**HR Analysis Project – Power BI**

**Business Understanding:**

* The purpose is to understand salary differences, termination reasons, and Performance trends in each department by creating visuals to help Managers and HR department gain insights to help managers on making decision that will help in employee retention and manage its recourses.

**Business Questions:**

1. What is the difference in salary between Positions, and genders in the organization?
2. What are the common reasons of termination in each department?
3. How has the performance score of each department changed over time?

**Data Understanding:**

* Dataset File: HRDataset\_v14.csv
* Dataset URL: <https://www.kaggle.com/datasets/rhuebner/human-resources-data-set/data>
* HR Dataset contain: 36 Columns & 311 Rows “Excluding Header”.

**Data Types:**

|  |  |  |
| --- | --- | --- |
| Employee\_Name (Text) | EmpID (Whole Number) | MarriedID (Whole Number) |
| MaritalStatusID (Whole Number) | GenderID (Whole Number) | EmpStatusID (Whole Number) |
| DeptID (Whole Number) | PerfScoreID (Whole Number) | FromDiversityJobFairID (Whole Number) |
| Salary (Whole Number) | Termd (Whole Number) | PositionID (Whole Number) |
| Position (Text) | State (Text) | Zip (Whole Number) |
| DOB (Text) | Sex (Text) | MaritalDesc (Text) |
| CitizenDesc (Text) | HispanicLatino (Text) | RaceDesc (Text) |
| DateofHire (Date) | DateofTermination (Date) | TermReason (Text) |
| EmploymentStatus (Text) | Department (Text) | ManagerName (Text) |
| ManagerID (Whole Number) | RecruitmentSource (Text) | PerformanceScore (Text) |
| EngagementSurvey (Decimal Number) | EmpSatisfaction (Whole Number) | SpecialProjectsCount (Whole Number) |
| LastPerformanceReview\_Date (Date) | DaysLateLast30 (Whole Number) | Absences (Whole Number) |

**Data Preparation:**

* Remove columns that I won’t be using: “MarriedID”, “MaritalStatusID”, “GenderID”, “EmpStatusID”, “DeptID”, “FromDiversityJobFairID”, “Termd”, “PositionID”, “HispanicLatino”, “RaceDesc”, “DateofTermination”, “ManagerID”.
* Splitting column “Employee\_Name” into two columns to separate full name into “FirstName”, and “LastName”, then rename columns and reorder the new columns.
* Remove Duplicates in all columns.
* Change the data type of “Zip” column from Whole Number to Text.
* Change the data type of “DOB” column from Text to Date data type, but an error occurred, therefore I created a column from examples/selections to make sure the year is 4 digits instead of 2 to capture year’s dates without an error, then removed the original “DOB” column and renamed the new custom column to DOB and then reorder it, and changed its type to date data type.
* Add custom column named “Age” to capture the date without the time and calculate the duration between current date and employee’s DOB divided by the number of days within a year 365 to get their current Age in years, then reorder the new column.
* Formula used: Age = Duration.Days(Date.From(DateTime.LocalNow()) - Date.From([DOB])) / 365.
* Add two new Measures to calculate the total salary by genders (Male & Female).
* Formula used: TotalSalaryM = CALCULATE(SUM('HRDataset\_v14'[Salary]),'HRDataset\_v14'[Sex] = "M")
* Formula used: TotalSalaryF = CALCULATE(SUM('HRDataset\_v14'[Salary]),'HRDataset\_v14'[Sex] = "F")

**Modeling:**

* Facts: “Salary”, “Termd”, “PerfScoreID”, “EngagementSurvey”, “EmpSatisfaction”, “SpecialProjectsCount”, “DaysLateLast30”, “Absences”.
* Dimensions: “Employee\_Name”, “EmpID”, “MarriedID”, “MaritalStatusID”, “GenderID”, “EmpStatusID”, “DeptID”, “PerformanceScore”, “FromDiversityJobFairID”, “PositionID”, “Position”, “State”, “Zip”, “DOB”, “Sex”, “MaritalDesc”, “CitizenDesc”, “HispanicLatino”, “RaceDesc”, “DateofHire”, “DateofTermination”, “TermReason”, “EmploymentStatus”, “Department”, “ManagerName”, “ManagerID”, “RecruitmentSource”, “LastPerformanceReview\_Date”.
* Created two folders Facts, and Dimensions to reorganize the columns.

**Evaluation:**

1. What is the difference in salary between Positions, and genders in the organization?

* Provided a dashboard to identify gaps in salary, and found out that most of the Departments have a similar salary except for IT Department noticed Network Engineer, Database Administrator, and Sales Manager have a bigger gap so management should look into it.

1. What are the common reasons of termination in each department?

* Created a visual to provide the most common reasons of termination reasons and I am able to identify that the employees are more likely to leave for the following reasons: Another position, Unhappy, More money, Career change, and hours.

1. How has the performance score of each department changed over time?

* Admin Offices Department The average performance score didn’t change and been consistent for the past 8 years.
* IT Department had better score in 2010, 2011 but overall, they still Fully meet the requirements.
* Production Department Performance score is not consistent and in 2018 had their lowest performance score, so they need attention from the management.
* Sales Department started in 2006 with a perfect score, even though their score is not that bad its still not consistent they still need management attention.
* Software Engineering Department is consistent and their score have been on a positive on 2014.