

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [2]:

```
df=pd.read_csv("D:/Dixant/CDAC/Machine Learning/09-12-21/ign_n.csv")
```

In [3]:

```
df.head()
```

Out[3]:

	Unnamed: 0	score_phrase	title	url	platform	score	genre	e
0	0	Amazing	LittleBigPlanet PS Vita	/games/littlebigplanet-vita/vita-98907	PlayStation Vita	9.0	Platformer	
1	1	Amazing	LittleBigPlanet PS Vita -- Marvel Super Hero E...	/games/littlebigplanet-ps-vita-marvel-super-he...	PlayStation Vita	9.0	Platformer	
2	2	Great	Splice: Tree of Life	/games/splice/ipad-141070	iPad	8.5	Puzzle	
3	3	Great	NHL 13	/games/nhl-13/xbox-360-128182	Xbox 360	8.5	Sports	
4	4	Great	NHL 13	/games/nhl-13/ps3-128181	PlayStation 3	8.5	Sports	

In [4]:

```
df.shape
```

Out[4]:

```
(18625, 11)
```

In [5]:

```
df.isnull().sum()
```

Out[5]:

```
Unnamed: 0      0
score_phrase    0
title           0
url             0
platform        0
score           0
genre          36
editors_choice  0
release_year    0
release_month   0
release_day     0
dtype: int64
```

In [6]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18625 entries, 0 to 18624
Data columns (total 11 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   Unnamed: 0            18625 non-null  int64
 1   score_phrase          18625 non-null  object
 2   title                 18625 non-null  object
 3   url                   18625 non-null  object
 4   platform              18625 non-null  object
 5   score                 18625 non-null  float64
 6   genre                 18589 non-null  object
 7   editors_choice        18625 non-null  object
 8   release_year          18625 non-null  int64
 9   release_month         18625 non-null  int64
10  release_day           18625 non-null  int64
dtypes: float64(1), int64(4), object(6)
memory usage: 1.6+ MB
```

TASK 1

In [7]:

```
ps3=df[df['platform']=="PlayStation 3"]
ps3_c=ps3['score_phrase'].value_counts()
```

In [8]:

```
ps3_c
```

Out[8]:

Great	379
Good	320
Okay	196
Amazing	181
Mediocre	130
Bad	87
Awful	35
Painful	18
Masterpiece	9
Unbearable	1

Name: score_phrase, dtype: int64

In [9]:

```
ps3_c['Amazing']
```

Out[9]:

181

In [10]:

```
ps4=df[df['platform']=="PlayStation 4"]  
ps4_c=ps4['score_phrase'].value_counts()
```

In [11]:

```
ps4_c
```

Out[11]:

Great	83
Good	65
Amazing	55
Okay	36
Mediocre	19
Bad	7
Masterpiece	5
Awful	4
Painful	2
Unbearable	1

Name: score_phrase, dtype: int64

In [12]:

```
Xbox_360=df[df['platform']=="Xbox 360"]  
Xbox_360_c=Xbox_360['score_phrase'].value_counts()
```

In [13]:

```
Xbox_360_c
```

Out[13]:

Great	445
Good	375
Okay	256
Amazing	196
Mediocre	172
Bad	113
Awful	44
Painful	23
Masterpiece	5
Unbearable	2

Name: score_phrase, dtype: int64

In [14]:

```
Xbox_One=df[df['platform']=="Xbox One"]  
Xbox_One_c=Xbox_One['score_phrase'].value_counts()
```

In [15]:

```
Xbox_One_c
```

Out[15]:

Great	66
Good	50
Amazing	35
Okay	29
Mediocre	15
Bad	7
Masterpiece	4
Painful	2

Name: score_phrase, dtype: int64

In [16]:

```
PC=df[df['platform']=="PC"]  
PC_c=PC['score_phrase'].value_counts()
```

In [17]:

PC_c

Out[17]:

```

Great          967
Good           868
Okay          526
Amazing        351
Mediocre       321
Bad            179
Awful          90
Painful        47
Unbearable     15
Masterpiece    5
Disaster       1
Name: score_phrase, dtype: int64

```

In [18]:

```

amazing=[ps3_c['Amazing'],ps4_c['Amazing'],Xbox_360_c['Amazing'],Xbox_One_c['Amazing'],PC_c
great=[ps3_c['Great'],ps4_c['Great'],Xbox_360_c['Great'],Xbox_One_c['Great'],PC_c['Great']]
okay=[ps3_c['Okay'],ps4_c['Okay'],Xbox_360_c['Okay'],Xbox_One_c['Okay'],PC_c['Okay']]

```

In [19]:

```

print(amazing)
print(great)
print(okay)

```

```

[181, 55, 196, 35, 351]
[379, 83, 445, 66, 967]
[196, 36, 256, 29, 526]

```

In [20]:

```

data = pd.DataFrame(list(zip(amazing, great,okay)),
                    index=['PlayStation 3', 'PlayStation 4', 'Xbox 360', 'Xbox One', 'PC'],
                    columns=['Amazing', 'Great', 'Okay'])

```

In [21]:

data=data.transpose()

In [22]:

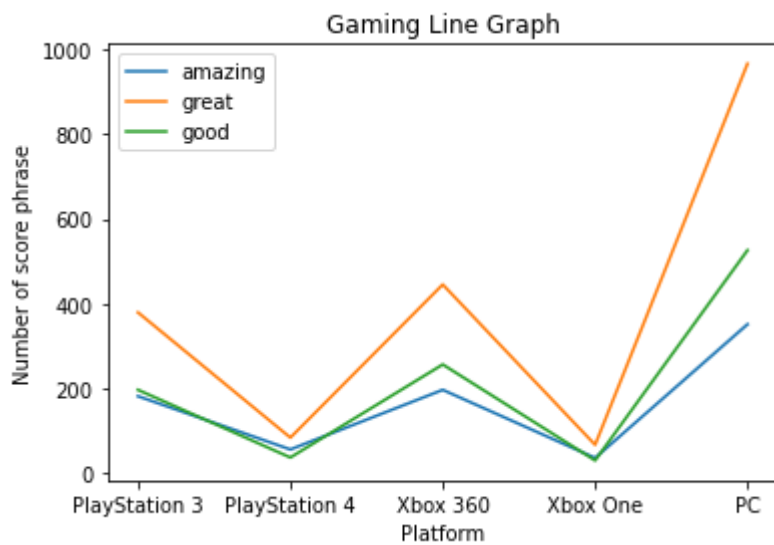
data

Out[22]:

	PlayStation 3	PlayStation 4	Xbox 360	Xbox One	PC
Amazing	181	55	196	35	351
Great	379	83	445	66	967
Okay	196	36	256	29	526

In [23]:

```
x=data.columns
y1=amazing
y2=great
y3=okay
plt.plot(x,y1,label='amazing')
plt.plot(x,y2,label='great')
plt.plot(x,y3,label='good')
plt.legend()
plt.xlabel("Platform")
plt.ylabel("Number of score phrase")
plt.title("Gaming Line Graph")
plt.show()
```



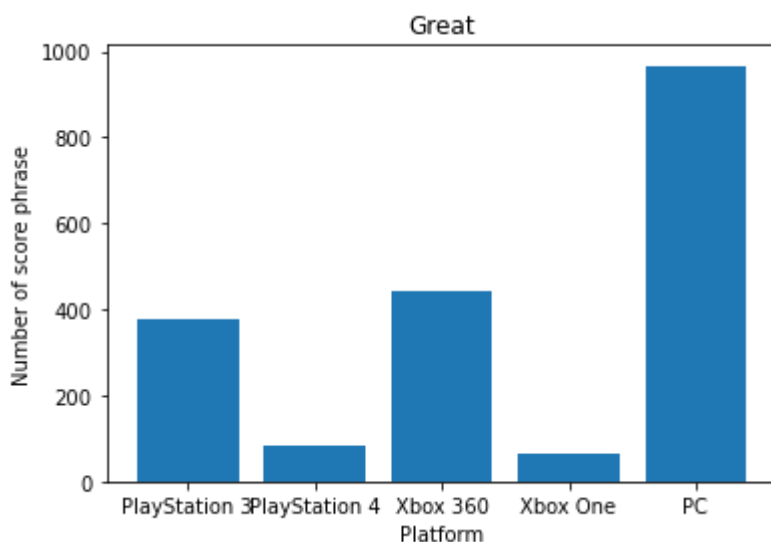
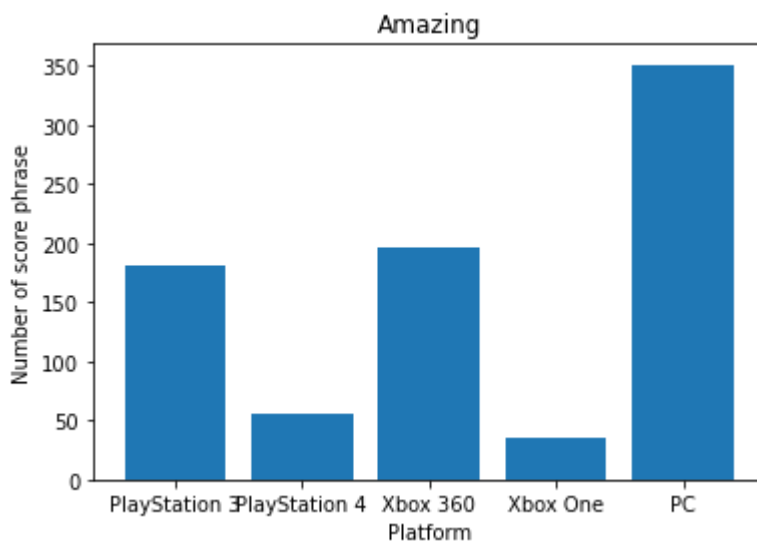
In [24]:

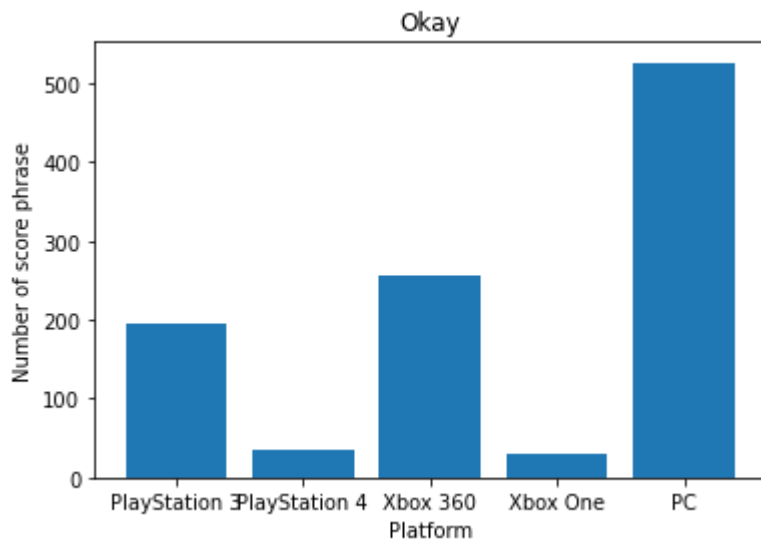
```
x=data.columns
y1=amazing
y2=great
y3=okay

plt.bar(x,y1,label='amazing')
plt.xlabel("Platform")
plt.ylabel("Number of score phrase")
plt.title("Amazing")
plt.show()

plt.bar(x,y2,label='great')
plt.xlabel("Platform")
plt.ylabel("Number of score phrase")
plt.title("Great")
plt.show()

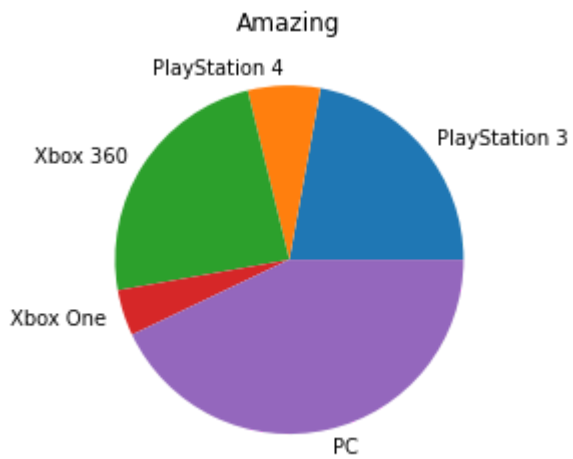
plt.bar(x,y3,label='okay')
plt.xlabel("Platform")
plt.ylabel("Number of score phrase")
plt.title("Okay")
plt.show()
```





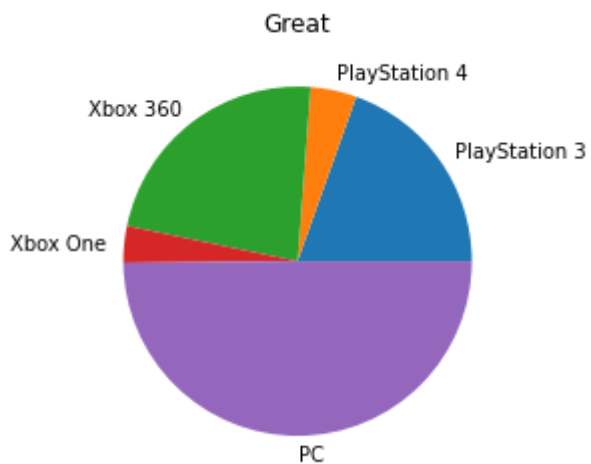
In [25]:

```
plt.title("Amazing")  
plt.pie(y1, labels = x)  
plt.show()
```



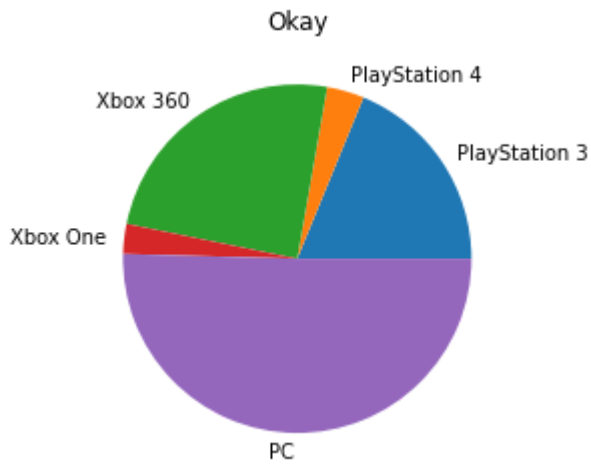
In [26]:

```
plt.title("Great")  
plt.pie(y2, labels = x)  
plt.show()
```



In [27]:

```
plt.title("Okay")  
plt.pie(y3, labels = x)  
plt.show()
```



TASK2

In [28]:

```
avg_score=df.score.mean()
```

In [29]:

```
ps3=df[df['platform']=="PlayStation 3"]  
ps3_filter1=ps3['score']>avg_score  
  
ps3_ab_avg=ps3.where(ps3_filter1)  
ps3_ab_avg.dropna(inplace=True)  
ab_ps3=ps3_ab_avg.shape[0]  
print(ab_ps3)
```

888

In [30]:

```
ps3_1=df[df['platform']=="PlayStation 3"]  
ps3_filter2=ps3['score']<=avg_score  
  
ps3_b_avg=ps3.where(ps3_filter2)  
ps3_b_avg.dropna(inplace=True)  
b_ps3=ps3_b_avg.shape[0]  
b_ps3
```

Out[30]:

467

In [31]:

```
t_ps3=ps3.shape[0]  
print(t_ps3)
```

1356

In [32]:

```
ps4=df[df['platform']=="PlayStation 4"]  
ps4_filter1=ps4['score']>avg_score  
  
ps4_ab_avg=ps4.where(ps4_filter1)  
ps4_ab_avg.dropna(inplace=True)  
ab_ps4=ps4_ab_avg.shape[0]  
print(ab_ps4)
```

208

In [33]:

```
ps4_1=df[df['platform']=="PlayStation 4"]  
ps4_filter2=ps4['score']<=avg_score  
  
ps4_b_avg=ps4.where(ps4_filter2)  
ps4_b_avg.dropna(inplace=True)  
b_ps4=ps4_b_avg.shape[0]  
b_ps4
```

Out[33]:

69

In [34]:

```
t_ps4=ps4.shape[0]  
print(t_ps4)
```

277

In [35]:

```
Xbox_360=df[df['platform']=="Xbox 360"]  
Xbox_360_filter1=Xbox_360['score']>avg_score  
  
Xbox_360_ab_avg=Xbox_360.where(Xbox_360_filter1)  
Xbox_360_ab_avg.dropna(inplace=True)  
ab_Xbox_360=Xbox_360_ab_avg.shape[0]  
print(ab_Xbox_360)
```

1021

In [36]:

```
Xbox_360_1=df[df['platform']=="Xbox 360"]
Xbox_360_filter2=Xbox_360['score']<=avg_score

Xbox_360_b_avg=Xbox_360.where(Xbox_360_filter2)
Xbox_360_b_avg.dropna(inplace=True)
b_Xbox_360=Xbox_360_b_avg.shape[0]
b_Xbox_360
```

Out[36]:

610

In [37]:

```
t_Xbox_360=Xbox_360.shape[0]
print(t_Xbox_360)
```

1631

In [38]:

```
Xbox_One=df[df['platform']=="Xbox One"]
Xbox_One_filter1=Xbox_One['score']>avg_score

Xbox_One_ab_avg=Xbox_One.where(Xbox_One_filter1)
Xbox_One_ab_avg.dropna(inplace=True)
ab_Xbox_One=Xbox_One_ab_avg.shape[0]
print(ab_Xbox_One)
```

155

In [39]:

```
Xbox_One_1=df[df['platform']=="Xbox One"]
Xbox_One_filter2=Xbox_One['score']<=avg_score

Xbox_One_b_avg=Xbox_One.where(Xbox_One_filter2)
Xbox_One_b_avg.dropna(inplace=True)
b_Xbox_One=Xbox_One_b_avg.shape[0]
b_Xbox_One
```

Out[39]:

53

In [40]:

```
t_Xbox_One=Xbox_One.shape[0]
print(t_Xbox_One)
```

208

In [41]:

```
PC=df[df['platform']=="PC"]
PC_filter1=PC['score']>avg_score

PC_ab_avg=PC.where(PC_filter1)
PC_ab_avg.dropna(inplace=True)
ab_PC=PC_ab_avg.shape[0]
print(ab_PC)
```

2190

In [42]:

```
PC_1=df[df['platform']=="PC"]
PC_filter2=PC['score']<=avg_score

PC_b_avg=PC.where(PC_filter2)
PC_b_avg.dropna(inplace=True)
b_PC=PC_b_avg.shape[0]
b_PC
```

Out[42]:

1177

In [43]:

```
t_PC=PC.shape[0]
print(t_PC)
```

3370

In [44]:

```
Above_avg=[ab_ps3,ab_ps4,ab_Xbox_360,ab_Xbox_One,ab_PC]
Below_avg=[b_ps3,b_ps4,b_Xbox_360,b_Xbox_One,b_PC]
Total=[t_ps3,t_ps4,t_Xbox_360,t_Xbox_One,t_PC]
```

In [45]:

```
print(Above_avg)
print(Below_avg)
print(Total)
```

```
[888, 208, 1021, 155, 2190]
[467, 69, 610, 53, 1177]
[1356, 277, 1631, 208, 3370]
```

In [46]:

```
data1 = pd.DataFrame(list(zip(Above_avg, Below_avg,Total)),
                      index=['PlayStation 3', 'PlayStation 4', 'Xbox 360', 'Xbox One', 'PC'],
                      columns=['Above Average', 'Below Average', 'Total'])
```

In [47]:

data1

Out[47]:

	Above Average	Below Average	Total
PlayStation 3	888	467	1356
PlayStation 4	208	69	277
Xbox 360	1021	610	1631
Xbox One	155	53	208
PC	2190	1177	3370

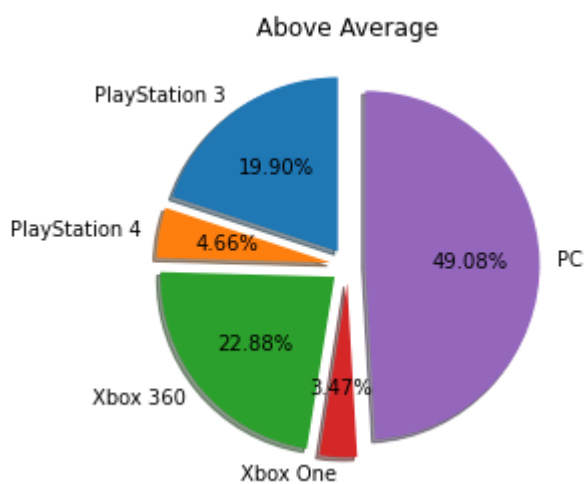
In [48]:

data1=data1.transpose()

In [49]:

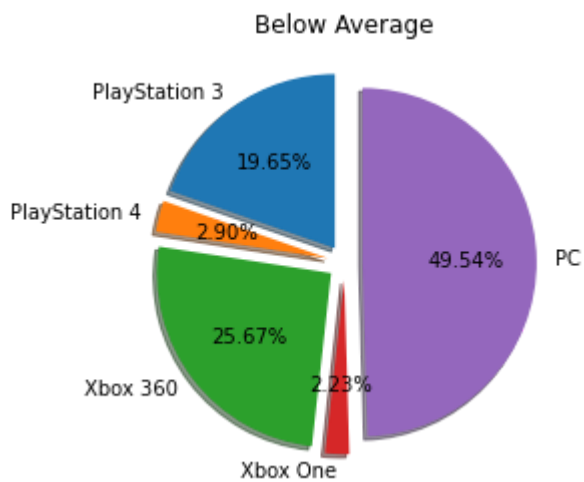
```
x= data1.columns
y1= Above_avg
y2= Below_avg
y3= Total

plt.pie(y1,labels=x, startangle=90, shadow=True,explode=(0.1,0.1, 0.1, 0.1, 0.1), autopct='
plt.title('Above Average')
plt.show()
```



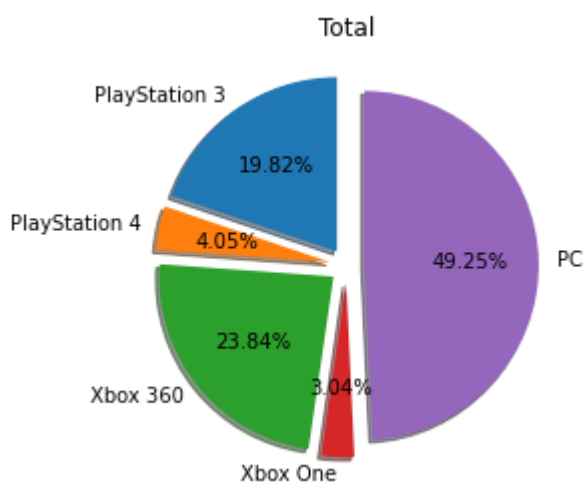
In [50]:

```
plt.pie(y2,labels=x, startangle=90, shadow=True,explode=(0.1,0.1, 0.1, 0.1, 0.1), autopct='  
plt.title('Below Average')  
plt.show()
```



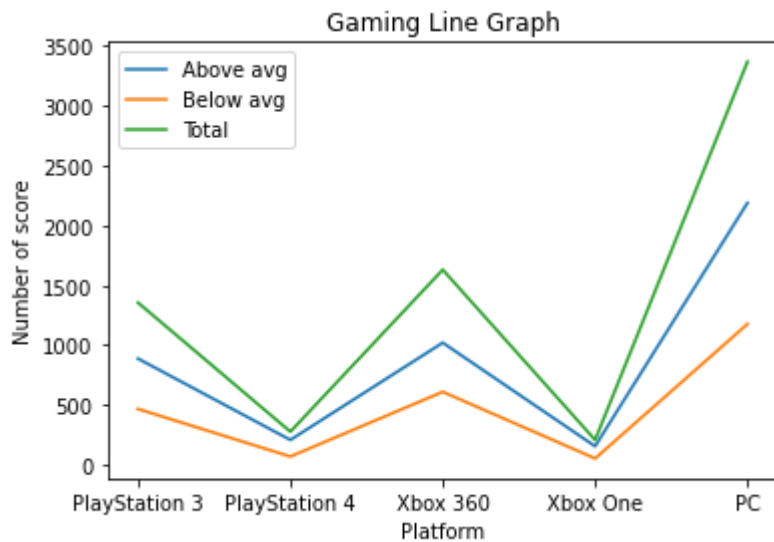
In [51]:

```
plt.pie(y3,labels=x, startangle=90, shadow=True,explode=(0.1,0.1, 0.1, 0.1, 0.1), autopct='  
plt.title('Total')  
plt.show()
```



In [52]:

```
plt.plot(x,y1,label='Above avg')
plt.plot(x,y2,label='Below avg')
plt.plot(x,y3,label='Total')
plt.legend()
plt.xlabel("Platform")
plt.ylabel("Number of score")
plt.title("Gaming Line Graph")
plt.show()
```

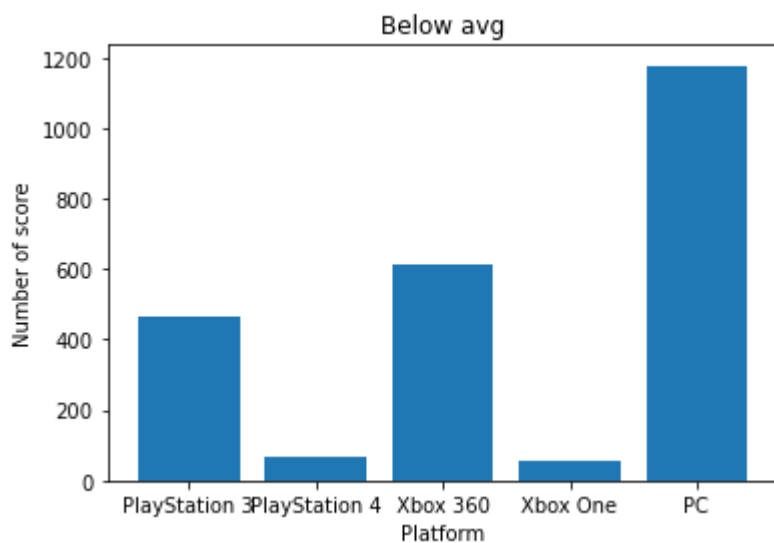
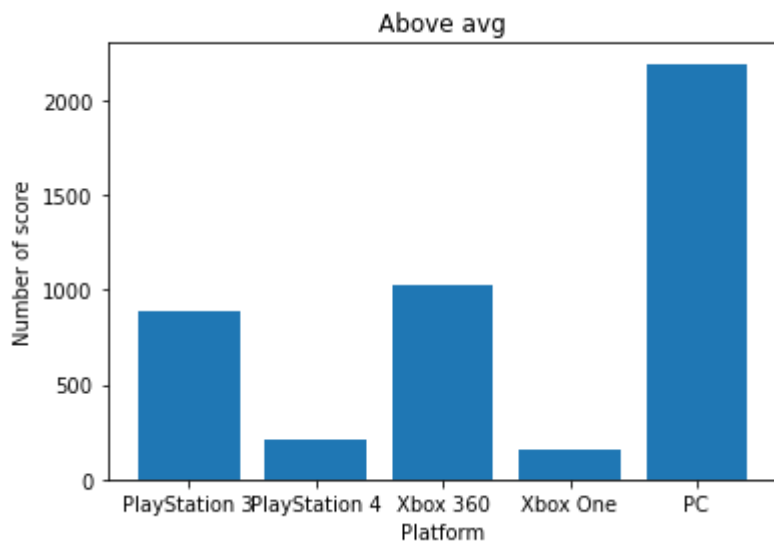


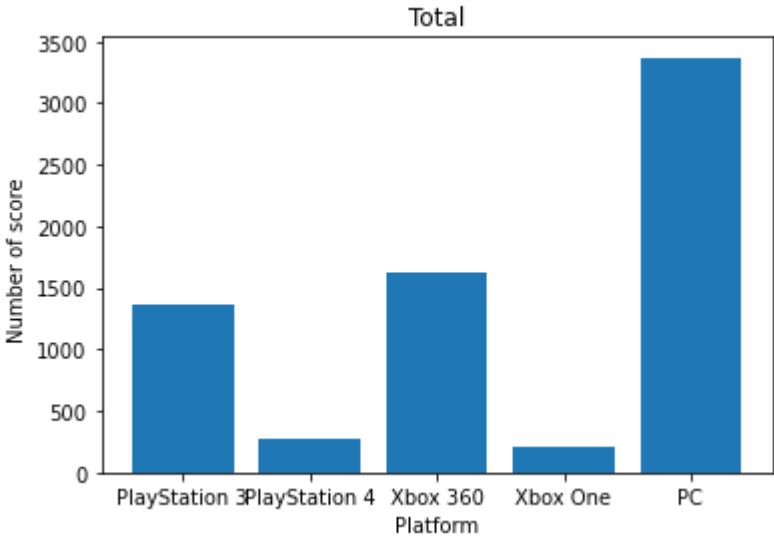
In [53]:

```
plt.bar(x,y1,label='Above avg')  
plt.xlabel("Platform")  
plt.ylabel("Number of score")  
plt.title("Above avg")  
plt.show()
```

```
plt.bar(x,y2,label='Below avg')  
plt.xlabel("Platform")  
plt.ylabel("Number of score")  
plt.title("Below avg")  
plt.show()
```

```
plt.bar(x,y3,label='Total')  
plt.xlabel("Platform")  
plt.ylabel("Number of score ")  
plt.title("Total")  
plt.show()
```





In []: