HR ANALYTICS PROJECT

Introduction

This project delves into the realm of **HR attrition**, meticulously examining a dataset **comprising 50,000 employee** records across two comprehensive sheets. My primary objective is to decipher the underlying reasons behind employee departures and devise strategic measures to mitigate attrition rates. Through meticulous data cleaning, merging, analysis, and presentation, we aim to deliver actionable insights to our client, facilitating informed decision-making and fostering a more resilient workforce.

Data

- Age of employee
- Employee ID
- Business Travel
- Department
- Education Field
- Environment Satisfaction
- Gender
- Job Level
- Job Role
- Attrition rate

- Job Satisfaction
- Monthly Income
  - Performance Rating
- Total Working Years
  - Work Life Balance
  - Years In Current Role
- Years Since Last Promotion

Tools Used – MS EXCEL, POWER BI, TABLEAU, MYSQL

MS EXCEL

Power Bi

Tableau

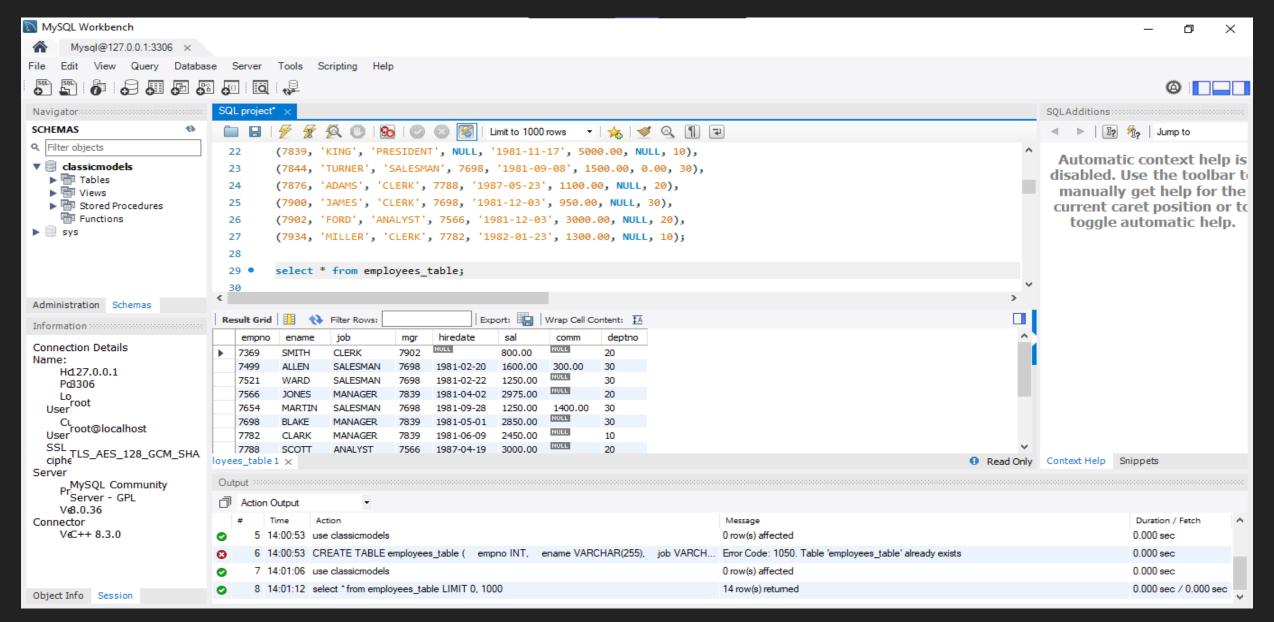| empno	ename	job	mgr	hiredate	sal	comm	deptno
7369	SMITH	CLERK	7902	1890-12-17	800.00	NULL	20
7499	ALLEN	SALESMAN	7698	1981-02-20	1600.00	300.00	30
7521	WARD	SALESMAN	7698	1981-02-22	1250.00	500.00	30
7566	JONES	MANAGER	7839	1981-04-02	2975.00	NULL	20
7654	MARTIN	SALESMAN	7698	1981-09-28	1250.00	1400.00	30
7698	BLAKE	MANAGER	7839	1981-05-01	2850.00	NULL	30
7782	CLARK	MANAGER	7839	1981-06-09	2450.00	NULL	10
7788	SCOTT	ANALYST	7566	1987-04-19	3000.00	NULL	20
7839	KING	PRESIDENT	NULL	1981-11-17	5000.00	NULL	10
7844	TURNER	SALESMAN	7698	1981-09-08	1500.00	0.00	30
7876	ADAMS	CLERK	7788	1987-05-23	1100.00	NULL	20
7900	JAMES	CLERK	7698	1981-12-03	950.00	NULL	30
7902	FORD	ANALYST	7566	1981-12-03	3000.00	NULL	20
7934	MILLER	CLERK	7782	1982-01-23	1300.00	NULL	10
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

My SQL

MS EXCEL

- <u>Preparing KPIs</u>: Define key performance indicators to measure number in precise way.

- <u>Clean the data</u>: Remove errors and inconsistencies to ensure accuracy.

- <u>Use pivot table</u>: Organize and summarize data for analysis.

- <u>Analyze the data</u>: Interpret trends and patterns to draw insights.

- <u>Visualize the data</u>: Present findings in a clear and compelling way.

POWER BI

- <u>Importing the data from Excel</u>: Power BI allows you to seamlessly import data from Excel spreadsheets, enabling you to work with your existing data easily.
- <u>Cleaning the data</u>: With Power BI, you can clean your data by removing duplicates, correcting errors, and transforming data into a usable format for analysis.
- <u>Analyzing the data:</u> Power BI provides robust tools for data analysis, including the ability to create relationships between datasets, perform complex calculations, and derive insights from your data.
- <u>Creating calculations and DAX</u>: Power BI's Data Analysis Expressions (DAX) language allows you to create custom calculations and measures, enabling you to perform advanced analysis on your data.
- <u>Visualizing the data with attractive charts</u>: Power BI offers a wide range of visualization options, including bar charts, line charts, and pie charts, allowing you to create visually appealing and informative reports.

Tableau

- <u>Importing Data from Excel</u>: Tableau allows users to seamlessly import data from Excel spreadsheets, simplifying the process of integrating external data sources into your analysis.
- <u>Analyzing Data</u>: With Tableau, users can perform in-depth analysis on their data, uncovering insights and trends that can inform decision-making and drive business strategies.
- <u>Creating Calculated Fields</u>: Tableau provides the ability to create calculated fields, allowing users to perform complex calculations on their data, further enhancing the depth of analysis possible.
- <u>Visualizing Data with Attractive Charts</u>: Tableau offers a wide range of visualization options, from basic charts to advanced graphs, enabling users to create visually appealing and informative representations of their data.
- <u>Publishing Dashboard on Tableau Public</u>: Tableau allows users to easily publish their dashboards to Tableau Public, making them accessible to a wider audience and enabling easy sharing and collaboration.

MySQL Workbench

Mysql@127.0.0.1:3306

File   Edit   View   Query   Database   Server   Tools   Scripting   Help

Navigator

SCHEMAS

Filter objects

▼ classicmodels
  ▶ Tables
  ▶ Views
  ▶ Stored Procedures
    Functions
▶ sys

SQL project*

Limit to 1000 rows

```
30
31    -- 2.List the Names and salary of the employees whose salary is greater than 1000
32
33  ● SELECT ename, sal
34    FROM employees_table
35    WHERE sal > 1000;
36
37
38
```

SQL Additions

◀ ▶ | Jump to

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Administration   Schemas

Information

Connection Details
Name:
    Ho127.0.0.1
    Po3306
    Lo
User root
    Cu
User root@localhost
    SSL
ciphe TLS_AES_128_GCM_SHA
Server
    MySQL Community
    Pr Server - GPL
    V8.0.36
Connector
    V8 C++ 8.3.0

Object Info   Session

Result Grid | Filter Rows: Export: Wrap Cell Content: ⬚A

ename	sal
ALLEN	1600.00
WARD	1250.00
JONES	2975.00
MARTIN	1250.00
BLAKE	2850.00
CLARK	2450.00
SCOTT	3000.00
KING	5000.00

employees_table 2

                                              Read Only   Context Help   Snippets

Output

Action Output

#	Time	Action	Message	Duration / Fetch
❌ 6	14:00:53	CREATE TABLE employees_table (empno INT, ename VARCHAR(255), job VARCH...	Error Code: 1050. Table 'employees_table' already exists	0.000 sec
✅ 7	14:01:06	use classicmodels	0 row(s) affected	0.000 sec
✅ 8	14:01:12	select * from employees_table LIMIT 0, 1000	14 row(s) returned	0.000 sec / 0.000 sec
✅ 9	14:02:22	SELECT ename, sal FROM employees_table WHERE sal > 1000 LIMIT 0, 1000	12 row(s) returned	0.000 sec / 0.000 sec

```
46
47    -- 4.   List average monthly salary for each job within each department
48
49  • SELECT deptno, job, AVG(sal) AS avg_salary
50    FROM employees_table
51    GROUP BY deptno, job;
52
53
54
```

deptno	job	avg_salary
20	CLERK	950.000000
30	SALESMAN	1400.000000
20	MANAGER	2975.000000
30	MANAGER	2850.000000
10	MANAGER	2450.000000
20	ANALYST	3000.000000
10	PRESIDENT	5000.000000
30	CLERK	950.000000

#	Time	Action	Message	Duration / Fetch
8	14:01:12	select * from employees_table LIMIT 0, 1000	14 row(s) returned	0.000 sec / 0.000 sec
9	14:02:22	SELECT ename, sal FROM employees_table WHERE sal > 1000 LIMIT 0, 1000	12 row(s) returned	0.000 sec / 0.000 sec
10	14:03:15	SELECT ename FROM employees_table WHERE ename LIKE 'A%' OR ename LIKE 'S%' LI...	4 row(s) returned	0.000 sec / 0.000 sec
11	14:03:48	SELECT deptno, job, AVG(sal) AS avg_salary FROM employees_table GROUP BY deptno, j...	9 row(s) returned	0.000 sec / 0.000 sec

MySQL Workbench

Mysql@127.0.0.1:3306

File   Edit   View   Query   Database   Server   Tools   Scripting   Help

Navigator

SCHEMAS

Filter objects

▼ classicmodels
  ▶ Tables
  ▶ Views
  ▶ Stored Procedures
    Functions
▶ sys

SQL project*

Limit to 1000 rows

```
52
53
54
55       -- 5.   Display the empno, Emp name and the Manager name under whom the Employee works.
56
57 ●    SELECT empno, ename, ename AS manager_name
58        FROM employees_table
59        LEFT JOIN employees m ON mgr = empno;
60
```

SQL Additions

Jump to

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Result Grid   Filter Rows: Export: Wrap Cell Content:

empno	ename	manager_name
7369	SMITH	SMITH
7499	ALLEN	ALLEN
7521	WARD	WARD
7566	JONES	JONES
7654	MARTIN	MARTIN
7698	BLAKE	BLAKE
7782	CLARK	CLARK
7788	SCOTT	SCOTT

Result 5

Read Only   Context Help   Snippets

Administration   Schemas

Information

Connection Details
Name:
    Hd127.0.0.1
    Pd3306
    Lo
       root
User
    Cu
       root@localhost
User
    SSL
    ciphe TLS_AES_128_GCM_SHA
Server
    MySQL Community
    Pr Server - GPL
    V8.0.36
Connector
    V&C++ 8.3.0

Object Info   Session

Output

Action Output

#	Time	Action	Message	Duration / Fetch
❌ 14	14:04:55	SELECT e.empno, e.ename, e.ename AS manager_name FROM employees_table e LEFT J...	Error Code: 1054. Unknown column 'm.empno' in 'on clause'	0.000 sec
❌ 15	14:05:55	SELECT e.empno, e.ename, m.ename AS manager_name FROM employees_table e LEFT J...	Error Code: 1054. Unknown column 'm.ename' in 'field list'	0.000 sec
❌ 16	14:06:11	SELECT empno, ename, ename AS manager_name FROM employees_table LEFT JOIN em...	Error Code: 1054. Unknown column 'e.mgr' in 'on clause'	0.000 sec
✅ 17	14:06:17	SELECT empno, ename, ename AS manager_name FROM employees_table LEFT JOIN em...	14 row(s) returned	0.016 sec / 0.000 sec

MySQL Workbench

Mysql@127.0.0.1:3306 ✕

File   Edit   View   Query   Database   Server   Tools   Scripting   Help

Navigator

SCHEMAS

Filter objects

▼ 🛢 **classicmodels**
  ▶ 📇 Tables
  ▶ 📇 Views
  ▶ 📇 Stored Procedures
    📇 Functions
▶ 🛢 sys

Administration   Schemas

Information

Connection Details
Name:
  H127.0.0.1
  P3306
  Lo
User root
  Cu
User root@localhost
  SSL
ciphe TLS_AES_128_GCM_SHA
Server
  MySQL Community
Pr Server - GPL
V8.0.36
Connector
  V C++ 8.3.0

Object Info   Session

SQL project* ✕

Limit to 1000 rows

```
64        -- 6.    Display Empname, department no, salary , Rank of salary in Organisation ,
65        -- Rank of Salary in their respective department
66
67 ⊟  WITH ranked_salaries AS (
68          SELECT empno,ename,deptno,sal,
69              RANK() OVER (ORDER BY sal DESC) AS org_rank,
70              RANK() OVER (PARTITION BY deptno ORDER BY sal DESC) AS dept_rank
71          FROM employees_table)
72      SELECT ename,deptno,sal,org_rank,
73          dept_rank FROM ranked_salaries;
74
```

Result Grid | Filter Rows: Export: Wrap Cell Content:

ename	deptno	sal	org_rank	dept_rank
▶ KING	10	5000.00	1	1
CLARK	10	2450.00	6	2
MILLER	10	1300.00	9	3
SCOTT	20	3000.00	2	1
FORD	20	3000.00	2	1

Result 8 ✕

❶ Read Only   Context Help   Snippets

SQLAdditions

◀ ▶   Jump to

Automatic context help is
disabled. Use the toolbar t
manually get help for the
current caret position or t
toggle automatic help.

Output

Action Output

#	Time	Action	Message	Duration / Fetch
✅	17	14:06:17	SELECT empno, ename, ename AS manager_name FROM employees_table LEFT JOIN em...	14 row(s) returned	0.016 sec / 0.000 sec
✅	18	14:07:32	WITH ranked_salaries AS (SELECT empno,ename,deptno,sal, RANK() OVER (ORD...	14 row(s) returned	0.000 sec / 0.000 sec
✅	19	14:08:16	WITH ranked_salaries AS (SELECT empno,ename,deptno,sal, RANK() OVER (ORD...	14 row(s) returned	0.000 sec / 0.000 sec
✅	20	14:08:37	WITH ranked_salaries AS (SELECT empno,ename,deptno,sal, RANK() OVER (ORD...	14 row(s) returned	0.000 sec / 0.000 sec

```
77   -- 7. Display each department highest salary holder employee name with his salary in descending order and
78   -- arrange department according to salary in descending order
79
80 ⊠⊖ WITH ranked_salaries AS (
81        SELECT empno,ename,deptno,sal,
82            RANK() OVER (PARTITION BY deptno ORDER BY sal DESC) AS dept_rank
83        FROM employees_table)
84    SELECT ename,sal,deptno
85    FROM ranked_salaries WHERE dept_rank = 1
86    ORDER BY sal DESC,deptno DESC;
87
```

ename	sal	deptno
KING	5000.00	10
SCOTT	3000.00	20
FORD	3000.00	20
BLAKE	2850.00	30

#	Time	Action	Message	Duration / Fetch
18	14:07:32	WITH ranked_salaries AS (SELECT empno,ename,deptno,sal,	RANK() OVER (ORD... 14 row(s) returned	0.000 sec / 0.000 sec
19	14:08:16	WITH ranked_salaries AS (SELECT empno,ename,deptno,sal,	RANK() OVER (ORD... 14 row(s) returned	0.000 sec / 0.000 sec
20	14:08:37	WITH ranked_salaries AS (SELECT empno,ename,deptno,sal,	RANK() OVER (ORD... 14 row(s) returned	0.000 sec / 0.000 sec
21	14:09:42	WITH ranked_salaries AS (SELECT empno,ename,deptno,sal,	RANK() OVER (PAR... 4 row(s) returned	0.000 sec / 0.000 sec

Creating Table:

- Used to define a new table with columns, data types, and constraints.

- Establishes the structure for storing data in a database.

Greater Than Function:

- The > operator compares two values.

- Returns true if the left operand is greater than the right operand, otherwise false.

Name Starts with Alphabet: Uses the **LIKE** operator with the pattern 'A%' to filter names starting with 'A'.

Average and Group By Function:

- The **AVG()** function calculates the average value of a numeric column.

- When used with **GROUP BY**, it calculates the average for each group.

Left Join Function:

- Combines rows from two or more tables based on a related column.

- Returns all rows from the left table and matched rows from the right table.

Ranking Function with Descending and Partition Function:

- Provides ranking functions (**ROW_NUMBER(), RANK(), DENSE_RANK()**) to assign a rank to each row.

- Used with **ORDER BY** and **PARTITION BY** to rank rows based on specific criteria within partitions.

Insights generated

Agenda- Reason for attrition in a company

•Findings- Overall job satisfaction rate in all departments is 66%.

•The department with the maximum attrition rate is Research & Development (51.2%), while the minimum is in the Hardware department (49.4%).

•The main reasons for the high attrition rate in the Research & Development department are:

- Low work-life balance (2.5 out of 4 rating).

- Low total working years (20.25 years).

- Low monthly income (26,007 only).

- Low promotion rate (on average 5.85 years since the last promotion).

•The attrition rate for females in the Research & Development department is highest at 51%, primarily due to low monthly income (25,800 only compared to sales 26,297) and a low promotion rate (5.8 years on average).

•Job satisfaction ratings are best in the Healthcare Representative department and worst in the Sales Representative department.

Suggestions for the client

Agenda- Reason for attrition in a company

Based on the analysis of the data, as a data analyst, I would suggest the following recommendations to the client to address the high attrition rate and improve job satisfaction in the Research & Development department:

- **Improve Work-Life Balance**: Implement policies and initiatives to improve work-life balance for employees in the Research & Development department. This could include flexible working hours, remote work options, and promoting a culture that values work-life balance.

- **Increase Monthly Income**: Review and adjust the salary structure to ensure that employees in the Research & Development department are being fairly compensated. Consider conducting a salary benchmarking exercise to ensure that salaries are competitive within the industry.

- **Enhance Promotion Opportunities**: Develop a clear and transparent promotion policy that provides employees in the Research & Development department with a clear path for career advancement. Provide regular feedback and opportunities for skill development to help employees progress in their careers.

- **Address Gender Disparities**: Address the gender pay gap by ensuring that female employees in the Research & Development department are being paid equally for equal work. Implement policies to support gender diversity and inclusion in the department.

- **Improve Job Satisfaction**: Conduct regular surveys to assess job satisfaction levels and identify areas for improvement. Use the feedback to implement changes that enhance job satisfaction, such as improving communication channels, recognition programs, and opportunities for professional growth.

- **Provide Support for Total Working Years**: Recognize and reward employees for their loyalty and contribution to the company. Provide opportunities for long-term employees to mentor and share their knowledge with newer employees, which can help improve job satisfaction and retention.

- **Monitor and Evaluate**: Continuously monitor the effectiveness of these recommendations and evaluate their impact on attrition rates and job satisfaction in the Research & Development department. Adjust strategies as needed to ensure they are meeting the desired outcomes.

By implementing these recommendations, the client can work towards reducing attrition rates and improving job satisfaction in the Research & Development department, ultimately leading to a more engaged and productive workforce.