HUMAN RESOURCES

Data Analysis Report

PRESENTED BY

Mohamed Waled, Nourhan Farag,
Basant Yasser, Doha Mostafa

TABLE OF CONTENTS

INTRODUCTION

Dataset Introduction The dataset comprises employee demographic, performance, satisfaction, and organizational data, designed to support human resource analytics and workforce decision-making. It integrates information from multiple dimensions including personal characteristics (e.g., age, gender, education), job attributes (e.g., department, business travel, salary), performance reviews (e.g., manager and self-ratings), and satisfaction metrics (e.g., job, environment, relationship). Additionally, it captures employment dynamics such as tenure, promotions, and attrition status. The dataset is structured across several related tables
— Employee, PerformanceRating, EducationLevel, RatingLevel, and SatisfiedLevel
— enabling a multifaceted analysis of employee behavior and organizational outcomes. This comprehensive view is valuable for identifying drivers of attrition, assessing employee engagement, evaluating performance trends, and informing strategies for talent development and retention.

وزارة الاتصـــــالات
وتكنولوجيا المعلومات

DATASET DESCRIPTION

The dataset consists of several interconnected tables:

- Employee Table: Contains demographic and employment details of employees such as EmployeeID, Gender, Age, Department, Marital Status, Salary, OverTime status, and Years at Company.

- EducationLevel Table: Maps numerical education levels to their descriptive labels.

- PerformanceRating Table: Includes multiple performance metrics including Job Satisfaction, Environment Satisfaction, Work-Life Balance, Self and Manager Ratings, and number of Training Opportunities.

- SatisfiedLevel Table: Translates satisfaction scores (1-5) to human-readable categories.

- RatingLevel Table: Translates performance scores (1-5) to descriptions such as "Meets Expectation" and "Above and Beyond".

The dataset covers an 11-year period and includes 1,470 employee records and over 6,700 performance reviews.

وزارة الاتصـــالات
وتكنولوجيا المعلومات

PREPROCESSING STEPS

- Removed extra spaces and fixed spelling inconsistencies in categorical values.

- Converted date columns to datetime format.

- Created additional features such as age groups, salary levels, and travel categories.

- Merged data from multiple tables using foreign keys.

- Ensured no null values in the core dataframes.

وزارة الاتصـــــــــالات
وتكنولوجيا المعلومات

DATASET DESCRIPTION & NECESSARY DESCRIPTIVE STATISTICS

Employee

This sheet contains core employee demographic and employment-related data. It includes attributes such as name, gender, age, department, distance from home, salary, marital status, stock options, overtime status, and attrition. It also tracks employment duration metrics like years at the company, time since last promotion, and years with the current manager. This sheet is essential for profiling the workforce and analyzing patterns in retention, diversity, and compensation.

RatingLevel

This sheet maps numerical RatingID values to qualitative performance evaluation categories (e.g., Meets Expectation, Exceeds Expectation). It supports standardization of rating terminology across the organization and is used to interpret and visualize performance data in a meaningful way.

DATASET DESCRIPTION & NECESSARY DESCRIPTIVE STATISTICS

PerformanceRating

This sheet captures employee performance evaluations and engagement indicators. It includes self and manager ratings, review dates, and satisfaction scores across work environment, job role, and workplace relationships. It also records training opportunities and participation levels, making it valuable for analyzing performance trends, identifying training gaps, and correlating satisfaction with productivity.

EducationLevel

This lookup table defines the education level categories associated with employees. It maps EducationLevelID to descriptive labels (e.g., High School, Bachelors, Masters). It is used to standardize and interpret the educational background data in the employee records and enables analysis of the impact of education on performance, compensation, and attrition.

وزارة الاتصـــــالات
وتكنولوجيا المعلومات

DATASET DESCRIPTION & NECESSARY DESCRIPTIVE STATISTICS

SatisfiedLevel

This sheet provides a mapping between SatisfactionID and descriptive levels of satisfaction (e.g., Very Dissatisfied to Very Satisfied). It allows for consistent interpretation of satisfaction-related fields across the dataset and supports analysis of workforce engagement and morale.

SatisfactionID	Unique identifier for satisfaction level
SatisfactionLevel	Description of the employee's overall satisfaction level
RatingID	Unique identifier for rating level
RatingLevel	Description of the employee's performance rating
EducationLevelID	Unique identifier for education level
EducationLevel	Highest education qualification attained by the employee
PerformanceID	Unique identifier for the performance review
EmployeeID	Unique identifier for each employee
ReviewDate	Date of the performance review
EnvironmentSatisfaction	Employee's satisfaction with the work environment (scale 1–5)
JobSatisfaction	Employee's satisfaction with their job role (scale 1–5)
RelationshipSatisfaction	Employee's satisfaction with colleague/manager relationships (scale 1–5)
TrainingOpportunitiesWithinYear	Number of training opportunities available in the current year
TrainingOpportunitiesTaken	Number of training opportunities the employee has taken
WorkLifeBalance	Employee's perception of their work-life balance (scale 1–5)
SelfRating	Employee's self-assessment score
ManagerRating	Manager's assessment score of the employee
FirstName	Employee's first name
LastName	Employee's last name
Gender	Gender identity of the employee
Age	Age of the employee
BusinessTravel	Employee's business travel frequency
Department	Department in which the employee works
DistanceFromHome (KM)	Distance between employee's residence and workplace in kilometers
State	State of employment
Ethnicity	Ethnic background of the employee
MaritalStatus	Marital status of the employee
Salary	Annual salary of the employee in USD
StockOptionLevel	Stock options provided to the employee (ordinal level)
OverTime	Whether the employee works overtime (Yes/No)
HireDate	Date the employee was hired
Attrition	Whether the employee has left the company (Yes/No)
YearsAtCompany	Total years the employee has worked at the company
YearsInMostRecentRole	Years the employee has spent in their most recent role
YearsSinceLastPromotion	Years since the employee's last promotion
YearsWithCurrManager	Years the employee has worked with their current manager

EXPLORATORY DATA ANALYSIS

Gender Distribution:
- Female: ~46%
- Male: ~44%
- Non-binary: ~8%
- Prefer not to say: ~1%

Age Groups:
- Youth (18-25): ~44%
- Young Adult (26-34): ~32%
- Adult (35-44): ~18%
- Mature Adult (45-51): ~6%

Marital Status:
- Married: 42%
- Single: 37%
- Divorced: 20%

Education:
- Bachelor's: 39%
- Master's: 27%
- High School: 18%
- No Formal Qualification: 12%
- Doctorate: 3%

Departments:
- Technology: 65%
- Sales: 30%
- Human Resources: 4%

Attrition Rate:
- Total attrition: ~16% (237 employees)
- Most attrition occurred in 2021 and 2022.

QUESTIONS TO BE ASKED

1. Ethnicity by Employees

2. What is the average number of training opportunities taken in each department ?

3. Count Employees by Job Role

4. How does the level of education affect salary?

5. Attrition Rate by Education Level

6. Attrition rate by Marital status

7. Average Salary by Department

8. the correlation between years at the company, years in the most recent role, years since last promotion, and years with the current manager.

9. Distribution of Education Field.

10. Distribution of Marital Satus

وزارة الاتصـــــــالات
وتكنولوجيا المعلومات

KEY INSIGHTS

1. High Attrition in Recent Years: Around 16% of employees left the company, with the highest rates in 2021 and 2022. This highlights a recent retention challenge, especially post-pandemic.

2. Salary Inequality Noted: Salaries are significantly skewed, with a few outliers (mainly managers/seniors) earning much more than the majority. This may cause dissatisfaction among lower-level staff.

3. Promotion and Career Growth Concerns: 75% of employees haven't been promoted in the last 6 years, and most haven't transferred roles in more than 4 years, indicating limited internal mobility and career advancement.

4. Gender & Age Distribution: A relatively balanced gender distribution, with a notable 8% identifying as non-binary. The workforce is mainly young (44% between 18–25), suggesting a need for structured growth and mentoring.

5. Departmental Imbalance: Technology department dominates the workforce (65%), while HR and Sales are underrepresented, possibly requiring hiring focus in non-tech roles for organizational balance.

6. Education Impact on Salary: Higher education levels correlate with better salaries, but further analysis is needed to check for fairness and consistency across roles.

7. Marital Status & Attrition: Single employees show higher attrition rates compared to married ones, potentially indicating different stability or work-life balance expectations.

VISUAL

- Ben Hired

- Hired by Year

- Department

- location

- Gender

- Map Stats

- Age vs Education

- Age group

- Education Level

- Education vs Performance

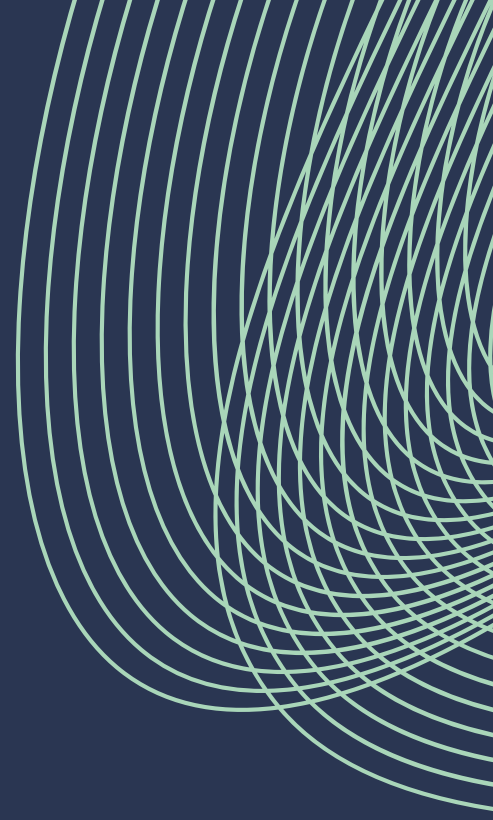- Gender VS education level

- Age And Salary

وزارة الاتصـــــــالات
وتكنولوجيا المعلومات

FINAL REPORT & PRESENTATION

o Prepared a Detailed report covering the whole process done

o Created a presentation demonstrating the key steps taken in this project.

o Project Files:

- HR Folder (Data Source – Consists of a set of Excel files)
- HR-Final Project-DEPI .ipynb
- HR - Final Project.twbx
- HR- Final Project Report.pdf
- HR- Final Project.pptx
- HR- Final Project proposal.pdf

TEAM ROLE

Member Name	Role	Responsibilities
Mohamed Waled	Data Analyst, Visualization Specialist	Data cleaning, EDA, visualization
Nourhan Farg	Visualization Specialist, Data Analyst,	Creating graphs using Seaborn/Plotly
Basant Yasser	Documentation Lead, Data Analyst.	Preparing report and presentation Visualization
Doha Mostafa	Visualization Specialist, Data Analyst.	EDA, Visualization

THANK YOU !