HR Analytics (Data Analyst)

A Mahesh Reddy

[Unified Mentor]

1.Introduction

- **Objective**: The main goal of this project is to analyze the HR dataset using data visualization techniques to understand key trends and factors that influence employee behavior and outcomes (such as retention or performance).
- **Data Source**: The dataset is sourced from Kaggle and includes various employee metrics like age, department, satisfaction levels, performance scores, and whether or not the employee has left the organization.

Objectives

The main objectives of this HR Analytics Visualization project are:

1. **Analyze Employee Data**: To explore the key characteristics of employee data (e.g., satisfaction levels, performance, years at the company) and their relationship with turnover.
2. **Identify Patterns**: To use visual techniques to discover trends and patterns in employee satisfaction, performance, and salary distribution across different departments.
3. **Predict Turnover Risk**: To highlight which employees are more likely to leave based on their profile (e.g., satisfaction level, salary, department).
4. **Provide Recommendations**: To offer data-driven insights for improving employee retention, focusing on high-risk employees with low satisfaction or poor performanc

3. Literature Survey

S.No	Paper/Study	Authors/Year	Key Insights	Techniques/Methods
1	Predicting Employee Attrition Using HR Data	Nguyen et al. (2020)	Machine learning models can predict employee attrition with high accuracy based on HR datasets.	Decision Trees, Random Forest, Logistic Regression
2	Visualization Techniques in HR Analytics	Bhardwaj et al. (2019)	Visualization is crucial for HR managers to interpret patterns in employee behavior, such as turnover trends.	Data visualization techniques (bar charts, heatmaps, etc.)
3	Employee Retention Strategies	Kaur & Sharma (2021)	Employee retention is influenced by factors like engagement, workplace culture, and career development.	Employee engagement analysis, qualitative methods
4	Understanding Satisfaction and Turnover	Singh & Gupta (2018)	Satisfaction levels and compensation are closely linked to employee turnover, especially in certain departments.	Satisfaction surveys, salary band analysis, turnover rates
5	Performance and Retention in Organizations	Martinez et al. (2022)	High performers are more likely to stay, while low satisfaction is a	Performance score analysis, retention modeling

	strong predictor of turnover.

4.Data Description

- The dataset contains several key features:
  - **Employee ID**: Unique identifier for each employee.
  - **Department**: Department in which the employee works.
  - **Satisfaction Level**: Employee satisfaction on a scale.
  - **Performance**: A numerical score of employee performance.
  - **Years at Company**: How long the employee has been with the company.
  - **Salary**: Categorical variable indicating salary level (low, medium, high).
  - **Status**: Whether the employee is still working or has left the company.

5. Key Insights and Visualizations

- **Employee Satisfaction Distribution**: A histogram showing the distribution of satisfaction levels among employees. The majority of employees have moderate to high satisfaction, but a small group is highly dissatisfied.
- **Satisfaction vs Turnover**: A scatter plot demonstrating that employees with lower satisfaction tend to leave the company more frequently.

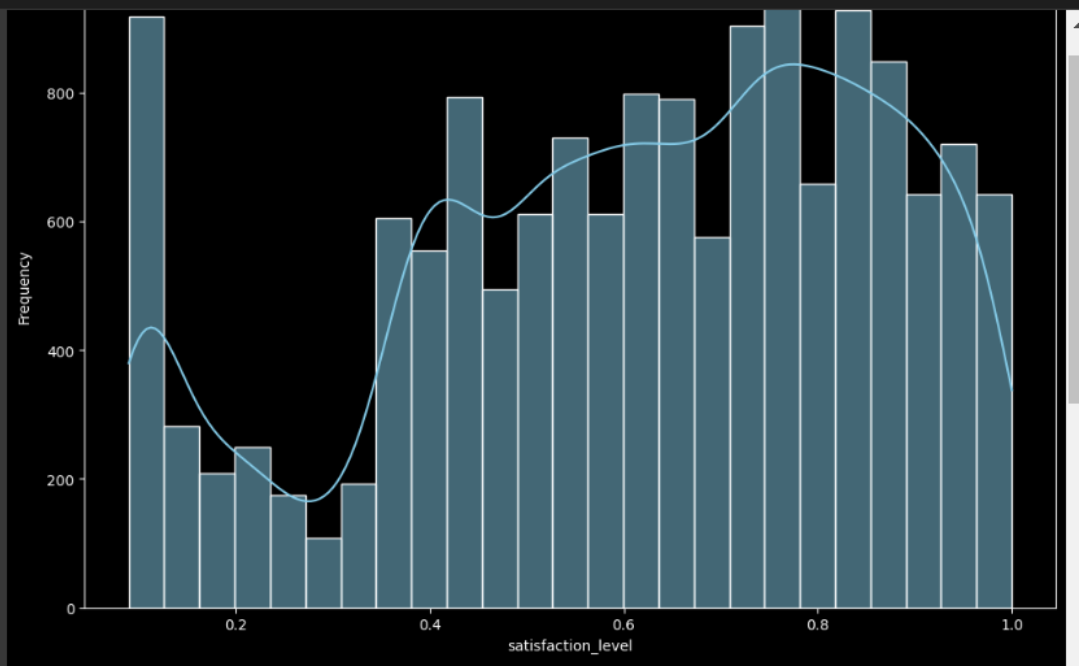
- **Salary and Turnover**: A bar chart showing the turnover rate across different salary bands. Employees with lower salaries are more likely to leave the company.
- **Department-wise Distribution**: A pie chart showing the distribution of employees across departments, with the largest number in departments like Sales, IT, and HR.
- **Years at Company**: A line graph showing how employee satisfaction and performance change with years at the company. There's often a dip in satisfaction after the first few years, indicating a potential challenge in employee retention.
- **Heatmap of Correlations**: A heatmap visualizing the correlations between different features, highlighting the strongest relationships between satisfaction, performance, and turnover.

➢ **Recommendations**

- **Employee Satisfaction Programs**: Introduce initiatives to improve employee satisfaction, particularly for those in lower salary bands and with lower performance scores.
- **Retention Strategies**: Focus on retaining high performers and those with moderate satisfaction by offering incentives, promotions, and salary increases.
- **Monitor Long-term Employees**: Employees who have been with the company for 3-5 years may be at higher risk of turnover due to declining satisfaction levels.

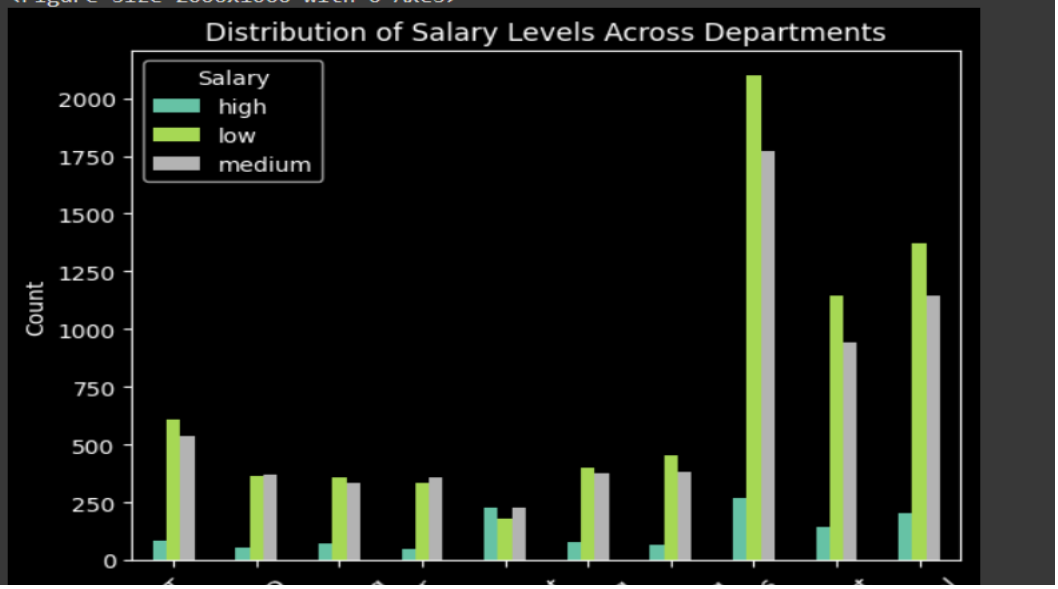
Outpts:

```python
# Distribution analysis
for var in col:
 plt.figure(figsize=(12, 8))
 sns.histplot(my_data[var], bins=25, kde=True,color='skyblue')
 plt.title(f'Distribution of {var}')
 plt.xlabel(var)
 plt.ylabel('Frequency')
 plt.show()
```




```python
plt.figure(figsize=(20, 10))
grouped_data.plot(kind='bar', cmap='Set2')
plt.title('Distribution of Salary Levels Across Departments')
plt.xlabel('Department')
plt.ylabel('Count')
plt.xticks(rotation=45)
plt.legend(title='Salary')
plt.show()
```

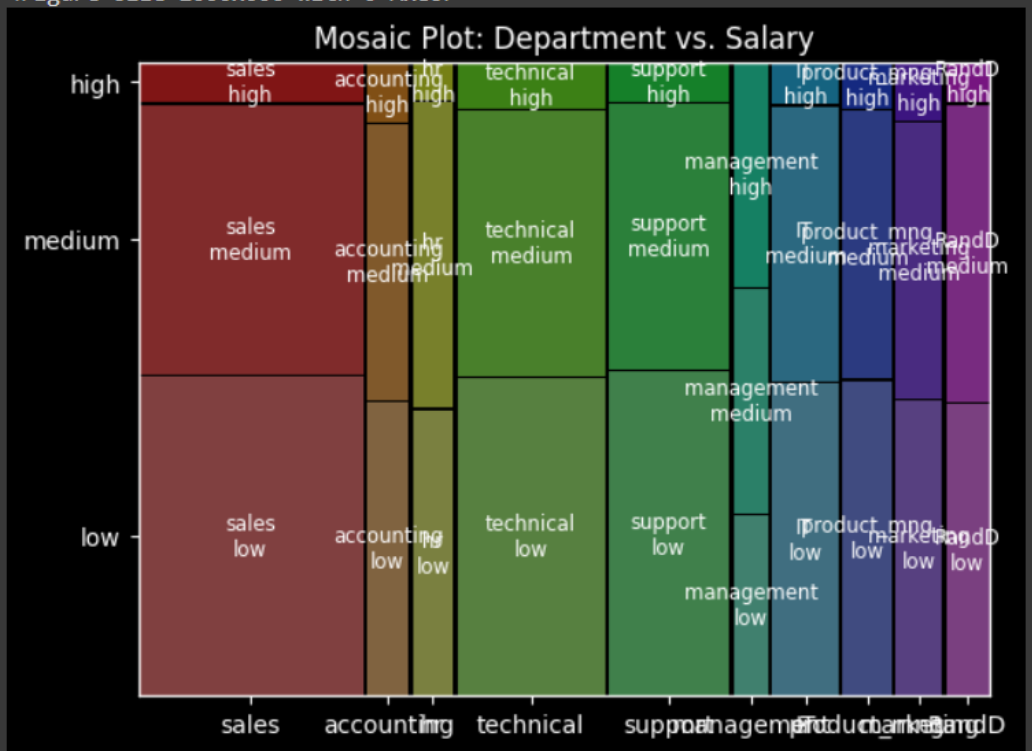
```
<Figure size 2000x1000 with 0 Axes>
```



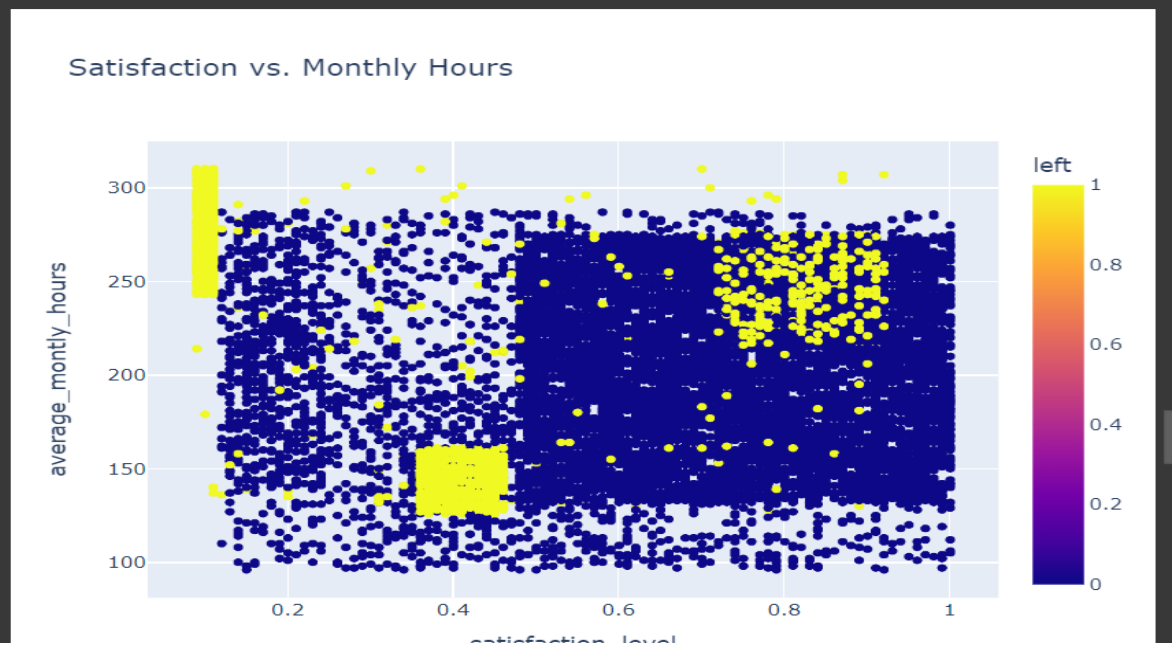
```python
from statsmodels.graphics.mosaicplot import mosaic

# Create a mosaic plot
plt.figure(figsize=(10, 6))
mosaic(my_data, ['Department', 'salary'], title='Mosaic Plot: Department vs. Salary')
plt.show()
```

```
<Figure size 1000x600 with 0 Axes>
```




```python
# Interactive visualization with Plotly
fig = px.scatter(my_data, x='satisfaction_level', y='average_montly_hours', color='lef
fig.show()
```



Conclusion

- The analysis of the HR data provided key insights into the factors affecting employee turnover and satisfaction. By focusing on specific areas like employee satisfaction and salary, companies can improve retention and overall performance.