

HR Employee Master & Training Dataset.

Dataset Overview

Item	Details
Dataset Title	HR Employee Master & Training Dataset
Domain / Theme	Human Resources – Employee Demographics, Performance, Training, Payroll
Source	Internal HR System Export (HR_DataAnalysiss.xlsx)
File Type	Excel (.xlsx)
No. of Rows	2845
No. of Columns	28

Data Structure Summary

Column Name	Data Type	Dimension / Measure	Description
Employee ID	Integer	Dimension	Unique employee identifier
StartDate	Date	Dimension	Employee joining date
Title	String	Dimension	Job title / designation
BusinessUnit	String	Dimension	Functional business unit
EmployeeStatus	String	Dimension	Active / Terminated status
EmployeeType	String	Dimension	Full-Time / Part-Time / Contract
PayZone	String	Dimension	Salary pay-grade category
EmployeeClassificationType	String	Dimension	Employee classification (Temp/PT/FT)
DepartmentType	String	Dimension	Department the employee belongs to
Division	String	Dimension	Division under the business unit
Gender	String	Dimension	Employee gender
MaritalStatus	String	Dimension	Employee marital status
Functional Specialization	String	Dimension	Area of expertise

Column Name	Data Type	Dimension / Measure	Description
Location State	String	Dimension	Work location state
Salary	Float	Measure	Monthly salary
Annual Salary CTC	Float	Measure	Annual compensation
Bonus	Float	Measure	Bonus amount
Termination Year	Integer / Float	Dimension	Year of termination (if applicable)
Reason For Termination	String	Dimension	Reason for employee exit
Engagement Score	Integer	Measure	Employee engagement rating
Performance Score	String	Dimension	Performance category (Meets, Exceeds, PIP)
Absences	Integer	Measure	Number of absences
Training Program Name	String	Dimension	Name of training attended
Training Type	String	Dimension	Internal / External
Training Outcome	String	Dimension	Completed / Failed / Incomplete
Training Duration (Days)	Integer	Measure	Duration of training
Training Cost	Float	Measure	Cost of training
Age	Integer	Measure	Age of employee

Data Quality Assessment

Aspect Checked	Observations	Action Needed
Missing Data	Missing values in Termination Year, Termination Reason, some Training Outcomes ,Tenure years	Impute “Not Applicable” for active employees; clean inconsistent blanks
Duplicates	No duplicate Employee IDs observed	No action required
Outliers	Few extremely high Salary, Bonus, and Training Cost values	Investigate with HR or apply capping
Incorrect Data Types	Termination Year appears as float; some numeric values formatted as text	Convert to correct numeric/type formats
Consistency Issues	Training Outcome labels vary; Division names inconsistent in some entries	Standardize naming formats

Key Metrics & KPIs Identified

KPI	Why It Matters
Employee Headcount	Shows total workforce size
Attrition Rate	Measures employee turnover
Termination Count by Reason	Helps identify HR issues & causes
Average Salary by Business Unit	Supports payroll planning
Performance Score Distribution	Identifies high/low performers
Training Completion Rate	Helps evaluate training effectiveness
Average Engagement Score	Measures employee morale
Absence Rate	Helps understand attendance trends

Assumptions / Risks

- Dataset may not include latest updates for training or performance.
- Missing termination fields likely indicate active employees.
- Salary and cost outliers may distort KPI results if not cleaned.
- Training records may not apply to all employees (some never trained).