Import the necessary libraries

In [4]:
```
import pandas as pd
```

Create the DataFrame with the following values:

In [5]:
```
raw_data = {'regiment': ['Nighthawks', 'Nighthawks', 'Nighthawks', 'Nighthawks', 'Dragoons', 'Dragoons'
, 'Dragoons', 'Dragoons', 'Scouts', 'Scouts', 'Scouts', 'Scouts'],
        'company': ['1st', '1st', '2nd', '2nd', '1st', '1st', '2nd', '2nd','1st', '1st', '2nd', '2nd'],

        'name': ['Miller', 'Jacobson', 'Ali', 'Milner', 'Cooze', 'Jacon', 'Ryaner', 'Sone', 'Sloan', 'P
iger', 'Riani', 'Ali'],
        'preTestScore': [4, 24, 31, 2, 3, 4, 24, 31, 2, 3, 2, 3],
        'postTestScore': [25, 94, 57, 62, 70, 25, 94, 57, 62, 70, 62, 70]}
```

Assign it to a variable called regiment.

In [6]:
```
regiment=pd.DataFrame(raw_data)
regiment
```

Out[6]:

	regiment	company	name	preTestScore	postTestScore
0	Nighthawks	1st	Miller	4	25
1	Nighthawks	1st	Jacobson	24	94
2	Nighthawks	2nd	Ali	31	57
3	Nighthawks	2nd	Milner	2	62
4	Dragoons	1st	Cooze	3	70
5	Dragoons	1st	Jacon	4	25
6	Dragoons	2nd	Ryaner	24	94
7	Dragoons	2nd	Sone	31	57
8	Scouts	1st	Sloan	2	62
9	Scouts	1st	Piger	3	70
10	Scouts	2nd	Riani	2	62
11	Scouts	2nd	Ali	3	70

From each regiment,what is the mean preTestScore ?

In [9]:
```
regiment["preTestScore"].mean()
```

Out[9]:
```
11.083333333333334
```

Present general statistics by company

In [26]:
```
p = regiment.groupby(regiment['company']).describe()
print(p)
```

```
        preTestScore                                                    \
               count       mean        std  min   25%   50%     75%   max
company
1st              6.0   6.666667   8.524475  2.0  3.00   3.5   4.00  24.0
2nd              6.0  15.500000  14.652645  2.0  2.25  13.5  29.25  31.0

        postTestScore
                count       mean        std   min    25%   50%   75%   max
company
1st               6.0  57.666667  27.485754  25.0  34.25  66.0  70.0  94.0
2nd               6.0  67.000000  14.057027  57.0  58.25  62.0  68.0  94.0
```

What is the mean of each company's preTestScore?

In [23]:
```
regiment['preTestScore'].groupby([regiment['company']]).mean()
```

Out[23]:
```
company
1st     6.666667
2nd    15.500000
Name: preTestScore, dtype: float64
```

Present the mean preTestScores grouped by regiment and company

In [24]:
```
regiment.groupby(['regiment','company']).mean()
```

Out[24]:

		preTestScore	postTestScore
regiment	**company**		
Dragoons	**1st**	3.5	47.5
	2nd	27.5	75.5
Nighthawks	**1st**	14.0	59.5
	2nd	16.5	59.5
Scouts	**1st**	2.5	66.0
	2nd	2.5	66.0

Present the mean postTestScores grouped by regiment and company

In [27]:
```
regiment.groupby(['regiment','company']).mean()
```

Out[27]:

		preTestScore	postTestScore
regiment	**company**		
Dragoons	**1st**	3.5	47.5
	2nd	27.5	75.5
Nighthawks	**1st**	14.0	59.5
	2nd	16.5	59.5
Scouts	**1st**	2.5	66.0
	2nd	2.5	66.0

Group the entire dataframe by regiment and company

In [37]:
```
regiment.groupby(['regiment','company']).mean()
```

Out[37]:

		preTestScore	postTestScore
regiment	**company**		
Dragoons	**1st**	3.5	47.5
	2nd	27.5	75.5
Nighthawks	**1st**	14.0	59.5
	2nd	16.5	59.5
Scouts	**1st**	2.5	66.0
	2nd	2.5	66.0

What is the number of observations in each regiment and company

In [39]:
```
regiment.groupby(['regiment','company']).size()
```

Out[39]:
```
regiment    company
Dragoons    1st        2
            2nd        2
Nighthawks  1st        2
            2nd        2
Scouts      1st        2
            2nd        2
dtype: int64
```

In [ ]:
```

```