

UNDERSTANDING THE DATASET

This project is based on an understanding of the factors to keep employees at the Company and which prompt others to leave. The data can be downloaded from the [Dataset](#): We need to use only (3 files - General_data.csv, Employee_Survey_Data.csv, Manager_Survey_data.csv) for the current project.

Let us try to understand each field of the data (general_data.csv)

Below are the values each column has. The column names are pretty self-explanatory.

1. AGE Numerical Value
2. ATTRITION Employee leaving the company (0=no, 1=yes)
3. BUSINESS TRAVEL (1=No Travel, 2=Travel Frequently, 3=Travel Rarely)
4. DEPARTMENT (1=HR, 2=R&D, 3=Sales)
5. DISTANCE FROM HOME Numerical Value - THE DISTANCE FROM WORK TO HOME
6. EDUCATION Numerical Value. (1 'Below College' 2 'College' 3 'Bachelor' 4 'Master' 5 'Doctor')
7. EDUCATION FIELD (1=HR, 2=LIFE SCIENCES, 3=MARKETING, 4=MEDICAL SCIENCES, 5=OTHERS, 6=TECHNICAL)
8. EMPLOYEE COUNT Numerical Value
9. EMPLOYEE ID Numerical Value
10. GENDER (1=FEMALE, 2=MALE)
11. JOB LEVEL Numerical Value
12. JOB ROLE (1=HR REP, 2=HR, 3=LAB TECHNICIAN, 4=MANAGER, 5= MANAGING DIRECTOR, 6= RESEARCH DIRECTOR, 7= RESEARCH SCIENTIST, 8=SALES EXECUTIVE, 9= SALES REPRESENTATIVE)
13. MARITAL STATUS (1=DIVORCED, 2=MARRIED, 3=SINGLE)
14. MONTHLY INCOME Numerical Value - MONTHLY SALARY
15. NUMCOMPANIES WORKED Numerical Value - NO. OF COMPANIES WORKED AT
16. OVER 18 (1=YES, 2=NO)
17. PERCENT SALARY HIKE Numerical Value - PERCENTAGE INCREASE IN SALARY

18. STANDARD HOURS Numerical Value - STANDARD HOURS
19. STOCK OPTIONS LEVEL Numerical Value - STOCK OPTIONS (Higher the number, the more stock option an employee has)
20. TOTAL WORKING YEARS Numerical Value - TOTAL YEARS WORKED
21. TRAINING TIMES LAST YEAR Numerical Value - HOURS SPENT TRAINING
22. YEARS AT COMPANY Numerical Value - TOTAL NUMBER OF YEARS AT THE COMPANY
23. YEARS SINCE LAST PROMOTION Numerical Value - LAST PROMOTION
24. YEARS WITH CURRENT MANAGER Numerical Value - YEARS SPENT WITH CURRENT MANAGER

b. Let us try to understand about each field of the data
(employee_survey_data.csv)

1. Employee ID
2. Environment Satisfaction (1 'Low' 2 'Medium' 3 'High' 4 'Very High')
3. Job Satisfaction (1 'Low' 2 'Medium' 3 'High' 4 'Very High')
4. Work Life Balance (1 'Bad', 2 'Good', 3 'Better', 4 'Best')

c. Let us try to understand about each field of the data
(manager_survey_data.csv)

1. Employee ID
2. Job Involvement (1 'Low' 2 'Medium' 3 'High' 4 'Very High')
3. Performance Rating (1 'Low', 2 'Good', 3 'Excellent', 4 'Outstanding')

DATA DICTIONARY :

Variable	Meaning	Levels
Age	Age of the employee	
Attrition	Whether the employee left in the previous year or not	
BusinessTravel	How frequently the employees travelled for business purposes in the last year	
Department	Department in company	
DistanceFromHome	Distance from home in kms	
Education	Education Level	1 'Below College' 2 'College' 3 'Bachelor' 4 'Master' 5 'Doctor'
EducationField	Field of education	
EmployeeCount	Employee count	
EmployeeNumber	Employee number/id	
EnvironmentSatisfaction	Work Environment Satisfaction Level	1 'Low' 2 'Medium' 3 'High' 4 'Very High'
Gender	Gender of employee	
JobInvolvement	Job Involvement Level	1 'Low' 2 'Medium' 3 'High' 4 'Very High'
JobLevel	Job level at company on a scale of 1 to 5	
JobRole	Name of job role in company	
		1 'Low'

