```python
import pandas as pd
import numpy as np

data = pd.read_excel("MODULE PROJECT EXCEL FILE.xlsx")

data["Height"] = np.random.randint(150, 181, size=len(data))

print(data.isnull().sum())
print(data.duplicated().sum())
```

[1] ✓ 6.8s

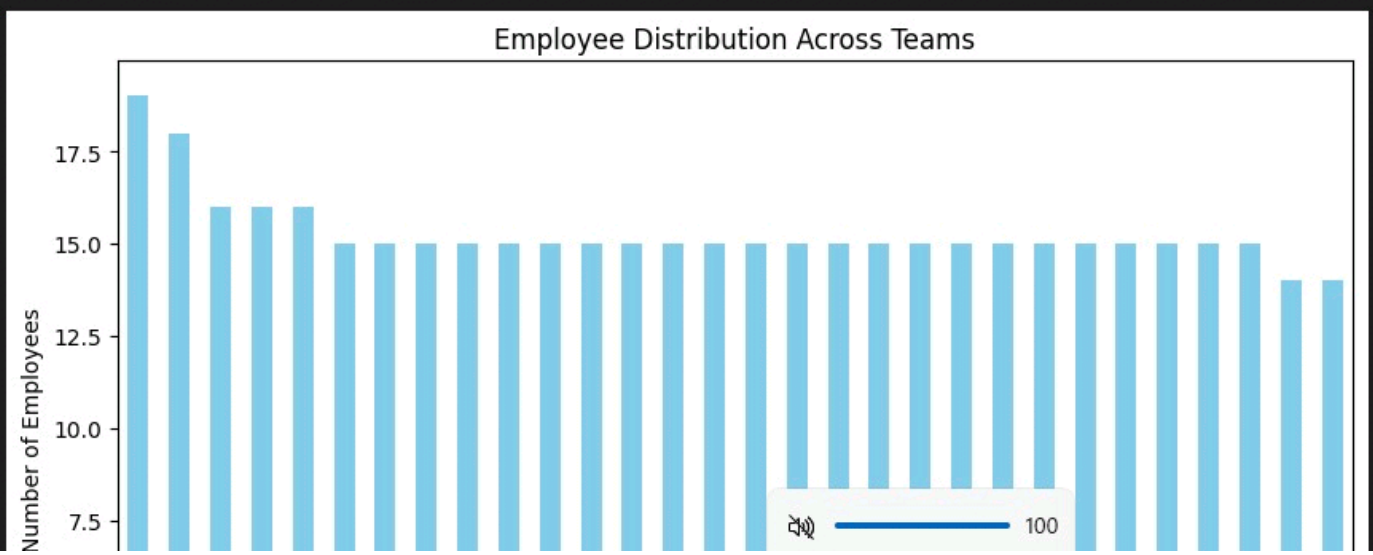
```
Name        0
Team        0
Number      0
Position    0
Age         0
Height      0
Weight      0
College     0
Salary      0
dtype: int64
0
```

```python
team_distribution = data["Team"].value_counts()
team_percentage = (team_distribution / len(data)) * 100

import matplotlib.pyplot as plt

plt.figure(figsize=(10, 6))
team_distribution.plot(kind="bar", color="skyblue")
plt.title("Employee Distribution Across Teams")
plt.xlabel("Team")
plt.ylabel("Number of Employees")
plt.show()
```

✓ 1.0s

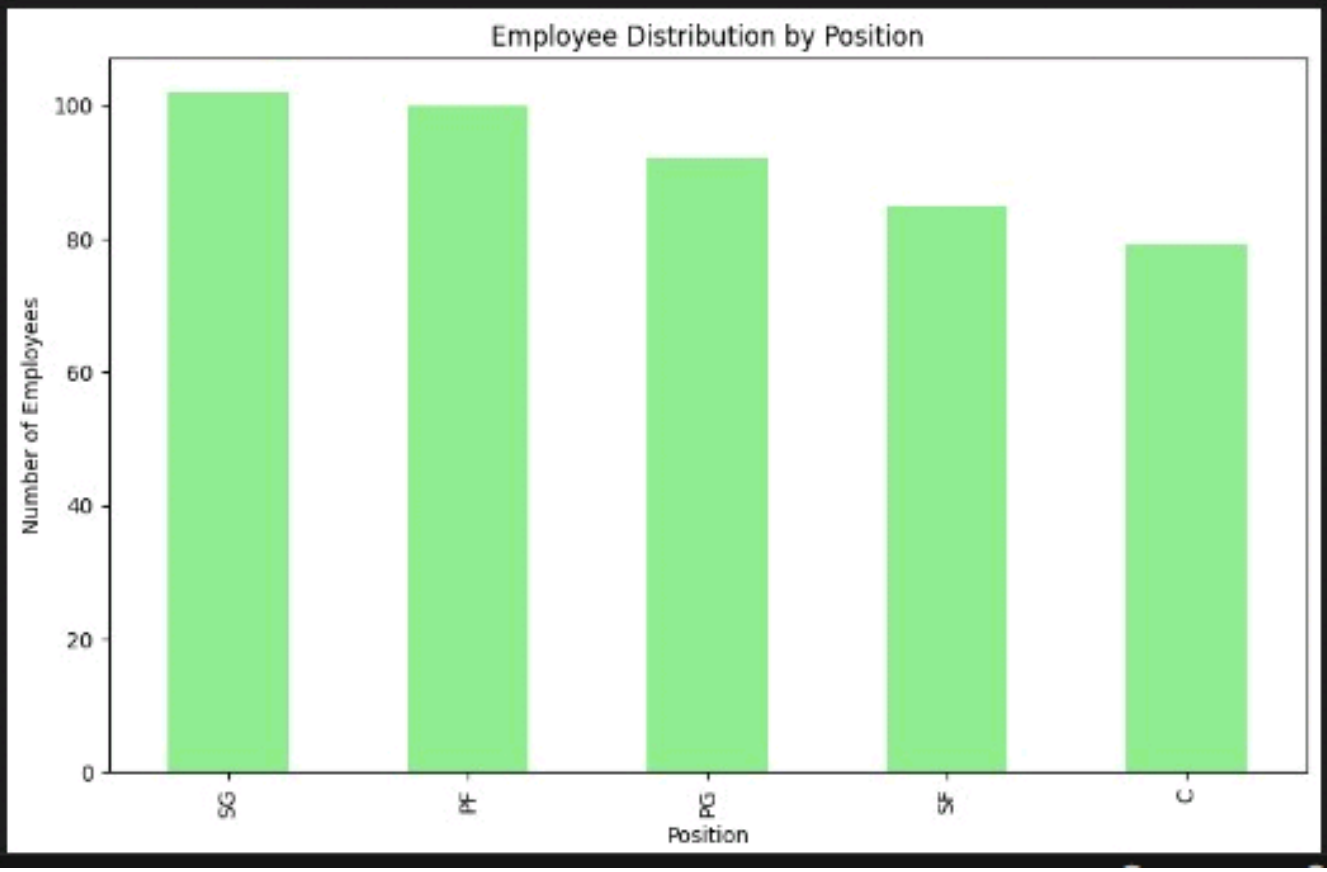


```
position_distribution = data["Position"].value_counts()
```

[5] ✓ 0.2s

```
plt.figure(figsize=(10, 6))
position_distribution.plot(kind="bar", color="lightgreen")
plt.title("Employee Distribution by Position")
plt.xlabel("Position")
plt.ylabel("Number of Employees")
plt.show()
```

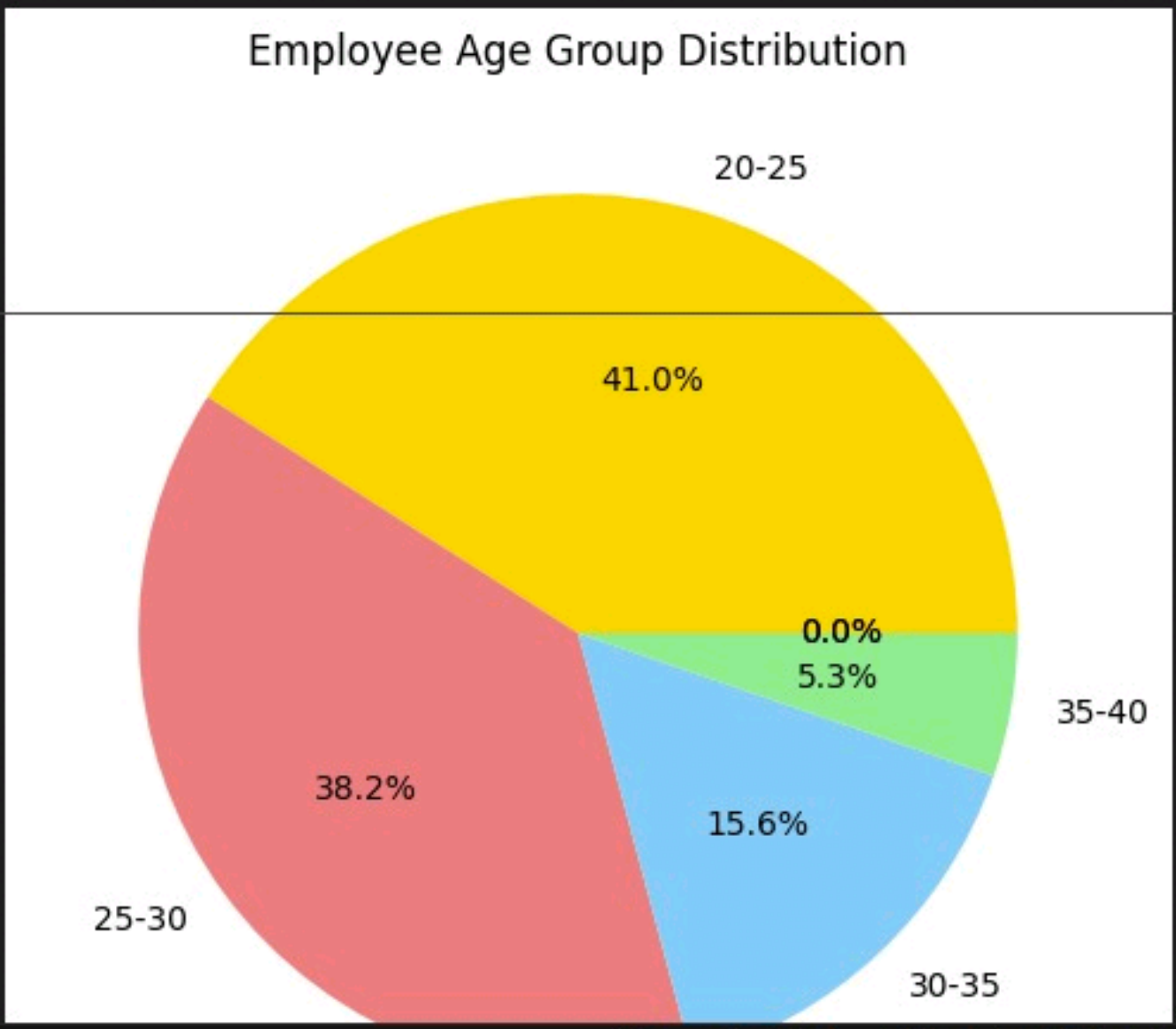
[6] ✓ 0.7s



100

```
plt.figure(figsize=(8, 6))
age_group_distribution.plot(kind="pie", autopct="%1.1f%%", colors=["
plt.title("Employee Age Group Distribution")
plt.ylabel("")
plt.show()
```

✓ 0.4s Python



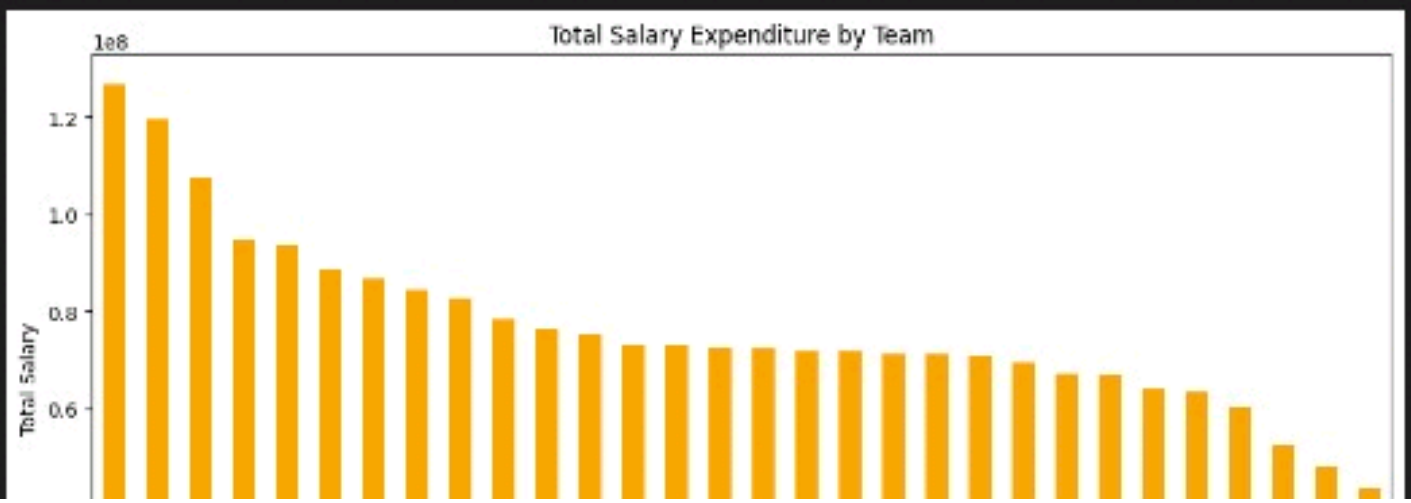
18 Spaces: 2 CRLF Cell 7 of 7 Go Live

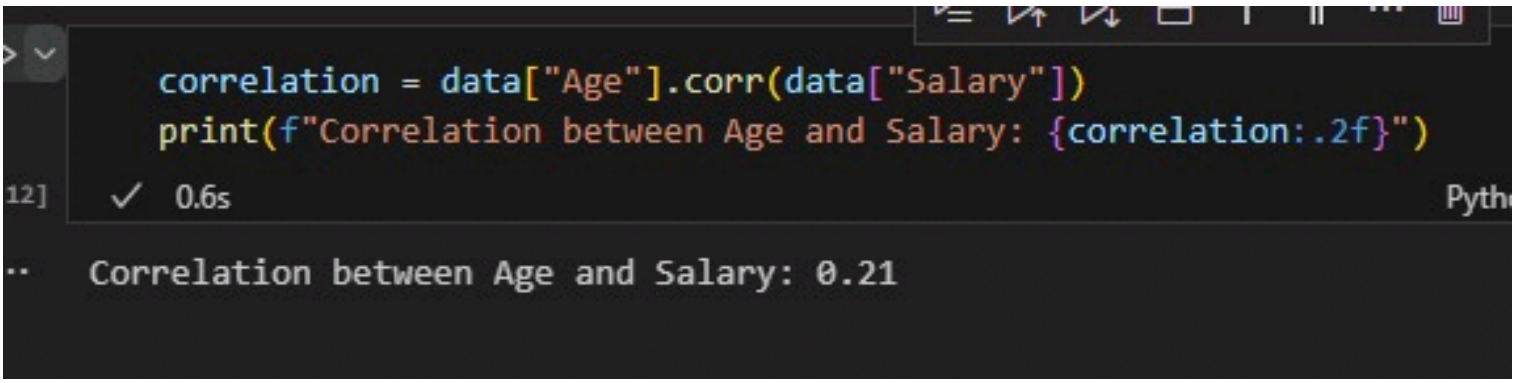
Click to add a breakpoint

```
(12, 6))
team_salary.plot(kind="bar", color="orange")
plt.title("Total Salary Expenditure by Team")
plt.xlabel("Team")
plt.ylabel("Total Salary")
plt.show()

plt.figure(figsize=(12, 6))
position_salary.plot(kind="bar", color="purple")
plt.title("Total Salary Expenditure by Position")
plt.xlabel("Position")
plt.ylabel("Total Salary")
plt.show()
```

✓ 4.0s Python




correlation = data["Age"].corr(data["Salary"])
print(f"Correlation between Age and Salary: {correlation:.2f}")
12]
0.6s
Correlation between Age and Salary: 0.21

```
plt.figure(figsize=(8, 6))
plt.scatter(data["Age"], data["Salary"], alpha=0.5, color="blue")
plt.title("Age vs Salary")
plt.xlabel("Age")
plt.ylabel("Salary")
plt.show()
```

✓ 1.0s