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

In [2]:
```
df=pd.read_csv('Desktop/قائمة الانجازات/athlete_events.csv')
```

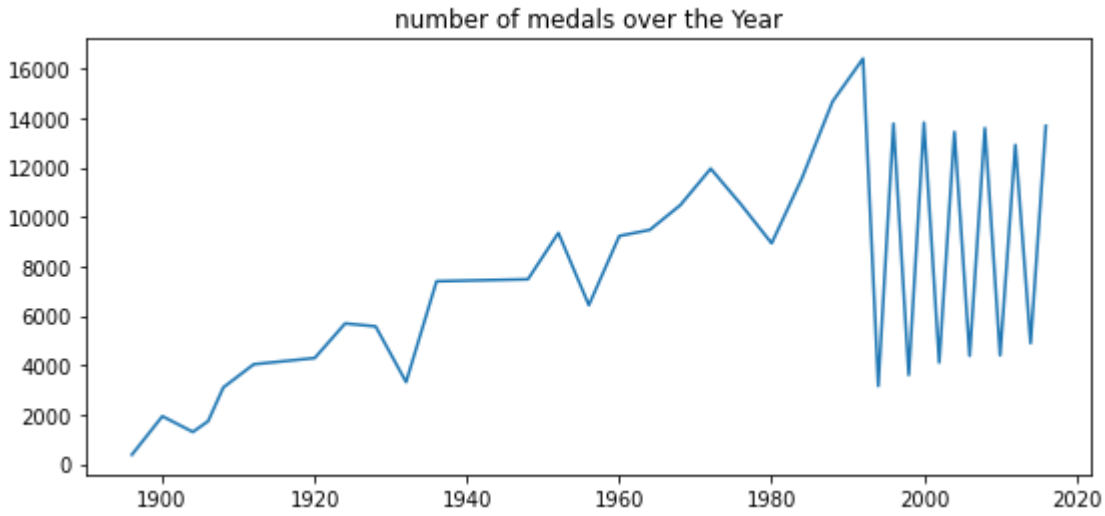
In [40]:
```
df.head()
```

Out[40]:

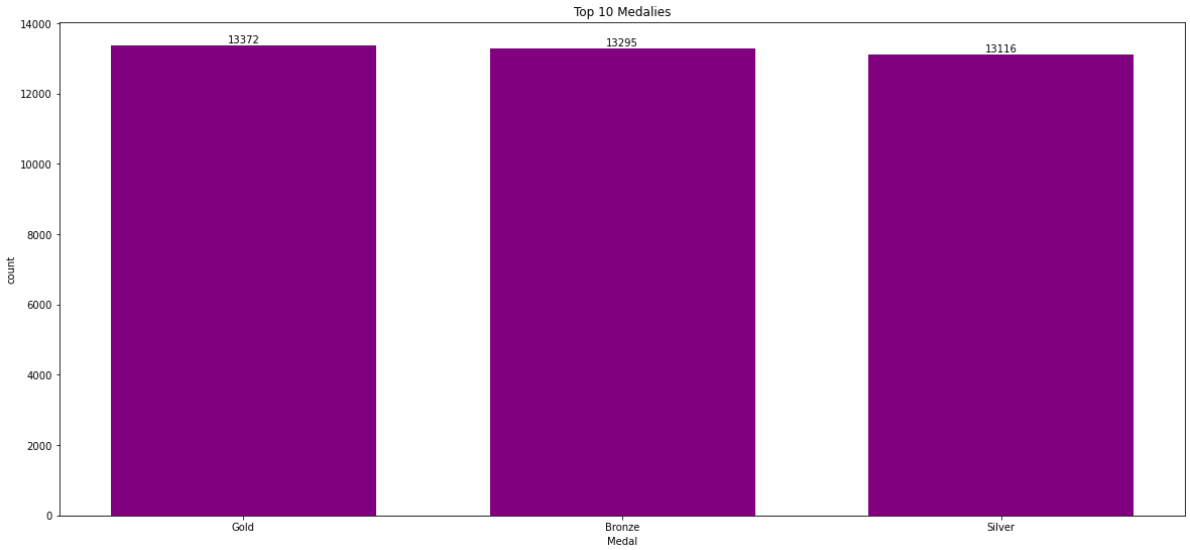
	ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year	Season	
0	1	A Dijiang	M	24.0	180.0	80.0	China	CHN	1992 Summer	1992	Summer	Bar
1	2	A Lamusi	M	23.0	170.0	60.0	China	CHN	2012 Summer	2012	Summer	L
2	3	Gunnar Nielsen Aaby	M	24.0	NaN	NaN	Denmark	DEN	1920 Summer	1920	Summer	Antw
3	4	Edgar Lindenau Aabye	M	34.0	NaN	NaN	Denmark/Sweden	DEN	1900 Summer	1900	Summer	
4	5	Christine Jacoba Aaftink	F	21.0	185.0	82.0	Netherlands	NED	1988 Winter	1988	Winter	C

