```
from google.colab import files

# Upload the file
uploaded = files.upload()
```

Choose Files Day_13_Ph…a_data.csv

- **Day_13_Pharma_data.csv**(text/csv) - 679 bytes, last modified: 1/24/2025 - 100% done

```
Saving Day_13_Pharma_data.csv to Day_13_Pharma_data.csv
```

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns


fp='Day_13_Pharma_data.csv'
Pharma_data=pd.read_csv(fp)


df = pd.DataFrame(Pharma_data)


print(df.isnull().sum())

# Check for duplicates
print(df.duplicated().sum())
```

```
Product_ID         0
Product_Name       0
Region             0
Marketing_Spend    0
Sales              0
Effectiveness      0
Side_Effects       0
Age_Group          0
Trial_Period       0
dtype: int64
0
```

NameError

Please explain the error:

```
NameError: name 'data' is not defined
```

Hypothesis

The error message "NameError: name 'data' is not defined" indicates that the variable `data` is being used in the line `df = pd.DataFrame(data)` without being defined or assigned a value previously. In other words, Python doesn't know what 'data' refers to.

This is happening likely because you are attempting to create a pandas DataFrame from a variable called `data`, but this variable has not been initialized or assigned any value. You probably intended to use the DataFrame you read in from the csv file 'Day_13_Pharma_data.csv' which you assigned to `Pharma_data`.

Suggested Changes

```
df = pd.DataFrame(Pharma_data) # Use the Pharma
```

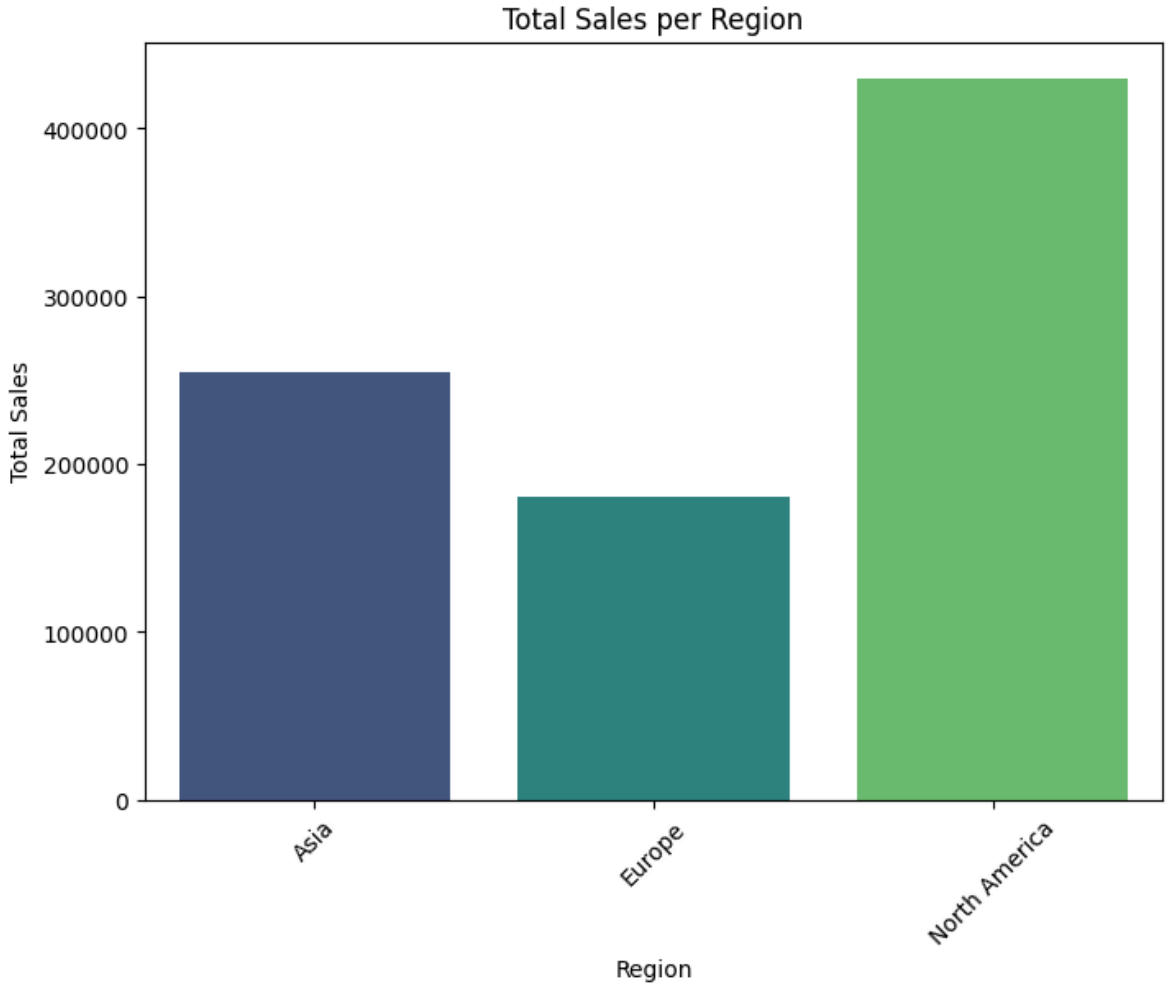
Use code with caution

▶ Sources

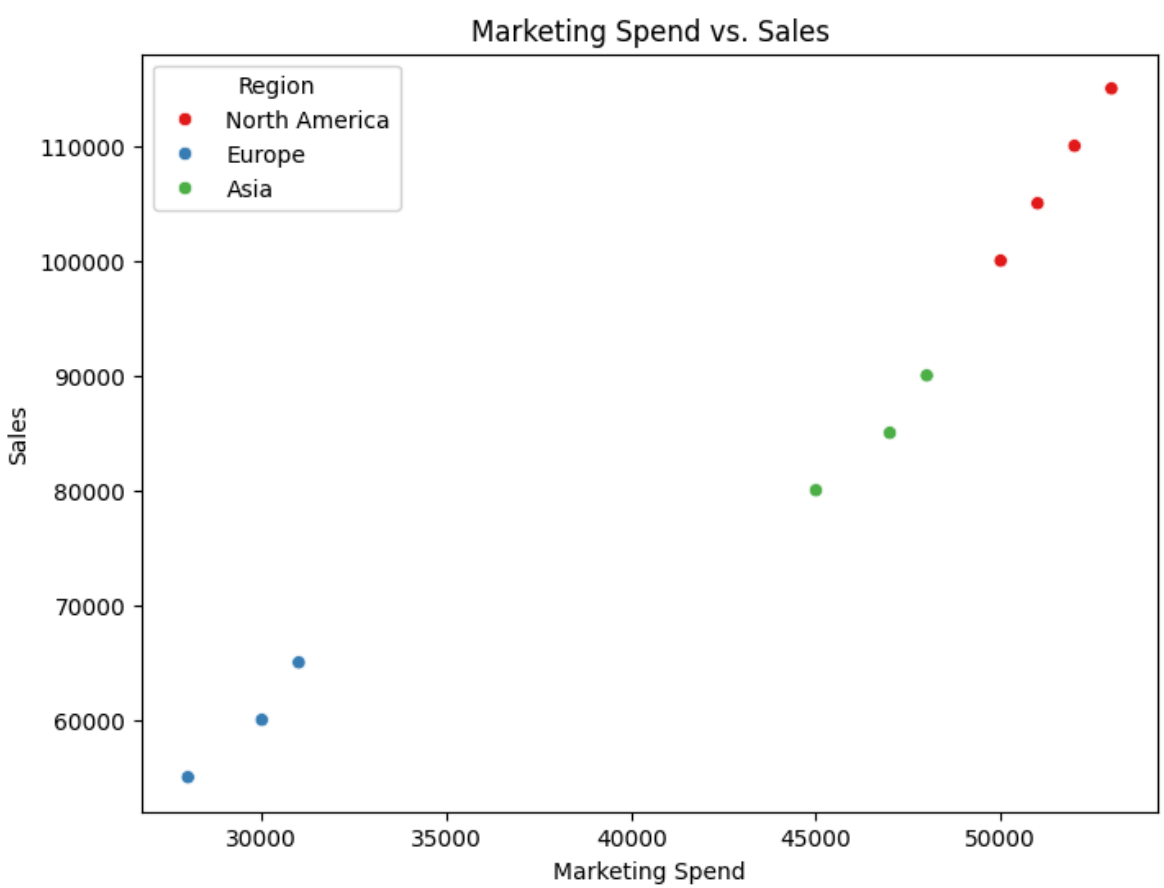
Rate this answer

```
region_sales = df.groupby('Region')['Sales'].sum().reset_index()

# Plot
plt.figure(figsize=(8, 6))
sns.barplot(x='Region', y='Sales', data=region_sales, palette='viridis')
plt.title('Total Sales per Region')
plt.xlabel('Region')
plt.ylabel('Total Sales')
plt.xticks(rotation=45)
plt.show()
```

```
<ipython-input-8-a0cbaa2ddf46>:5: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed i

  sns.barplot(x='Region', y='Sales', data=region_sales, palette='viridis')
```



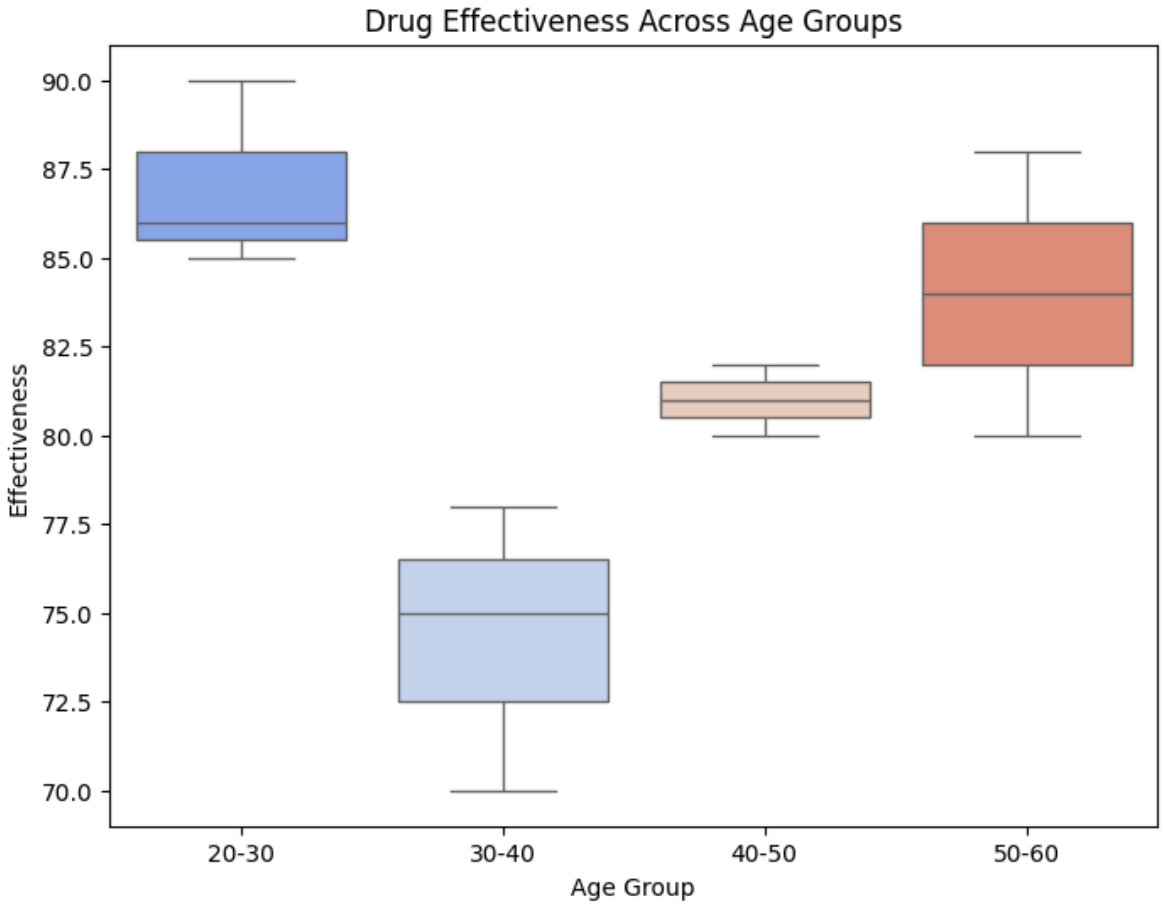
```
plt.figure(figsize=(8, 6))
sns.scatterplot(x='Marketing_Spend', y='Sales', data=df, hue='Region', palette='S
plt.title('Marketing Spend vs. Sales')
plt.xlabel('Marketing Spend')
plt.ylabel('Sales')
plt.show()
```



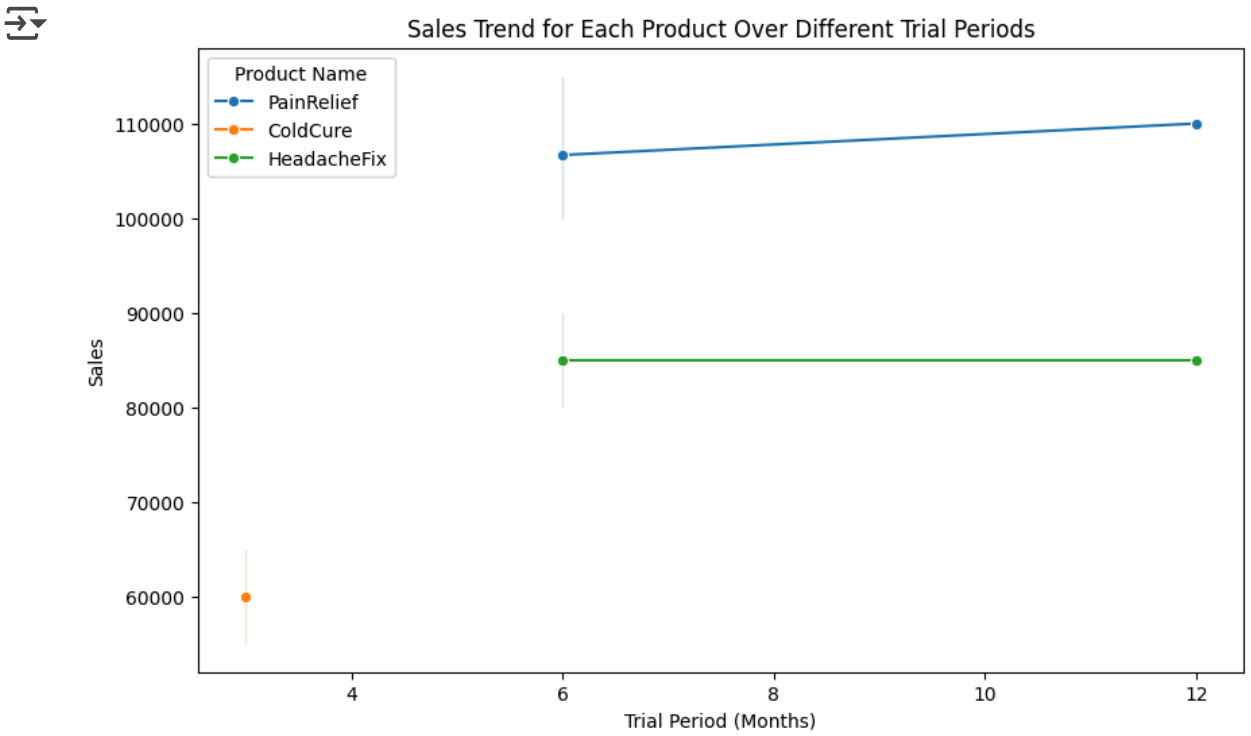

```
plt.figure(figsize=(8, 6))
sns.boxplot(x='Age_Group', y='Effectiveness', data=df, palette='coolwarm')
plt.title('Drug Effectiveness Across Age Groups')
plt.xlabel('Age Group')
```

```
plt.ylabel('Effectiveness')
plt.show()
```

```
<ipython-input-10-e32995401a82>:2: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed i

  sns.boxplot(x='Age_Group', y='Effectiveness', data=df, palette='coolwarm')
```

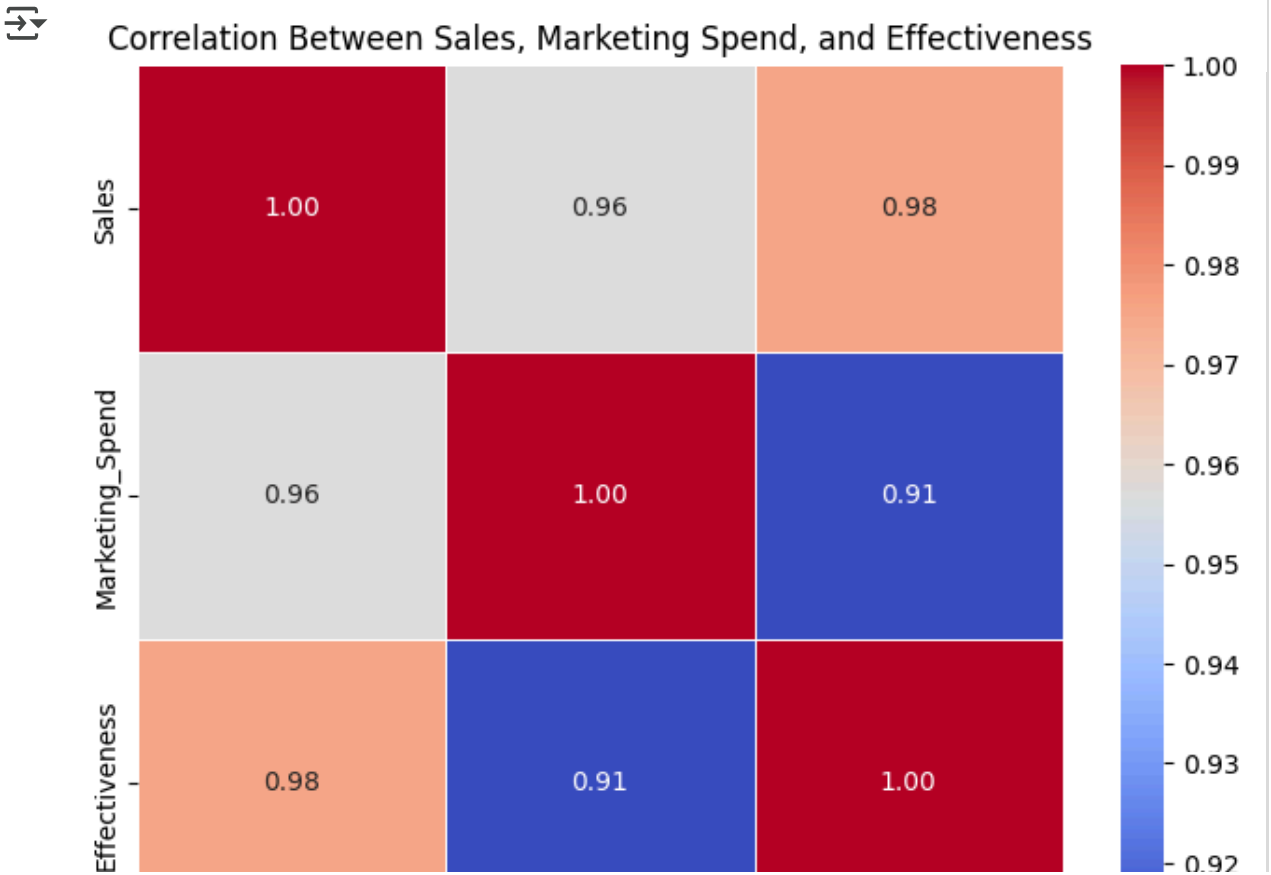


```
df['Trial_Period_Num'] = df['Trial_Period'].apply(lambda x: int(x.split()[0]))

plt.figure(figsize=(10, 6))
sns.lineplot(x='Trial_Period_Num', y='Sales', hue='Product_Name', data=df, marker
plt.title('Sales Trend for Each Product Over Different Trial Periods')
plt.xlabel('Trial Period (Months)')
plt.ylabel('Sales')
plt.legend(title='Product Name')
plt.show()
```




```
corr_data = df[['Sales', 'Marketing_Spend', 'Effectiveness']]

# Compute correlation matrix
corr_matrix = corr_data.corr()
```

```
# Plot
plt.figure(figsize=(8, 6))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt='.2f', linewidths=0.5)
plt.title('Correlation Between Sales, Marketing Spend, and Effectiveness')
plt.show()
```



Enter a prompt here

0 / 2000

Responses may display inaccurate or offensive information that doesn't represent Google's views. Learn more