

```
In [6]: import pandas as np
```

```
In [7]: df = np.read_csv('Copy of Copy of Salaries.csv.csv')
```

```
In [8]: df
```

Out[8]:		Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency
	0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco
	1	2	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	538909.28	2011	NaN	San Francisco
	2	3	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	335279.91	2011	NaN	San Francisco
	3	4	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	NaN	332343.61	332343.61	2011	NaN	San Francisco
	4	5	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)	134401.60	9737.00	182234.59	NaN	326373.19	326373.19	2011	NaN	San Francisco

	148649	148650	Roy I Tillery	Custodian	0.00	0.00	0.00	0.0	0.00	0.00	2014	NaN	San Francisco
	148650	148651	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	0.00	2014	NaN	San Francisco
	148651	148652	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	0.00	2014	NaN	San Francisco

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency
148652	148653	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	0.00	2014	NaN	San Francisco
148653	148654	Joe Lopez	Counselor, Log Cabin Ranch	0.00	0.00	-618.13	0.0	-618.13	-618.13	2014	NaN	San Francisco

148654 rows × 13 columns



In [9]: `df.head()`

Out[9]:

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco	NaN
1	2	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	538909.28	2011	NaN	San Francisco	NaN
2	3	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	335279.91	2011	NaN	San Francisco	NaN
3	4	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	NaN	332343.61	332343.61	2011	NaN	San Francisco	NaN
4	5	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)	134401.60	9737.00	182234.59	NaN	326373.19	326373.19	2011	NaN	San Francisco	NaN

In [10]: `df.tail()`

Out[10]:

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
--	-----------	---------------------	-----------------	----------------	--------------------	-----------------	-----------------	-----------------	-------------------------	-------------	--------------	---------------	---------------

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
148649	148650	Roy I Tillery	Custodian	0.0	0.0	0.00	0.0	0.00	0.00	2014	NaN	San Francisco	NaN
148650	148651	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	0.00	2014	NaN	San Francisco	NaN
148651	148652	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	0.00	2014	NaN	San Francisco	NaN
148652	148653	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	0.00	2014	NaN	San Francisco	NaN
148653	148654	Joe Lopez	Counselor, Log Cabin Ranch	0.0	0.0	-618.13	0.0	-618.13	-618.13	2014	NaN	San Francisco	NaN

In [11]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148654 entries, 0 to 148653
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Id                    148654 non-null  int64
1   EmployeeName          148654 non-null  object
2   JobTitle               148654 non-null  object
3   BasePay               148045 non-null  float64
4   OvertimePay           148650 non-null  float64
5   OtherPay              148650 non-null  float64
6   Benefits              112491 non-null  float64
7   TotalPay              148654 non-null  float64
8   TotalPayBenefits      148654 non-null  float64
9   Year                  148654 non-null  int64
10  Notes                 0 non-null       float64
11  Agency                148654 non-null  object
12  Status                0 non-null       float64
dtypes: float64(8), int64(2), object(3)
memory usage: 14.7+ MB
```

In [13]: `df.groupby('Year')`

Out[13]: `<pandas.core.groupby.generic.DataFrameGroupBy object at 0x000000EDD0DA1A30>`

```
In [14]: df.groupby('Year').mean()
```

```
Out[14]:
```

	Id	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Notes	Status
Year									
2011	18080.0	63595.956517	4531.065429	3617.081926	NaN	71744.103871	71744.103871	NaN	NaN
2012	54542.5	65436.406857	5023.417824	3653.437583	26439.966967	74113.262265	100553.229232	NaN	NaN
2013	91728.5	69630.030216	5281.641980	3819.969007	23829.076572	77611.443142	101440.519714	NaN	NaN
2014	129593.0	66564.421924	5401.993737	3505.421251	24789.601756	75463.918140	100250.918884	NaN	NaN

```
In [ ]:
```