LEVERAGE DATA-DRIVEN INSIGHTS

EMPLOYEES ATTRITION

Empower Businesses to make Data-Driven Decisions

By
Ali Sherif

TABLE OF CONTENTS

→ DATA UNDERSTANDING

GET TO KNOW ABOUT DATA

HR Attrition Dataset Overview

This dataset contains information about employees within a company, focusing on various factors that may influence employee attrition (i.e., whether an employee leaves the company). The data includes both numerical and categorical variables that describe employee demographics, job characteristics, and satisfaction levels. By analysing these columns, we can identify trends and patterns that contribute to employee turnover and retention.

→ COLUMNS CHARACTERISTICS

DEFINING THE COLUMNS AND DATA TYPES

Overall Column Overview

The dataset contains a comprehensive set of features related to employees within the company, focusing on various factors that can influence employee attrition and overall job satisfaction. It includes demographic information, job characteristics, compensation details, work environment and conditions, employee experience, and work-life balance factors. By analyzing these columns, we aim to identify patterns that help in predicting employee turnover, understanding satisfaction levels, and improving retention strategies.

→ COLUMNS CHARACTERISTICS

DEFINING THE COLUMNS AND DATA TYPES

- Age: Represents the employee's age in years, which can influence career stage and job satisfaction.
- Attrition: Indicates whether the employee left the company (Yes) or stayed (No).
- BusinessTravel: Describes the frequency of business travel required for the employee's role.
- DailyRate: The daily rate of pay for the employee, affecting job satisfaction and retention.
- Department: The department in which the employee works, reflecting their role within the organization.
- DistanceFromHome: The distance (in miles) between the employee's home and their workplace, affecting commute time and work-life balance.
- Education: Represents the employee's level of education, which can correlate with job performance and career advancement.
- EducationField: The field of study in which the employee's education was completed, influencing job fit and satisfaction.
- EmployeeCount: Represents the number of employees in the dataset (usually fixed to 1).
- EmployeeNumber: A unique identifier for each employee, used for tracking and data analysis.
- EnvironmentSatisfaction: Measures the employee's satisfaction with the work environment (1-4 scale).
- Gender: The gender of the employee, which may influence job dynamics and diversity factors.
- HourlyRate: The employee's hourly rate of pay, influencing job satisfaction and retention.
- JobInvolvement: The level of employee involvement in their job, which impacts engagement and performance.
- JobLevel: The hierarchical level of the employee within the company (1-5 scale).
- JobRole: The specific role or position held by the employee (e.g., Sales Executive, Research Scientist).
- JobSatisfaction: The employee's self-reported satisfaction with their job (1-4 scale).
- MaritalStatus: The marital status of the employee, which can influence work-life balance and retention.

→ COLUMNS CHARACTERISTICS

DEFINING THE COLUMNS AND DATA TYPES

- MonthlyIncome: The employee's monthly salary, which impacts job satisfaction and retention.
- MonthlyRate: A value representing the employee's monthly rate of pay, contributing to overall compensation.
- NumCompaniesWorked: The number of companies the employee has previously worked for, indicating job stability.
- Over18: Indicates whether the employee is over 18 years of age, generally fixed to "Yes".
- OverTime: Whether the employee works overtime (Yes/No), impacting work-life balance and job satisfaction.
- PercentSalaryHike: The percentage increase in the employee's salary, reflecting growth opportunities.
- PerformanceRating: The employee's performance rating, reflecting their job performance and potential for advancement.
- RelationshipSatisfaction: The employee's satisfaction with their relationships at work (1-4 scale).
- StandardHours: The standard number of hours the employee is expected to work per week (usually fixed to 80).
- StockOptionLevel: The level of stock options granted to the employee, indicating long-term incentives.
- TotalWorkingYears: The total number of years the employee has worked, influencing career stability and experience.
- TrainingTimesLastYear: The number of training sessions the employee attended last year, reflecting opportunities for growth.
- YearsAtCompany: The number of years the employee has worked at the current company, indicating loyalty and tenure.
- YearsInCurrentRole: The number of years the employee has spent in their current role, influencing satisfaction and career progression.
- YearsSinceLastPromotion: The number of years since the employee's last promotion, potentially impacting job satisfaction.
- YearsWithCurrManager: The number of years the employee has worked with their current manager, influencing job satisfaction and retention.

→ DATA MODELLING

BUILDING STAR SCHEMA

A star schema is a type of data modelling technique commonly used in data warehousing and business intelligence. It consists of a central fact table surrounded by related dimension tables.

- Fact Table: Contains quantitative data
- Dimension Tables: Contain descriptive attributes or characteristics

- This data model is designed to analyze employee attrition, focusing on various factors that influence it. It uses a star schema, with a central Fact Table and multiple Dimension Tables:

- Fact_HR_Employee_Attrition: The central table containing key metrics related to employee attrition, such as attrition rate, average salary, and tenure.

- Dimension Tables:
  Dim_Age: Contains age groups and corresponding ages.
  Dim_Distance: Categorizes distance information.
  Dim_MaritalStatus: Describes marital statuses.
  Dim_Department: Lists organizational departments.
  Dim_CareerExperience: Represents different levels of career experience.
  Dim_JobRole: Describes various job roles within the organization.
  Dim_BusinessTravel: Represents business travel frequencies.

- Relationships: These tables are connected via foreign keys (e.g., "Age FK," "Department ID FK," and "JobRole FK") to enable detailed analysis, such as attrition by age, department, job role, marital status, experience level, and salary.

- The purpose of this model is to help analysts identify patterns and factors that contribute to employee attrition by examining the relationships between these dimensions and key attrition metrics.

→ DATA MODELLING

BUILDING STAR SCHEMA

A star schema is a type of data modelling technique commonly used in data warehousing and business intelligence. It consists of a central fact table surrounded by related dimension tables.

- Fact Table: Contains quantitative data
- Dimension Tables: Contain descriptive attributes or characteristics

The data model I built incorporates several custom attributes to enhance analysis and improve performance:

- Custom Career Experience Levels:

For the Total Working Years column, I created custom categories based on the number of years, including "Entry-Level," "Junior," "Mid-Level," "Senior," "Expert," and "Veteran." This allows for a more granular analysis of employee experience.

- Custom Age Categories:

The Age column was divided into categories like "Young," "Youth," "Mature," "Senior," and "Seasoned." This helps in understanding how attrition varies across different age groups.

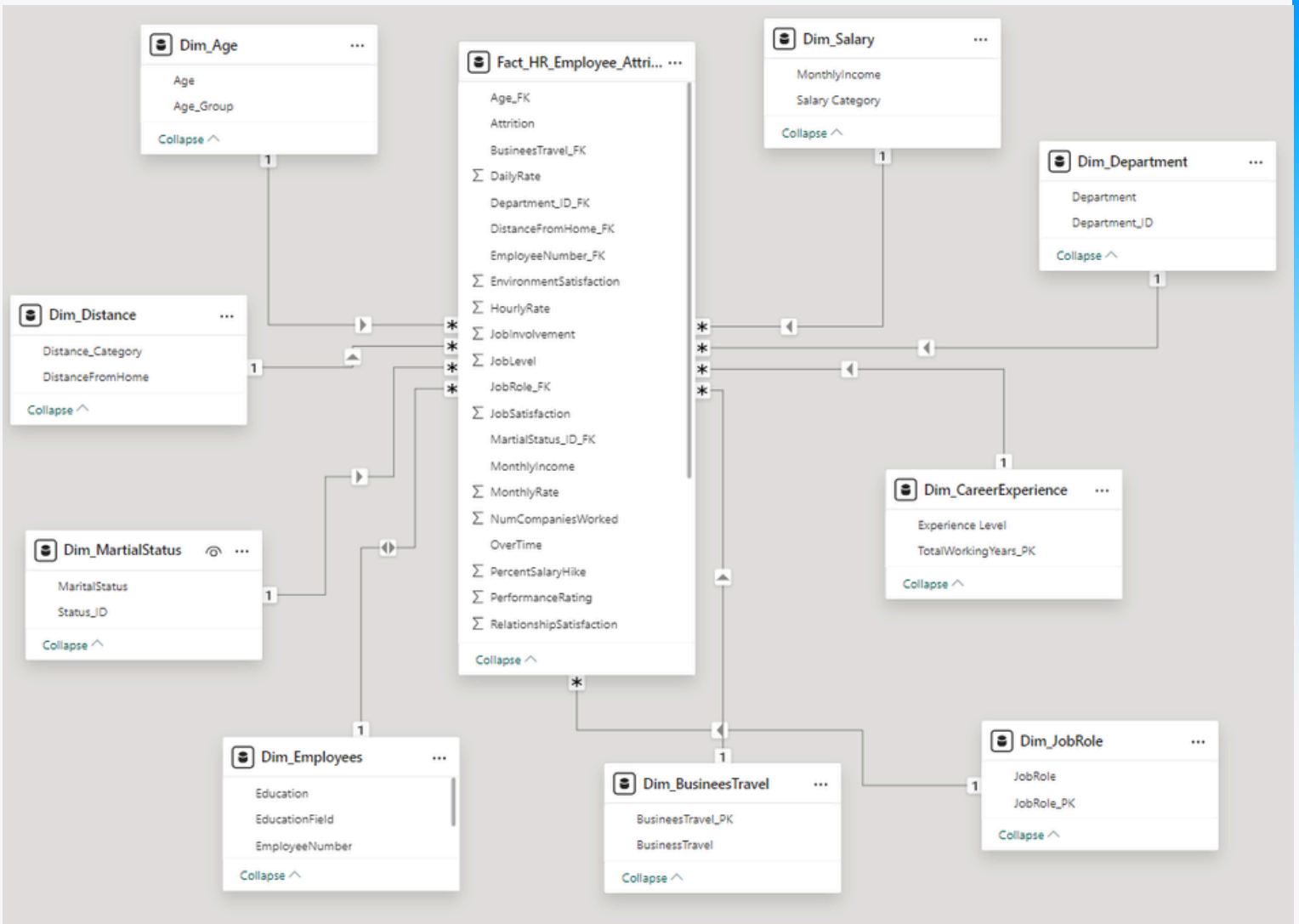
- Indexing for Faster Joins:

To optimize performance, I avoided using strings for the relationships between the fact and dimension tables. Instead, I created custom integer indexes for the primary and foreign key relationships. This method enhances the speed of joins and ensures more efficient querying.

By using these custom attributes and indexing strategies, the model ensures more efficient data processing and better categorization, which leads to faster and more insightful analysis of employee attrition.

→ DATA MODELLING

BUILDING STAR SCHEMA



→ CREATING MEASURES

MEASURES USING VARIABLES

The measures are designed to calculate specific metrics based on the data in the fact and dimension tables.
Here's a brief overview of how these measures contribute to the dashboard:

- Total Attrition:

This calculates the total number of employees who left the company.

```
Total Attrition =
VAR Total_Attrition = CALCULATE(DISTINCTCOUNT(Fact_HR_Employee_Attrition
  [EmployeeNumber_FK]), Fact_HR_Employee_Attrition[Attrition] = "YES")
RETURN Total_Attrition
```

- Avg. Salary Hike (Attrition):

This calculates the average salary hike for employees who left the company.

```
Avg. Salary Hike (Attrition) =
var AVG_SAL_HIKE = AVERAGE(Fact_HR_Employee_Attrition[PercentSalaryHike])
var AVG_SAL_HIKE_ATTR = CALCULATE(AVG_SAL_HIKE, Fact_HR_Employee_Attrition
  [Attrition] = "YES")/100
RETURN AVG_SAL_HIKE_ATTR
```

- Attrition Performance:

This measures the overall rate of employee attrition.

```
Attrition Performance =
var Total_employees = DISTINCTCOUNTNOBLANK(Fact_HR_Employee_Attrition
  [EmployeeNumber_FK])
var Attrition_Percentage = DIVIDE([Total Attrition],Total_employees)
RETURN Attrition_Percentage
```

→ CREATING MEASURES

MEASURES USING VARIABLES

- Attrition with Overtime (%):

This calculates the percentage of employees who left the company while working overtime.

```
Attrition with Overtime (%) =
var Total_EMPS = DISTINCTCOUNTNOBLANK(Fact_HR_Employee_Attrition[EmployeeNumber_FK])
var TOTAL_EMP_ATTRITON = CALCULATE(Total_EMPS, Fact_HR_Employee_Attrition[Attrition]
="YES")
var TOTal_Attrition_Work_OverTime = CALCULATE(Total_EMPS,Fact_HR_Employee_Attrition
[OverTime]="YES",Fact_HR_Employee_Attrition[Attrition]="YES")
var ATTRITION_OVERTIME_PERCENTAGE = DIVIDE(TOTal_Attrition_Work_OverTime,
TOTAL_EMP_ATTRITON)
RETURN ATTRITION_OVERTIME_PERCENTAGE
```

- Average Environment Satisfaction (Attrition):

This measures the average level of satisfaction with the work environment among employees who left the company.

```
Average Environment Satisfaction (Attrition) =
var AVG_SATSFICATION = AVERAGE(Fact_HR_Employee_Attrition[EnvironmentSatisfaction])
var AVG_FOR_ATTRITION = CALCULATE(AVG_SATSFICATION, Fact_HR_Employee_Attrition
[Attrition]="YES")
RETURN AVG_FOR_ATTRITION
```

- Average Job Satisfaction (Attrition):

This measures the average level of job satisfaction among employees who left the company.

```
Average Job Satisfaction (Attrition) =
var AVG_SATSFICATION = AVERAGE(Fact_HR_Employee_Attrition[JobSatisfaction])
var AVG_FOR_ATTRITION = CALCULATE(AVG_SATSFICATION, Fact_HR_Employee_Attrition
[Attrition]="YES")
RETURN AVG_FOR_ATTRITION
```

→ CREATING MEASURES

MEASURES USING VARIABLES

- Average Tenure:

This calculates the average length of employment for Attrition employees.

```
Average Tenure =
var Average_Years_At_Company = AVERAGE(Fact_HR_Employee_Attrition[YearsAtCompany])
var Average_Tenure_Left = CALCULATE(Average_Years_At_Company,
  Fact_HR_Employee_Attrition[Attrition]="YES")
RETURN Average_Tenure_Left
```

- Average Years With Manager:

This measures the average number of years employees have worked with their current manager.

```
Average Years With Manager =
var avg_years_with_manager = AVERAGE(Fact_HR_Employee_Attrition
  [YearsWithCurrManager])
var AVG_FOR_ATTRITION = CALCULATE(avg_years_with_manager, Fact_HR_Employee_Attrition
  [Attrition]="YES")
RETURN AVG_FOR_ATTRITION
```

- Promotion Gap (Years):

This measures the average number of years between promotions for employees.

