HR Data Analysis: Insights and Implications for Decision-Making

Exploring key insights from HR data to inform strategic decision-making processes.

Meet Our Team

- Mostafa Bakry AboElhamd
- Esraa Ayman Abd-Elgany Elhofy
- Menna-allah Taha Saadeldin
- Rahma Essam Hassan

Project's Overview

Our graduation project focuses on HR Data Analysis.
The goal of the project was to understand the key factors influencing employee behavior especially attrition and to support HR decision-making using data-driven insights.

We worked with two datasets: employee information and performance ratings.
Our objective was to clean, analyze, and visualize this data to discover patterns that can help organizations reduce employee turnover, improve satisfaction, and strengthen overall workforce management."

Objectives of the Presentation

Present Key Findings from HR Data Analysis

The HR data analysis project aims to showcase critical findings that reveal patterns and trends in workforce behavior, performance metrics, and overall organizational health. This analysis serves as the foundation for informed decision-making within the HR department and beyond.

Discuss Implications for HR Decision-Making

Understanding the implications of the HR data analysis results is crucial for strategic HR decision-making. This includes identifying opportunities for workforce optimization, enhancing talent acquisition strategies, and improving retention efforts, ultimately aligning HR practices with organizational goals.

Provide Recommendations for Future Analysis

Based on the insights gathered, this presentation will provide actionable recommendations for future HR data analysis initiatives. These recommendations will focus on improving data collection practices, leveraging new analytical tools, and fostering a data-driven culture within the HR team to enhance overall effectiveness.

Present Key Findings from HR Data Analysis

The HR data analysis project aims to showcase critical findings that reveal patterns and trends in workforce behavior, performance metrics, and overall organizational health. This analysis serves as the foundation for informed decision-making within the HR department and beyond.

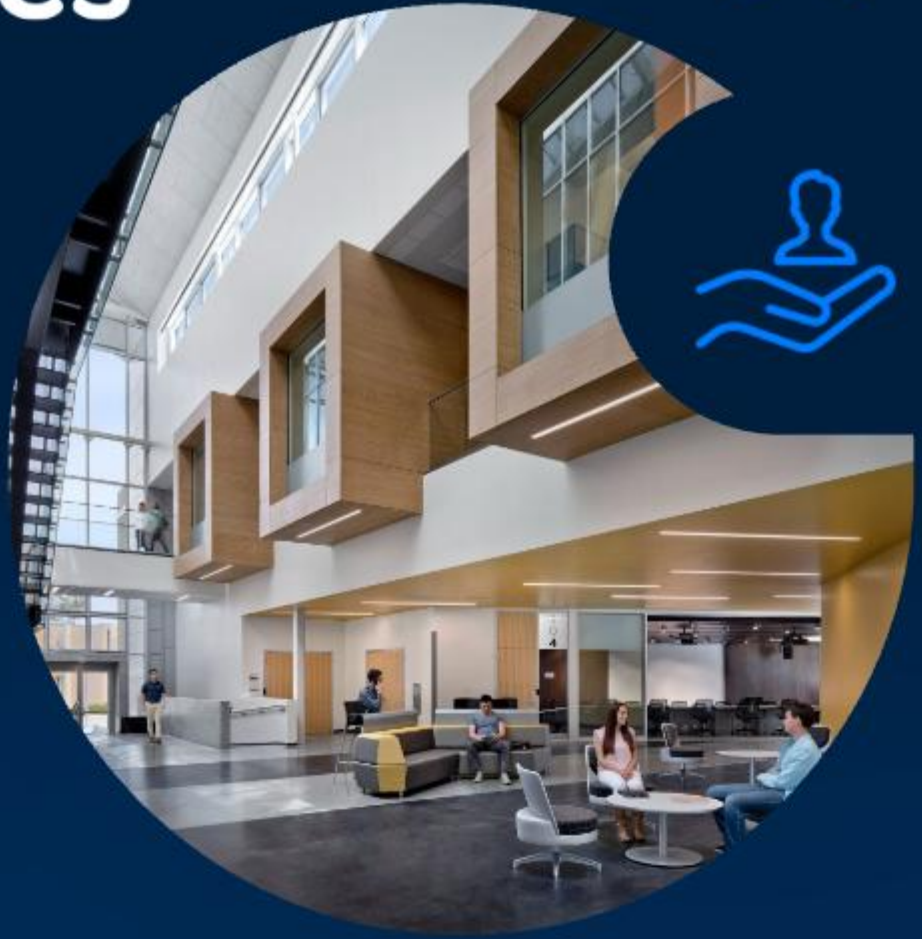
Discuss Implications for HR Decision-Making

Understanding the implications of the HR data analysis results is crucial for strategic HR decision-making. This includes identifying opportunities for workforce optimization, enhancing talent acquisition strategies, and improving retention efforts, ultimately aligning HR practices with organizational goals.

Provide Recommendations for Future Analysis

Based on the insights gathered, this presentation will provide actionable recommendations for future HR data analysis initiatives. These recommendations will focus on improving data collection practices, leveraging new analytical tools, and fostering a data-driven culture within the HR team to enhance overall effectiveness.

Project Explanation & Key Features



Part 1 — Data Cleaning

Part 2 — Data Analysis

Part 3 — Dashboard

Data Cleaning

Ensured that both datasets were clean and reliable.

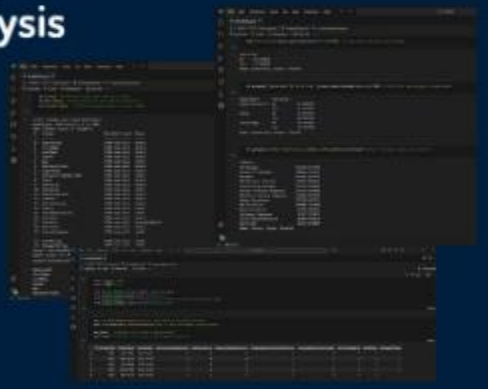
Checked:

- Data structure and column consistency
- Unique values in categorical columns
- Duplicate records
- Missing values
- Outliers in numerical fields

Also corrected inconsistent values, such as fixing duplicated 'Marketing' variations in the Education Field column. After cleaning, the datasets became fully consistent and ready for merging."

Data Analysis

- Merged datasets
- Attrition analysis
- Salary trends
- Age groups & attrition
- Promotion impact
- Business travel effect
- Correlation insights



Rigorous Data Cleaning and Preparation

Prior to analysis, data underwent rigorous cleaning and preparation to ensure accuracy and reliability. This process involved eliminating duplicates, correcting errors, and standardizing formats to create a clean dataset.

Key Insights

- Attrition rate: 33.7%
- Highest attrition: Sales
- Lower attrition: Technology
- Younger employees leave more
- No promotion □ higher attrition
- Frequent travel □ higher attrition
- Salary has mild negative correlation with attrition

Project Breakdown & Key Findings

Data Cleaning

Ensured that both datasets were clean and reliable.

Checked:

- Data structure and column consistency
- Unique values in categorical columns
- Duplicate records
- Missing values
- Outliers in numerical fields

Also corrected inconsistent values, such as fixing duplicated ‘Marketing’ variations in the Education Field column. After cleaning, the datasets became fully consistent and ready for merging.”

HR Data Analysis

In [17]:
```
import pandas as pd

df = pd.read_csv('Employee.csv')              # to load csv files
df_second = pd.read_csv('PerformanceRating.csv')
```

Cleaning process for : Employee.csv

1- Initial Review

to confirm the data is loaded successfully and to have an idea about the data we are working with

In [18]:
```
print(df.head())                  # show first few rows
```

to spot any typos or inconsistencies

In [20]:
```
print(df.columns)                 #Lists all column names
```
```
Index(['EmployeeID', 'FirstName', 'LastName', 'Gender', 'Age',
       'BusinessTravel', 'Department', 'DistanceFromHome (KM)', 'State',
       'Ethnicity', 'Education', 'EducationField', 'JobRole', 'MaritalStatus',
       'Salary', 'StockOptionLevel', 'OverTime', 'HireDate', 'Attrition',
       'YearsAtCompany', 'YearsInMostRecentRole', 'YearsSinceLastPromotion',
       'YearsWithCurrManager'],
      dtype='object')
```

define columns with categorical values (non_numeric values)

In [21]:
```
categorical_cols_to_check = [
    'Gender', 'BusinessTravel', 'Department', 'State',
    'Ethnicity', 'EducationField', 'JobRole', 'MaritalStatus',
    'OverTime', 'Attrition'
]
```

Loop through the list to Check for Unique Values to find (Inconsistencies & Typos)

2- Statistical Overview

In [24]:
```
print(df.shape)            #number of rows,columns
```
```
(1470, 23)
```

to spot outliers

In [25]:
```
print(df.describe())              #Provides summary statistics
```

3- checking for duplicates and missing values

In [26]:
```
# Handling duplicates
df.duplicated().sum()
```
Out[26]:
```
np.int64(0)
```

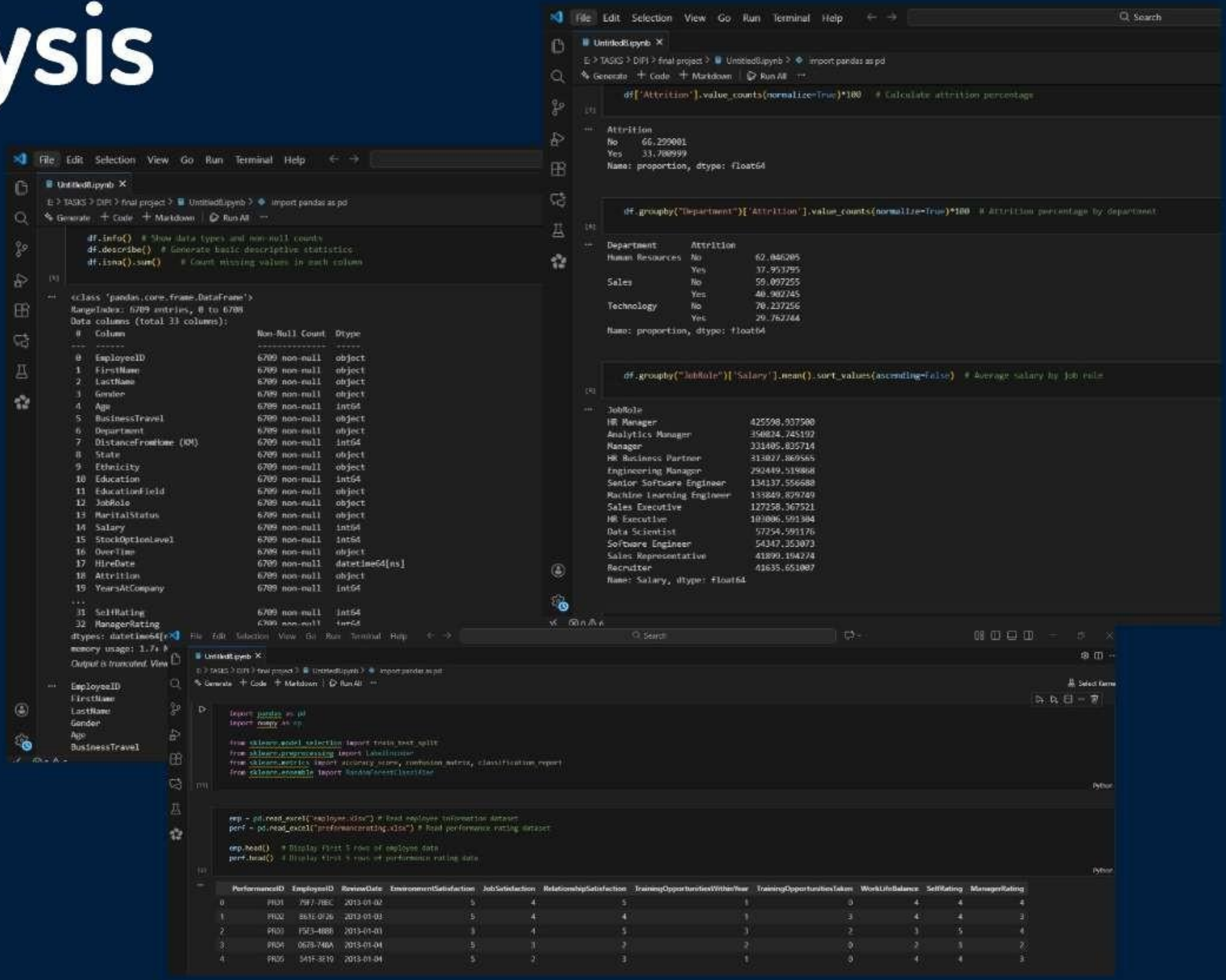
there are no duplicates

In [27]:
```
# checking missing value
df.isnull().sum()
```

Prezi

Data Analysis

- Merged datasets
- Attrition analysis
- Salary trends
- Age groups & attrition
- Promotion impact
- Business travel effect
- Correlation insights



Key Insights

- Attrition rate: 33.7%
- Highest attrition: Sales
- Lower attrition: Technology
- Younger employees leave more
- No promotion ▯ higher attrition
- Frequent travel ▯ higher attrition
- Salary has mild negative correlation with attrition

Dashboard

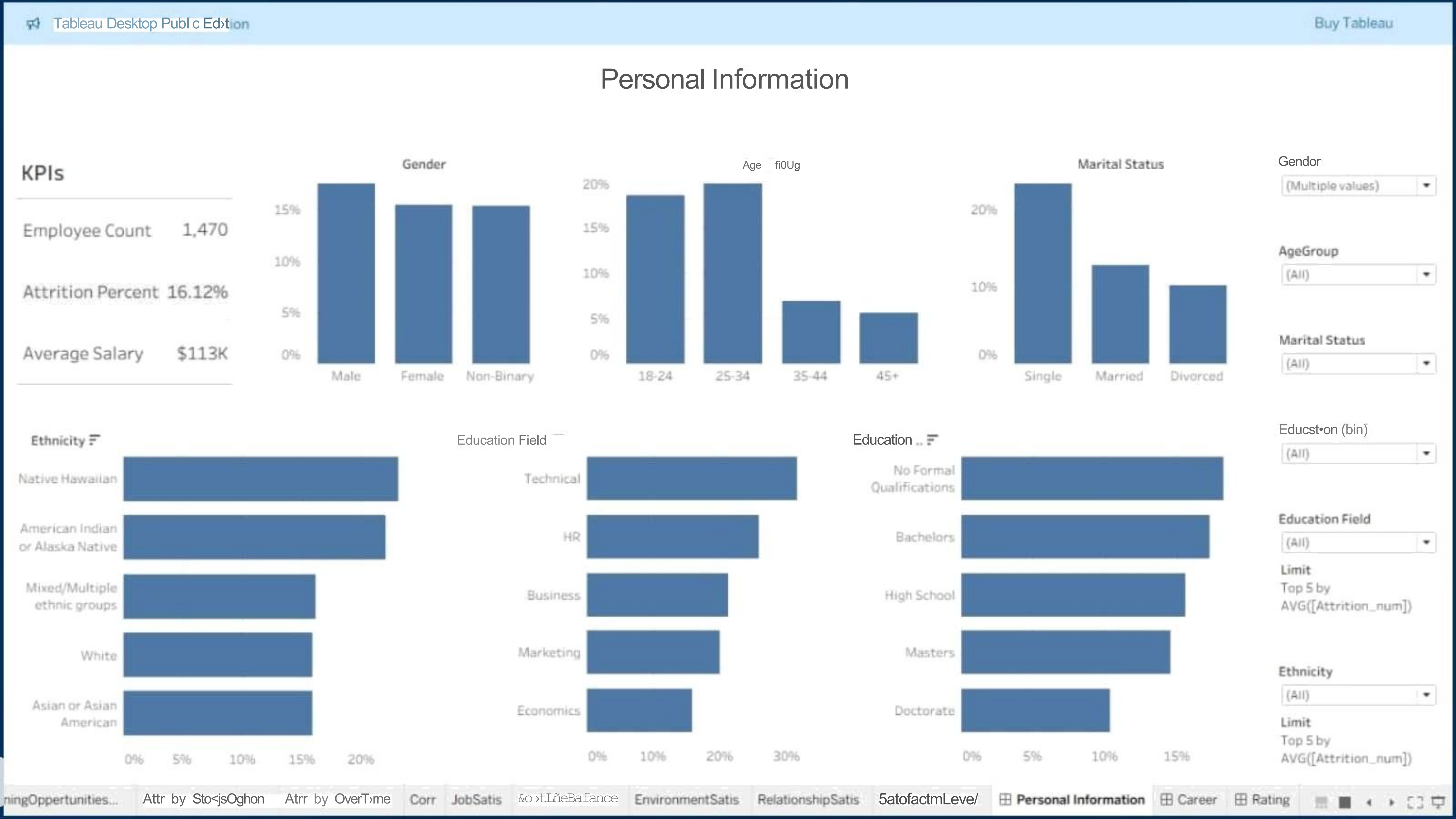


Tableau Desktop Public Edition
Buy Tableau
Personal Information
KPIs
Employee Count 1,470
Attrition Percent 16.12%
Average Salary $113K
Gender
Male
Female
Non-Binary
Age
18-24
25-34
35-44
45+
Marital Status
Single
Married
Divorced
Ethnicity
Native Hawaiian
American Indian or Alaska Native
Mixed/Multiple ethnic groups
White
Asian or Asian American
Education Field
Technical
HR
Business
Marketing
Economics
Education
No Formal Qualifications
Bachelors
High School
Masters
Doctorate
Gendor
(Multiple values)
AgeGroup
(All)
Marital Status
(All)
Education (bin)
(All)
Education Field
(All)
Limit
Top 5 by AVG([Attrition_num])
Ethnicity
(All)
Limit
Top 5 by AVG([Attrition_num])
Corr
JobSatis
EnvironmentSatis
RelationshipSatis
Personal Information
Career
Rating

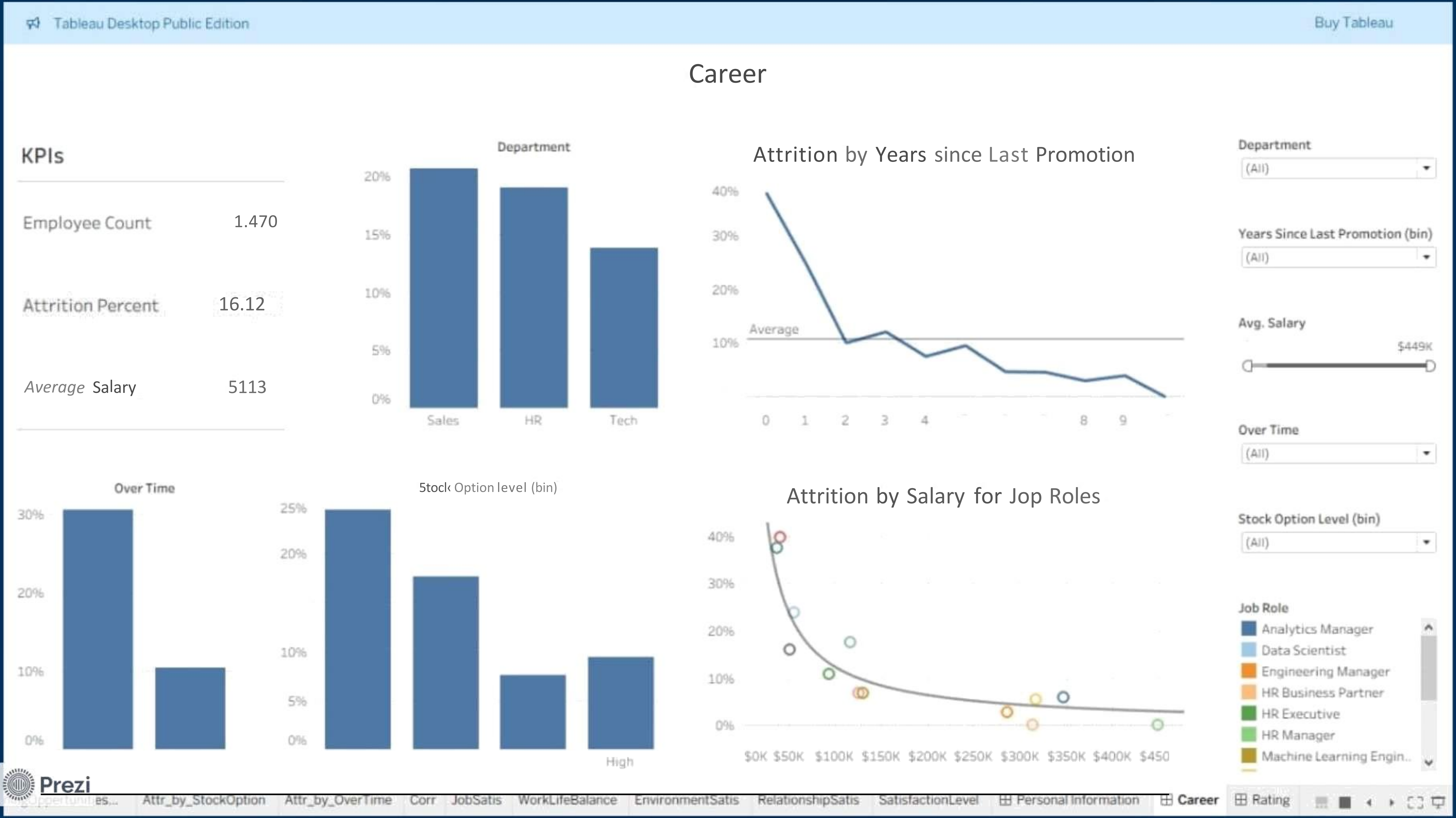
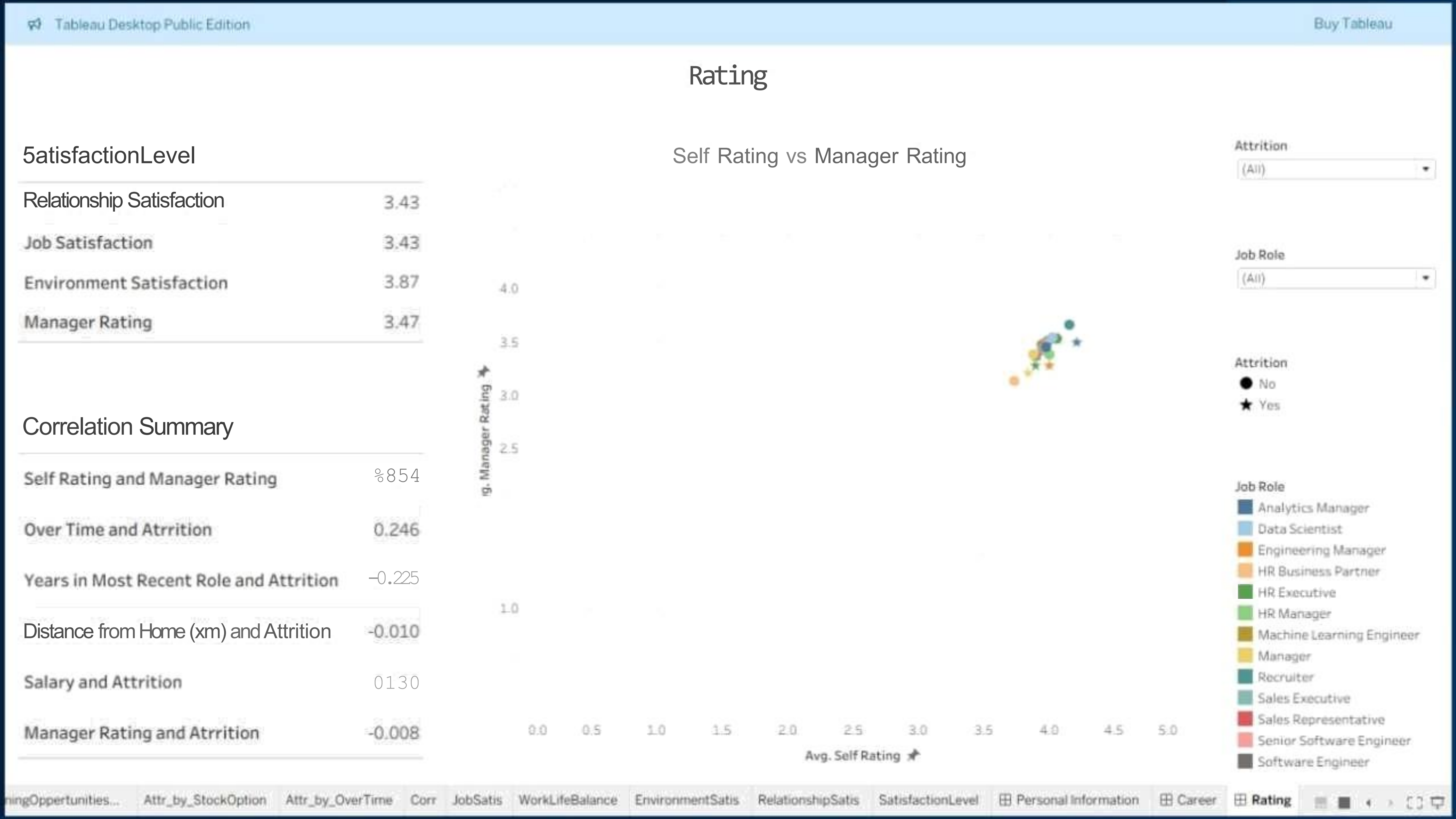


Tableau Desktop Public Edition
Buy Tableau
Career
KPIs
Employee Count 1.470
Attrition Percent 16.12
Average Salary 5113
Department
20%
15%
10%
5%
0%
Sales
HR
Tech
Attrition by Years since Last Promotion
40%
30%
20%
10%
Average
0
1
2
3
4
8
9
Over Time
30%
20%
10%
0%
Stock Option level (bin)
25%
20%
10%
5%
0%
High
Attrition by Salary for Jop Roles
40%
30%
20%
10%
0%
$0K $50K $100K $150K $200K $250K $300K $350K $400K $450
Department
(All)
Years Since Last Promotion (bin)
(All)
Avg. Salary
$449K
Over Time
(All)
Stock Option Level (bin)
(All)
Job Role
Analytics Manager
Data Scientist
Engineering Manager
HR Business Partner
HR Executive
HR Manager
Machine Learning Engin..
Prezi
Attr_by_StockOption
Attr_by_OverTime
Corr
JobSatis
WorkLifeBalance
EnvironmentSatis
RelationshipSatis
SatisfactionLevel
Personal Information
Career
Rating

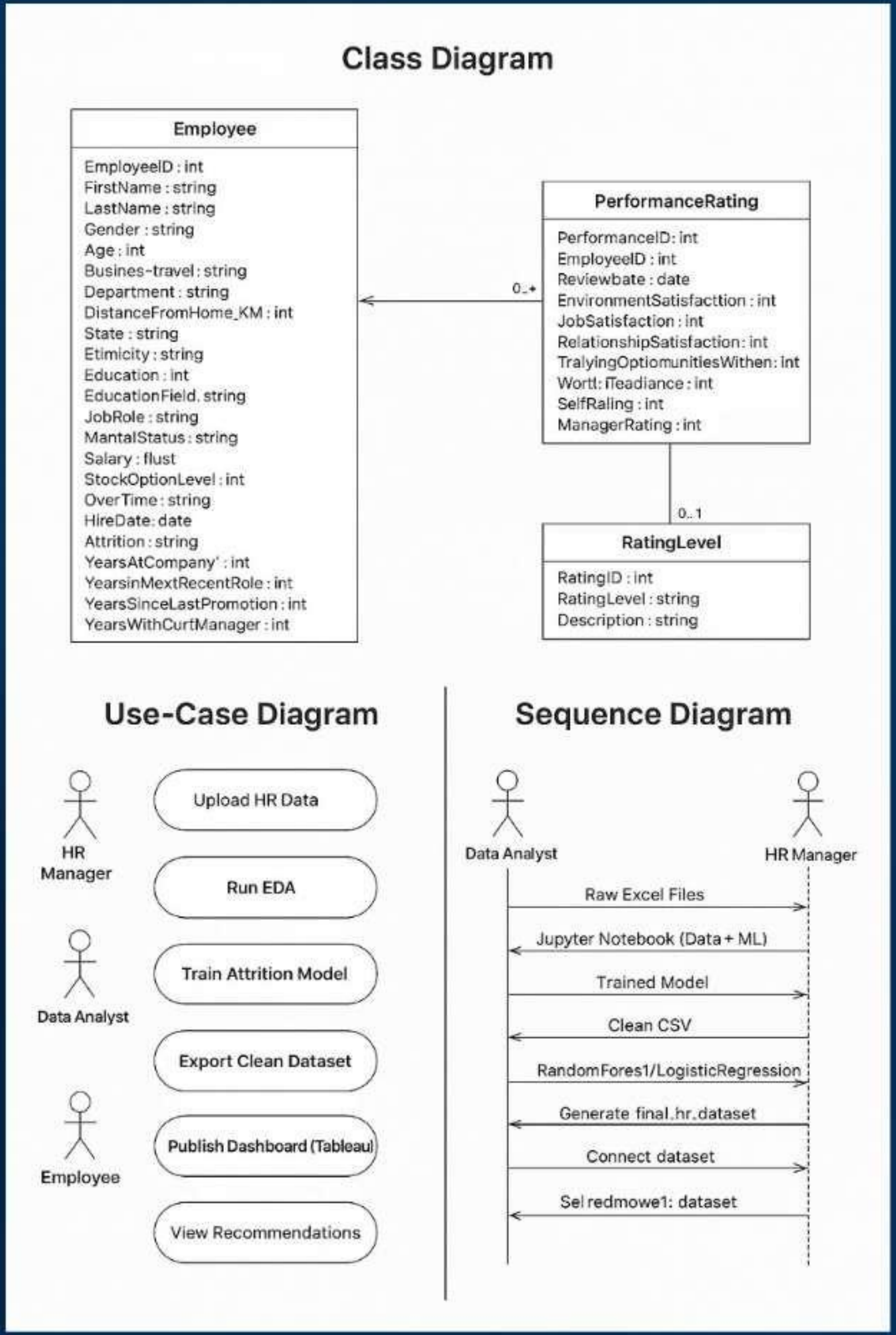
Dashboard

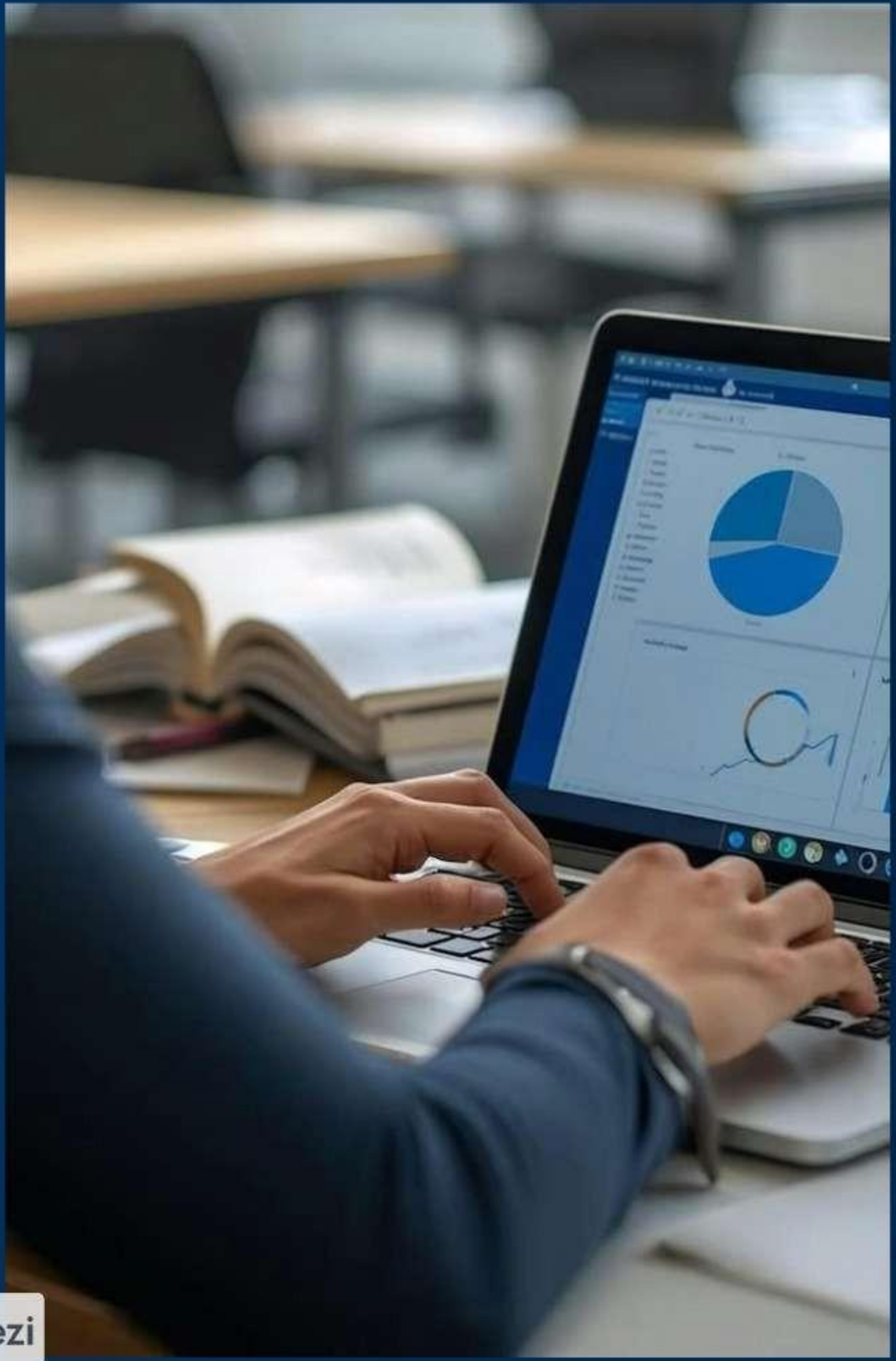
- Interactive KPIs
- Attrition by department & role
- Salary visualization
- Satisfaction metrics
- Promotions & travel insights



Future Development

- Predict attrition with ML
- HR system automation
- Alerts & recommendations
- Enterprise-level analytics product
- Dashboard integration with HR databases





Conclusion

- Full HR analytics pipeline
- Clear insights
- Ready for real-world application
- Strong foundation for future expansion

Thank You