**Employee Performance Analysis – HR Analytics Project Report**

**1. Introduction**

Employee performance is one of the most critical factors that determine an organization’s success. Understanding what influences performance helps companies make informed decisions on training, motivation, and workforce planning.

In today’s data-driven HR environment, analytics provides valuable insights into workforce behavior, engagement, and productivity. This project applies **HR analytics** to explore the factors that affect employee performance — focusing on **training quality, recognition, education, and tenure** — using data from the *Employees Performance for HR Analytics* dataset on Kaggle.

**2. Problem Statement**

Many organizations struggle to identify the specific factors that contribute to high employee performance. Without data-based insights, HR decisions on training, recruitment, and recognition can be inconsistent and ineffective.

This project aims to answer the question:

“What key factors drive employees to meet or exceed their KPI performance targets?”

This analysis seeks to uncover patterns that explain why some employees consistently achieve their goals while others do not, providing clear recommendations for HR improvement.

**SMART framing:**

* **Specific:** Identify variables that most influence KPI success.
* **Measurable:** Use quantifiable HR metrics such as training scores and KPI rates.
* **Achievable:** Analyze a real-world dataset with 17K records.
* **Relevant:** Focus on actionable HR performance insights.
* **Time-bound:** Project completed within the capstone term timeline.

**3. Objectives**

1. Evaluate the relationship between **training effectiveness** and KPI success.
2. Assess how **recognition and past performance ratings** impact outcomes.
3. Identify **top-performing departments** and analyze performance variations.
4. Examine how **demographic factors** (age, education, service length) affect performance.
5. Compare **recruitment channels** to determine which sources attract high-performing employees.
6. Assess **gender fairness** in performance outcomes.
7. Provide **data-backed HR recommendations** to improve overall workforce effectiveness.

**4. Target Audience**

The project is primarily aimed at:

* **HR Leaders and Managers** – to enhance decision-making on employee development.
* **Business Executives** – to understand workforce performance trends.
* **Talent Development Teams** – to design training programs aligned with performance needs.

**Fictional example:**

*Fatima, an HR Manager, wants to use data analytics to understand which training programs and recruitment strategies drive high-performing employees.*

**5. Dataset Overview**

**Dataset Source:** [Employees Performance for HR Analytics – Kaggle](https://www.kaggle.com/datasets/sanjanchaudhari/employees-performance-for-hr-analytics?utm_source=chatgpt.com)

**Dataset Profile:**

* **Total Records:** 17,417 employees
* **Total Columns:** 14 original + new derived features
* **Key Columns:**
  + employee\_id, department, education, gender, age, length\_of\_service
  + no\_of\_trainings, previous\_year\_rating, avg\_training\_score
  + recruitment\_channel, awards\_won, KPIs\_met\_more\_than\_80

**Data Types:**

* Categorical: Department, Gender, Education, Recruitment Channel, Awards Won
* Numerical: Age, Length of Service, Avg Training Score, No. of Trainings, Previous Year Rating
* Binary: KPIs\_met\_more\_than\_80 (0 or 1)

**6. Data Handling & Preprocessing**

**Data Cleaning Steps:**

* Removed duplicates and verified unique employee\_id values.
* Handled missing data in previous\_year\_rating by imputing based on department mean.
* Filled nulls in education and recruitment\_channel using mode imputation.
* Verified and standardized categorical entries (e.g., “Bachelors”, “Masters & above”).

**New Derived Columns:**

1. **Age Category:**
   * <28 (Junior), 28–35 (Young), 36–45 (Mid-career), 46+ (Senior)
2. **Service Category:**
   * 0–2 years (New), 3–7 years (Mid-tenure), 8+ years (Experienced)
3. **Training Score Category:**
   * Low (≤50), Medium (51–75), High (>75)
4. **Awards Won (Yes/No):** Converted binary flag into a readable category.
5. **KPI Percentage Column:** Converted KPIs\_met\_more\_than\_80 binary into % form.

**Software & Tools Used:**

* **Python:** pandas, matplotlib for EDA and visualization
* **Excel:** Pivot tables and charts for validation
* **PowerPoint:** Dashboard storytelling

**7. Analysis and Findings**

**Objective 1 – Training Effectiveness**

* Employees with **high training scores (>75)** had the highest KPI success rate.
* Performance improved with **more training sessions**, but plateaued after 3–4.

**Objective 2 – KPI Performance & Recognition**

* Only **35.9%** of employees met their KPI targets.
* Those who received **awards or recognition** showed higher KPI achievement.
* High **previous-year ratings** correlated with consistent KPI success.

**Objective 3 – Departmental Insights**

* **Analytics, R&D, and Operations** departments performed best overall.
* **Sales & Marketing** had the largest workforce but moderate KPI success.

**Objective 4 – Demographic Impact on KPI**

* **Age 28–35** (young professionals) achieved the best KPI success rates.
* **Mid-tenure employees (3–7 years)** performed strongest, balancing experience and engagement.
* **Master’s degree holders** showed slightly higher KPI rates than Bachelor’s graduates.

**Objective 5 – Recruitment Channels**

* Most hires came from the **“Other”** channel, but **referral hires** performed best.
* Referrals had higher KPI success due to stronger cultural fit and pre-screening.

**Objective 6 – Diversity & Fairness**

* The workforce is **70% male and 30% female**.
* KPI performance between genders was nearly identical — indicating fairness in performance evaluation.

**Objective 7 – Summary Dashboard Highlights**

* **Total Employees:** 17,417
* **KPI Success:** 35.9%
* **Avg Age:** 34.8 years | **Avg Training Score:** 63.2
* **Top Departments:** Sales, Operations, Procurement
* **Awards:** 2.3% of employees
* **Avg Tenure:** 5.8 years

**8. Recommendations**

* **Enhance training programs** with role-specific learning to boost KPI success.
* **Expand recognition and rewards** to motivate consistent high performance.
* **Focus on mid-tenure retention** to leverage high-performing groups.
* **Develop gender diversity initiatives** to balance workforce representation.
* **Prioritize referral and sourcing channels** to improve quality of hires.

**9. Limitations & Assumptions**

* Missing data in previous\_year\_rating may have slightly impacted accuracy.
* Dataset lacks details on **job role, salary, and engagement surveys**, which could improve depth.
* KPI definition is limited to “>80% target met,” not continuous or department-adjusted.
* It is assumed that training scores and KPI outcomes are measured consistently across departments.

**10. References**

* Kaggle Dataset: [Employees Performance for HR Analytics](https://www.kaggle.com/datasets/sanjanchaudhari/employees-performance-for-hr-analytics?utm_source=chatgpt.com)
* HR Analytics Research: CIPD Learning and Performance Frameworks
* Python Libraries: pandas, matplotlib
* General Assembly Data Analytics Course Materials