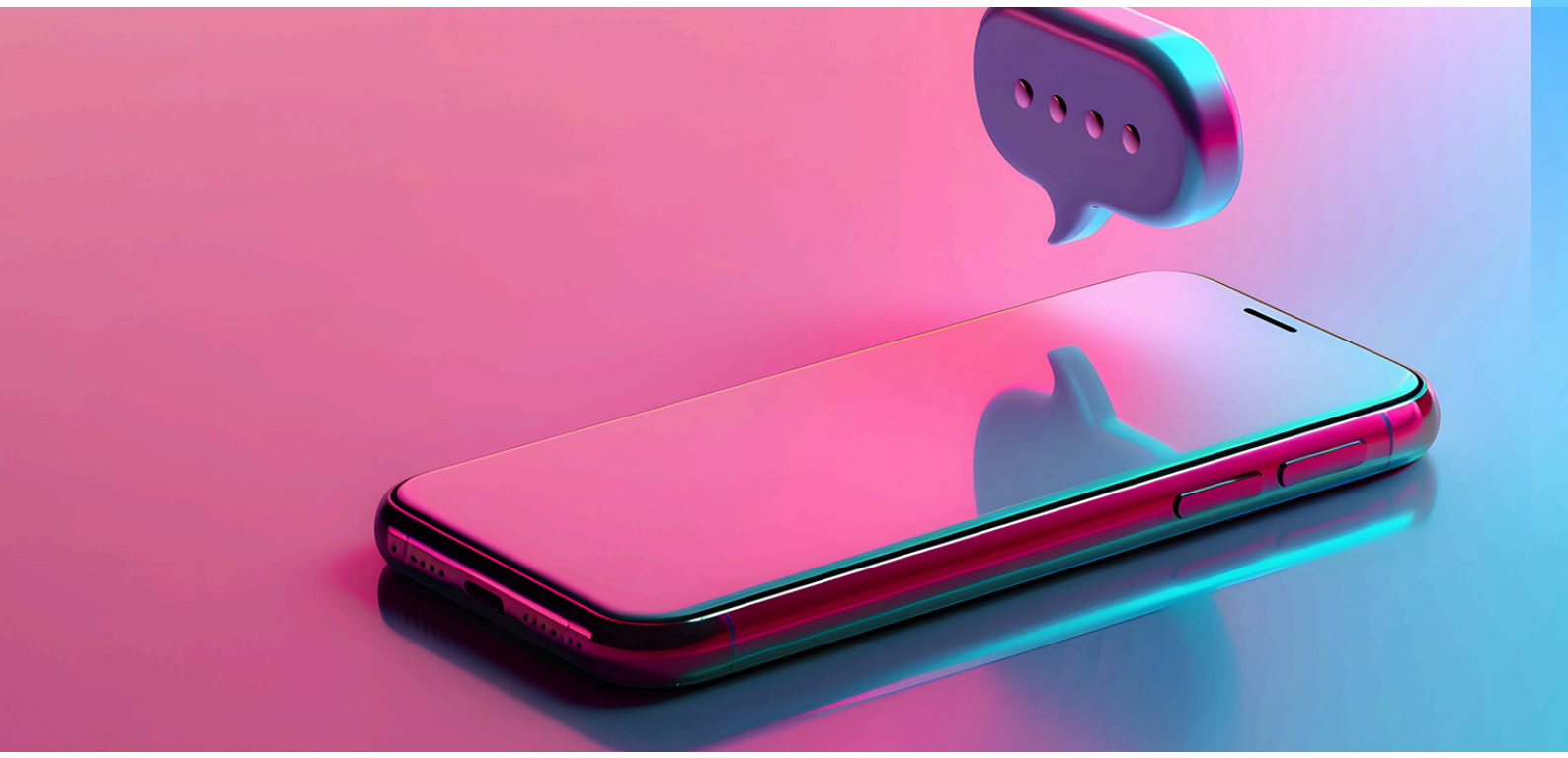
```
Promotion Gap (Years) =
var AVG_Years_Since_Last_Prom =AVERAGE(Fact_HR_Employee_Attrition
  [YearsSinceLastPromotion])
var AVG_YEARS_GAP_For_Attrition = CALCULATE(AVG_Years_Since_Last_Prom,
  Fact_HR_Employee_Attrition[Attrition]="YES")
RETURN AVG_YEARS_GAP_For_Attrition
```

→ DASHBOARD BUILDING

BUILDING DYNAMIC INTERACTIVE DASHBOARD

The Power BI dashboard is designed to analyze employee attrition by leveraging custom measures and interactive features. Key measures such as Total Attrition Rate, Attrition Performance, and Promotion Gap are calculated using DAX and displayed in various visualizations like bar charts and KPIs. Interactive elements, including filters and slicers, allow users to drill down into specific attributes like Age, Experience Level, Department, and Salary, enabling deeper insights into attrition patterns. The dashboard is designed for easy navigation, with features like drillthrough for detailed analysis and conditional formatting to highlight key metrics, helping HR teams identify trends and make data-driven decisions.

STAY IN TOUCH WITH ME!



Linkedin
linkedin.com/in/alisherif98/

Phone
+201090231677

E-mail
alisherif560@gmail.com