Project Name	Corporate Employee Attrition Analytics
TeamId	PNT2022TMID20674

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.

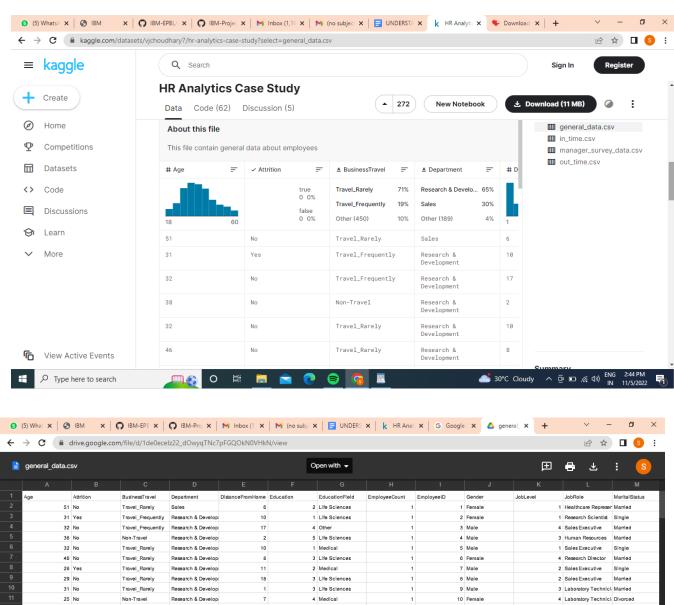
Let us try to understand each field of the data (general_data.csv)
Below are the values each column has. The column names are pretty self-explanatory.

- 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

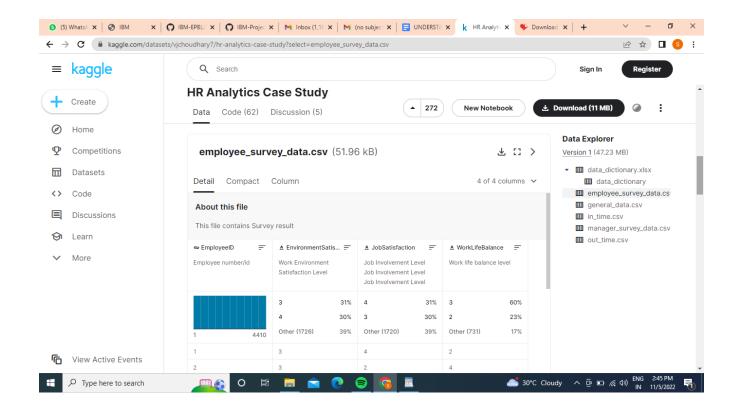
24. YEARS WITH CURRENT MANAGER Numerical Value - YEARS SPENT WITH CURRENT MANAGER

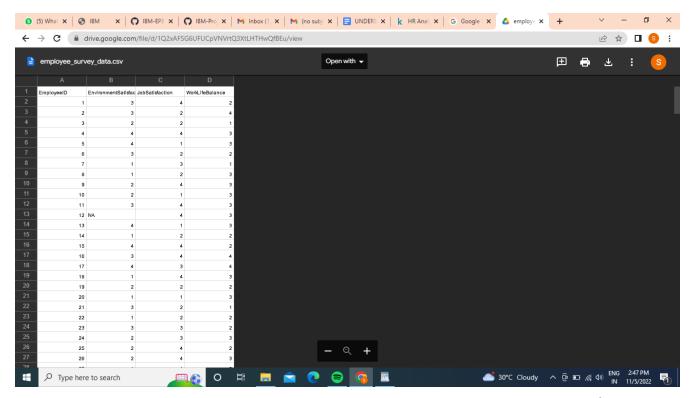


- Travel_Rarely 2 Medical 11 Male 45 No Research & Develop 2 Laboratory Technici Married Travel_Rarely 1 Laboratory Technici Married Research & Develor 55 No Travel_Rarely 4 Life Sciences 13 Female 1 Sales Executive Single 47 Yes Non-Travel Research & Develop 1 Medical 14 Male 1 Research Scientist Married Research & Develop 3 Life Sciences 15 Male 1 Manufacturing Direc Married 28 No Travel_Rarely 3 Life Sciences 16 Male 2 Healthcare Represer Married Research & Develop 1 Laboratory Technici Single 21 No Travel_Rarely Research & Develop 2 Life Sciences 17 Male 37 No Non-Travel Research & Develop 3 Medical 18 Male 2 Sales Executive Divorced Travel_Rarely 4 Life Sciences 19 Male 1 Sales Representative Divorced 35 No 38 No Travel_Rarely 3 Life Sciences 20 Female 26 No Travel Frequently Research & Develop 4 Other 21 Male 2 Laboratory Technici Divorced 22 Male 50 No Travel_Rarely Sales 4 Life Sciences 1 Research Scientist Divorced 53 No Travel_Rarely Research & Develop 4 Life Sciences 23 Female 2 Research Scientist Married 42 No Travel_Rarely Research & Devel 24 Male 1 Manufacturing Direc Married 29 No Travel Frequently Research & Develop 25 Male 1 Laboratory Technici Single Q ^ ② ② □ //(Φ)) ENG Type here to search 0 29°C Cloudy
- 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)
 - 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')

