

Data Visualization

Data in csv format followed by a graphical representation:

ID,Name,Age,Position,Department,Salary,Joining Date,Performance Score
1,Alice Smith,28,Software Engineer,IT,75000,2022-01-15,85
2,Bob Johnson,35,Data Scientist,Data,90000,2019-05-10,90
3,Charlie Brown,40,Manager,HR,95000,2018-03-20,88
4,David Williams,29,UX Designer,Design,70000,2021-09-01,80
5,Eva Davis,32,Database Admin,IT,78000,2020-02-25,82
6,Frank Miller,45,CTO,Executive,150000,2017-06-15,92
7,Grace Lee,38,Product Manager,Product,85000,2019-11-05,87
8,Hank Garcia,30,Data Engineer,Data,82000,2021-04-10,84
9,Ivy Wilson,27,Frontend Developer,IT,72000,2022-03-01,81
10,Jack White,42,CEO,Executive,200000,2016-07-20,95

ID	Name	Age	Position	Department	Salary	Joining Date	Performance Score
1	Alice Smith	28	Software Engineer	IT	75000	2022-01-15	85
2	Bob Johnson	35	Data Scientist	Data	90000	2019-05-10	90
3	Charlie Brown	40	Manager	HR	95000	2018-03-20	88
4	David Williams	29	UX Designer	Design	70000	2021-09-01	80
5	Eva Davis	32	Database Admin	IT	78000	2020-02-25	82
6	Frank Miller	45	CTO	Executive	150000	2017-06-15	92
7	Grace Lee	38	Product Manager	Product	85000	2019-11-05	87
8	Hank Garcia	30	Data Engineer	Data	82000	2021-04-10	84
9	Ivy Wilson	27	Frontend Developer	IT	72000	2022-03-01	81
10	Jack White	42	CEO	Executive	200000	2016-07-20	95

Prompt(s):

The answers will be generated in the form of a Python script.

Departmental Distribution: Create a bar chart visualizing the number of employees in each department.

Salary Distribution: Can you generate a box plot showcasing the salary distribution of employees, highlighting any outliers?

Performance Heatmap: Design a heatmap that visualizes the performance scores of employees across different departments.

Tenure vs. Performance: Plot a scatter plot to visualize the relationship between an employee's tenure in the company and their performance score.

Age Distribution: Create a histogram showcasing the age distribution of employees in the company.

GPT-4

Departmental Distribution:

1. Bar chart for Departmental Distribution

```
plt.figure(figsize=(10, 5))
df['Department'].value_counts().plot(kind='bar', color='skyblue')
plt.title('Departmental Distribution')
plt.ylabel('Number of Employees')
plt.xlabel('Department')
plt.show()
```

Salary Distribution:

2. Salary Distribution with Outliers

```
plt.figure(figsize=(10, 5))
sns.boxplot(x=df['Salary'], color='salmon')
plt.title('Salary Distribution with Outliers')
plt.xlabel('Salary')
plt.show()
```

Performance Heatmap:

3. Performance Heatmap

```
heatmap_data = df.pivot_table(index='Department', values='Performance Score', aggfunc='mean')
plt.figure(figsize=(10, 5))
sns.heatmap(heatmap_data, cmap='viridis', annot=True)
plt.title('Performance Scores Across Different Departments')
plt.show()
```

Tenure vs. Performance:

4. Tenure vs. Performance

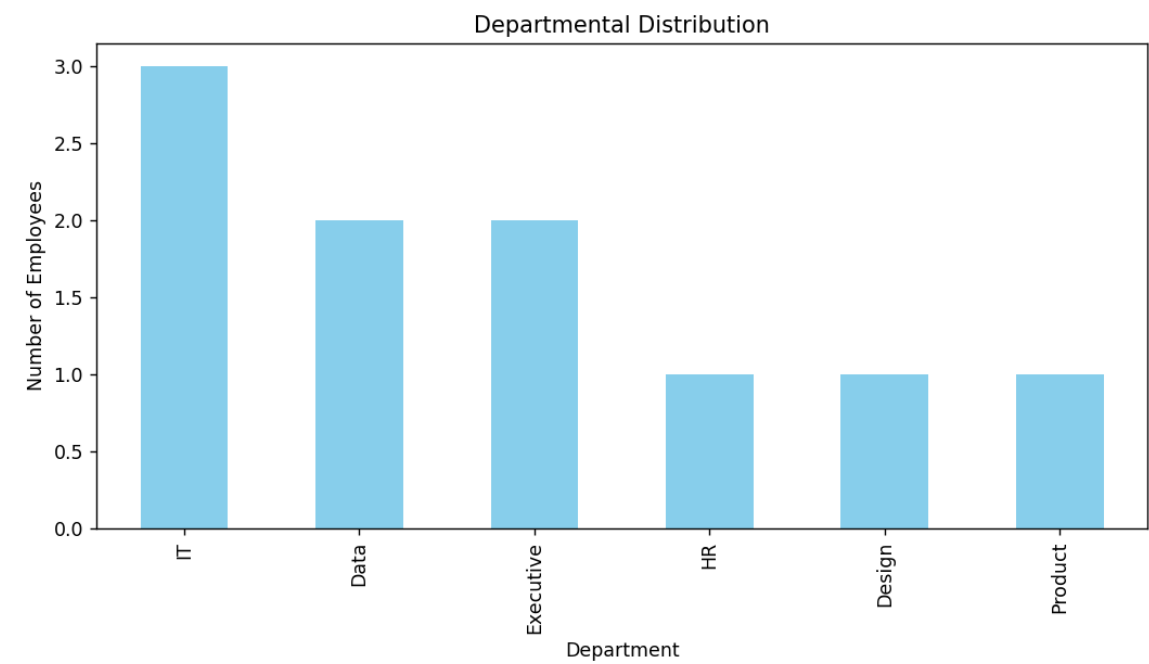
```
df['Joining Date'] = pd.to_datetime(df['Joining Date'])
current_year = 2023
df['Tenure'] = current_year - df['Joining Date'].dt.year
plt.figure(figsize=(10, 5))
sns.scatterplot(x=df['Tenure'], y=df['Performance Score'], hue=df['Department'], palette='tab10', s=100)
plt.title('Tenure vs. Performance Score')
plt.xlabel('Years of Tenure')
plt.ylabel('Performance Score')
plt.show()
```

Age Distribution:

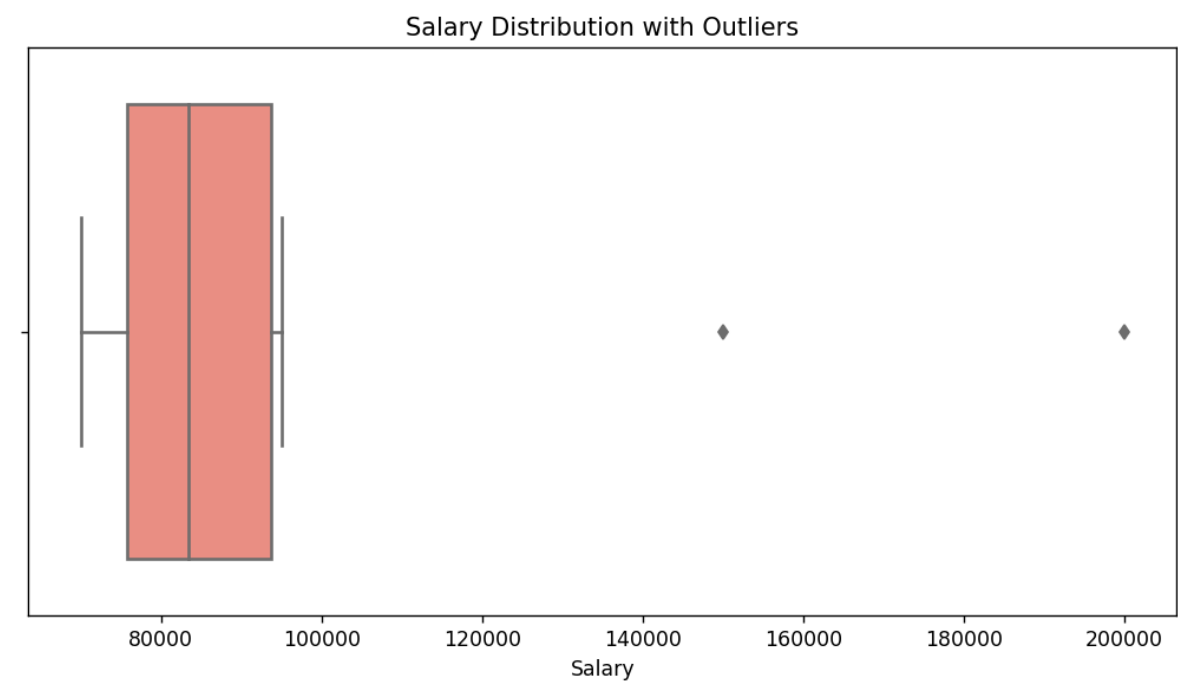
5. Age Distribution

```
plt.figure(figsize=(10, 5))
sns.histplot(df['Age'], bins=15, color='mediumseagreen')
plt.title('Age Distribution of Employees')
plt.xlabel('Age')
plt.ylabel('Number of Employees')
plt.show()
```

Departmental Distribution:



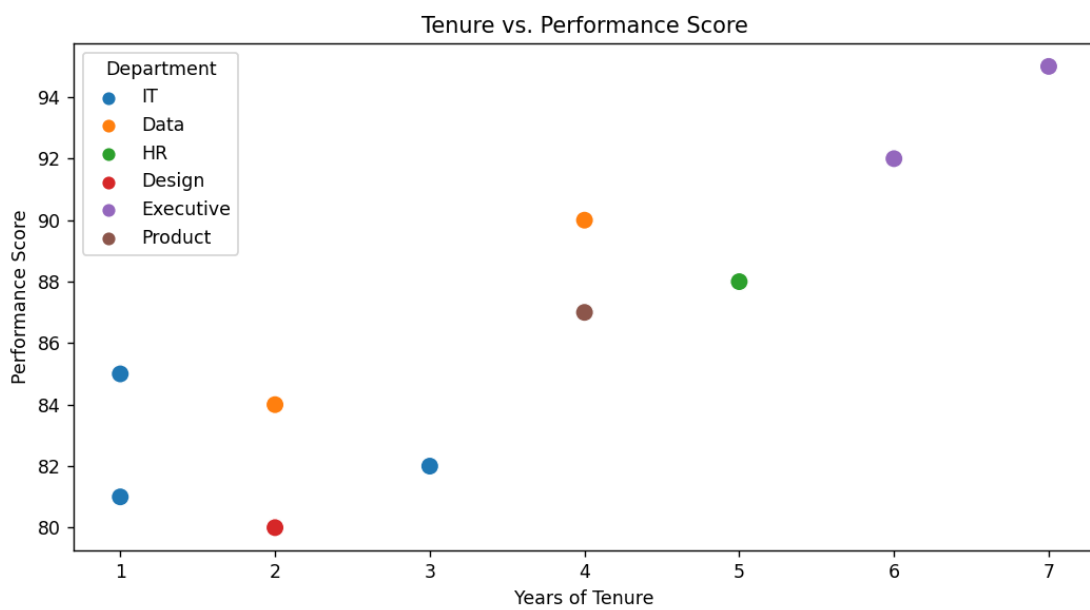
Salary Distribution:



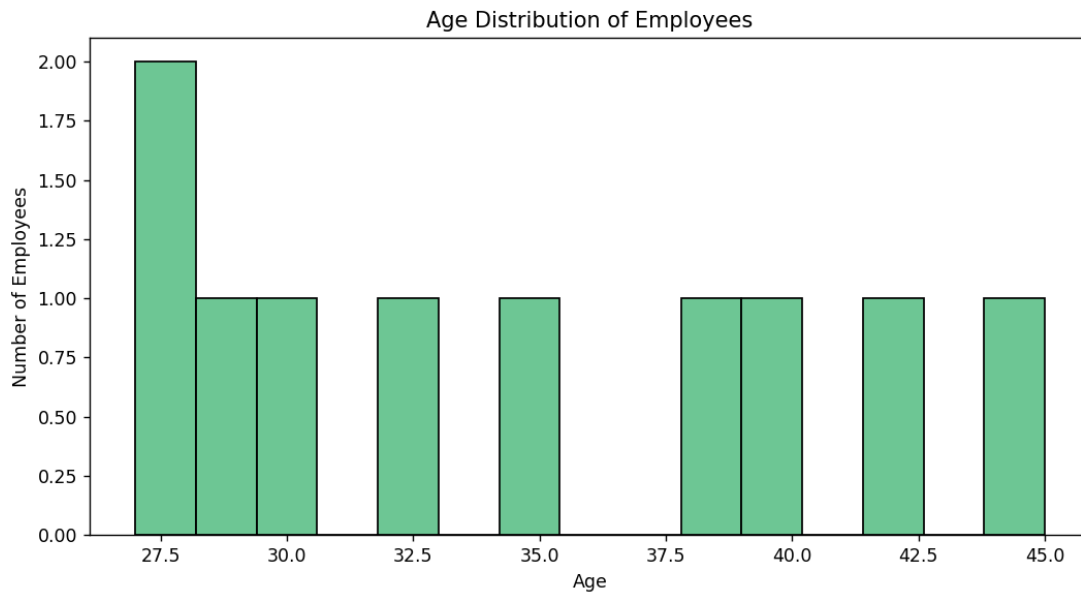
Performance Heatmap:



Tenure vs. Performance:



Age Distribution:



GPT-3.5

Departmental Distribution:

Prompt 1: Departmental Distribution - Bar Chart

```
department_counts = data['Department'].value_counts()
department_counts.plot(kind='bar', xlabel='Department', ylabel='Number of Employees',
title='Departmental Distribution')
plt.show()
```

Salary Distribution:

Prompt 2: Salary Distribution - Box Plot

```
plt.figure(figsize=(8, 6))
sns.boxplot(x='Salary', data=data, showfliers=True)
plt.xlabel('Salary')
plt.title('Salary Distribution (with Outliers)')
plt.show()
```

Performance Heatmap:

Prompt 3: Performance Heatmap

```
heatmap_data = data.pivot_table(values='Performance Score', index='Department', columns='Name',
fill_value=0)
plt.figure(figsize=(10, 6))
sns.heatmap(heatmap_data, cmap='coolwarm', annot=True, fmt=".1f", cbar=True)
plt.title('Performance Heatmap by Department')
plt.show()
```

Tenure vs. Performance:

Prompt 4: Tenure vs. Performance - Scatter Plot

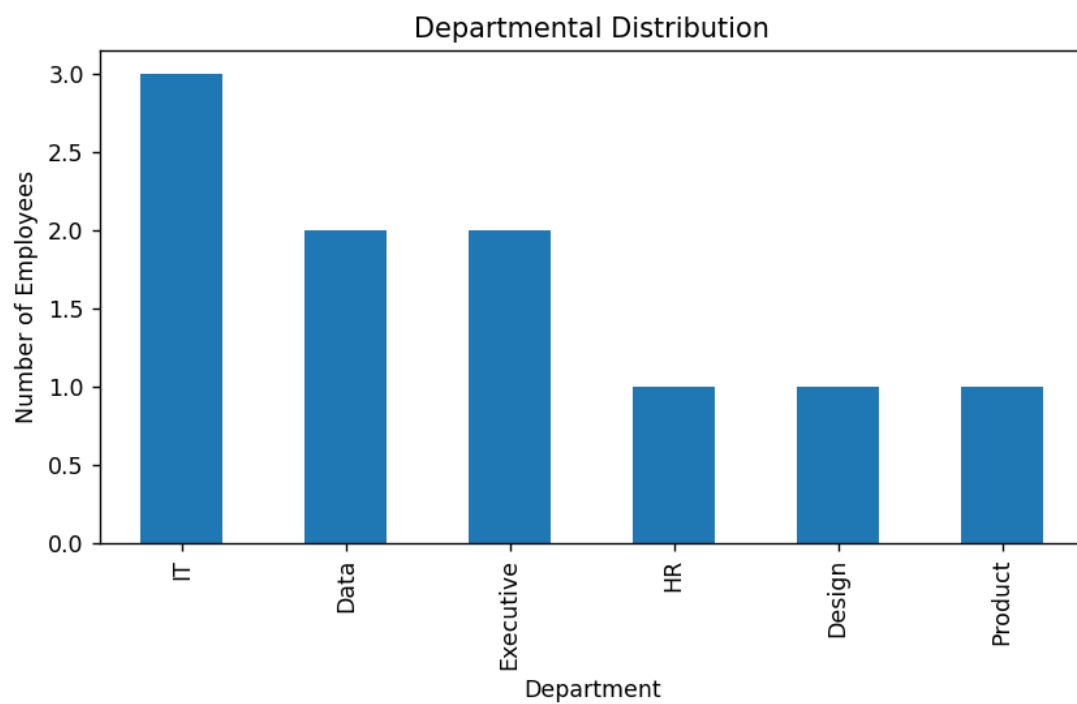
```
data['Joining Date'] = pd.to_datetime(data['Joining Date'])
data['Tenure'] = (pd.to_datetime('now') - data['Joining Date']).dt.days / 365.0
plt.figure(figsize=(8, 6))
plt.scatter(data['Tenure'], data['Performance Score'], alpha=0.5)
plt.xlabel('Tenure (Years)')
plt.ylabel('Performance Score')
plt.title('Tenure vs. Performance Score')
```

```
plt.grid(True)
plt.show()
```

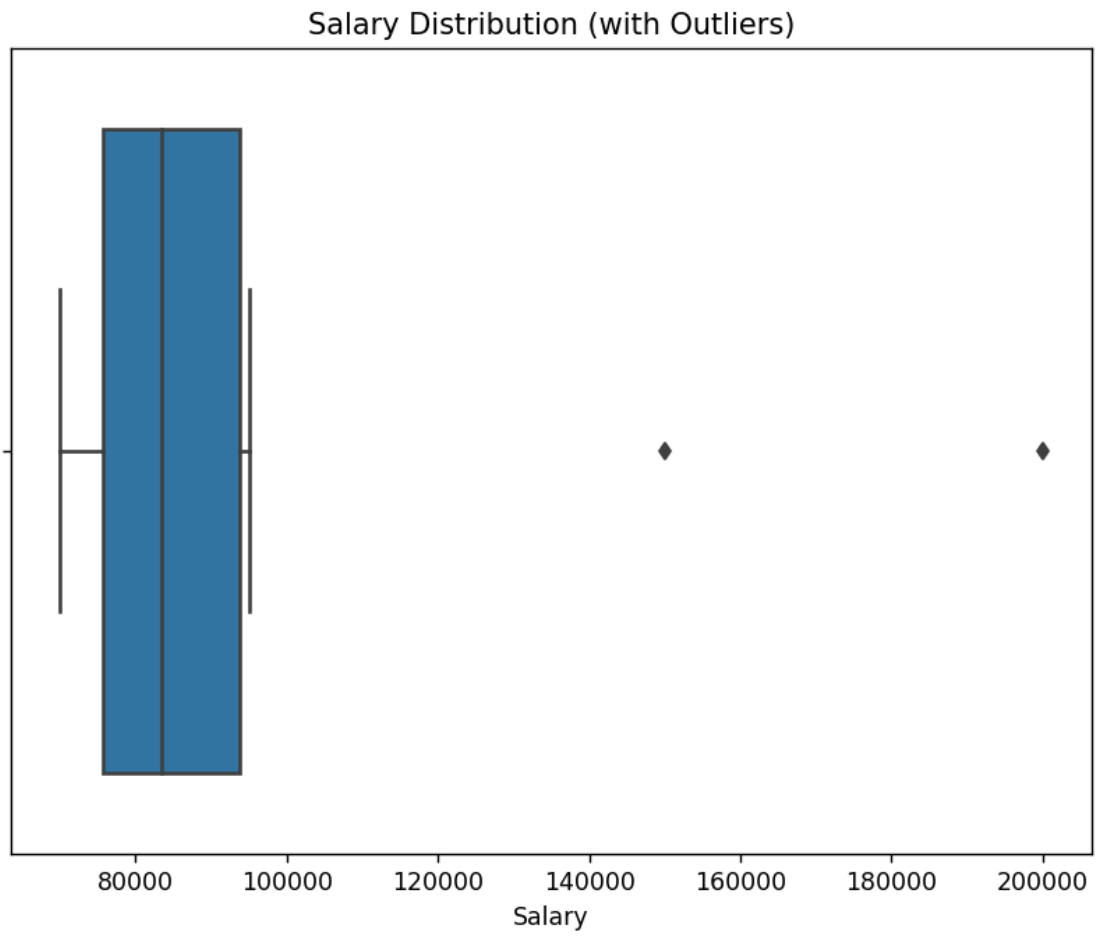
Age Distribution:

```
# Prompt 5: Age Distribution - Histogram
plt.figure(figsize=(8, 6))
plt.hist(data['Age'], bins=10, edgecolor='k')
plt.xlabel('Age')
plt.ylabel('Number of Employees')
plt.title('Age Distribution of Employees')
plt.show()
```

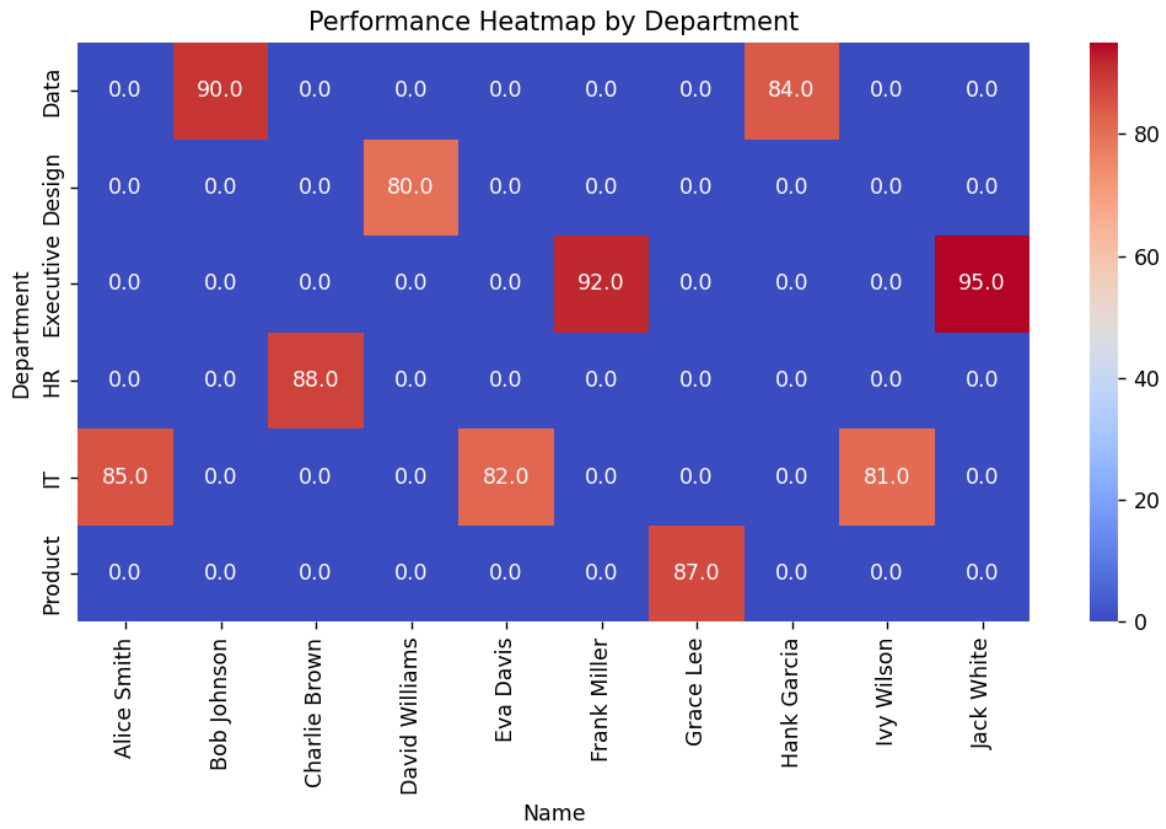
Departmental Distribution:



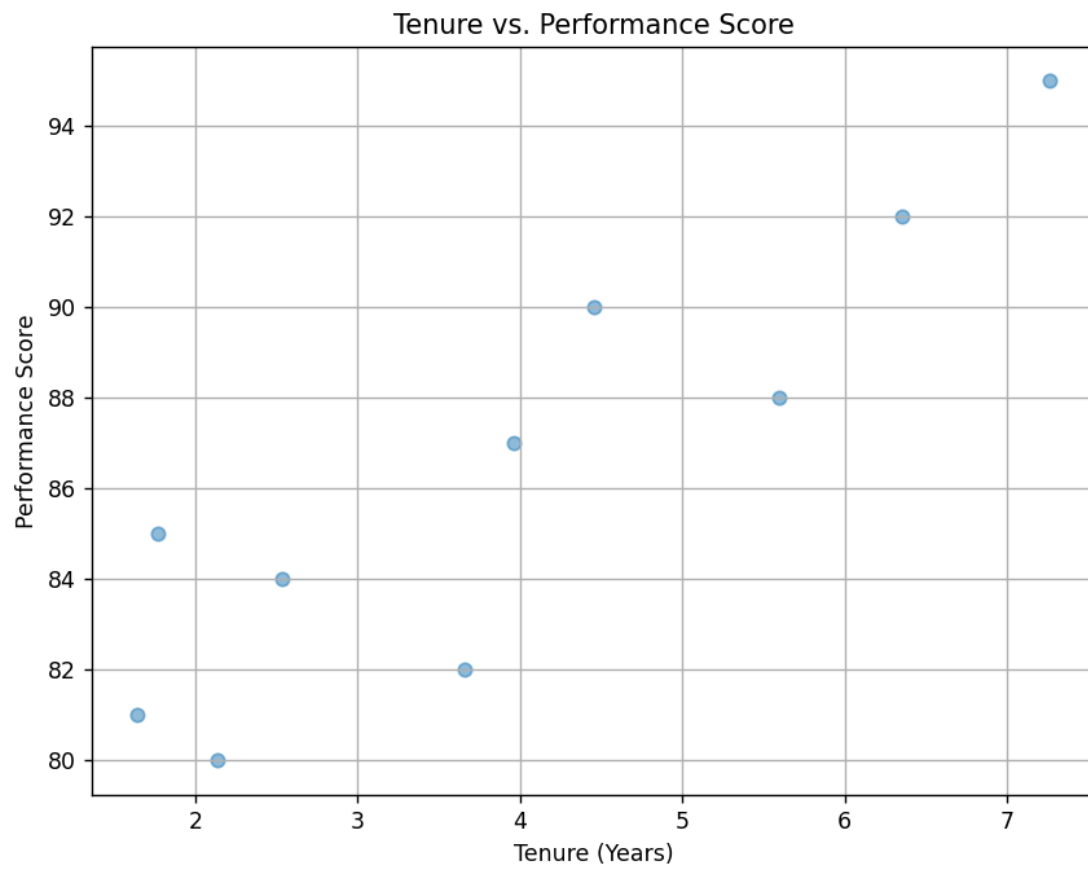
Salary Distribution:



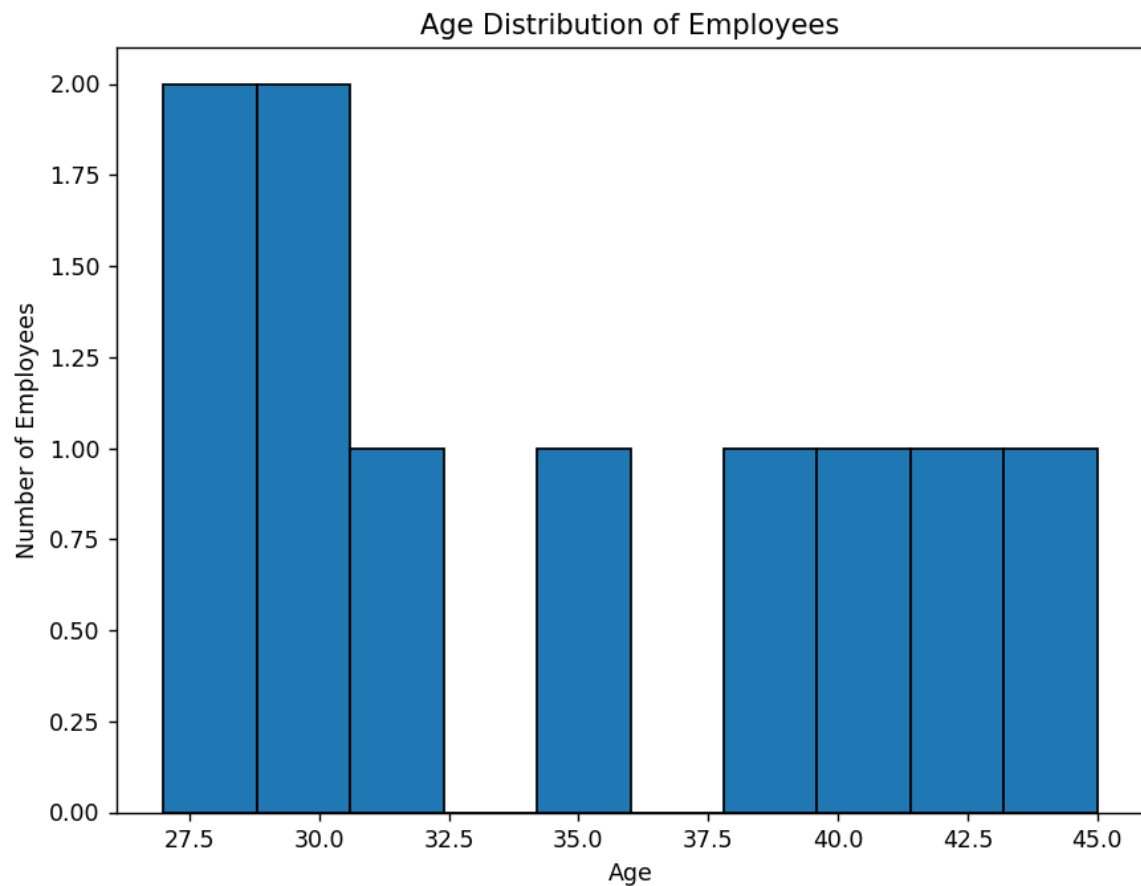
Performance Heatmap:
Unnecessarily complex



Tenure vs. Performance:



Age Distribution:



BingAI

Departmental Distribution:

```
# Departmental Distribution
plt.figure(figsize=(10,6))
sns.countplot(x='Department', data=df)
plt.title('Departmental Distribution')
plt.show()
```

Salary Distribution:

```
# Salary Distribution
plt.figure(figsize=(10,6))
sns.boxplot(x='Salary', data=df)
plt.title('Salary Distribution')
plt.show()
```

Performance Heatmap:

```
# Performance Heatmap
performance = df.pivot_table(index='Name', columns='Department', values='Performance Score')
plt.figure(figsize=(10,6))
sns.heatmap(performance)
plt.title('Performance Heatmap')
plt.show()
```

Tenure vs. Performance:

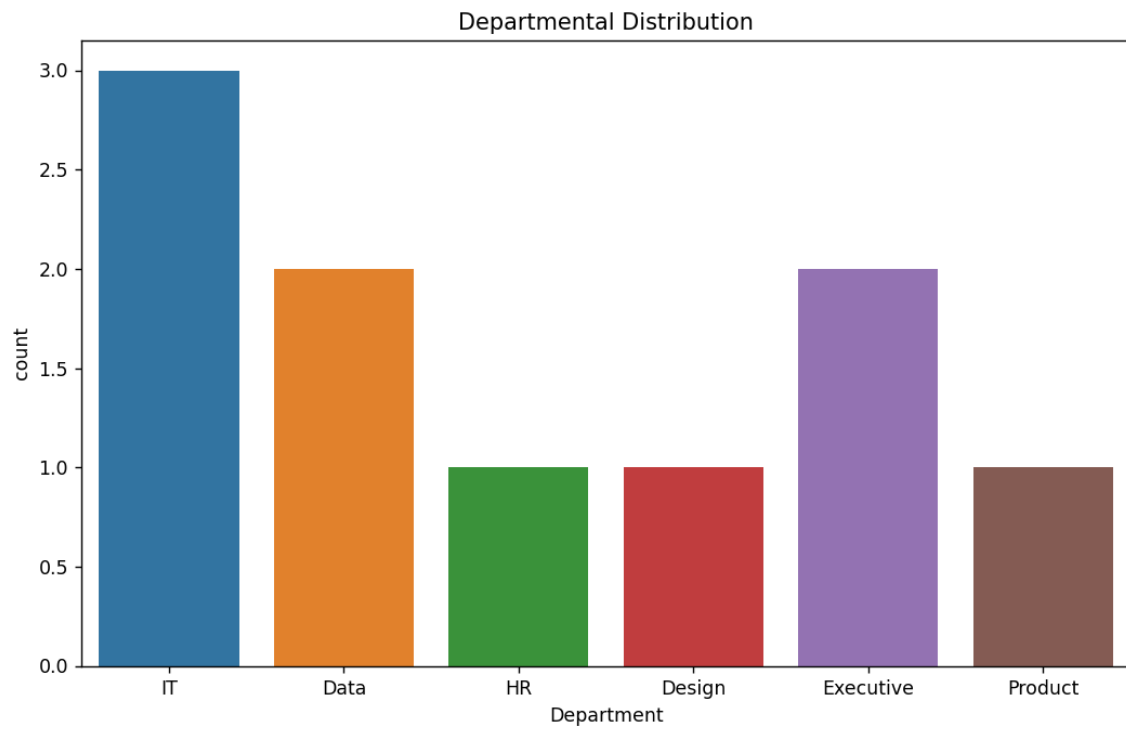
```
# Tenure vs. Performance
plt.figure(figsize=(10,6))
sns.scatterplot(x='Tenure', y='Performance Score', data=df)
```

```
plt.title('Tenure vs. Performance')  
plt.show()
```

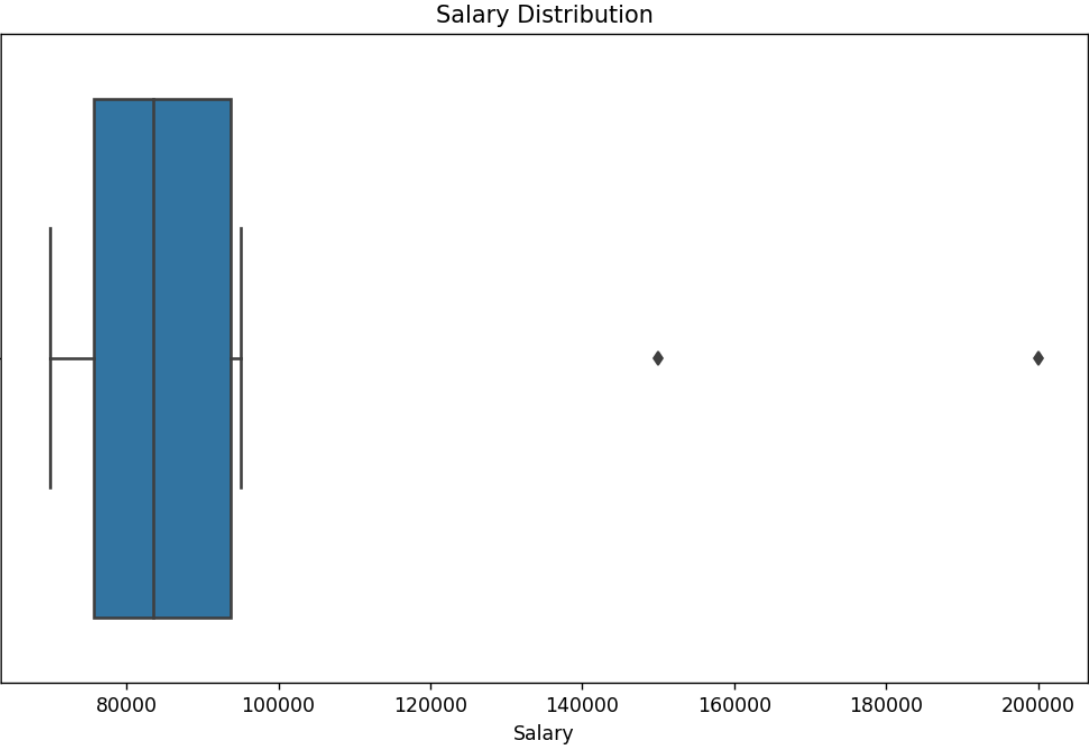
Age Distribution:

```
# Age Distribution  
plt.figure(figsize=(10,6))  
sns.histplot(df['Age'], bins=10)  
plt.title('Age Distribution')  
plt.show()
```

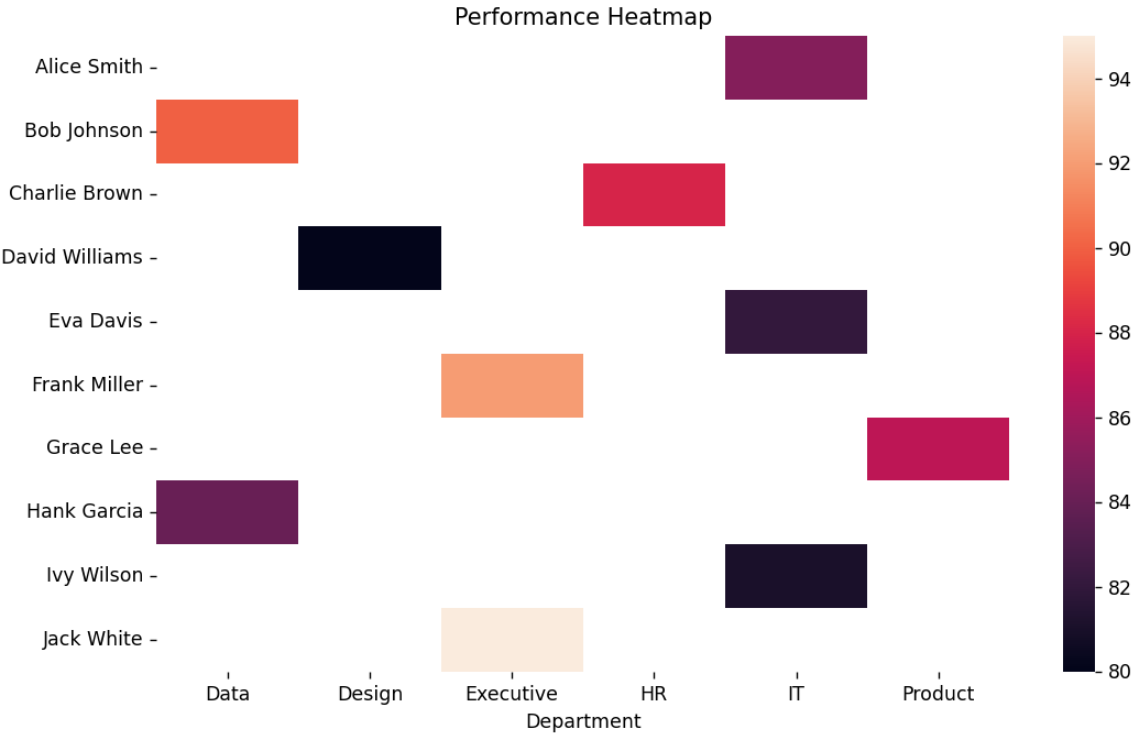
Departmental Distribution:



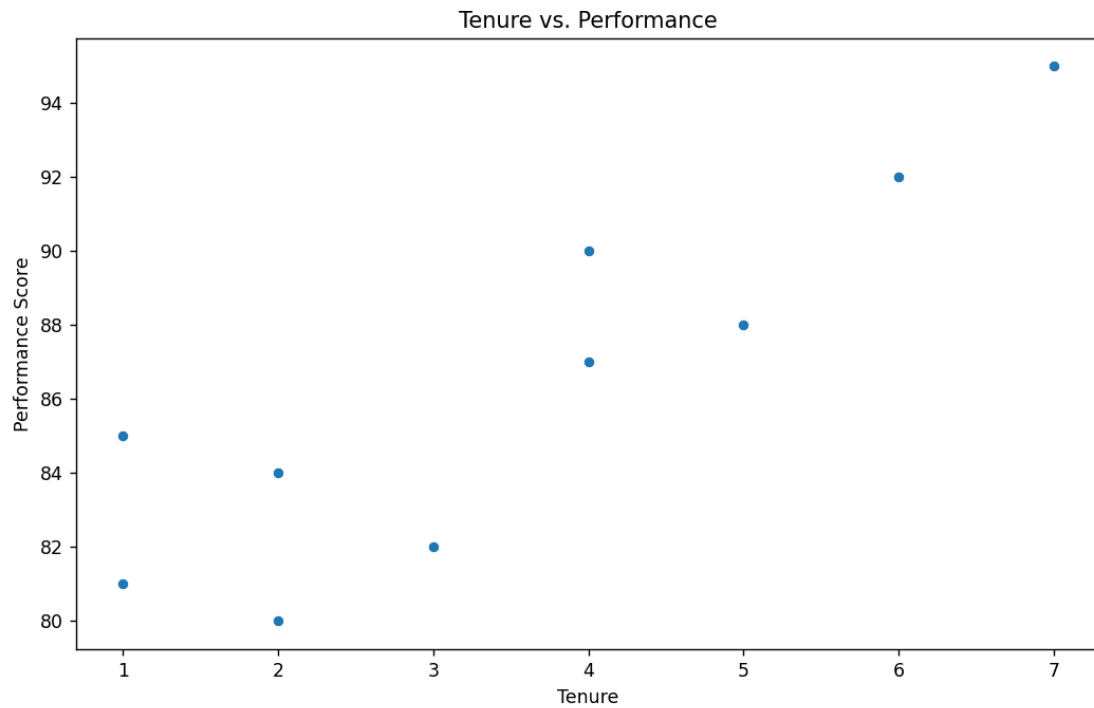
Salary Distribution:



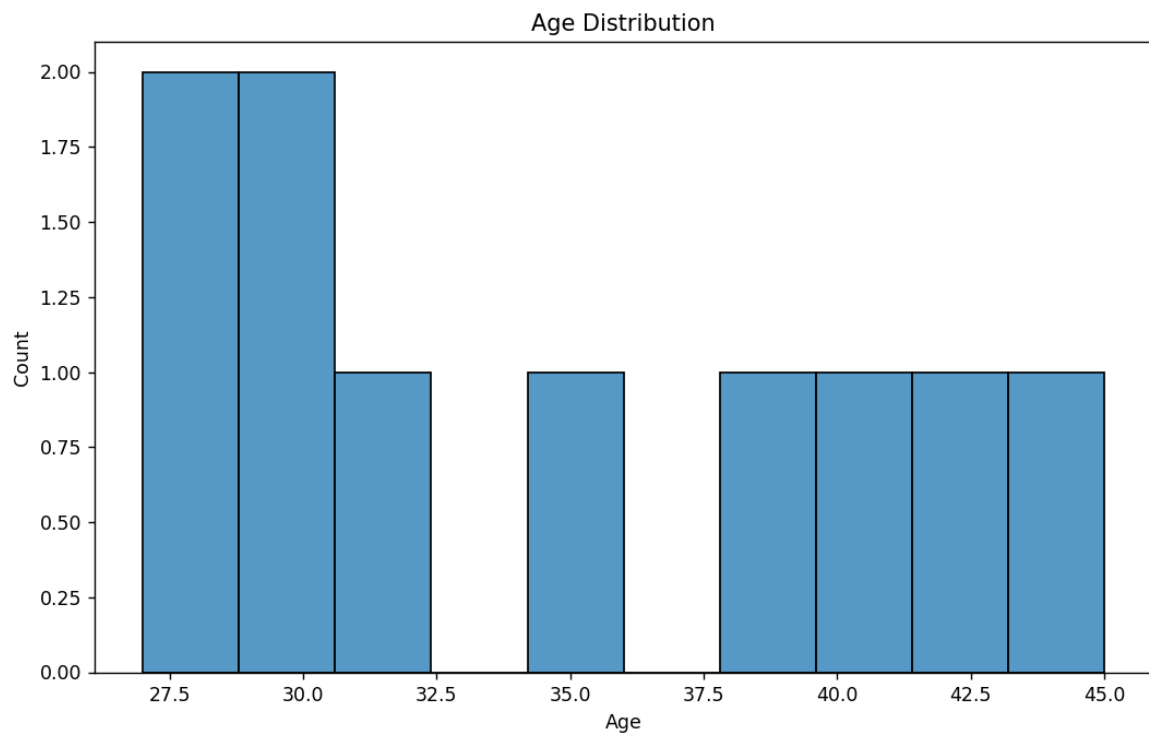
Performance Heatmap:



Tenure vs. Performance:



Age Distribution:



You

Departmental Distribution:

```
# Departmental Distribution: Bar Chart
department_counts = df['Department'].value_counts()
plt.bar(department_counts.index, department_counts.values)
plt.xlabel('Department')
plt.ylabel('Number of Employees')
```

```
plt.title('Departmental Distribution of Employees')
plt.show()
```

Salary Distribution:

```
# Salary Distribution: Box Plot
plt.boxplot(df['Salary'], vert=False)
plt.xlabel('Salary')
plt.title('Salary Distribution of Employees')
plt.show()
```

Performance Heatmap:

```
# Performance Heatmap
heatmap_data = df.pivot_table(index='Department', columns='Performance Score', aggfunc='size',
                                fill_value=0)
sns.heatmap(heatmap_data, cmap='YlGnBu', annot=True, fmt='d')
plt.xlabel('Performance Score')
plt.ylabel('Department')
plt.title('Performance Heatmap')
plt.show()
```

Tenure vs. Performance:

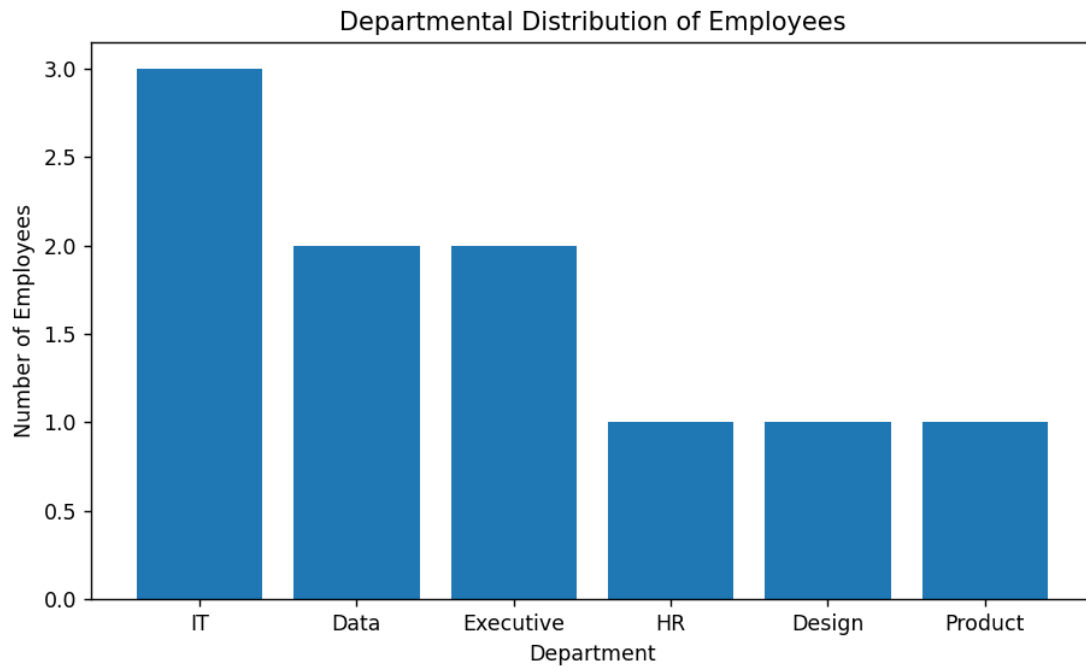
```
# Convert the Joining Date column to datetime
df['Joining Date'] = pd.to_datetime(df['Joining Date'])

# Calculate the tenure as the difference between the current date and the joining date
df['Tenure'] = (datetime.now() - df['Joining Date']).dt.days / 365
# Tenure vs. Performance: Scatter Plot
plt.scatter(df['Tenure'], df['Performance Score'])
plt.xlabel('Tenure (years)')
plt.ylabel('Performance Score')
plt.title('Tenure vs. Performance')
plt.show()
```

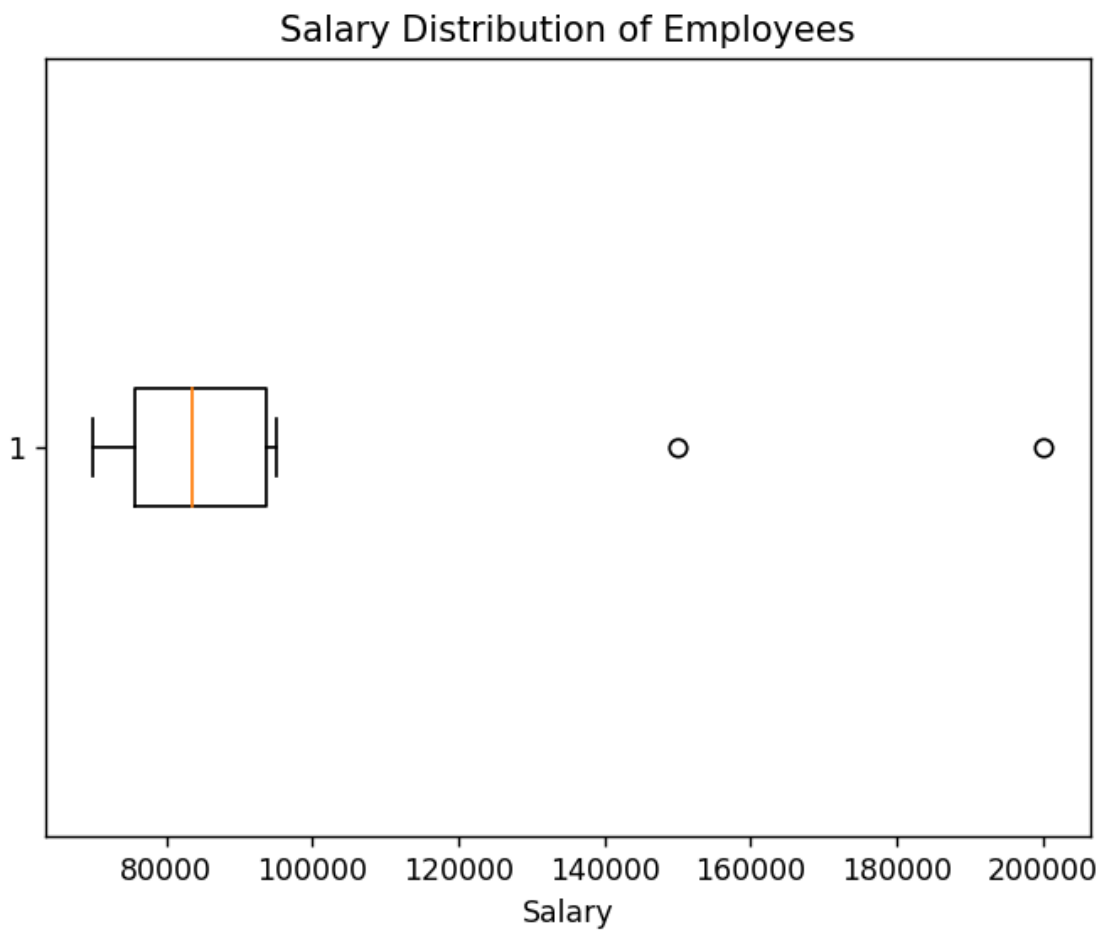
Age Distribution:

```
# Age Distribution: Histogram
plt.hist(df['Age'], bins=10, edgecolor='black')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.title('Age Distribution of Employees')
plt.show()
```

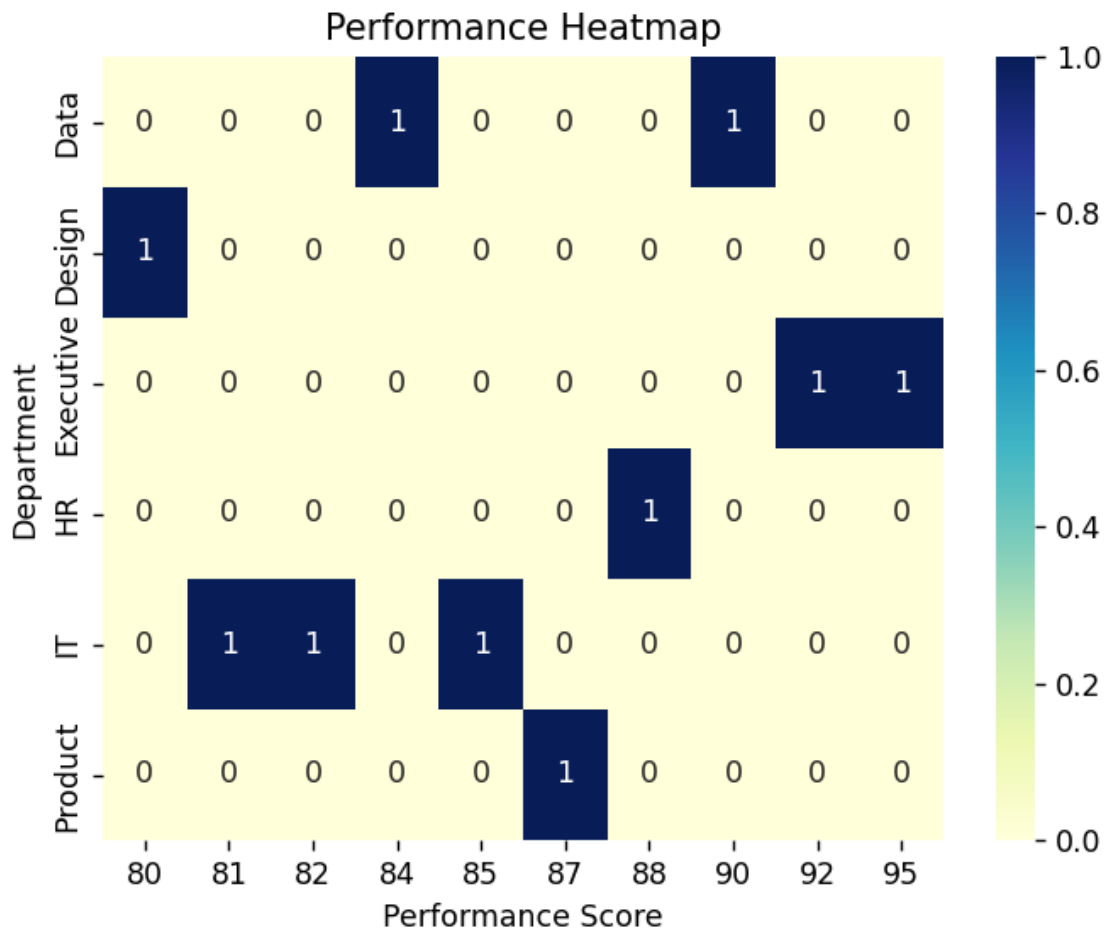
Departmental Distribution:



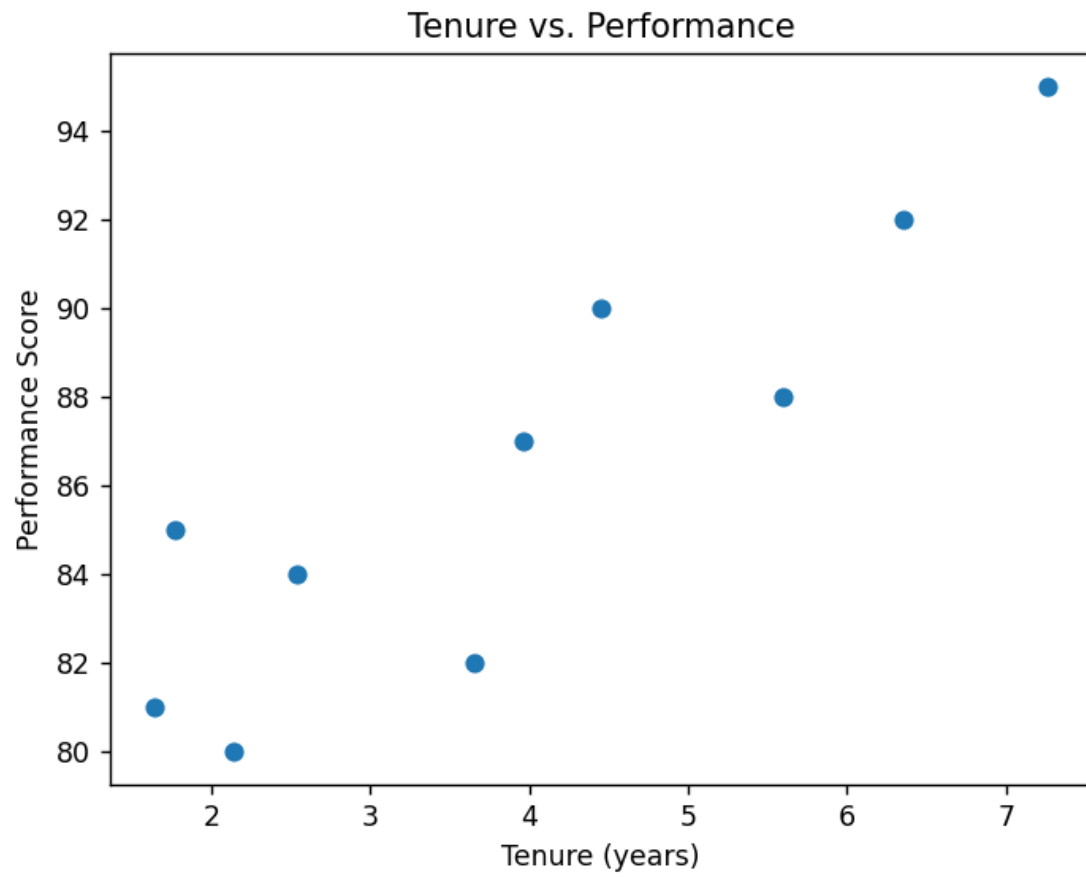
Salary Distribution:



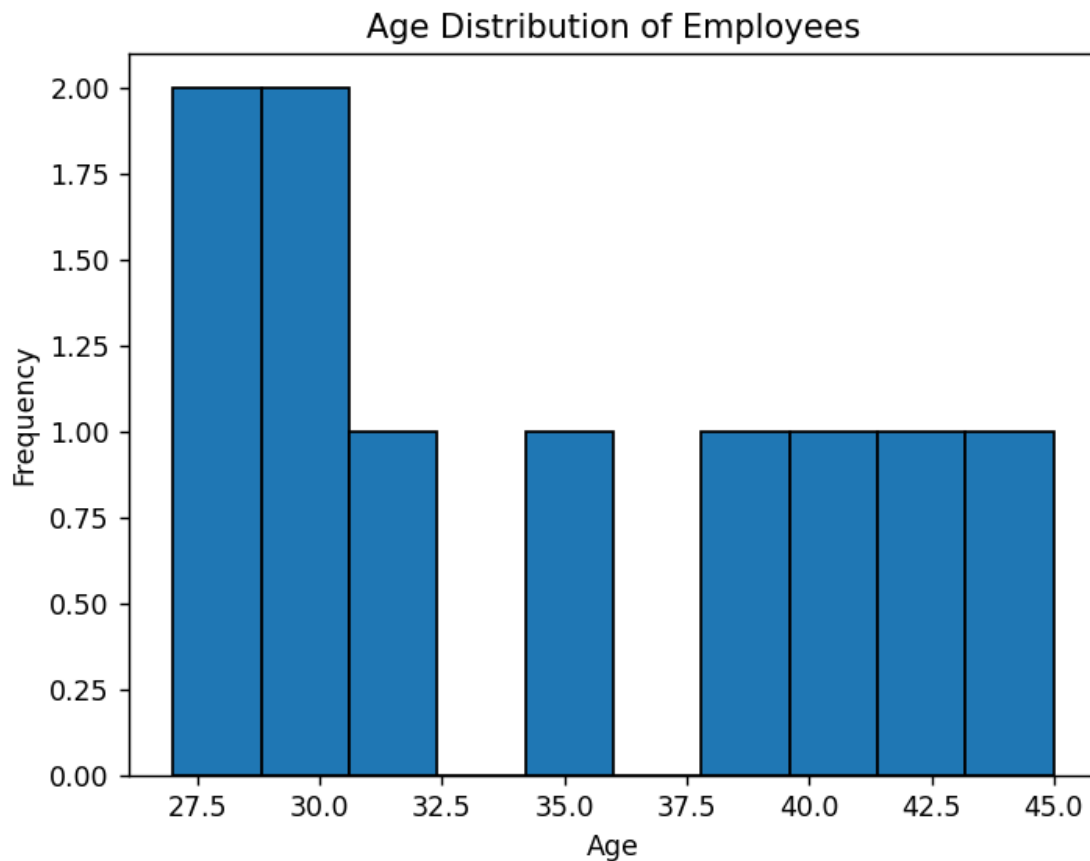
Performance Heatmap:



Tenure vs. Performance:



Age Distribution:



Google Bard

Departmental Distribution:

```
# **Departmental Distribution**
```

```
# Get the number of employees in each department
department_counts = df['Department'].value_counts()
```

```
# Create a bar chart of the department counts
plt.bar(department_counts.index, department_counts.values)
plt.xlabel('Department')
plt.ylabel('Number of Employees')
plt.title('Departmental Distribution')
plt.show()
```

Salary Distribution:

```
# **Salary Distribution**
```

```
# Create a box plot of the salary distribution
sns.boxplot(
    x = 'Department',
    y = 'Salary',
    showmeans=True,
    data=df
)
plt.title('Salary Distribution by Department')
plt.show()
```

Performance Heatmap:

```

# **Performance Heatmap**

# Create a pivot table of the performance scores by department
performance_pivot = df.pivot_table(
    index='Department',
    columns='Name',
    values='Performance Score'
)

# Create a heatmap of the performance pivot table
sns.heatmap(
    performance_pivot,
    cmap='RdYlGn',
    annot=True
)
plt.title('Performance Heatmap by Department')
plt.show()

Tenure vs. Performance:
# **Tenure vs. Performance**

# Calculate the tenure of each employee in years
df['Tenure'] = (pd.Timestamp.today() - df['Joining Date']).apply(pd.Timestamp)).dt.days / 365.25

# Create a scatter plot of tenure vs. performance score
plt.scatter(df['Tenure'], df['Performance Score'])
plt.xlabel('Tenure (years)')
plt.ylabel('Performance Score')
plt.title('Tenure vs. Performance')
plt.show()

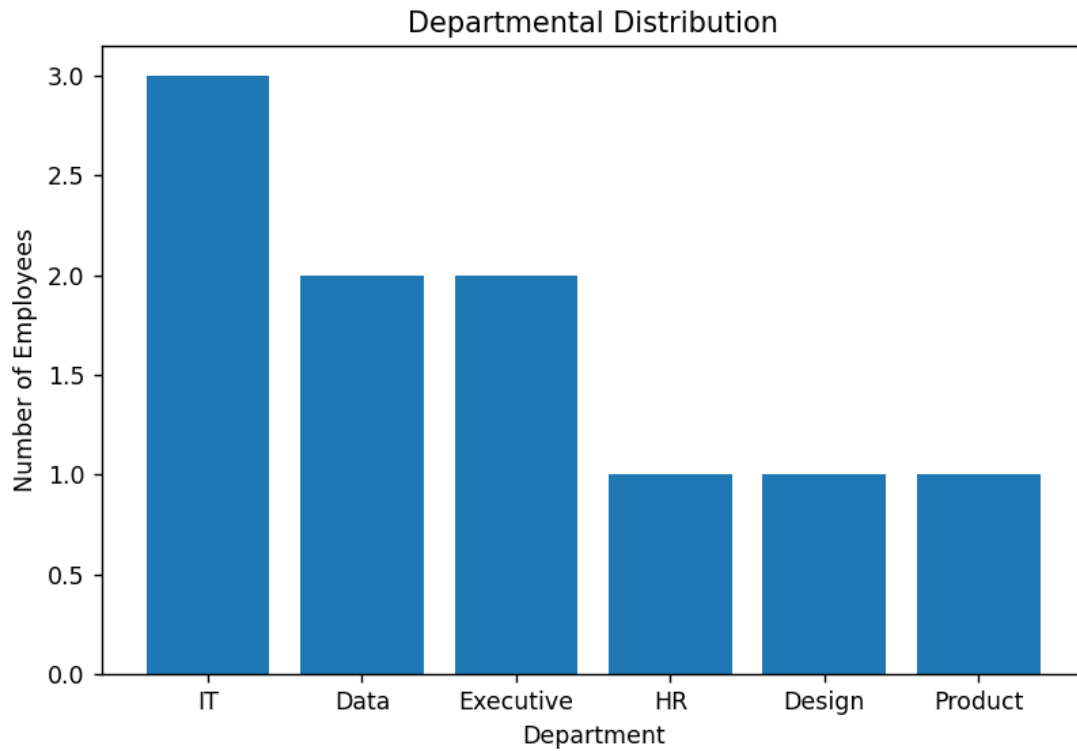
Age Distribution:
# **Age Distribution**

# Create a histogram of the employee ages
plt.hist(df['Age'], bins=10)
plt.xlabel('Age')
plt.ylabel('Number of Employees')
plt.title('Age Distribution')
plt.show()

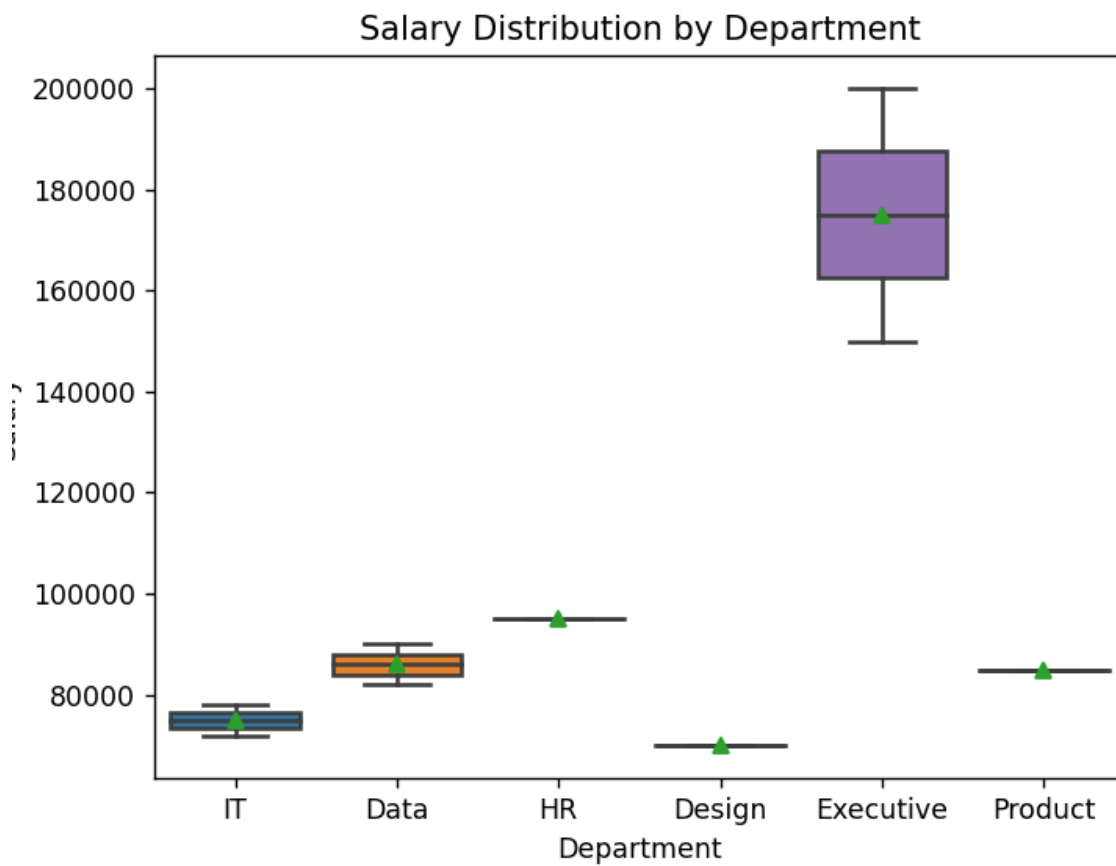
```

Output:

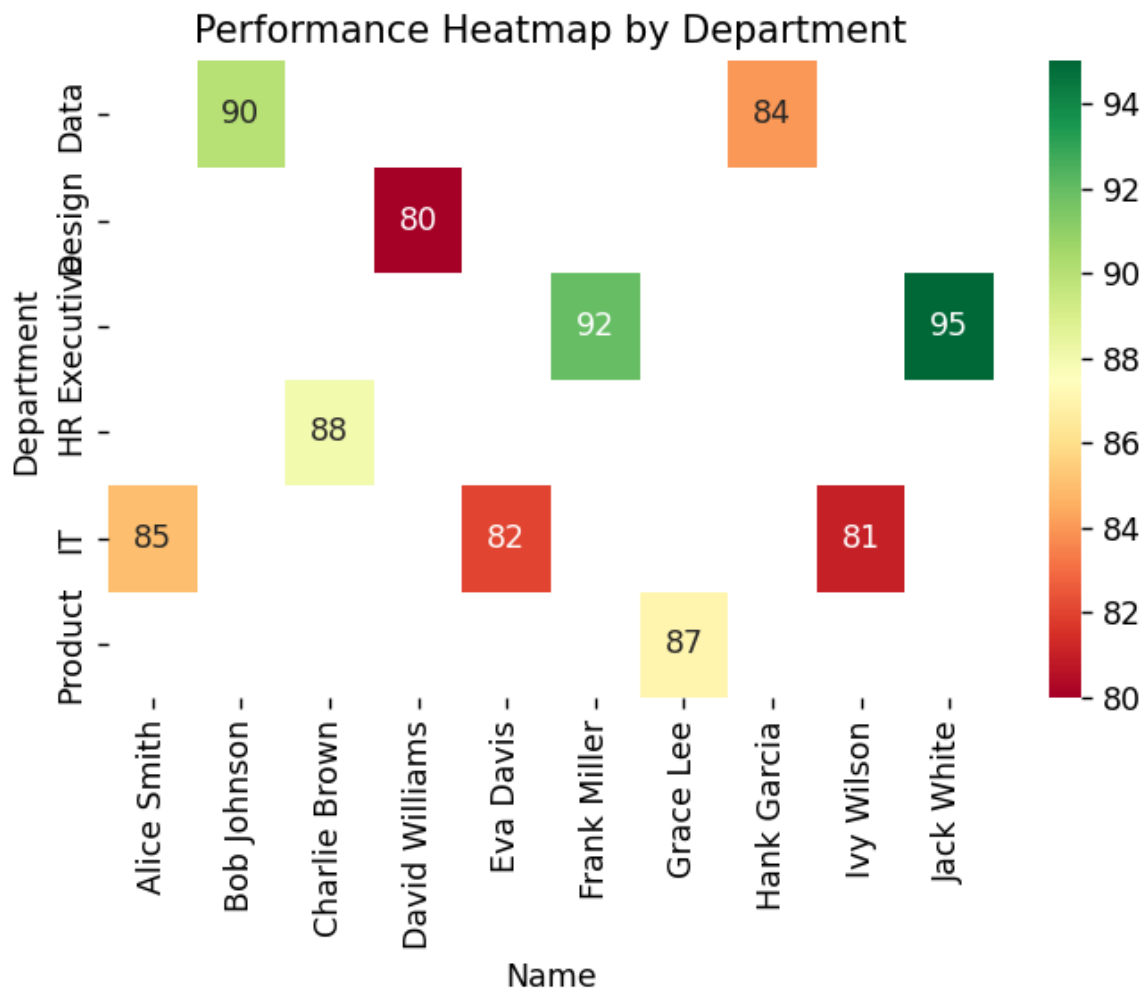
Departmental Distribution:



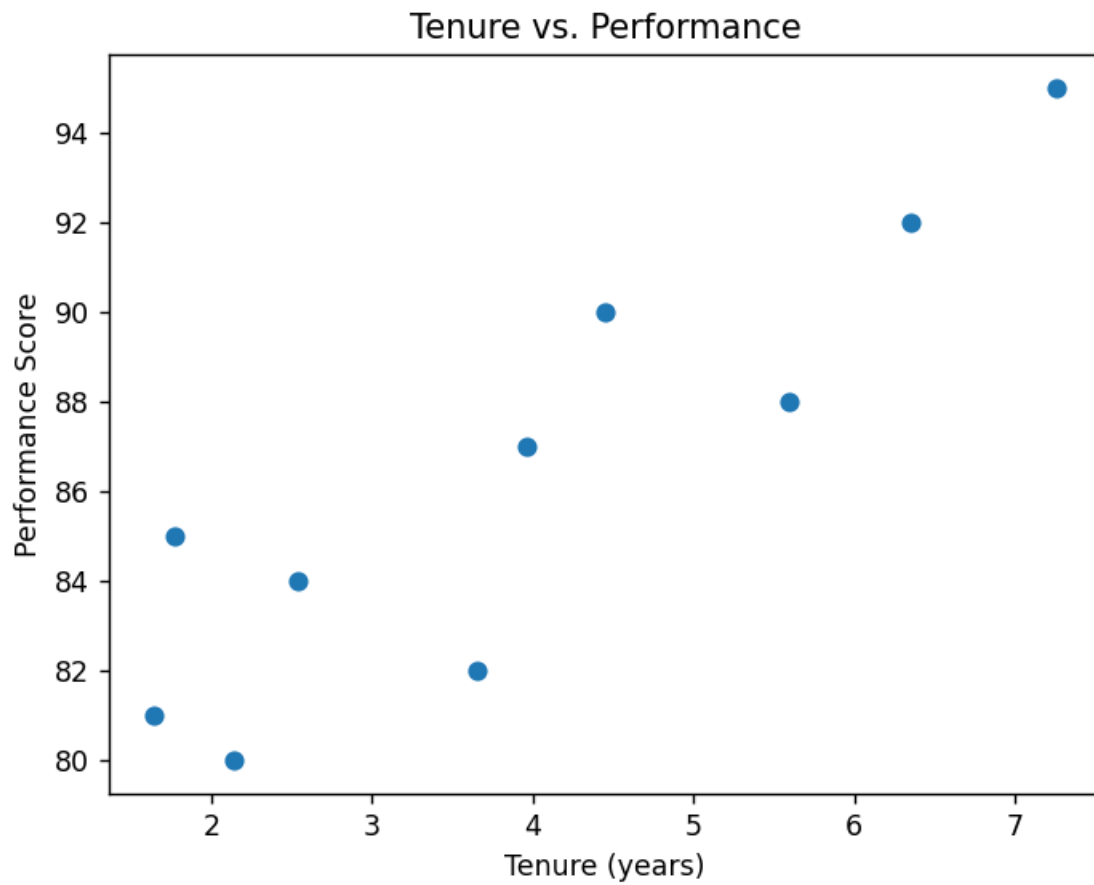
Salary Distribution:



Performance Heatmap:



Tenure vs. Performance:



Age Distribution:

