

# Analysis

## Initial Observations:

- Total Records: ~1200 rows
- Key columns:
  - Employee ID, Age, Gender, Education Background, Joining Year, Time of Service, Department, Past Year Rating, Performance Rating, Work-Life Balance, Training Score, Awards, Salary Hike, etc.

## Data Cleaning:

- Removed null/missing values (some columns had NaNs).
- Converted date-related fields for time-based analysis.
- Ensured numerical encoding for categorical fields for ML processing.

## Exploratory Data Insights:

- **Correlation Observed:**
  - Training Score and Awards have positive correlation with Performance Rating.
  - Departments like R&D and Technology have higher average performance.
  - Work-Life Balance scores affect employee ratings significantly.
- **Gender Bias:**
  - No significant bias observed in performance across genders.
- **Experience vs Rating:**
  - Employees with >5 years of service tend to have slightly higher performance scores.

## Machine Learning Model Insights:

- Models Tested: Logistic Regression, Random Forest, XGBoost.
- Best Performing: **Random Forest**
  - Accuracy: ~87%
  - Important Features: Training Score, Time of Service, Past Rating, Work-Life Balance, Awards