bigthree-eda

June 11, 2024

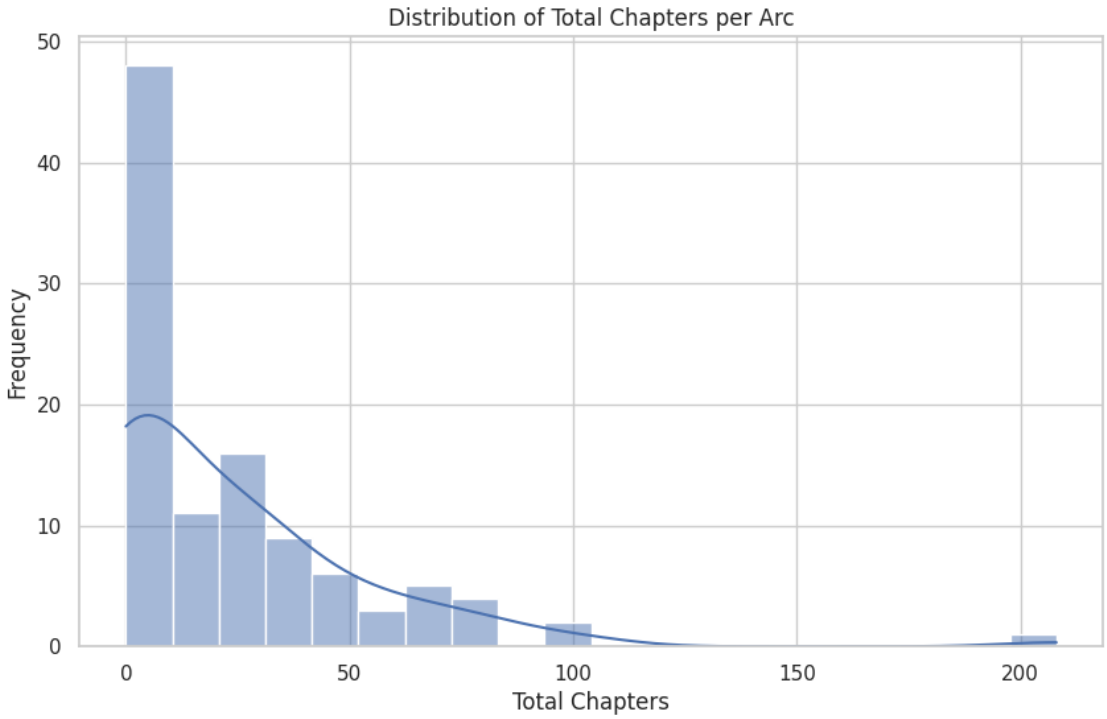
1 Exploratory Data Analysis on BigThree

[2]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set(style="whitegrid")
```

[3]:
```
df = pd.read_csv('/content/BigThree.csv')
df['Manga%'] = df['Manga%'].str.rstrip('%').astype(float)
df['Anime%'] = df['Anime%'].str.rstrip('%').astype(float)
```

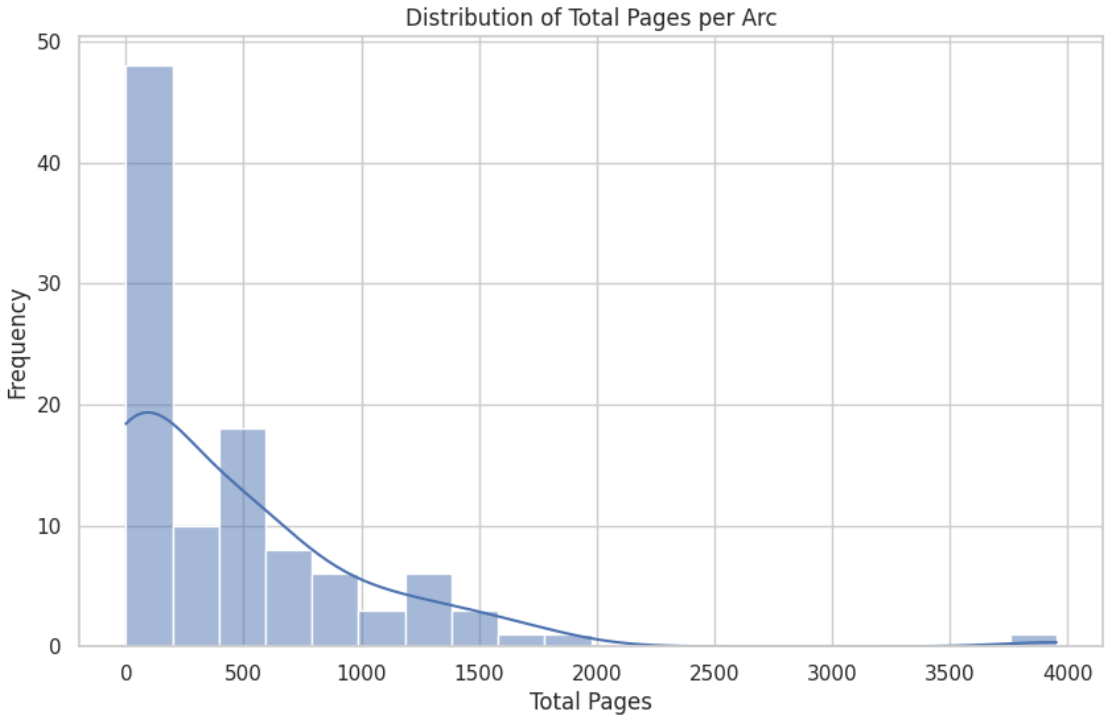
1.1 Plot 1: Distribution of Total Chapters per Arc

[4]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['TotalChapters'], bins=20, kde=True)
plt.title('Distribution of Total Chapters per Arc')
plt.xlabel('Total Chapters')
plt.ylabel('Frequency')
plt.show()
```



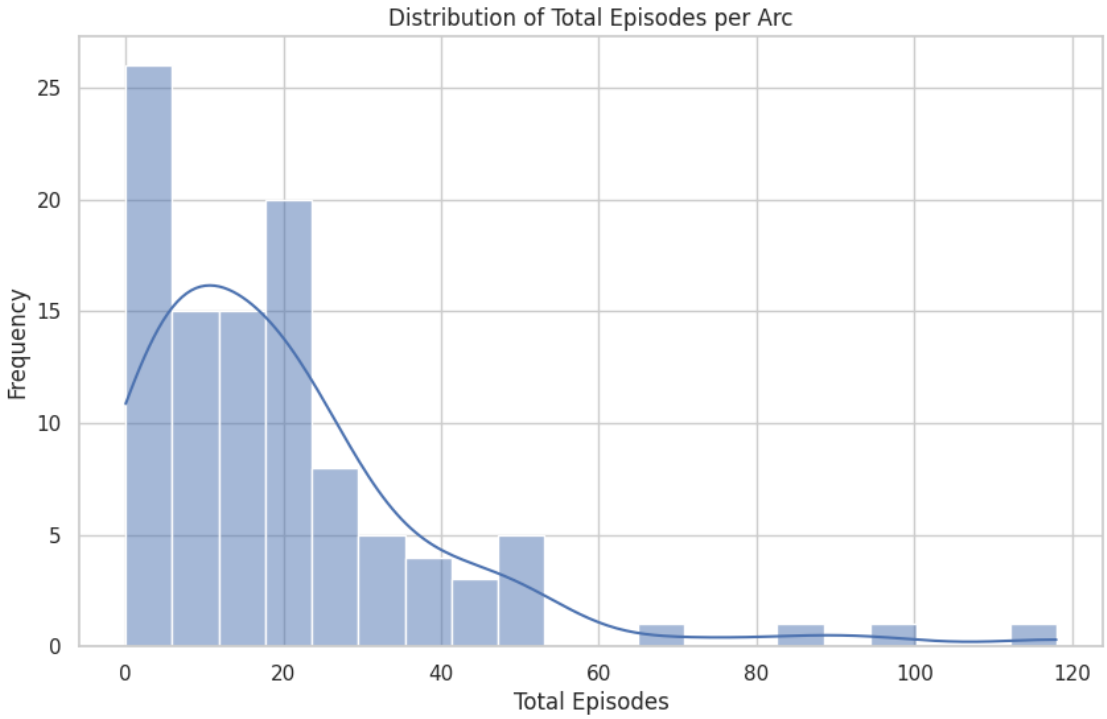
1.2 Plot 2: Distribution of Total Pages per Arc

[6]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['TotalPages'], bins=20, kde=True)
