DATA ANALYTICS AND REPORTING

EMPLOYEE ATTRITION ANALYSIS USING IBM HR DATASET

NISTHA ARORA: 202410101150078

DIVYATA MAURYA: 202410101150082

SUBMITTED TO: MS.DEEPIKA TIWARI

CONTENT :

USING PANDAS:

- Data cleaning
- Summary of the dataset

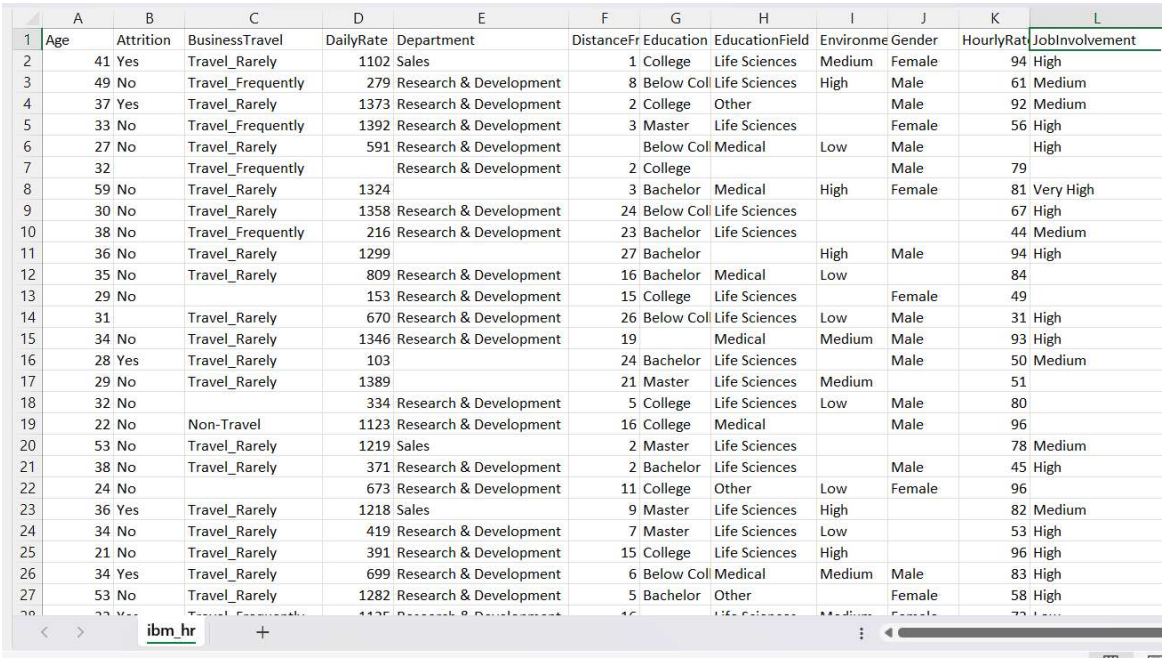
USING MATPLOTLIB:

- Data visualization
- Customizing plots

USING EXCEL:

- Conditional formatting
- Pivot tables
- Filtering and Sorting

IBM HR DATASET



	A	B	C	D	E	F	G	H	I	J	K	L
1	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFr	Education	EducationField	Environme	Gender	HourlyRat	JobInvolvement
2	41	Yes	Travel_Rarely	1102	Sales	1	College	Life Sciences	Medium	Female	94	High
3	49	No	Travel_Frequently	279	Research & Development	8	Below Coll	Life Sciences	High	Male	61	Medium
4	37	Yes	Travel_Rarely	1373	Research & Development	2	College	Other		Male	92	Medium
5	33	No	Travel_Frequently	1392	Research & Development	3	Master	Life Sciences		Female	56	High
6	27	No	Travel_Rarely	591	Research & Development		Below Coll	Medical	Low	Male		High
7	32		Travel_Frequently		Research & Development	2	College			Male	79	
8	59	No	Travel_Rarely	1324		3	Bachelor	Medical	High	Female	81	Very High
9	30	No	Travel_Rarely	1358	Research & Development	24	Below Coll	Life Sciences			67	High
10	38	No	Travel_Frequently	216	Research & Development	23	Bachelor	Life Sciences			44	Medium
11	36	No	Travel_Rarely	1299		27	Bachelor		High	Male	94	High
12	35	No	Travel_Rarely	809	Research & Development	16	Bachelor	Medical	Low		84	
13	29	No		153	Research & Development	15	College	Life Sciences		Female	49	
14	31		Travel_Rarely	670	Research & Development	26	Below Coll	Life Sciences	Low	Male	31	High
15	34	No	Travel_Rarely	1346	Research & Development	19		Medical	Medium	Male	93	High
16	28	Yes	Travel_Rarely	103		24	Bachelor	Life Sciences		Male	50	Medium
17	29	No	Travel_Rarely	1389		21	Master	Life Sciences	Medium		51	
18	32	No		334	Research & Development	5	College	Life Sciences	Low	Male	80	
19	22	No	Non-Travel	1123	Research & Development	16	College	Medical		Male	96	
20	53	No	Travel_Rarely	1219	Sales	2	Master	Life Sciences			78	Medium
21	38	No	Travel_Rarely	371	Research & Development	2	Bachelor	Life Sciences		Male	45	High
22	24	No		673	Research & Development	11	College	Other	Low	Female	96	
23	36	Yes	Travel_Rarely	1218	Sales	9	Master	Life Sciences	High		82	Medium
24	34	No	Travel_Rarely	419	Research & Development	7	Master	Life Sciences	Low		53	High
25	21	No	Travel_Rarely	391	Research & Development	15	College	Life Sciences	High		96	High
26	34	Yes	Travel_Rarely	699	Research & Development	6	Below Coll	Medical	Medium	Male	83	High
27	53	No	Travel_Rarely	1282	Research & Development	5	Bachelor	Other		Female	58	High

ibm_hr

DATA CLEANING

(USING PANDAS)

LOADING THE DATASET FOR DATA CLEANING

IBM HR Dataset

```
import pandas as pd

# Loading the Dataset
Ibm_data=pd.read_csv("/content/ibm_hr.csv")

# Displaying first 5 Rows
Ibm_data.head()
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EnvironmentSatisfaction	Gender	...	PerformanceRating	RelationshipSa
0	41.0	Yes	Travel_Rarely	1102.0	Sales	1.0	College	Life Sciences	Medium	Female	...	Excellent	
1	49.0	No	Travel_Frequently	279.0	Research & Development	8.0	Below College	Life Sciences	High	Male	...	Outstanding	
2	37.0	Yes	Travel_Rarely	1373.0	Research & Development	2.0	College	Other	NaN	Male	...	Excellent	
3	33.0	No	Travel_Frequently	1392.0	Research & Development	3.0	Master	Life Sciences	NaN	Female	...	Excellent	
4	27.0	No	Travel_Rarely	591.0	Research & Development	NaN	Below College	Medical	Low	Male	...	Excellent	

5 rows × 31 columns

STATISTICAL SUMMARY OF THE DATASET

