HR Employee Attrition Analysis

OVERVIEW

This project analyzes employee attrition data to understand the factors that contribute to employee turnover.

Using a dataset from an HR department, I created a comprehensive Tableau dashboard to visualize key metrics such as attrition rate, total number of employee and so on.

OBJECTIVE

THE MAIN GOALS OF THIS PROJECT ARE:

TO IDENTIFY TRENDS AND PATTERNS IN EMPLOYEE ATTRITION.

TO PROVIDE ACTIONABLE INSIGHTS FOR HR TO REDUCE TURNOVER.

TO GAIN HANDS-ON EXPERIENCE WITH POWER BI FOR DATA VISUALIZATION.

Importing the required Libraries and cleaning the Dataset

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')

data = pd.read_csv(r"C:\Users\91931\Downloads\Intern\Dataset\HR-Employee
Attrition.csv")

#Checking for duplicated values
data.duplicated().sum()

#Checking for null values
data.isnull().sum()
```

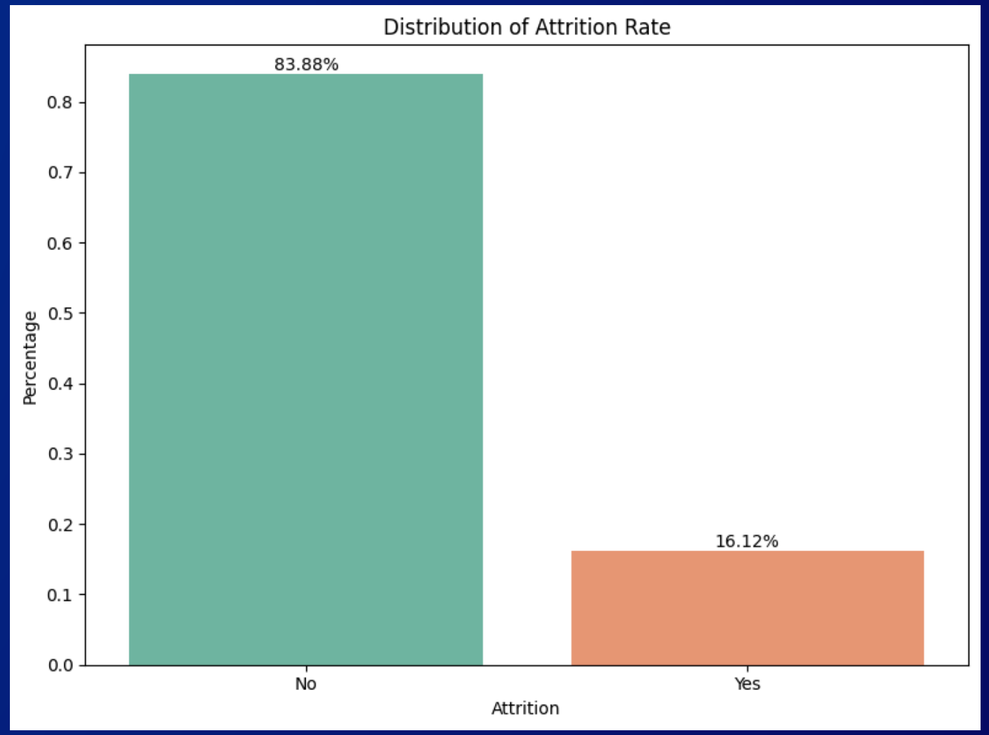
Exploratory Data Analysis

```
# Attritionrate: The attrition rate measures the percent age of employees
who leave the company in a given period of time. It is usually calculated
with in a year and is expressed as a percentage of the total number of
employees.

data['Attrition'].value_counts(normalize=True)

attrition = data['Attrition'].value_counts(normalize=True)

plt.figure(figsize=(8,6))
ax = sns.barplot(x=attrition.index, y=attrition,palette='Set2')
for p in ax.patches:ax.annotate(f'{p.get_height() * 100:.2f}%',
(p.get_x() + p.get_width() / 2.,p.get_height()),
ha='center', va='bottom')
plt.title('Distribution of Attrition Rate')
plt.xlabel('Attrition')
plt.ylabel('Percentage')
plt.tight_layout()
plt.show()
```



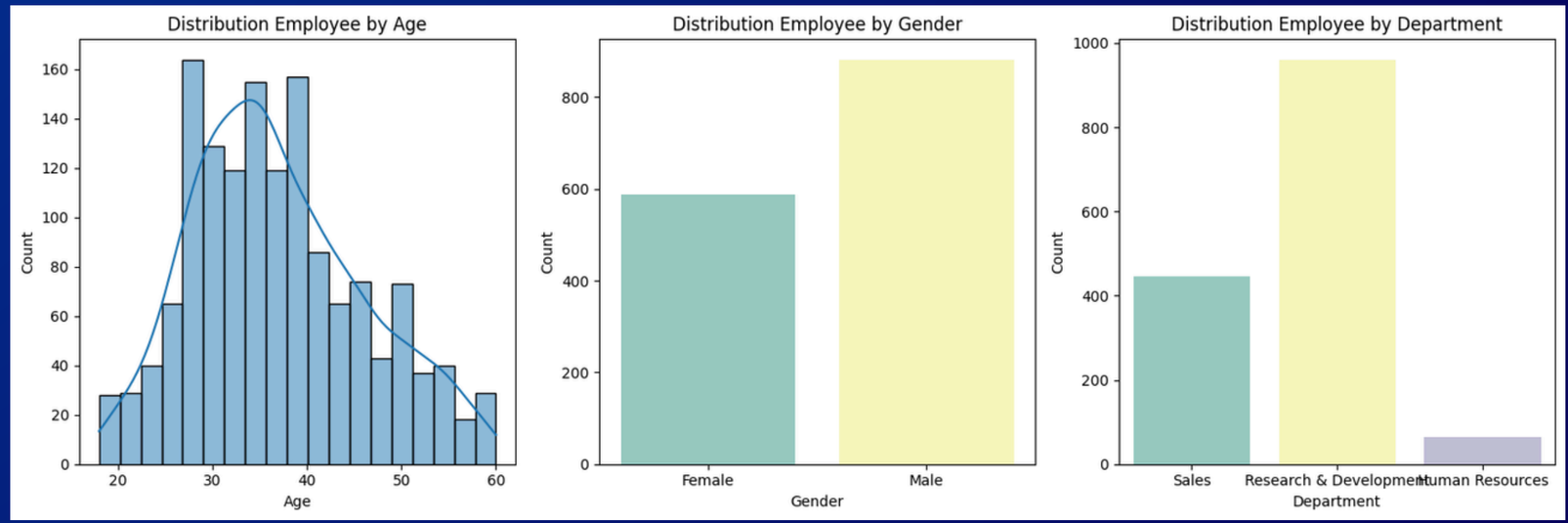
Employee's Demographics

```
fig, axes = plt.subplots(nrows=1, ncols=3, figsize=(15,5))
sns.histplot(data=data, x='Age', kde=True, ax=axes[0])
axes[0].set_title('Distribution Employee by Age')
axes[0].set_xlabel('Age')
axes[0].set_ylabel('Count')

sns.countplot(data=data, x='Gender', ax=axes[1], palette='Set3')
axes[1].set_title('Distribution Employee by Gender')
axes[1].set_xlabel('Gender')
axes[1].set_ylabel('Count')

sns.countplot(data=data, x='Department', ax=axes[2], palette='Set3')
axes[2].set_title('Distribution Employee by Department')
axes[2].set_xlabel('Department')
axes[2].set_ylabel('Count')

plt.tight_layout()
plt.show()
```



Creating a New Function

```
# Make a new function called calculation attrition rate
def calculate_attrition_rate(data,column):
  attrition_counts =
data.groupby([column,'Attrition']).size().unstack(fill_value=0)
  attrition_rate = attrition_counts['Yes'] / attrition_counts.sum(axis=1) *
100
  attrition_rate_data = attrition_rate.reset_index()
  attrition_rate_data.columns = [column, 'AttritionRate']
  return attrition_rate_data
```

Attrition by Age and Gender

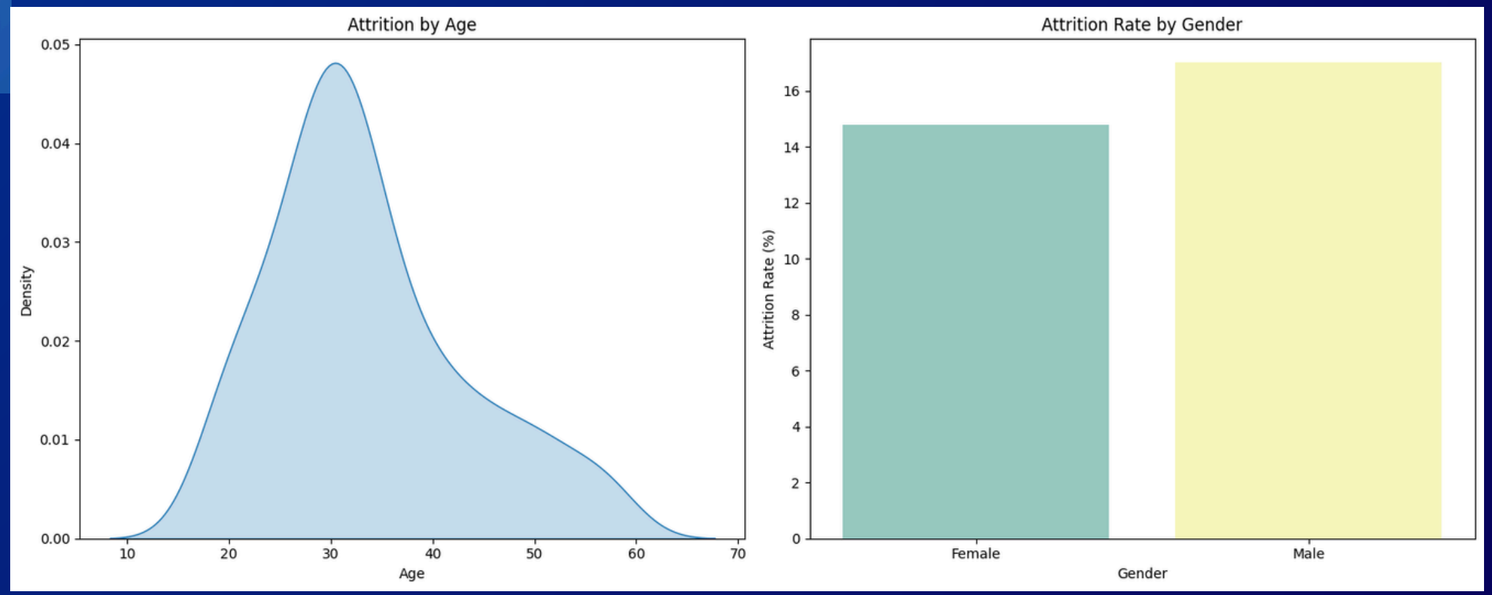
```
fig, axes = plt.subplots(nrows=1, ncols=2, figsize=(15,6))

# Plot 1: KDE plot of Age with Attrition hue

sns.kdeplot(data=data_attrition, x='Age', fill=True, ax=axes[0])
axes[0].set_title('Attrition by Age')
axes[0].set_xlabel('Age')
axes[0].set_ylabel('Density')

 # Plot 2: Bar plot of Gender count with Attrition hue

attrition_rate_data = calculate_attrition_rate(data, 'Gender')
sns.barplot(data=attrition_rate_data, x='Gender',
y='AttritionRate', ax=axes[1], palette='Set3')
axes[1].set_title(f'Attrition Rate by Gender')
axes[1].set_xlabel('Gender')
axes[1].set_ylabel('Attrition Rate (%)')
plt.tight_layout()
plt.show()
```



CONCLUSION

- The analysis of the HR dataset reveals important insights regarding employee attrition and job satisfaction within the organization.
- A high attrition rate was observed among employees aged 25-34, indicating a potential need for enhanced career development opportunities, better compensation packages, and improved work-life balance initiatives.
- By implementing targeted strategies such as leadership development programs, improved employee engagement initiatives, and enhanced support for career progression, the organization can reduce turnover and boost overall job satisfaction.
- These actions are expected to lead to increased employee retention, better performance, and a healthier organizational culture.
- The findings and recommendations from this analysis serves as a foundation for to make data-driven decisions aimed at improving employee experience and driving sustainable growth.

Dashboard

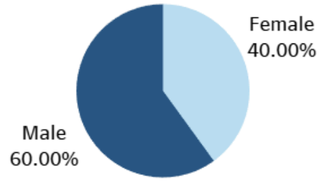
Avg age by department and job role

		Female	Male
Human Resources	Manager	50.50	47.71
	Human Resources	34.44	35.97
Research & Development	Manager	47.50	44.73
	Research Director	41.91	45.47
	Healthcare Representati..	38.90	40.39
	Manufacturing Director	38.64	37.96
	Research Scientist	33.99	34.39
	Laboratory Technician	35.27	33.52
Sales	Manager	47.32	47.39
	Sales Executive	37.83	36.25
	Sales Representative	30.05	30.62

Average years at company
7.008

Total employees
1,470

Total employees by gender



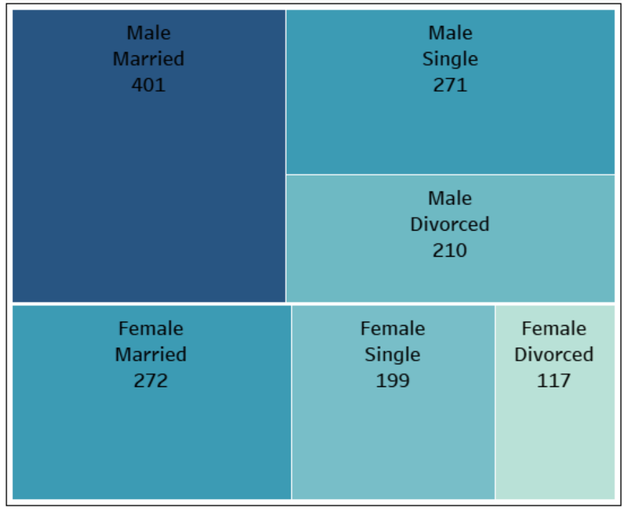
Average total working years
11.28

Average age
36.92

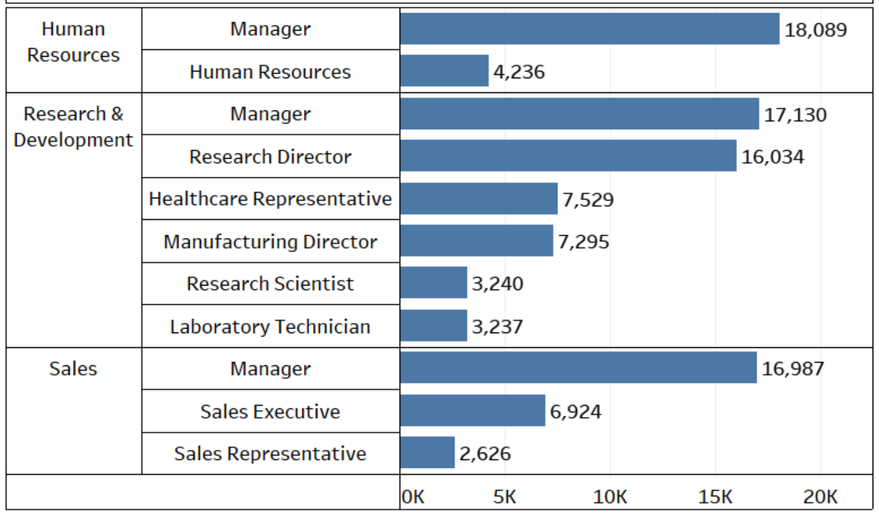
Gender and job role

	Female	Male
Sales Executive	132	194
Research Scientist	114	178
Laboratory Technician	85	174
Manufacturing Director	72	73
Healthcare Representati..	51	80
Manager	47	55
Sales Representative	38	45
Research Director	33	47
Human Resources	16	36

Gender and marital status



Avg monthly income by department and job role



Total employees by department and job role

Human Resources	Human Resources	52
	Manager	11
Research & Development	Research Scientist	292
	Laboratory Technician	259
	Manufacturing Director	145
	Healthcare Representative	131
	Research Director	80
	Manager	54
Sales	Sales Executive	326
	Sales Representative	83
	Manager	37