EMPLOYEE SURVEY DATA :

EmployeeID	EnvironmentSatisfaction	JobSatisfaction	WorkLifeBalance
1	1	3	4
2	2	3	2
3	3	2	2
4	4	4	4
5	5	4	1
6	6	3	2
7	7	1	3
8	8	1	2
9	9	2	4
10	10	2	1
11	11	3	4
12	12	NA	4
13	13	4	1
14	14	1	2
15	15	4	4
16	16	3	4
17	17	4	3
18	18	1	4
19	19	2	2
20	20	1	1
21	21	3	2
22	22	1	2

GENERAL DATA :

File

Home

Insert

References

Data Tools

Developer

Cut

Copy

Paste

Format Painter

Clipboard

Calibri

11

A^A

Font

Alignment

Wrap Text

Merge & Center

Number

General

\$

%

#

.

/

Styles

Conditional Formatting

Format as Table

Cell Styles

Cells

Insert

Delete

Format

Editing

AutoSum

Fill

Clear

Sort & Filter

Find & Select

A1

Age

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Age	Attrition	Business Department	Distance	Education	Education	Employee Gender	Joblevel	JobRole	MaritalStatus	MonthlyIncome	NumComplaints	Over18	PercentSatisfied	StandardScore	StockOptions	TotalWorkTime	TrainingTimeInHours			
2	51 No	Travel_Ra Sales		6	2 Life Scienc	1	Female	1	Healthcar	Married	131160	1 Y	11	8	0	1					
3	31 Yes	Travel_Fr Research		10	1 Life Scienc	1	Female	1	Research	Single	41890	0 Y	23	8	1	6					
4	32 No	Travel_Fr Research		17	4 Other	1	Male	4	Sales Exec	Married	193280	1 Y	15	8	3	5					
5	38 No	Non-Trav Research		2	5 Life Scienc	1	Male	3	Human Re	Married	83210	3 Y	11	8	3	13					
6	32 No	Travel_Ra Research		10	1 Medical	1	Male	1	Sales Exec	Single	23420	4 Y	12	8	2	9					
7	46 No	Travel_Ra Research		8	3 Life Scienc	1	Female	4	Research	Married	40710	3 Y	13	8	0	28					
8	28 Yes	Travel_Ra Research		11	2 Medical	1	Male	2	Sales Exec	Single	58130	2 Y	20	8	1	5					
9	29 No	Travel_Ra Research		18	3 Life Scienc	1	Male	2	Sales Exec	Married	31430	2 Y	22	8	3	10					
10	31 No	Travel_Ra Research		1	3 Life Scienc	1	Male	3	Laborator	Married	20440	0 Y	21	8	0	10					
11	25 No	Non-Trav Research		7	4 Medical	1	Female	4	Laborator	Divorced	134640	1 Y	13	8	1	6					
12	45 No	Travel_Ra Research		17	2 Medical	1	Male	2	Laborator	Married	79910	0 Y	13	8	2	21					
13	36 No	Travel_Ra Research		28	1 Life Scienc	1	Male	1	Laborator	Married	33770	0 Y	12	8	2	16					
14	55 No	Travel_Ra Research		14	4 Life Scienc	1	Female	1	Sales Exec	Single	55380	0 Y	17	8	0	37					
15	47 Yes	Non-Trav Research		1	1 Medical	1	Male	1	Research	Married	57620	1 Y	11	8	2	10					
16	28 No	Travel_Ra Research		1	3 Life Scienc	1	Male	1	Manufactur	Married	25920	1 Y	14	8	0	5					
17	37 No	Travel_Ra Research		1	3 Life Scienc	1	Male	2	Healthcar	Married	53460	4 Y	11	8	0	7					
18	21 No	Travel_Ra Research		3	2 Life Scienc	1	Male	1	Laborator	Single	42130	1 Y	12	8	3	3					
19	37 No	Non-Trav Research		1	3 Medical	1	Male	2	Sales Exec	Divorced	41270	2 Y	13	8	1	15					
20	35 No	Travel_Ra Sales		7	4 Life Scienc	1	Male	1	Sales Rep	Divorced	24380	7 Y	16	8	0	10					
21	38 No	Travel_Ra Research		8	3 Life Scienc	1	Female	1	Manager	Divorced	68700	1 Y	11	8	1	8					
22	26 No	Travel_Fr Research		1	4 Other	1	Male	2	Laborator	Divorced	104470	1 Y	18	8	0	26					
23	50 No	Travel_Ra Sales		8	4 Life Scienc	1	Male	2	Laborator	Divorced	96670	3 Y	23	8	0	28					

IN TIME :

Microsoft Excel

File Home Insert Layout References Send To Back View Formulas Data Tools Developer Help

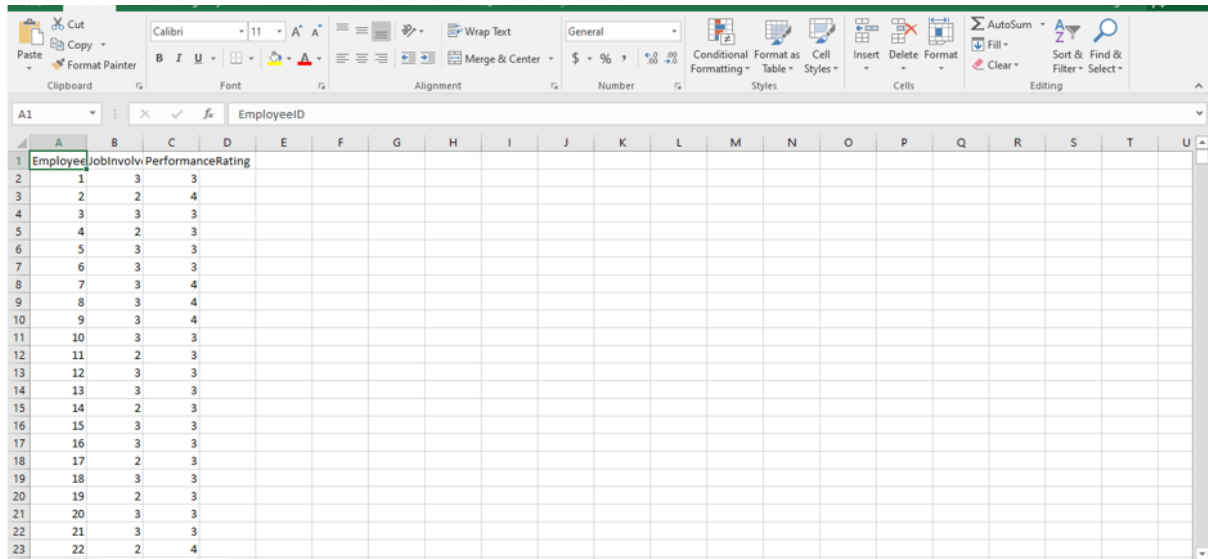
Clipboard Font Alignment Number Styles Cells Editing

Calibri 11 A^A Wrap Text General Conditional Format as Cell Insert Delete Format AutoSum Sort & Find & Filter Select

B I U - [] [] [] Merge & Center \$ % + .00 .00 Formatting Table Styles Cells Fill Clear

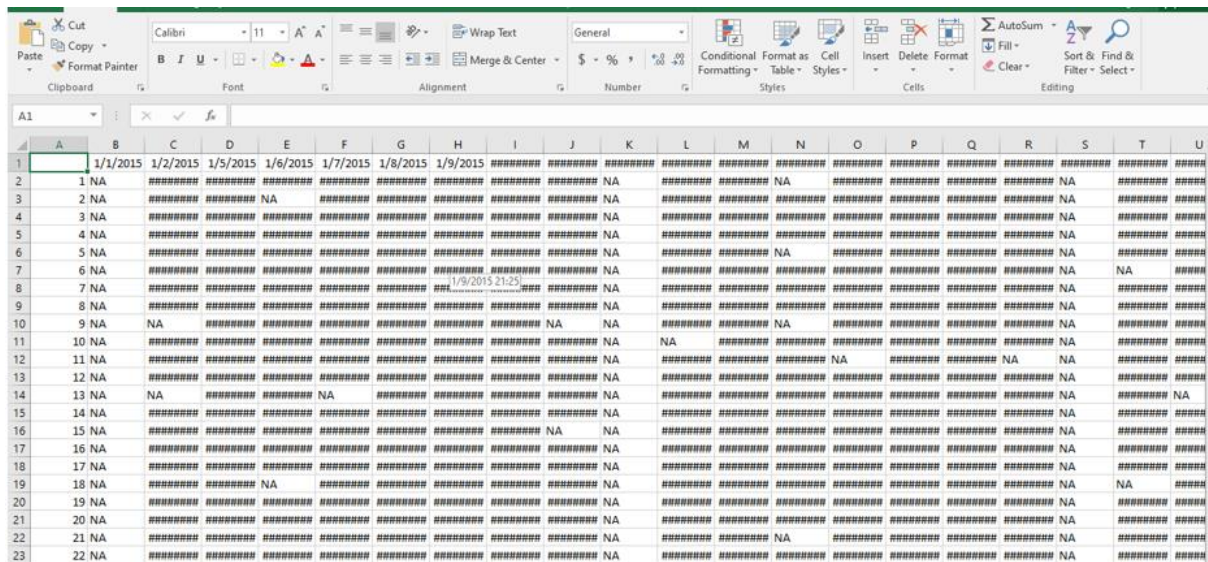
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1		1/1/2015	1/2/2015	1/5/2015	1/6/2015	1/7/2015	1/8/2015	1/9/2015	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####
2		1 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	NA	#####	#####	#####	#####	NA	#####	#####
3		2 NA	#####	#####	NA	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
4		3 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
5		4 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
6		5 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	NA	#####	#####	#####	#####	NA	#####	#####
7		6 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	NA	#####
8		7 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
9		8 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
10		9 NA	NA	#####	#####	#####	#####	#####	#####	NA	NA	#####	#####	NA	#####	#####	#####	#####	NA	#####	#####
11		10 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	NA	#####	#####	#####	#####	#####	#####	NA	#####	#####
12		11 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	NA	#####	#####	NA	NA	#####	#####
13		12 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
14		13 NA	NA	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####
15		14 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
16		15 NA	#####	#####	#####	#####	#####	#####	#####	NA	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
17		16 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
18		17 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
19		18 NA	#####	#####	NA	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	NA	#####
20		19 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
21		20 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####
22		21 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	NA	#####	#####	#####	#####	NA	#####	#####
23		22 NA	#####	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####	#####	#####	#####	#####	#####	NA	#####	#####

MANAGER SURVEY DATA :



EmployeeID	JobInvolvement	PerformanceRating
1	3	3
2	2	4
3	3	3
4	2	3
5	3	3
6	3	3
7	3	4
8	3	4
9	3	4
10	3	3
11	2	3
12	3	3
13	3	3
14	2	3
15	3	3
16	3	3
17	2	3
18	3	3
19	2	3
20	3	3
21	3	3
22	3	3
23	2	4

OUT TIME :



1/1/2015	1/2/2015	1/5/2015	1/6/2015	1/7/2015	1/8/2015	1/9/2015
1 NA						
2 NA						
3 NA						
4 NA						
5 NA						
6 NA						
7 NA						
8 NA						
9 NA						
10 NA						
11 NA						
12 NA						
13 NA						
14 NA						
15 NA						
16 NA						
17 NA						
18 NA						
19 NA						
20 NA						
21 NA						
22 NA						
23 NA						