

Assignment 4 22nd Sept

September 28, 2023

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

```
[2]: dataset=pd.read_csv("Employee-Attrition.csv")
```

```
[3]: dataset
```

```
[3]:      Age Attrition  BusinessTravel  DailyRate  Department \
0      41      Yes      Travel_Rarely      1102      Sales
1      49      No  Travel_Frequently      279  Research & Development
2      37      Yes      Travel_Rarely     1373  Research & Development
3      33      No  Travel_Frequently     1392  Research & Development
4      27      No      Travel_Rarely      591  Research & Development
...  ...  ...  ...  ...  ...
1465   36      No  Travel_Frequently      884  Research & Development
1466   39      No      Travel_Rarely      613  Research & Development
1467   27      No      Travel_Rarely      155  Research & Development
1468   49      No  Travel_Frequently     1023      Sales
1469   34      No      Travel_Rarely      628  Research & Development
```

```
      DistanceFromHome  Education  EducationField  EmployeeCount \
0                    1          2  Life Sciences              1
1                    8          1  Life Sciences              1
2                    2          2          Other              1
3                    3          4  Life Sciences              1
4                    2          1          Medical              1
...  ...  ...  ...  ...
1465          23          2          Medical              1
1466           6          1          Medical              1
1467           4          3  Life Sciences              1
1468           2          3          Medical              1
1469           8          3          Medical              1
```

```
      EmployeeNumber  ...  RelationshipSatisfaction  StandardHours \
0                    1  ...                        1              80
1                    2  ...                        4              80
```

2	4	...	2	80
3	5	...	3	80
4	7	...	4	80
...
1465	2061	...	3	80
1466	2062	...	1	80
1467	2064	...	2	80
1468	2065	...	4	80
1469	2068	...	1	80

	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	\
0	0	8	0	
1	1	10	3	
2	0	7	3	
3	0	8	3	
4	1	6	3	
...	
1465	1	17	3	
1466	1	9	5	
1467	1	6	0	
1468	0	17	3	
1469	0	6	3	

	WorkLifeBalance	YearsAtCompany	YearsInCurrentRole	\
0	1	6	4	
1	3	10	7	
2	3	0	0	
3	3	8	7	
4	3	2	2	
...	
1465	3	5	2	
1466	3	7	7	
1467	3	6	2	
1468	2	9	6	
1469	4	4	3	

	YearsSinceLastPromotion	YearsWithCurrManager
0	0	5
1	1	7
2	0	0
3	3	0
4	2	2
...
1465	0	3
1466	1	7
1467	0	3
1468	0	8

1469

1

2

[1470 rows x 35 columns]

```
[4]: dataset.head()
```

```
[4]:   Age Attrition   BusinessTravel   DailyRate   Department \
0    41      Yes   Travel_Rarely    1102      Sales
1    49      No  Travel_Frequently    279  Research & Development
2    37      Yes   Travel_Rarely    1373  Research & Development
3    33      No  Travel_Frequently    1392  Research & Development
4    27      No   Travel_Rarely    591   Research & Development

   DistanceFromHome   Education EducationField   EmployeeCount   EmployeeNumber \
0                1         2   Life Sciences             1             1
1                8         1   Life Sciences             1             2
2                2         2         Other             1             4
3                3         4   Life Sciences             1             5
4                2         1         Medical             1             7

   ... RelationshipSatisfaction   StandardHours   StockOptionLevel \
0   ...                1             80             0
1   ...                4             80             1
2   ...                2             80             0
3   ...                3             80             0
4   ...                4             80             1

   TotalWorkingYears   TrainingTimesLastYear   WorkLifeBalance   YearsAtCompany \
0                8                0                1             6
1               10                3                3            10
2                7                3                3             0
3                8                3                3             8
4                6                3                3             2

   YearsInCurrentRole   YearsSinceLastPromotion   YearsWithCurrManager
0                4                0                5
1                7                1                7
2                0                0                0
3                7                3                0
4                2                2                2
```

[5 rows x 35 columns]

```
[5]: dataset.tail()
```

```
[5]:   Age Attrition   BusinessTravel   DailyRate   Department \
1465   36      No  Travel_Frequently    884  Research & Development
```

1466	39	No	Travel_Rarely	613	Research & Development
1467	27	No	Travel_Rarely	155	Research & Development
1468	49	No	Travel_Frequently	1023	Sales
1469	34	No	Travel_Rarely	628	Research & Development

	DistanceFromHome	Education	EducationField	EmployeeCount	\
1465	23	2	Medical	1	
1466	6	1	Medical	1	
1467	4	3	Life Sciences	1	
1468	2	3	Medical	1	
1469	8	3	Medical	1	

	EmployeeNumber	...	RelationshipSatisfaction	StandardHours	\
1465	2061	...	3	80	
1466	2062	...	1	80	
1467	2064	...	2	80	
1468	2065	...	4	80	
1469	2068	...	1	80	

	StockOptionLevel	TotalWorkingYears	TrainingTimesLastYear	\
1465	1	17	3	
1466	1	9	5	
1467	1	6	0	
1468	0	17	3	
1469	0	6	3	

	WorkLifeBalance	YearsAtCompany	YearsInCurrentRole	\
1465	3	5	2	
1466	3	7	7	
1467	3	6	2	
1468	2	9	6	
1469	4	4	3	

	YearsSinceLastPromotion	YearsWithCurrManager
1465	0	3
1466	1	7
1467	0	3
1468	0	8
1469	1	2

[5 rows x 35 columns]

```
[7]: dataset.shape
```

```
[7]: (1470, 35)
```

```
[8]: dataset.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1470 entries, 0 to 1469
Data columns (total 35 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Age                                   1470 non-null   int64
1   Attrition                           1470 non-null   object
2   BusinessTravel                       1470 non-null   object
3   DailyRate                            1470 non-null   int64
4   Department                           1470 non-null   object
5   DistanceFromHome                     1470 non-null   int64
6   Education                             1470 non-null   int64
7   EducationField                       1470 non-null   object
8   EmployeeCount                        1470 non-null   int64
9   EmployeeNumber                       1470 non-null   int64
10  EnvironmentSatisfaction               1470 non-null   int64
11  Gender                               1470 non-null   object
12  HourlyRate                           1470 non-null   int64
13  JobInvolvement                       1470 non-null   int64
14  JobLevel                             1470 non-null   int64
15  JobRole                              1470 non-null   object
16  JobSatisfaction                       1470 non-null   int64
17  MaritalStatus                        1470 non-null   object
18  MonthlyIncome                        1470 non-null   int64
19  MonthlyRate                          1470 non-null   int64
20  NumCompaniesWorked                   1470 non-null   int64
21  Over18                              1470 non-null   object
22  OverTime                             1470 non-null   object
23  PercentSalaryHike                    1470 non-null   int64
24  PerformanceRating                    1470 non-null   int64
25  RelationshipSatisfaction              1470 non-null   int64
26  StandardHours                        1470 non-null   int64
27  StockOptionLevel                     1470 non-null   int64
28  TotalWorkingYears                    1470 non-null   int64
29  TrainingTimesLastYear                1470 non-null   int64
30  WorkLifeBalance                      1470 non-null   int64
31  YearsAtCompany                       1470 non-null   int64
32  YearsInCurrentRole                   1470 non-null   int64
33  YearsSinceLastPromotion              1470 non-null   int64
34  YearsWithCurrManager                 1470 non-null   int64
dtypes: int64(26), object(9)
memory usage: 402.1+ KB

```

```

[11]: numeric_columns = dataset.select_dtypes(include=['number'])
      corr = numeric_columns.corr()
      corr

```

```
[11]:
```

	Age	DailyRate	DistanceFromHome	Education	\
Age	1.000000	0.010661	-0.001686	0.208034	
DailyRate	0.010661	1.000000	-0.004985	-0.016806	
DistanceFromHome	-0.001686	-0.004985	1.000000	0.021042	
Education	0.208034	-0.016806	0.021042	1.000000	
EmployeeCount	NaN	NaN	NaN	NaN	
EmployeeNumber	-0.010145	-0.050990	0.032916	0.042070	
EnvironmentSatisfaction	0.010146	0.018355	-0.016075	-0.027128	
HourlyRate	0.024287	0.023381	0.031131	0.016775	
JobInvolvement	0.029820	0.046135	0.008783	0.042438	
JobLevel	0.509604	0.002966	0.005303	0.101589	
JobSatisfaction	-0.004892	0.030571	-0.003669	-0.011296	
MonthlyIncome	0.497855	0.007707	-0.017014	0.094961	
MonthlyRate	0.028051	-0.032182	0.027473	-0.026084	
NumCompaniesWorked	0.299635	0.038153	-0.029251	0.126317	
PercentSalaryHike	0.003634	0.022704	0.040235	-0.011111	
PerformanceRating	0.001904	0.000473	0.027110	-0.024539	
RelationshipSatisfaction	0.053535	0.007846	0.006557	-0.009118	
StandardHours	NaN	NaN	NaN	NaN	
StockOptionLevel	0.037510	0.042143	0.044872	0.018422	
TotalWorkingYears	0.680381	0.014515	0.004628	0.148280	
TrainingTimesLastYear	-0.019621	0.002453	-0.036942	-0.025100	
WorkLifeBalance	-0.021490	-0.037848	-0.026556	0.009819	
YearsAtCompany	0.311309	-0.034055	0.009508	0.069114	
YearsInCurrentRole	0.212901	0.009932	0.018845	0.060236	
YearsSinceLastPromotion	0.216513	-0.033229	0.010029	0.054254	
YearsWithCurrManager	0.202089	-0.026363	0.014406	0.069065	

	EmployeeCount	EmployeeNumber	\
Age	NaN	-0.010145	
DailyRate	NaN	-0.050990	
DistanceFromHome	NaN	0.032916	
Education	NaN	0.042070	
EmployeeCount	NaN	NaN	
EmployeeNumber	NaN	1.000000	
EnvironmentSatisfaction	NaN	0.017621	
HourlyRate	NaN	0.035179	
JobInvolvement	NaN	-0.006888	
JobLevel	NaN	-0.018519	
JobSatisfaction	NaN	-0.046247	
MonthlyIncome	NaN	-0.014829	
MonthlyRate	NaN	0.012648	
NumCompaniesWorked	NaN	-0.001251	
PercentSalaryHike	NaN	-0.012944	
PerformanceRating	NaN	-0.020359	
RelationshipSatisfaction	NaN	-0.069861	
StandardHours	NaN	NaN	

StockOptionLevel	NaN	0.062227
TotalWorkingYears	NaN	-0.014365
TrainingTimesLastYear	NaN	0.023603
WorkLifeBalance	NaN	0.010309
YearsAtCompany	NaN	-0.011240
YearsInCurrentRole	NaN	-0.008416
YearsSinceLastPromotion	NaN	-0.009019
YearsWithCurrManager	NaN	-0.009197

	EnvironmentSatisfaction	HourlyRate	JobInvolvement	\
Age	0.010146	0.024287	0.029820	
DailyRate	0.018355	0.023381	0.046135	
DistanceFromHome	-0.016075	0.031131	0.008783	
Education	-0.027128	0.016775	0.042438	
EmployeeCount	NaN	NaN	NaN	
EmployeeNumber	0.017621	0.035179	-0.006888	
EnvironmentSatisfaction	1.000000	-0.049857	-0.008278	
HourlyRate	-0.049857	1.000000	0.042861	
JobInvolvement	-0.008278	0.042861	1.000000	
JobLevel	0.001212	-0.027853	-0.012630	
JobSatisfaction	-0.006784	-0.071335	-0.021476	
MonthlyIncome	-0.006259	-0.015794	-0.015271	
MonthlyRate	0.037600	-0.015297	-0.016322	
NumCompaniesWorked	0.012594	0.022157	0.015012	
PercentSalaryHike	-0.031701	-0.009062	-0.017205	
PerformanceRating	-0.029548	-0.002172	-0.029071	
RelationshipSatisfaction	0.007665	0.001330	0.034297	
StandardHours	NaN	NaN	NaN	
StockOptionLevel	0.003432	0.050263	0.021523	
TotalWorkingYears	-0.002693	-0.002334	-0.005533	
TrainingTimesLastYear	-0.019359	-0.008548	-0.015338	
WorkLifeBalance	0.027627	-0.004607	-0.014617	
YearsAtCompany	0.001458	-0.019582	-0.021355	
YearsInCurrentRole	0.018007	-0.024106	0.008717	
YearsSinceLastPromotion	0.016194	-0.026716	-0.024184	
YearsWithCurrManager	-0.004999	-0.020123	0.025976	

	JobLevel	...	RelationshipSatisfaction	\
Age	0.509604	...	0.053535	
DailyRate	0.002966	...	0.007846	
DistanceFromHome	0.005303	...	0.006557	
Education	0.101589	...	-0.009118	
EmployeeCount	NaN	...	NaN	
EmployeeNumber	-0.018519	...	-0.069861	
EnvironmentSatisfaction	0.001212	...	0.007665	
HourlyRate	-0.027853	...	0.001330	
JobInvolvement	-0.012630	...	0.034297	

JobLevel	1.000000	...	0.021642
JobSatisfaction	-0.001944	...	-0.012454
MonthlyIncome	0.950300	...	0.025873
MonthlyRate	0.039563	...	-0.004085
NumCompaniesWorked	0.142501	...	0.052733
PercentSalaryHike	-0.034730	...	-0.040490
PerformanceRating	-0.021222	...	-0.031351
RelationshipSatisfaction	0.021642	...	1.000000
StandardHours	NaN	...	NaN
StockOptionLevel	0.013984	...	-0.045952
TotalWorkingYears	0.782208	...	0.024054
TrainingTimesLastYear	-0.018191	...	0.002497
WorkLifeBalance	0.037818	...	0.019604
YearsAtCompany	0.534739	...	0.019367
YearsInCurrentRole	0.389447	...	-0.015123
YearsSinceLastPromotion	0.353885	...	0.033493
YearsWithCurrManager	0.375281	...	-0.000867

	StandardHours	StockOptionLevel	TotalWorkingYears	\
Age	NaN	0.037510	0.680381	
DailyRate	NaN	0.042143	0.014515	
DistanceFromHome	NaN	0.044872	0.004628	
Education	NaN	0.018422	0.148280	
EmployeeCount	NaN	NaN	NaN	
EmployeeNumber	NaN	0.062227	-0.014365	
EnvironmentSatisfaction	NaN	0.003432	-0.002693	
HourlyRate	NaN	0.050263	-0.002334	
JobInvolvement	NaN	0.021523	-0.005533	
JobLevel	NaN	0.013984	0.782208	
JobSatisfaction	NaN	0.010690	-0.020185	
MonthlyIncome	NaN	0.005408	0.772893	
MonthlyRate	NaN	-0.034323	0.026442	
NumCompaniesWorked	NaN	0.030075	0.237639	
PercentSalaryHike	NaN	0.007528	-0.020608	
PerformanceRating	NaN	0.003506	0.006744	
RelationshipSatisfaction	NaN	-0.045952	0.024054	
StandardHours	NaN	NaN	NaN	
StockOptionLevel	NaN	1.000000	0.010136	
TotalWorkingYears	NaN	0.010136	1.000000	
TrainingTimesLastYear	NaN	0.011274	-0.035662	
WorkLifeBalance	NaN	0.004129	0.001008	
YearsAtCompany	NaN	0.015058	0.628133	
YearsInCurrentRole	NaN	0.050818	0.460365	
YearsSinceLastPromotion	NaN	0.014352	0.404858	
YearsWithCurrManager	NaN	0.024698	0.459188	

TrainingTimesLastYear WorkLifeBalance \

Age	-0.019621	-0.021490
DailyRate	0.002453	-0.037848
DistanceFromHome	-0.036942	-0.026556
Education	-0.025100	0.009819
EmployeeCount	NaN	NaN
EmployeeNumber	0.023603	0.010309
EnvironmentSatisfaction	-0.019359	0.027627
HourlyRate	-0.008548	-0.004607
JobInvolvement	-0.015338	-0.014617
JobLevel	-0.018191	0.037818
JobSatisfaction	-0.005779	-0.019459
MonthlyIncome	-0.021736	0.030683
MonthlyRate	0.001467	0.007963
NumCompaniesWorked	-0.066054	-0.008366
PercentSalaryHike	-0.005221	-0.003280
PerformanceRating	-0.015579	0.002572
RelationshipSatisfaction	0.002497	0.019604
StandardHours	NaN	NaN
StockOptionLevel	0.011274	0.004129
TotalWorkingYears	-0.035662	0.001008
TrainingTimesLastYear	1.000000	0.028072
WorkLifeBalance	0.028072	1.000000
YearsAtCompany	0.003569	0.012089
YearsInCurrentRole	-0.005738	0.049856
YearsSinceLastPromotion	-0.002067	0.008941
YearsWithCurrManager	-0.004096	0.002759

	YearsAtCompany	YearsInCurrentRole \
Age	0.311309	0.212901
DailyRate	-0.034055	0.009932
DistanceFromHome	0.009508	0.018845
Education	0.069114	0.060236
EmployeeCount	NaN	NaN
EmployeeNumber	-0.011240	-0.008416
EnvironmentSatisfaction	0.001458	0.018007
HourlyRate	-0.019582	-0.024106
JobInvolvement	-0.021355	0.008717
JobLevel	0.534739	0.389447
JobSatisfaction	-0.003803	-0.002305
MonthlyIncome	0.514285	0.363818
MonthlyRate	-0.023655	-0.012815
NumCompaniesWorked	-0.118421	-0.090754
PercentSalaryHike	-0.035991	-0.001520
PerformanceRating	0.003435	0.034986
RelationshipSatisfaction	0.019367	-0.015123
StandardHours	NaN	NaN
StockOptionLevel	0.015058	0.050818

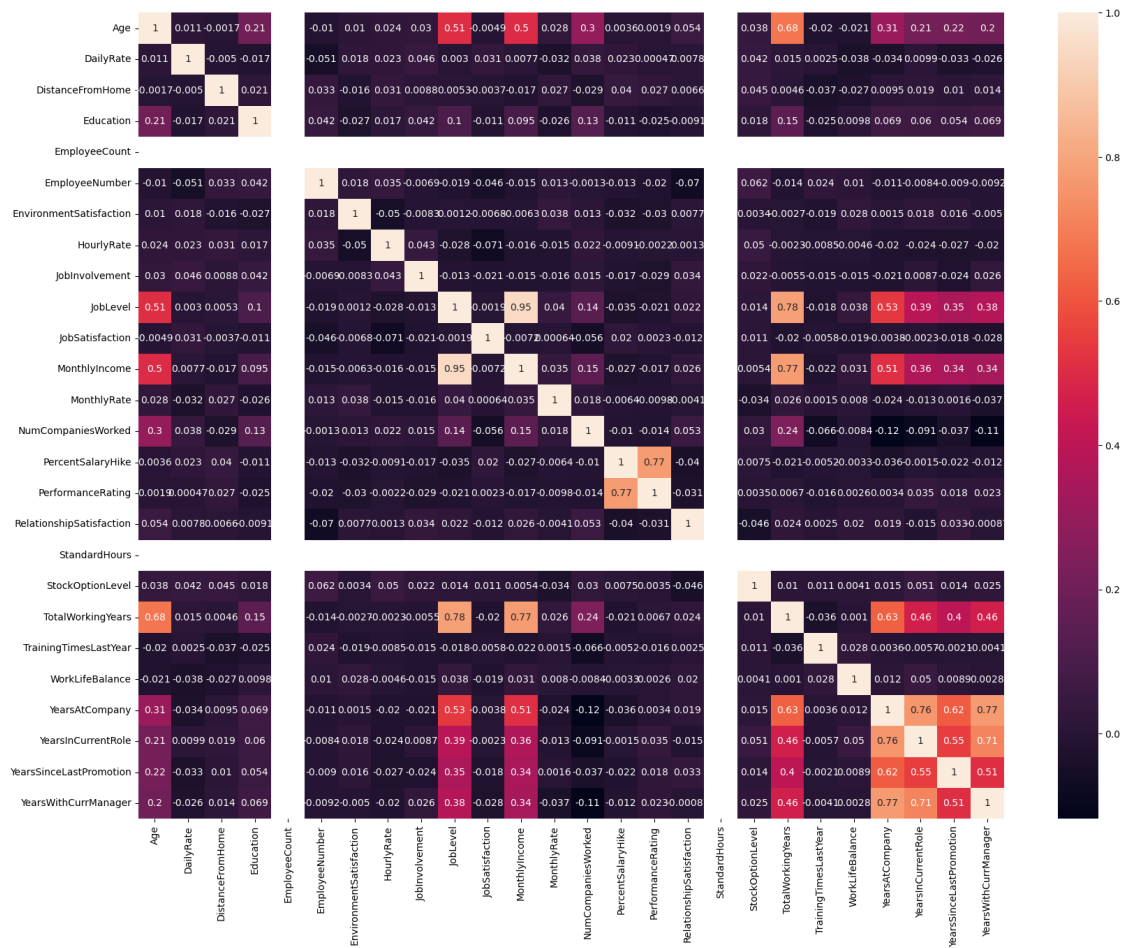
TotalWorkingYears	0.628133	0.460365
TrainingTimesLastYear	0.003569	-0.005738
WorkLifeBalance	0.012089	0.049856
YearsAtCompany	1.000000	0.758754
YearsInCurrentRole	0.758754	1.000000
YearsSinceLastPromotion	0.618409	0.548056
YearsWithCurrManager	0.769212	0.714365

	YearsSinceLastPromotion	YearsWithCurrManager
Age	0.216513	0.202089
DailyRate	-0.033229	-0.026363
DistanceFromHome	0.010029	0.014406
Education	0.054254	0.069065
EmployeeCount	NaN	NaN
EmployeeNumber	-0.009019	-0.009197
EnvironmentSatisfaction	0.016194	-0.004999
HourlyRate	-0.026716	-0.020123
JobInvolvement	-0.024184	0.025976
JobLevel	0.353885	0.375281
JobSatisfaction	-0.018214	-0.027656
MonthlyIncome	0.344978	0.344079
MonthlyRate	0.001567	-0.036746
NumCompaniesWorked	-0.036814	-0.110319
PercentSalaryHike	-0.022154	-0.011985
PerformanceRating	0.017896	0.022827
RelationshipSatisfaction	0.033493	-0.000867
StandardHours	NaN	NaN
StockOptionLevel	0.014352	0.024698
TotalWorkingYears	0.404858	0.459188
TrainingTimesLastYear	-0.002067	-0.004096
WorkLifeBalance	0.008941	0.002759
YearsAtCompany	0.618409	0.769212
YearsInCurrentRole	0.548056	0.714365
YearsSinceLastPromotion	1.000000	0.510224
YearsWithCurrManager	0.510224	1.000000

[26 rows x 26 columns]

```
[17]: plt.subplots(figsize=(20,15))
      sns.heatmap(corr,annot=True)
```

```
[17]: <Axes: >
```



```
[18]: dataset.isnull().any()
```

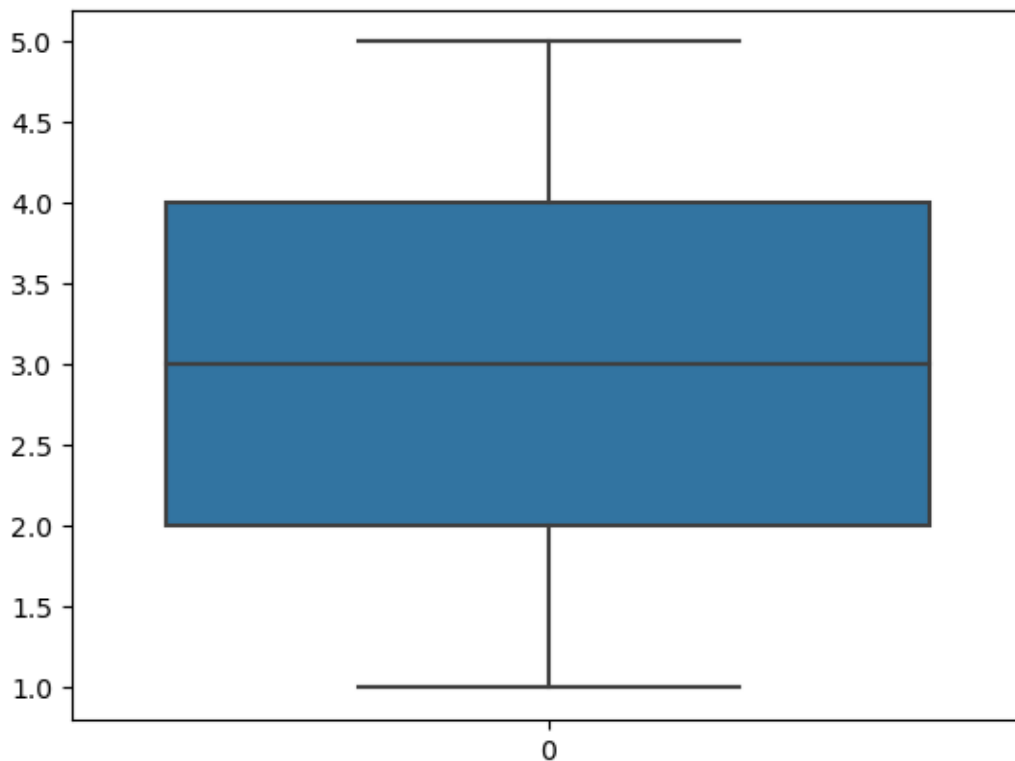
```
[18]: Age                False
      Attrition           False
      BusinessTravel      False
      DailyRate           False
      Department          False
      DistanceFromHome    False
      Education           False
      EducationField       False
      EmployeeCount       False
      EmployeeNumber       False
      EnvironmentSatisfaction False
      Gender              False
      HourlyRate          False
      JobInvolvement      False
      JobLevel            False
```

JobRole	False
JobSatisfaction	False
MaritalStatus	False
MonthlyIncome	False
MonthlyRate	False
NumCompaniesWorked	False
Over18	False
OverTime	False
PercentSalaryHike	False
PerformanceRating	False
RelationshipSatisfaction	False
StandardHours	False
StockOptionLevel	False
TotalWorkingYears	False
TrainingTimesLastYear	False
WorkLifeBalance	False
YearsAtCompany	False
YearsInCurrentRole	False
YearsSinceLastPromotion	False
YearsWithCurrManager	False

dtype: bool

```
[21]: sns.boxplot(dataset.Education)
```

```
[21]: <Axes: >
```



```
[22]: x=dataset.iloc[:,3:13]
      y=dataset.iloc[:,13:14]
```

```
[23]: x.head()
```

```
[23]:   DailyRate      Department  DistanceFromHome  Education \
0      1102             Sales                1         2
1       279  Research & Development                8         1
2      1373  Research & Development                2         2
3      1392  Research & Development                3         4
4       591  Research & Development                2         1

      EducationField  EmployeeCount  EmployeeNumber  EnvironmentSatisfaction \
0  Life Sciences             1             1             2
1  Life Sciences             1             2             3
2      Other             1             4             4
3  Life Sciences             1             5             4
4      Medical             1             7             1

      Gender  HourlyRate
0  Female         94
1   Male         61
2   Male         92
3  Female         56
4   Male         40
```

```
[24]: y.head()
```

```
[24]:   JobInvolvement
0             3
1             2
2             2
3             3
4             3
```

```
[25]: from sklearn.preprocessing import LabelEncoder
```

```
[27]: le=LabelEncoder()
```

```
[28]: x["DailyRate"]=le.fit_transform(x["DailyRate"])
```

```
[29]: x["DailyRate"]
```

```
[29]: 0      624
      1      113
```

```

2      805
3      820
4      312
...
1465   494
1466   327
1467    39
1468   579
1469   336
Name: DailyRate, Length: 1470, dtype: int64

```

```
[30]: Department=pd.get_dummies(x["Department"],drop_first=True)
```

```
[31]: Department
```

```
[31]:      Research & Development  Sales
0                False    True
1                 True   False
2                 True   False
3                 True   False
4                 True   False
...
1465                True   False
1466                True   False
1467                True   False
1468                False    True
1469                True   False

```

```
[1470 rows x 2 columns]
```

```
[32]: from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3,random_state=0)
```

```
[34]: x_train.shape,x_test.shape,y_train.shape,y_test.shape
```

```
[34]: ((1029, 10), (441, 10), (1029, 1), (441, 1))
```

```
[38]: x_train
```

```
[38]:      DailyRate      Department  DistanceFromHome  Education  \
338      295           Sales           5           3
363      159           Sales           5           3
759      437  Human Resources          24           4
793      502  Research & Development          15           2
581      516  Research & Development           1           3
...
763      778           Sales          10           4
835      266  Human Resources           8           4

```

1216	671	Sales	2	3
559	109	Research & Development	2	5
684	358	Sales	10	4

	EducationField	EmployeeCount	EmployeeNumber	\
338	Marketing	1	456	
363	Marketing	1	485	
759	Medical	1	1049	
793	Life Sciences	1	1102	
581	Life Sciences	1	806	
...	
763	Life Sciences	1	1055	
835	Technical Degree	1	1164	
1216	Medical	1	1706	
559	Medical	1	773	
684	Marketing	1	954	

	EnvironmentSatisfaction	Gender	HourlyRate
338	4	Female	30
363	4	Female	34
759	2	Male	36
793	1	Male	50
581	4	Male	38
...
763	3	Female	87
835	3	Male	100
1216	4	Male	73
559	4	Male	92
684	1	Male	67

[1029 rows x 10 columns]

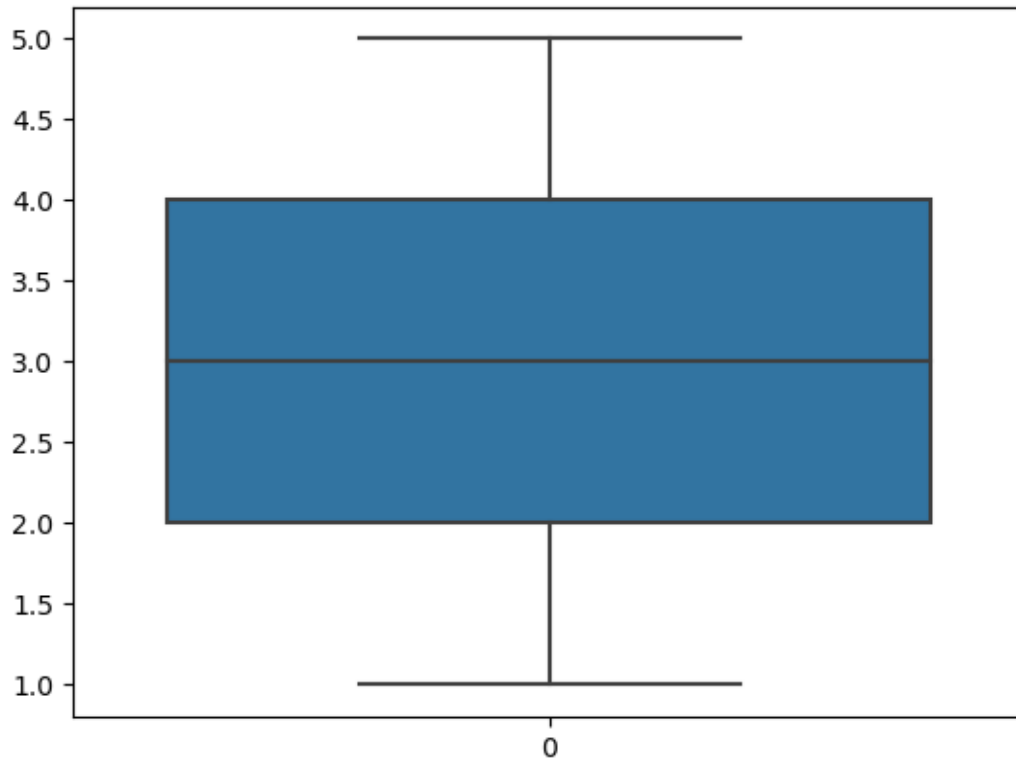
```
[39]: y_train
```

```
[39]: JobInvolvement
338    2
363    3
759    3
793    3
581    1
...
763    3
835    3
1216   3
559    3
684    2
```

[1029 rows x 1 columns]

```
[43]: sns.boxplot(dataset.Education)
```

[43]: <Axes: >



```
[44]: x.head()
```

```
[44]:   DailyRate      Department  DistanceFromHome  Education  \
0      624          Sales                1          2
1      113  Research & Development                8          1
2      805  Research & Development                2          2
3      820  Research & Development                3          4
4      312  Research & Development                2          1

   EducationField  EmployeeCount  EmployeeNumber  EnvironmentSatisfaction  \
0  Life Sciences                1                1                      2
1  Life Sciences                1                2                      3
2      Other                1                4                      4
3  Life Sciences                1                5                      4
4    Medical                1                7                      1
```


	Gender	HourlyRate
0	Female	94
1	Male	61
2	Male	92
3	Female	56
4	Male	40

```
[45]: x=dataset.iloc[:,1:4]
      x.head()
```

```
[45]: Attrition      BusinessTravel  DailyRate
0      Yes      Travel_Rarely      1102
1      No  Travel_Frequently      279
2      Yes      Travel_Rarely      1373
3      No  Travel_Frequently      1392
4      No      Travel_Rarely      591
```

```
[46]: y=dataset.Education
      y.head()
```

```
[46]: 0      2
      1      1
      2      2
      3      4
      4      1
      Name: Education, dtype: int64
```

```
[47]: from sklearn.linear_model import LogisticRegression
      from sklearn.tree import DecisionTreeClassifier
      from sklearn.ensemble import RandomForestClassifier
```

```
[49]: x_train
```

```
[49]:      DailyRate      Department  DistanceFromHome  Education  \
338      295      Sales      5      3
363      159      Sales      5      3
759      437  Human Resources      24      4
793      502  Research & Development      15      2
581      516  Research & Development      1      3
...      ...      ...      ...      ...
763      778      Sales      10      4
835      266  Human Resources      8      4
1216      671      Sales      2      3
559      109  Research & Development      2      5
684      358      Sales      10      4

      EducationField  EmployeeCount  EmployeeNumber  \
```

338	Marketing	1	456
363	Marketing	1	485
759	Medical	1	1049
793	Life Sciences	1	1102
581	Life Sciences	1	806
...
763	Life Sciences	1	1055
835	Technical Degree	1	1164
1216	Medical	1	1706
559	Medical	1	773
684	Marketing	1	954

	EnvironmentSatisfaction	Gender	HourlyRate
338	4	Female	30
363	4	Female	34
759	2	Male	36
793	1	Male	50
581	4	Male	38
...
763	3	Female	87
835	3	Male	100
1216	4	Male	73
559	4	Male	92
684	1	Male	67

[1029 rows x 10 columns]

```
[50]: y_train
```

```
[50]: JobInvolvement
338      2
363      3
759      3
793      3
581      1
...
763      3
835      3
1216     3
559      3
684      2
```

[1029 rows x 1 columns]

```
[59]: x_train_encoded = pd.get_dummies(x_train,
    ↪ columns=['Department', 'EducationField', 'Gender'])
```

```
x_test_encoded = pd.get_dummies(x_test,
↳ columns=['Department', 'EducationField', 'Gender'])
```

```
[71]: from sklearn.linear_model import LinearRegression
```

```
[72]: lr=LinearRegression()
```

```
[85]: lr.fit(x_train_encoded, y_train)
y_pred = lr.predict(x_test_encoded)
```

```
[86]: y_pred
```

```
[86]: array([[2.65690437],
[2.70547807],
[2.7138974 ],
[2.74666002],
[2.73285664],
[2.72540929],
[2.6648645 ],
[2.68485674],
[2.64987438],
[2.75082749],
[2.8994411 ],
[2.87226009],
[2.72388743],
[2.64496468],
[2.9251039 ],
[2.82780923],
[2.57881295],
[2.52157381],
[2.60853448],
[2.8285084 ],
[2.66983154],
[2.79030593],
[2.85068347],
[2.64136564],
[2.7120441 ],
[2.48227984],
[2.82410547],
[2.71781759],
[2.58795571],
[2.585702 ],
[2.68906158],
[2.75731703],
[2.78851163],
[2.75157253],
[2.85017783],
```

[2.93029264],
[2.57803406],
[2.77664803],
[2.68996047],
[2.83752071],
[2.7281984],
[2.73666833],
[2.7964555],
[2.61883512],
[2.79694042],
[2.54195275],
[2.63231051],
[2.69898013],
[2.8169835],
[2.79113546],
[2.87421893],
[2.71526732],
[2.7039422],
[2.81726545],
[2.63157328],
[2.56180143],
[2.74702624],
[2.46225191],
[2.80882088],
[2.69860852],
[2.7745101],
[2.72655269],
[2.65531091],
[2.82708695],
[2.8147311],
[2.55777179],
[2.73559218],
[2.80491248],
[2.89380854],
[2.63547128],
[2.90164508],
[2.89382689],
[2.77740666],
[2.8102358],
[2.77064052],
[2.7750251],
[2.59804447],
[2.70775401],
[2.67871374],
[2.63492452],
[2.47320394],
[2.59059037],

[2.74737969],
[2.64888067],
[2.6662233],
[2.51636909],
[2.84734503],
[2.75565356],
[2.76642736],
[2.82476777],
[2.73626222],
[2.98631005],
[2.75643181],
[2.71336886],
[2.68673295],
[2.66777255],
[2.7232133],
[2.64813125],
[2.79978051],
[2.78284088],
[2.76916402],
[2.82664145],
[2.92419789],
[2.70949359],
[2.73462779],
[2.91363679],
[2.6196038],
[2.82807025],
[2.62630241],
[2.95347995],
[2.86084464],
[2.71054853],
[2.89325636],
[2.75431012],
[2.89716446],
[2.61541111],
[2.74208213],
[2.68265815],
[2.74982596],
[2.73806274],
[2.95878263],
[2.76662783],
[2.53275042],
[2.68698483],
[2.563474],
[2.61887253],
[2.69999617],
[2.70503374],
[2.67710475],

[2.45997005],
[2.74916385],
[2.84032501],
[2.89711563],
[2.7707884],
[2.7809892],
[2.78283475],
[2.59709711],
[2.66665411],
[2.70686347],
[2.69573176],
[2.7800801],
[2.74984485],
[2.57117252],
[2.661436],
[2.63240123],
[2.75908962],
[2.62313454],
[2.56977155],
[2.52171064],
[2.58578729],
[2.92858841],
[2.8843186],
[2.87075247],
[2.77308172],
[2.79243844],
[2.63114783],
[2.91720577],
[2.67747768],
[2.8079399],
[2.59865614],
[2.5852289],
[2.79487919],
[2.79848132],
[2.82748554],
[2.65828764],
[2.77044192],
[2.5689588],
[2.69352854],
[2.72991192],
[2.83679309],
[2.71360532],
[2.53123572],
[2.72340663],
[2.83013749],
[2.63290522],
[2.77951433],

[2.7305658],
[2.72618211],
[2.70063186],
[2.62542312],
[2.68547676],
[2.93569967],
[2.72790781],
[2.74357384],
[2.86440706],
[2.75473956],
[2.82414652],
[2.63169968],
[2.67967962],
[2.68367012],
[2.83404615],
[2.77570809],
[2.78579774],
[2.69487253],
[2.79586309],
[2.47886379],
[2.65075036],
[2.66652061],
[2.75186963],
[2.75080093],
[2.89505587],
[2.63570242],
[2.58195114],
[2.65791651],
[2.79473434],
[2.63540709],
[2.6860769],
[2.74406788],
[2.60977667],
[2.67734162],
[2.88756621],
[2.84233674],
[2.68961684],
[2.67353675],
[2.90414502],
[2.87824714],
[2.77211469],
[2.81249067],
[2.86258492],
[2.79441125],
[2.78289707],
[2.45380242],
[2.76377768],

[2.60522786],
[2.66218774],
[2.76685078],
[2.66813126],
[2.70003401],
[2.59041023],
[2.79458184],
[2.73706049],
[2.66395758],
[2.7208538],
[2.78720271],
[2.82142702],
[2.81036053],
[2.69842874],
[2.92498947],
[2.86802729],
[2.74374854],
[2.81003789],
[2.66003744],
[2.7279001],
[2.80333435],
[2.62828371],
[2.77553142],
[2.85696258],
[2.67819063],
[2.74341464],
[2.78123458],
[2.72502353],
[2.73578076],
[2.71217628],
[2.78247617],
[2.69554507],
[2.81549872],
[2.84092976],
[2.81298483],
[2.61748068],
[2.72288601],
[2.73993353],
[2.59710791],
[2.9006467],
[2.84293668],
[2.83474789],
[2.73503206],
[2.6804666],
[2.71829661],
[2.91234452],
[2.62940982],

[2.65182935],
[2.74524996],
[2.89186438],
[2.73514173],
[2.47725779],
[2.60942372],
[2.78640014],
[2.72847718],
[2.88625652],
[2.75004307],
[2.72386],
[2.81026816],
[2.77804458],
[2.73328649],
[2.82424901],
[2.66934697],
[2.84836027],
[2.9255639],
[2.4865508],
[2.71057859],
[2.69781914],
[2.68341454],
[2.90086908],
[2.696865],
[2.65580373],
[2.79419069],
[2.7097804],
[2.78256214],
[2.62734036],
[2.65149191],
[2.67704167],
[2.81557005],
[2.64749773],
[2.65974078],
[2.58683552],
[2.79183845],
[2.78788576],
[2.68719858],
[2.74301825],
[2.61292049],
[2.79273502],
[2.80222166],
[2.72519036],
[2.70219607],
[2.7865236],
[2.75608774],
[2.75094567],

[2.75816641],
[2.64951228],
[2.74773069],
[2.53990284],
[2.74410757],
[2.53136888],
[2.60684799],
[2.81481212],
[2.59903989],
[2.78152994],
[2.76016213],
[2.67475246],
[2.79232585],
[2.67643696],
[2.63422552],
[2.68783578],
[2.67031465],
[2.71075019],
[2.81302875],
[2.69118786],
[2.63981686],
[2.73368328],
[2.71188022],
[2.6891315],
[2.57897746],
[2.65559525],
[2.65317542],
[2.8812143],
[2.78241854],
[2.57018791],
[2.59878599],
[2.68069764],
[2.82171925],
[2.69508736],
[2.61895336],
[2.66104848],
[2.81193719],
[2.79636091],
[2.86052723],
[2.834928],
[2.57743045],
[2.52346257],
[2.68243456],
[2.81573375],
[2.71241731],
[2.7443939],
[2.77867028],

[2.84786851],
[2.74780526],
[2.72194401],
[2.73146974],
[2.73379408],
[2.71235713],
[2.76082159],
[2.8489066],
[2.59463108],
[2.62587308],
[2.79287367],
[2.70342596],
[2.64648233],
[2.82195512],
[2.70132046],
[2.6163191],
[2.70377449],
[2.84781997],
[2.66733716],
[2.69825902],
[2.92704719],
[2.52697656],
[2.76495476],
[2.70585974],
[2.68612714],
[2.87275882],
[2.71159576],
[2.73583103],
[2.68398645],
[2.76018038],
[2.61760178],
[2.5808226],
[2.73841174],
[2.8023075],
[2.57458944],
[2.71724053],
[2.57323506],
[2.77015998],
[2.79073644],
[2.95172136],
[2.73575908],
[2.78738463],
[2.62308223],
[2.81158371],
[2.65616292],
[2.70116301],
[2.63425112],

```
[2.8403481 ],
[2.72073343],
[2.49830627],
[2.77089509],
[2.66955744],
[2.75352651],
[2.80152442],
[2.84407099],
[2.70986489],
[2.90262079],
[2.72640021],
[2.89228691],
[2.73060537],
[2.78711215],
[2.73567476],
[2.8060471 ],
[2.71444588],
[2.61221292],
[2.77062827],
[2.7257265 ],
[2.39630991],
[2.65963299],
[2.76736437],
[2.49731139],
[2.65936424],
[2.77200409],
[2.76451633],
[2.76189316],
[2.69650685],
[2.71125328]])
```

```
[87]: y_test
```

```
[87]:      JobInvolvement
442                3
1091               2
981                3
785                3
1332               2
...
817                3
399                2
458                1
406                2
590                3
```

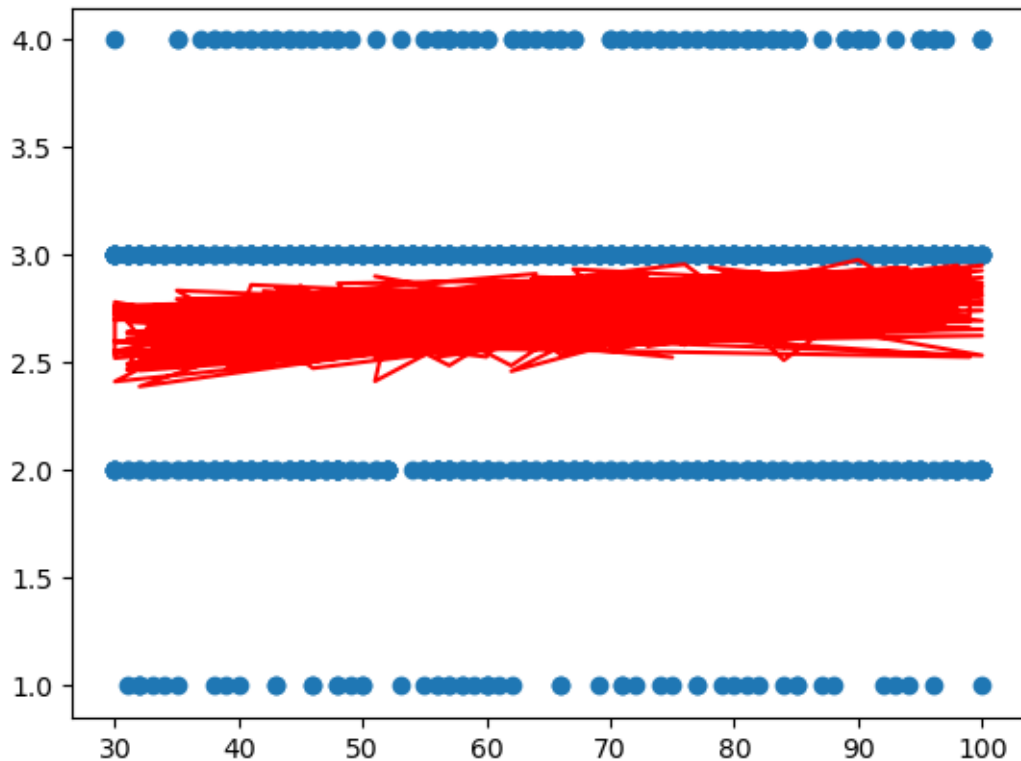
```
[441 rows x 1 columns]
```

```
[88]: from sklearn.metrics import r2_score  
r2_score(y_test,y_pred)
```

```
[88]: -0.04175099289695883
```

