

## UNDERSTANDING THE DATASET

```
import pandas as pd
df = pd.read_csv('manager_survey_data.csv')
df
```



	EmployeeID	JobInvolvement	PerformanceRating
0	1	3	3
1	2	2	4
2	3	3	3
3	4	2	3
4	5	3	3
...	...	...	...
4405	4406	3	3
4406	4407	2	3
4407	4408	3	4
4408	4409	2	3
4409	4410	4	3

4410 rows x 3 columns

#load the data

```
from google.colab import files # use to load data on Google colab
uploaded = files.upload()# use to load data on colab
```

manager\_su...y\_data.csv

• **manager\_survey\_data.csv**(text/csv) - 43038 bytes, last modified: 11/15/2022 - 100% done  
 Saving manager\_survey\_data.csv to manager\_survey\_data (1).csv

# store the data into the df variable

```
df = pd . read_csv('manager_survey_data.csv')
df.head(7)#print the first 7 rows
```

	EmployeeID	JobInvolvement	PerformanceRating
0	1	3	3
1	2	2	4
2	3	3	3
3	4	2	3
4	5	3	3
5	6	3	3
6	7	3	4

```
# Get the number of rows and number of columns in the data  
df.shape
```

```
(4410, 3)
```