```
# Statistical summary of the dataset
Ibm_data.describe()
```

	Age	DailyRate	DistanceFromHome	HourlyRate	MonthlyIncome	MonthlyRate	NumCompaniesWorked	PercentSalaryHike	StockOptionLevel
count	1470.000000	1470.000000	1470.000000	1470.000000	1470.000000	1470.000000	1470.000000	1470.000000	1470.000000
mean	36.834694	781.765508	9.192517	66.142177	6275.493197	13941.248299	2.698639	15.208844	0.793878
std	9.311449	414.109138	8.106780	20.263198	4806.532649	7360.245317	2.503027	3.660442	0.852077
min	18.000000	1.000000	1.000000	30.000000	0.000000	9.000000	0.000000	11.000000	0.000000
25%	30.000000	444.000000	2.000000	49.000000	2756.750000	7654.750000	1.000000	12.000000	0.000000
50%	35.000000	781.882754	7.000000	66.000000	4845.500000	13911.000000	2.000000	14.000000	1.000000
75%	43.000000	1144.000000	14.000000	84.000000	8150.750000	20325.250000	4.000000	18.000000	1.000000
max	60.000000	1499.000000	29.000000	100.000000	19999.000000	26999.000000	9.000000	25.000000	3.000000

CHECKING FOR NULL VALUES (USING ISNULL())

```
# Cleaning the dataset

# Checking for null values
Ibm_data.isnull().sum()
```

	0
Age	57
Attrition	3
BusinessTravel	7
DailyRate	3
Department	4
DistanceFromHome	2
Education	5
EducationField	6
EnvironmentSatisfaction	447
Gender	2
HourlyRate	20
JobInvolvement	146
JobLevel	2

REMOVING NULL VALUES FROM DATASET

```
# Removing the null values
Ibm_data_dropped= Ibm_data.dropna(axis=0,how="any")
Ibm_data_dropped
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EnvironmentSatisfaction	Gender	...	PerformanceRating
14	28.0	Yes	Travel_Rarely	103.0	Research & Development	24.0	Bachelor	Life Sciences	High	Male	...	Excellent
21	36.0	Yes	Travel_Rarely	1218.0	Sales	9.0	Master	Life Sciences	High	Male	...	Outstanding
22	34.0	No	Travel_Rarely	419.0	Research & Development	7.0	Master	Life Sciences	Low	Female	...	Excellent
36	50.0	Yes	Travel_Rarely	869.0	Sales	3.0	College	Marketing	Low	Male	...	Excellent
46	34.0	No	Non-Travel	1065.0	Sales	23.0	Master	Marketing	Medium	Male	...	Outstanding
...	...	...	...	...	...	...	...	...	...	...	...	...
1441	56.0	No	Non-Travel	667.0	Research & Development	1.0	Master	Life Sciences	High	Male	...	Outstanding
1443	42.0	No	Travel_Rarely	300.0	Research & Development	2.0	Bachelor	Life Sciences	Low	Male	...	Excellent
1445	41.0	No	Travel_Rarely	582.0	Research & Development	28.0	Master	Life Sciences	Low	Female	...	Outstanding

FILLING THE NULL VALUES (USING FILLNA())

```
# Filling the null values
Ibm_data['DailyRate']=Ibm_data['DailyRate'].fillna(mean)
```

```
Ibm_data.isnull().sum()
```

Attrition	3
BusinessTravel	7
DailyRate	0
Department	4
DistanceFromHome	2
Education	5
EducationField	6
EnvironmentSatisfaction	447

Filtering the data

```
# Filtering the data by applying different conditions
filtered_data= (Ibm_data["DailyRate"]>1000)&(Ibm_data["Education"]=="Master")
Ibm_data[filtered_data]
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EnvironmentSatisfaction	Gender	...	PerformanceRating
3	33.0	No	Travel_Frequently	1392.0	Research & Development	3.0	Master	Life Sciences	High	Female	...	Excellent
15	29.0	No	Travel_Rarely	1389.0	Research & Development	21.0	Master	Life Sciences	Medium	Female	...	Excellent
18	53.0	No	Travel_Rarely	1219.0	Sales	2.0	Master	Life Sciences	Low	Female	...	Excellent
21	36.0	Yes	Travel_Rarely	1218.0	Sales	9.0	Master	Life Sciences	High	Male	...	Outstanding
31	44.0	No	Travel_Rarely	1459.0	Research & Development	10.0	Master	Other	High	Male	...	Excellent
...	...	...	...	...	...	...	...	...	...	...	...	...
1450	35.0	No	Travel_Rarely	1146.0	Human Resources	26.0	Master	Life Sciences	High	Female	...	Excellent
1453	36.0	No	Travel_Rarely	1120.0	Sales	11.0	Master	Marketing	Medium	Female	...	Excellent
1455	45.0	No	Travel_Rarely	1322.0	Research & Development	2.0	Master	Life Sciences	High	Male	...	Excellent
1456	35.0	No	Travel_Frequently	1199.0	Research & Development	18.0	Master	Life Sciences	High	Male	...	Excellent

Sorting (using sort_values())

```
# Sorting the dataset
Ibm_data.sort_values(by="Age")
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EnvironmentSatisfaction	Gender	...	PerformanceRating
457	18.0	Yes	Travel_Frequently	1306.0	Sales	5.0	Bachelor	Marketing	Medium	Male	...	Excellent
458	18.0	No	Non-Travel	1094.0	Sales	28.0	Bachelor	Other	High	Male	...	Excellent
301	18.0	No	Travel_Rarely	812.0	Sales	10.0	Bachelor	Medical	High	Female	...	Excellent
296	18.0	Yes	Travel_Rarely	230.0	Research & Development	3.0	Bachelor	Life Sciences	High	Male	...	Excellent
828	18.0	Yes	Non-Travel	247.0	Research & Development	8.0	Below College	Medical	High	Male	...	Excellent
...	...	...	...	...	...	...	...	...	...	...	...	...
427	60.0	No	Travel_Frequently	1499.0	Sales	28.0	Bachelor	Marketing	High	Female	...	Excellent

Removing Duplicate rows

```
# Removing duplictate rows
Ibm_data.drop_duplicates()
```

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EnvironmentSatisfaction	Gender	...	PerformanceRating
0	41.0	Yes	Travel_Rarely	1102.0	Sales	1.0	College	Life Sciences	Medium	Female	...	Excellent
1	49.0	No	Travel_Frequently	279.0	Research & Development	8.0	Below College	Life Sciences	High	Male	...	Outstanding
2	37.0	Yes	Travel_Rarely	1373.0	Research & Development	2.0	College	Other	High	Male	...	Excellent
3	33.0	No	Travel_Frequently	1392.0	Research & Development	3.0	Master	Life Sciences	High	Female	...	Excellent
4	27.0	No	Travel_Rarely	591.0	Research & Development	3.0	Below College	Medical	Low	Male	...	Excellent
...	...	...	...	...	...	...	...	...	...	...	...	...
1465	36.0	No	Travel_Frequently	884.0	Research & Development	23.0	College	Medical	High	Male	...	Excellent

DATA VISUALIZATION

(USING MATPLOTLIB)

Importing required libraries

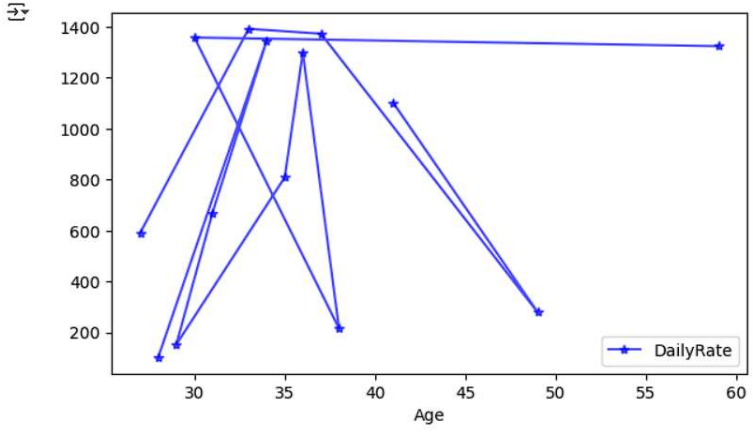
DATA VISUALIZATION

```
import pandas as pd
import matplotlib.pyplot as plt

# Loading the dataset
Ibm_data= pd.read_csv("/content/ibm_hr.csv").head(15)
```

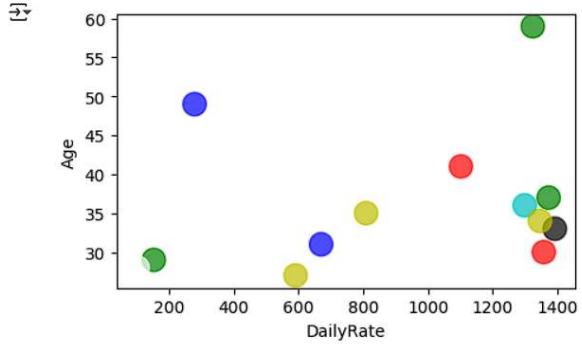
Line Plot

```
# Line plot
Ibm_data.plot(x="Age",y="DailyRate",figsize=(7,4),color="Blue",marker="*",alpha=0.7)
plt.xlabel="Age"
plt.ylabel="DailyRate"
plt.title="Daily Rate by Age"
plt.show()
```



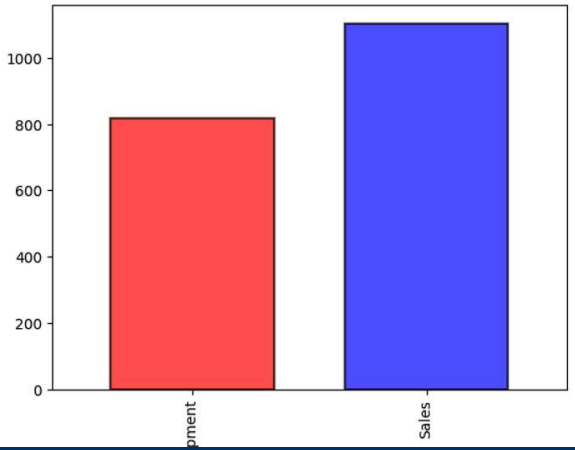
Scatter Plot

```
[6]
✓ 0s
#scatter plot
color=["r","b","g","k","y","m","g","r","w","c","y","g","b","y","w"]
Ibm_data.plot(x="DailyRate",y="Age",kind="scatter",figsize=(5,3),alpha=0.7,color=color,s=200)
plt.show()
```



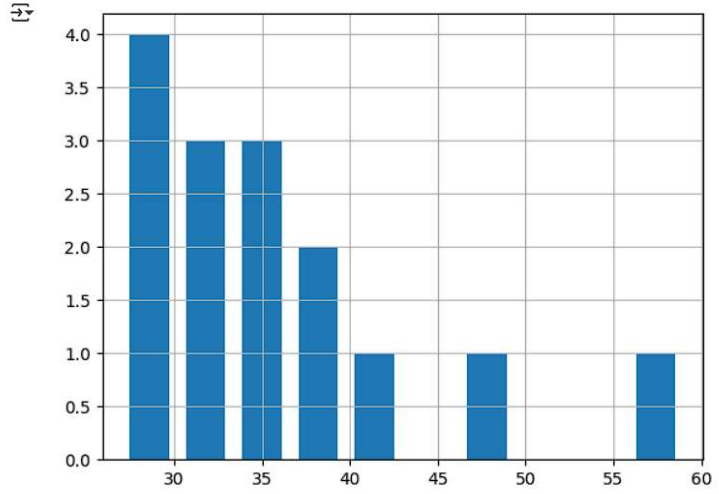
Bar Plot

```
Daily_Rate=Ibm_data.groupby("Department")["DailyRate"].mean().sort_values()
colors=["red","b"]
Daily_Rate.plot(kind="bar", edgecolor="Black",width= 0.7,color=colors,alpha=0.7,linewidth=1.7)
plt.show()
```



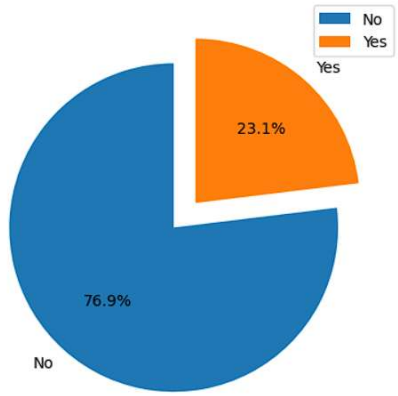
Histogram

```
plt.hist(Ibm_data["Age"],bins=10,align="mid",rwidth=0.7)
plt.grid()
plt.show()
```



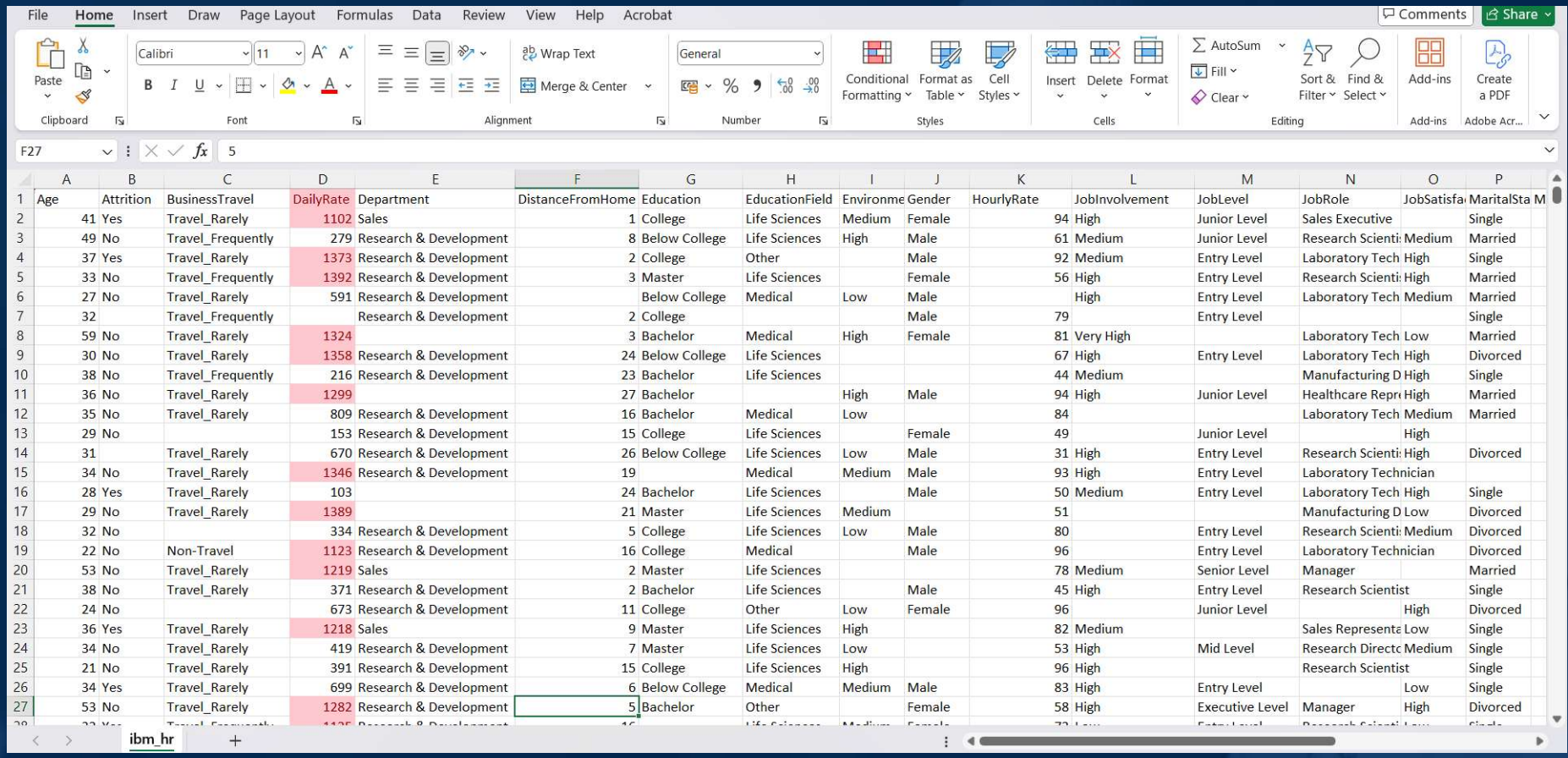
Pie Plot

```
attrition_counts = Ibm_data['Attrition'].value_counts()
ex=[0.2,0.0]
plt.pie(attrition_counts, labels=attrition_counts.index, autopct='%1.1f%%', startangle=90,explode=ex)
plt.legend()
plt.show()
```



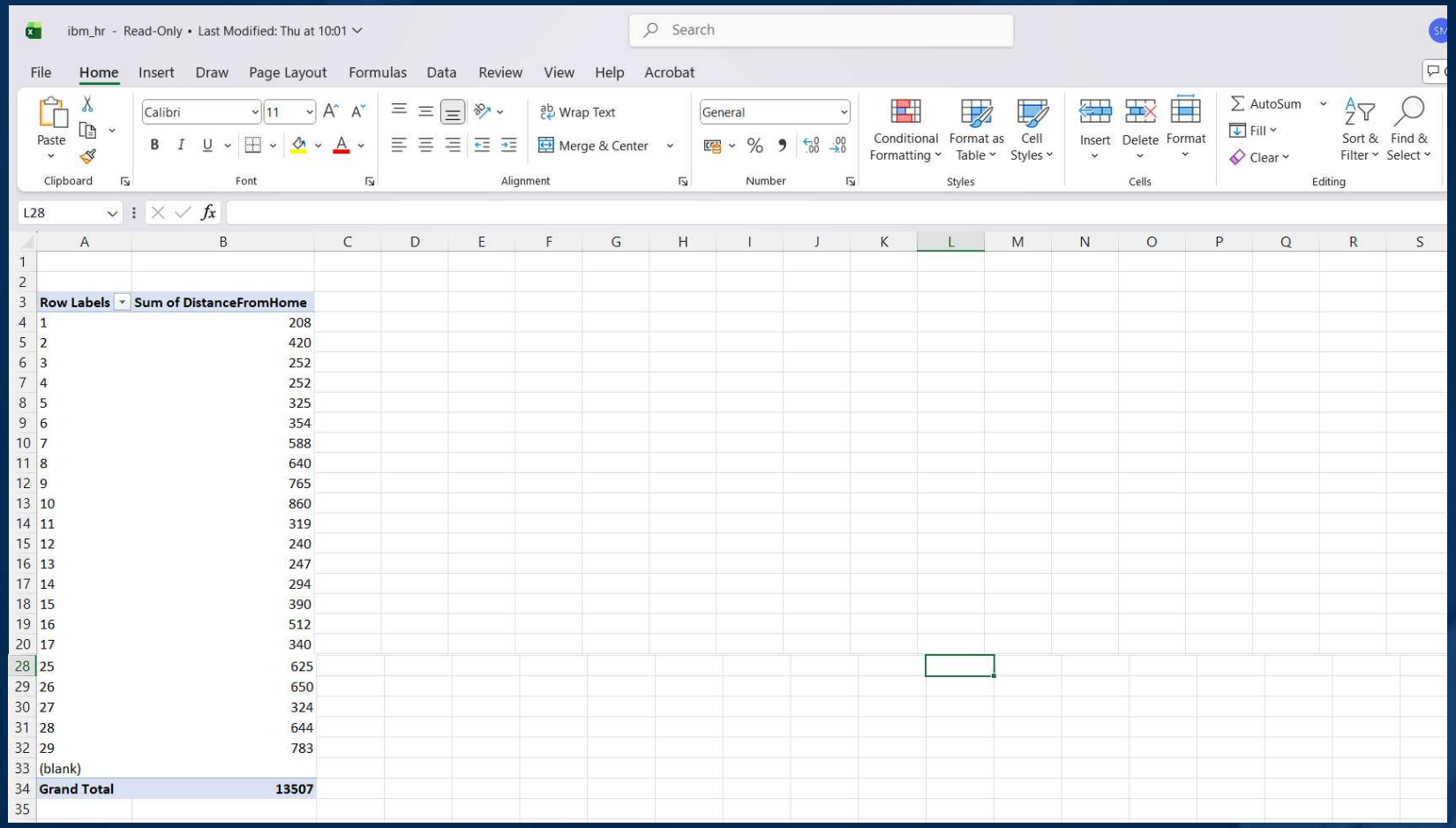
WORKING ON EXCEL

CONDITIONAL FORMATTING



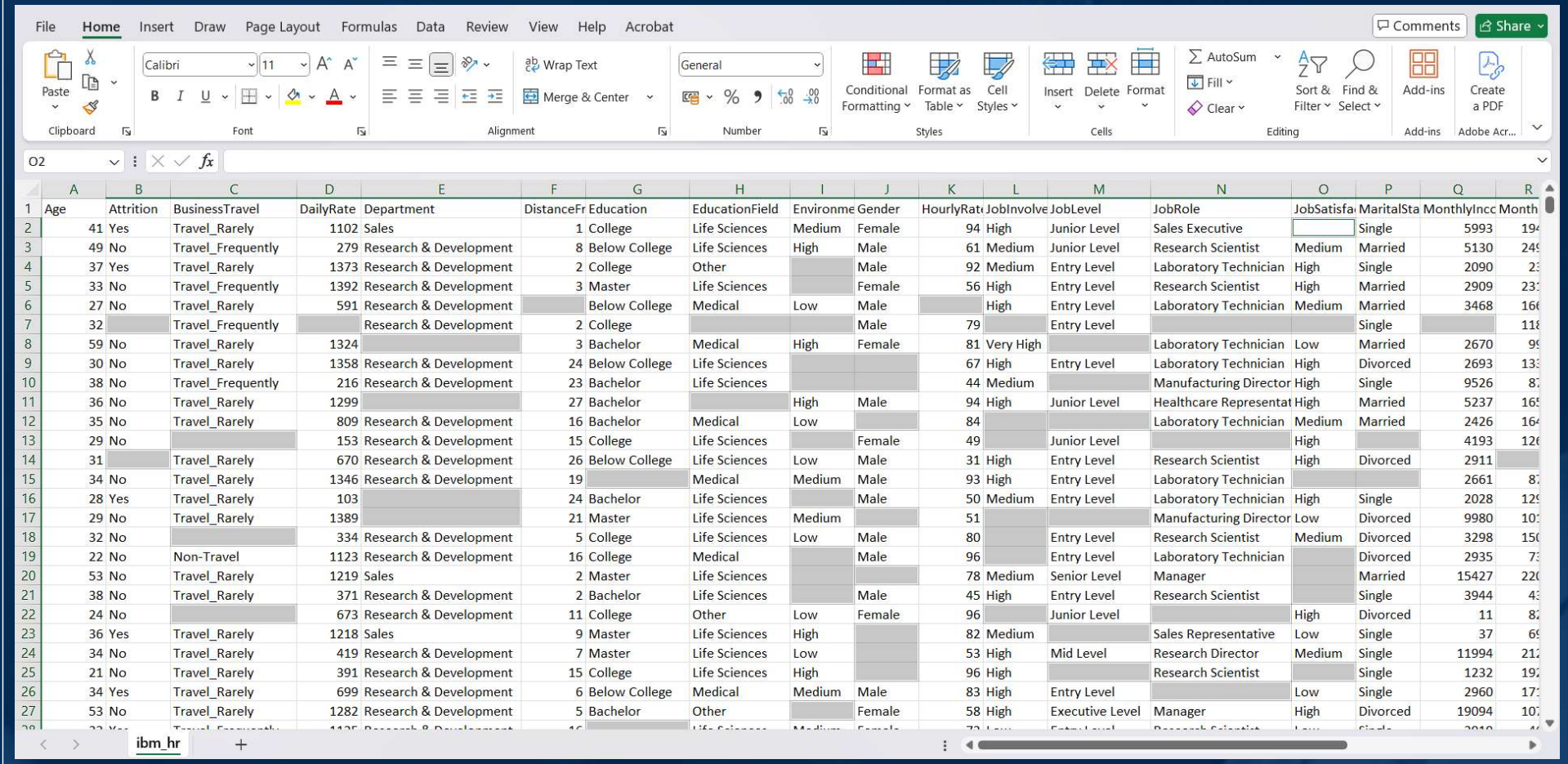
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	Environme	Gender	HourlyRate	JobInvolvement	JobLevel	JobRole	JobSatisfa	MaritalSta M
2	41	Yes	Travel_Rarely	1102	Sales	1	College	Life Sciences	Medium	Female	94	High	Junior Level	Sales Executive		Single
3	49	No	Travel_Frequently	279	Research & Development	8	Below College	Life Sciences	High	Male	61	Medium	Junior Level	Research Scienti	Medium	Married
4	37	Yes	Travel_Rarely	1373	Research & Development	2	College	Other		Male	92	Medium	Entry Level	Laboratory Tech	High	Single
5	33	No	Travel_Frequently	1392	Research & Development	3	Master	Life Sciences		Female	56	High	Entry Level	Research Scienti	High	Married
6	27	No	Travel_Rarely	591	Research & Development		Below College	Medical	Low	Male		High	Entry Level	Laboratory Tech	Medium	Married
7	32		Travel_Frequently		Research & Development	2	College			Male	79		Entry Level			Single
8	59	No	Travel_Rarely	1324		3	Bachelor	Medical	High	Female	81	Very High		Laboratory Tech	Low	Married
9	30	No	Travel_Rarely	1358	Research & Development	24	Below College	Life Sciences			67	High	Entry Level	Laboratory Tech	High	Divorced
10	38	No	Travel_Frequently	216	Research & Development	23	Bachelor	Life Sciences			44	Medium		Manufacturing D	High	Single
11	36	No	Travel_Rarely	1299		27	Bachelor		High	Male	94	High	Junior Level	Healthcare Repr	High	Married
12	35	No	Travel_Rarely	809	Research & Development	16	Bachelor	Medical	Low		84			Laboratory Tech	Medium	Married
13	29	No		153	Research & Development	15	College	Life Sciences		Female	49		Junior Level		High	
14	31		Travel_Rarely	670	Research & Development	26	Below College	Life Sciences	Low	Male	31	High	Entry Level	Research Scienti	High	Divorced
15	34	No	Travel_Rarely	1346	Research & Development	19		Medical	Medium	Male	93	High	Entry Level	Laboratory Technician		
16	28	Yes	Travel_Rarely	103		24	Bachelor	Life Sciences		Male	50	Medium	Entry Level	Laboratory Tech	High	Single
17	29	No	Travel_Rarely	1389		21	Master	Life Sciences	Medium		51			Manufacturing D	Low	Divorced
18	32	No		334	Research & Development	5	College	Life Sciences	Low	Male	80		Entry Level	Research Scienti	Medium	Divorced
19	22	No	Non-Travel	1123	Research & Development	16	College	Medical		Male	96		Entry Level	Laboratory Technician		Divorced
20	53	No	Travel_Rarely	1219	Sales	2	Master	Life Sciences			78	Medium	Senior Level	Manager		Married
21	38	No	Travel_Rarely	371	Research & Development	2	Bachelor	Life Sciences		Male	45	High	Entry Level	Research Scientist		Single
22	24	No		673	Research & Development	11	College	Other	Low	Female	96		Junior Level		High	Divorced
23	36	Yes	Travel_Rarely	1218	Sales	9	Master	Life Sciences	High		82	Medium		Sales Represente	Low	Single
24	34	No	Travel_Rarely	419	Research & Development	7	Master	Life Sciences	Low		53	High	Mid Level	Research Directc	Medium	Single
25	21	No	Travel_Rarely	391	Research & Development	15	College	Life Sciences	High		96	High		Research Scientist		Single
26	34	Yes	Travel_Rarely	699	Research & Development	6	Below College	Medical	Medium	Male	83	High	Entry Level		Low	Single
27	53	No	Travel_Rarely	1282	Research & Development	5	Bachelor	Other		Female	58	High	Executive Level	Manager	High	Divorced

PIVOT TABLE



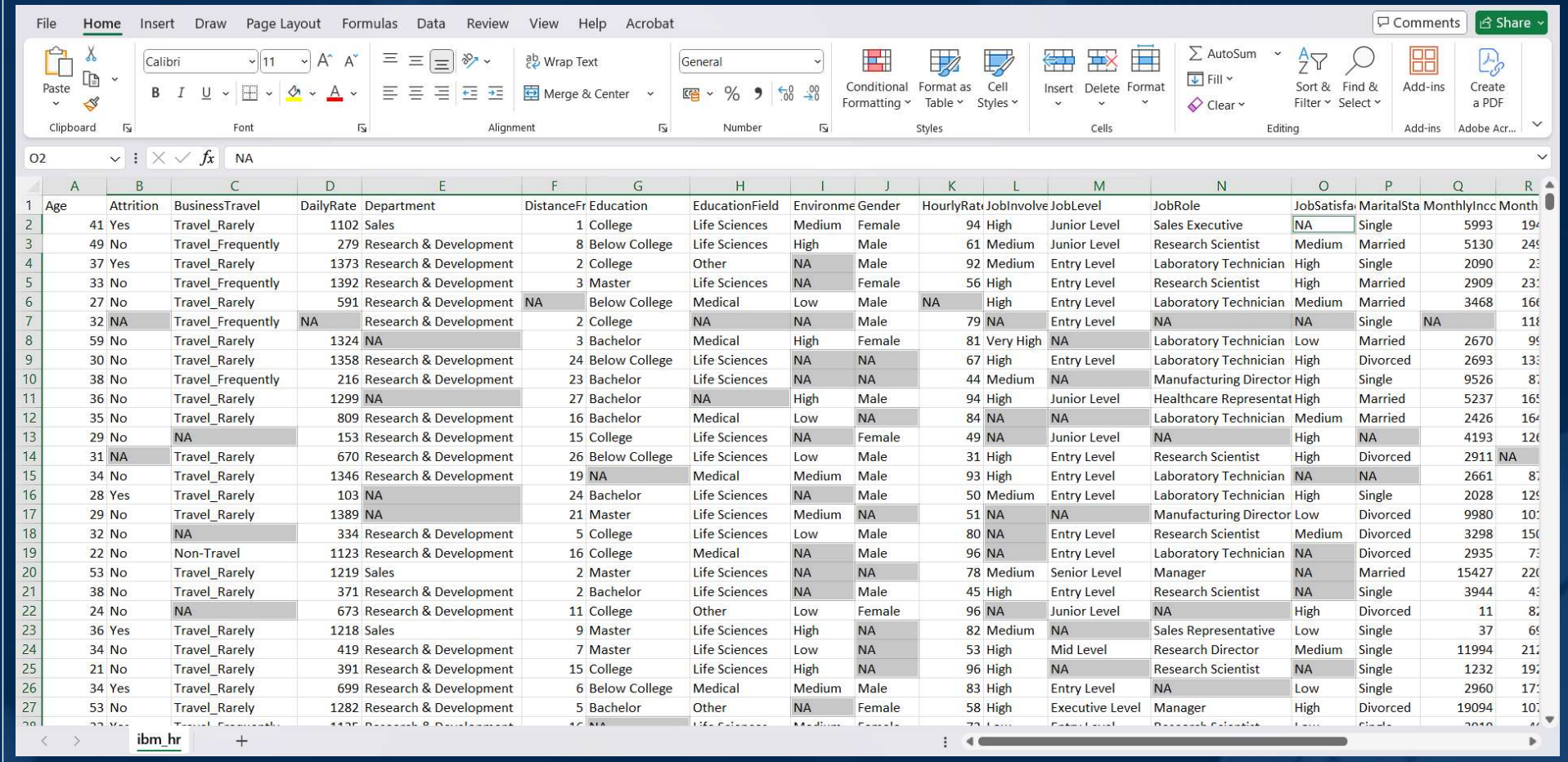
Row Labels	Sum of DistanceFromHome
1	208
2	420
3	252
4	252
5	325
6	354
7	588
8	640
9	765
10	860
11	319
12	240
13	247
14	294
15	390
16	512
17	340
25	625
26	650
27	324
28	644
29	783
(blank)	
Grand Total	**13507**

HIGHLIGHTING THE BLANK VALUES



Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFr	Education	EducationField	Environme	Gender	HourlyRat	JobInvolve	JobLevel	JobRole	JobSatisfa	MaritalSta	MonthlyIncc	Month
41	Yes	Travel_Rarely	1102	Sales	1	College	Life Sciences	Medium	Female	94	High	Junior Level	Sales Executive		Single	5993	194
49	No	Travel_Frequently	279	Research & Development	8	Below College	Life Sciences	High	Male	61	Medium	Junior Level	Research Scientist	Medium	Married	5130	249
37	Yes	Travel_Rarely	1373	Research & Development	2	College	Other		Male	92	Medium	Entry Level	Laboratory Technician	High	Single	2090	23
33	No	Travel_Frequently	1392	Research & Development	3	Master	Life Sciences		Female	56	High	Entry Level	Research Scientist	High	Married	2909	231
27	No	Travel_Rarely	591	Research & Development		Below College	Medical	Low	Male		High	Entry Level	Laboratory Technician	Medium	Married	3468	166
32		Travel_Frequently		Research & Development	2	College			Male	79		Entry Level			Single		118
59	No	Travel_Rarely	1324		3	Bachelor	Medical	High	Female	81	Very High		Laboratory Technician	Low	Married	2670	99
30	No	Travel_Rarely	1358	Research & Development	24	Below College	Life Sciences			67	High	Entry Level	Laboratory Technician	High	Divorced	2693	133
38	No	Travel_Frequently	216	Research & Development	23	Bachelor	Life Sciences			44	Medium		Manufacturing Director	High	Single	9526	87
36	No	Travel_Rarely	1299		27	Bachelor		High	Male	94	High	Junior Level	Healthcare Representat	High	Married	5237	165
35	No	Travel_Rarely	809	Research & Development	16	Bachelor	Medical	Low		84			Laboratory Technician	Medium	Married	2426	164
29	No		153	Research & Development	15	College	Life Sciences		Female	49		Junior Level		High		4193	126
31		Travel_Rarely	670	Research & Development	26	Below College	Life Sciences	Low	Male	31	High	Entry Level	Research Scientist	High	Divorced	2911	
34	No	Travel_Rarely	1346	Research & Development	19		Medical	Medium	Male	93	High	Entry Level	Laboratory Technician			2661	87
28	Yes	Travel_Rarely	103		24	Bachelor	Life Sciences		Male	50	Medium	Entry Level	Laboratory Technician	High	Single	2028	129
29	No	Travel_Rarely	1389		21	Master	Life Sciences	Medium		51			Manufacturing Director	Low	Divorced	9980	101
32	No		334	Research & Development	5	College	Life Sciences	Low	Male	80		Entry Level	Research Scientist	Medium	Divorced	3298	150
22	No	Non-Travel	1123	Research & Development	16	College	Medical		Male	96		Entry Level	Laboratory Technician		Divorced	2935	73
53	No	Travel_Rarely	1219	Sales	2	Master	Life Sciences			78	Medium	Senior Level	Manager		Married	15427	220
38	No	Travel_Rarely	371	Research & Development	2	Bachelor	Life Sciences		Male	45	High	Entry Level	Research Scientist		Single	3944	43
24	No		673	Research & Development	11	College	Other	Low	Female	96		Junior Level		High	Divorced	11	82
36	Yes	Travel_Rarely	1218	Sales	9	Master	Life Sciences	High		82	Medium		Sales Representative	Low	Single	37	69
34	No	Travel_Rarely	419	Research & Development	7	Master	Life Sciences	Low		53	High	Mid Level	Research Director	Medium	Single	11994	212
21	No	Travel_Rarely	391	Research & Development	15	College	Life Sciences	High		96	High		Research Scientist		Single	1232	192
34	Yes	Travel_Rarely	699	Research & Development	6	Below College	Medical	Medium	Male	83	High	Entry Level		Low	Single	2960	171
53	No	Travel_Rarely	1282	Research & Development	5	Bachelor	Other		Female	58	High	Executive Level	Manager	High	Divorced	19094	107

REPLACING BLANK VALUES WITH " NA "



File Home Insert Draw Page Layout Formulas Data Review View Help Acrobat — Comments — Share

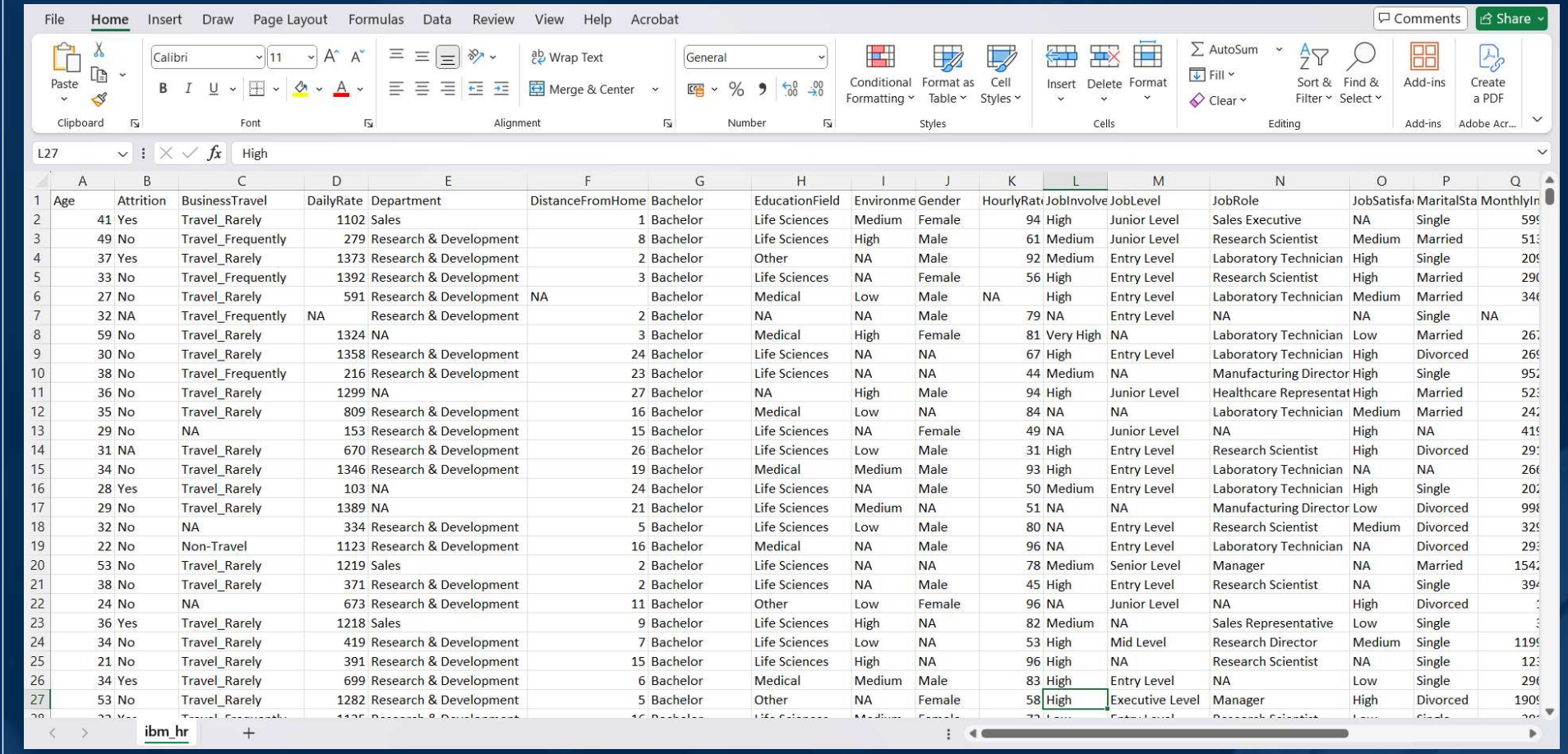
Paste — Clipboard — Calibri 11 — Font — Wrap Text — Merge & Center — Alignment — General — Number — Conditional Formatting — Format as Table — Cell Styles — Styles — Insert — Delete — Format — Cells — AutoSum — Fill — Clear — Sort & Filter — Find & Select — Editing — Add-ins — Create a PDF — Adobe Acr...

O2 | NA

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFr	Education	EducationField	Environme	Gender	HourlyRat	JobInvolve	JobLevel	JobRole	JobSatisfa	MaritalSta	MonthlyInco	Month
2	41	Yes	Travel_Rarely	1102	Sales	1	College	Life Sciences	Medium	Female	94	High	Junior Level	Sales Executive	NA	Single	5993	19
3	49	No	Travel_Frequently	279	Research & Development	8	Below College	Life Sciences	High	Male	61	Medium	Junior Level	Research Scientist	Medium	Married	5130	24
4	37	Yes	Travel_Rarely	1373	Research & Development	2	College	Other	NA	Male	92	Medium	Entry Level	Laboratory Technician	High	Single	2090	2
5	33	No	Travel_Frequently	1392	Research & Development	3	Master	Life Sciences	NA	Female	56	High	Entry Level	Research Scientist	High	Married	2909	23
6	27	No	Travel_Rarely	591	Research & Development	NA	Below College	Medical	Low	Male	NA	High	Entry Level	Laboratory Technician	Medium	Married	3468	16
7	32	NA	Travel_Frequently	NA	Research & Development	2	College	NA	NA	Male	79	NA	Entry Level	NA	NA	Single	NA	11
8	59	No	Travel_Rarely	1324	NA	3	Bachelor	Medical	High	Female	81	Very High	NA	Laboratory Technician	Low	Married	2670	9
9	30	No	Travel_Rarely	1358	Research & Development	24	Below College	Life Sciences	NA	NA	67	High	Entry Level	Laboratory Technician	High	Divorced	2693	13
10	38	No	Travel_Frequently	216	Research & Development	23	Bachelor	Life Sciences	NA	NA	44	Medium	NA	Manufacturing Director	High	Single	9526	8
11	36	No	Travel_Rarely	1299	NA	27	Bachelor	NA	High	Male	94	High	Junior Level	Healthcare Representat	High	Married	5237	16
12	35	No	Travel_Rarely	809	Research & Development	16	Bachelor	Medical	Low	NA	84	NA	NA	Laboratory Technician	Medium	Married	2426	16
13	29	No	NA	153	Research & Development	15	College	Life Sciences	NA	Female	49	NA	Junior Level	NA	High	NA	4193	12
14	31	NA	Travel_Rarely	670	Research & Development	26	Below College	Life Sciences	Low	Male	31	High	Entry Level	Research Scientist	High	Divorced	2911	NA
15	34	No	Travel_Rarely	1346	Research & Development	19	NA	Medical	Medium	Male	93	High	Entry Level	Laboratory Technician	NA	NA	2661	8
16	28	Yes	Travel_Rarely	103	NA	24	Bachelor	Life Sciences	NA	Male	50	Medium	Entry Level	Laboratory Technician	High	Single	2028	12
17	29	No	Travel_Rarely	1389	NA	21	Master	Life Sciences	Medium	NA	51	NA	NA	Manufacturing Director	Low	Divorced	9980	10
18	32	No	NA	334	Research & Development	5	College	Life Sciences	Low	Male	80	NA	Entry Level	Research Scientist	Medium	Divorced	3298	15
19	22	No	Non-Travel	1123	Research & Development	16	College	Medical	NA	Male	96	NA	Entry Level	Laboratory Technician	NA	Divorced	2935	7
20	53	No	Travel_Rarely	1219	Sales	2	Master	Life Sciences	NA	NA	78	Medium	Senior Level	Manager	NA	Married	15427	22
21	38	No	Travel_Rarely	371	Research & Development	2	Bachelor	Life Sciences	NA	Male	45	High	Entry Level	Research Scientist	NA	Single	3944	4
22	24	No	NA	673	Research & Development	11	College	Other	Low	Female	96	NA	Junior Level	NA	High	Divorced	11	8
23	36	Yes	Travel_Rarely	1218	Sales	9	Master	Life Sciences	High	NA	82	Medium	NA	Sales Representative	Low	Single	37	6
24	34	No	Travel_Rarely	419	Research & Development	7	Master	Life Sciences	Low	NA	53	High	Mid Level	Research Director	Medium	Single	11994	21
25	21	No	Travel_Rarely	391	Research & Development	15	College	Life Sciences	High	NA	96	High	NA	Research Scientist	NA	Single	1232	19
26	34	Yes	Travel_Rarely	699	Research & Development	6	Below College	Medical	Medium	Male	83	High	Entry Level	NA	Low	Single	2960	17
27	53	No	Travel_Rarely	1282	Research & Development	5	Bachelor	Other	NA	Female	58	High	Executive Level	Manager	High	Divorced	19094	10

