**Report: Employee Performance and Attrition**

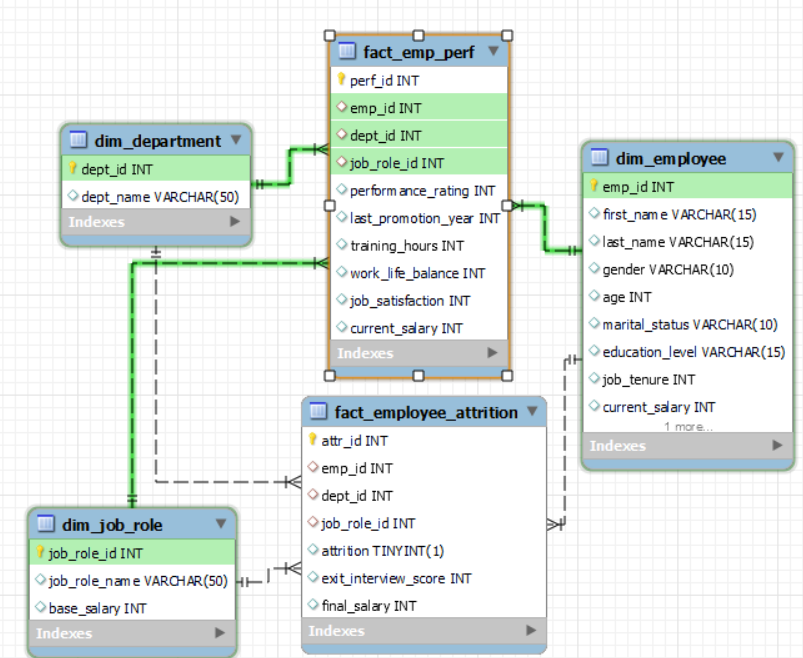
**Problem Statement**:

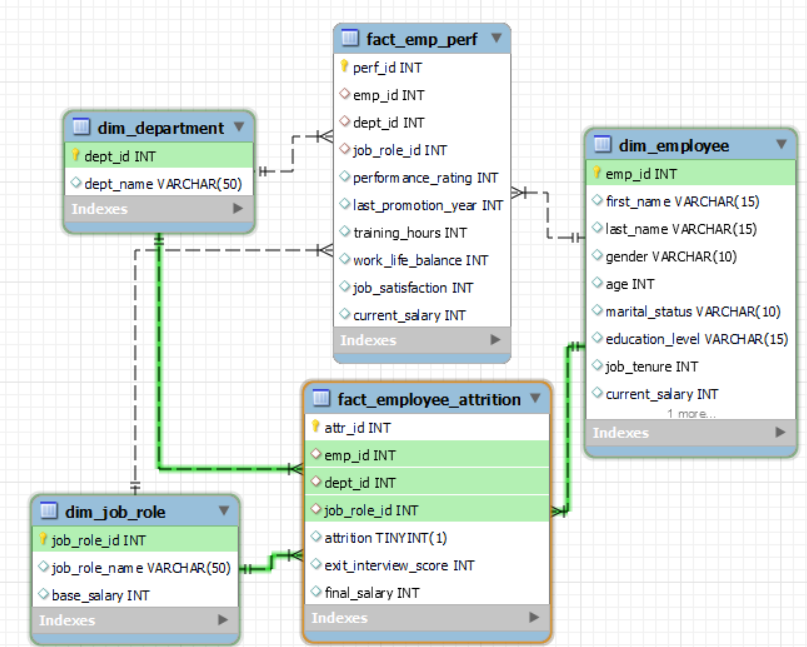
Organizations face significant challenges in managing employee performance and retention. High employee attrition not only leads to increased hiring costs but also results in the loss of institutional knowledge, reduced productivity, and potential disruptions in business operations. Many organizations struggle to identify the root causes of attrition and performance issues due to the lack of comprehensive data analysis.

**Work Flow**

**Database Creation**:

* Created a database ‘emp\_perf’ in MYSQL Workbench.
* **Star Schema Design –** Build a warehouse schema to optimize queries for inventory and supply chain management.
* I have created tables in the database using ‘star schema’ design.
* I created total 5 tables – 3 dimension tables and 2 fact table.

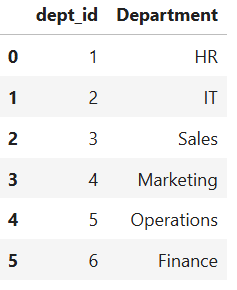




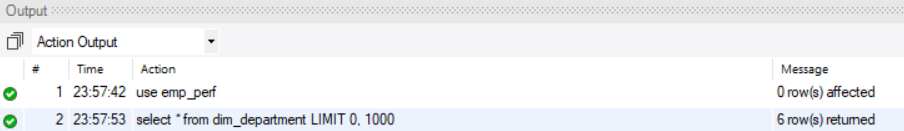
* dim\_department, dim\_employee, dim\_job\_role, fact\_emp\_perf, fact\_employee\_attrition are 5 tables.

**Importing data into Tables using Python**:

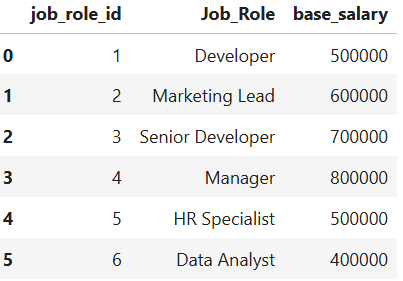
* Now I have connected the python and database tables using mysql connector library.
* The cursor is used to execute SQL queries.
* **department** data is residing in ‘employee\_data 1.csv’.
* loading data into department DataFrame.
* I found unique department values and mapped department id.

 Findings: there are total 6 departments.

* Converted the DataFrame into a list of tuples for database insertion.
* Executed an SQL INSERT query to insert the data into the dim\_department table.
* Checked the dim\_department table in the MySQL Workbench all the records have been inserted.



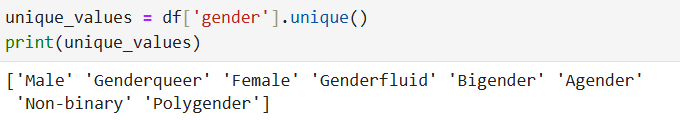
* Findings: All the 6 records are imported.
* **Job role data** is residing in ‘employee\_data 1.csv’.
* Loading Job role data into Job role DataFrame.
* I found unique Job role values and mapped Job role id.
* Note: I created dummy data - base salary for each job role.

 Findings: there are total 6 job roles.

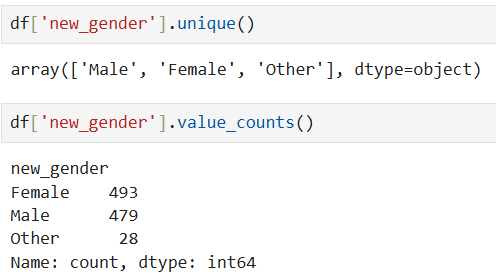
* Converted the DataFrame into a list of tuples for database insertion.
* Executed an SQL INSERT query to insert the data into the dim\_job\_role table.
* Checked the dim\_job\_role table in the MySQL Workbench all the records have been inserted.



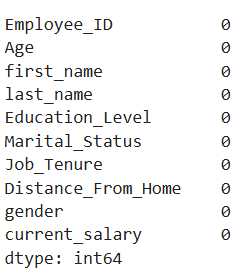
* Findings: All the 6 records are imported.
* **Employee data** is residing in ‘employee\_data 1.csv’.
* Loading data into employee DataFrame.
* There are 2 columns for genders ['gender','Gender'].
* I found unique gender column values



* Findings: there are many genders.
* I considered only ‘male’ and ‘female’ values from ‘gender’ column and for other values I considered ‘Gender’ column and created a new gender column.



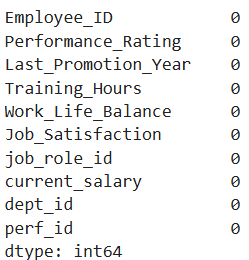
* Findings: there are 3 genders and most employees are female.
* I dropped the original two gender columns.
* calculated current salary based on base salary and job tenure considering yearly 20% growth. By merging job role DataFrame and employee DataFrame.
* Droped unnecessary columns.
* Checked for null values and duplicates in the DataFrame.

 Findings: There are no null and duplicate values.

* Converted the DataFrame into a list of tuples for database insertion.
* Executed an SQL INSERT query to insert the data into the dim\_employee table.
* Checked the dim\_employee table in the MySQL Workbench all the records have been inserted.



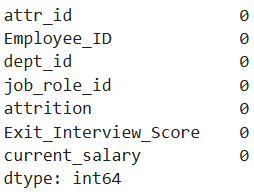
* Findings: All the 1000 records are imported.
* **Employee Performance** data is residing in ‘employee\_performance\_data 1.csv’.
* Loading data into performance DataFrame.
* Merged performance DataFrame and employee DataFrame for emp id, job role id and current salary.
* Droped unnecessary columns.
* Merged performance DataFrame and department DataFrame for dept id.
* Droped unnecessary columns.
* Mapped perf id for each record.
* Checked for null values and duplicates in the DataFrame.

 Findings: There are no null and duplicate values.

* Converted the DataFrame into a list of tuples for database insertion.
* Executed an SQL INSERT query to insert the data into the fact\_emp\_perf table.
* Checked the fact\_emp\_perf table in the MySQL Workbench all the records have been inserted.



* Findings: All the 1000 records are imported.
* **Attrition data** is residing in ‘Attrition 1.csv’.
* Loading data into attrition DataFrame.
* Merged attrition DataFrame and employee DataFrame for emp id, job role id and current salary.
* Droped unnecessary columns.
* Merged attrition DataFrame and department DataFrame for dept id.
* Droped unnecessary columns.
* Mapped attr id for each record.
* Checked for null values and duplicates in the DataFrame.

 Findings: There are no null and 2 duplicate values.

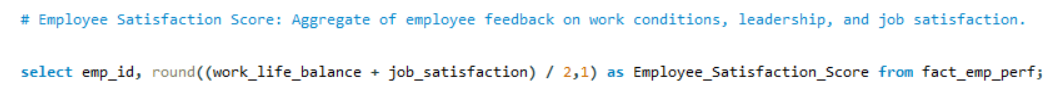
* Dropped the duplicate records.
* Converted the DataFrame into a list of tuples for database insertion.
* Executed an SQL INSERT query to insert the data into the fact\_ employee\_attrition table.
* Checked the fact\_employee\_attrition table in the MySQL Workbench all the records have been inserted.



* Findings: All the 142 records are imported.

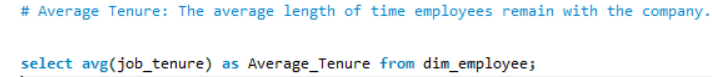
**SQL Tasks:**

1. **Employee Satisfaction Score**: Aggregate of employee feedback on work conditions, leadership, and job satisfaction.



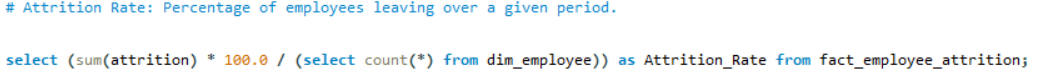
* I calculated Employee Satisfaction Score based on work life balance and job satisfaction.

1. **Average Tenure**: The average length of time employees remain with the company.



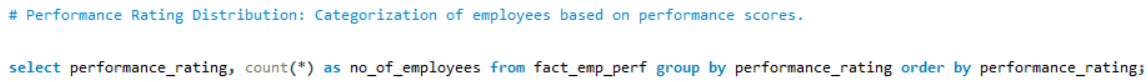
 Findings: average tenure is 10 years.

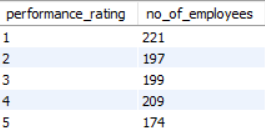
1. **Attrition Rate:** Percentage of employees leaving over a given period.



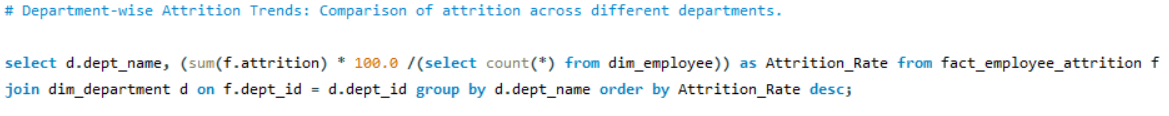
 Findings: attrition rate is 7%.

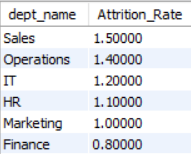
1. **Performance Rating Distribution**: Categorization of employees based on performance scores.



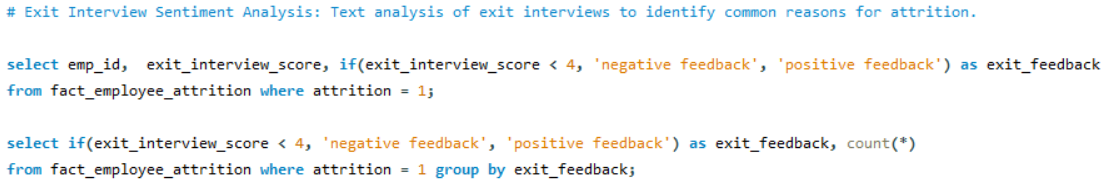
 Findings: there are many employees whose performance ratings is 1and very few employees with 5 rating.

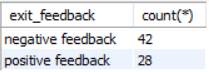
1. **Department-wise Attrition Trends**: Comparison of attrition across different departments.



 Findings: Sales department has highest attrition rate and Finance the lowest.

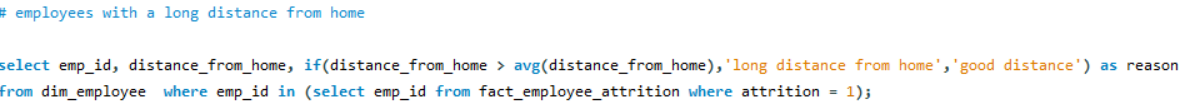
1. **Exit Interview Sentiment Analysis**: Text analysis of exit interviews to identify common reasons for attrition.



 Findings: Attrition is more for employees who gave negative feedback.

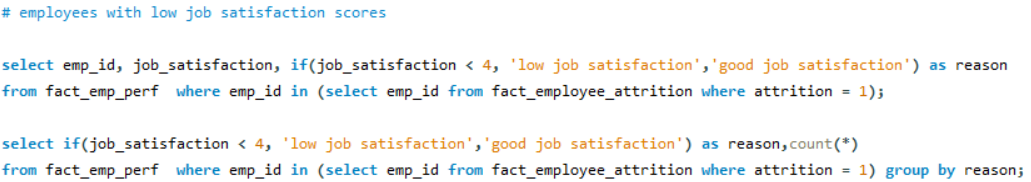
* I considered exit interview score below 4 as negative feedback.

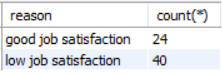
1. **employees with a long distance** from home based on average distance.



 Findings: Only one employee (42 id) attrition can be long distance from home to office.

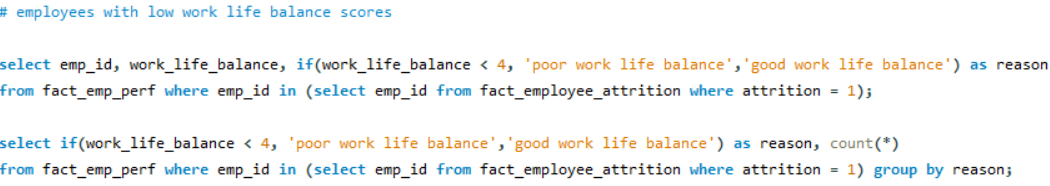
1. **employees with low job satisfaction scores**.

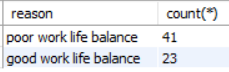


Findings: Can consider job satisfaction scores as a strong reason for attrition as employees with low job satisfaction scores are in attrition.

* I considered job satisfaction score below 4 as low job satisfaction.

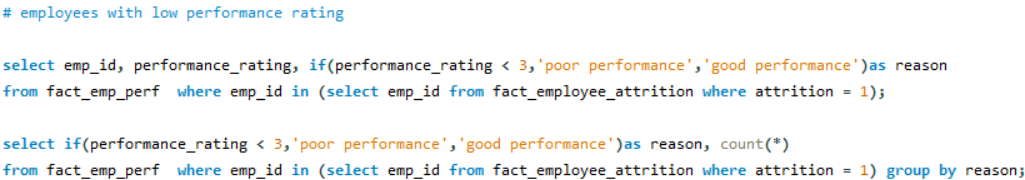
1. **employees with low work life balance scores**

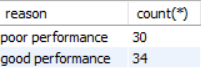


Findings: Can consider work life balance scores as a strong reason for attrition as employees with poor work life balance scores are in attrition.

* I considered work life balance score below 4 as poor work life balance.

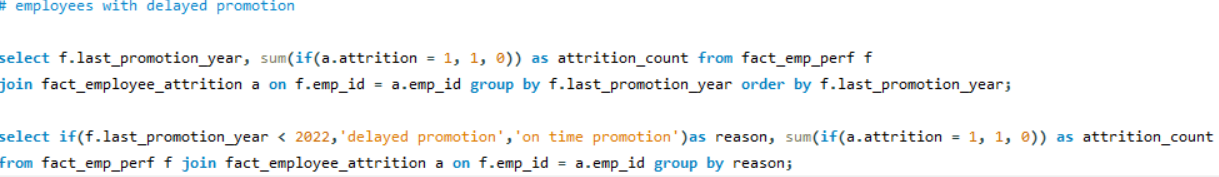
1. **employees with low performance rating**

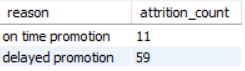


 Findings: Can consider performance rating as a strong reason for attrition as employees with good performance rating are in attrition. They can be expecting more salary or got better offering salary than current salary or not getting promotions.

* I considered performance rating below 3 as poor performance rating.

1. **employees with delayed promotion**

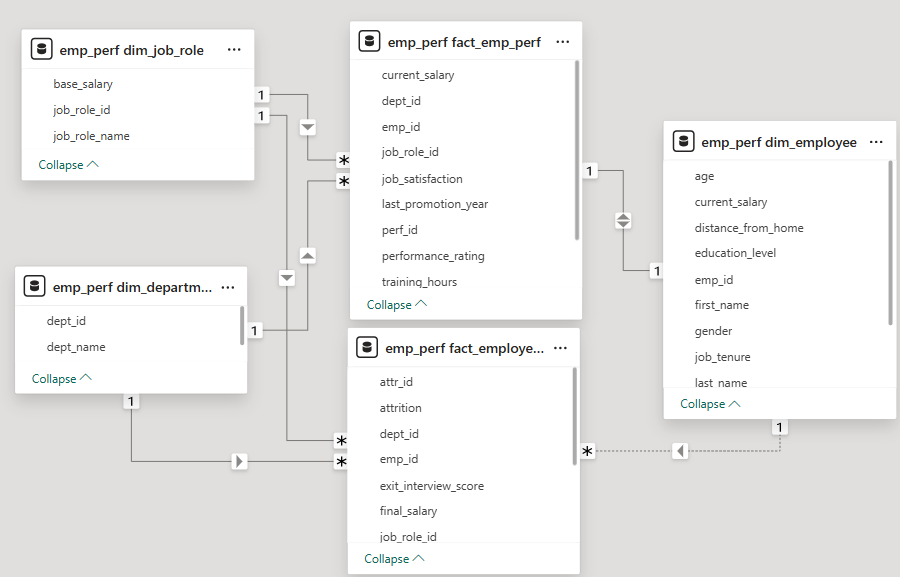


 Findings: Can consider promotion year as a strong reason for attrition as employees with delayed promotion are in attrition. They can be expecting promotion.

* I considered last\_promoted\_year below 2022 as delayed promotion.
* Out of all criteria **last promoted year** makes clear distinguishing among the attrition employees.

**Power BI Reporting and Analytics**

* The 5 Tables in MYSQL Workbench are loaded into power query for transformation.
* The data types are changed for ID columns and item number from number to text as there is no need of performing mathematical calculations.
* After transformation data is loaded into report.
* Goto Model view and check the relationships.



* I added title ‘**Employee Performance and Attrition**’ using text box.
* Created measure ‘**total employees’**. Displayed using card visual under emp perf dim employee table.

 Findings: total number of employees are 1000.

* Created measure ‘**average tenure’**. Displayed using card visual under emp perf dim employee table. formatted with “years” string.

 Findings: Average tenure of employees is 10 years.

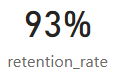
* Created measure ‘**total attrition count’**. Displayed using card visual under emp perf fact employee attrition table.

 Findings: total number of employees left the company are 70.

* Created measure ‘**attrition rate’**. Displayed using card visual under emp perf fact employee attrition table. Calculated by total attrition count/ total employees and format to percentage.

 Findings: attrition rate is 7%.

