Pandas Groupby

Data Frame

```python
import pandas as pd

data = {'Company':['GOOG','GOOG','MSFT','MSFT','FB','FB'],
        'Person':['Sam','Charlie','Amy','Vanessa','Carl','Sarah'],
        'Sales':[200,120,340,124,243,350]}
df = pd.DataFrame(data)
```

Creatine sample data frame to work with.

	Company	Person	Sales
0	GOOG	Sam	200
1	GOOG	Charlie	120
2	MSFT	Amy	340
3	MSFT	Vanessa	124
4	FB	Carl	243
5	FB	Sarah	350

Pandas Groupby

```python
by_comp = df.groupby("Company")
```

Returns DataFrameGroupBy object

	Company	Person	Sales
0	GOOG	Sam	200
1	GOOG	Charlie	120
2	MSFT	Amy	340
3	MSFT	Vanessa	124
4	FB	Carl	243
5	FB	Sarah	350

Aggregate function

any function that takes many values and then spits out or outputs a single value like ...

```python
by_comp.min()
```

Company	Person	Sales
FB	Carl	243
GOOG	Charlie	120
MSFT	Amy	124

```python
by_comp.max()
```

Company	Person	Sales
FB	Sarah	350
GOOG	Sam	200
MSFT	Vanessa	340

```python
by_comp.count()
```

Company	Person	Sales
FB	2	2
GOOG	2	2
MSFT	2	2

Pandas Groupby

Describe

Gives you bunch of useful information all at once like ***count, mean, min max .. etc***

```
by_comp.describe()
```

	Sales							
	count	mean	std	min	25%	50%	75%	max
Company								
FB	2.0	296.5	75.660426	243.0	269.75	296.5	323.25	350.0
GOOG	2.0	160.0	56.568542	120.0	140.00	160.0	180.00	200.0
MSFT	2.0	232.0	152.735065	124.0	178.00	232.0	286.00	340.0

- Just another way of representation

```
by_comp.describe().transpose()
```

	Company	FB	GOOG	MSFT
Sales	count	2.000000	2.000000	2.000000
	mean	296.500000	160.000000	232.000000
	std	75.660426	56.568542	152.735065
	min	243.000000	120.000000	124.000000
	25%	269.750000	140.000000	178.000000
	50%	296.500000	160.000000	232.000000
	75%	323.250000	180.000000	286.000000
	max	350.000000	200.000000	340.000000