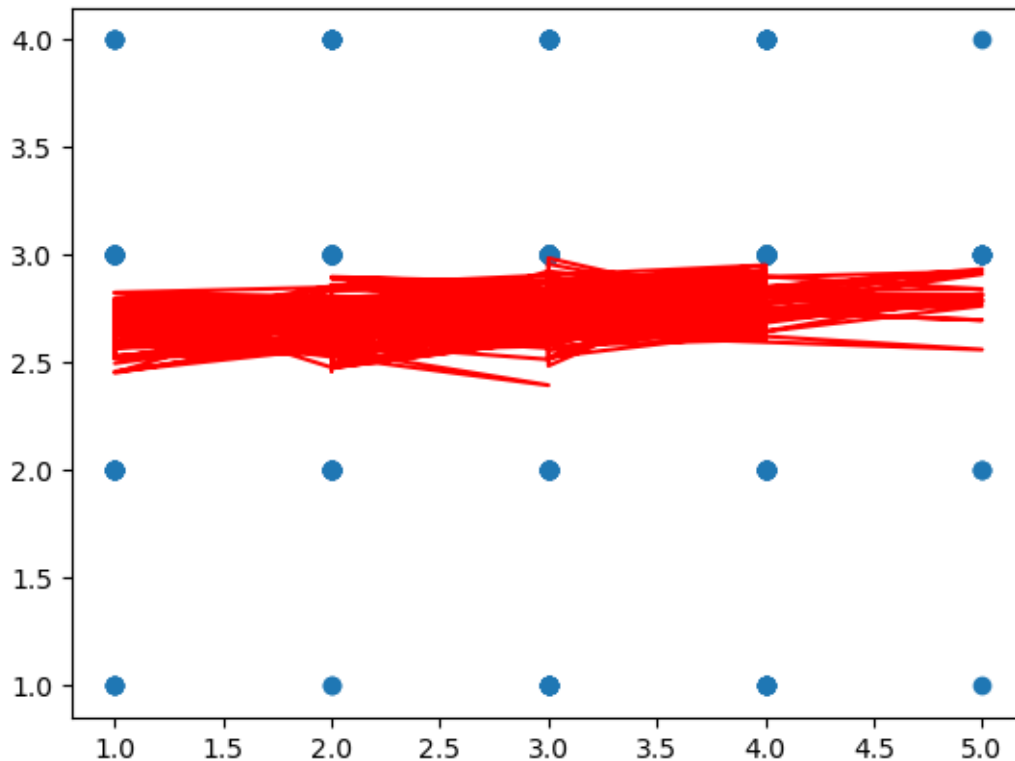
```
[104]: x_scatter = x_train_encoded['HourlyRate']  
y_scatter = y_train  
plt.scatter(x_scatter, y_scatter)  
plt.plot(x_scatter,lr.predict(x_train_encoded), color="red")
```

```
[104]: [<matplotlib.lines.Line2D at 0x27f55efffd0>]
```



```
[116]: plt.scatter(x_test_encoded['Education'],y_test)  
plt.plot(x_test_encoded['Education'],lr.predict(x_test_encoded),color="red")
```

```
[116]: [<matplotlib.lines.Line2D at 0x27f54881210>]
```



```
[118]: lr.predict([[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18]])
```

C:\Users\asus\anaconda3\Lib\site-packages\sklearn\base.py:464: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names

```
warnings.warn(
```

```
[118]: array([[2.3155183]])
```

```
[119]: from sklearn.metrics import_
        accuracy_score, confusion_matrix, classification_report, roc_auc_score, roc_curve
```

```
[123]: from sklearn.linear_model import LogisticRegression
        model=LogisticRegression()
```

```
[125]: model.fit(x_train_encoded,y_train)
```

C:\Users\asus\anaconda3\Lib\site-packages\sklearn\utils\validation.py:1184: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().

```
y = column_or_1d(y, warn=True)
```

C:\Users\asus\anaconda3\Lib\site-packages\sklearn\linear_model_logistic.py:460:

```
ConvergenceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.
```

Increase the number of iterations (max_iter) or scale the data as shown in:

<https://scikit-learn.org/stable/modules/preprocessing.html>

Please also refer to the documentation for alternative solver options:

https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression

```
n_iter_i = _check_optimize_result(
```

```
[125]: LogisticRegression()
```

```
[127]: pred=model.predict(x_test_encoded)
```

```
[129]: y_test
```

```
[129]: JobInvolvement
```

442	3
1091	2
981	3
785	3
1332	2
...	...
817	3
399	2
458	1
406	2
590	3

```
[441 rows x 1 columns]
```

```
[131]: from sklearn.metrics import accuracy_score, confusion_matrix, classification_report, roc_auc_score, roc_curve
```

```
[132]: accuracy_score(y_test, pred)
```

```
[132]: 0.5895691609977324
```

```
[133]: confusion_matrix(y_test, pred)
```

```
[133]: array([[ 0,  0, 26,  0],
        [ 0,  0, 109,  0],
        [ 0,  0, 260,  0],
        [ 0,  0,  46,  0]], dtype=int64)
```

```
[151]: print(classification_report(y_test, pred))
```

```
precision    recall  f1-score   support
```

1	0.00	0.00	0.00	26
2	0.00	0.00	0.00	109
3	0.59	1.00	0.74	260
4	0.00	0.00	0.00	46
accuracy			0.59	441
macro avg	0.15	0.25	0.19	441
weighted avg	0.35	0.59	0.44	441

```
C:\Users\asus\anaconda3\Lib\site-
packages\sklearn\metrics\_classification.py:1469: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
_warn_prf(average, modifier, msg_start, len(result))
C:\Users\asus\anaconda3\Lib\site-
packages\sklearn\metrics\_classification.py:1469: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
_warn_prf(average, modifier, msg_start, len(result))
C:\Users\asus\anaconda3\Lib\site-
packages\sklearn\metrics\_classification.py:1469: UndefinedMetricWarning:
Precision and F-score are ill-defined and being set to 0.0 in labels with no
predicted samples. Use `zero_division` parameter to control this behavior.
_warn_prf(average, modifier, msg_start, len(result))
```

```
[153]: probability=model.predict_proba(x_test_encoded)[: ,1]
```

```
[154]: probability
```

```
[154]: array([0.26956807, 0.2430042 , 0.24086044, 0.22049629, 0.23100066,
0.26808826, 0.24627925, 0.27604556, 0.26514207, 0.23539603,
0.20241986, 0.2143144 , 0.26030742, 0.27735474, 0.20456273,
0.26369472, 0.29107501, 0.2932736 , 0.26543938, 0.24273852,
0.27106826, 0.26136365, 0.25554543, 0.25360673, 0.21937664,
0.25736676, 0.21150135, 0.26364052, 0.29107854, 0.28059562,
0.26058912, 0.24837516, 0.21410762, 0.27831267, 0.25255731,
0.24098723, 0.23982778, 0.25491436, 0.24981637, 0.22209124,
0.23247939, 0.27025048, 0.25035332, 0.25950496, 0.27682358,
0.29672344, 0.27482862, 0.27445897, 0.24020014, 0.26857724,
0.22084739, 0.21182976, 0.2536024 , 0.24563018, 0.29345857,
0.27219032, 0.25526369, 0.26955625, 0.25561561, 0.27986276,
0.269 , 0.20241929, 0.26586281, 0.24013507, 0.2284598 ,
0.29199588, 0.27793045, 0.22046354, 0.22393385, 0.27008102,
0.2027003 , 0.20562869, 0.23086558, 0.22604539, 0.24769106,
0.23504568, 0.27703673, 0.27510138, 0.26127609, 0.27293351,
0.28867177, 0.24587447, 0.23911958, 0.29908854, 0.26178921,
0.28628296, 0.24061055, 0.25080934, 0.26273846, 0.18434105,
```


0.26928727, 0.17519976, 0.22843586, 0.26946922, 0.28749625,
0.23528613, 0.22779266, 0.26808747, 0.24274483, 0.25055385,
0.24982879, 0.19636134, 0.22157554, 0.21691942, 0.27099293,
0.24198115, 0.29586372, 0.20828559, 0.24223711, 0.24515754,
0.21957138, 0.29226202, 0.25635868, 0.24079529, 0.22132164,
0.25710966, 0.25287962, 0.24912244, 0.24138803, 0.27290978,
0.18050737, 0.28174288, 0.27879486, 0.21717208, 0.26249491,
0.29926829, 0.27310546, 0.22950031, 0.27078463, 0.27819535,
0.24857794, 0.24728407, 0.23013386, 0.26575957, 0.25586794,
0.26218855, 0.26207445, 0.28763539, 0.26317967, 0.27690665,
0.27214442, 0.25908307, 0.28787379, 0.27392747, 0.27258812,
0.26040274, 0.29876279, 0.2681578 , 0.28513088, 0.28183404,
0.24603238, 0.22089128, 0.26524846, 0.23540375, 0.21254407,
0.25058379, 0.18435994, 0.26182955, 0.23447936, 0.28142113,
0.26879213, 0.2523383 , 0.25431545, 0.21096484, 0.28184637,
0.21838612, 0.28236489, 0.23434372, 0.22468087, 0.24610174,
0.20870716, 0.28909208, 0.25185816, 0.21409392, 0.24544339,
0.25706923, 0.23940609, 0.2484739 , 0.26263278, 0.27450676,
0.23394664, 0.24364493, 0.29711473, 0.22180265, 0.22421746,
0.26224331, 0.21584099, 0.28900494, 0.22851564, 0.28262091,
0.22213865, 0.24420858, 0.23562952, 0.27455932, 0.2471027 ,
0.28718807, 0.27789776, 0.23321922, 0.22431242, 0.27932783,
0.21565477, 0.25858956, 0.26783409, 0.29474432, 0.21608008,
0.2414745 , 0.27535367, 0.26671 , 0.23916197, 0.28450209,
0.22153662, 0.17373564, 0.25897781, 0.23920089, 0.23934051,
0.22680326, 0.24468868, 0.25990107, 0.23265861, 0.26275433,
0.26473222, 0.2683412 , 0.21640506, 0.27834618, 0.25714326,
0.24296612, 0.2698034 , 0.29182743, 0.29519086, 0.27303274,
0.2648296 , 0.26915756, 0.2528014 , 0.24068023, 0.26242577,
0.26640389, 0.28973894, 0.19774879, 0.22541091, 0.22559658,
0.26615013, 0.25736512, 0.24827118, 0.23720072, 0.28858157,
0.26640792, 0.22817941, 0.29302217, 0.27701378, 0.27067655,
0.22317641, 0.27554664, 0.21052523, 0.22684526, 0.27222191,
0.22860891, 0.22960233, 0.23087262, 0.25790646, 0.26359525,
0.27521537, 0.28003737, 0.21801997, 0.24327672, 0.25931986,
0.26037568, 0.27319692, 0.2607984 , 0.2061589 , 0.29520023,
0.2682459 , 0.26134809, 0.18194081, 0.22773388, 0.28122784,
0.29137149, 0.22161628, 0.26008932, 0.21635463, 0.2549834 ,
0.22399741, 0.2617155 , 0.27205757, 0.2525977 , 0.26209474,
0.27065293, 0.25174316, 0.21285374, 0.29213707, 0.26553481,
0.2670428 , 0.2507417 , 0.237957 , 0.25617313, 0.23651505,
0.1873933 , 0.21963317, 0.23491579, 0.22406451, 0.28382827,
0.27571661, 0.25314822, 0.2426788 , 0.27615467, 0.27428274,
0.24957971, 0.22551155, 0.235094 , 0.2772315 , 0.29715493,
0.21223628, 0.26624873, 0.27376115, 0.25443999, 0.25716504,
0.21466437, 0.26721825, 0.25506092, 0.19699402, 0.25303916,
0.27995859, 0.29298312, 0.24958106, 0.28685607, 0.25697988,

```
0.28347155, 0.24682999, 0.25711843, 0.27392554, 0.27112444,
0.27693488, 0.27346932, 0.2223736 , 0.25532082, 0.28099639,
0.18277447, 0.26942512, 0.27262991, 0.2665004 , 0.2627904 ,
0.23037117, 0.29652058, 0.27324508, 0.29173393, 0.25492416,
0.25496567, 0.23830898, 0.27065865, 0.23521753, 0.27409072,
0.25231172, 0.26947773, 0.27978384, 0.21583441, 0.2643591 ,
0.25184933, 0.22994516, 0.26092426, 0.28702179, 0.27806589,
0.24037511, 0.24490783, 0.29450048, 0.26886933, 0.23812792,
0.27087559, 0.28285804, 0.22942595, 0.27655121, 0.28203042,
0.27998587, 0.24575235, 0.30656558, 0.29194056, 0.20673365,
0.2426442 , 0.29557104, 0.26096321, 0.23368586, 0.26978592,
0.26724104, 0.23966547, 0.23552782, 0.24673616, 0.18486419,
0.24772804, 0.2560365 , 0.24461123, 0.27684373, 0.22313183,
0.25349109, 0.24327435, 0.27927174, 0.25388237, 0.26395288,
0.27926954, 0.25869017, 0.24995174, 0.25782938, 0.22778793,
0.28158782, 0.28378829, 0.26255617, 0.22837657, 0.25962119,
0.24762507, 0.30806649, 0.22412744, 0.26633609, 0.29775899,
0.2678604 , 0.26353398, 0.29609588, 0.30156894, 0.23786505,
0.26523454, 0.25653632, 0.26207681, 0.22156106, 0.27156895,
0.24047834, 0.25986641, 0.24184177, 0.25641924, 0.22575366,
0.26979551, 0.20784101, 0.26535035, 0.22775406, 0.26271069,
0.25907923, 0.25487998, 0.2695432 , 0.19738736, 0.2347263 ,
0.25024584, 0.2362352 , 0.25460508, 0.21094276, 0.27721962,
0.26922654])
```

```
[158]: accuracy = accuracy_score(y_test, pred)
```

```
[159]: accuracy
```

```
[159]: 0.5895691609977324
```

```
[162]: from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, confusion_matrix
        precision = precision_score(y_test, pred, average='weighted')
        recall = recall_score(y_test, pred, average='weighted')
        f1 = f1_score(y_test, pred, average='weighted')
        confusion = confusion_matrix(y_test, pred)
```

```
C:\Users\asus\anaconda3\Lib\site-
packages\sklearn\metrics\_classification.py:1469: UndefinedMetricWarning:
Precision is ill-defined and being set to 0.0 in labels with no predicted
samples. Use `zero_division` parameter to control this behavior.
    _warn_prf(average, modifier, msg_start, len(result))
```

```
[163]: precision
```

```
[163]: 0.3475917955995701
```

```
[164]: recall
```

```
[164]: 0.5895691609977324
```

```
[165]: f1
```

```
[165]: 0.43734088975580715
```

```
[166]: confusion
```

```
[166]: array([[ 0,  0, 26,  0],  
          [ 0,  0, 109,  0],  
          [ 0,  0, 260,  0],  
          [ 0,  0,  46,  0]], dtype=int64)
```