

```
In [1]: #Data Analysis With Python Pandas ----- Employee Salary DataSet
```

```
In [2]: import numpy as np  
import pandas as pd
```

```
In [3]: pd.set_option("display.max_columns",None)
df=pd.read_csv("E:\\KVR-PANDAS-EXCLUSIVE\\EMPLOYEE SALARAY ANALYSIS PROJECT\\Salaries.csv")
print(df)
```

	Id	EmployeeName	\
0	1	NATHANIEL FORD	
1	2	GARY JIMENEZ	
2	3	ALBERT PARDINI	
3	4	CHRISTOPHER CHONG	
4	5	PATRICK GARDNER	
...	
148649	148650	Roy I Tillery	
148650	148651	Not provided	
148651	148652	Not provided	
148652	148653	Not provided	
148653	148654	Joe Lopez	

	JobTitle	BasePay	\
0	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	
1	CAPTAIN III (POLICE DEPARTMENT)	155966.02	
2	CAPTAIN III (POLICE DEPARTMENT)	212739.13	
3	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	
4	DEPUTY CHIEF OF DEPARTMENT,(FIRE DEPARTMENT)	134401.60	
...	
148649	Custodian	0.00	
148650	Not provided	NaN	
148651	Not provided	NaN	
148652	Not provided	NaN	
148653	Counselor, Log Cabin Ranch	0.00	

	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	\
0	0.00	400184.25	NaN	567595.43	567595.43	2011	
1	245131.88	137811.38	NaN	538909.28	538909.28	2011	
2	106088.18	16452.60	NaN	335279.91	335279.91	2011	
3	56120.71	198306.90	NaN	332343.61	332343.61	2011	
4	9737.00	182234.59	NaN	326373.19	326373.19	2011	
...	
148649	0.00	0.00	0.0	0.00	0.00	2014	
148650	NaN	NaN	NaN	0.00	0.00	2014	
148651	NaN	NaN	NaN	0.00	0.00	2014	
148652	NaN	NaN	NaN	0.00	0.00	2014	
148653	0.00	-618.13	0.0	-618.13	-618.13	2014	

	Notes	Agency	Status
0	NaN	San Francisco	NaN
1	NaN	San Francisco	NaN
2	NaN	San Francisco	NaN
3	NaN	San Francisco	NaN
4	NaN	San Francisco	NaN
...
148649	NaN	San Francisco	NaN
148650	NaN	San Francisco	NaN
148651	NaN	San Francisco	NaN
148652	NaN	San Francisco	NaN
148653	NaN	San Francisco	NaN

[148654 rows x 13 columns]

In [4]: `df.head(2)`

Out[4]:

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBei
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	5675
1	2	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	5389

In [5]: `#1. Display Top 10 Rows of The Dataset`
`df.head(10)`

Out[5]:

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBei
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	5675
1	2	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	5389
2	3	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	3352
3	4	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	NaN	332343.61	3323
4	5	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)	134401.60	9737.00	182234.59	NaN	326373.19	3263
5	6	DAVID SULLIVAN	ASSISTANT DEPUTY CHIEF II	118602.00	8601.00	189082.74	NaN	316285.74	3162
6	7	ALSON LEE	BATTALION CHIEF, (FIRE DEPARTMENT)	92492.01	89062.90	134426.14	NaN	315981.05	3159
7	8	DAVID KUSHNER	DEPUTY DIRECTOR OF INVESTMENTS	256576.96	0.00	51322.50	NaN	307899.46	3078
8	9	MICHAEL MORRIS	BATTALION CHIEF, (FIRE DEPARTMENT)	176932.64	86362.68	40132.23	NaN	303427.55	3034
9	10	JOANNE HAYES-WHITE	CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)	285262.00	0.00	17115.73	NaN	302377.73	3023

In [6]: *#2. Check Last 10 Rows of The Dataset*
`df.tail(10)`

Out[6]:

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	Total
148644	148645	Randy D Winn	Stationary Eng, Sewage Plant	0.0	0.0	0.00	0.0	0.00	
148645	148646	Carolyn A Wilson	Human Services Technician	0.0	0.0	0.00	0.0	0.00	
148646	148647	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	
148647	148648	Joann Anderson	Communications Dispatcher 2	0.0	0.0	0.00	0.0	0.00	
148648	148649	Leon Walker	Custodian	0.0	0.0	0.00	0.0	0.00	
148649	148650	Roy I Tillery	Custodian	0.0	0.0	0.00	0.0	0.00	
148650	148651	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	
148651	148652	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	
148652	148653	Not provided	Not provided	NaN	NaN	NaN	NaN	0.00	
148653	148654	Joe Lopez	Counselor, Log Cabin Ranch	0.0	0.0	-618.13	0.0	-618.13	

In [7]: *#3. Find Shape of Our Dataset (Number of Rows And Number of Columns)*
`df.shape`

Out[7]: (148654, 13)

In [8]: `print("Number of Rows=",df.shape[0])`
`print("Number of Columns=",df.shape[1])`

Number of Rows= 148654
 Number of Columns= 13

In [9]: *#4. Getting Information About Our Dataset Like Total Number Rows, Total Number of Columns, Datatypes of Each Column And Memory Requirement*
`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 [10]: `df.index`

Out[10]: RangeIndex(start=0, stop=148654, step=1)

In [11]: `df.columns`

Out[11]: Index(['Id', 'EmployeeName', 'JobTitle', 'BasePay', 'OvertimePay', 'OtherPay', 'Benefits', 'TotalPay', 'TotalPayBenefits', 'Year', 'Notes', 'Agency', 'Status'], dtype='object')

In [13]: *#5. Check Null Values In The Dataset*
`df.isnull().sum()`

```
Out[13]: Id                    0
EmployeeName                  0
JobTitle                      0
BasePay                       609
OvertimePay                    4
OtherPay                       4
Benefits                      36163
TotalPay                       0
TotalPayBenefits               0
Year                           0
Notes                        148654
Agency                        0
Status                       148654
dtype: int64
```

```
In [14]: #6. Drop ID, Notes, Agency, and Status Columns
df.columns
```

```
Out[14]: Index(['Id', 'EmployeeName', 'JobTitle', 'BasePay', 'OvertimePay', 'OtherPay',
               'Benefits', 'TotalPay', 'TotalPayBenefits', 'Year', 'Notes', 'Agency',
               'Status'],
              dtype='object')
```

```
In [15]: df=df.drop(["Id","Notes","Agency","Status"],axis=1)
```

```
In [16]: print(df)
```

	EmployeeName	JobTitle
0	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY
1	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)
2	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)
3	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC
4	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT,(FIRE DEPARTMENT)
...
148649	Roy I Tillery	Custodian
148650	Not provided	Not provided
148651	Not provided	Not provided
148652	Not provided	Not provided
148653	Joe Lopez	Counselor, Log Cabin Ranch

	BasePay	OvertimePay	OtherPay	Benefits	TotalPay
0	167411.18	0.00	400184.25	NaN	567595.43
1	155966.02	245131.88	137811.38	NaN	538909.28
2	212739.13	106088.18	16452.60	NaN	335279.91
3	77916.00	56120.71	198306.90	NaN	332343.61
4	134401.60	9737.00	182234.59	NaN	326373.19
...
148649	0.00	0.00	0.00	0.0	0.00
148650	NaN	NaN	NaN	NaN	0.00
148651	NaN	NaN	NaN	NaN	0.00
148652	NaN	NaN	NaN	NaN	0.00
148653	0.00	0.00	-618.13	0.0	-618.13

	TotalPayBenefits	Year
0	567595.43	2011
1	538909.28	2011
2	335279.91	2011
3	332343.61	2011
4	326373.19	2011
...
148649	0.00	2014
148650	0.00	2014
148651	0.00	2014
148652	0.00	2014
148653	-618.13	2014

```
[148654 rows x 9 columns]
```

In [18]: `df.head()`

Out[18]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefit
0	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	567595.4
1	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	538909.2
2	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	335279.9
3	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	NaN	332343.61	332343.6
4	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)	134401.60	9737.00	182234.59	NaN	326373.19	326373.1

In [19]: `df.columns`

Out[19]: Index(['EmployeeName', 'JobTitle', 'BasePay', 'OvertimePay', 'OtherPay', 'Benefits', 'TotalPay', 'TotalPayBenefits', 'Year'], dtype='object')

In [20]: `len(df.columns)`

Out[20]: 9

In [21]: *#7. Get Overall Statistics About The Dataframe(Descriptive Statistics)*
`df.describe() # for Numerical Columns only`

Out[21]:

	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	
count	148045.000000	148650.000000	148650.000000	112491.000000	148654.000000	148654.000000	148654
mean	66325.448840	5066.059886	3648.767297	25007.893151	74768.321972	93692.554811	2014
std	42764.635495	11454.380559	8056.601866	15402.215858	50517.005274	62793.533483	
min	-166.010000	-0.010000	-7058.590000	-33.890000	-618.130000	-618.130000	2014
25%	33588.200000	0.000000	0.000000	11535.395000	36168.995000	44065.650000	2014
50%	65007.450000	0.000000	811.270000	28628.620000	71426.610000	92404.090000	2014
75%	94691.050000	4658.175000	4236.065000	35566.855000	105839.135000	132876.450000	2014
max	319275.010000	245131.880000	400184.250000	96570.660000	567595.430000	567595.430000	2014

In [22]: *#7. Get Overall Statistics About The Dataframe(Descriptive Statistics)*
`df.describe(include="all")` *# for Numerical and Non-numerical Columns*

Out[22]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay
count	148654	148654	148045.000000	148650.000000	148650.000000	112491.000000	148654.000000
unique	110811	2159	NaN	NaN	NaN	NaN	NaN
top	Kevin Lee	Transit Operator	NaN	NaN	NaN	NaN	NaN
freq	13	7036	NaN	NaN	NaN	NaN	NaN
mean	NaN	NaN	66325.448840	5066.059886	3648.767297	25007.893151	74768.321900
std	NaN	NaN	42764.635495	11454.380559	8056.601866	15402.215858	50517.005200
min	NaN	NaN	-166.010000	-0.010000	-7058.590000	-33.890000	-618.130000
25%	NaN	NaN	33588.200000	0.000000	0.000000	11535.395000	36168.995000
50%	NaN	NaN	65007.450000	0.000000	811.270000	28628.620000	71426.610000
75%	NaN	NaN	94691.050000	4658.175000	4236.065000	35566.855000	105839.135000
max	NaN	NaN	319275.010000	245131.880000	400184.250000	96570.660000	567595.430000

In [25]: *#8. Find Occurrence of The Employee Names (Top 5)*
`df["EmployeeName"].value_counts()`

Out[25]: EmployeeName
 Kevin Lee 13
 Richard Lee 11
 Steven Lee 11
 William Wong 11
 Stanley Lee 9
 ..
 Jeffrey Skover 1
 Ken Dever 1
 Kevin Whitfield 1
 Granville McCollough III 1
 Joe Lopez 1
 Name: count, Length: 110811, dtype: int64

In [27]: *#8. Find Occurrence of The Employee Names (Top 5)*
`df["EmployeeName"].value_counts().head()`

Out[27]: EmployeeName
 Kevin Lee 13
 Richard Lee 11
 Steven Lee 11
 William Wong 11
 Stanley Lee 9
 Name: count, dtype: int64

In [29]: *#9. Find The Number of Unique Job Titles*
`df['JobTitle'].nunique()`

Out[29]: 2159


```
In [35]: #10.Total Number of Job Titles Contain Captain
df[df['JobTitle'].str.contains("CAPTAIN")]
```

Out[35]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBe
1	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	5389
2	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	3352
11	PATRICIA JACKSON	CAPTAIN III (POLICE DEPARTMENT)	99722.00	87082.62	110804.30	NaN	297608.92	2976
17	SEBASTIAN WONG	CAPTAIN, EMERGENCYCY MEDICAL SERVICES	140546.87	119397.26	18625.08	NaN	278569.21	2785
22	GEORGE GARCIA	CAPTAIN, FIRE SUPPRESSION	140546.88	93200.58	39955.25	NaN	273702.71	2737
...
8684	JEANNE SEYLER	CAPTAIN, FIRE SUPPRESSION	95055.34	0.00	9197.14	NaN	104252.48	1042
10485	JANE SMITH	CAPTAIN, EMERGENCYCY MEDICAL SERVICES	74592.00	1538.59	18804.68	NaN	94935.27	949
11198	KATHRYN BROWN	CAPTAIN III (POLICE DEPARTMENT)	10684.50	0.00	81244.87	NaN	91929.37	919
31297	MARCO CARNIGLIA	CAPTAIN, EMERGENCYCY MEDICAL SERVICES	9839.72	0.00	1203.77	NaN	11043.49	110
34124	JOHN FORBES- 3	CAPTAIN, FIRE SUPPRESSION	0.00	982.06	2277.34	NaN	3259.40	32

141 rows × 9 columns



```
In [36]: #10.Total Number of Job Titles Contain Captain
df[df['JobTitle'].str.contains("CAPTAIN")].count()
```

```
Out[36]: EmployeeName      141
JobTitle      141
BasePay      141
OvertimePay  141
OtherPay     141
Benefits      0
TotalPay     141
TotalPayBenefits 141
Year         141
dtype: int64
```

```
In [40]: #10.Total Number of Job Titles Contain Captain
df[df['JobTitle'].str.contains("CAPTAIN",case=False)].count()
```

```
Out[40]: EmployeeName      552
JobTitle      552
BasePay      551
OvertimePay   552
OtherPay      552
Benefits      411
TotalPay      552
TotalPayBenefits 552
Year      552
dtype: int64
```

```
In [43]: #10.Total Number of Job Titles Contain Captain
len(df[df['JobTitle'].str.contains("CAPTAIN",case=False)])
```

```
Out[43]: 552
```

```
In [54]: #Display ALL the Employee Names From Fire Department
df.loc[df['JobTitle'].str.contains("Fire",case=False)][["EmployeeName", 'JobTitle']]
```

```
Out[54]:
```

	EmployeeName	JobTitle
4	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT,(FIRE DEPARTMENT)
6	ALSON LEE	BATTALION CHIEF, (FIRE DEPARTMENT)
8	MICHAEL MORRIS	BATTALION CHIEF, (FIRE DEPARTMENT)
9	JOANNE HAYES-WHITE	CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)
10	ARTHUR KENNEY	ASSISTANT CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)
...
145956	Kenneth C Farris	Firefighter
147556	Edward A Dunn	Firefighter
148021	Kari A Johnson	Firefighter
148209	Sheryl K Lee	Firefighter
148554	Lawrence F Gatt	Fire Alarm Dispatcher

5879 rows × 2 columns

```
In [55]: #Display Number Employee Names From Fire Department
len(df.loc[df['JobTitle'].str.contains("Fire",case=False)][["EmployeeName", 'JobTitle']])
```

```
Out[55]: 5879
```

```
In [58]: #12. Find Minimum, Maximum, and Average BasePay
df['BasePay'].max()
```

```
Out[58]: 319275.01
```

```
In [59]: #12. Find Minimum, Maximum, and Average BasePay
df['BasePay'].min()
```

Out[59]: -166.01

```
In [60]: df['BasePay'].mean()
```

Out[60]: 66325.4488404877

```
In [61]: #OR
df.describe()
```

Out[61]:

	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	
count	148045.000000	148650.000000	148650.000000	112491.000000	148654.000000	148654.000000	148654
mean	66325.448840	5066.059886	3648.767297	25007.893151	74768.321972	93692.554811	2011
std	42764.635495	11454.380559	8056.601866	15402.215858	50517.005274	62793.533483	2011
min	-166.010000	-0.010000	-7058.590000	-33.890000	-618.130000	-618.130000	2011
25%	33588.200000	0.000000	0.000000	11535.395000	36168.995000	44065.650000	2011
50%	65007.450000	0.000000	811.270000	28628.620000	71426.610000	92404.090000	2011
75%	94691.050000	4658.175000	4236.065000	35566.855000	105839.135000	132876.450000	2011
max	319275.010000	245131.880000	400184.250000	96570.660000	567595.430000	567595.430000	2011

```
In [63]: #13. Replace 'Not Provided' in EmployeeName' Column to NaN
df.loc[df['EmployeeName']=='Not Provided']
```

Out[63]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayB
13797	Not Provided	PUBLIC SAFETY COMMUNICATIONS DISPATCHER	57963.39	19683.91	4373.86	NaN	82021.16	82021.16
44151	Not Provided	Public SafetyComm Disp	82196.52	25934.20	3752.49	38024.3	111883.21	149817.71

```
In [66]: import numpy as np
df['EmployeeName']=df['EmployeeName'].str.replace('Not Provided',"np.NaN")
```

```
In [68]: df.loc[df['EmployeeName']=="np.NaN"]
```

Out[68]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayB
13797	np.NaN	PUBLIC SAFETY COMMUNICATIONS DISPATCHER	57963.39	19683.91	4373.86	NaN	82021.16	82021.16
44151	np.NaN	Public SafetyComm Disp	82196.52	25934.20	3752.49	38024.3	111883.21	149817.71

```
In [69]: df['EmployeeName']=df['EmployeeName'].str.replace("np.NaN", "MS DHONI")
```

```
In [70]: df.loc[df['EmployeeName']=="MS DHONI"]
```

Out[70]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayB
13797	MS DHONI	PUBLIC SAFETY COMMUNICATIONS DISPATCHER	57963.39	19683.91	4373.86	NaN	82021.16	82
44151	MS DHONI	Public SafetyComm Disp	82196.52	25934.20	3752.49	38024.3	111883.21	149

```
In [77]: #14. Drop The Rows Having 4 Missing Values
```

```
df[df.isnull().sum(axis=1)==4]
```

Out[77]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Ye
148646	Not provided	Not provided	NaN	NaN	NaN	NaN	0.0	0.0	20
148650	Not provided	Not provided	NaN	NaN	NaN	NaN	0.0	0.0	20
148651	Not provided	Not provided	NaN	NaN	NaN	NaN	0.0	0.0	20
148652	Not provided	Not provided	NaN	NaN	NaN	NaN	0.0	0.0	20

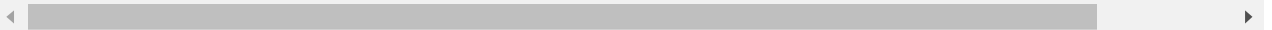
```
In [79]: #df.drop(df[df.isnull().sum(axis=1)==4],axis=0) # Gives Key Error
```

```
In [80]: df.drop(df[df.isnull().sum(axis=1)==4].index,axis=0)
```

Out[80]:

	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayE
0	NATHANIEL FORD	GENERAL MANAGER- METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	567
1	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	538
2	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	335
3	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	NaN	332343.61	332
4	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)	134401.60	9737.00	182234.59	NaN	326373.19	326
...
148645	Carolyn A Wilson	Human Services Technician	0.00	0.00	0.00	0.0	0.00	
148647	Joann Anderson	Communications Dispatcher 2	0.00	0.00	0.00	0.0	0.00	
148648	Leon Walker	Custodian	0.00	0.00	0.00	0.0	0.00	
148649	Roy I Tillery	Custodian	0.00	0.00	0.00	0.0	0.00	
148653	Joe Lopez	Counselor, Log Cabin Ranch	0.00	0.00	-618.13	0.0	-618.13	

148650 rows × 9 columns



```
In [81]: df.drop(df[df.isnull().sum(axis=1)==4].index,axis=0,inplace=True)
```

In [82]: `print(df)`

```

      EmployeeName      JobTitle \
0      NATHANIEL FORD  GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY
1      GARY JIMENEZ      CAPTAIN III (POLICE DEPARTMENT)
2      ALBERT PARDINI      CAPTAIN III (POLICE DEPARTMENT)
3      CHRISTOPHER CHONG      WIRE ROPE CABLE MAINTENANCE MECHANIC
4      PATRICK GARDNER  DEPUTY CHIEF OF DEPARTMENT,(FIRE DEPARTMENT)
...
148645  Carolyn A Wilson      Human Services Technician
148647  Joann Anderson      Communications Dispatcher 2
148648  Leon Walker      Custodian
148649  Roy I Tillery      Custodian
148653  Joe Lopez      Counselor, Log Cabin Ranch

      BasePay  OvertimePay  OtherPay  Benefits  TotalPay \
0      167411.18      0.00  400184.25      NaN  567595.43
1      155966.02      245131.88  137811.38      NaN  538909.28
2      212739.13      106088.18   16452.60      NaN  335279.91
3       77916.00      56120.71  198306.90      NaN  332343.61
4      134401.60      9737.00  182234.59      NaN  326373.19
...
148645      0.00      0.00      0.00      0.0      0.00
148647      0.00      0.00      0.00      0.0      0.00
148648      0.00      0.00      0.00      0.0      0.00
148649      0.00      0.00      0.00      0.0      0.00
148653      0.00      0.00     -618.13      0.0     -618.13

      TotalPayBenefits  Year
0      567595.43  2011
1      538909.28  2011
2      335279.91  2011
3      332343.61  2011
4      326373.19  2011
...
148645      0.00  2014
148647      0.00  2014
148648      0.00  2014
148649      0.00  2014
148653     -618.13  2014

```

[148650 rows x 9 columns]

In []:

In []: