

pandas

10/10

May 28, 2024

```
[ ]: import pandas as pd
```

```
[ ]: country_data = pd.DataFrame(data={
    'Country': ['Belgium', 'India', 'Brazil'],
    'Capital': ['Brussels', 'New Delhi', 'Brasilia'],
    'Population': [11190846, 1303171035, 207847528]
})
country_data
```

```
[ ]:
Country    Capital    Population
0 Belgium  Brussels    11190846
1   India   New Delhi  1303171035
2   Brazil  Brasilia   207847528
```

```
[ ]: #How to use pd.DataFrame.iat to modify or replace a specific value in a
      ↪ DataFrame
country_data.iat[1,1]= 'Mumbai'
print(country_data)
```

```
Country    Capital    Population
0 Belgium  Brussels    11190846
1   India   Mumbai    1303171035
2   Brazil  Brasilia   207847528
```

```
[ ]: #How to select a number of values
country_data.iloc[0:2]
```

```
[ ]:
Country    Capital    Population
0 Belgium  Brussels    11190846
1   India   Mumbai    1303171035
```

```
[ ]: #How to sort and rank data
student_data = pd.DataFrame(data={
    'Name': ['Peter', 'Betty', 'John', 'Joan', 'Grace'],
    'Age' : [15,14,16,15,17],
    'Score': [80,70,90,85,98]
})
student_data
```

```
[ ]:      Name  Age  Score
0  Peter   15    80
1  Betty   14    70
2   John   16    90
3   Joan   15    85
4   Grace  17    98
```

```
[ ]: #How to sort
sorted_data=student_data.sort_values(by='Score',ascending=False)
print(sorted_data)
```

```
      Name  Age  Score
4  Grace   17    98
2   John   16    90
3   Joan   15    85
0  Peter   15    80
1  Betty   14    70
```

```
[ ]: # How to rank
sorted_data['Rank']=sorted_data['Score'].rank(ascending=False)
print(sorted_data)
```

```
      Name  Age  Score  Rank
4  Grace   17    98    1.0
2   John   16    90    2.0
3   Joan   15    85    3.0
0  Peter   15    80    4.0
1  Betty   14    70    5.0
```

```
[ ]: #How to concatenate two or more DataFrames by column
pd.concat([student_data,country_data],axis=1)
```

```
[ ]:      Name  Age  Score  Country  Capital  Population
0  Peter   15    80  Belgium  Brussels  1.119085e+07
1  Betty   14    70   India   Mumbai   1.303171e+09
2   John   16    90  Brazil  Brasilia  2.078475e+08
3   Joan   15    85     NaN     NaN     NaN
4   Grace  17    98     NaN     NaN     NaN
```


```
[ ]: #How to get the mean
student_data.Score.mean()
```

```
[ ]: 84.6
```

```
[ ]: #How to get the median
student_data.Score.median()
```


[]: 85.0

```
[ ]: #How to get the Cumulative sum  
student_data.Score.cumsum()
```




```
[ ]: 0      80  
     1     150  
     2     240  
     3     325  
     4     423  
     Name: Score, dtype: int64
```

```
[ ]: #How to get the minimum  
student_data.Score.min()
```



[]: 70

```
[ ]: #How to get the Maximum  
student_data.Score.max()
```



[]: 98