In [4]:
```
plt.figure(figsize=(9,4))
plt.title('number of medals over the Year')
df.Year.value_counts().sort_index().plot()
```

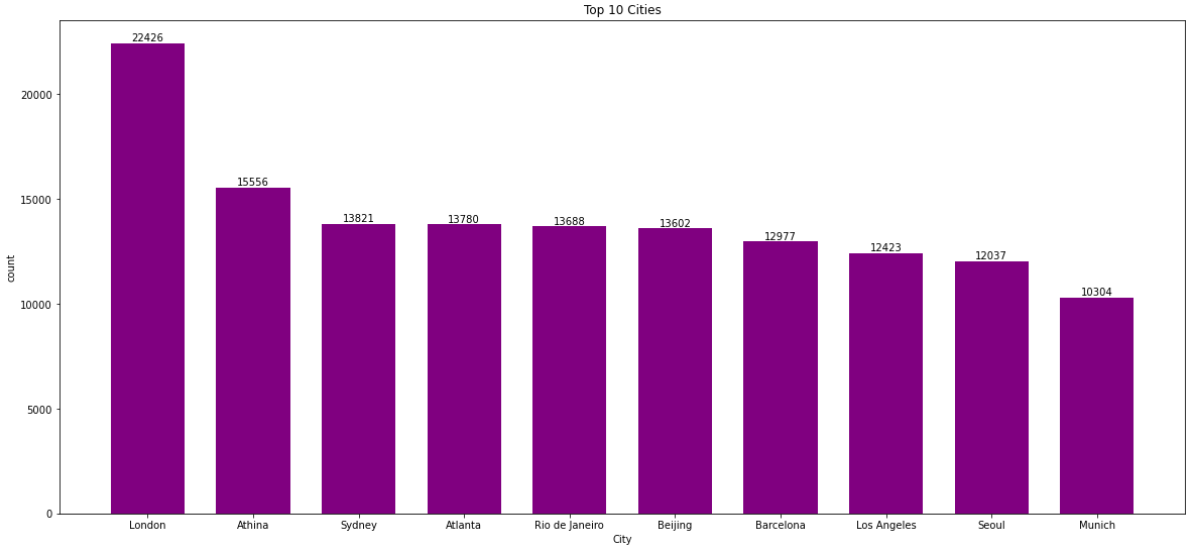
Out[4]:
```
<AxesSubplot:title={'center':'number of medals over the Year'}>
```



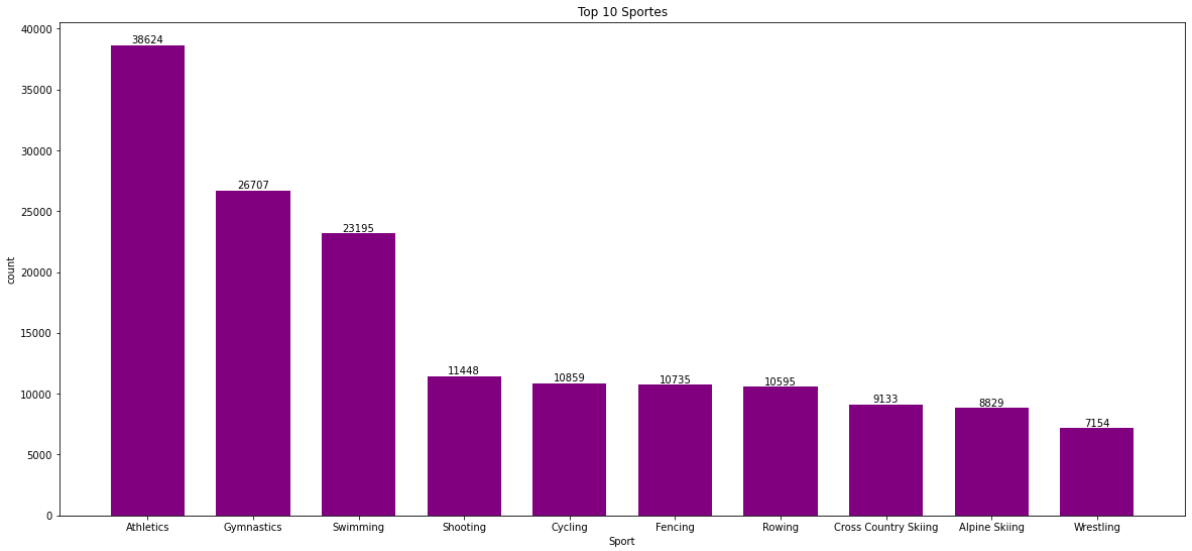
In [6]:
```
itemNames = df['Medal'].value_counts().index[:10]
itemValues = df['Medal'].value_counts().values[:10]
plt.figure(figsize=(20,9))
plt.ylabel('count', fontsize='medium')
plt.xlabel('Medal', fontsize='medium')
plt.title('Top 10 Medalies')
plt.bar(itemNames,itemValues, width = 0.7,color='purple',linewidth=0.4)
for i in range(len(itemNames)):
    plt.text(i,itemValues[i],itemValues[i],ha='center',va='bottom')
plt.show()
```



In [9]:
```
itemNames = df['City'].value_counts().index[:10]
itemValues = df['City'].value_counts().values[:10]
plt.figure(figsize=(20,9))
plt.ylabel('count', fontsize='medium')
plt.xlabel('City', fontsize='medium')
plt.title('Top 10 Cities')
plt.bar(itemNames,itemValues, width = 0.7,color='purple',linewidth=0.4)
for i in range(len(itemNames)):
    plt.text(i,itemValues[i],itemValues[i],ha='center',va='bottom')
plt.show()
```

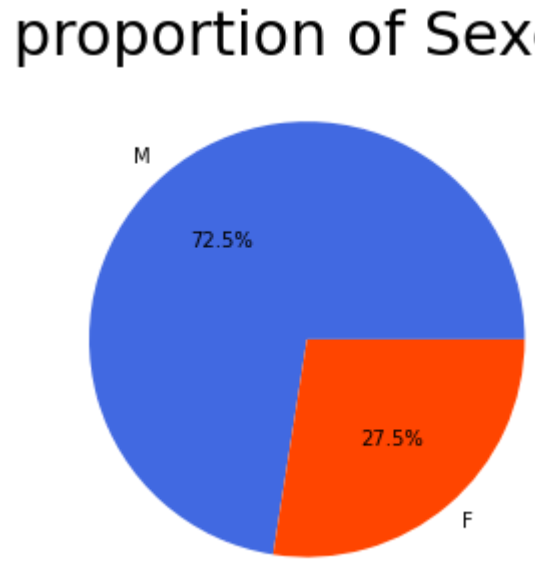


In [8]:
```
itemNames = df['Sport'].value_counts().index[:10]
itemValues = df['Sport'].value_counts().values[:10]
plt.figure(figsize=(20,9))
plt.ylabel('count', fontsize='medium')
plt.xlabel('Sport', fontsize='medium')
plt.title('Top 10 Sportes')
plt.bar(itemNames,itemValues, width = 0.7,color='purple',linewidth=0.4)
for i in range(len(itemNames)):
    plt.text(i,itemValues[i],itemValues[i],ha='center',va='bottom')
plt.show()
```



In [15]:
```
labels = df.Sex.value_counts().index
colors = ['royalblue','orangered']
sex = df.Sex.value_counts().values
plt.figure(figsize = (5,5))
plt.pie(sex, labels=labels, colors=colors, autopct='%1.1f%%')
plt.title('proportion of Sexes',color = 'black',fontsize = 30)
```

Out[15]: Text(0.5, 1.0, 'proportion of Sexes')



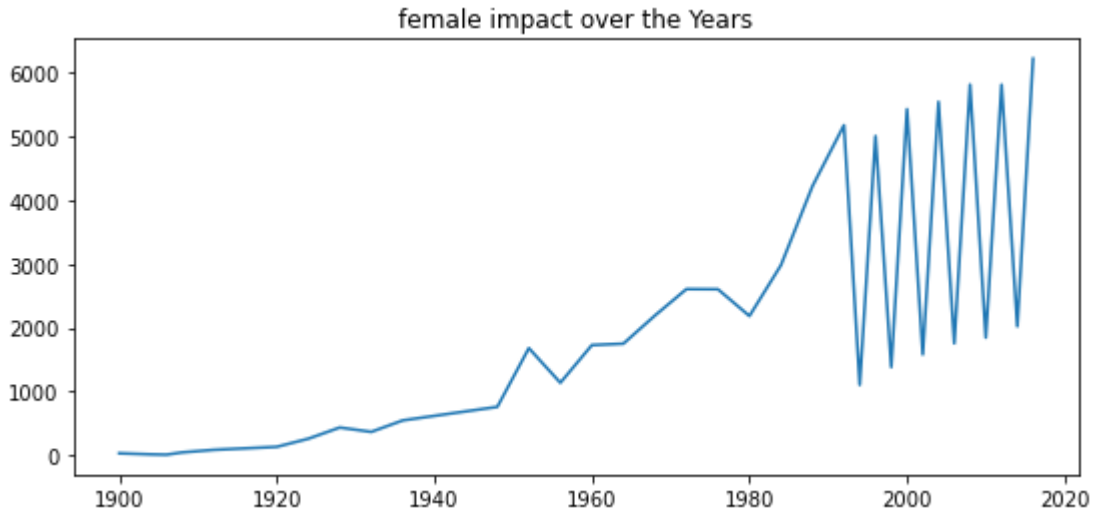
In [16]:
```
female_df=df[df['Sex']== 'F']
female_df.head()
```

Out[16]:

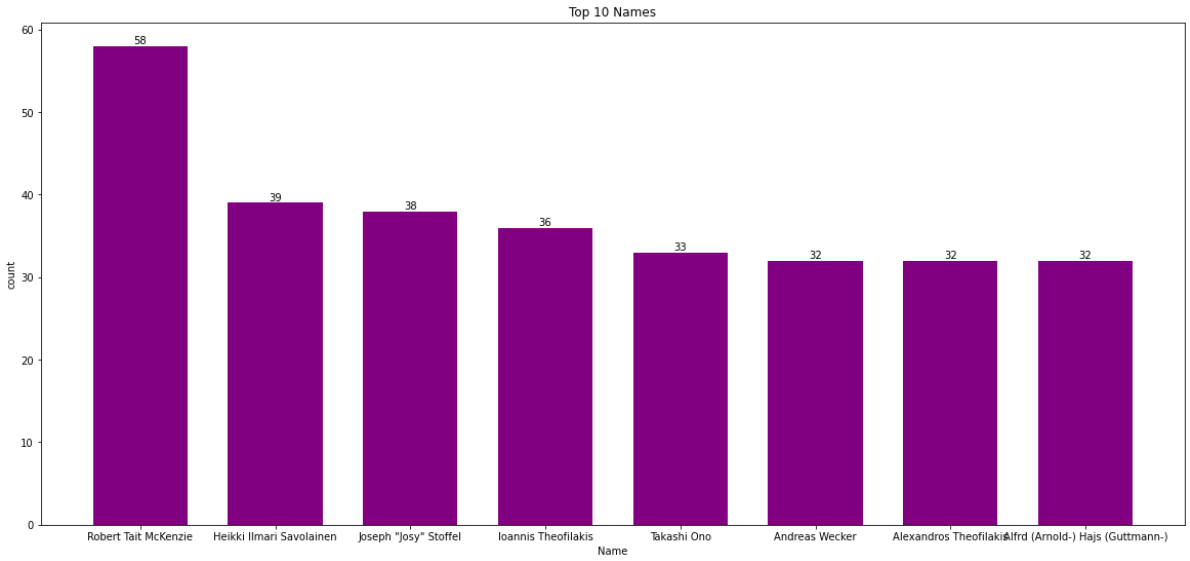
	ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year	Season	City
4	5	Christine Jacoba Aaftink	F	21.0	185.0	82.0	Netherlands	NED	1988 Winter	1988	Winter	Calgary
5	5	Christine Jacoba Aaftink	F	21.0	185.0	82.0	Netherlands	NED	1988 Winter	1988	Winter	Calgary
6	5	Christine Jacoba Aaftink	F	25.0	185.0	82.0	Netherlands	NED	1992 Winter	1992	Winter	Albertville
7	5	Christine Jacoba Aaftink	F	25.0	185.0	82.0	Netherlands	NED	1992 Winter	1992	Winter	Albertville
8	5	Christine Jacoba Aaftink	F	27.0	185.0	82.0	Netherlands	NED	1994 Winter	1994	Winter	Lillehammer