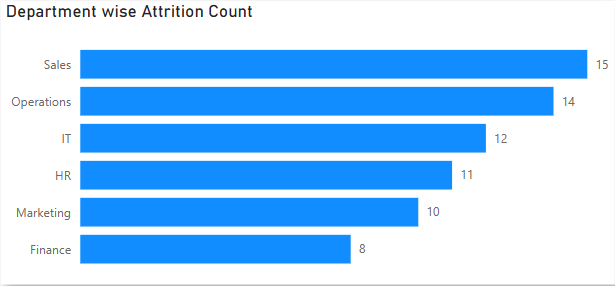
* Created measure ‘**retention rate’**. Displayed using card visual under emp perf fact employee attrition table. Calculated by 1-attrition rate and format to percentage.

 Findings: retention rate is 7%.

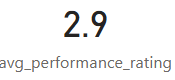
* Created measure ‘**Average Exit Interview Satisfaction Score’**. Displayed using card visual under emp perf fact employee attrition table.

 Findings: Average Exit Interview Satisfaction Score is 3.1 which can be considered okay.

* Created measure ‘dept attrition rate’ for ‘**Department wise Attrition Rate’** under emp perf fact employee attrition table.
* Displayed using clustered bar chart visual.

 Findings: Sales department has highest attrition count and finance has the lowest.

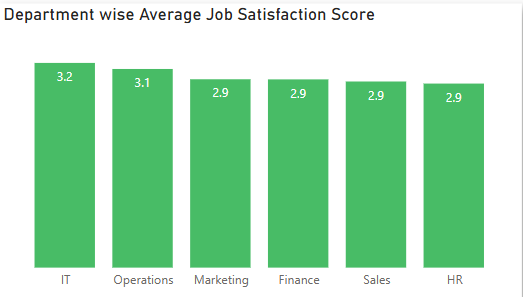
* Created measure ‘**Average Performance Rating’** under emp perf fact emp perf table. Displayed using card visual.

 Findings: Average Performance Rating is 2.9 which is bit low.

**Department-wise Employee Score**

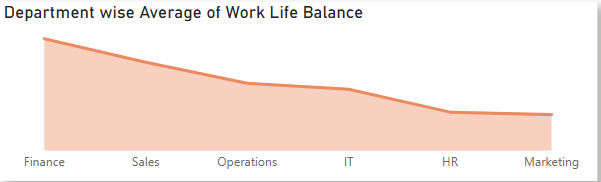
1. Department wise Average Job Satisfaction Score:

* Created measure ‘dept\_avg\_job\_satisfaction’ under emp perf dim department table. Displayed using clustered column chart.

 Findings: IT department has highest Average Job Satisfaction Score and HR has the lowest.

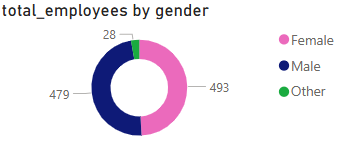
1. Department wise Average of Work Life Balance:

* Plotted department and average of work life balance score using area chart.

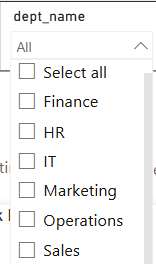


Findings: finance department has highest Average work life balance score and Marketing has the lowest.

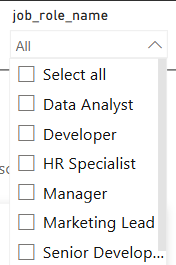
----- Total employees by gender



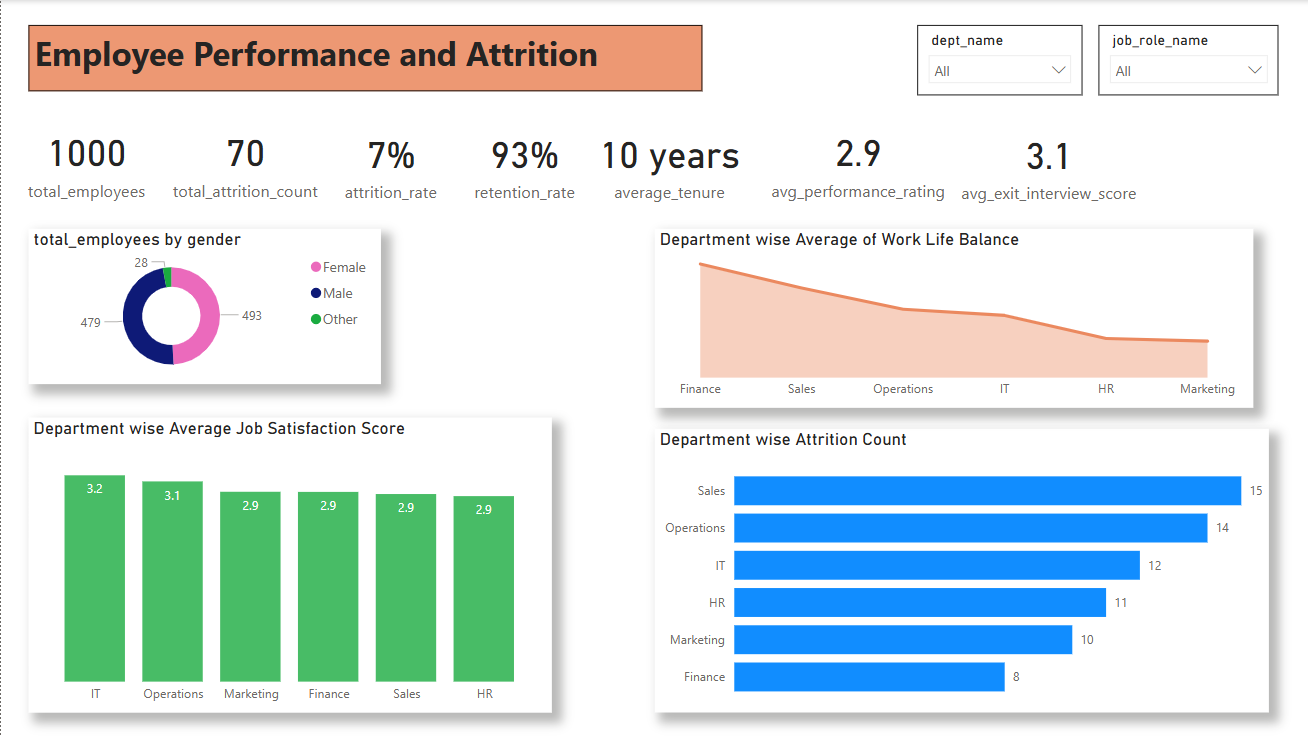
* Findings: there are 3 genders and most employees are female.
* Created slicer for department to filter all the visuals.

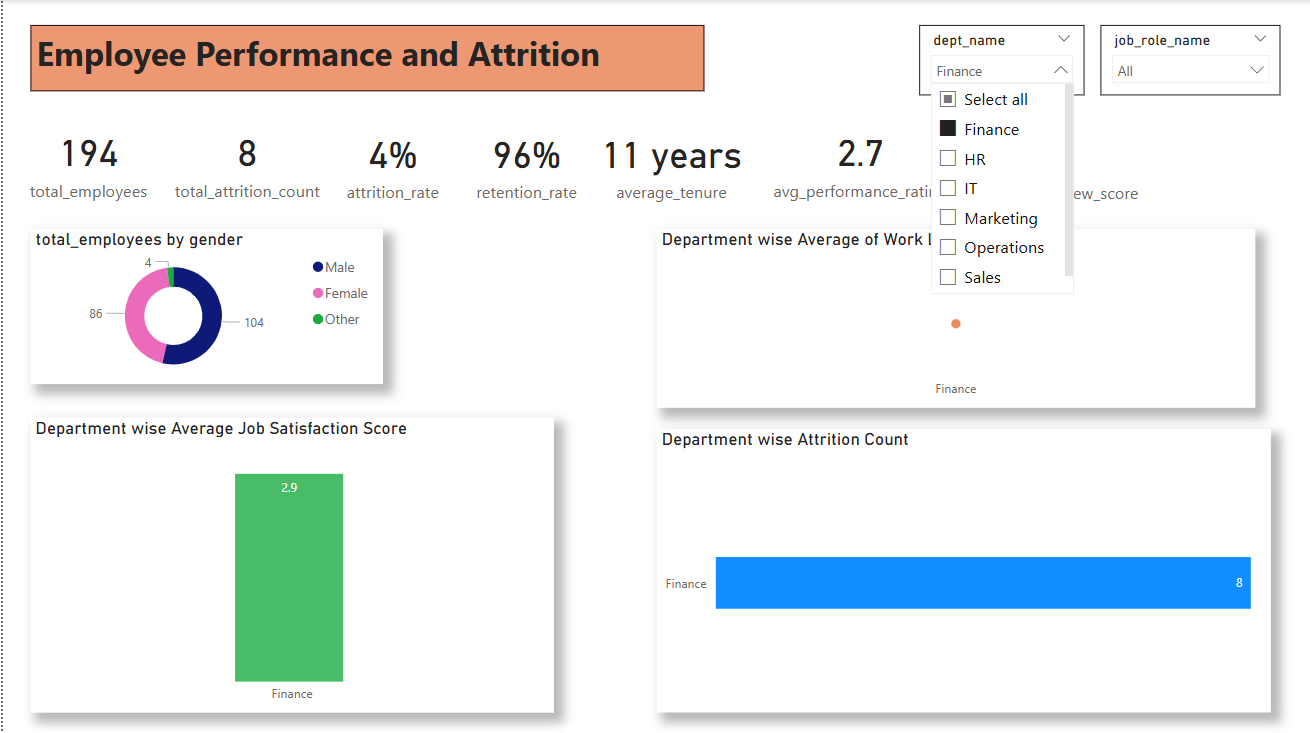


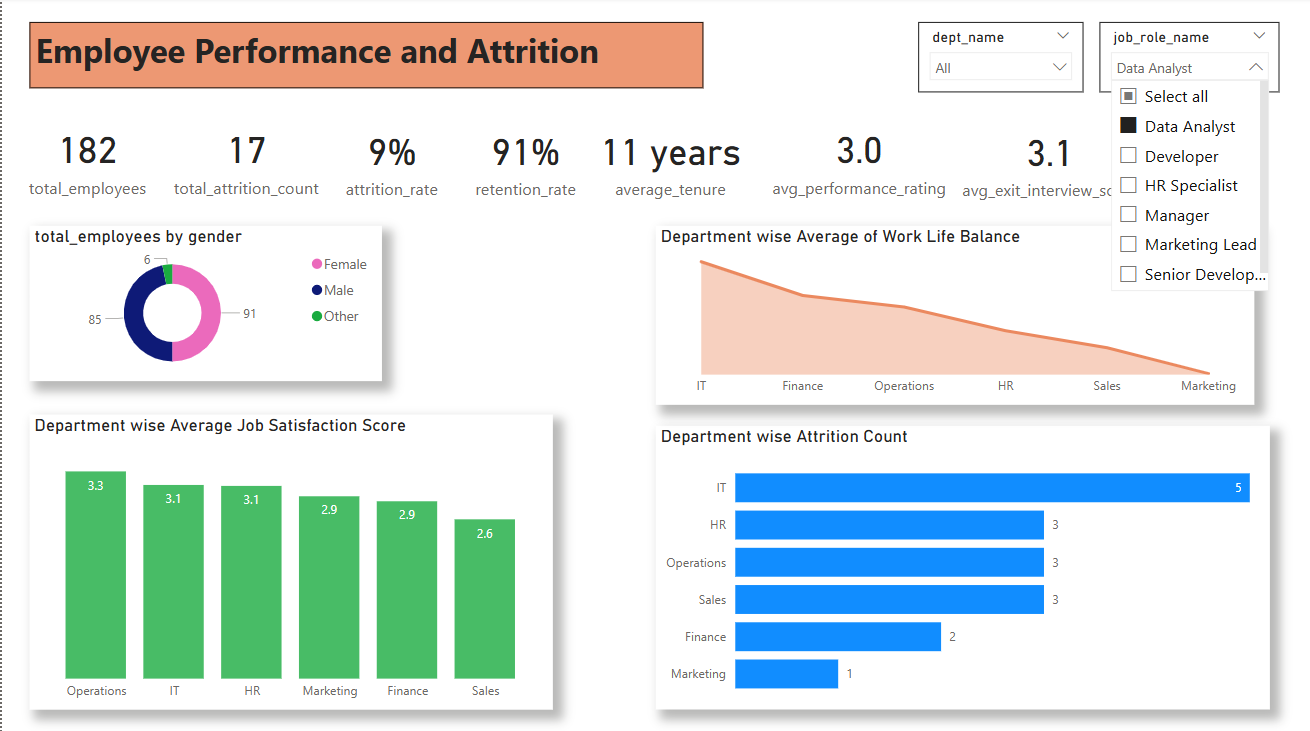
* Created slicer for job role to filter all the visuals.



**Complete Report**







* By improving the work life balance and giving promotion to employees can reduce the attrition.
* All the files are attached in folder.
* A clear understanding of factors affecting employee performance and attrition.
* Actionable insights for HR teams to enhance employee retention.
* A user-friendly dashboard for continuous monitoring and decision-making.