ibm_hr

SORTING THE COLUMNS IN ALPHABETICAL ORDER



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Bachelor	EducationField	Environme	Gender	HourlyRat	JobInvolve	JobLevel	JobRole	JobSatisfa	MaritalSta	MonthlyIn
2	41	Yes	Travel_Rarely	1102	Sales	1	Bachelor	Life Sciences	Medium	Female	94	High	Junior Level	Sales Executive	NA	Single	59
3	49	No	Travel_Frequently	279	Research & Development	8	Bachelor	Life Sciences	High	Male	61	Medium	Junior Level	Research Scientist	Medium	Married	513
4	37	Yes	Travel_Rarely	1373	Research & Development	2	Bachelor	Other	NA	Male	92	Medium	Entry Level	Laboratory Technician	High	Single	209
5	33	No	Travel_Frequently	1392	Research & Development	3	Bachelor	Life Sciences	NA	Female	56	High	Entry Level	Research Scientist	High	Married	290
6	27	No	Travel_Rarely	591	Research & Development	NA	Bachelor	Medical	Low	Male	NA	High	Entry Level	Laboratory Technician	Medium	Married	346
7	32	NA	Travel_Frequently	NA	Research & Development	2	Bachelor	NA	NA	Male	79	NA	Entry Level	NA	NA	Single	NA
8	59	No	Travel_Rarely	1324	NA	3	Bachelor	Medical	High	Female	81	Very High	NA	Laboratory Technician	Low	Married	267
9	30	No	Travel_Rarely	1358	Research & Development	24	Bachelor	Life Sciences	NA	NA	67	High	Entry Level	Laboratory Technician	High	Divorced	269
10	38	No	Travel_Frequently	216	Research & Development	23	Bachelor	Life Sciences	NA	NA	44	Medium	NA	Manufacturing Director	High	Single	952
11	36	No	Travel_Rarely	1299	NA	27	Bachelor	NA	High	Male	94	High	Junior Level	Healthcare Representat	High	Married	523
12	35	No	Travel_Rarely	809	Research & Development	16	Bachelor	Medical	Low	NA	84	NA	NA	Laboratory Technician	Medium	Married	242
13	29	No	NA	153	Research & Development	15	Bachelor	Life Sciences	NA	Female	49	NA	Junior Level	NA	High	NA	419
14	31	NA	Travel_Rarely	670	Research & Development	26	Bachelor	Life Sciences	Low	Male	31	High	Entry Level	Research Scientist	High	Divorced	291
15	34	No	Travel_Rarely	1346	Research & Development	19	Bachelor	Medical	Medium	Male	93	High	Entry Level	Laboratory Technician	NA	NA	266
16	28	Yes	Travel_Rarely	103	NA	24	Bachelor	Life Sciences	NA	Male	50	Medium	Entry Level	Laboratory Technician	High	Single	202
17	29	No	Travel_Rarely	1389	NA	21	Bachelor	Life Sciences	Medium	NA	51	NA	NA	Manufacturing Director	Low	Divorced	998
18	32	No	NA	334	Research & Development	5	Bachelor	Life Sciences	Low	Male	80	NA	Entry Level	Research Scientist	Medium	Divorced	329
19	22	No	Non-Travel	1123	Research & Development	16	Bachelor	Medical	NA	Male	96	NA	Entry Level	Laboratory Technician	NA	Divorced	293
20	53	No	Travel_Rarely	1219	Sales	2	Bachelor	Life Sciences	NA	NA	78	Medium	Senior Level	Manager	NA	Married	1542
21	38	No	Travel_Rarely	371	Research & Development	2	Bachelor	Life Sciences	NA	Male	45	High	Entry Level	Research Scientist	NA	Single	394
22	24	No	NA	673	Research & Development	11	Bachelor	Other	Low	Female	96	NA	Junior Level	NA	High	Divorced	1
23	36	Yes	Travel_Rarely	1218	Sales	9	Bachelor	Life Sciences	High	NA	82	Medium	NA	Sales Representative	Low	Single	3
24	34	No	Travel_Rarely	419	Research & Development	7	Bachelor	Life Sciences	Low	NA	53	High	Mid Level	Research Director	Medium	Single	1199
25	21	No	Travel_Rarely	391	Research & Development	15	Bachelor	Life Sciences	High	NA	96	High	NA	Research Scientist	NA	Single	123
26	34	Yes	Travel_Rarely	699	Research & Development	6	Bachelor	Medical	Medium	Male	83	High	Entry Level	NA	Low	Single	296
27	53	No	Travel_Rarely	1282	Research & Development	5	Bachelor	Other	NA	Female	58	High	Executive Level	Manager	High	Divorced	1909

CONCLUSION

This project uses data analytics in understanding and addressing employee attrition.

By using tools like Python (Pandas & Matplotlib) and Microsoft Excel, we cleaned, visualized, and explored key patterns in the IBM HR dataset.
We cleaned the data, visualizes it and further worked in excel to prepare it for further analysis.
Our analysis helped identify factors influencing employee turnover, enabling HR teams to make informed decisions.


THANK YOU