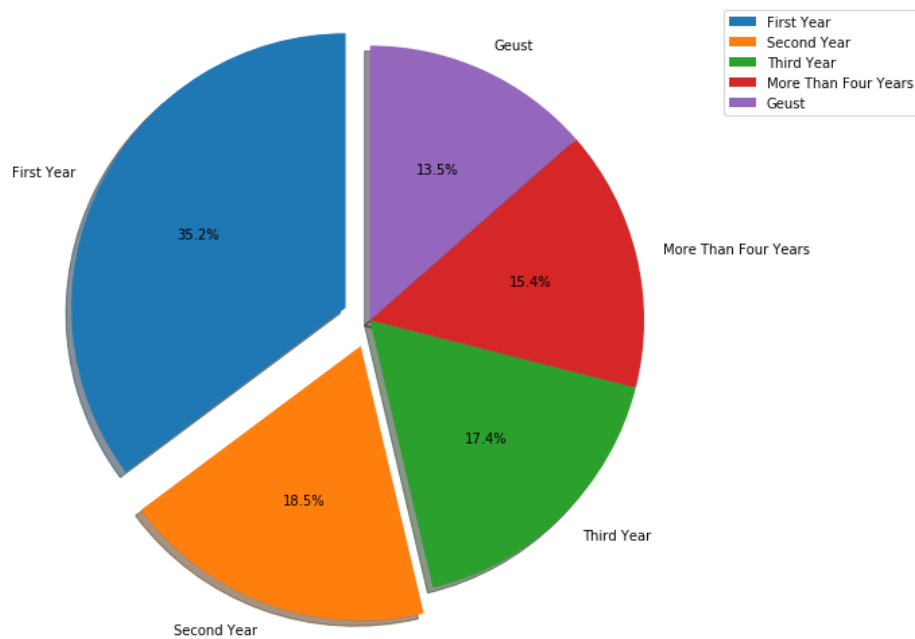
In [16]:

```python
plot('Stay_In_Current_City_Years','Purchase','bar')
```
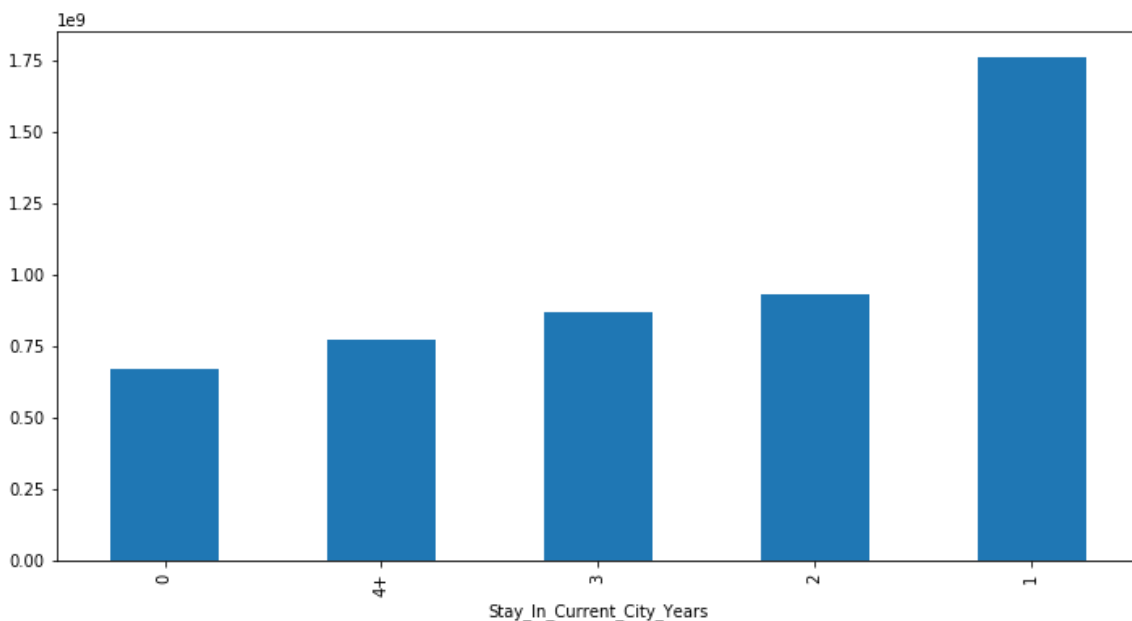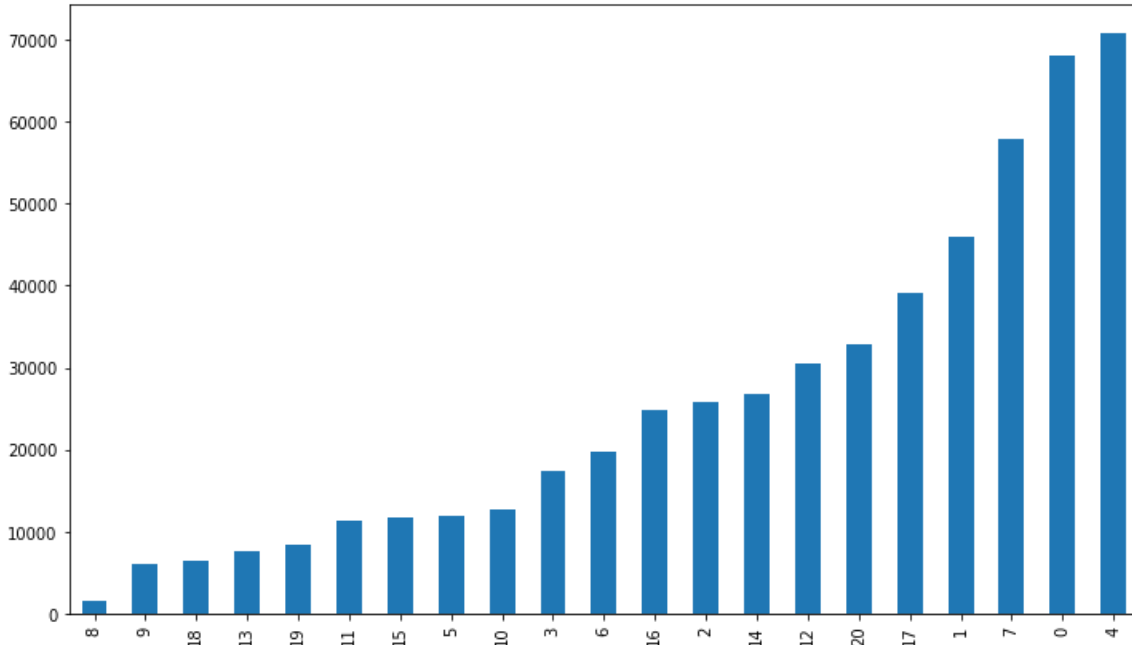
In [18]:

```python
# Occupation
fig1, ax1 = plt.subplots(figsize=(12,7))
df['Occupation'].value_counts().sort_values().plot(kind='bar')
```
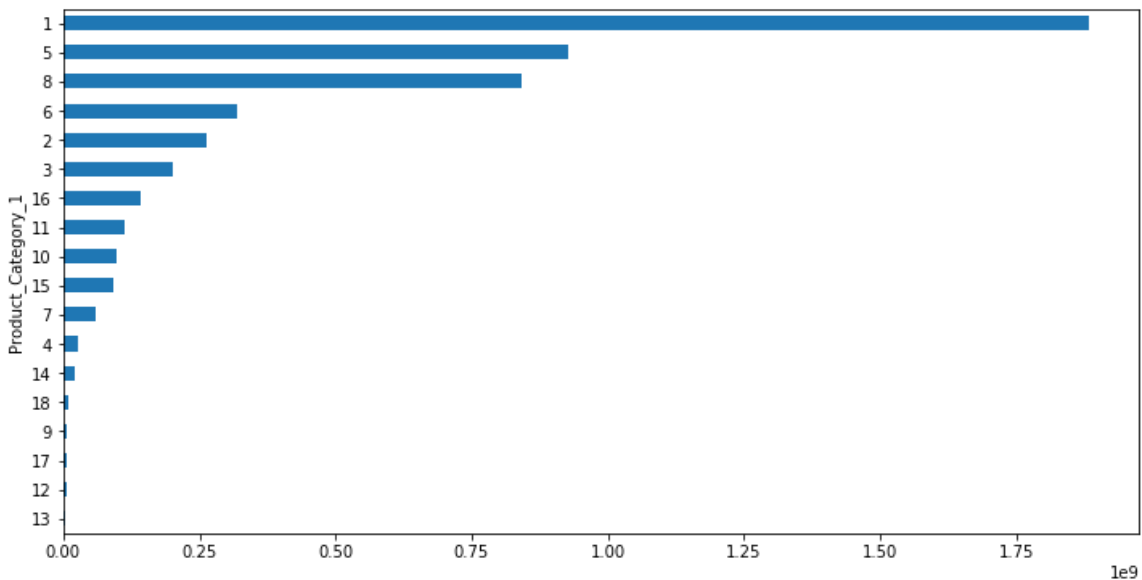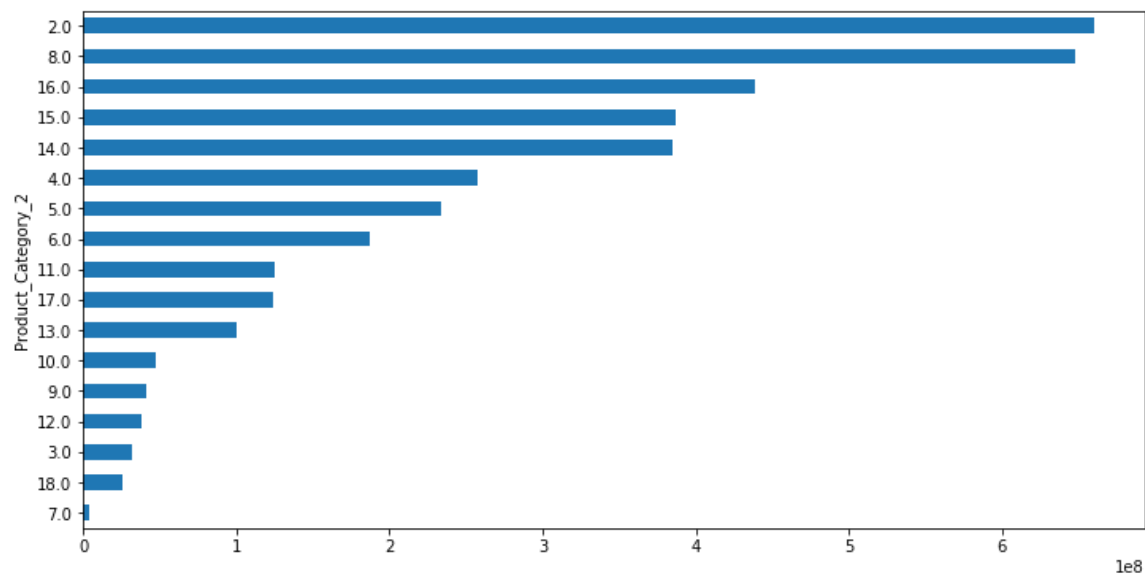
Out[18]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x2164d9d6788>
```


In [19]:

```python
# Products and Categories
plot('Product_Category_1','Purchase','barh')
```
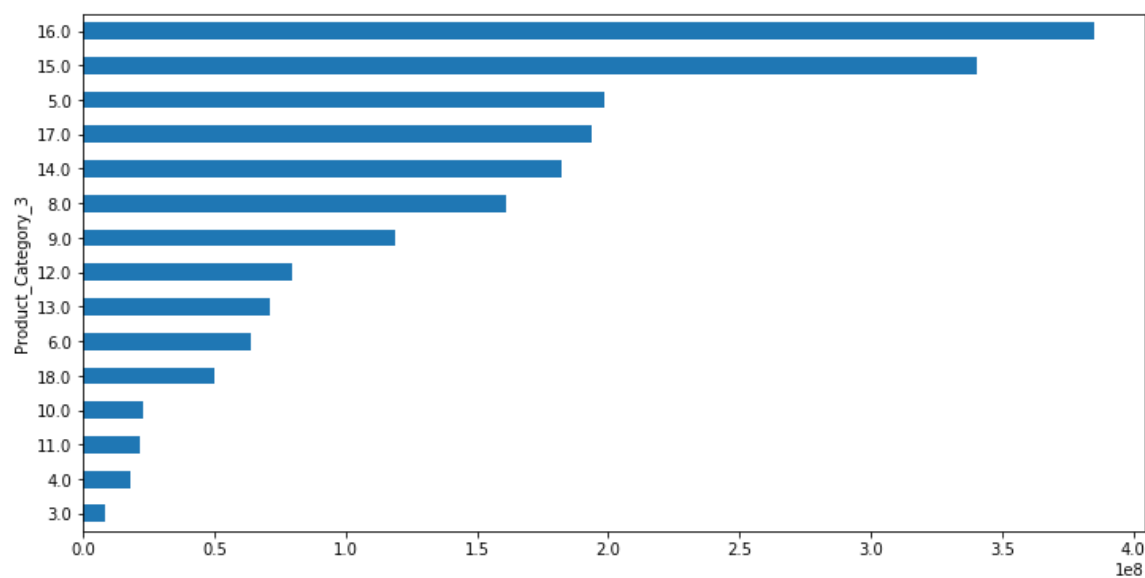
In [20]:

```python
plot('Product_Category_2','Purchase','barh')
```


In [21]:

```python
plot('Product_Category_3','Purchase','barh')
```
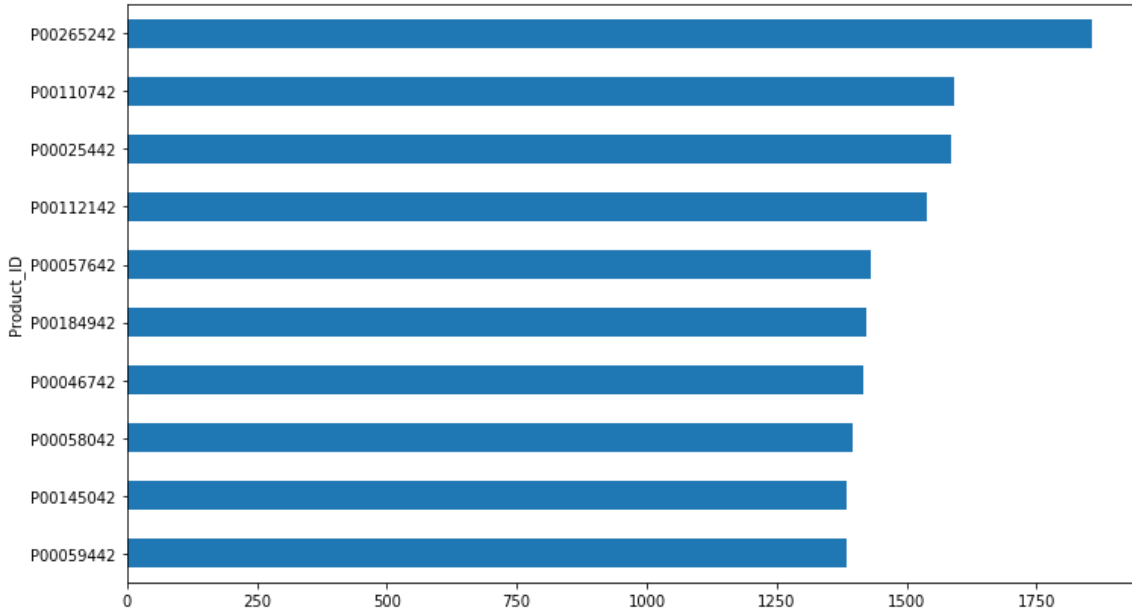
In [23]:

```python
fig1, ax1 = plt.subplots(figsize=(12,7))
df.groupby('Product_ID')['Purchase'].count().nlargest(10).sort_values().plot(kind='bar
h')
```

