In [2]:
```
import pandas as pd
bios=pd.read_excel("C:/Users/LENOVO/Downloads/olympics-data.xlsx")
```

In [ ]:
```
# **** value.count() شمردن تعداد تکرار هر مقدار در یک ستون ****

# **** رتبه بندی براساس تعداد تکرار نام شهر ها ****
# bios['born_city'].value_counts()

# **** رتبه بندی براساس تعداد تکرار نام مناطق usa ****
# bios[bios['born_country']=='USA']['born_region'].value_counts()
# bios[bios['born_country']=='USA']['born_region'].value_counts().head()


# bios[(bios['height_cm']==180) | (bios['height_cm']>=180)]['born_city'].value_counts()

# bios[bios['born_country']=='USA']

# bios
```

In [ ]:
```
bios[bios['born_country'] == 'USA']['born_region'].value_counts()
```

Out[ ]:

born_region	
California	1634
New York	990
Illinois	585
Pennsylvania	530
Massachusetts	530
New Jersey	381
Texas	368
Minnesota	365
Ohio	328
Michigan	319
Washington	240
Florida	235
Wisconsin	209
Colorado	207
Connecticut	156
Indiana	150
Oregon	132
Georgia	129
Virginia	121
Maryland	117
District of Columbia	107
Iowa	102
Hawai‘i	95
Kansas	94
Oklahoma	93
Louisiana	92
Utah	91
Missouri	91
North Carolina	86
Arizona	83
New Hampshire	83
Vermont	68
Mississippi	66
Alabama	64
Kentucky	62
Tennessee	62
Nebraska	60
Rhode Island	56
Montana	55
South Carolina	50
Maine	50

```
Alaska                    45
Arkansas                  42
Idaho                     41
New Mexico                38
Nevada                    36
South Dakota              27
West Virginia             24
Delaware                  22
North Dakota              16
Wyoming                   14
Name: count, dtype: int64
```

In [ ]:
```
coffee=pd.read_csv('https://raw.githubusercontent.com/KeithGalli/complete-pandas-tutorial/refs/heads/master/warmup-da
coffee
```

Out[ ]:

	Day	Coffee Type	Units Sold
0	Monday	Espresso	25
1	Monday	Latte	15
2	Tuesday	Espresso	30
3	Tuesday	Latte	20
4	Wednesday	Espresso	35
5	Wednesday	Latte	25
6	Thursday	Espresso	40
7	Thursday	Latte	30
8	Friday	Espresso	45
9	Friday	Latte	35
10	Saturday	Espresso	45
11	Saturday	Latte	35
12	Sunday	Espresso	45
13	Sunday	Latte	35

In [ ]:
```
# **** Grouping ****
# coffee.groupby(['Coffee Type'])['Units Sold'].sum()
# coffee.groupby('Coffee Type')['Units Sold'].sum()
# coffee.groupby('Coffee Type')['Units Sold'].sum()['Latte']
# coffee[coffee['Coffee Type'] == 'Latte']['Units Sold'].sum()
coffee
```

Out[ ]:

	Day	Coffee Type	Units Sold
0	Monday	Espresso	25
1	Monday	Latte	15
2	Tuesday	Espresso	30
3	Tuesday	Latte	20
4	Wednesday	Espresso	35
5	Wednesday	Latte	25
6	Thursday	Espresso	40
7	Thursday	Latte	30
8	Friday	Espresso	45
9	Friday	Latte	35
10	Saturday	Espresso	45
11	Saturday	Latte	35
12	Sunday	Espresso	45
13	Sunday	Latte	35

In [ ]:
```
# coffee.groupby('Coffee Type')['Units Sold'].mean()
coffee.groupby('Coffee Type').agg({'Units Sold': 'sum'})
```

Out[ ]:

Coffee Type	Units Sold
Espresso	265
Latte	195

In [ ]:
```
pivot=coffee.pivot_table(columns='Coffee Type', index='Day', values='Units Sold', aggfunc='sum')

pivot
```

Out[ ]:

Day	Coffee Type	Espresso	Latte
Friday		45	35
Monday		25	15
Saturday		45	35
Sunday		45	35
Thursday		40	30
Tuesday		30	20
Wednesday		35	25

In [ ]:
```
pivot=coffee.pivot_table(columns='Day', index='Coffee Type', values='Units Sold', aggfunc='sum',fill_value='-')

pivot
```

Out[ ]:

Coffee Type	Day	Friday	Monday	Saturday	Sunday	Thursday	Tuesday	Wednesday
Espresso		45	25	45	45	40	30	35
Latte		35	15	35	35	30	20	25

In [ ]:
```
# result = bios.groupby(bios['born_date'].dt.year)['name'].count()
# result=bios.dtype
# print(result)
# result
# print(bios.dtypes)
# **** تبدیل به تاریخ ****
bios['born_date']=pd.to_datetime(bios['born_date'])
```

In [32]:
```
# ***** reset_index()=Generate a new DataFrame or Series with the index reset. *****
# ***** صرفا برای شمارش از ستون name استفاده میکنه *****
result = bios.groupby(bios['born_date'].dt.year)['name'].count().reset_index().sort_values(by='name', ascending=Fals
result
```

Out[32]:

	born_date	name
139	1972.0	2231
152	1985.0	2227
140	1973.0	2216
138	1971.0	2205
137	1970.0	2174
...	...	...
5	1838.0	1
4	1837.0	1
3	1836.0	1
2	1833.0	1
176	2009.0	1

177 rows × 2 columns

In [ ]: