EDA

**Employee Attrition Analysis**

**Description:** This data set contains various attributes about many employees in a company. It includes both the employees who are currently working and who left the job; Uncover the factors that lead to employee attrition

**Column Description**

|  |  |  |
| --- | --- | --- |
| **No.** | **Column name** | **Description/Comments** |
| 1 | Age | Age of the employee |
| 2 | Attrition | Left the organization or not |
| 3 | BusinessTravel | How frequently the employee travel |
| 4 | Department | Department of the employee |
| 5 | DistanceFromHome | What is the distance of employee's home from office ( in KM) |
| 6 | Education | Education background; Refer to next table for mapping |
| 7 | EducationField | Education field |
| 8 | EmployeeNumber | Unique ID for the employee |
| 9 | EnvironmentSatisfaction | Is employee satisfied with office environment; Refer to the next table for mapping |
| 10 | Gender | Gender |
| 11 | JobInvolvement | Refer to the next table for mapping |
| 12 | JobRole | Designation |
| 13 | JobSatisfaction | Refer to the next table for mapping |
| 14 | MaritalStatus | MaritalStatus |
| 15 | MonthlyIncome | Montly income for the employee |
| 16 | NumCompaniesWorked | No. of companies previously worked |
| 17 | PercentSalaryHike | Last salary hike percentage |
| 18 | PerformanceRating | Refer to the next table for mapping |
| 19 | RelationshipSatisfaction | Refer to the next table for mapping |
| 20 | TotalWorkingYears | Total years of experience |
| 21 | WorkLifeBalance | Refer to the next table for mapping |
| 22 | YearsAtCompany | No. of years working in the current company |
| 23 | YearsInCurrentRole | No. of years working in the current designation |
| 24 | YearsSinceLastPromotion | No. of years since last promotion was given to the employee |
| 25 | YearsWithCurrManager | No. of years the employee working with his/her current manager |

**Mapping Table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Column Name** | **Label name** |  | **Column Name** | **Label name** |
| Education | 1 'Below College' |  | RelationshipSatisfaction | 1 'Low' |
| 2 'College' |  | 2 'Medium' |
| 3 'Bachelor' |  | 3 'High' |
| 4 'Master' |  | 4 'Very High' |
| 5 'Doctor' |  |  |  |
|  |  |  | WorkLifeBalance | 1 'Bad' |
| EnvironmentSatisfaction | 1 'Low' |  | 2 'Good' |
| 2 'Medium' |  | 3 'Better' |
| 3 'High' |  | 4 'Best' |
| 4 'Very High' |  |  |  |
|  |  |  |  |  |
| JobInvolvement | 1 'Low' |  |  |  |
| 2 'Medium' |  |  |  |
| 3 'High' |  |  |  |
| 4 'Very High' |  |  |  |
|  |  |  |  |  |
| JobSatisfaction | 1 'Low' |  |  |  |
| 2 'Medium' |  |  |  |
| 3 'High' |  |  |  |
| 4 'Very High' |  |  |  |
|  |  |  |  |  |
| PerformanceRating | 1 'Low' |  |  |  |
| 2 'Good' |  |  |  |
| 3 'Excellent' |  |  |  |
| 4 'Outstanding' |  |  |  |

**Task: Few Question which I can ask**

1. What proportion of our staff are leaving?
2. What role does gender and age play in this workplace?
3. How much does income matter?
4. Where is attrition occurring?
5. How does Relationship Satisfaction affect termination?
6. Which Education field is more likely to leave?
7. What factors contribute to retaining high performing employees?
8. What, if anything, else contributes to it?

**Task: Answer the following question based on Univariate Analysis**

1. What is the average age of the employees in each department excluding the outliers?
2. Which education field has the highest attrition rate for the employees with less than 5 years of work experience and monthly salary between 2000 and 4000 USD.
3. What is the average salary hike for the employees with work experience between 8 years to 10 years
4. Is the “number of companies worked” normally distributed for the employees with work experience between 8 years and 10 years.
5. Divided the whole data into two groups – Group1: people with monthly salary more than the average salary of all employees and people with monthly salary less than the average salary of all employees. For each group analysis the attrition rate in each Job Role. Find the Job role which has the highest attrition rate gap across the group.

Please create a R Mark Down file with all the steps required for performing the above steps and generate an html