plt.title('Distribution of Total Pages per Arc')
plt.xlabel('Total Pages')
plt.ylabel('Frequency')
plt.show()
```



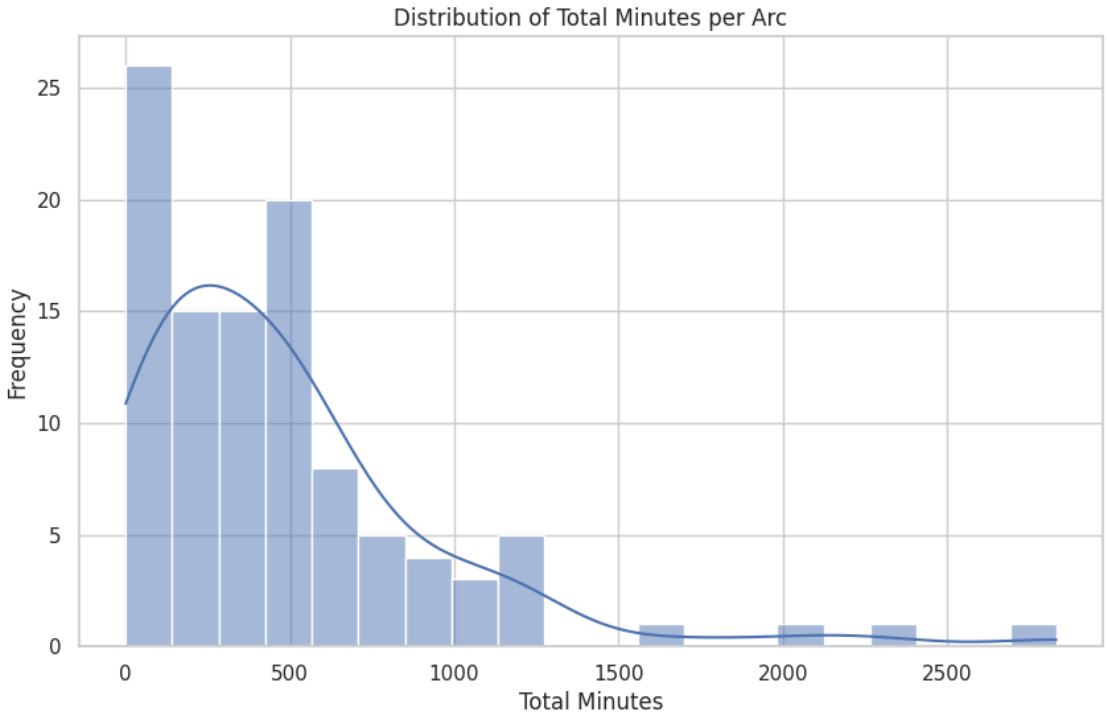
1.3 Plot 3: Distribution of Total Episodes per Arc

[8]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['TotalEpisodes'], bins=20, kde=True)
plt.title('Distribution of Total Episodes per Arc')
plt.xlabel('Total Episodes')
plt.ylabel('Frequency')
plt.show()
```



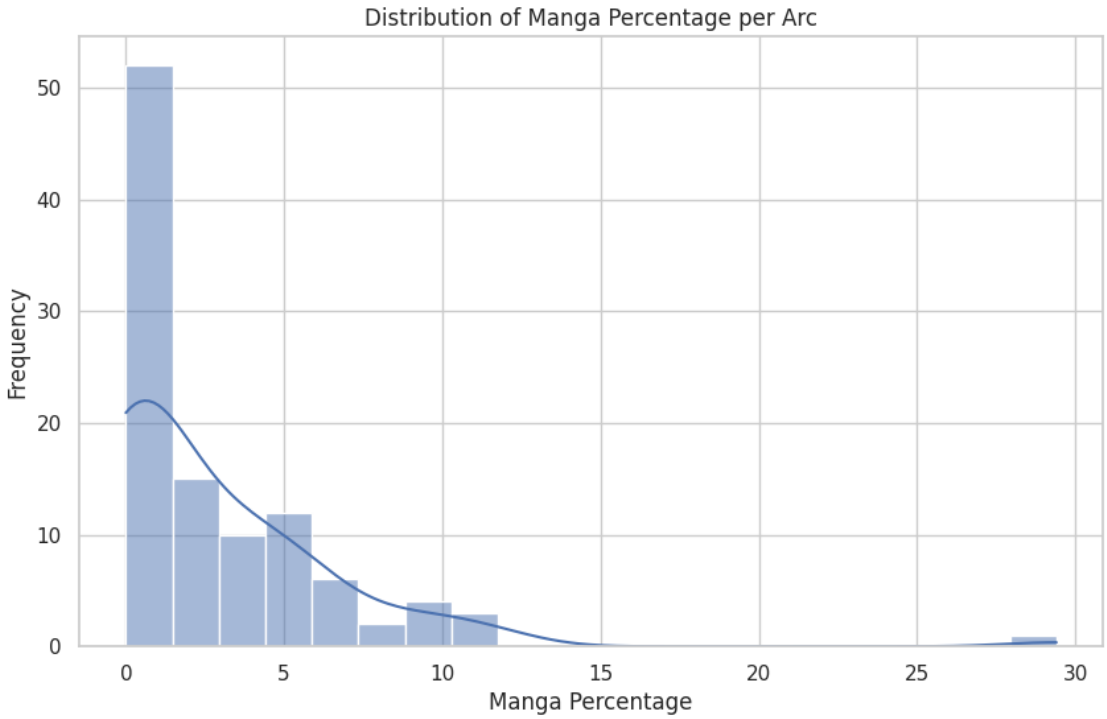
1.4 Plot 4: Distribution of Total Minutes per Arc

[9]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['TotalMinutes(avg 24)'], bins=20, kde=True)
plt.title('Distribution of Total Minutes per Arc')
plt.xlabel('Total Minutes')
plt.ylabel('Frequency')
plt.show()
```



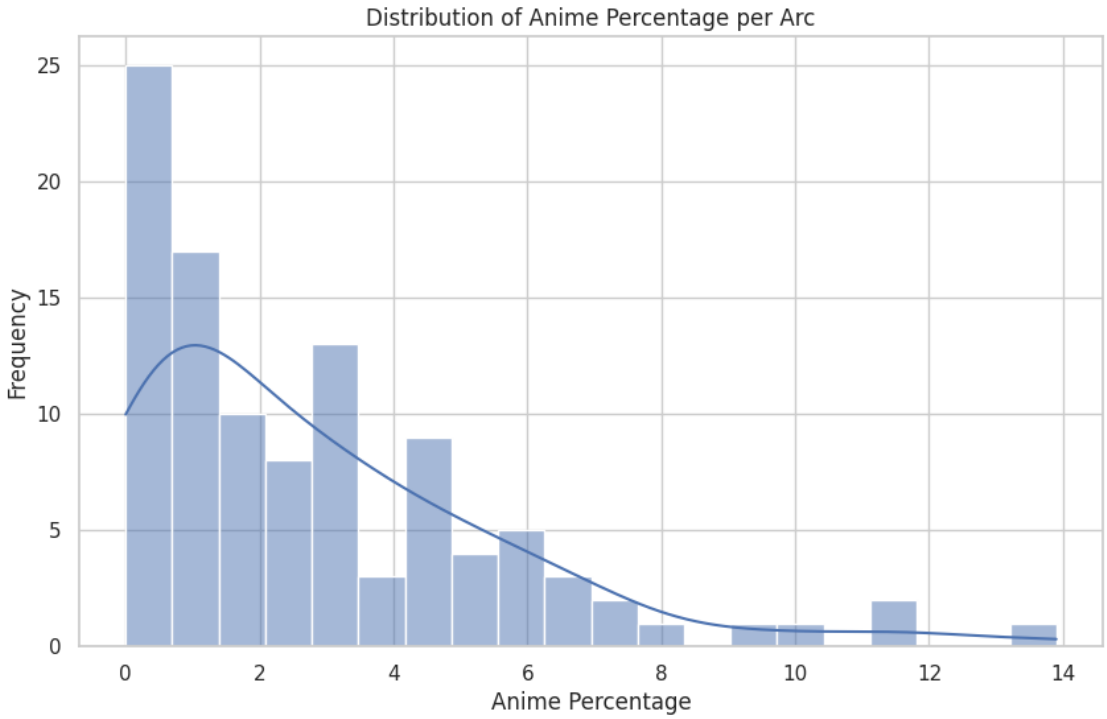
1.5 Plot 5: Distribution of Manga Percentage per Arc

[10]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['Manga%'], bins=20, kde=True)
plt.title('Distribution of Manga Percentage per Arc')
plt.xlabel('Manga Percentage')
plt.ylabel('Frequency')
plt.show()
```



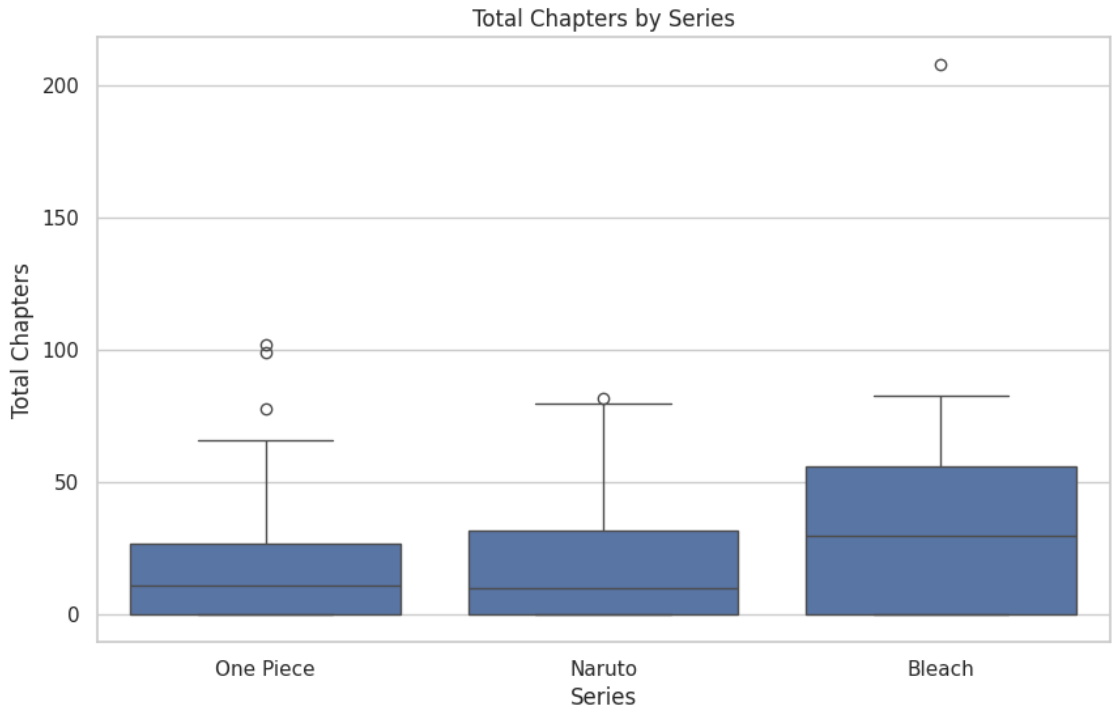
1.6 Plot 6: Distribution of Anime Percentage per Arc

[12]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['Anime%'], bins=20, kde=True)
plt.title('Distribution of Anime Percentage per Arc')
plt.xlabel('Anime Percentage')
