FACEBOOK DATASET

In [ ]:
```
1
```

In [1]:
```
1 import pandas as pd
2 import matplotlib.pyplot as plt
```

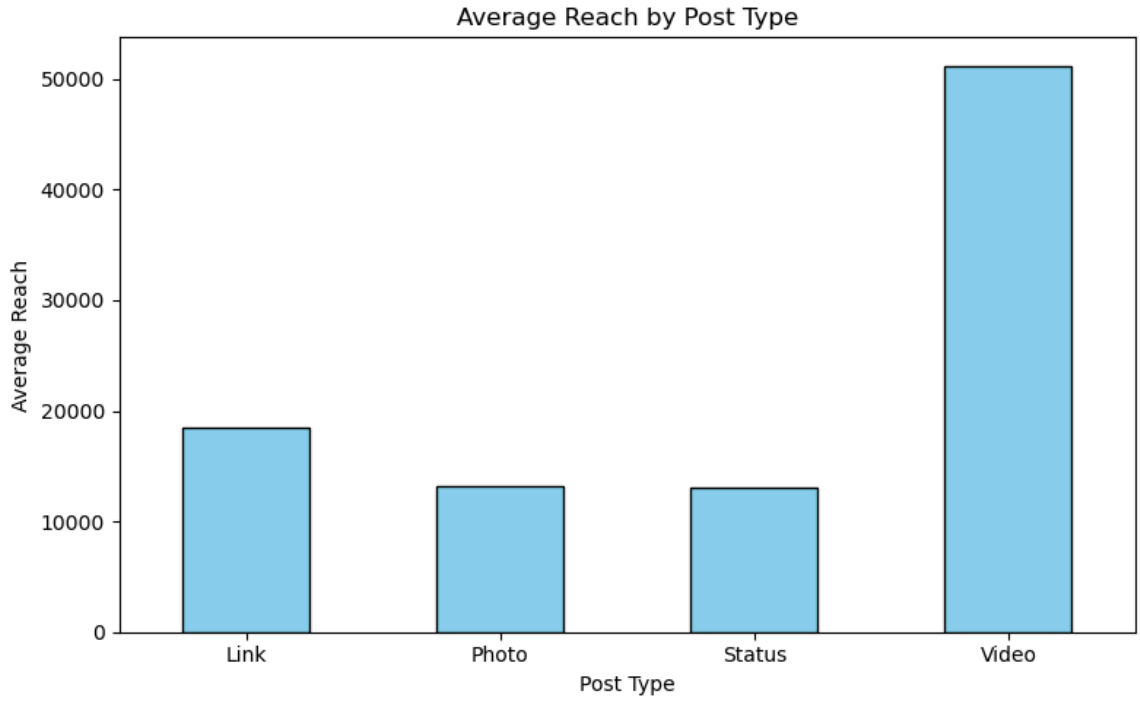
In [2]:
```
1 df=pd.read_csv("dataset_facebook.csv", sep=';')
2 df.head()
```

Out[2]:

	Page total likes	Type	Category	Post Month	Post Weekday	Post Hour	Paid	Lifetime Post Total Reach	Lifetime Post Total Impressions	Lifetime Engaged Users	C
0	139441	Photo	2	12	4	3	0.0	2752	5091	178	
1	139441	Status	2	12	3	10	0.0	10460	19057	1457	
2	139441	Photo	3	12	3	3	0.0	2413	4373	177	
3	139441	Photo	2	12	2	10	1.0	50128	87991	2211	
4	139441	Photo	2	12	2	3	0.0	7244	13594	671	

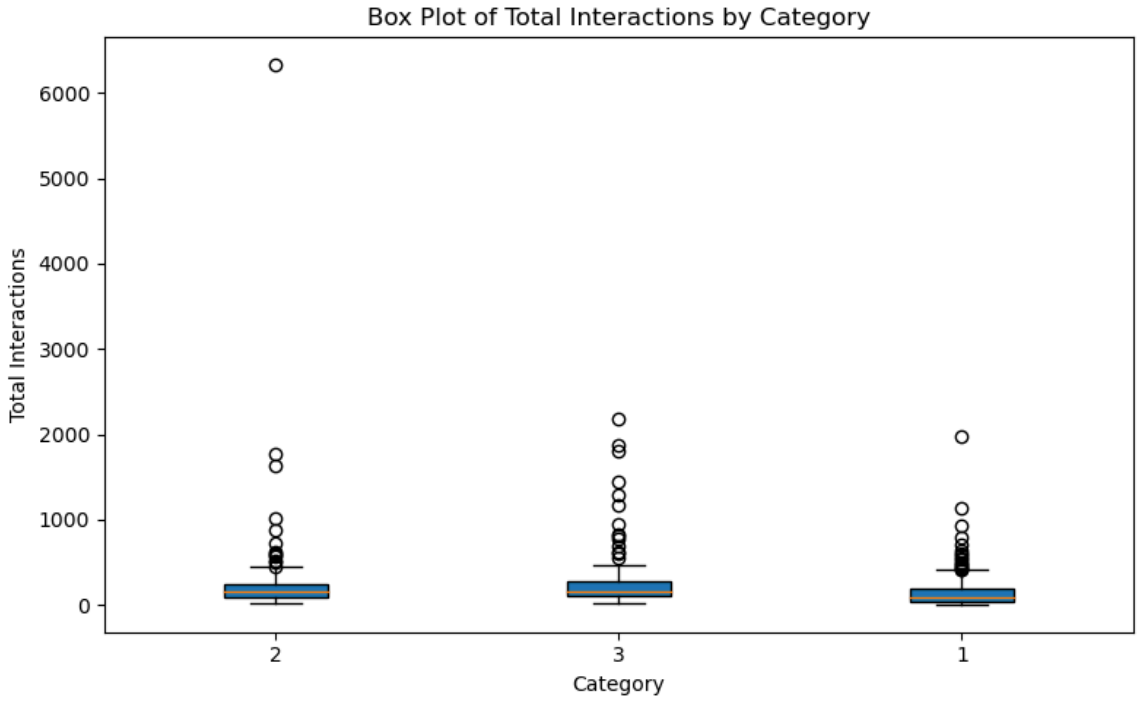
BAR PLOT

In [3]:
```
avg_reach = df.groupby('Type')['Lifetime Post Total Reach'].mean()
plt.figure(figsize=(8, 5))
avg_reach.plot(kind='bar', color='skyblue', edgecolor='black')
plt.title('Average Reach by Post Type')
plt.xlabel('Post Type')
plt.ylabel('Average Reach')
plt.xticks(rotation=0)
plt.tight_layout()
plt.show()
```



In [4]:

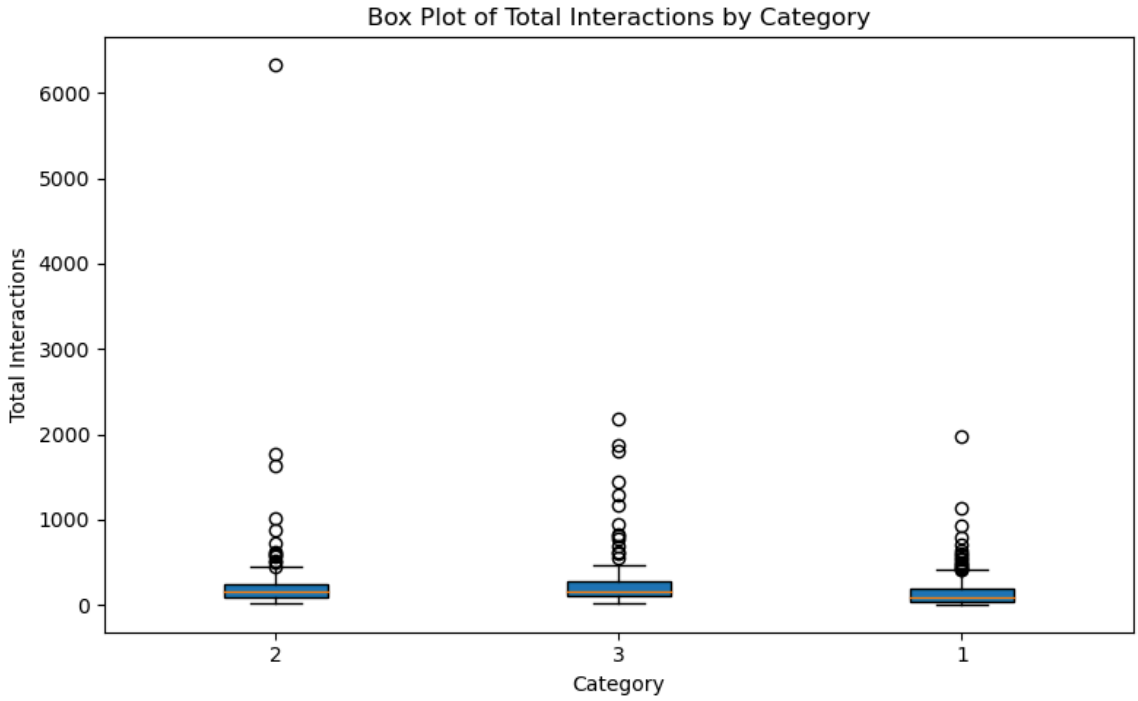
```
df = df[['Category', 'Total Interactions']].dropna()


df['Category'] = df['Category'].astype(str)


categories = df['Category'].unique()
data_to_plot = [df[df['Category'] == cat]['Total Interactions'] for cat


plt.figure(figsize=(8, 5))
plt.boxplot(data_to_plot, labels=categories, patch_artist=True)
plt.title('Box Plot of Total Interactions by Category')
plt.xlabel('Category')
plt.ylabel('Total Interactions')
plt.tight_layout()
plt.show()

```



BOX PLOT

In [5]:

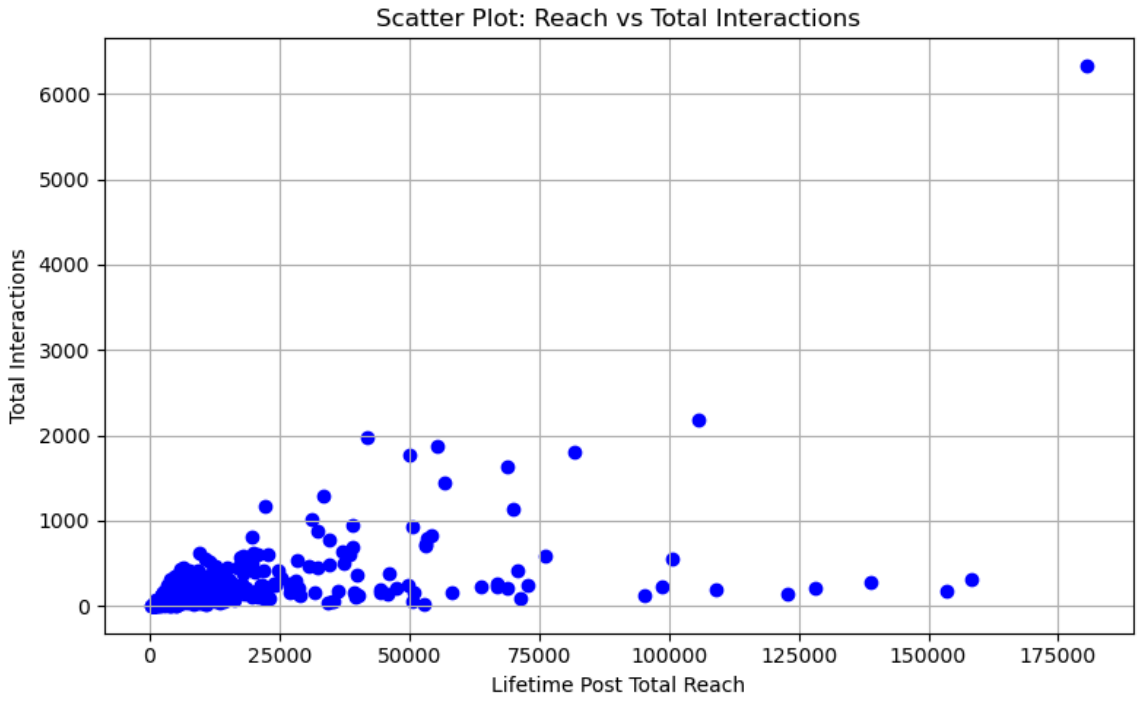
```
df = df[['Category', 'Total Interactions']].dropna()


df['Category'] = df['Category'].astype(str)


categories = df['Category'].unique()
data_to_plot = [df[df['Category'] == cat]['Total Interactions'] for cat


plt.figure(figsize=(8, 5))
plt.boxplot(data_to_plot, labels=categories, patch_artist=True)
plt.title('Box Plot of Total Interactions by Category')
plt.xlabel('Category')
plt.ylabel('Total Interactions')
plt.tight_layout()
plt.show()

```



SCATTER PLOT

In [11]:

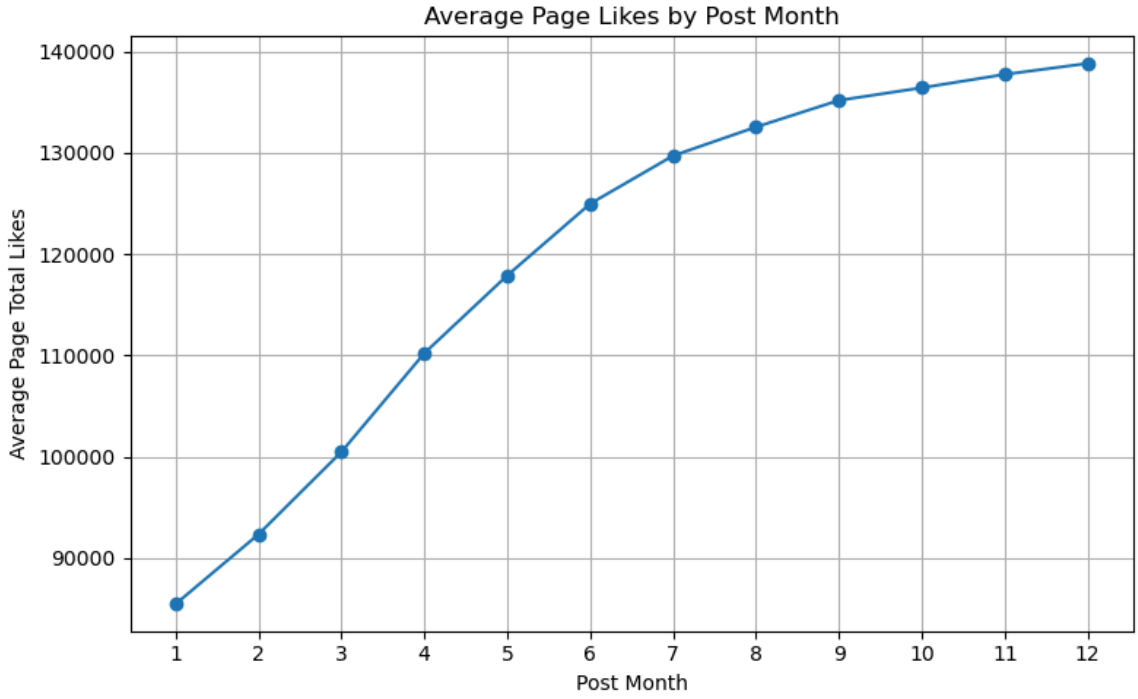
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("dataset_Facebook.csv", sep=';')
s
df = df[['Lifetime Post Total Reach', 'Total Interactions']].dropna()

df['Lifetime Post Total Reach'] = pd.to_numeric(df['Lifetime Post Total
df['Total Interactions'] = pd.to_numeric(df['Total Interactions'], erro

df = df.dropna()

plt.figure(figsize=(8, 5))
plt.scatter(df['Lifetime Post Total Reach'], df['Total Interactions'],
plt.title('Scatter Plot: Reach vs Total Interactions')
plt.xlabel('Lifetime Post Total Reach')
plt.ylabel('Total Interactions')
plt.grid(True)
plt.tight_layout()
plt.show()
```



LINE PLOT

In [14]:

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

df = pd.read_csv("dataset_Facebook.csv", sep=';')

df = df[['Post Month', 'Page total likes']].dropna()

df['Post Month'] = pd.to_numeric(df['Post Month'], errors='coerce')
df['Page total likes'] = pd.to_numeric(df['Page total likes'], errors='
df = df.dropna()

monthly_avg_likes = df.groupby('Post Month')['Page total likes'].mean()
# Plotting
plt.figure(figsize=(8, 5))
plt.plot(monthly_avg_likes.index, monthly_avg_likes.values, marker='o')
plt.title('Average Page Likes by Post Month')
plt.xlabel('Post Month')
plt.ylabel('Average Page Total Likes')
plt.grid(True)
plt.xticks(monthly_avg_likes.index)
plt.tight_layout()
plt.show()
```



HEART DATASET

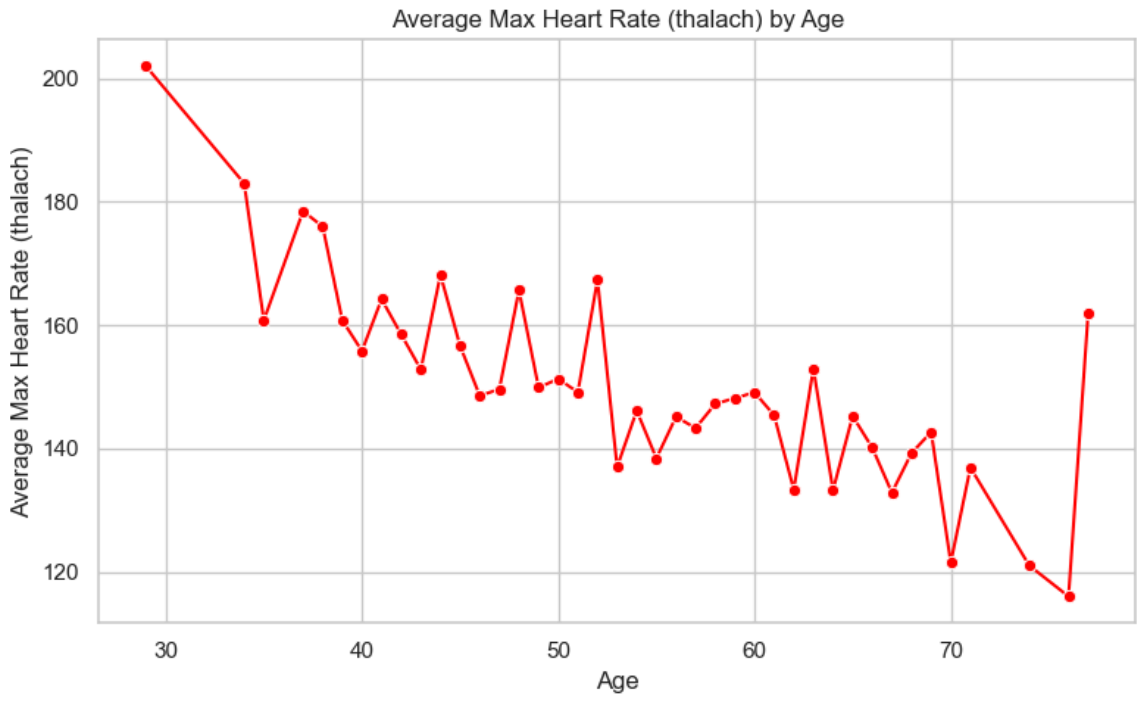
In [15]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

df = pd.read_csv('heart.csv')
df.head()
```

Out[15]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	targe
0	52	1	0	125	212	0	1	168	0	1.0	2	2	3	
1	53	1	0	140	203	1	0	155	1	3.1	0	0	3	
2	70	1	0	145	174	0	1	125	1	2.6	0	0	3	
3	61	1	0	148	203	0	1	161	0	0.0	2	1	3	
4	62	0	0	138	294	1	1	106	0	1.9	1	3	2	

LINE PLOT

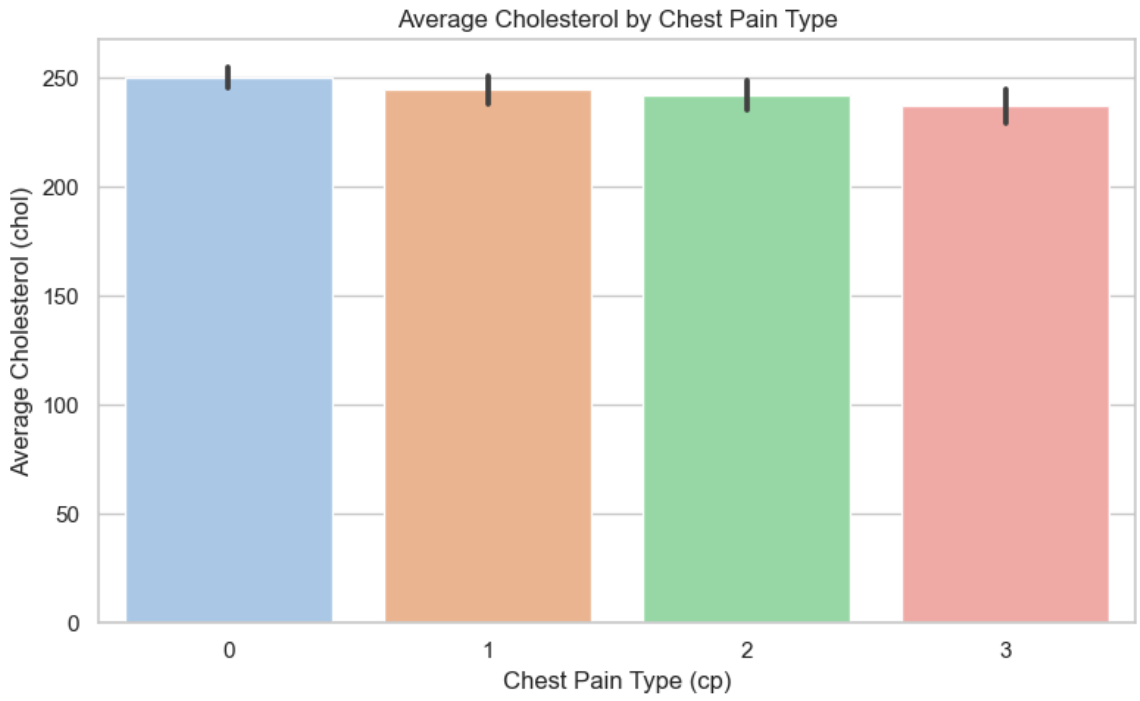
In [16]:
```
df = pd.read_csv('heart.csv')
df = df[['age', 'thalach']].dropna()
df['age'] = pd.to_numeric(df['age'], errors='coerce')
df['thalach'] = pd.to_numeric(df['thalach'], errors='coerce')
df = df.dropna()

df_grouped = df.groupby('age')['thalach'].mean().reset_index()

sns.set(style='whitegrid')
plt.figure(figsize=(8, 5))
sns.lineplot(data=df_grouped, x='age', y='thalach', marker='o', color='
plt.title('Average Max Heart Rate (thalach) by Age')
plt.xlabel('Age')
plt.ylabel('Average Max Heart Rate (thalach)')
plt.tight_layout()
```



BAR PLOT

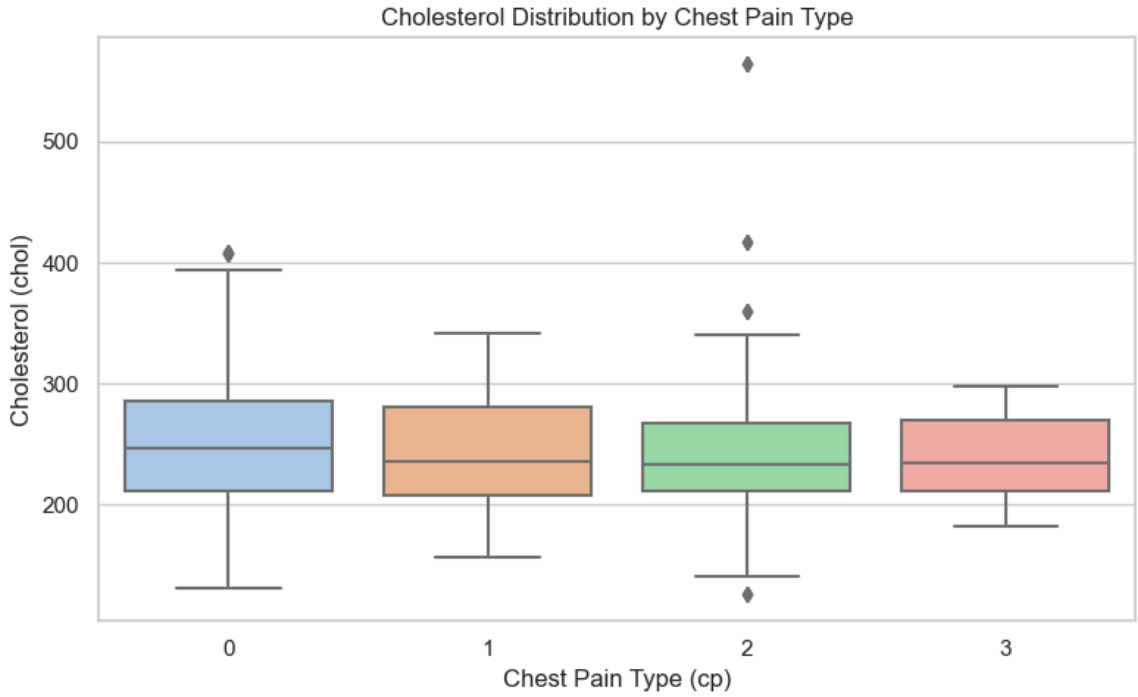
In [17]:

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

df = pd.read_csv('heart.csv')

df = df[['cp', 'chol']].dropna()
df['cp'] = pd.to_numeric(df['cp'], errors='coerce')DA
df['chol'] = pd.to_numeric(df['chol'], errors='coerce')
df = df.dropna()

sns.set(style='whitegrid')

plt.figure(figsize=(8, 5))
sns.barplot(data=df, x='cp', y='chol', estimator='mean', palette='paste
plt.title('Average Cholesterol by Chest Pain Type')
plt.xlabel('Chest Pain Type (cp)')
plt.ylabel('Average Cholesterol (chol)')
plt.tight_layout()
plt.show()
```



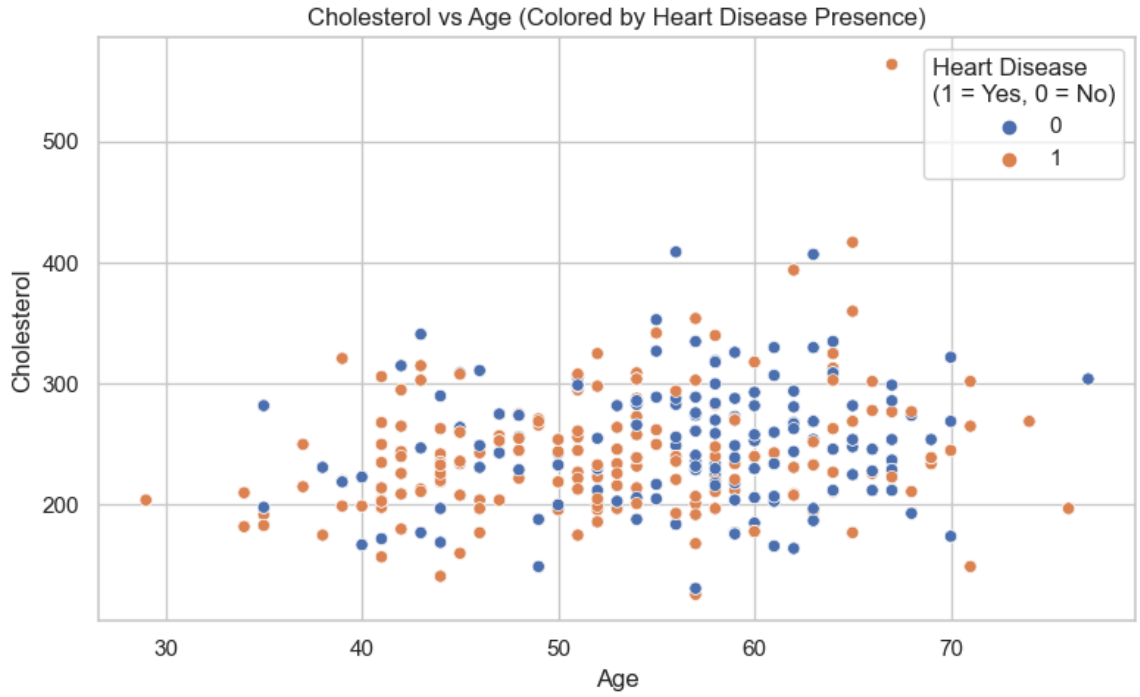
BOX PLOT

In [18]:
```
df = df[['cp', 'chol']].dropna()

sns.set(style='whitegrid')

plt.figure(figsize=(8, 5))
sns.boxplot(data=df, x='cp', y='chol', palette='pastel')
plt.title('Cholesterol Distribution by Chest Pain Type')
plt.xlabel('Chest Pain Type (cp)')
plt.ylabel('Cholesterol (chol)')
plt.tight_layout()
```



SCATTER PLOT

In [20]:
```
df = df[['age', 'chol', 'target']].dropna()
sns.set(style="whitegrid")
plt.figure(figsize=(8, 5))
sns.scatterplot(data=df, x='age', y='chol', hue='target', palette='deep
plt.title('Cholesterol vs Age (Colored by Heart Disease Presence)')
plt.xlabel('Age')
plt.ylabel('Cholesterol')
plt.legend(title='Heart Disease\n(1 = Yes, 0 = No)')
plt.tight_layout()
plt.show()
```



In [ ]:
```
1
```