In [17]:
```
plt.figure(figsize=(9,4))
plt.title('female impact over the Years')
female_df.Year.value_counts().sort_index().plot()
```

Out[17]:
```
<AxesSubplot:title={'center':'female impact over the Years'}>
```



In [21]:
```
itemNames = df['Name'].value_counts().index[:8]
itemValues = df['Name'].value_counts().values[:8]
plt.figure(figsize=(20,9))
plt.ylabel('count', fontsize='medium')
plt.xlabel('Name', fontsize='medium')
plt.title('Top 10 Names')
plt.bar(itemNames,itemValues, width = 0.7,color='purple',linewidth=0.4)
for i in range(len(itemNames)):
    plt.text(i,itemValues[i],itemValues[i],ha='center',va='bottom')
plt.show()
```



In [23]:
```
df_gold=df[df['Medal']== 'Gold']
df_gold.head()
```

Out[23]:

	ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year	Season	
3	4	Edgar Lindenau Aabye	M	34.0	NaN	NaN	Denmark/Sweden	DEN	1900 Summer	1900	Summer	
42	17	Paavo Johannes Aaltonen	M	28.0	175.0	64.0	Finland	FIN	1948 Summer	1948	Summer	
44	17	Paavo Johannes Aaltonen	M	28.0	175.0	64.0	Finland	FIN	1948 Summer	1948	Summer	
48	17	Paavo Johannes Aaltonen	M	28.0	175.0	64.0	Finland	FIN	1948 Summer	1948	Summer	
60	20	Kjetil Andr Aamodt	M	20.0	176.0	85.0	Norway	NOR	1992 Winter	1992	Winter	All

In [33]:
```
df_silver=df[df['Medal']== 'Silver']
df_silver.head()
```

Out[33]:

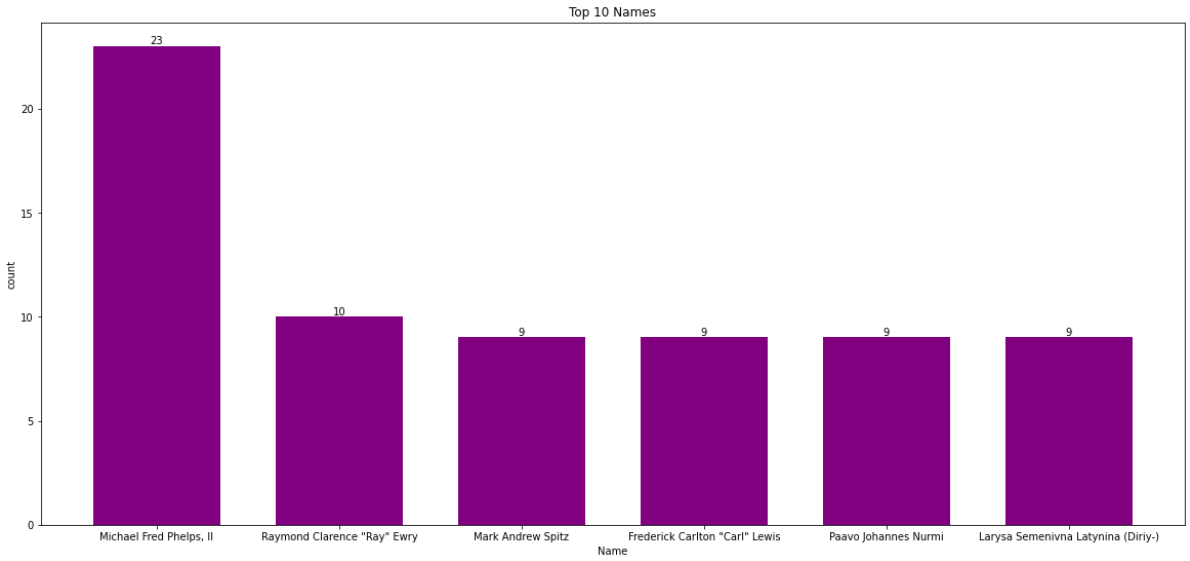
	ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year	Season	
63	20	Kjetil Andr Aamodt	M	22.0	176.0	85.0	Norway	NOR	1994 Winter	1994	Winter	Lilleha
67	20	Kjetil Andr Aamodt	M	22.0	176.0	85.0	Norway	NOR	1994 Winter	1994	Winter	Lilleha
86	25	Alf Lied Aanning	M	24.0	NaN	NaN	Norway	NOR	1920 Summer	1920	Summer	Antw
92	30	Pepijn Aardewijn	M	26.0	189.0	72.0	Netherlands	NED	1996 Summer	1996	Summer	A
106	38	Karl Jan Aas	M	20.0	NaN	NaN	Norway	NOR	1920 Summer	1920	Summer	Antw

In [38]:
```
df_bronze=df[df['Medal']== 'Bronze']
df_bronze.head()
```

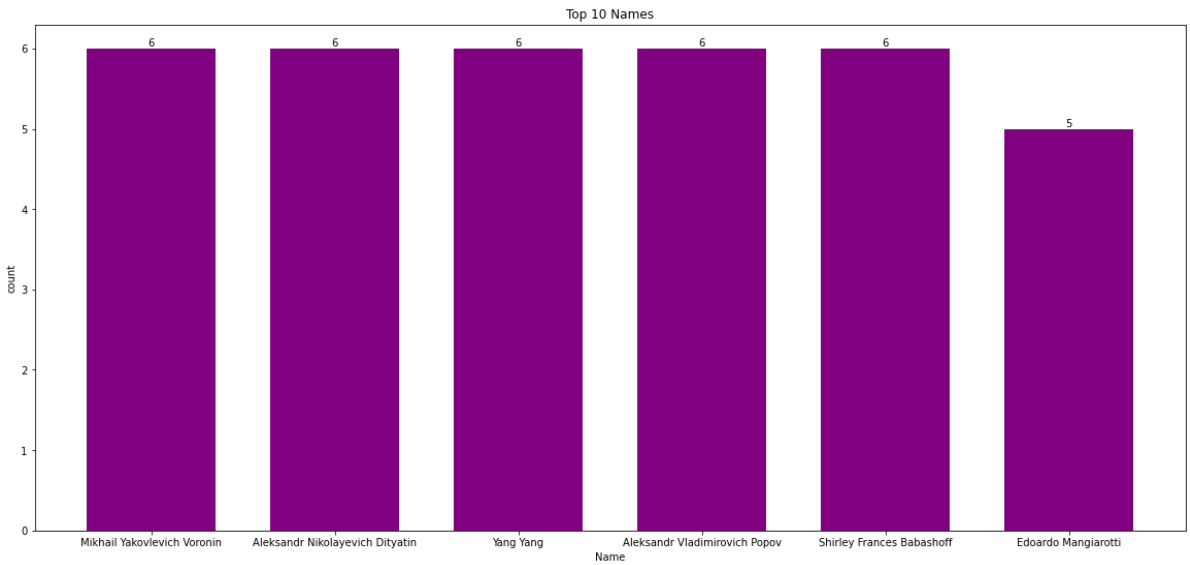
Out[38]:

	ID	Name	Sex	Age	Height	Weight	Team	NOC	Games	Year	Season	City
37	15	Arvo Ossian Aaltonen	M	30.0	NaN	NaN	Finland	FIN	1920 Summer	1920	Summer	Antwerpen
38	15	Arvo Ossian Aaltonen	M	30.0	NaN	NaN	Finland	FIN	1920 Summer	1920	Summer	Antwerpen
40	16	Juhamatti Tapio Aaltonen	M	28.0	184.0	85.0	Finland	FIN	2014 Winter	2014	Winter	Sochi
41	17	Paavo Johannes Aaltonen	M	28.0	175.0	64.0	Finland	FIN	1948 Summer	1948	Summer	London
50	17	Paavo Johannes Aaltonen	M	32.0	175.0	64.0	Finland	FIN	1952 Summer	1952	Summer	Helsinki

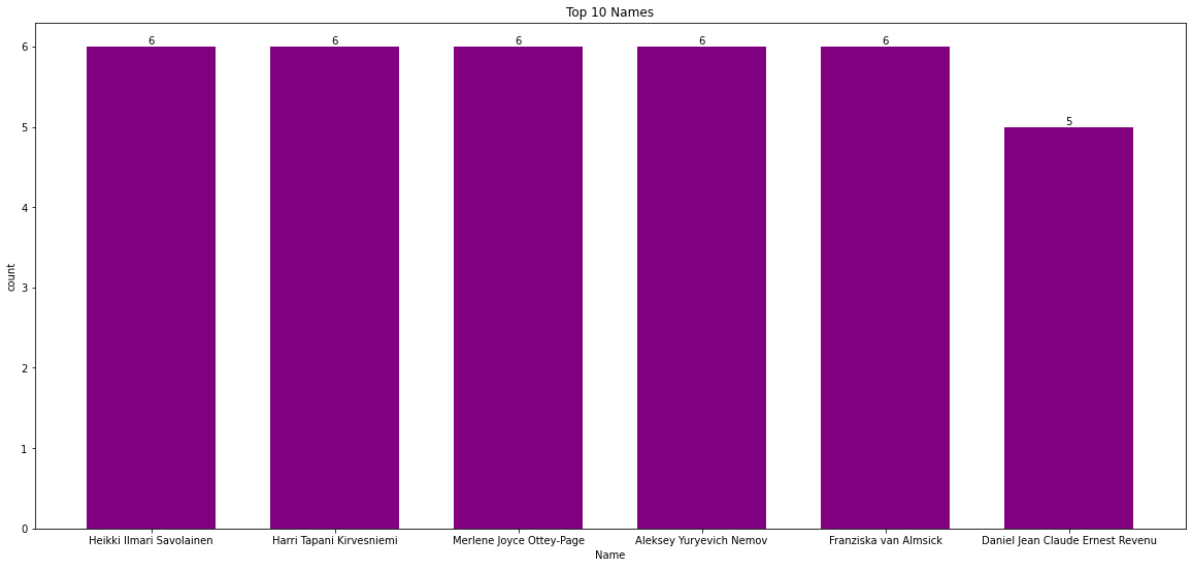
In [29]:
```
itemNames = df_gold['Name'].value_counts().index[:6]
itemValues = df_gold['Name'].value_counts().values[:6]
plt.figure(figsize=(20,9))
plt.ylabel('count', fontsize='medium')
plt.xlabel('Name', fontsize='medium')
plt.title('Top 10 Names')
plt.bar(itemNames,itemValues, width = 0.7,color='purple',linewidth=0.4)
for i in range(len(itemNames)):
    plt.text(i,itemValues[i],itemValues[i],ha='center',va='bottom')
plt.show()
```



In [35]:
```
itemNames = df_silver['Name'].value_counts().index[:6]
itemValues = df_silver['Name'].value_counts().values[:6]
plt.figure(figsize=(20,9))
plt.ylabel('count', fontsize='medium')
plt.xlabel('Name', fontsize='medium')
plt.title('Top 10 Names')
plt.bar(itemNames,itemValues, width = 0.7,color='purple',linewidth=0.4)
for i in range(len(itemNames)):
    plt.text(i,itemValues[i],itemValues[i],ha='center',va='bottom')
plt.show()
```



In [36]:
```
itemNames = df_bronze['Name'].value_counts().index[:6]
itemValues = df_bronze['Name'].value_counts().values[:6]
plt.figure(figsize=(20,9))
plt.ylabel('count', fontsize='medium')
plt.xlabel('Name', fontsize='medium')
plt.title('Top 10 Names')
plt.bar(itemNames,itemValues, width = 0.7,color='purple',linewidth=0.4)
for i in range(len(itemNames)):
    plt.text(i,itemValues[i],itemValues[i],ha='center',va='bottom')
plt.show()
```



Thank You