Working With The Dataset

Understanding The Dataset

Date	30 October 2022
Team ID	PNT2022TMID40488
Project Name	Corporate Employee Attrition Analysis

1. **General Data CSV file**:

   The general_data.csv has totally 24 columns and the have their respected datatypes.

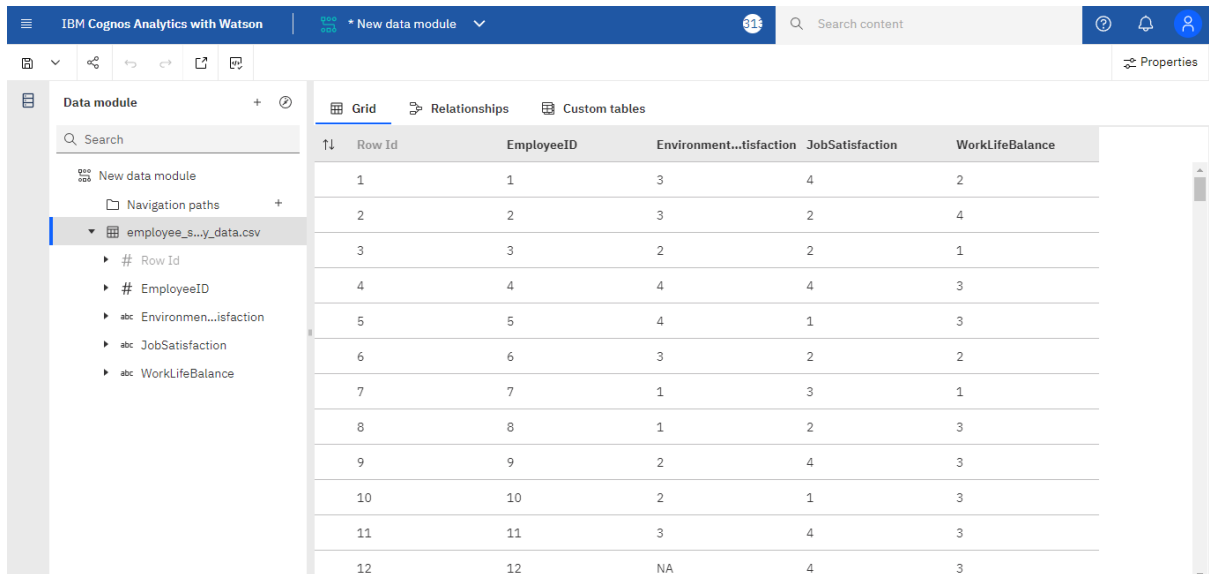Below are the values for each column:.
   1. **AGE:** Numerical Value
   2. **ATTRITION:** Employee leaving the company (0=no, 1=yes)
   3. **BUSINESS TRAVEL:**
      1=No Travel
      2=Travel Frequently
      3=Travel Rarely
   4. **DEPARTMENT:**
      1=HR
      2=R&D
      3=Sales)
   5. **DISTANCE 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.   NUMBER OF COMPANIES 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

2. Employee Survey Data CSV file:

The employee_survey_data.csv has totally 4 columns and the have their respected datatypes.

Row Id	EmployeeID	Environment...tisfaction	JobSatisfaction	WorkLifeBalance
1	1	3	4	2
2	2	3	2	4
3	3	2	2	1
4	4	4	4	3
5	5	4	1	3
6	6	3	2	2
7	7	1	3	1
8	8	1	2	3
9	9	2	4	3
10	10	2	1	3
11	11	3	4	3
12	12	NA	4	3

Below are the values for each column:.
  1. Employee ID
  2. Environment Satisfaction
    1 'Low'
    2 'Medium'
    3 'High'
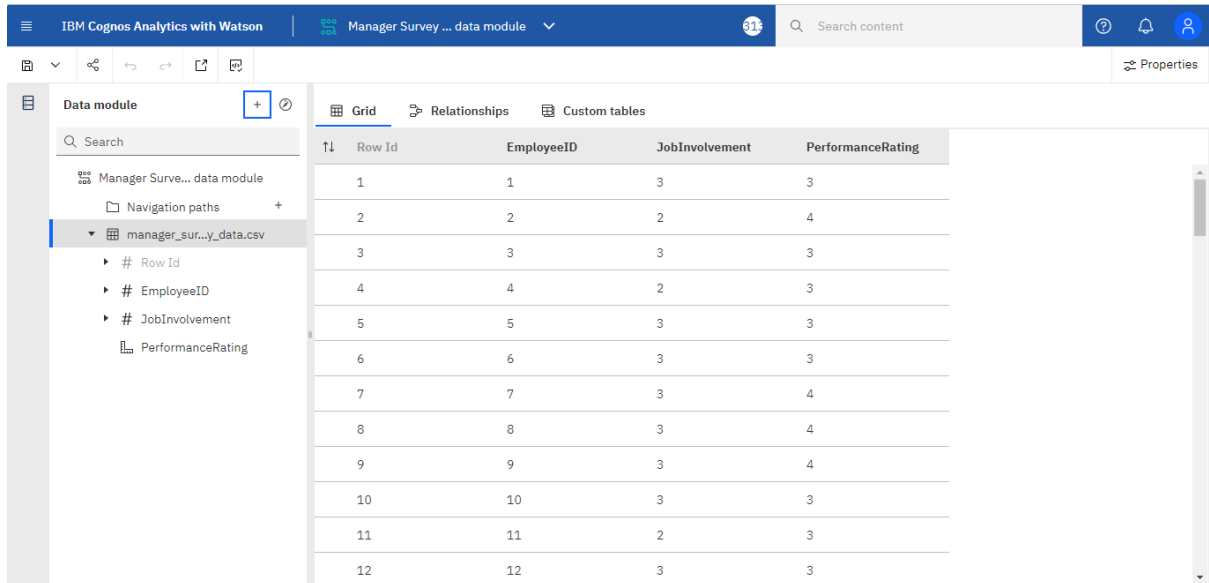   4 'Very High'
  3. Job Satisfaction
    1 'Low'
    2 'Medium'
    3 'High'
    4 'Very High'
  4. Work Life Balance
    1 'Bad'
    2 'Good'
    3 'Better'
    4 'Best'

3. Manager Survey Data CSV file:

The manager_survey_data.csv has totally 4 columns and the have their respected datatypes.

Below are the values for each column:
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'