Out[23]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21650732348>
```
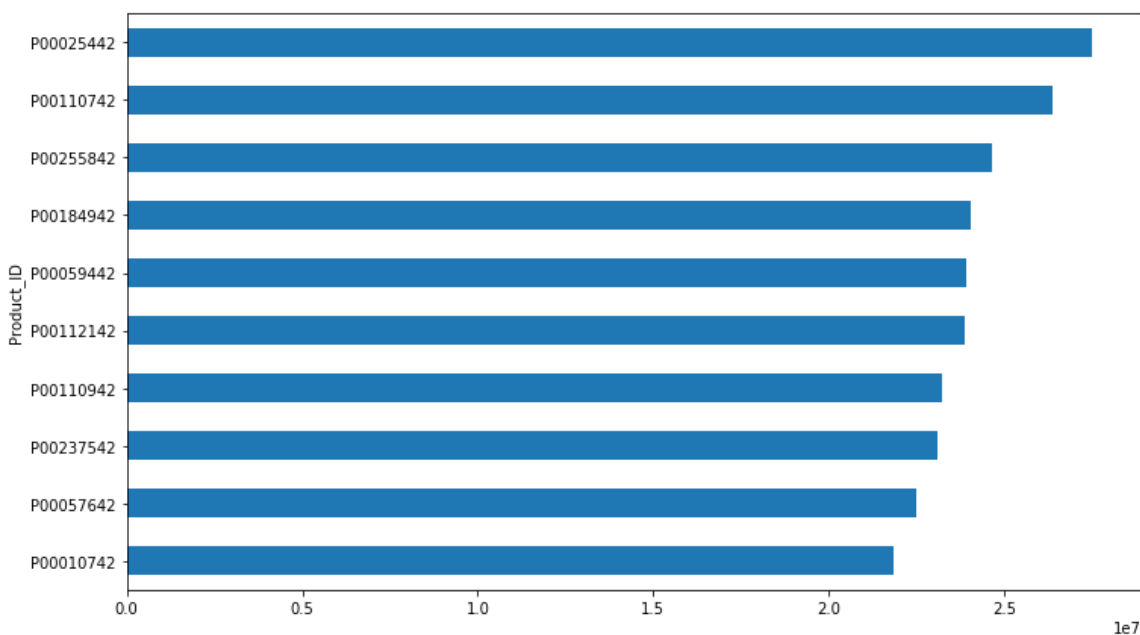

In [24]:

```python
fig1, ax1 = plt.subplots(figsize=(12,7))
df.groupby('Product_ID')['Purchase'].sum().nlargest(10).sort_values().plot(kind='barh')
```

Out[24]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21650a9acc8>
```

In [ ]: