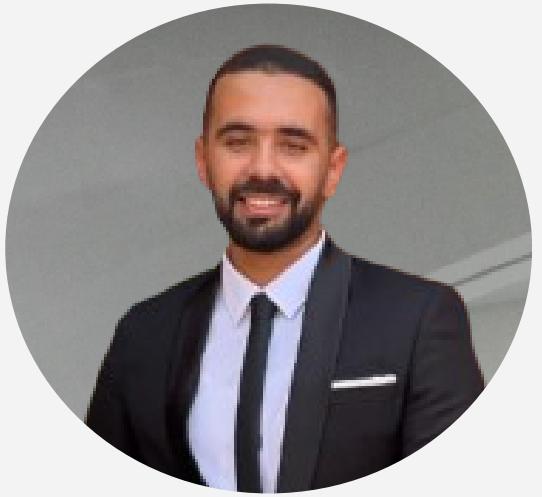


Human Resources Analysis

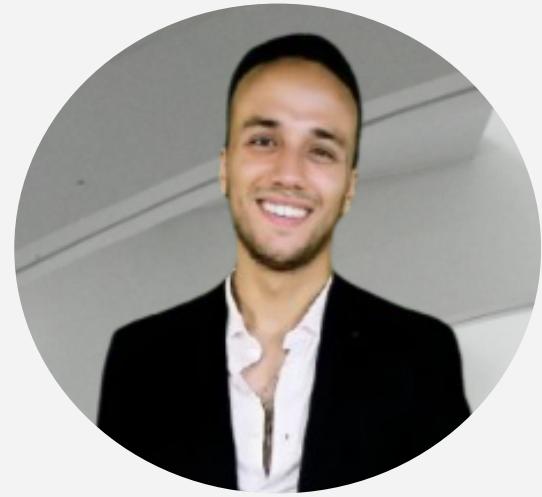


Meet Our Team

Abdelrhman Mohamed



Mostafa Mohamed



Hazem Ahmed



Omar Hegazy



Table Of Contents



01

SUMMARY

02

INTRO & GOALS

03

DATA
PREPROCESING

04

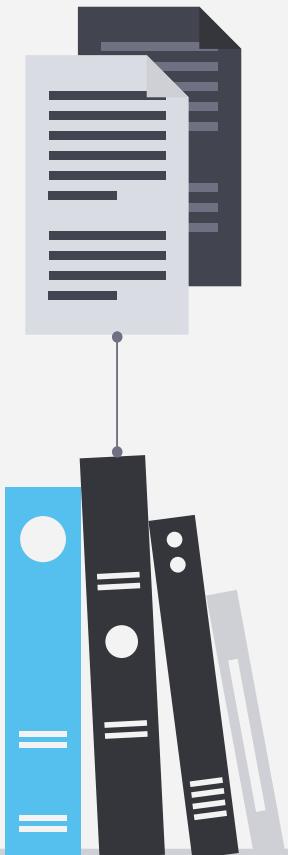
INSIGTHS

05

DATA ANALYSIS

06

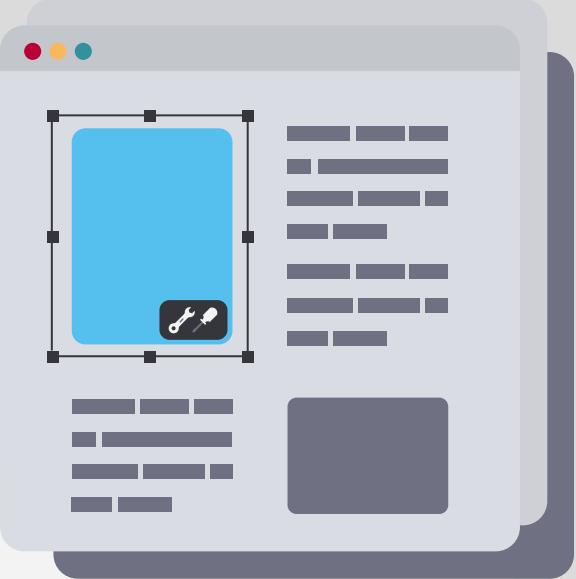
DASHBOARD





01

Summary



Summary

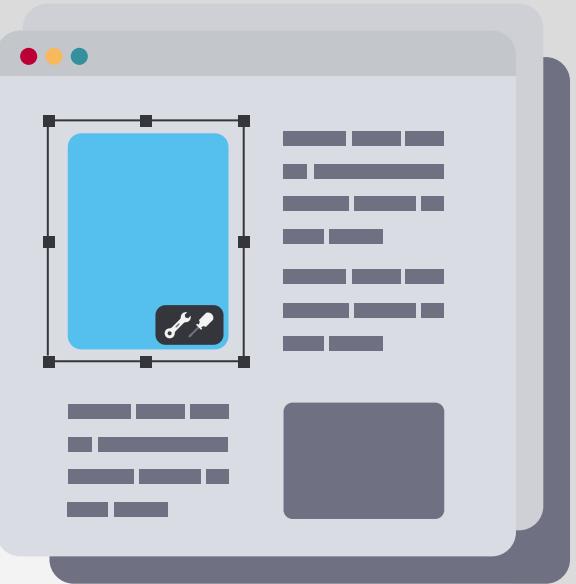


This project focuses on analyzing employee performance and relationships within the company through HR data. It examines employee evaluations based on their salaries, feedback from their managers, and annual promotions. The goal is to assess the overall performance of staff and identify trends or correlations between compensation, managerial evaluations, and career progression.



02

Intro & Goals



Introduction

This project focuses on analyzing the HR data of a company to gain a comprehensive understanding of employee demographics, salary trends, promotions, and work-life factors. By examining key areas such as employee turnover, salary distributions across gender and ethnicity, and correlations between education, marital status, and overtime, the project aims to provide data-driven insights. It will also explore employee-manager relationships through performance ratings and the impact of training opportunities on promotions and salaries. Various statistical methods and visual representations will be used to identify trends and patterns in the data.



Objectives



Analyze Age and Tenure:

- Calculate the average, highest, and lowest employee ages.
- Identify the year with the highest employee turnover.
- Calculate the average duration employees have been with the company.
- Explore the correlation between years at the company and marital status, as well as years at the company and salary.

Examine Salary Distribution:

- Analyze salaries based on gender and ethnicity.
- Identify the highest-paid department and job role.
- Investigate the relationship between marital status and salaries, as well as distance from home and overtime.

Objectives



Department and Role Analysis:

- Determine the department with the fastest promotions.
- Assess the average age of employees by state.
- Analyze the correlation between work-life balance and marital status.

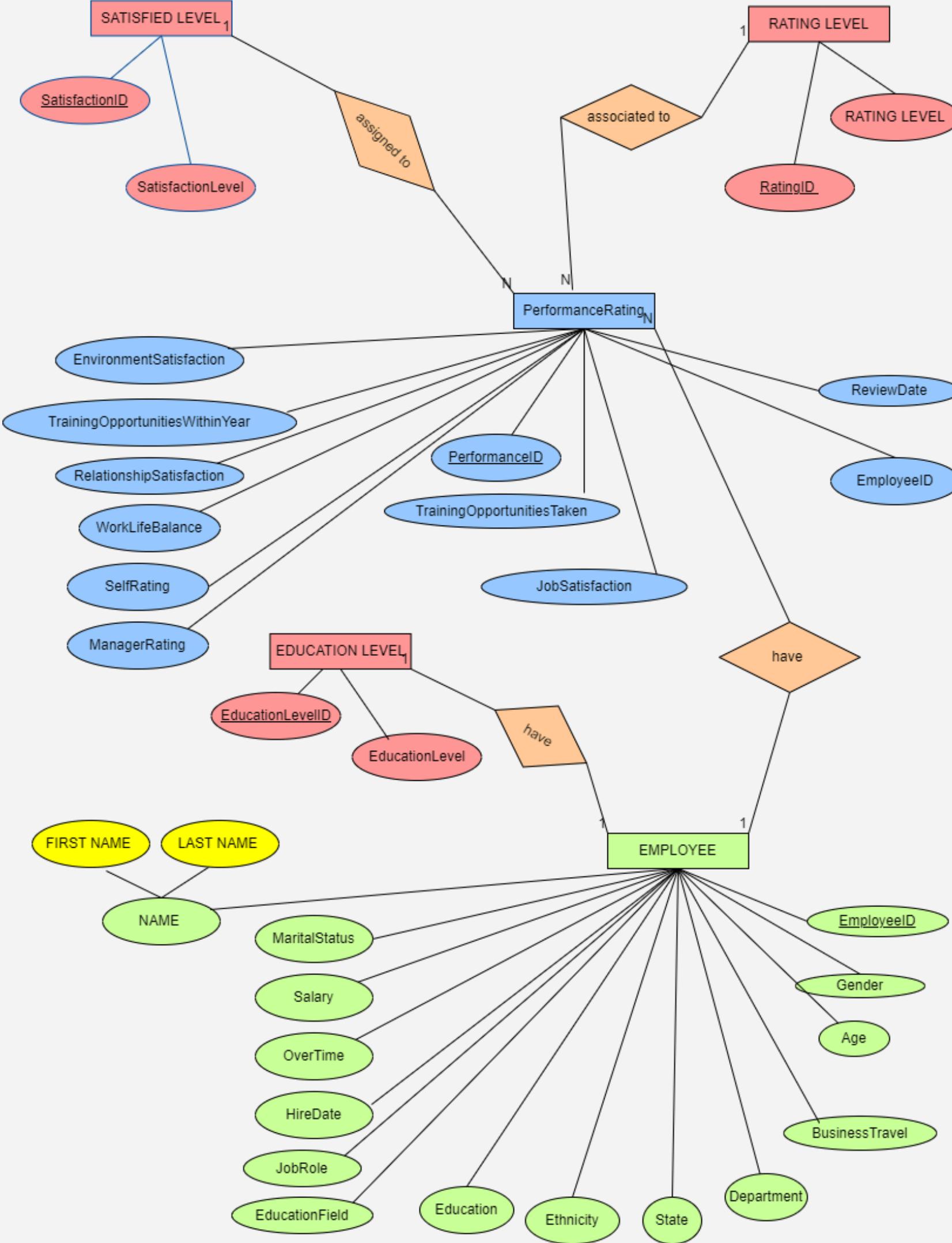
Performance and Promotions:

- Examine the relationship between manager ratings and promotions.
- Calculate the difference between self-ratings and manager ratings.
- Evaluate the correlation between training opportunities taken and salaries or promotions.

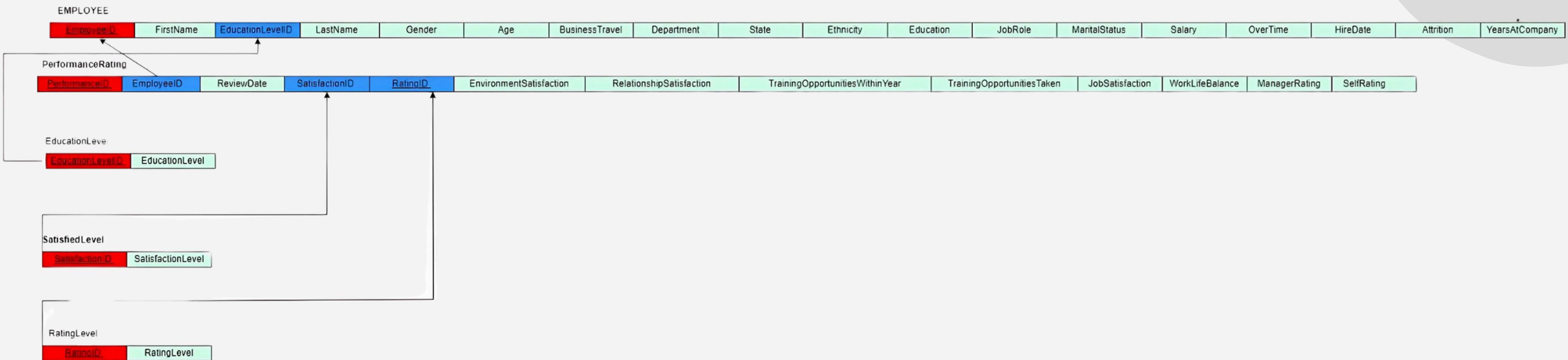
Visualize Employment Trends:

- Use a line chart to display hire dates by employee count.
- Present marital status distribution with a pie chart, highlighting any correlation with salaries and overtime.

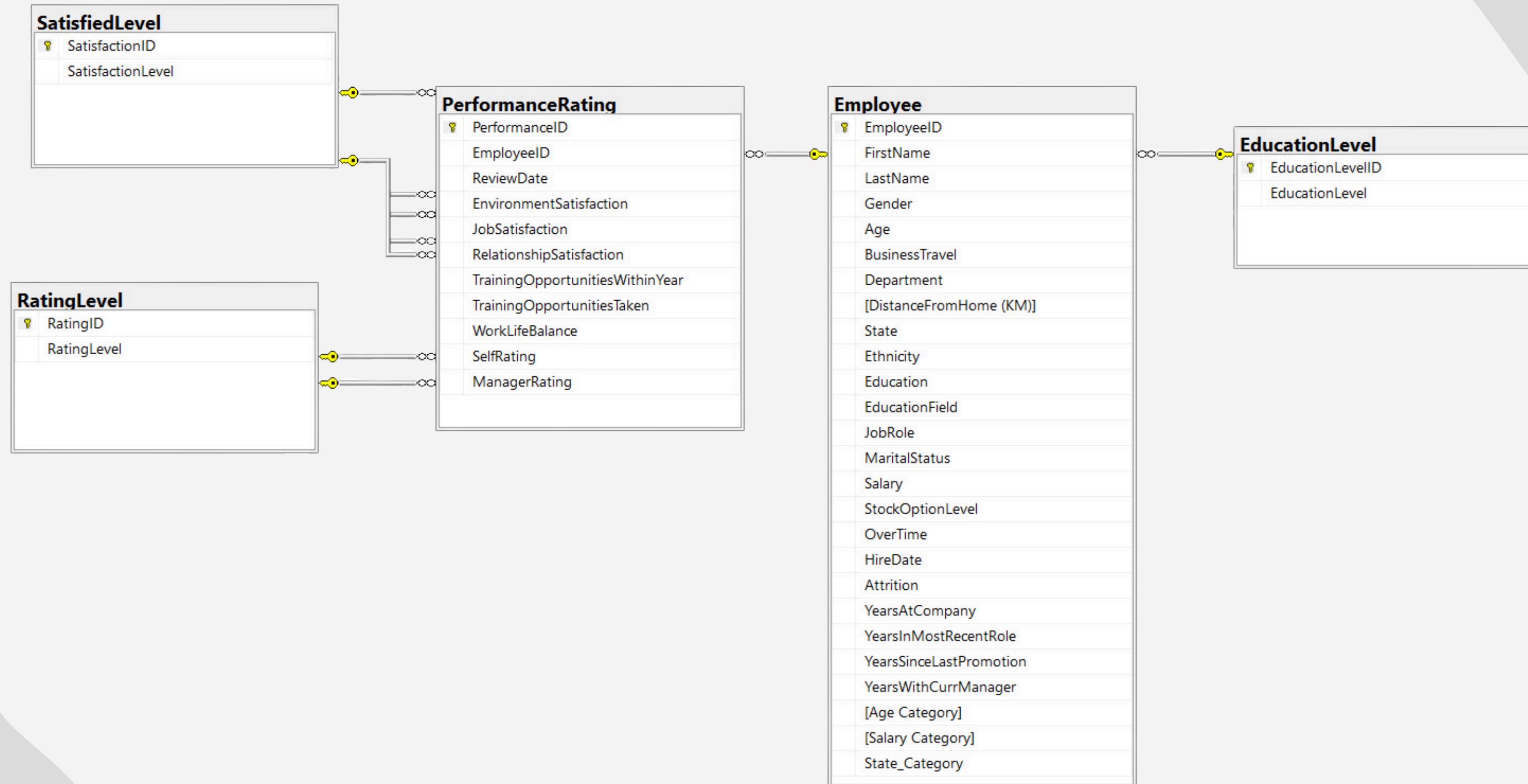
ERD



Mapping

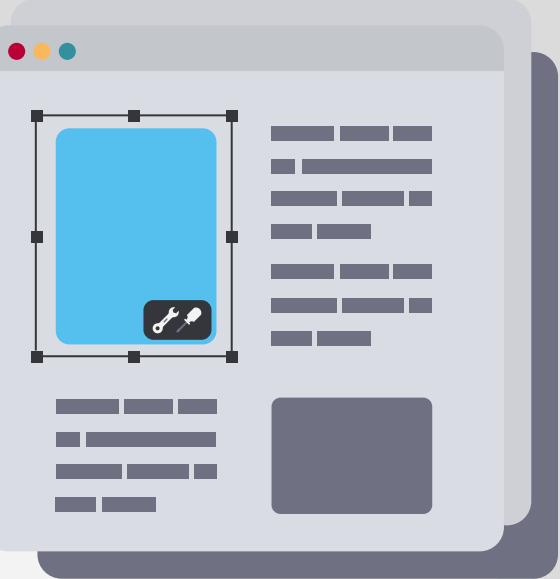


DB – Diagram



03

Data Cleaning



Data Cleaning with Python



- Performed **data cleaning** using Python to prepare the dataset for analysis.
- Corrected the **data types** for the "Salary" and "Hire Date" columns.

```
df.columns = df.columns.str.strip()  
df.info()
```

```
[10]: df['Salary'] = df['Salary'].astype(str).str.strip() # Convert to string and remove spaces  
df['Salary'] = df['Salary'].str.replace(',', '') # Remove commas  
df['Salary'] = df['Salary'].astype(float)  
df.info()
```

Data Cleaning with Python



- **Removed duplicates** to ensure the dataset was free of redundancy.

```
employee_data_cleaned = df.drop_duplicates  
df.info()
```

- Applied **standardization** to columns like "Gender" and "Ethnicity" for consistency.

```
[229]: df['Gender'] = df['Gender'].str.title()  
        df['Ethnicity'] = df['Ethnicity'].str.title()  
        df
```

Data Cleaning with Python



- Grouping the State column into "Head Quarter" and "Branch."

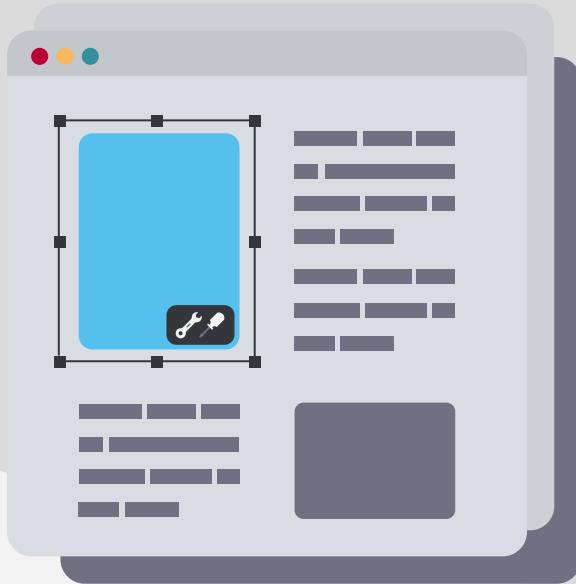
```
[18]: df['State_Category'] = np.where(df['State'] == 'CA', 'HQ', 'Branch')
       df[["State", "State_Category"]]
```

- Created **calculated columns**, including : Dividing **age** into categories ('20s', '30s', '40s', '50s').

```
[16]: # Create the age category column with 4 categories
       df['Age Category'] = pd.cut(df['Age'], bins=[18, 29, 39, 49, 59],
                                     labels=['20s', '30s', '40s', '50s'])

       # Create the salary category column (remains the same as the previous example)
       df['Salary Category'] = pd.cut(df['Salary'], bins=[19999, 49999, 99999, 199999, 547000],
                                         labels=['Low', 'Medium', 'High', 'Very High'])
       df[['Age', 'Age Category', 'Salary', 'Salary Category']]
```

04 Insights

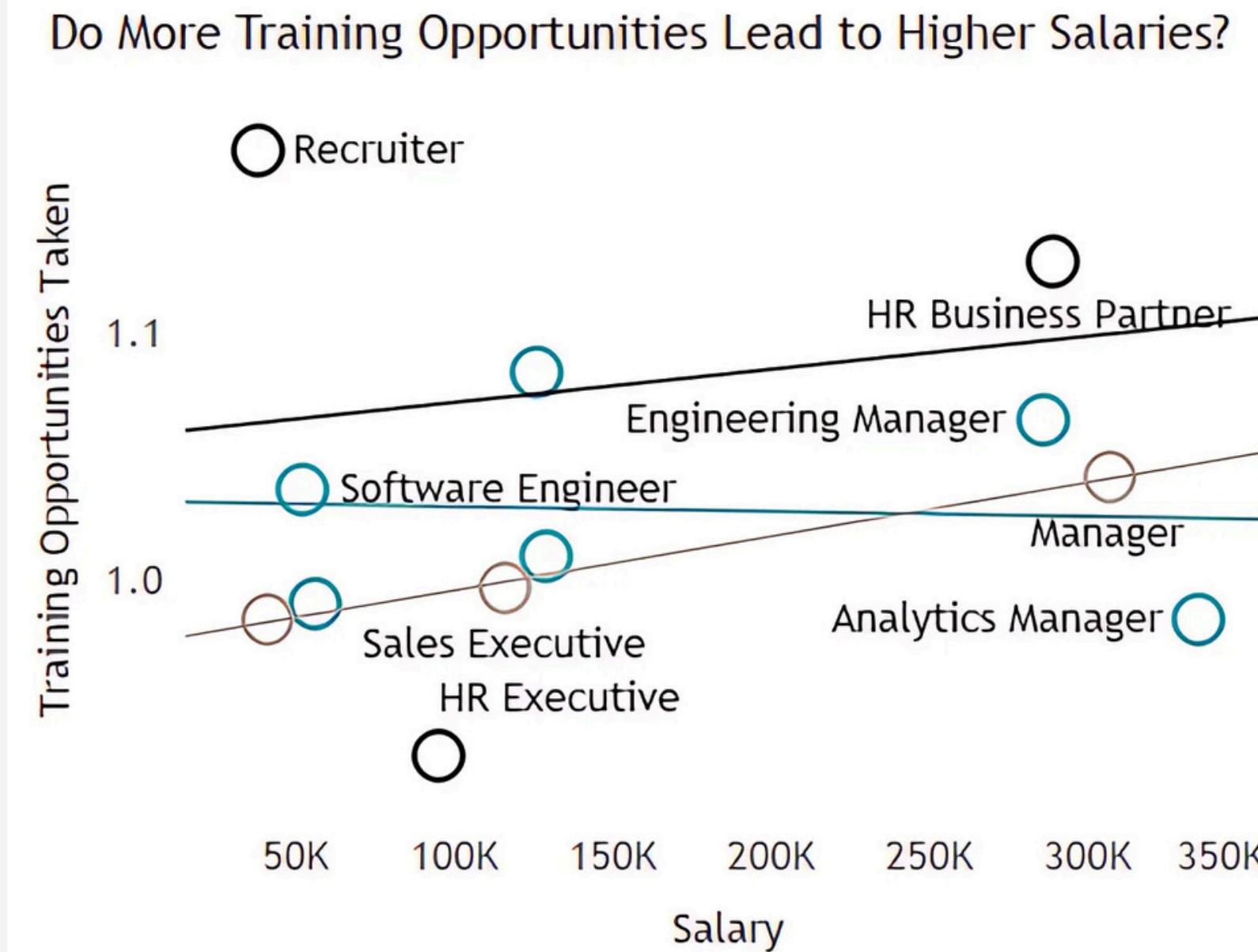


Insights

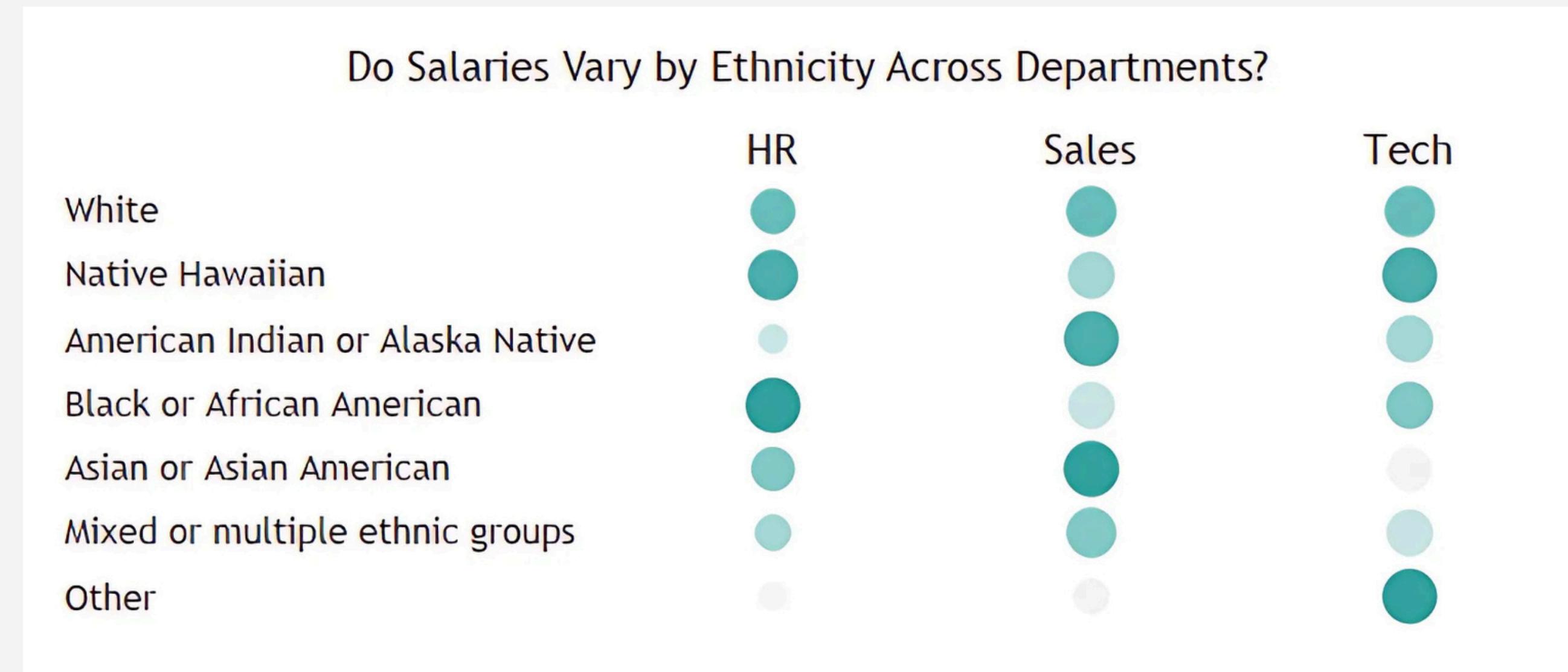


- Average Age, Highest, and Lowest Age
- Highest Department Salary and Job Role
- Average Salary by Ethnicity
- State by Salary and Count of Employees
- Correlation between Education Levels and Salary
- Highest Paid Role and Department
- Pie Chart of Marital Status and Correlation with Salaries
- Correlation between Marital Status and Overtime
- Total Salaries
- Correlation between Distance from Home and Overtime
- Hire Date by Count of Employees (Line Chart)
- Identify the year with the highest employee turnover.
- Analyze salaries based on gender.
- Calculate the average duration employees have been with the company.
- Explore the correlation between years at the company and marital status.
- Investigate the relationship between years at the company and salaries.
- Determine the department with the fastest promotions.
- Assess the average age of employees by state.
- Analyze the correlation between work-life balance and marital status.
- Examine the relationship between manager ratings and promotions.
- Calculate the difference between self-ratings and manager ratings.
- Evaluate the correlation between training opportunities taken and salaries or promotions.

This chart illustrates that the impact of training on salary varies by job role. In some cases, like HR and senior engineering roles, more training correlates with higher salaries. In other roles, such as machine learning or analytics management, other factors might play a more significant role in determining salary.



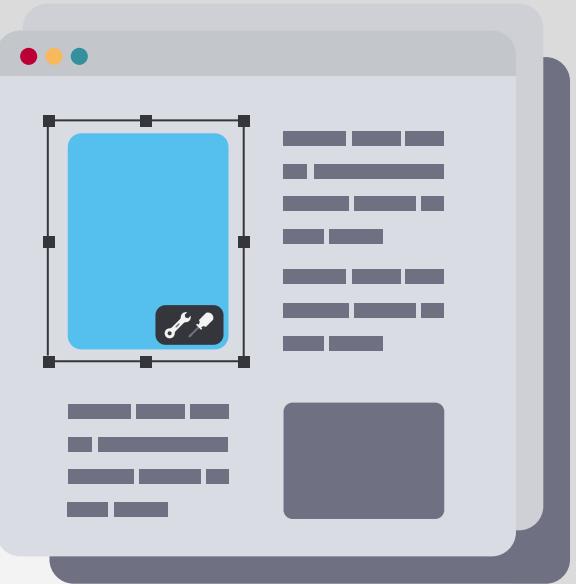
There are clear differences in salary based on ethnicity and department. Sales appears to offer the highest earning potential across different ethnicities, while Human Resources shows more disparity with lower salaries for certain groups like American Indian and Mixed or multiple ethnic groups. Technology presents a mix, with some groups like Other earning more, while Asian and Black or African American employees earn less than in Sales.



05

Data

Analysis



Data Analysis by SQL



```
SELECT top 1
    Department,
    JobRole,
    MAX(CAST(Salary AS DECIMAL(10, 2))) AS HighestSalary
FROM Employee
GROUP BY Department, JobRole
ORDER BY HighestSalary DESC
```

```
SELECT top 1
    YEAR(HireDate) AS TurnoverYear,
    COUNT(EmployeeID) AS TurnoverCount
FROM Employee
WHERE Attrition = 'Yes'
GROUP BY YEAR(HireDate)
ORDER BY TurnoverCount DESC
```

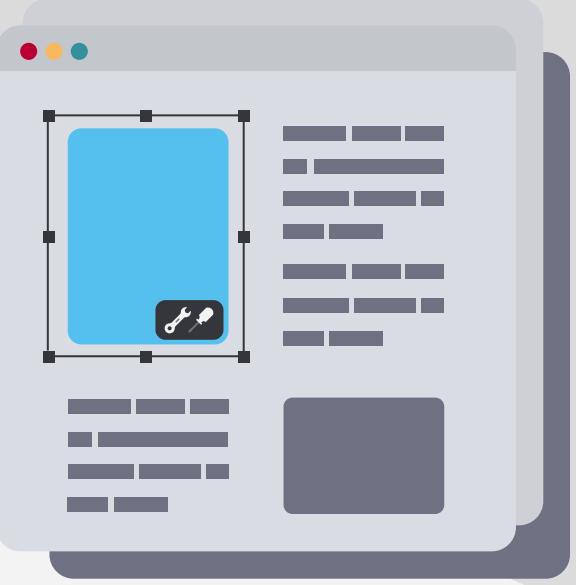
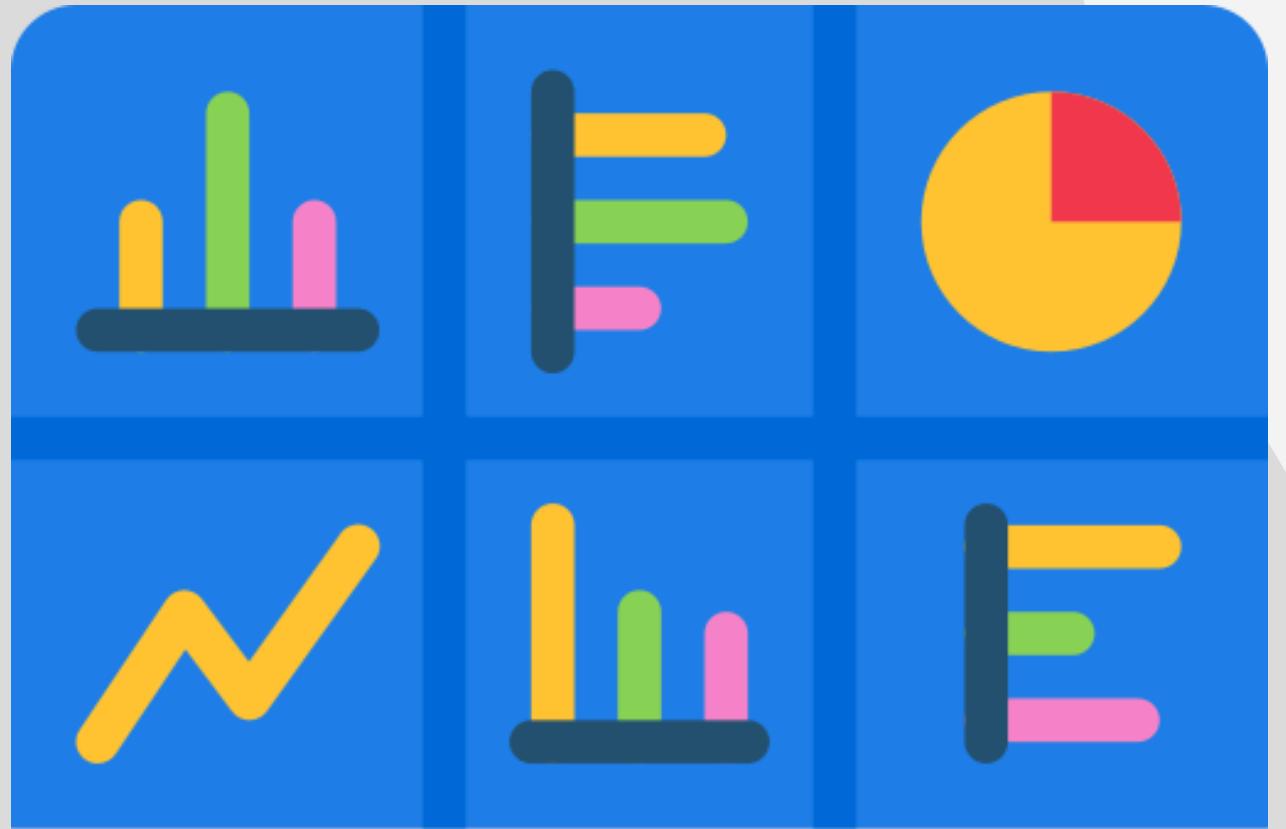
```
SELECT top 1
    Department,
    JobRole,
    MAX(CAST(Salary AS DECIMAL(10, 2))) AS HighestSalary
FROM Employee
GROUP BY Department, JobRole
ORDER BY HighestSalary DESC
```

```
SELECT
    YEAR(HireDate) AS HireYear,
    COUNT(EmployeeID) AS EmployeeCount
FROM Employee
GROUP BY YEAR(HireDate)
ORDER BY HireYear;
```

```
SELECT
    E.EducationLevel,
    AVG(CAST(Empl.Salary AS DECIMAL(10, 2))) AS AvgSalary
FROM EducationLevel E
JOIN Employee Empl ON E.EducationLevelID = Empl.Education
GROUP BY E.EducationLevel;
```

06

Dashboard



Human Resource Dashboard | Overview

Overview

In-Depth Analysis

Employee List



Active Employees

1,233

Term. Employees

237

Total Salaries

155.03M

Avg. Years Tenure

4.6

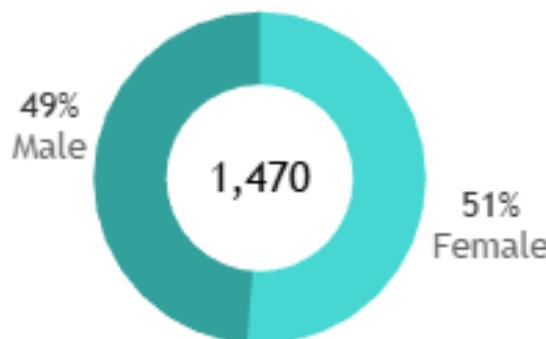
Avg. Age

29

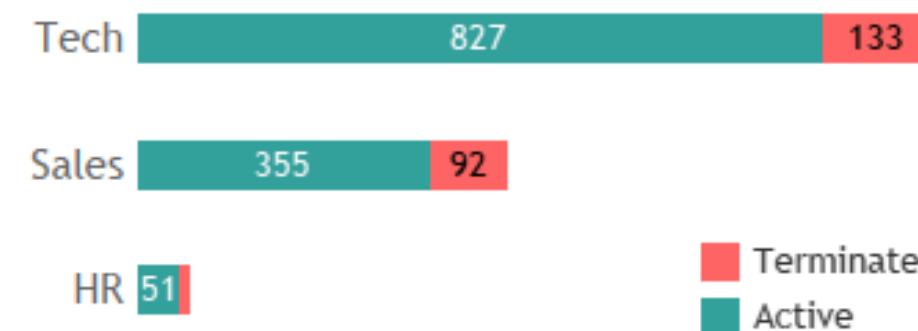
Turn Over

16.1%

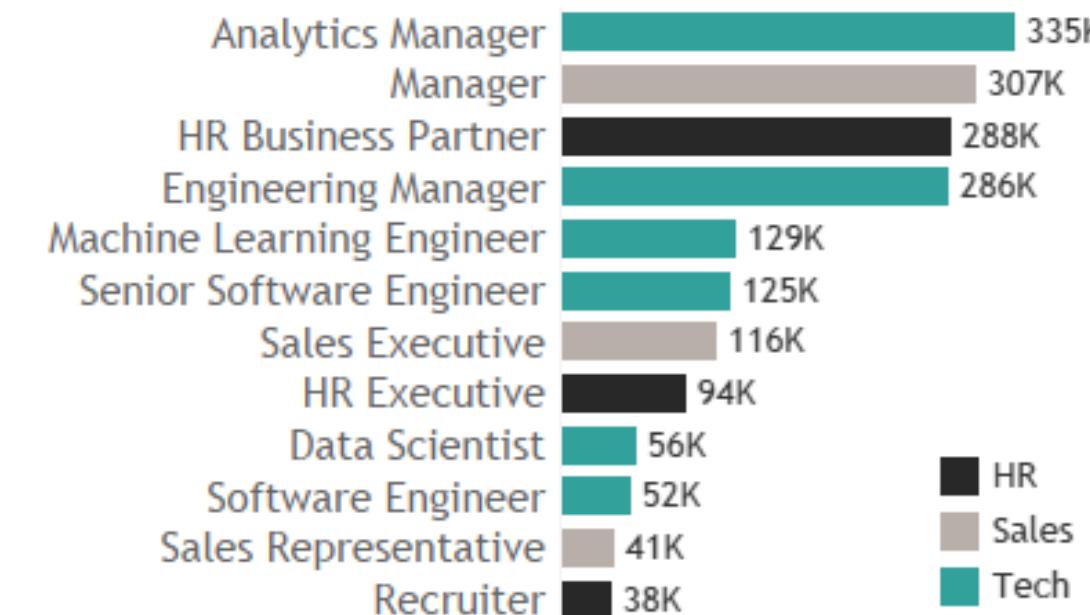
Gender



Total Employees by Department



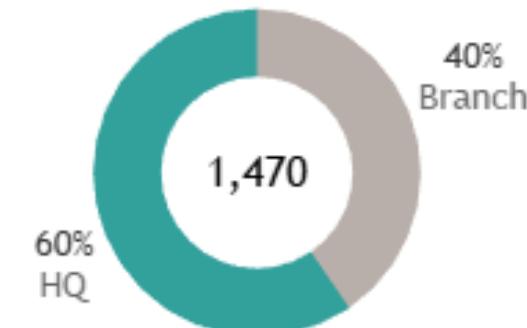
Avg. Salary by Job Role



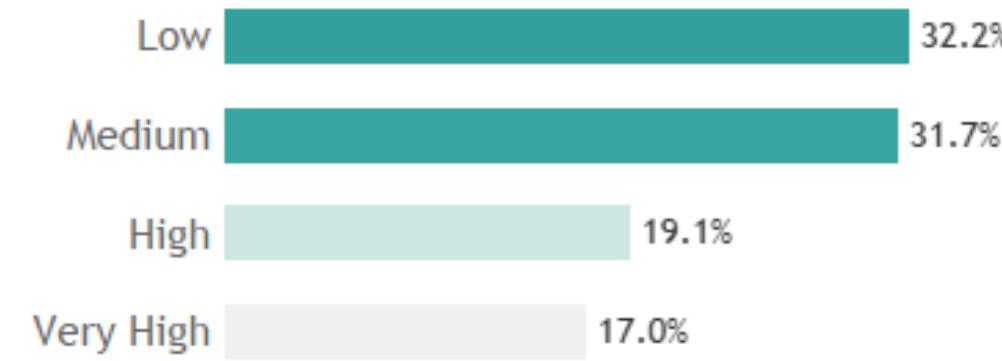
Avg. Years Since Last Promotion by Dep.



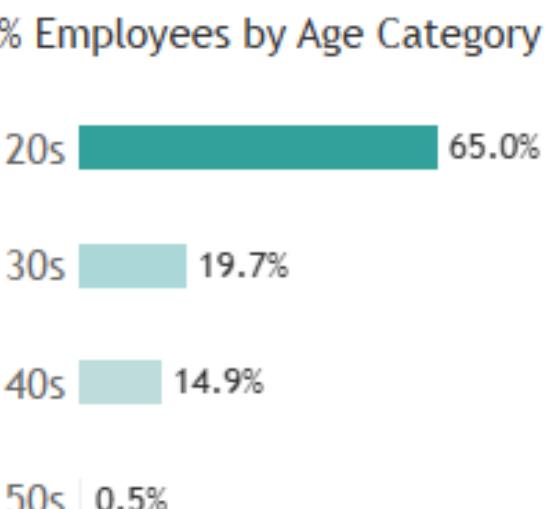
Location



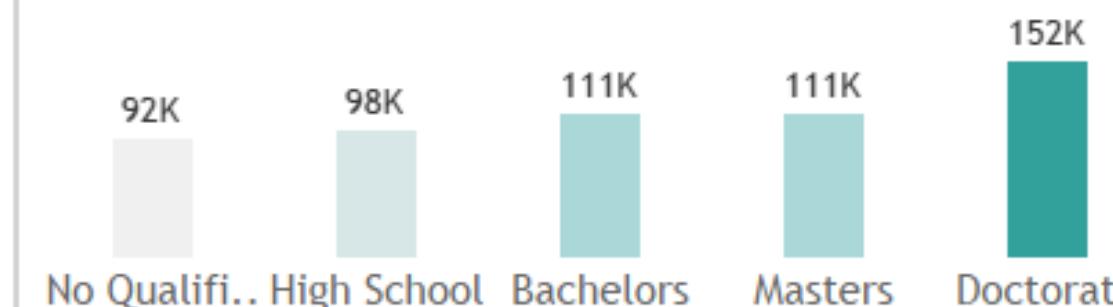
% Employees By Salary Category



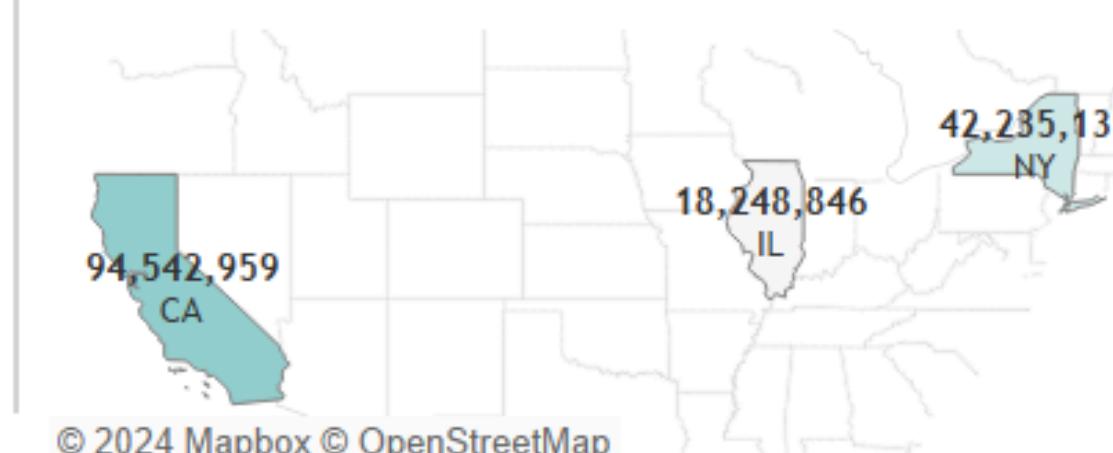
Total Employees by Hire Date



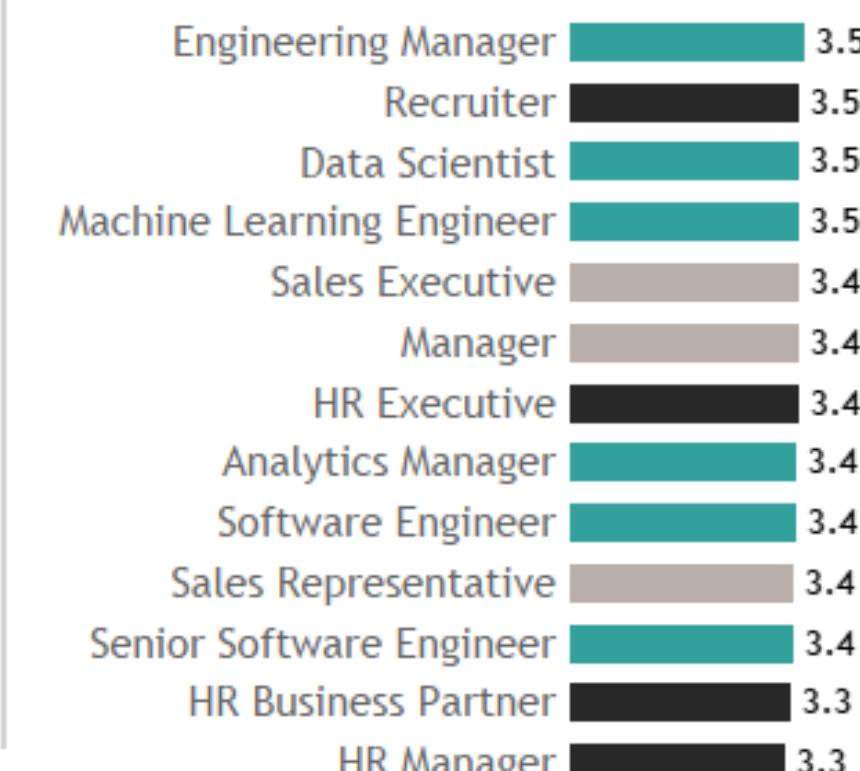
Avg. Salary by Education Level

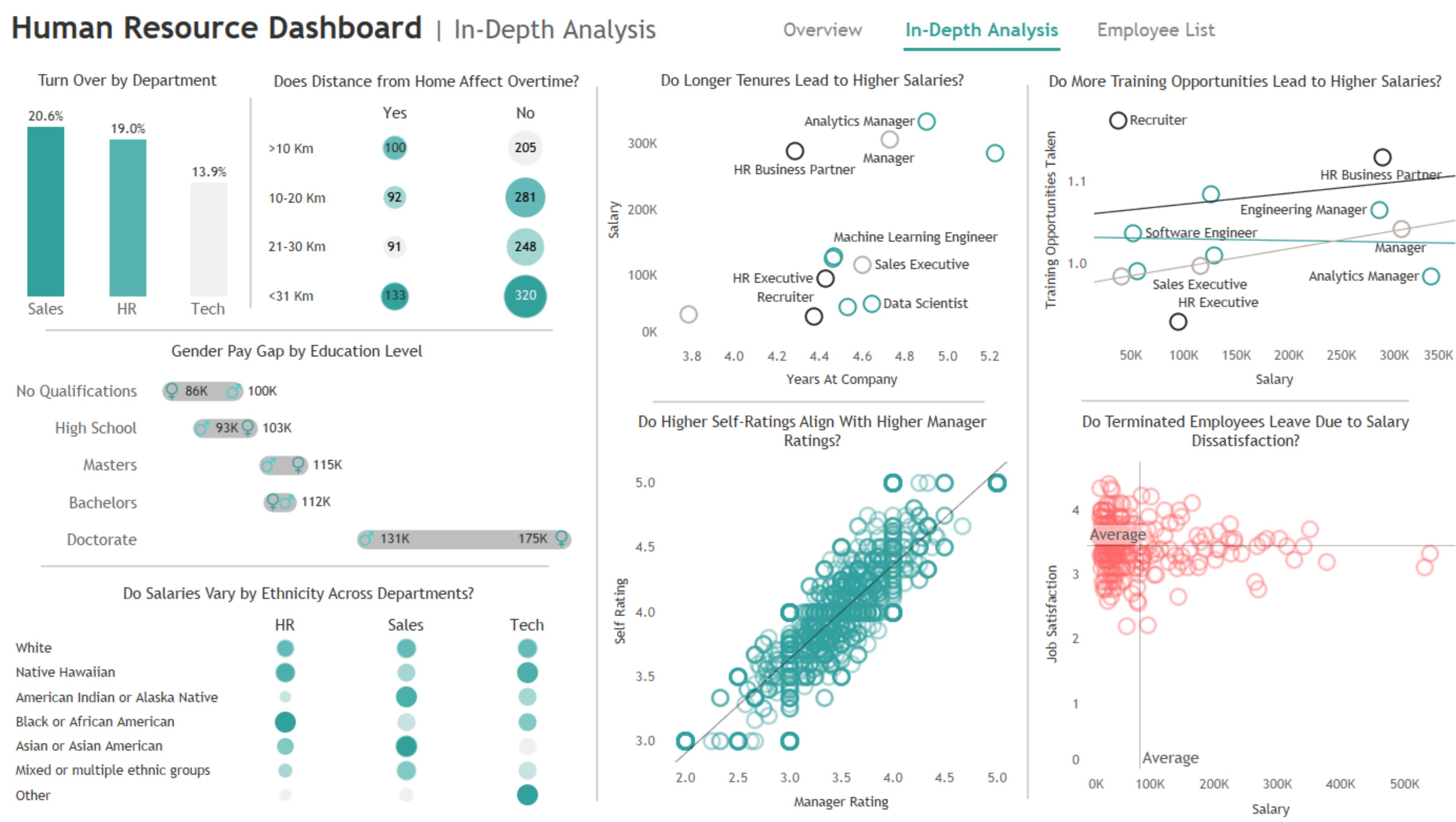


Total Employees & Salary by State



Avg. Job Satisfaction Rating by Job Role





Human Resource Dashboard | Employee List

[Overview](#)
[In-Depth Analysis](#)
[Employee List](#)


ID	Gender	Name	Age	Degree	Role	Department	Location	Salary	Hire Date	Employment Status	Contract Length
0A4E-1A4A	♂	Aloysius Holdren	22	Bachelors	Sales Executive	Sales	CA	\$53,014	Hired 2/4/2019	3 Years	
0A05-3943	♀	Sherie Pattington	26	Bachelors	Software Engineer	Tech	CA	\$33,273	Terminated 1/20/2015	2 Years	
0AA2-881A	♂	Jeffrey Mildmott	25	Masters	Software Engineer	Tech	NY	\$29,126	Hired 7/18/2016	6 Years	
0AE7-DDE2	♂	Martyn Rathke	23	High School	Machine Learning Engineer	Tech	CA	\$46,850	Hired 10/20/2018	4 Years	
0AF4-EC15	♀	Gwenore Bernucci	24	Bachelors	Sales Executive	Sales	CA	\$91,915	Hired 7/23/2018	4 Years	
0B9A-B6DE	♂	Mikol Ortner	19	Bachelors	Machine Learning Engineer	Tech	CA	\$60,077	Hired 7/11/2021	1 Years	
0B19-EFB2	♀	Ada Bynold	39	Masters	Engineering Manager	Tech	CA	\$300,594	Hired 4/24/2018	4 Years	
0B49-32CD	♀	Nettle Manwaring	20	High School	Data Scientist	Tech	CA	\$41,179	Hired 11/4/2020	2 Years	
0B72-0EE6	♂	Noll Antoniotti	24	Masters	Senior Software Engineer	Tech	NY	\$70,849	Hired 9/3/2016	6 Years	
0BAF-7454	♀	Ninnette Braganza	26	Masters	Software Engineer	Tech	NY	\$37,530	Hired 7/30/2017	5 Years	
0BEE-FC20	♀	Sherill Covotti	49	Masters	Manager	Sales	IL	\$413,946	Hired 10/19/2021	1 Years	
0C5B-2EF4	♂	Amery Theunissen	20	Bachelors	Software Engineer	Tech	CA	\$28,506	Hired 12/26/2021	1 Years	
0C6D-2825	♂	Tann Lettley	21	Bachelors	Software Engineer	Tech	CA	\$37,164	Hired 3/6/2019	3 Years	
0C7B-2EA6	♂	Oswell McCloid	34	High School	Data Scientist	Tech	NY	\$39,015	Hired 9/15/2013	9 Years	
0C23-26DD	♀	Bathsheba Ablewhite	35	Bachelors	Sales Representative	Sales	NY	\$81,389	Hired 3/11/2013	9 Years	
0C37-B7CF	♂	Maximilien Bodimeade	22	Masters	Senior Software Engineer	Tech	NY	\$68,743	Hired 3/23/2019	3 Years	
0C54-705D	♀	Grata Yannoni	50	High School	Manager	Sales	NY	\$291,477	Hired 12/10/2021	1 Years	

Thank You!

