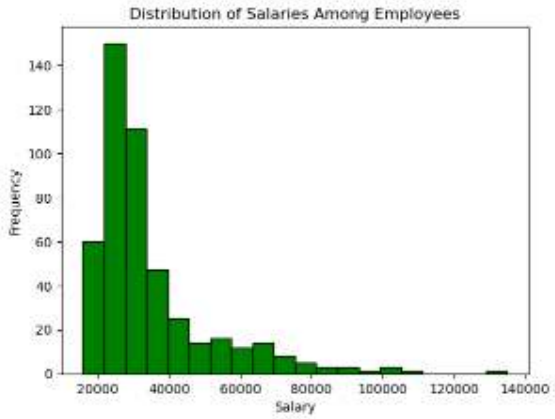
DELIVERABLES – ANSH SHARMA

1) Generate a histogram using any library to visualize the distribution of salaries among employees in the dataset.

Answer:

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt
%matplotlib inline
#%matplotlib inline sets the backend of matplotlib to the 'inline' backend
```

In [2]:
```
#Q1

# Load the dataset
df = pd.read_csv(r'C:\Users\ADMIN\Downloads\Employee data.csv')

# Extract the 'salary' column
salaries = df['salary']

# Create a histogram
plt.hist(salaries, bins=20, color='green', edgecolor='black')

# Add title and labels
plt.title('Distribution of Salaries Among Employees')
plt.xlabel('Salary')
plt.ylabel('Frequency')

# Display the plot
plt.show()
```



DELIVERABLES – ANSH SHARMA

2) Generate a bar plot to compare the average salary of male and female employees using any library.

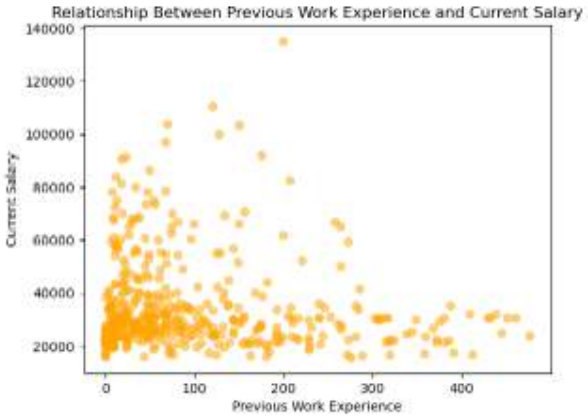
Answer:

```
In [3]: #Q2

# Group the data by gender and calculate the average salary
average_salary_by_gender = df.groupby('gender')['salary'].mean()

# Create a bar plot
average_salary_by_gender.plot(kind='bar', color=['blue', 'pink'])

# Add title and labels
plt.title('Average Salary Comparison by Gender')
plt.xlabel('Gender')
plt.ylabel('Average Salary')

# Show the plot
plt.show()
```

3. Create a scatter plot using any library to illustrate the relationship between previous work experience (prevexp) and the current salary of employees.

Answer:

```
In [4]: # Extract the relevant columns
prevexp = df['prevexp']
salary = df['salary']

# Create a scatter plot
plt.scatter(prevexp, salary, color='orange', alpha=0.5)

# Add title and labels
plt.title('Relationship Between Previous Work Experience and Current Salary')
plt.xlabel('Previous Work Experience')
plt.ylabel('Current Salary')

# Show the plot
plt.show()
```

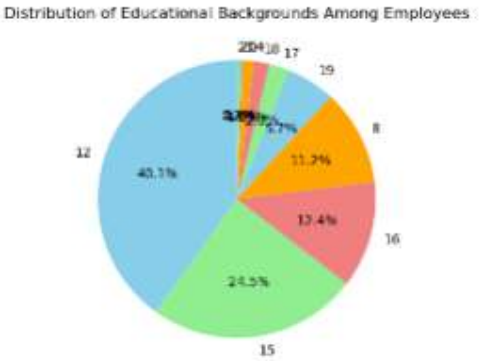


DELIVERABLES – ANSH SHARMA

4. Generate a pie chart to visualize the distribution of educational backgrounds among employees.

Answer:

```
In [5]: # Count the occurrences of each education level
education_counts = df['educ'].value_counts()

# Create a pie chart
plt.pie(education_counts, labels=education_counts.index, autopct='%1.1f%%', startangle=90, colors=['skyblue', 'lightgreen', 'lig

# Add title
plt.title('Distribution of Educational Backgrounds Among Employees')

# Show the plot
plt.show()
```



5. Use any AI library such as sweetviz or dtale to generate the summary of the data.

Answer: