HR Analytics To Track Employee Performance

L. Rajashree[1], Dr. S. Prakasam[2]
[1]Dept of Computer Science and Applications
[2]Associate Professor, Dept of Computer Science and Applications
[1, 2] SCSVMV University

Abstract- *Human Resource (HR) Analytics is transforming workforce management by enabling data-driven decision-making. This paper presents an HR Analytics system designed to track and analyze employee performance using a structured dataset comprising metrics such as department, job role, age, experience, training hours, education, and performance ratings. Developed using Python, Jupyter Notebook, and Power BI, the system includes modules for data preprocessing, exploratory data analysis (EDA), performance metric evaluation, KPI dashboard creation, and result interpretation. The resulting dashboard provides HR professionals with actionable insights to monitor performance trends, identify top performers, and address under performance.Experimental results demonstrate the system's ability to visualize key HR metrics, such as KPI achievement rates and training effectiveness, with departments like Analytics and R&D achieving average training scores of 84.56 and 84.43, respectively. The system offers a scalable, user-friendly solution to replace traditional manual performance tracking, enhancing organizational efficiency and strategic decision-making.*

Keywords- HR Analytics, Employee Performance, Data Visualization, KPI Dashboard, Python, Power BI, Human Resource Management

I. INTRODUCTION

In today's competitive business landscape, organizations increasingly rely on data-driven insights to optimize workforce performance and achieve strategic goals. Human Resource (HR) Analytics, or People Analytics, involves the collection, analysis, and interpretation of HR data to enhance employee productivity, retention, and organizational outcomes. Traditional performance tracking systems often depend on manual processes or subjective evaluations, which are prone to bias, inefficiency, and lack of real-time insights. These limitations hinder proactive decision-making in areas such as talent management, training allocation, and workforce planning.

This research proposes an HR Analytics system to track employee performance using a data-driven approach. The system leverages Python for data processing, Jupyter Notebook for analysis, and Power BI for interactive visualization. It analyzes employee metrics, including age, experience, education, department, performance ratings, and training scores, to provide actionable insights. By addressing the shortcomings of traditional systems, this solution empowers HR professionals to make informed decisions, optimize training programs, and enhance overall organizational performance.

The motivation for this study stems from the growing significance of HR Analytics in modern organizations, as emphasized by Panwar and Chaudhary (2023), who highlight its transformative potential in HR management. This paper outlines the development process, key findings, and proposed improvements for the system, focusing on its practical applicability in real-world HR management scenarios.

II. IDENTIFY,RESEARCHANDCOLLECT IDEA

The development of this HR Analytics system was motivated by the need to address inefficiencies in traditional employee performance tracking methods, which often rely on manual reviews and lack predictive capabilities. A comprehensive review of existing HR systems revealed that many organizations still use spreadsheets or basic software for performance management, resulting in delayed feedback and limited analytical depth.

To design an effective solution, we explored academic literature and industry practices in HR Analytics. Key references, such as Panwar and Chaudhary (2023), highlighted the potential of data-driven tools to improve decision-making. Platforms like Kaggle were used to source relevant employee datasets, which included metrics such as KPIs met, training scores, and previous year ratings. Additionally, industry tools like Power BI and Python libraries (Pandas, NumPy, Matplotlib, Seaborn) were identified as suitable for building a robust analytics system.

The proposed system integrates data preprocessing, analysis, and visualization to create a user-friendly dashboard. It combines Python's analytical capabilities with Power BI's visualization strengths to provide HR

professionals with a scalable solution for performance tracking. The idea evolved into a modular system that processes structured employee data, identifies performance trends, and presents insights through interactive visualizations, making it accessible to users with minimal technical expertise.

III. WRITEDOWNYOURSTUDIESAND FINDINGS

This research aimed to develop a real-time HR Analytics system to track employee performance, replacing manual processes with a data-driven approach. The project was divided into five key modules: data preprocessing, exploratory data analysis (EDA), performance metric evaluation, KPI dashboard creation, and result interpretation. The system was implemented during an internship at VCodez from February to April 2025.

1.Dataset Collection

"Uncleaned_employees_final_dataset.csv," contained employee metrics such as department, gender, age, education, recruitment channel, training scores, KPIs met (>80%), previous year ratings, and length of service. The dataset comprised 17,415 records after preprocessing, covering diverse departments (e.g., Analytics, R&D, Technology) and regions.

2.Data Preprocessing

Preprocessing was performed using Python's Pandas library in Jupyter Notebook. Key steps included:

Loading Data: Imported using pd.read_csv()

Removing Duplicates: Eliminated duplicate rows, reducing the dataset to 17,415 unique records.

Handling Missing Values: Removed rows with missing or irrelevant data in numeric columns (e.g., age, training scores).
Data Validation: Ensured realistic ranges for columns (e.g., age: 20–60 years, training scores: 30–100, KPIs: binary 0/1).
Data Transformation: Standardized formats and derived new features, such as tenure categories

3. Exploratory Data Analysis (EDA)

EDA was conducted to extract key HR metrics:

Average Age by Department and Gender: Ranged between 30–40 years, with balanced distribution in Procurement and Operations but a slight skew toward younger males in R&D and Analytics.
Training Analytics: (84.56), R&D (84.43), and Technology (79.81) led in average training scores.
Award-Winning Employees: Region 5 had the highest percentage (6%), while Region 18 had none.
KPIs Met (>80%): Analyzed by recruitment channel and education level, with the "2K" channel and Bachelor's degree holders showing the highest counts.

4.Dashboard Development

The preprocessed data was imported into Power BI to create an interactive dashboard with the following visualizations:

Histogram: Average age by department and gender.
Area Chart: Percentage of award-winning employees by region.
Pie Chart: Top three departments by average training scores (Analytics: 33.99%, R&D: 33.99%, Technology: 32.02%).
Area Chart: Average length of service for employees with >3 trainings, grouped by department and gender.
Heatmap: Top five regions with highest average previous year ratings (Regions 25, 23, 4, 28, 8).
Bar Chart: Employees meeting >80% KPIs by recruitment channel and education level.

The dashboard included slicers and filters for dynamic exploration (e.g., by department or gender) and was published to the Power BI Service for stakeholder access.

5.Performance Evaluation

The system successfully automated performance tracking, reducing manual effort and enabling real-time insights. Key findings include:

High training scores in Analytics and R&D indicate effective training programs.
Region 5's high award rate suggests strong performance recognition practices.
Female employees in R&D and Technology showed slightly higher tenure post-training, indicating better retention.

6.Implementation Details

The system was developed using Python 3.9, Jupyter Notebook, and Power BI. Python libraries included Pandas, NumPy, Matplotlib, and Seaborn for data processing and initial visualization. Power BI was used for final dashboard creation, leveraging its ETL (extract, transform, load) capabilities and interactive visualization features.

IV. GETPEERREVIEWED

1. Limited Predictive Capabilities
The system focuses on descriptive analytics but lacks predictive models for forecasting performance trends.

2.Dataset Scope
The dataset is comprehensive but lacks real-time data integration from live HR systems.

3.Scalability Concerns
Performance on large datasets (>100,000 records) was not tested.

4.User Training
The dashboard assumes basic technical knowledge, which may limit adoption by non-technical HR staff.

5.Alert Mechanisms:
No automated alerts for critical performance drops or anomalies.

6.Integration Gaps
Limited integration with other HR tools (e.g., payroll, recruitment systems).

7.Diversity in Metrics
Additional metrics, such as employee engagement or turnover risk, could enhance insights.

8.Real-World Validation
The system was tested in a controlled environment, not in a live organizational setting.

9.Explainability
The dashboard lacks explanations for how metrics are derived, which may reduce trust among users.

10.Mobile Accessibility
No mobile version of the dashboard was developed.

V. IMPROVEMENT AS PER REVIEWER COMMENTS

1.Predictive Analytics
Incorporate machine learning models (e.g., Logistic Regression, Random Forest) to predict performance trends and identify potential high-performers.

2.Real-Time Integration
Enable real-time data updates via API integration with HR systems.

3.Scalability Testing
Test the system on larger datasets to ensure performance efficiency.

4.User Training Materials
Develop tutorials or tooltips to guide non-technical HR users..

5.Custom Alerts
Add automated alerts for significant performance drops or anomalies (e.g., sudden decline in training scores).

6.System Integration
Integrate with payroll and recruitment tools for a holistic HR analytics platform.

7.Expanded Metrics
Include sentiment analysis on employee feedback to assess engagement and turnover risk.

8.Real-World Testing
Conduct pilot testing in an organizational setting to validate effectiveness.

9.Explainability Tools
Use tools like SHAP to provide feature importance and metric explanations.

10.Mobile Version
Develop a mobile-compatible dashboard using Power BI's mobile app features.

VI. CONCLUSION

This study successfully developed an HR Analytics system to track employee performance, addressing the inefficiencies of traditional manual methods through a data-driven approach. By leveraging Python for data processing, Jupyter Notebook for analysis, and Power BI for interactive visualization, the system processes a comprehensive dataset encompassing metrics such as age, experience, education, department, performance ratings, and training scores. The resulting interactive dashboard provides HR professionals with actionable insights, enabling real-time monitoring of performance trends, identification of top performers, and targeted interventions for underperformance. Experimental results underscore the system's efficacy, with departments like Analytics and R&D achieving average training scores of 84.56 and 84.43, respectively, and high KPI achievement rates (>80%) among employees recruited through the "2K" channel. These findings highlight the effectiveness of targeted training programs and recruitment strategies, offering a scalable solution to enhance organizational efficiency.

The system's automation and visualization capabilities reduce bias, streamline performance tracking, and support strategic workforce planning. By replacing subjective evaluations with objective metrics, it empowers organizations to optimize training allocation, improve retention, and align workforce capabilities with business goals. The dashboard's user-friendly interface, with features like slicers and filters, ensures accessibility for HR professionals with varying technical expertise, making it a practical tool for diverse organizational contexts.

Despite its strengths, the system faces limitations, including a lack of predictive analytics and real-time data integration, which restrict its ability to forecast performance trends or adapt to live HR systems. Peer feedback also highlighted scalability concerns and the need for mobile accessibility. Future enhancements include incorporating machine learning models (e.g., Random Forest) for predictive insights, enabling API-driven real-time data updates, and developing a mobile-compatible dashboard. Additionally, integrating metrics like employee engagement or turnover risk could further enrich insights. Pilot testing in real-world organizational settings will validate the system's effectiveness and scalability. Ultimately, this HR Analytics system represents a significant step toward modernizing HR management, with the potential to transform workforce strategies through data-driven decision-making.

APPENDIX

The appendix includes supplementary information that supports the research, such as system configurations, software tools used, code snippets, and design diagrams referenced throughout the study. This section is useful for readers who want to replicate or build upon the system.

A. System Configuration

Hardware:

- Processor: Intel Core i5 / i7 (64-bit)
- RAM: 8 GB (minimum 4 GB)
- Hard Disk: 500 GB+

Software:

- Operating System: Windows 10 (64-bit)
- Programming Language: Python 3.9
- IDEs: Jupyter Notebook
- Libraries: Pandas, Numpy, Matplotlib, Seaborn
- Visualization Tool: Power BI
- Database: MySQL, Excel

B. Functional Modules Overview

1. Data Preprocessing

Load, clean, and validate employee dataset using Pandas.

2. Exploratory Data Analysis

Extract HR metrics (e.g., KPIs, training scores) using Python.

3. Dashboard Development Create interactive visualizations in Power BI.
4. Result Interpretation

Analyze trends and provide actionable insights

C. Code Snippet (Example: Data Preprocessing)

```
import pandas as pd
# Load dataset
df =pd.read_csv('Uncleaned_employees_final_dataset.csv')
# Remove duplicates
df = df.drop_duplicates()
print("Number of rows after removing duplicates:", len(df))
# Convert numeric columns and remove invalid data
numeric_columns = ['no_of_trainings', 'age','previous_year_rating', 'length_of_service', 'awards_won', 'avg_training_score']
df[numeric_columns]=df[numeric_columns].apply(pd.to_numeric,errors='coerce')
df = df.dropna(subset=numeric_columns)
# Validate ranges
df = df[(df['age'] >= 20) & (df['age'] <= 60)]
df = df[(df['previous_year_rating'] >= 1) & (df['previous_year_rating'] <= 5)]
df = df[(df['avg_training_score'] >= 30) & (df['avg_training_score'] <= 100)]
# Export cleaned dataset
df.to_excel('cleaned_employees_dataset.xlsx', index=False)defcompute_ear(eye):
```

D. Dashboard Visualizations

Histogram: Average age by department and gender.
Pie Chart: Top three departments by training scores.
Heatmap: Top five regions by previous year ratings.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my project guide, Dr. S. Prakasam, for his invaluable guidance, constant encouragement, and dedicated support throughout the duration of this research work. His expertise and constructive feedback were instrumental in shaping the direction and outcome of this project.

I also extend my thanks to the faculty members of the Department of Information Technology for providing the

necessary facilities and a conducive environment for research. Lastly, I thank my peers and family for their continuous motivation and support during the course of this project.

REFERENCES

[1] [1] Panwar, S., & Chaudhary, T. (2023). HR Analytics: Transforming Human Resource Management through Data-Driven Insights. International Journal of Human Resource Studies.

[2] [2] Mishra, S. N., Lama, D. R., & Pal, Y. (2016). Human Resource Predictive Analytics (HRPA) for HR Management in Organizations. International Journal of Scientific & Technology Research, 5(5), 33-35.

[3] [3] Hoffman, M., &Tadelis, S. (2018). People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis. National Bureau of Economic Research.

[4] [4] Frye, A., Boomhower, C., Smith, M., Vitovsky, L., & Fabricant, S. (2018). Employee Attrition: What Makes an Employee Quit? MU Data Science Review, 1(1).

[5] [5] RabiyathulBasariya, S., &Rzgar Ahmed, R. (2019). A Study on Attrition Turnover Intentions of Employees. International Journal of Civil Engineering and Technology (IJCIET), 10(1), 2594-2601.

[6] [6] Halkos, G., &Bousinakis, D. (2017). The Effect of Stress and Dissatisfaction on Employees During Crisis. Economic Analysis and Policy, 55(C), 25-34.

[7] [7] Glavas, A., &Willness, C. (2020). Employee (Dis)engagement in Corporate Social Responsibility. In Employee Engagement in Corporate Social Responsibility (pp. 10-27). Sage Publications Ltd.

[8] [8] Yadav, S., Jain, A., & Singh, D. (2018). Early Prediction of Employee Attrition using Data Mining Techniques. IEEE 8th International Advance Computing Conference (IACC), 349-354.