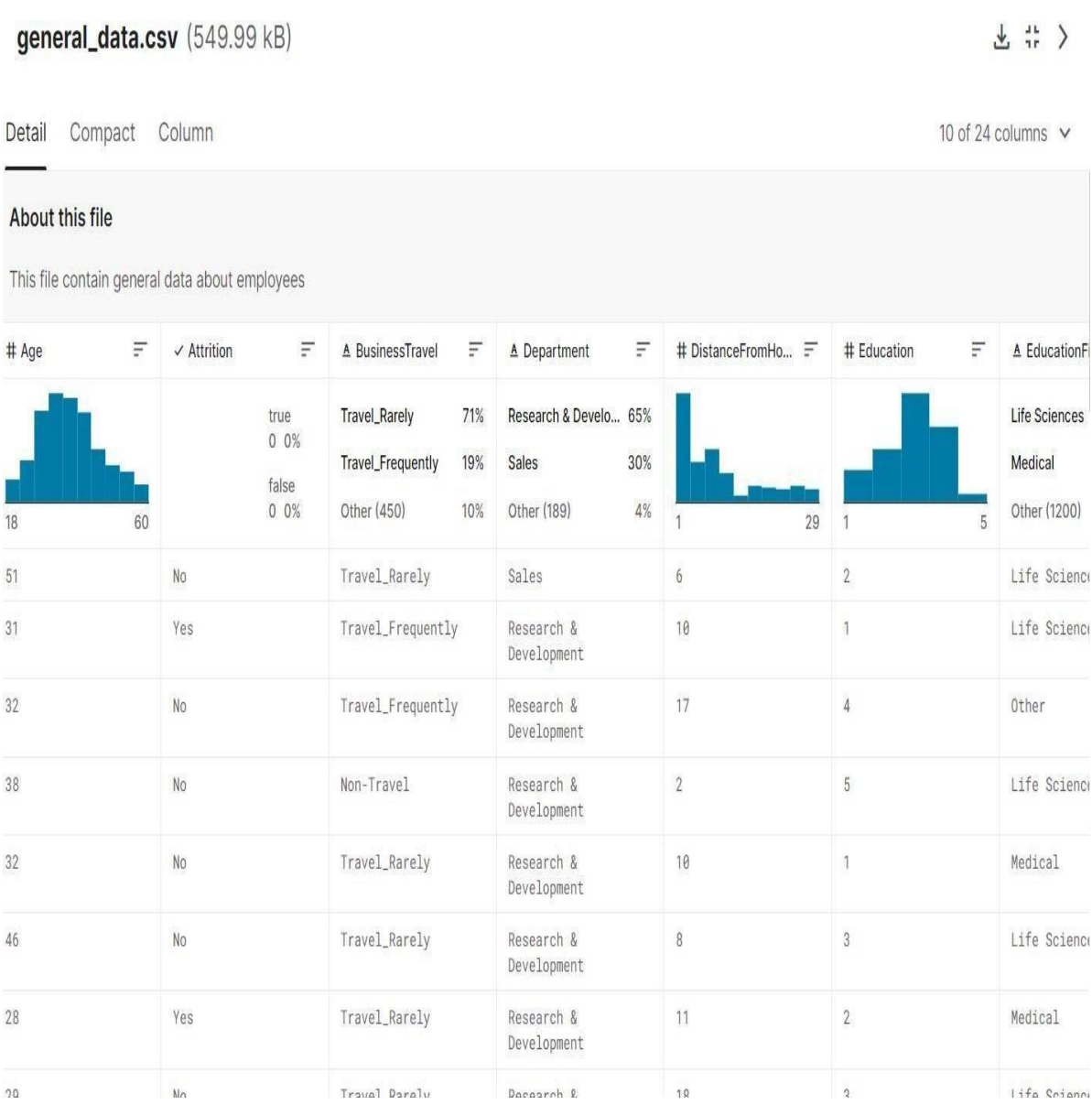
UNDERSTAND THE DATASET

Team ID	PNT2022TMID36382
Project Name	Corporate Employee Attrition Analysis

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



general_data.csv (549.99 kB)

Detail Compact Column 10 of 24 columns

About this file

This file contain general data about employees

# Age	✓ Attrition	A BusinessTravel	A Department	# DistanceFromHo...	# Education	A EducationF
18 60	true 0 0% false 0 0%	Travel_Rarely 71% Travel_Frequently 19% Other (450) 10%	Research & Develo... 65% Sales 30% Other (189) 4%	1 29	1 5	Life Sciences Medical Other (1200)
51	No	Travel_Rarely	Sales	6	2	Life Scienc
31	Yes	Travel_Frequently	Research & Development	10	1	Life Scienc
32	No	Travel_Frequently	Research & Development	17	4	Other
38	No	Non-Travel	Research & Development	2	5	Life Scienc
32	No	Travel_Rarely	Research & Development	10	1	Medical
46	No	Travel_Rarely	Research & Development	8	3	Life Scienc
28	Yes	Travel_Rarely	Research & Development	11	2	Medical
29	No	Travel_Rarely	Research &	10	3	Life Scienc

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')

employee_survey_data.csv (51.96 kB)

Detail Compact Column

About this file

This file contains Survey result

EmployeeID Employee number/id	EnvironmentSatis... Work Environment Satisfaction Level	JobSatisfaction Job Involvement Level Job Involvement Level Job Involvement Level	WorkLifeBalance Work life balance level
1 4410	3 31% 4 30% Other (1726) 39%	4 31% 3 30% Other (1720) 39%	3 60% 2 23% Other (731) 17%
1	3	4	2
2	3	2	4
3	2	2	1
4	4	4	3
5	4	1	3
6	3	2	2
7	1	3	1
8	1	2	3

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

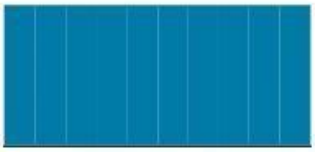
1. Employee ID
2. Job Involvement (1 'Low' 2 'Medium' 3 'High' 4 'Very High')
3. Performance Rating (1 'Low', 2 'Good', 3 'Excellent', 4 'Outstanding')

manager_survey_data.csv (43.04 kB)

Detail Compact Column

About this file

This file contains employees feedback survey about their managers

EmployeeID	JobInvolvement	PerformanceRating
 1 4410	 1 4	 3 4
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