

# HR Attrition Analytics: Unlocking Insights on Employee Satisfaction & Performance

DEPI Initiative - Data Analyst Specialist Track

Group: ALX3\_DAT1\_G2

Under the Supervision of: Eng./Dina Ezzat



# Team Members

## (Group 3)

- Moataz Tammam Abdelaziz —> Data Cleaning & Analysis
- Yasmeen Farouq Saleh —> Data Cleaning & Analysis
- Muhammed Nasseim El-Saied —> Data Cleaning & Analysis
- Sara Ayed Wahba —> Data Visualization & Presentation Preparation
- Malak Tarek Muhammed —> Data Visualization & Presentation Preparation

# Project Overview

This project focuses on uncovering the underlying factors that lead to ***Employee Attrition*** within the company. Our analysis examines three key evaluation dimensions:

- **Employees' Satisfaction and Engagement.**
- **Manager Assessments of Employees.**
- **Employees' Self-evaluation.**

These insights are derived from HR dataset across **three** core departments: **Sales**, **Technology**, and **Human Resources**.

# Data Analysis Process

1

**Data Exploration**

2

**Data Cleaning & Preparation**

3

**Data Analysis (EDA)**

4

**Data Visualization**

5

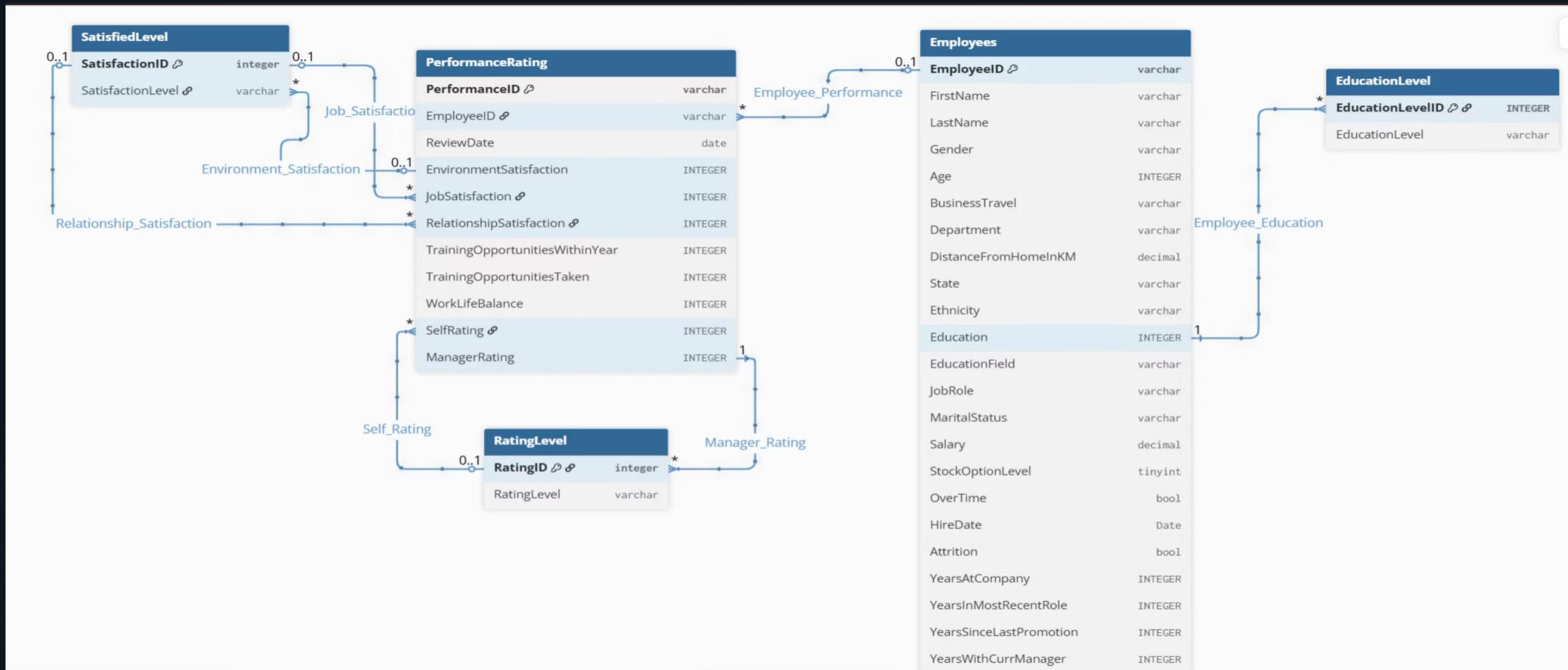
**Data Insights & Reporting**

# Understanding Data

Our HR Dataset is built on a relational database consisting of **five interconnected tables**:

1. Employee table —> Master employee table (Demographics & Employment data, Employee Attrition & tenure fields). **Primary key: EmployeeID**
2. Performance Rating —> Employee Performance & Satisfaction scores per review (ReviewDate) and training participations.  
**Primary Key: PerformanceID, Foreign Key: EmployeeID.**
3. Education Level —> Employees' Educational Background.
  - Lookup table: mapping **EducationLevelID “PK”** to Employees table “**FK: Education**”
4. Rating Level —> Categorical description for each rating.
  - Lookup table: **RatingID “PK”** → **Manager & Self-Rating “FK”** in PerformanceRating table.
5. Satisfied Level —> Categorical description for each Satisfaction Criteria.
  - Lookup table: **SatisfactionID “PK”** → **Environment, Job & Relationship Sats. “FK”** in PerformanceRating table.

# Database Schema



# Data Loading

```
1 -- Data Exploration PerformanceRating  
2 SELECT * FROM PerformanceRating;  
3 -- Data Exploration Employee  
4 SELECT * FROM Employees;  
5 -- Data Exploration SatisfiedLevel  
6 SELECT * FROM SatisfiedLevel;  
7 -- Data Exploration RatingLevel  
8 SELECT * FROM RatingLevel;  
9 -- Data Exploration EducationLevel  
10 SELECT * FROM EducationLevel;
```

# Data Cleaning & Preparation



# Handling Duplicates

```
16 -- Checking Duplicates in Employees table  
17 SELECT EmployeeID, COUNT (*) FROM Employees  
18 GROUP BY EmployeeID  
19 HAVING COUNT(*) > 1;  
20 -- Checking Duplicates in PerPerformanceRating table  
21 SELECT PerformanceID, COUNT (*) FROM PerformanceRating  
22 GROUP BY PerformanceID  
23 HAVING COUNT(*) > 1;
```

- It appears that there are no duplicates in our dataset.

# Handling Null Values

```
24 -- Searching for Missing Values & Removing Trailing Whitespaces at key Fields in Employees table
25 SELECT
26   COUNT(IF(EmployeeID IS NULL OR TRIM(EmployeeID) = '', 1, NULL)) AS Missing_EmpID,
27   COUNT(IF(Gender IS NULL OR TRIM(Gender) = '', 1, NULL)) AS Missing_Gender,
28   COUNT(IF(Age IS NULL OR TRIM(Age) = '', 1, NULL)) AS Missing_Age,
29   COUNT(IF(Department IS NULL OR TRIM(Department) = '', 1, NULL)) AS Missing_Dept,
30   COUNT(IF(Salary IS NULL OR TRIM(Salary) = '', 1, NULL)) AS Missing_Salary,
31   COUNT(IF(YearsAtCompany IS NULL OR TRIM(YearsAtCompany) = '', 1, NULL)) AS Missing_Exp_Yrs
32 FROM Employees
33 -- Searching for Missing Values & Removing Trailing Whitespaces at key Fields in PerformanceRating table
34 SELECT
35   COUNT(IF(PerformanceID IS NULL OR TRIM(PerformanceID) = '', 1, NULL)) AS Missing_PerfID,
36   COUNT(IF(EmployeeID IS NULL OR TRIM(EmployeeID) = '', 1, NULL)) AS Missing_EmpID,
37   COUNT(IF(SelfRating IS NULL OR TRIM(SelfRating) = '', 1, NULL)) AS Missing_Self_Rate,
38   COUNT(IF(ManagerRating IS NULL OR TRIM(ManagerRating) = '', 1, NULL)) AS Missing_Mangager_Rate,
39   COUNT(IF(JobSatisfaction IS NULL OR TRIM(JobSatisfaction) = '', 1, NULL)) AS Missing_Job_Sat,
40   COUNT(IF(TrainingOpportunitiesWithinYear IS NULL OR TRIM(TrainingOpportunitiesWithinYear) = '', 1, NULL)) AS Miss_TrOppWithinYear,
41   COUNT(IF(RelationshipSatisfaction IS NULL OR TRIM(RelationshipSatisfaction) = '', 1, NULL)) AS Missing_Rel_Sat,
42   COUNT(IF(TrainingOpportunitiesTaken IS NULL OR TRIM(TrainingOpportunitiesTaken) = '', 1, NULL)) AS Missing_TrOppTaken
43 FROM PerformanceRating;
```

- ❑ It appears that there are no Null Values in our dataset.

# Checking Invalid Data (Ensure Data Quality)

```
63 -- Checking Data Quality if there are any Invalid Values in Employees Table (Age & Salary)
64 SELECT COUNT ( * ) FROM Employee
65 WHERE Age < 0;
66 SELECT COUNT ( * ) FROM Employee
67 WHERE salary < 0;
68 -- Checking invalid data in PerformanceRating Table
69 SELECT COUNT(*) AS InvalidRatings
70 FROM PerformanceRating
71 WHERE SelfRating NOT BETWEEN 1 AND 5
72     OR ManagerRating NOT BETWEEN 1 AND 5
73     OR EnvironmentSatisfaction NOT BETWEEN 1 AND 5
74     OR JobSatisfaction NOT BETWEEN 1 AND 5
75     OR RelationshipSatisfaction NOT BETWEEN 1 AND 5
76     OR WorkLifeBalance NOT BETWEEN 1 AND 5;
77 -- Identifying Outliers in Employee Table
78 SELECT EmployeeID, Age
79 FROM Employees
80 WHERE Age < 18 OR Age > 60;
```

# Exploratory Data Analysis (EDA)

- To examine & investigate datasets to discover patterns, trends & spot anomalies.

```
86 -- Summary Statistics
87 -- Salary & Employee Tenure by their Education Level
88 SELECT ed.EducationLevel,
89     COUNT(e.EmployeeID) AS cnt,
90     ROUND(AVG(COALESCE(e.Salary,0)), 2) AS Avg_Salary,
91     ROUND(AVG(COALESCE(e.YearsAtCompany,0)), 2) AS Avg_Years_At_Company
92 FROM Employees e
93 LEFT JOIN EducationLevel ed
94     ON e.Education = ed.EducationLevelID
95 GROUP BY ed.EducationLevel
96 ORDER BY Avg_Salary DESC;
97 -- Salary Distribution Statistics
98 SELECT MIN(Salary) AS MinSalary,
99     MAX(Salary) AS MaxSalary,
100    AVG(Salary) AS AvgSalary
101 FROM Employees;
102 -- Age Statistics
103 SELECT MIN(Age) AS MinAge,
104     MAX(Age) AS MaxAge,
105    AVG(Age) AS AvgAge
106 FROM Employees;
```

# Focus Areas of Analysis

1

**Factors influencing  
Employee Satisfaction and  
Attrition**

2

**Overtime Impact on  
Employee WLB &  
Performance**

3

**Correlation between Salary  
and Satisfaction or  
Performance**

4

**Training Impact on Performance & Promotion  
Potential**

5

**Common Characteristics of Attritors**

# Performance Analysis (Performance Correlation with Age)

```
210 -- Correlation
211 -- Performance analysis and Age / Years At Company.
212 -- Employee Performance & Age
213 WITH perf AS (
214     SELECT
215         EmployeeID,
216         AVG(JobSatisfaction) AS Avg_JobSatisfaction,
217         AVG(SelfRating) AS Avg_SelfRating,
218         AVG(ManagerRating) AS Avg_ManagerRating
219     FROM PerformanceRating
220     GROUP BY EmployeeID
221 )
222 SELECT
223     CASE
224         WHEN e.Age < 25 THEN '<25'
225         WHEN e.Age BETWEEN 25 AND 35 THEN '25-35'
226         WHEN e.Age BETWEEN 36 AND 45 THEN '36-45'
227         ELSE '>45'
228     END AS Age_Range,
229     ROUND(AVG(p.Avg_JobSatisfaction),2) AS Avg_JobSatisfaction,
230     ROUND(AVG(p.Avg_SelfRating),2) AS Avg_SelfRating,
231     ROUND(AVG(p.Avg_ManagerRating),2) AS Avg_ManagerRating,
232     COUNT(*) AS Employee_Count
233 FROM Employee e
234 LEFT JOIN perf p
235     ON e.EmployeeID = p.EmployeeID
236 GROUP BY Age_Range
237 ORDER BY Age_Range;
238
```

# Performance Analysis (Performance Correlation with tenure)

```
239 --Employee Performance & Years At Company
240 WITH perf AS (
241     SELECT
242         EmployeeID,
243         AVG(JobSatisfaction) AS JobSatisfaction,
244         AVG(SelfRating) AS SelfRating,
245         AVG(ManagerRating) AS ManagerRating
246     FROM PerformanceRating
247     GROUP BY EmployeeID
248 )
249 SELECT
250     CASE
251         WHEN YearsAtCompany < 2 THEN '<2'
252         WHEN YearsAtCompany BETWEEN 2 AND 5 THEN '2-5'
253         WHEN YearsAtCompany BETWEEN 6 AND 10 THEN '6-10'
254         ELSE '>10'
255     END AS YearsAtCompany_Range,
256     AVG(p.JobSatisfaction) AS Avg_JobSatisfaction,
257     AVG(p.SelfRating) AS Avg_SelfRating,
258     AVG(p.ManagerRating) AS Avg_ManagerRating,
259     COUNT(*) AS Employee_Count
260 FROM Employee e
261 JOIN perf p ON e.EmployeeID = p.EmployeeID
262 GROUP BY YearsAtCompany_Range
263 ORDER BY YearsAtCompany_Range;
264
```

# Factors influencing Employee Satisfaction and Attrition

```
88 -- Identify the factors most associated with Low Employee Satisfaction or High Attrition
89 SELECT
90     e.Department,
91     ROUND(AVG(p.EnvironmentSatisfaction),2) AS Avg_Env_Satisfaction,
92     ROUND(AVG(p.JobSatisfaction),2) AS Avg_Job_Satisfaction,
93     ROUND(AVG(p.RelationshipSatisfaction),2) AS Avg_Relationship_Satisfaction,
94     ROUND(AVG(p.WorkLifeBalance),2) AS Avg_WorkLifeBalance,
95     ROUND(AVG(p.SelfRating),2) AS Avg_SelfRating,
96     ROUND(AVG(p.TrainingOpportunitiesWithinYear),2) AS Avg_Training_Opportunities
97 FROM Employees AS e
98 LEFT JOIN PerformanceRating AS p
99     ON e.EmployeeID = p.EmployeeID
100 GROUP BY e.Department
101 ORDER BY Avg_Job_Satisfaction ASC;
```

# Overtime Impact on Work-Life Balance and Employee Performance:

```
122 -- Determine the Impact of Overtime on Work-Life balance and Overall Performance
123 SELECT
124     e.OverTime,
125     COUNT(DISTINCT(e.EmployeeID)) AS Num_Employees,
126     ROUND(AVG(p.WorkLifeBalance),2) AS Avg_WorkLife_Balance,
127     ROUND(AVG(p.JobSatisfaction),2) AS Avg_Job_Satisfaction,
128     ROUND(ROUND(AVG(p.ManagerRating),2) AS Avg_Manager_Rating,
129     ROUND(AVG(p.SelfRating),2) AS Avg_Self_Rating
130 FROM Employees AS e
131 LEFT JOIN PerformanceRating AS p
132     ON e.EmployeeID = p.EmployeeID
133 GROUP BY e.OverTime
134 ORDER BY Avg_WorkLife_Balance ASC;
```

# Overtime Impact on Employee Performance



## Overtime and Job Satisfaction

Overtime had a noticeable effect on job satisfaction. Employees who **worked extra hours** reported **higher satisfaction** compared to those who did not (3.51 vs 3.38).



## Overtime and Performance

Employees who **did not work overtime** have **higher manager rating** average of 3.47 compared to 3.45 for those who do.

# Relation between Salary and Employee Satisfaction & Performance

```
135 -- Relation between Salary and Employee Satisfaction & Performance
136 SELECT
137     e.Department,
138     ROUND(AVG(e.Salary), 2) AS Avg_Salary,
139     ROUND(AVG(p.ManagerRating), 2) AS Avg_Manager_Rating,
140     ROUND(AVG(p.SelfRating), 2) AS Avg_Self_Rating,
141     ROUND(AVG(p.JobSatisfaction), 2) AS Avg_Job_Satisfaction
142 FROM Employees AS e
143 LEFT JOIN PerformanceRating AS p
144     ON e.EmployeeID = p.EmployeeID
145 GROUP BY e.Department
146 ORDER BY Avg_Salary DESC;
147 SELECT
148     CASE
149         WHEN Salary < 80000 THEN '< 80k'
150         WHEN Salary BETWEEN 80000 AND 120000 THEN '80k-120k'
151         WHEN Salary BETWEEN 120001 AND 180000 THEN '120k-180k'
152         WHEN Salary BETWEEN 180001 AND 300000 THEN '180k-300k'
153         ELSE '> 300k'
154     END AS Salary_Band,
155     COUNT(*) AS Employee_Count
156 FROM Employees
157 GROUP BY Salary_Band
158 ORDER BY Employee_Count DESC;
```

# Training Opportunities Impact on Performance & Promotion Potential

```
207 -- The Impact of Training Opportunities on Performance and Promotion Rate
208 SELECT
209     CASE
210         WHEN p.TrainingOpportunitiesWithinYear > 0 THEN 'Received Training'
211         ELSE 'No Training'
212     END AS Training_Status,
213     COUNT(DISTINCT e.EmployeeID) AS Num_Employees,
214     ROUND(AVG(p.ManagerRating), 2) AS Avg_Manager_Rating,
215     ROUND(AVG(p.SelfRating), 2) AS Avg_Self_Rating,
216     ROUND(AVG(e.YearsAtCompany), 2) AS Avg_Years_At_Company
217 FROM Employees e
218 LEFT JOIN PerformanceRating p
219     ON e.EmployeeID = p.EmployeeID
220 GROUP BY Training_Status
221 ORDER BY Avg_Years_At_Company DESC;
222 SELECT
223     e.Department,
224     ROUND(AVG(p.TrainingOpportunitiesWithinYear), 2) AS Avg_Training_Opportunities_WithinYear,
225     ROUND(AVG(p.TrainingOpportunitiesTaken), 2) AS Avg_Training_Opportunities_Taken,
226     ROUND(AVG(p.ManagerRating), 2) AS Avg_ManagerRating,
227     ROUND(AVG(p.SelfRating), 2) AS Avg_SelfRating,
228     ROUND(AVG(e.YearsAtCompany), 2) AS Avg_Years_At_Company,
229     ROUND(AVG((ManagerRating + SelfRating) / 2.0), 2) AS PromotionPotentialIndex
230 FROM Employees AS e
231 LEFT JOIN PerformanceRating AS p
232     ON e.EmployeeID = p.EmployeeID
233 GROUP BY e.Department
234 ORDER BY Avg_Training_Opportunities_Taken DESC;
```

- **HR Department** has the **lowest training opportunities** (Average: 0.96), but the **highest Performance "Manager Rating"** (Average: 3.49).
- **Sales Department** has **slightly higher training opportunities** (Average: 0.99), but the **lowest Performance** (Average: 3.43).
- **Technology Department** has the **highest training opportunities** (Average: 1.04) and **Manager rating of (Average: 3.47)**.

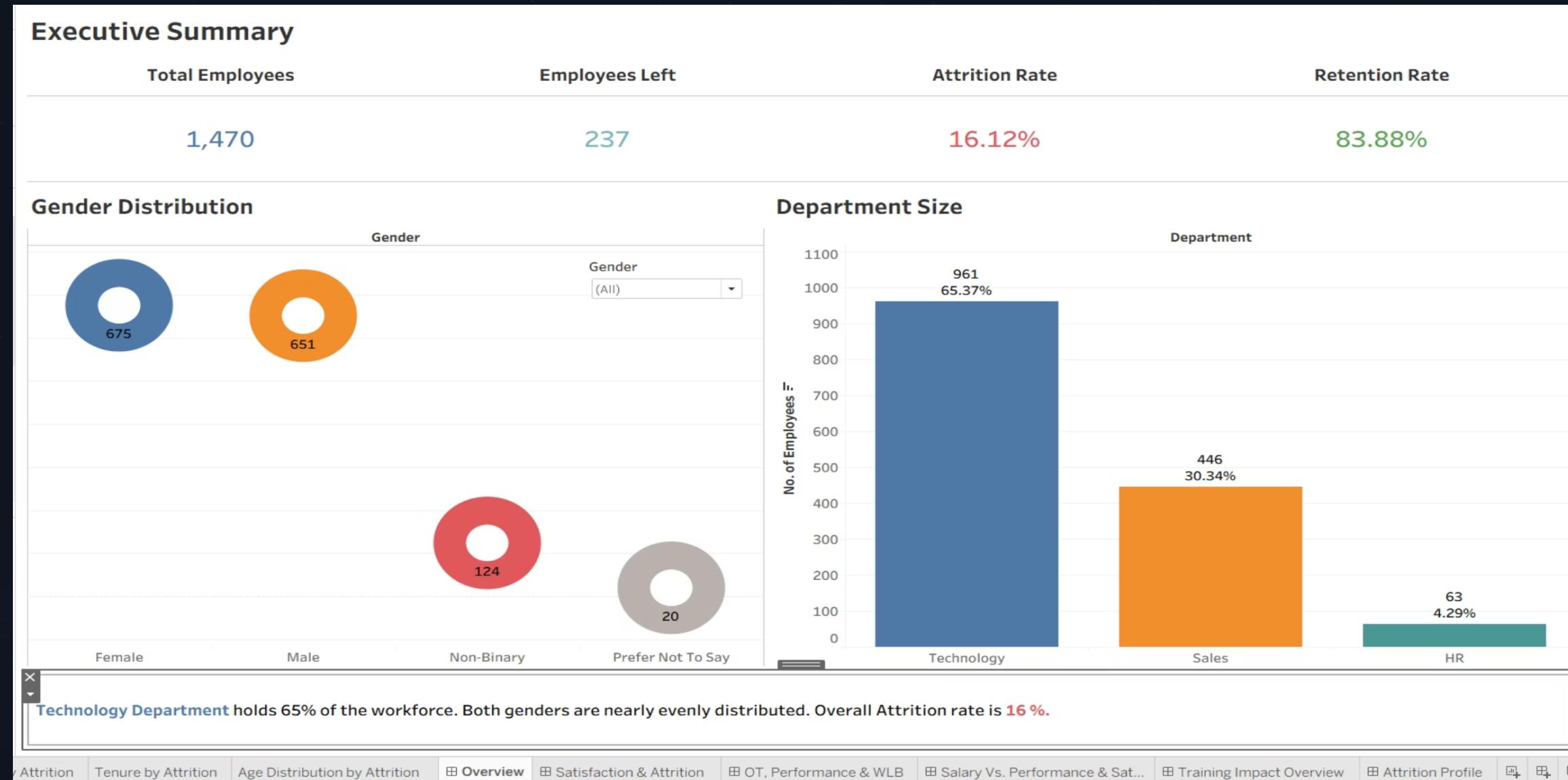
# Common Characteristics of Employees who Left

```
323 -- Common Characteristics of Employees who left (Attrition Rate = Yes)
324 SELECT
325   UPPER(TRIM(e.Department)) AS Department,
326   ROUND(AVG(e.Age), 2) AS Avg_Age,
327   ROUND(AVG(e.Salary), 2) AS Avg_Salary,
328   ROUND(AVG(e.YearsAtCompany), 2) AS Avg_YearsAtCompany,
329   ROUND(AVG(p.JobSatisfaction), 2) AS Avg_JobSatisfaction,
330   ROUND(AVG(p.ManagerRating), 2) AS Avg_ManagerRating,
331   ROUND(AVG(p.WorkLifeBalance), 2) AS Avg_WorkLifeBalance,
332   ROUND(AVG(p.EnvironmentSatisfaction), 2) AS Avg_EnvironmentSatisfaction,
333   ROUND(AVG(p.RelationshipSatisfaction), 2) AS Avg_RelationshipSatisfaction,
334   COUNT(DISTINCT e.EmployeeID) AS Attrition_Count,
335   ROUND(100.0 * COUNT(DISTINCT e.EmployeeID) /
336     (SELECT COUNT(DISTINCT EmployeeID) FROM Employees WHERE LOWER(TRIM(Attrition)) = 'yes'), 2) AS Attrition_Rate
337 FROM Employees e
338 LEFT JOIN PerformanceRating p ON e.EmployeeID = p.EmployeeID
339 WHERE LOWER(TRIM(e.Attrition)) = 'yes'
340 GROUP BY e.Department
341 ORDER BY Attrition_Count DESC;
```

Employees who are **younger, less incentivized**, and have **shorter tenure** tend to **leave at higher rates**.

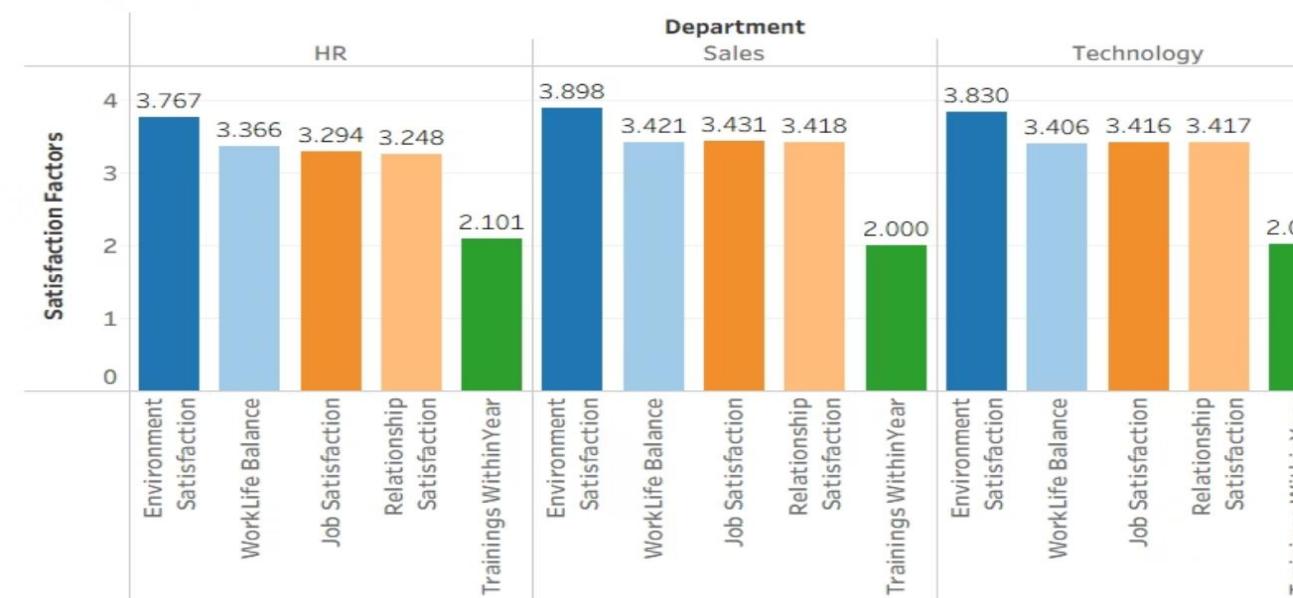
# Data Visualization (Tableau Dashboard)

## An Overview for HR Attrition Project

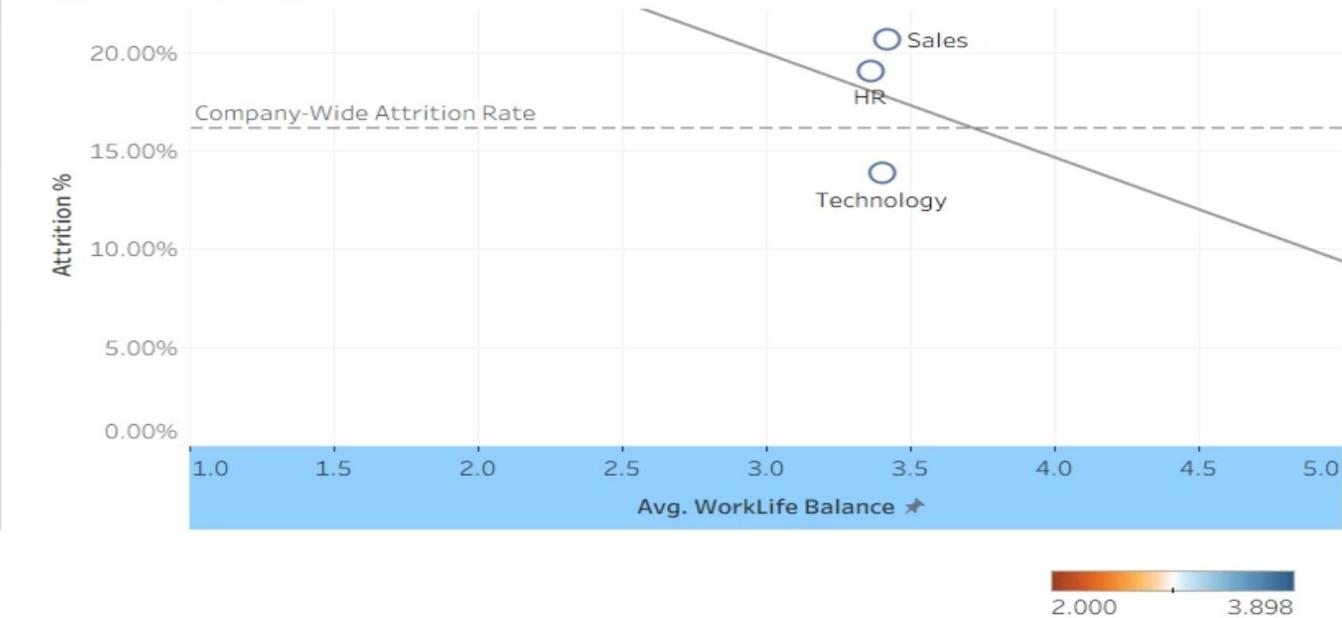


# Satisfaction Drivers Dashboard

## Employee Satisfaction & Attrition Drivers



## WLB Vs. Attrition



## Department Satisfaction Metrics

Department	Environment Satisfaction	Job Satisfaction	Relationship Satisfaction	WorkLife Balance	Trainings WithinYear
HR	3.767	3.294	3.248	3.366	2.101
Sales	3.898	3.431	3.418	3.421	2.000
Technology	3.830	3.416	3.417	3.406	2.016

**Insights:**  
Satisfaction Scores across departments vary only slightly, indicating that Attrition is not driven by satisfaction alone. Other factors such as WLB, Salary, and Overtime appear more impactful.

# Overtime Impact Dashboard

## Impact of Overtime on Attrition & Performance

Total Employees	Employees Left	Attrition Rate	WorkLife Balance
1,470	237	16.12%	3

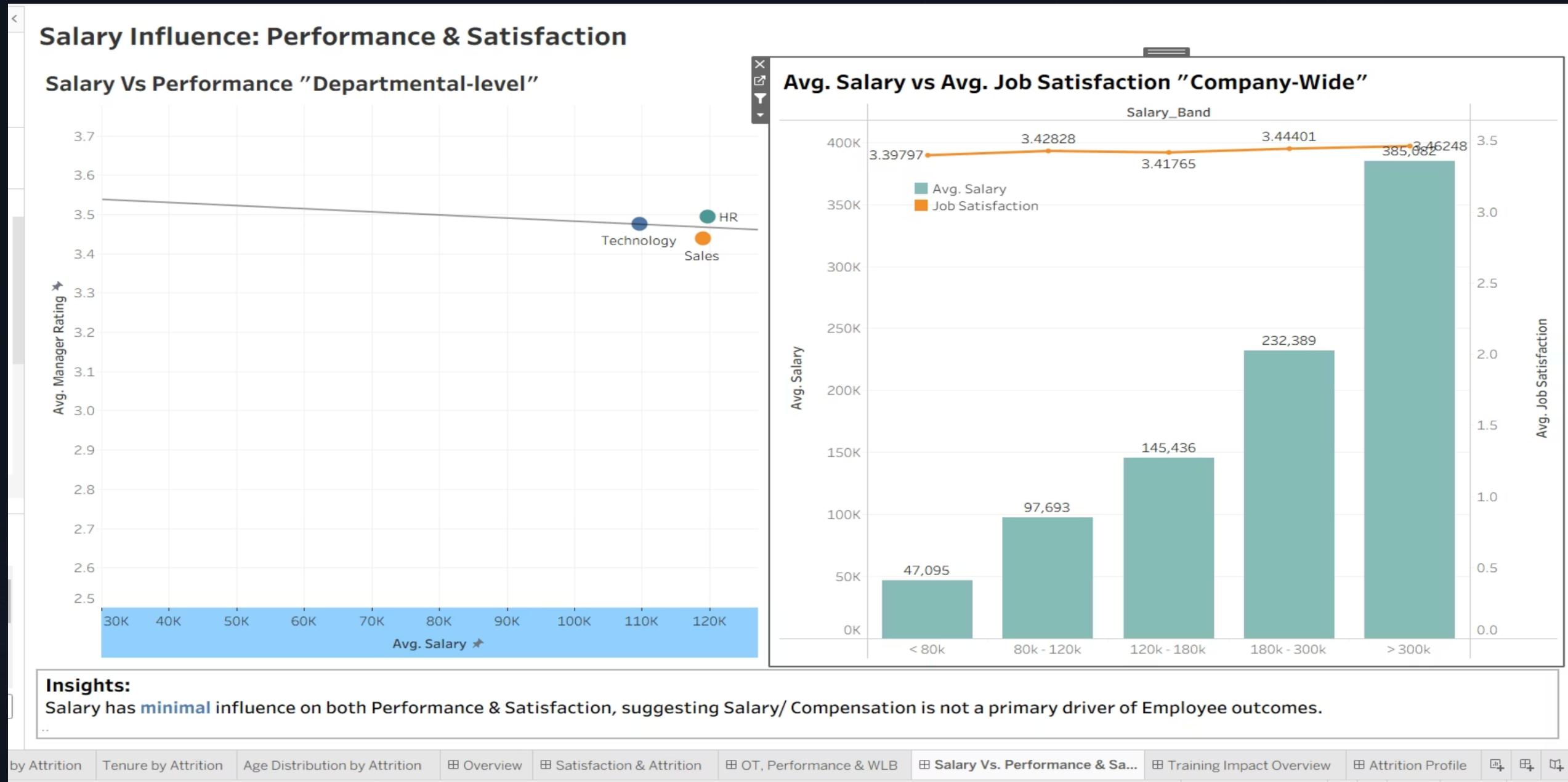
## Overtime Vs. WLB & Performance



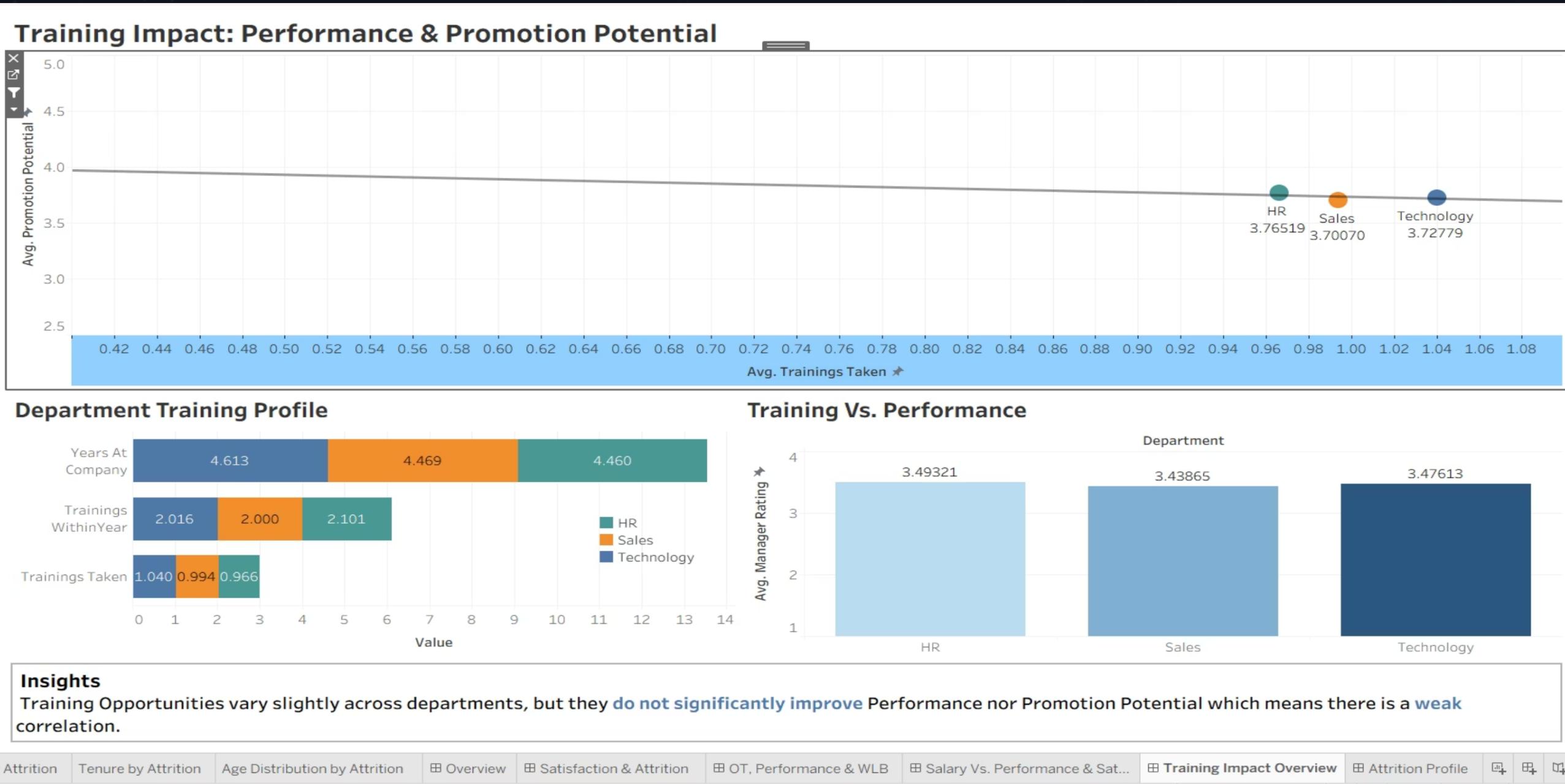
### Insights:

Employees working overtime have higher Job Satisfaction, however, lower Performance indicating overtime affects employee performance.

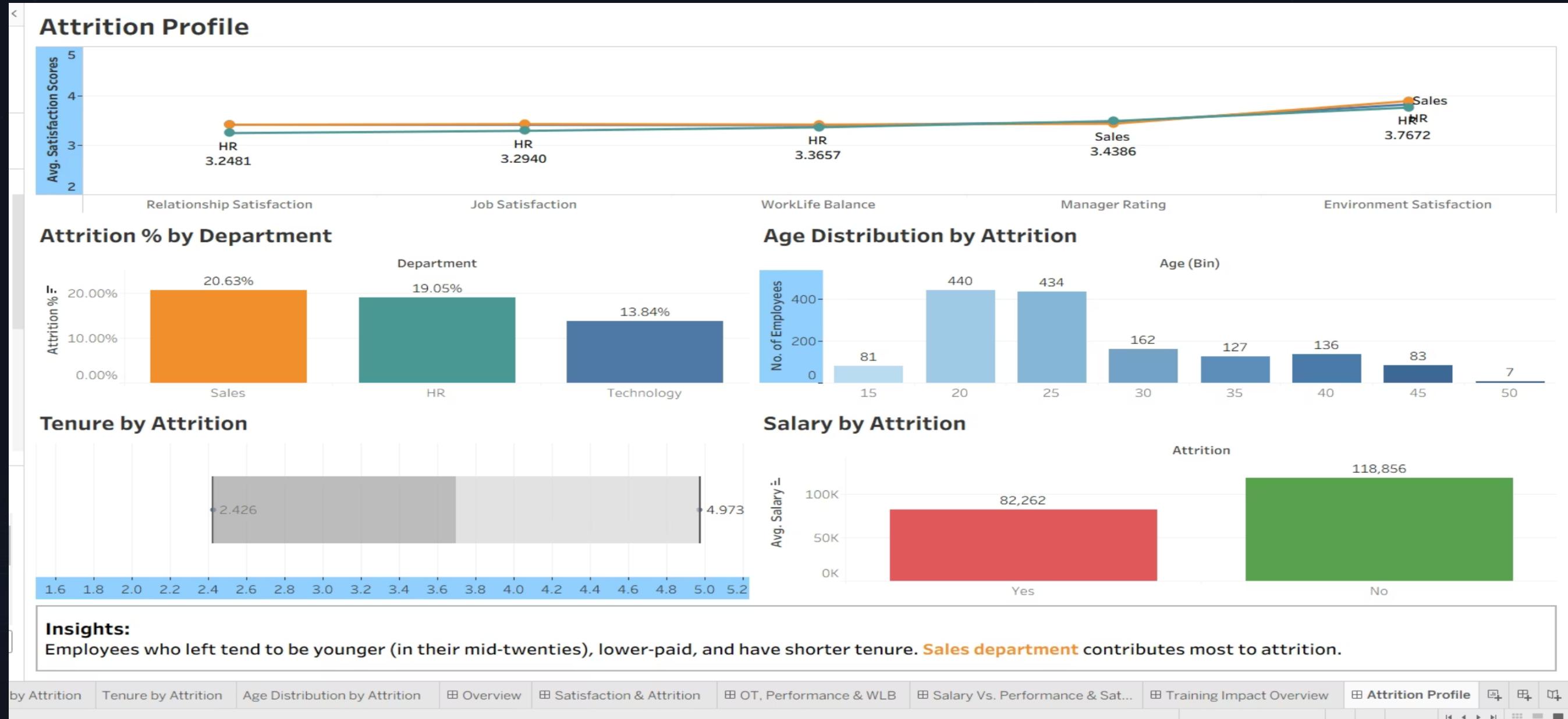
# Salary Influence Dashboard



# Training Impact Dashboard



# Attrition Profile Dashboard



# Conclusion

Our HR Analytics project revealed several key patterns shaping Employee Satisfaction and turnover:

- The **Technology department dominates the workforce (65%)**, and gender distribution is nearly **balanced** across the organization.
- Attrition is **moderately high at 16%**, but satisfaction scores across departments remain **relatively consistent**, suggesting that **Attrition is not driven by Satisfaction alone**, but rather by other operational and behavioral factors.
- **Salary level shows Limited influence** on Satisfaction or Performance, suggesting **compensation is not the main concern** for most employees.
- **Training opportunities show Weak impact** on Performance or Promotion readiness, indicating current programs **may lack effectiveness or alignment** with role needs.
- Employees who leave are generally **younger, earlier in their careers, lower-paid, and have shorter tenures**. The **Sales department contributes most** to the **overall attrition rate**.



# Recommendations to effectively reduce Employee turnover & strengthen Employee Experience

## Improve Early-Career Retention

- Create structured career paths, mentorship programs, and faster development cycles targeting younger, low-tenure employees who are most at risk.
- Create growth pathways to reduce early exits.



## Enhance WLB Company-Wide

- Introduce flexible scheduling, remote-work options, or mandatory time-off initiatives.
- Regularly monitor WLB metrics for early warning signs.

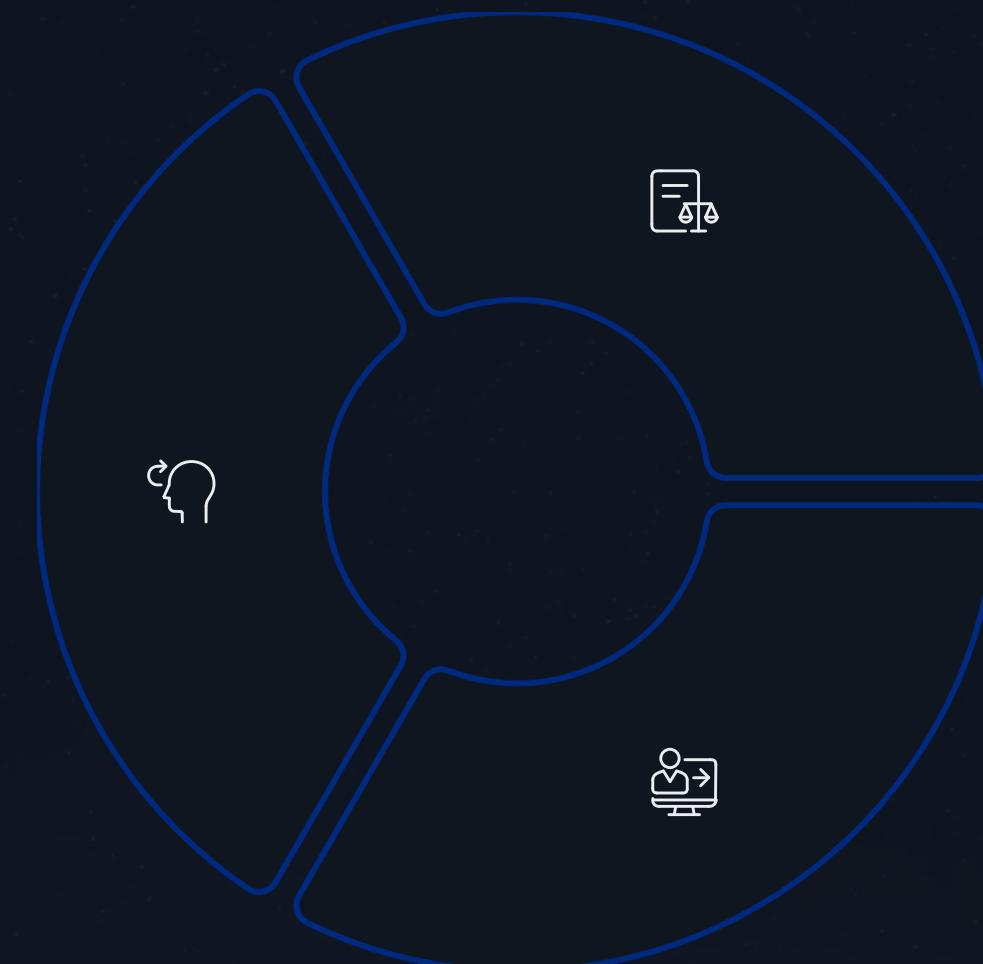
## Address High Attrition in Sales Department

- Review workload, targets, incentives, and managerial support practices.
- Conduct focused interviews or targeted pulse surveys to uncover department-specific stressors.

# Recommendations to effectively reduce Employee turnover & strengthen Employee Experience

## Reassess Training Strategies

- Shift from quantity to **quality and relevance**: design role-specific, skill-based, and progression-aligned development programs.
- Evaluate training effectiveness and ensure it contributes to skill growth and promotion readiness.



## Maintain Transparent Compensation Strategy

- While salary is not a primary turnover driver, ensure fairness across levels and roles to prevent future dissatisfaction.

## Adopt Predictive Analytics (Strategic Add-On)

- Build attrition prediction models to proactively identify high-risk employees and intervene early.

# Thank You!

Feel free to ask any questions!

