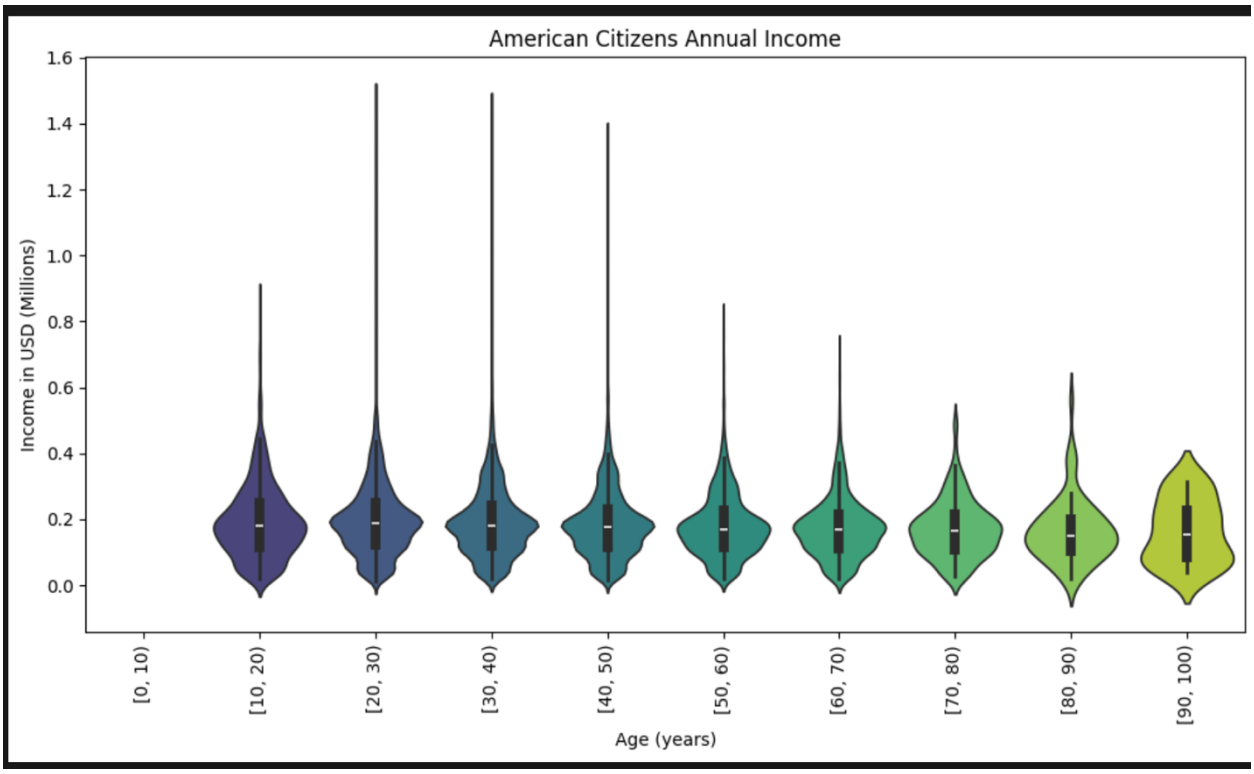
T4D2 Discussion

Bivariate Analysis on Income vs. Age of American Citizens

```python
# Import library
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load data from CSV file
csv_file = 'income.csv'
data = pd.read_csv(csv_file)

# Define age intervals in 10 year intervals
age_intervals = pd.cut(data['age'], bins=range(0, 101, 10), right=False)

# Create a dataframe
df = pd.DataFrame({
    'Age Intervals': age_intervals,
    'Income': data['fnlwgt'] / 1_000_000
})

# Create a violin plot
plt.figure(figsize=(10, 6))

# Plot x and y axis with corresponding labels and tiltles
sns.violinplot(x='Age Intervals', y='Income', data=df, palette='viridis')
plt.xlabel('Age (years)')
plt.ylabel('Income in USD (Millions)')
plt.title('American Citizens Annual Income')

# Show the plot
plt.xticks(rotation=90)  # Rotate x-axis labels for better readability
plt.tight_layout()
plt.show()
```

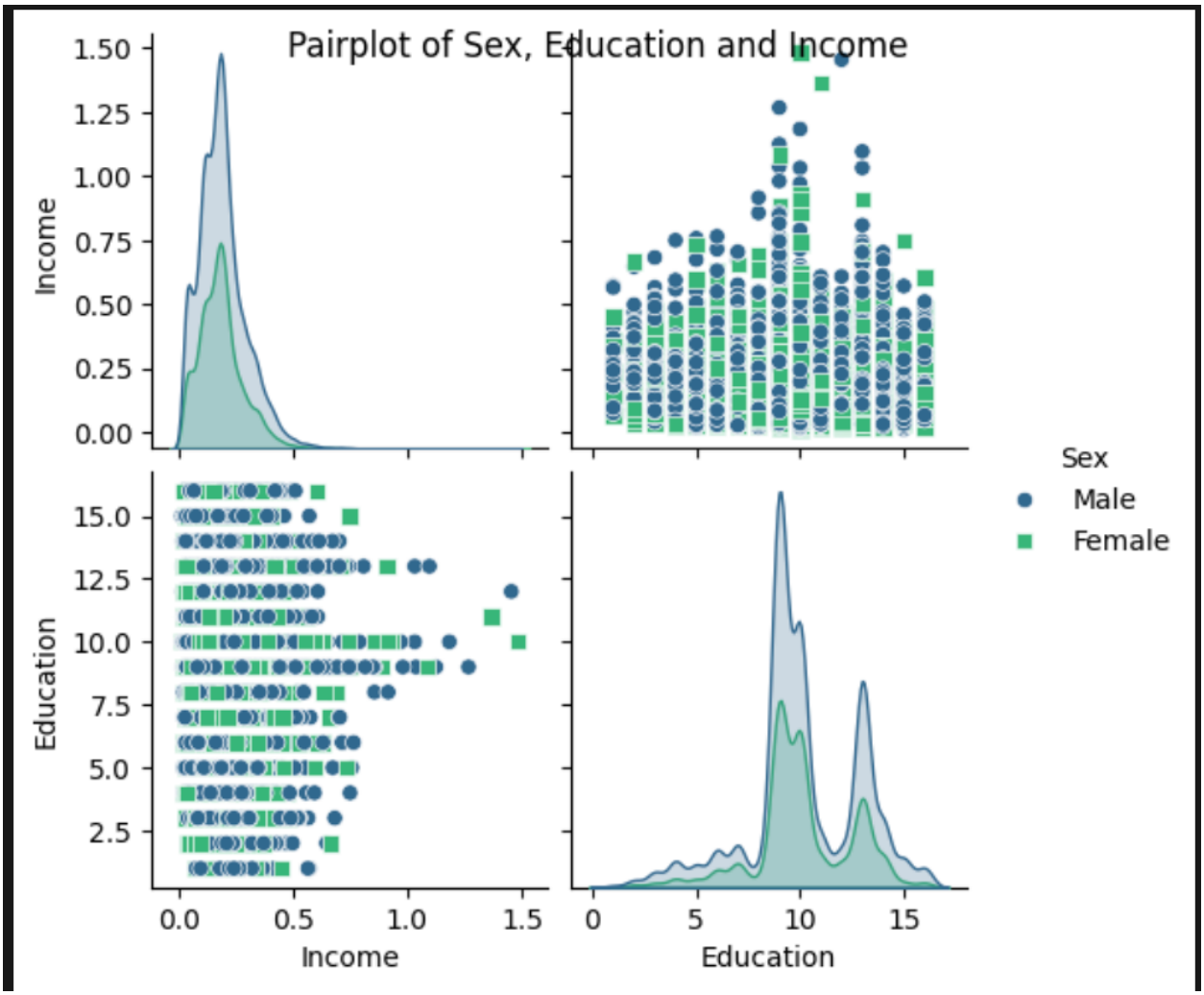
✓ 5.1s


American Citizens Annual Income
Income in USD (Millions)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
[0, 10)
[10, 20)
[20, 30)
[30, 40)
[40, 50)
[50, 60)
[60, 70)
[70, 80)
[80, 90)
[90, 100)
Age (years)

Multivariate Analysis between Sex, Education and Income

```python
# Import Libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load data from CSV file
csv_file = 'income.csv'
data = pd.read_csv(csv_file)

# Define age intervals in 10-year intervals
age_intervals = pd.cut(data['age'], bins=range(0, 101, 10), right=False)

# Create a dataframe
df = pd.DataFrame({
    'Age Intervals': age_intervals,
    'Income': data['fnlwgt'] / 1_000_000,
    'Education': data['education.num'],
    'Race': data['race'],
    'Sex': data['sex']
})

# Create a plot that would show relationship between the variables
sns.pairplot(df, hue='Sex', palette='viridis', hue_order=['Male', 'Female'], markers=['o', 's'], diag_kind='kde')
plt.suptitle('Pairplot of Sex, Education and Income')
plt.show()
```

✓ 10.6s


Pairplot of Sex, Education and Income
Income
Education
Sex
Male
Female
1.50
1.25
1.00
0.75
0.50
0.25
0.00
15.0
12.5
10.0
7.5
5.0
2.5
0.0
0.5
1.0
1.5
0
5
10
15