

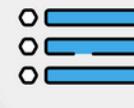
HR ANALYTICS

EMPLOYEE ATTRITION AND WORKFORCE ANALYSIS

BY: KASHISH PAL

Note: This project uses a publically available HR dataset based on IBM attrition designed for business analytics. Insights and findings reflect real-world workforce challenges.





Project Outline

■ Project overview	03
■ Data Cleaning & Validation	4-5
■ Feature Engineering	6-7
■ Data Integration : Table Joins	08
■ Exploratory Data Analysis & Insights	9-11
■ HR Analytics Dashboards	12-13
■ DAX Measures % KPIs	14-15
■ Key Insights & Deep-Dive Analysis	16-17
■ Business Recommendations & Actionable Steps	18
■ Summary & References	19



PROJECT OVERVIEW

This project uses SQL and Power BI to generate actionable insights on employee attrition and workforce dynamics within an organization. The goal is to uncover the main factors influencing staff departures, identify at-risk segments, and support data-driven strategies for improving retention, onboarding, and overall workforce stability.

◆ Business Problem Solved:

How can we reduce employee attrition and strengthen workforce retention by identifying key drivers—such as department, job satisfaction, compensation, work conditions, and career growth?

Through advanced data cleaning, feature engineering, and interactive dashboarding, this project equips HR leaders with clear insights into the causes and patterns of employee departures, enabling informed strategies that strengthen workforce engagement and organizational stability.



1. Data Cleaning & Validation : Ensuring Data Accuracy & Consistency

● Problem: Unrealistic Employee Ages Detected

Verified if any employees were underage (<18) or overaged (>60)

Query Used:

```
SELECT EmpID, Age FROM hr_basic_profile  
WHERE Age < 18 OR Age > 60;
```

Key takeaway: Outlier detection is a key part of HR data cleaning. Age values outside the 18–60 range should be flagged, investigated, or corrected based on business rules.

● Problem: Mismatched Education Levels vs. Age

Detected unrealistic "PHD" and "Post-Graduation" entries for employees under 20, and questionable "PHD" values for those aged 20–24

Query Used:

```
UPDATE hr_basic_profile SET Higher_Education = '12th'  
WHERE Age < 20 AND Higher_Education IN ('PhD', 'Post Graduation');  
UPDATE hr_basic_profile  
SET Higher_Education = 'Needs Review'  
WHERE Age < 25 AND Higher_Education = 'PhD';
```

Key takeaway: Age and education must be logically aligned. When data is clearly incorrect, fix it; when uncertain or unusual, flag it for review instead of making assumptions – this maintains data quality and integrity.

Continue....

● Problem: Find Duplicate Employee Records

Needed to confirm the uniqueness of employee records.

Query Used:

```
SELECT EmpID, COUNT(*) FROM hr_leave_recruitment  
GROUP BY EmpID HAVING COUNT(*) > 1;
```

Key takeaway: All employees in this table have unique IDs. Analysis on duplicate records can lead to misleading or entirely incorrect business conclusions.

● Problem: Checking for Null & Duplicate Values

Ensure all categorical columns are clean (no typos, inconsistent casing, or extra spaces) and that no critical fields are missing (NULL).

Query Used:

```
SELECT DISTINCT Higher_Education FROM hr_basic_profile;  
SELECT * FROM hr_basic_profile WHERE EmpID IS NULL;
```

Key takeaway: All distinct values were clean and consistently formatted. No action was needed using functions like TRIM() or UPPER(). No NULLs were found in key field but checks were done to confirm data quality.



2. Feature Engineering

Feature engineering transformed raw fields into grouped or business-relevant columns, enabling clearer visualization, cohort analysis, and actionable workforce insights.

SUMMARY OF DERIVED COLUMNS:

- **Age Group:** Categorizes employee ages into bands (e.g., Below 25, 25-35, etc.) to analyze attrition and retention patterns across different age demographics, enabling better workforce segmentation.
- **Promotion Status:** Classifies employees based on time since last promotion, highlighting those with delayed advancement who may be vulnerable to attrition due to career stagnation.
- **Hire Decade:** Categorized employee hire dates by decade, providing trend analysis on hiring surges or declines.
- **Tenure Band:** Segmented years at the company into lifecycle bands like "New Joiner," "Early Career," "Experienced," "Long-Term," and "Senior Most" to study workforce stability and attrition patterns.
- **Experience Level:** Grouped total professional experience into "Fresher," "Junior," "Mid-level," and "Senior" to facilitate targeted retention and tenure analysis.
- **Total Off Days:** Combines leaves and absenteeism into a single metric to better evaluate absence patterns and their possible correlation with employee turnover.

Example Feature Engineering Queries :

Below are selected SQL queries illustrating key feature engineering steps in this project.

● Query Used: Extracting HireYear & Hire Decade

```
ALTER TABLE tenure_attrition ADD COLUMN Hire_year INT;  
UPDATE tenure_attrition SET Hire_year = YEAR(Date_of_Hire);  
  
ALTER TABLE tenure_attrition ADD COLUMN Hire_decade VARCHAR(15);  
  
UPDATE tenure_attrition SET Hire_decade =  
CASE  
    WHEN Hire_year BETWEEN 1969 AND 1989 THEN 'Before 1990'  
    WHEN Hire_year BETWEEN 1990 AND 1999 THEN '1990-1999'  
    WHEN Hire_year BETWEEN 2000 AND 2009 THEN '2000-2009'  
    WHEN Hire_year BETWEEN 2010 AND 2019 THEN '2010-2019'  
    ELSE 'Latest Hirings'  
END;
```

● Query Used: For Age Group

```
ALTER TABLE hr_basic_profile ADD COLUMN age_group  
VARCHAR(20);  
  
UPDATE hr_basic_profile SET age_group =  
CASE  
    WHEN age < 25 THEN "Below 25"  
    WHEN age BETWEEN 25 AND 35 THEN "25-35"  
    WHEN age BETWEEN 35 AND 50 THEN "35-50"  
    ELSE "above 50"  
END;
```



3. Data Integration: Table Joins

Produced a single, consolidated table containing essential features for analysis—Employee demographics, compensation, tenure and attrition enabling smooth Power BI import and advanced EDA.

Join Strategy:

- Used INNER JOIN to retain all employee records.
- Joined on the unique employee identifier (e.g., EmplID) which is consistently available across all tables.

Business Impact:

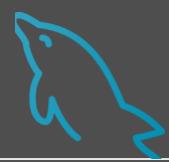
Streamlines Reporting: One unified dataset simplifies dashboard building and KPI calculation in Power BI.

Supports Rich Analysis: All relevant factors for attrition, tenure, compensation, and employee profile are available in one place.

Enables Data Quality Checks: Easy to identify and handle missing or mismatched records before analysis.

Query Used:

```
CREATE TABLE final_hr_analytics AS
SELECT
    p.EmplID, p.Gender, p.MaritalStatus, p.Higher_Education,
    p.Department, p.JobRole, p.BusinessTravel, p.Job_mode,
    p.Mode_of_work, p.OverTime, p.age_group, p.Exp_level,
    c.MonthlyIncome, c.JobSatisfaction, c.TrainingTimesLastYear,
    l.Source_of_hire, l.Work_accident,
    a.Attrition, a.Status_of_leaving, a.Hire_year, a.Hire_decade,
    a.YearsAtCompany, a.tenure_band, a.Promotion_status
FROM hr_basic_profile p
JOIN hr_compensation_performance c ON p.EmplID = c.EmplID
JOIN hr_leave_recruitment l ON p.EmplID = l.EmplID
JOIN tenure_attrition a ON p.EmplID = a.EmplID;
```



4. Explotarory Data Analysis & Insights

1. Which year had the highest employee hiring volume?

Query Used:

```
SELECT Hire_year, COUNT(*) AS Total_hires  
FROM final_hr_analytics  
GROUP BY Hire_year  
ORDER BY Total_hires DESC  
LIMIT 1;
```

Key Outcome: 2016 is the highest hiring year with 196 hires. It helps understand workforce growth & its impact on attrition trends.

2. Does overtime work increase the risk of attrition?

Query Used:

```
SELECT OverTime, COUNT(Attrition)  
FROM final_hr_analytics  
WHERE Attrition = 'Yes'  
GROUP BY OverTime;
```

Key Outcome: Employees working *overtime* have *higher attrition*, suggesting overtime contributes to turnover and needs retention focus.

Continue.....

3. What are the main reasons employees are leaving the company?

Query Used:

```
SELECT Status_of_leaving, COUNT(Attrition)
FROM final_hr_analytics
WHERE Attrition = 'Yes'
GROUP BY Status_of_leaving;
```

Key Outcome: Work accidents are the top exit reason, followed by salary and management issue highlighting focus on improving workplace safety, competitive salary, and departmental management to reduce employee exits.

4. what is the count and percentage of employees who have left the company?

Query Used:

```
SELECT Attrition, COUNT(*) AS Count,
ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*)
FROM final_hr_analytics), 1) AS Percentage
FROM final_hr_analytics
GROUP BY Attrition;
```

Key Outcome: The attrition count is 237 with a rate of 16.1%, slightly exceeding the 15% industry benchmark, indicating a need for targeted retention efforts.

Continue....

5. Which departments have the highest attrition rate?

Query Used:

```
SELECT Department,  
       ROUND(COUNT(CASE WHEN Attrition = 'Yes'  
                         THEN 1 END) * 100.0 / COUNT(*), 1)  
             AS AttritionRate  
      FROM final_hr_analytics  
     GROUP BY Department  
ORDER BY AttritionRate DESC;
```

Key Outcome: Sales (**20.6%**) and HR (**19.0%**) have the highest attrition rates, both above the benchmark of **15%**, indicating higher turnover risk.

6. What is the current distribution of active employees across each department?

Query Used:

```
SELECT Department, COUNT(Attrition) As Active_employees  
      FROM final_hr_analytics  
     WHERE Attrition = 'No'  
     GROUP BY Department  
ORDER BY Active_employees DESC;
```

Key Overview: Human Resources has only **51** active employees versus **828** in Research & Development and **354** in Sales. This small HR team may impact timely employee support and affect overall workforce management.

HR Analytics



Workforce Analysis

Attrition Analysis

Female

Male

Departm...

Human Re..

Research ...

Sales

Tenure B...

All

Job Role

All

Job Mode

All



237
Attrition count

1470
Total employees

1233
Active employees

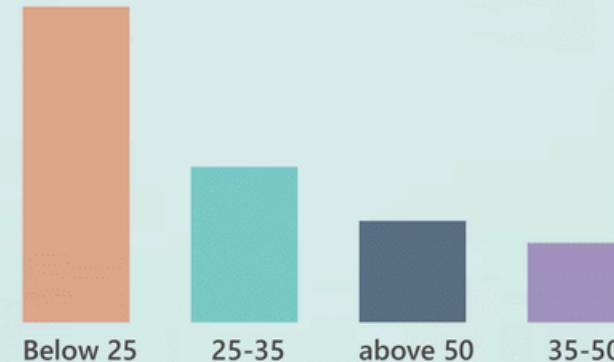
7.0
Avg tenure

6.5K
Avg salary

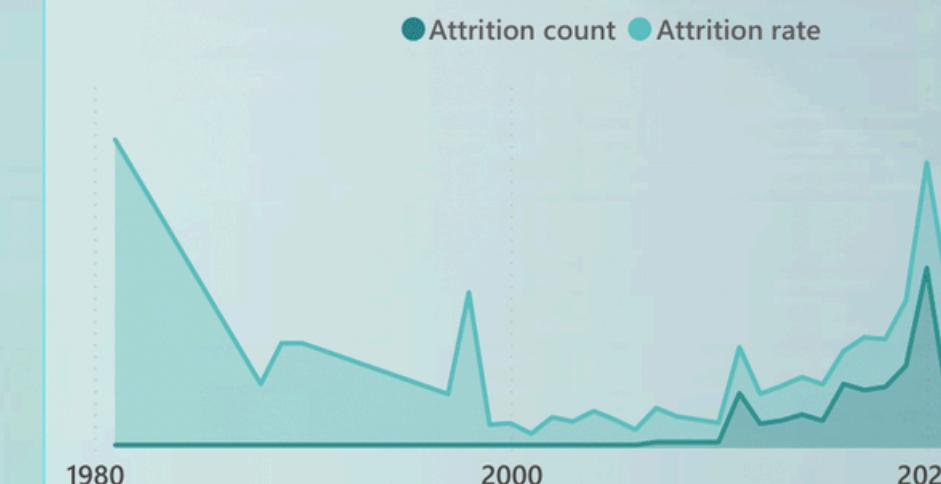
Attrition rate by Department



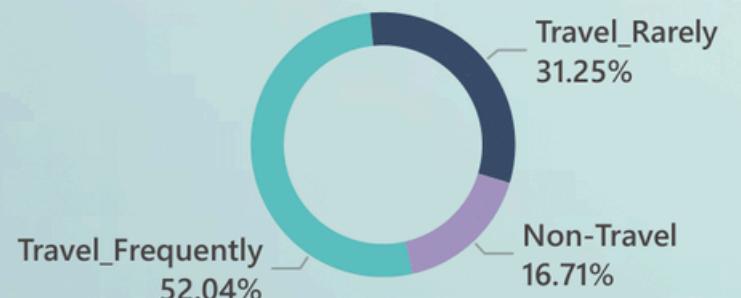
Attrition Rate By Age Group



Attrition Trends By Hire Years



Attrition rate by BusinessTravel



Attrition Rate By Over Time



Attrition Rate By Experience Level And Tenure Band

● New Joiner ● Early career ● Experienced ● Long-term ● Senior-most



JobRole

Attrition count

Laboratory Technician	62
Sales Executive	57
Research Scientist	47
Sales Representative	33
Human Resources	12
Total	237

Attrition Rate By Exit Reasons



HR Analytics



Attrition
Analysis

Workforce
Analysis

Depart...

All

Job Mode

All

Hire Dec...

All

Hiring S...

All

Tenure B...

All



1470
Total employees

1233
Active employees

1233
cumulative active emp

6.5K
Avq salary

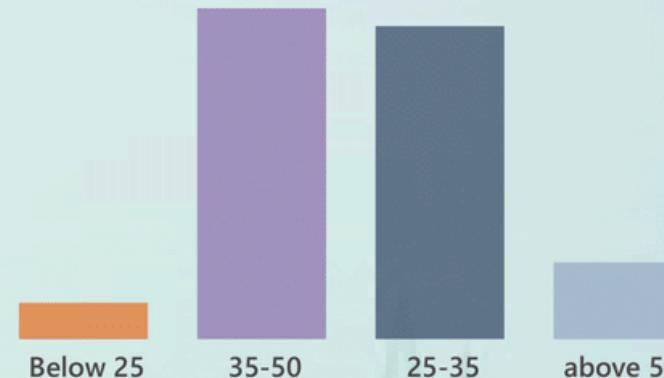
7.0
Avq tenure

Retention And Attrition Rate Over Time

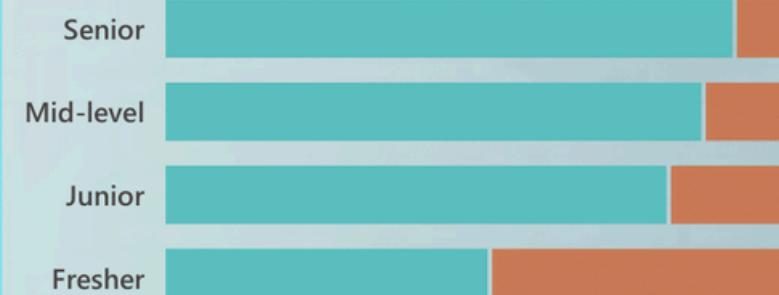
● Retention_rate ● Attrition rate

Before 1990 1990-1999 2000-2009 2010-2019 Latest
hirings

Active Employees By Age Group



Active vs Exited Employees By Experience level



Active employees by Department

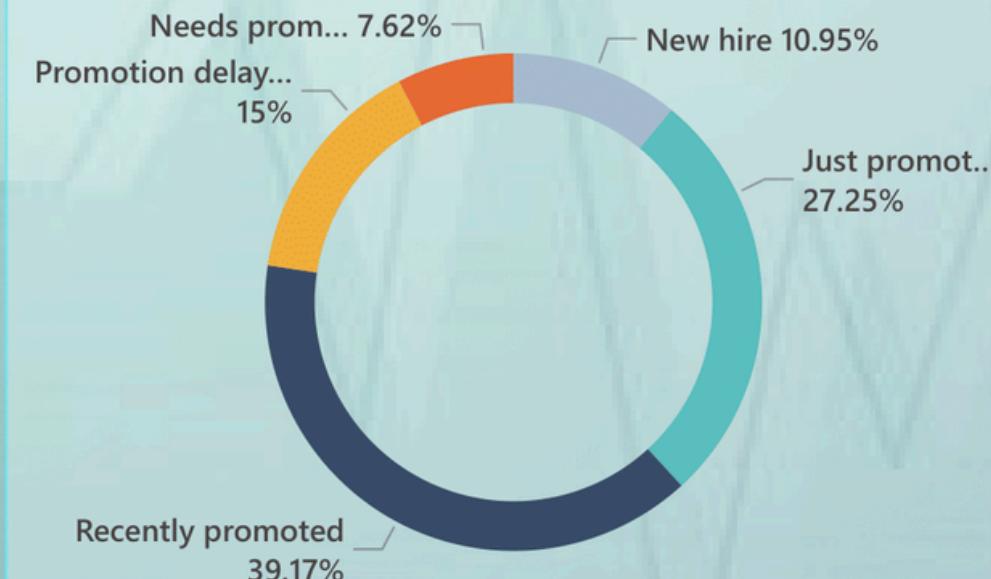
Research & Development



Sales

Human Resources

Active Employees By Promotion Status

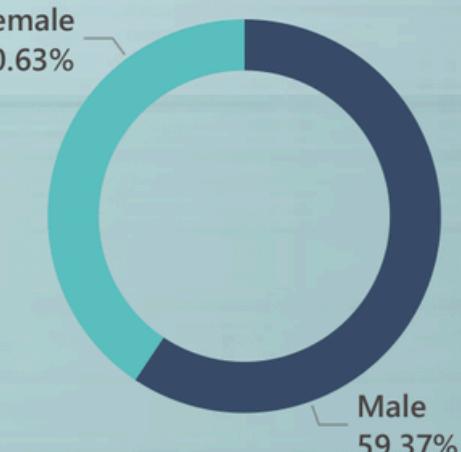


Tenure B...

Active Employees By Gender

Female

40.63%



Male

59.37%



DAX Measures & KPIs

These DAX measures drive the interactive HR dashboards, enabling dynamic business insights and clear data analysis. Each formula helps transform raw workforce data into actionable key performance indicators for better decision-making.

Attrition Rate:

Calculates % of employees who left the company & helps benchmark against the industry average.

Attrition rate = DIVIDE([Attrition count], [Total employees])*100

Average Salary:

Shows the mean salary of all employees for compensation and pay analysis.

Avg salary = AVERAGE(' final_hr_analytics'[MonthlyIncome])

Average Tenure:

Average number of years employees have worked at the company. Indicates overall workforce experience. Higher average tenure suggests a more stable or experienced workforce, while lower tenure can signal insatibility or rapid organizational growth.

Avg tenure = AVERAGE('final_hr_analytics'[YearsAtCompany])

Attrition Count

Displays the absolute number of employees who exited the company.

**Attrition count = CALCULATE(COUNT('final_hr_analytics'[EmpID]),
' final_hr_analytics'[Attrition]="Yes")**



DAX Measures & KPIs

Total Employees

Shows counts all employees (both current and former). Useful for understanding overall company size and as a base for HR KPIs.

```
Total employees = CALCULATE(COUNT('final_hr_analytics'[EmpID]))
```

Retention Rate:

Calculates the % of employees who stayed with the company, helping evaluate overall employee stability.

```
Retention_rate = DIVIDE([Active employees],[Total employees]) * 100
```

Cumulative Active Employees:

This advanced measure counts all employees still active up to the selected year & tracks the present size of the workforce. This is important for resource planning, workload distribution, and tracking hiring effectiveness.

```
cumulative active emp = CALCULATE(COUNTROWS('final_hr_analytics'),  
FILTER(ALL('final_hr_analytics'),'final_hr_analytics'[Hire_year]<= MAX  
('final_hr_analytics'[Hire_year])&&'final_hr_analytics'[Attrition] = "No" ))
```

Active Employees:

Counts only employees hired and retained in the selected year, highlighting current year new hires who remained active.

```
Active employees = CALCULATE(COUNT('final_hr_analytics'[EmpID]),  
'final_hr_analytics'[Attrition] = "NO")
```

Key Insights & Deep-Dive Analysis

▼ Attrition Patterns & Departmental Risk

- Attrition rates are highest in **Sales (20.6%)** and **HR (19.0%)**, both exceeding the **15% industry benchmark**. This highlights urgent retention risks in these functions.

Main drivers:

- Work accidents and low salaries remain leading exit reasons
- Overtime (frequent overtime linked to higher attrition)
- Employees traveling frequently have higher attrition—consider reviewing travel policies or offering improved benefits

Onboarding & Support: While HR headcount is appropriate, onboarding quality and ongoing support especially for junior and younger employees need improvement to strengthen early retention.

- **HR Department** is **male dominated (72.5%)** with very low new hires & no female hires under 25
- In **HR Department** attrition rate of **females (30%)** is very high as compared to **males (14%)**

Key Insights & Deep-Dive Analysis

▼ Workforce Tenure & Career Progression

Younger employees (<25 years) and entry-level (Fresher, Junior) staff are most likely to leave, highlighting the need for enhanced engagement and career support.

7.6% of active staff are waiting for promotion, and an additional 15% have experienced promotion delays indication future attrition risks.

▼ Attrition and Retention Over Time

Attrition rates have risen steadily for recent hire cohorts while retention has decreased, suggesting that changes in workplace culture, career growth, or pay over time may be affecting why employees stay or leave,

Historical analysis: Retention was stronger in earlier decades. Regular historical analysis is important to identify and address these ongoing challenges.



Business Recommendations & Actionable Steps

▲ Strengthen Onboarding & Training:

- Improve onboarding & peer-mentorship for new & junior staff to ensure they are better integrated into the company culture.

▲ Monitor Overtime & Workload:

- Set clearer boundaries for overtime work and encourage work-life balance, reducing burnout risk.
- Regularly assess teams or roles with excessive overtime for resource optimization.

▲ Engagement & Growth:

- Enhance job satisfaction through regular feedback, transparent communication, and visible recognition programs.
- Create faster and clearer promotion pathways for high-potential employees.

▲ Compensation & Benefits:

- Benchmark salaries to market standards and address any pay gaps contributing to attrition.
- Provide targeted benefits and support, especially for frequent travelers and front-line roles.

▲ Historical Trend Review:

- Monitor attrition and retention rates over time and review them regularly, so any negative trends can be detected early and addressed with targeted actions.

▲ Promote gender balance in HR:

- Conduct targeted exit interviews & engagement surveys to understand challenges faced by female HR staff.
- Revise hiring strategy to include more young professionals & promote diversity in HR roles.

▲ Workplace Safety:

- Invest in reducing workplace accidents through training and improved safety practices.



PROJECT SUMMARY:

This HR analytics project applies SQL and Power BI to analyze employee attrition and workforce trends. Key findings highlight that work accidents, salary dissatisfaction, overtime and onboarding issues are primary drivers of employee departures—especially among sales, HR, and junior staff. By transforming raw HR data into actionable KPIs and dashboards, the project delivers clear, practical insights and recommendations to help HR leaders improve retention, engagement, and organizational stability through data-driven decisions.

■ References & Acknowledgement

Dataset: [Kaggle](#)

Icons: [Flaticons](#)

Database & Data Processing: MySQL

Data Visualization: Power BI

Design & Presentation: Canva

 : [Lets connect on linkedin!](#)

 : [GitHub](#)

 : Kashishpal2020@gmail.com