

Final Project: Business Analytics

Note: Students need to attempt any **one** of the following projects.

Project A: Employee Performance and Retention Analysis

Overview

This project focuses on analysing employee performance and retention patterns within an organisation. By leveraging HR data, the goal is to identify key factors influencing employee productivity, performance, and retention, and provide actionable insights for improving workforce efficiency, reducing turnover, and fostering a positive work environment. The project will employ data analysis techniques, including exploratory data analysis (EDA), trend analysis, and visualisation to help HR managers make data-driven decisions for workforce management.

Problem Statement

The objective of this project is to analyse HR data to assess employee performance, identify trends influencing retention, and provide recommendations for improving employee productivity and reducing turnover. Specifically, the project aims to:

- Identify performance patterns based on factors such as age, education, department, and training.
- Examine retention trends based on employee characteristics such as tenure, performance ratings, and awards.
- Provide recommendations for improving employee engagement and retention strategies by analysing the relationship between key employee metrics.

Dataset: [Employees Dataset](#)

Dataset Information

The dataset includes records of employees with various demographic, performance, and organisational details. Key variables in the dataset are:

- **employee_id**: Unique identifier for each employee
- **department**: The department the employee works in
- **region**: Geographical region of the employee
- **education**: The education level of the employee
- **gender**: Gender of the employee
- **recruitment_channel**: The method by which the employee was recruited
- **no_of_trainings**: Number of training sessions attended by the employee
- **age**: Age of the employee

- **previous_year_rating**: Rating of the employee's performance from the previous year
- **length_of_service**: How long the employee has worked in the company
- **KPIs_met_more_than_80**: Whether the employee met more than 80% of their KPIs
- **awards_won**: Number of awards the employee has won
- **avg_training_score**: Average score of the employee in training programs

Deliverables

1. **Exploratory Data Analysis (EDA)**: Perform a thorough EDA to explore key trends and outliers. Generate descriptive statistics and visualisations to understand patterns in employee performance, retention, and demographics.
2. **Data Preprocessing**: Clean and preprocess the dataset by handling missing values, encoding categorical variables, and ensuring consistency in data formatting.
3. **Key Metrics Analysis**: Analyse key performance and retention metrics such as **length of service**, **average training score**, **awards won**, **previous year ratings**, and **KPIs met**.
4. **Retention Trends Analysis**: Assess retention trends based on **age**, **education**, **department**, and **training**.
5. **Predictive Insights**: Based on the analysis, provide actionable recommendations to HR for improving employee retention and performance management strategies.
6. **Final Report**: A detailed report documenting all steps of the analysis, findings, insights, and recommendations. This will include visualisations to support the conclusions.

Success Criteria

- Clear identification of key factors that influence employee performance and retention within the organisation.
- Effective analysis of the relationships between employee demographics, training, performance, and retention rates.
- Actionable recommendations for improving training programs, employee engagement, and retention strategies.
- A final, well-documented report with clear visualisations that help communicate insights to HR management.

Guidelines:

1. **Data Exploration**: Start with an overview of the dataset, checking for missing values, duplicates, and any anomalies.

2. **Key Metric Identification:** Focus on metrics that impact employee performance (e.g., KPIs, training scores) and retention (e.g., length of service, age, awards).
3. **Trend Analysis:** Use visualisations such as bar charts, histograms, and scatter plots to identify trends across different departments, regions, and training activities.
4. **Actionable Insights:** Provide data-backed recommendations to HR regarding the types of employees most likely to leave the organisation, and suggest improvements to engagement or retention strategies.
5. **Interpretability:** Ensure all insights are presented in a way that is easy to understand for HR and management, supporting decision-making with clear, data-driven evidence.

Tools Required

- **Python:** For data analysis, preprocessing, and visualisation (using libraries such as pandas, matplotlib, seaborn, and scikit-learn).
- **Jupyter Notebook or Excel:** For executing and documenting the analysis.
- **Visualisation Tools:** Tools such as **matplotlib** for creating compelling visual insights from the data.
- **Statistical Analysis:** Use statistical techniques to calculate correlations, averages, and other key metrics.

OR

Project B: Healthcare Analytics for Improving Patient Outcomes

Overview

This project focuses on analysing hospital operations and patient data to uncover trends, evaluate resource utilisation, and provide actionable insights for improving service quality and patient outcomes. By leveraging data analytics techniques, the goal is to optimise hospital processes, improve patient care, enhance operational efficiency, and identify areas for resource optimisation. The analysis will include descriptive statistics, trend analysis, and visualisation to aid decision-making.

Problem Statement

The objective of this project is to analyse hospital and patient data to identify trends and patterns that influence key outcomes, including patient length of stay, billing amounts, resource allocation, and service quality. The project will focus on the following areas:

- Assessing how patient demographics and medical conditions correlate with resource utilisation.
- Analysing billing patterns and identifying cost-saving opportunities.
- Evaluating hospital operations by looking at variables like admission type, discharge dates, and room utilisation.
- Providing insights for hospital management to improve operational efficiency and patient care.

Dataset: [Healthcare dataset](#)

Dataset Information:

The dataset includes hospital and patient records, capturing essential variables related to patient demographics, medical conditions, treatment details, and hospital operations. Sample variables include:

- **Name:** Patient's name
- **Age:** Patient's age
- **Gender:** Patient's gender
- **Blood Type:** Patient's blood type
- **Medical Condition:** Medical condition diagnosed (e.g., Cancer, Diabetes, etc.)
- **Date of Admission:** Date of admission
- **Doctor:** Attending doctor
- **Hospital:** Name of the hospital
- **Insurance Provider:** Insurance provider
- **Billing Amount:** Total billing for treatment
- **Room Number:** Room number during admission
- **Admission Type:** Type of admission (Urgent, Non-Urgent)
- **Discharge Date:** Discharge date
- **Medication:** Prescribed medication
- **Test Results:** Test results for the patient

Deliverables

1. **Exploratory Data Analysis (EDA):** Conduct an in-depth EDA to uncover key trends, identify outliers, and visualize the relationships between hospital variables and patient outcomes.

2. **Data Preprocessing:** Clean the data by handling missing values, transforming categorical variables, and standardising formats to ensure consistency in analysis.
3. **Trend Analysis:** Use descriptive statistics and visualisation techniques to analyse trends related to patient length of stay, admission type, billing amounts, and resource utilisation.
4. **Key Metrics:** Analyse and report on important metrics, such as:
 - Average patient billing amounts
 - Room utilisation rates
 - Length of stay analysis
 - Admission type distribution (urgent vs. non-urgent)
5. **Recommendations:** Provide actionable recommendations based on the data analysis to improve hospital efficiency, reduce costs, and enhance patient care.
6. **Final Report:** A comprehensive report documenting the methodology, insights, and actionable recommendations, including visualisations to support conclusions.

Success Criteria

- Clear insights into hospital operations and patient outcomes, with actionable recommendations for improvement.
- Strong data visualisations that help communicate trends and findings effectively. Effective identification of key metrics (e.g., billing patterns, length of stay, room utilisation) that can help hospital management optimise resources and improve services.
- A final report that includes well-documented analysis, code (if applicable), and visualisations.

Guidelines

1. **Data Exploration:** Start with a comprehensive analysis of the dataset, including identifying missing values, outliers, and exploring correlations between variables.
2. **Key Metrics Analysis:** Focus on metrics that drive operational decisions (e.g., cost per patient, length of stay, resource utilisation).
3. **Visualisation:** Use bar charts, histograms, and scatter plots to visualise trends, distributions, and correlations.
4. **Recommendations:** Based on the data analysis, make recommendations on how hospitals can improve efficiency, reduce unnecessary costs, and improve patient care and satisfaction.

Tools Required

- **Python:** For data analysis, cleaning, and visualisation (using libraries such as pandas, matplotlib, seaborn, etc.).

- **Jupyter Notebook or Excel:** For running and documenting the analysis, or any other preferred IDE for performing data analytics.
- **Statistical Analysis:** Use statistical tools for calculating key metrics such as averages, medians, and correlations.