In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sb
```

In [2]:
```
iris = pd.read_csv("/home/student/Desktop/Iris.csv")
```

In [3]:
```
iris.head(6)
```

Out[3]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa

In [4]:
```
iris.mean()
```

Out[4]:
```
Id               75.500000
SepalLengthCm     5.843333
SepalWidthCm      3.054000
PetalLengthCm     3.758667
PetalWidthCm      1.198667
dtype: float64
```

In [5]:
```
iris.loc[:,'SepalLengthCm'].mean()
```

Out[5]:
```
5.843333333333334
```

In [7]:
```
iris.mean(axis=1)[0:4]
```

Out[7]:
```
0    2.24
1    2.30
2    2.48
3    2.68
dtype: float64
```

In [8]:
```
iris.median()
```

Out[8]:
```
Id               75.50
SepalLengthCm     5.80
SepalWidthCm      3.00
PetalLengthCm     4.35
PetalWidthCm      1.30
dtype: float64
```

In [11]:
```
iris.loc[:,'Id'].median()
```

Out[11]:
```
75.5
```

In [13]:
```
iris.median(axis=1)[0:6]
```

Out[13]:
```
0    1.4
1    2.0
2    3.0
3    3.1
4    3.6
5    3.9
dtype: float64
```

In [14]:
```
iris.mode()
```

Out[14]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.0	3.0	1.5	0.2	Iris-setosa
1	2	NaN	NaN	NaN	NaN	Iris-versicolor
2	3	NaN	NaN	NaN	NaN	Iris-virginica
3	4	NaN	NaN	NaN	NaN	NaN
4	5	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...
145	146	NaN	NaN	NaN	NaN	NaN
146	147	NaN	NaN	NaN	NaN	NaN
147	148	NaN	NaN	NaN	NaN	NaN
148	149	NaN	NaN	NaN	NaN	NaN
149	150	NaN	NaN	NaN	NaN	NaN

150 rows × 6 columns

In [15]:
```
iris.loc[:,'PetalLengthCm'].mode()
```

Out[15]:
```
0    1.5
Name: PetalLengthCm, dtype: float64
```

In [17]:
```
iris.min()
```

Out[17]:
```
Id                      1
SepalLengthCm         4.3
SepalWidthCm          2.0
PetalLengthCm         1.0
PetalWidthCm          0.1
Species       Iris-setosa
dtype: object
```

In [18]:
```
iris.max()
```

Out[18]:
```
Id                    150
SepalLengthCm         7.9
SepalWidthCm          4.4
PetalLengthCm         6.9
PetalWidthCm          2.5
Species       Iris-virginica
dtype: object
```

In [19]:
```
iris.std()
```

Out[19]:
```
Id               43.445368
SepalLengthCm     0.828066
SepalWidthCm      0.433594
PetalLengthCm     1.764420
PetalWidthCm      0.763161
dtype: float64
```

In [20]:
```
iris.loc[:,'SepalWidthCm'].std()
```

Out[20]:
```
0.4335943113621737
```

In [22]:
```
iris.std(axis=1)[0:7]
```

Out[22]:
```
0    2.010721
1    1.772005
2    1.754138
3    1.813009
4    2.165179
5    2.391025
6    2.645373
dtype: float64
```

In [26]:
```
iris.groupby(['Id'])['SepalLengthCm'].mean()
```

Out[26]:
```
Id
1      5.1
2      4.9
3      4.7
4      4.6
5      5.0
      ...
146    6.7
147    6.3
148    6.5
149    6.2
150    5.9
Name: SepalLengthCm, Length: 150, dtype: float64
```

In [34]:
```
iris_u=iris.rename(columns={'PetalLengthCm':'P_Width'},inplace=True)
```

In [32]:
```
(iris_u.groupby(['SepalWidthCm']).P_Width.mean())
```

Out[32]:
```
SepalWidthCm
2.0    3.500000
2.2    4.500000
2.3    3.250000
2.4    3.600000
2.5    4.512500
2.6    4.880000
2.7    4.622222
2.8    5.042857
2.9    4.350000
3.0    4.234615
3.1    3.600000
3.2    3.753846
3.3    4.200000
3.4    2.466667
3.5    1.416667
3.6    2.833333
3.7    1.500000
3.8    3.300000
3.9    1.500000
4.0    1.200000
4.1    1.500000
4.2    1.400000
4.4    1.500000
Name: P_Width, dtype: float64
```

In [35]:
```
iris.head(5)
```

Out[35]:

	Id	SepalLengthCm	SepalWidthCm	P_Width	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

In [ ]:
```

```

In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [3]:
```
stud=pd.read_csv("/home/student/Desktop/Employee_Salary_Dataset.csv")
```

In [5]:
```
stud.mean()
```

Out[5]:
```
ID                  1.800000e+01
Experience_Years    9.200000e+00
Age                 3.548571e+01
Salary              2.059147e+06
dtype: float64
```

In [7]:
```
stud.loc[:,'Age'].mean()
```

Out[7]:
```
35.48571428571429
```

In [8]:
```
stud.mean(axis=1)[0:4]
```

Out[8]:
```
0    62508.50
1    12506.00
2    42507.25
3     6257.00
dtype: float64
```

In [9]:
```
stud.median()
```

Out[9]:
```
ID                     18.0
Experience_Years        6.0
Age                    29.0
Salary              250000.0
dtype: float64
```

In [10]:
```
stud.loc[:,'Age'].median()
```

Out[10]:
```
29.0
```

In [12]:
```
stud.median(axis=1)[0:4]
```

Out[12]:
```
0    16.5
1    11.5
2    13.0
3    13.0
dtype: float64
```

In [13]:
```
stud.mode()
```

Out[13]:

	ID	Experience_Years	Age	Gender	Salary
0	1	2.0	54.0	Female	25000.0
1	2	NaN	NaN	NaN	250000.0
2	3	NaN	NaN	NaN	NaN
3	4	NaN	NaN	NaN	NaN
4	5	NaN	NaN	NaN	NaN
5	6	NaN	NaN	NaN	NaN
6	7	NaN	NaN	NaN	NaN
7	8	NaN	NaN	NaN	NaN
8	9	NaN	NaN	NaN	NaN
9	10	NaN	NaN	NaN	NaN
10	11	NaN	NaN	NaN	NaN
11	12	NaN	NaN	NaN	NaN
12	13	NaN	NaN	NaN	NaN
13	14	NaN	NaN	NaN	NaN
14	15	NaN	NaN	NaN	NaN
15	16	NaN	NaN	NaN	NaN
16	17	NaN	NaN	NaN	NaN
17	18	NaN	NaN	NaN	NaN
18	19	NaN	NaN	NaN	NaN
19	20	NaN	NaN	NaN	NaN
20	21	NaN	NaN	NaN	NaN
21	22	NaN	NaN	NaN	NaN
22	23	NaN	NaN	NaN	NaN
23	24	NaN	NaN	NaN	NaN
24	25	NaN	NaN	NaN	NaN
25	26	NaN	NaN	NaN	NaN
26	27	NaN	NaN	NaN	NaN
27	28	NaN	NaN	NaN	NaN
28	29	NaN	NaN	NaN	NaN
29	30	NaN	NaN	NaN	NaN
30	31	NaN	NaN	NaN	NaN
31	32	NaN	NaN	NaN	NaN
32	33	NaN	NaN	NaN	NaN
33	34	NaN	NaN	NaN	NaN
34	35	NaN	NaN	NaN	NaN

In [14]:
```
stud.mode(axis=1)[0:4]
```

In [16]:
```
stud.loc[:,'Age'].mode()
```

Out[16]:
```
0    54
Name: Age, dtype: int64
```

In [17]:
```
stud.min()
```

Out[17]:
```
ID                     1
Experience_Years       1
Age                   17
Gender            Female
Salary              3000
dtype: object
```

In [18]:
```
stud.max()
```

Out[18]:
```
ID                     35
Experience_Years       27
Age                    62
Gender               Male
Salary           10000000
dtype: object
```

In [20]:
```
stud.loc[:,'Experience_Years'].min(skipna=False)
```

Out[20]:
```
1
```

In [22]:
```
stud.loc[:,'Experience_Years'].max(skipna=False)
```

Out[22]:
```
27
```

In [23]:
```
stud.std()
```

Out[23]:
```
ID                 1.024695e+01
Experience_Years   7.552950e+00
Age                1.464355e+01
Salary             3.170124e+06
dtype: float64
```

In [25]:
```
stud.loc[:,'Age'].std()
```

Out[25]:
```
14.643551940884361
```

In [26]:
```
stud.std(axis=1)[0:4]
```

Out[26]:
```
0    124994.333900
1     24996.001694
2     84995.167190
3     12495.336570
dtype: float64
```

In [28]:
```
stud.groupby(['Experience_Years'])['Age'].mean()
```

Out[28]:
```
Experience_Years
1     19.250000
2     21.600000
3     22.500000
4     25.500000
5     28.500000
6     29.000000
10    35.000000
11    40.000000
14    39.000000
15    54.000000
16    49.000000
19    53.666667
20    55.000000
25    62.000000
27    62.000000
Name: Age, dtype: float64
```

In [32]:
```
stud.rename(columns={"Experience_Years": "Emp_years"})
```

Out[32]:

	ID	Emp_years	Age	Gender	Salary
0	1	5	28	Female	250000
1	2	1	21	Male	50000
2	3	3	23	Female	170000
3	4	2	22	Male	25000
4	5	1	17	Male	10000
5	6	25	62	Male	5001000
6	7	19	54	Female	800000
7	8	2	21	Female	9000
8	9	10	36	Female	61500
9	10	15	54	Female	650000
10	11	4	26	Female	250000
11	12	6	29	Male	1400000
12	13	14	39	Male	6000050
13	14	11	40	Male	220100
14	15	2	23	Male	7500
15	16	4	27	Female	87000
16	17	10	34	Female	930000
17	18	15	54	Female	7900000
18	19	2	21	Male	15000
19	20	10	36	Male	330000
20	21	15	54	Male	6570000
21	22	4	26	Male	25000
22	23	5	29	Male	6845000
23	24	1	21	Female	6000
24	25	4	23	Female	8900
25	26	3	22	Female	20000
26	27	1	18	Male	3000
27	28	27	62	Female	10000000
28	29	19	54	Female	5000000
29	30	2	21	Female	6100
30	31	10	34	Male	80000
31	32	15	54	Male	900000
32	33	20	55	Female	1540000
33	34	19	53	Female	9300000
34	35	16	49	Male	7600000

In [33]:

```
from sklearn import preprocessing
enc=preprocessing.OneHotEncoder()
enc_stud=pd.DataFrame(enc.fit_transform(stud[['Gender']]).toarray())
enc_stud
```

Out[33]:

	0	1
0	1.0	0.0
1	0.0	1.0
2	1.0	0.0
3	0.0	1.0
4	0.0	1.0
5	0.0	1.0
6	1.0	0.0
7	1.0	0.0
8	1.0	0.0
9	1.0	0.0
10	1.0	0.0
11	0.0	1.0
12	0.0	1.0
13	0.0	1.0
14	0.0	1.0
15	1.0	0.0
16	1.0	0.0
17	1.0	0.0
18	0.0	1.0
19	0.0	1.0
20	0.0	1.0
21	0.0	1.0
22	0.0	1.0
23	1.0	0.0
24	1.0	0.0
25	1.0	0.0
26	0.0	1.0
27	1.0	0.0
28	1.0	0.0
29	1.0	0.0
30	0.0	1.0
31	0.0	1.0
32	1.0	0.0
33	1.0	0.0
34	0.0	1.0

In [36]:
```
stud_encode=stud.join(enc_stud)
stud_encode
```

Out[36]:

	ID	Experience_Years	Age	Gender	Salary	0	1
0	1	5	28	Female	250000	1.0	0.0
1	2	1	21	Male	50000	0.0	1.0
2	3	3	23	Female	170000	1.0	0.0
3	4	2	22	Male	25000	0.0	1.0
4	5	1	17	Male	10000	0.0	1.0
5	6	25	62	Male	5001000	0.0	1.0
6	7	19	54	Female	800000	1.0	0.0
7	8	2	21	Female	9000	1.0	0.0
8	9	10	36	Female	61500	1.0	0.0
9	10	15	54	Female	650000	1.0	0.0
10	11	4	26	Female	250000	1.0	0.0
11	12	6	29	Male	1400000	0.0	1.0
12	13	14	39	Male	6000050	0.0	1.0
13	14	11	40	Male	220100	0.0	1.0
14	15	2	23	Male	7500	0.0	1.0
15	16	4	27	Female	87000	1.0	0.0
16	17	10	34	Female	930000	1.0	0.0
17	18	15	54	Female	7900000	1.0	0.0
18	19	2	21	Male	15000	0.0	1.0
19	20	10	36	Male	330000	0.0	1.0
20	21	15	54	Male	6570000	0.0	1.0
21	22	4	26	Male	25000	0.0	1.0
22	23	5	29	Male	6845000	0.0	1.0
23	24	1	21	Female	6000	1.0	0.0
24	25	4	23	Female	8900	1.0	0.0
25	26	3	22	Female	20000	1.0	0.0
26	27	1	18	Male	3000	0.0	1.0
27	28	27	62	Female	10000000	1.0	0.0
28	29	19	54	Female	5000000	1.0	0.0
29	30	2	21	Female	6100	1.0	0.0
30	31	10	34	Male	80000	0.0	1.0
31	32	15	54	Male	900000	0.0	1.0
32	33	20	55	Female	1540000	1.0	0.0
33	34	19	53	Female	9300000	1.0	0.0
34	35	16	49	Male	7600000	0.0	1.0

In [37]:
```
import pandas as pd
```

In [40]:
```
iris=pd.read_csv("/home/student/Desktop/IRIS.csv")
```

In [44]:
```
irisSet = (iris['species']== 'Iris-setosa')
print('Iris-setosa')
print(iris[irisSet].describe())
```

```
Iris-setosa
       sepal_length  sepal_width  petal_length  petal_width
count      50.00000    50.000000     50.000000     50.00000
mean        5.00600     3.418000      1.464000      0.24400
std         0.35249     0.381024      0.173511      0.10721
min         4.30000     2.300000      1.000000      0.10000
25%         4.80000     3.125000      1.400000      0.20000
50%         5.00000     3.400000      1.500000      0.20000
75%         5.20000     3.675000      1.575000      0.30000
max         5.80000     4.400000      1.900000      0.60000
```

In [45]:
```
irisSet = (iris['species']== 'Iris-versicolor')
print('Iris-versicolor')
print(iris[irisSet].describe())
```

```
Iris-versicolor
       sepal_length  sepal_width  petal_length  petal_width
count     50.000000    50.000000     50.000000    50.000000
mean       5.936000     2.770000      4.260000     1.326000
std        0.516171     0.313798      0.469911     0.197753
min        4.900000     2.000000      3.000000     1.000000
25%        5.600000     2.525000      4.000000     1.200000
50%        5.900000     2.800000      4.350000     1.300000
75%        6.300000     3.000000      4.600000     1.500000
max        7.000000     3.400000      5.100000     1.800000
```

In [46]:
```
irisSet = (iris['species']== 'Iris-virginica')
print('Iris-virginica')
print(iris[irisSet].describe())
```

```
Iris-virginica
       sepal_length  sepal_width  petal_length  petal_width
count      50.00000    50.000000     50.000000     50.00000
mean        6.58800     2.974000      5.552000      2.02600
std         0.63588     0.322497      0.551895      0.27465
min         4.90000     2.200000      4.500000      1.40000
25%         6.22500     2.800000      5.100000      1.80000
50%         6.50000     3.000000      5.550000      2.00000
75%         6.90000     3.175000      5.875000      2.30000
max         7.90000     3.800000      6.900000      2.50000
```

In [ ]:
```
```