HR Analytics Project Report

Workforce Stability & Performance Analysis

---

1. Executive Summary

This project focuses on analyzing workforce data to understand employee attrition, engagement, and performance across the organization. The objective was to create a structured, end-to-end HR analytics solution—starting from raw data, cleaning and transformation, database storage, and finally interactive dashboards in Power BI. The outcome is a decision-ready HR Analytics Dashboard that helps leadership identify workforce risks, performance gaps, and hiring effectiveness.

---

2. Business Context

2.1 Business Problem

Organizations often face challenges related to: - High employee attrition without clear visibility into root causes - Inconsistent performance and engagement across departments - Difficulty identifying at-risk employee groups early - Limited insight into whether hiring channels are producing high-quality, long-term employees

Without data-driven insights, HR and leadership teams rely on assumptions rather than evidence-based decisions.

2.2 Business Goal

To improve workforce stability and performance by identifying the key drivers of employee attrition, engagement, and performance across departments and roles.

2.3 Project Objectives

- Analyze employee attrition trends and identify high-risk departments
- Understand the relationship between engagement, satisfaction, performance, and attrition
- Evaluate workforce composition and diversity
- Assess the effectiveness of recruitment sources
- Deliver clear, executive-friendly dashboards for decision-making

---

3. Dataset Overview

3.1 Data Description

The dataset represents HR employee-level data including: - Employee demographics (gender, age, race, location) - Employment details (department, position, manager, employment status) - Performance and engagement metrics - Attrition and termination details - Recruitment source information

Each row represents one employee record.

3.2 Key Variables

- **EmpID** – Unique employee identifier
- **Department / Position / ManagerName** – Organizational structure
- **Attrition** – Employee status (0 = Active, 1 = Left)
- **PerformanceScore** – Employee performance rating
- **EngagementSurvey / EmpSatisfaction** – Engagement and satisfaction scores
- **Absences** – Attendance indicator
- **RecruitmentSource** – Hiring channel
- **DOB, DateofHire, DateofTermination** – Time-based attributes

---

4. Data Analysis Pipeline

4.1 Tools Used

- **Python (Pandas, NumPy)** – Data cleaning, feature engineering
- **CSV** – Data exchange format
- **MySQL** – Centralized data storage (single source of truth)
- **Power BI** – Data modeling, visualization, and dashboarding

---

5. Data Cleaning & Preparation (Python)

5.1 Missing Value Analysis

- Identified missing values in DOB, DateofHire, DateofTermination
- Derived fields such as **Age** and **Tenure_Years** naturally contained nulls where source dates were missing
- Nulls were intentionally preserved to avoid creating artificial data

5.2 Feature Engineering

- **Age** calculated from DOB
- **Tenure_Years** calculated from DateofHire and DateofTermination

   • **Engagement Level** bucketed into Low / Medium / High
   • **Age Group** created for demographic analysis

5.3 Date Handling

   • All date columns standardized to ISO format (YYYY-MM-DD)
   • Ensured compatibility with SQL and Power BI

5.4 Export

   • Final cleaned dataset exported as CSV
   • Imported into MySQL to serve as the analytical data source

---

6. Database Layer (MySQL)

6.1 Purpose

   • Act as a single source of truth
   • Enable scalable reporting and BI integration
   • Maintain data consistency across dashboards

6.2 Validation

   • Row counts verified between Python and SQL
   • Data types validated for dates, numeric, and categorical fields

---

7. Power BI Data Modeling

7.1 Measures Created

   • Total Employees
   • Active Employees
   • Attrition Rate (%)
   • Average Performance Score
   • Average Engagement Score

7.2 Design Decisions

   • Gender used consistently as a legend across all visuals (Male/Female)
   • Consistent color usage across all pages for clarity and comparison
   • Limited filters per page to avoid clutter and improve usability

---

8. Dashboard Structure & Insights

Page 1: Executive Workforce Overview

Business Question: How healthy is our workforce right now?

Key insights: - Overall attrition rate and workforce size - Attrition trends over time - Department-level workforce distribution - Location-wise attrition patterns

---

Page 2: Attrition Drivers & Risk Analysis

Business Question: Who is leaving, and why?

Key insights: - High-attrition departments - Attrition by performance level - Attrition by engagement level - Top reasons for termination

---

Page 3: Performance & Engagement Insights

Business Question: How well are employees performing and how engaged are they?

Key insights: - Performance score distribution - Relationship between engagement and satisfaction - Identification of at-risk employee groups - Department-wise engagement and satisfaction comparison

---

Page 4: Workforce Demographics & Hiring Effectiveness

Business Question: Are we hiring and structuring the workforce effectively?

Key insights: - Gender, age, and race distribution - Recruitment sources with high attrition - Recruitment sources linked to stronger performance - Geographic workforce concentration

---

9. Business Solutions & Recommendations

Based on the analysis:

1. **Targeted Retention Strategies**
2. Focus on departments with consistently high attrition

3. Address engagement gaps before they lead to exits

4. **Performance Management Improvements**

5. Monitor low-performance and low-engagement clusters

6. Use early warning indicators such as absenteeism

7. **Hiring Optimization**

8. Prioritize recruitment sources associated with lower attrition and higher performance

9. Reduce dependence on channels producing high turnover

10. **Diversity & Workforce Planning**

11. Use demographic insights to support inclusive workforce planning
12. Align hiring with long-term organizational needs

---

10. Conclusion

This HR Analytics project demonstrates a complete, real-world data analysis workflow—from raw data to executive-ready insights. By integrating Python, SQL, and Power BI, the project delivers actionable intelligence that supports strategic HR decisions. The final dashboards enable leadership to proactively manage attrition, improve engagement, and strengthen workforce stability.

---

11. Future Enhancements

• Predictive attrition modeling using machine learning
• Employee-level risk scoring
• Automated data refresh and scheduled reporting
• Integration with payroll and performance management systems

---

End of Report