plt.ylabel('Frequency')
plt.show()
```



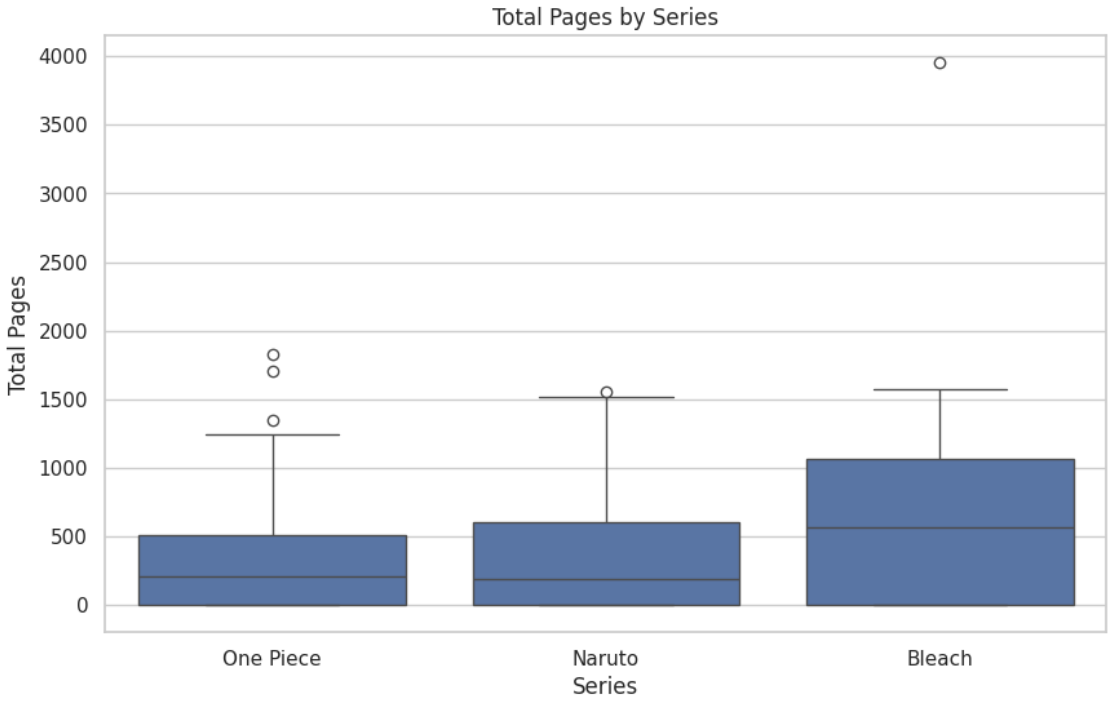
1.7 Plot 7: Total Chapters by Series

[13]:
```
plt.figure(figsize=(10, 6))
sns.boxplot(x='BigThree', y='TotalChapters', data=df)
plt.title('Total Chapters by Series')
plt.xlabel('Series')
plt.ylabel('Total Chapters')
plt.show()
```



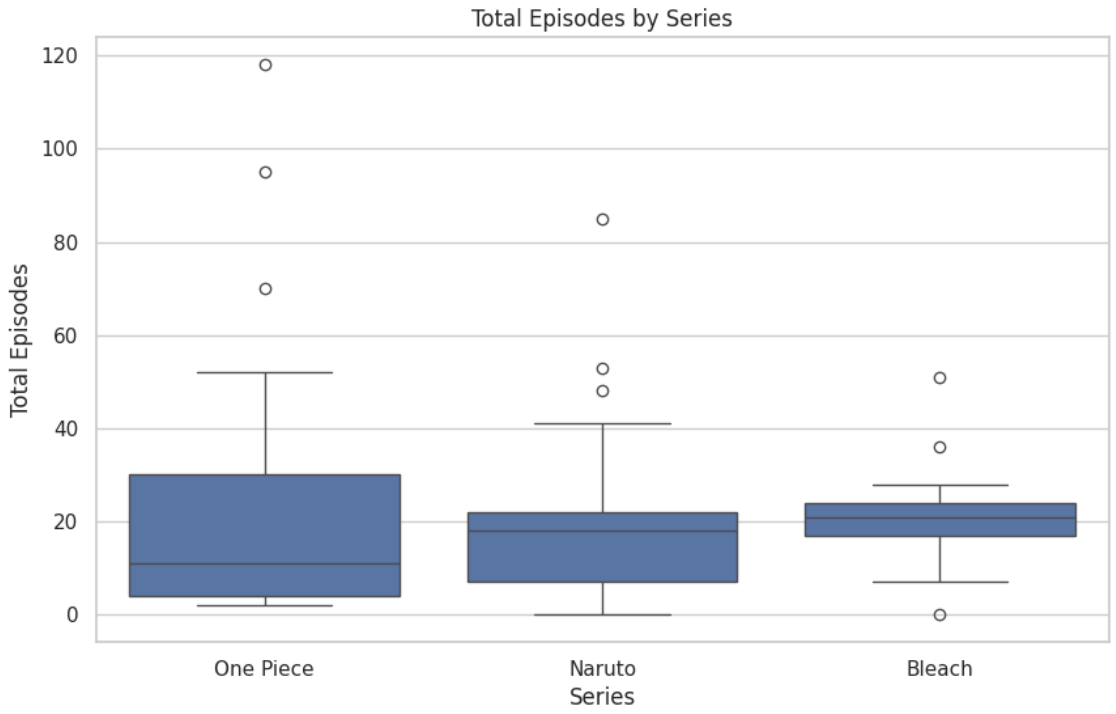
1.8 Plot 8: Total Pages by Series

[14]:
```
plt.figure(figsize=(10, 6))
sns.boxplot(x='BigThree', y='TotalPages', data=df)
plt.title('Total Pages by Series')
plt.xlabel('Series')
plt.ylabel('Total Pages')
plt.show()
```



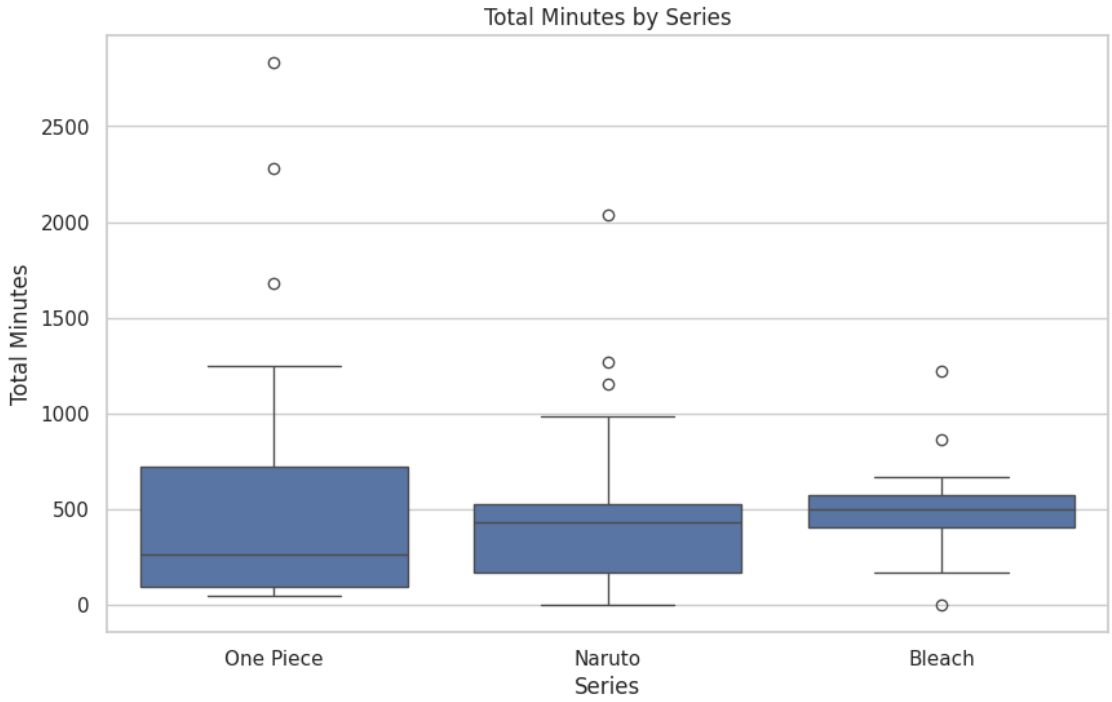
1.9 Plot 9: Total Episodes by Series

[15]:
```
plt.figure(figsize=(10, 6))
sns.boxplot(x='BigThree', y='TotalEpisodes', data=df)
plt.title('Total Episodes by Series')
plt.xlabel('Series')
plt.ylabel('Total Episodes')
plt.show()
```



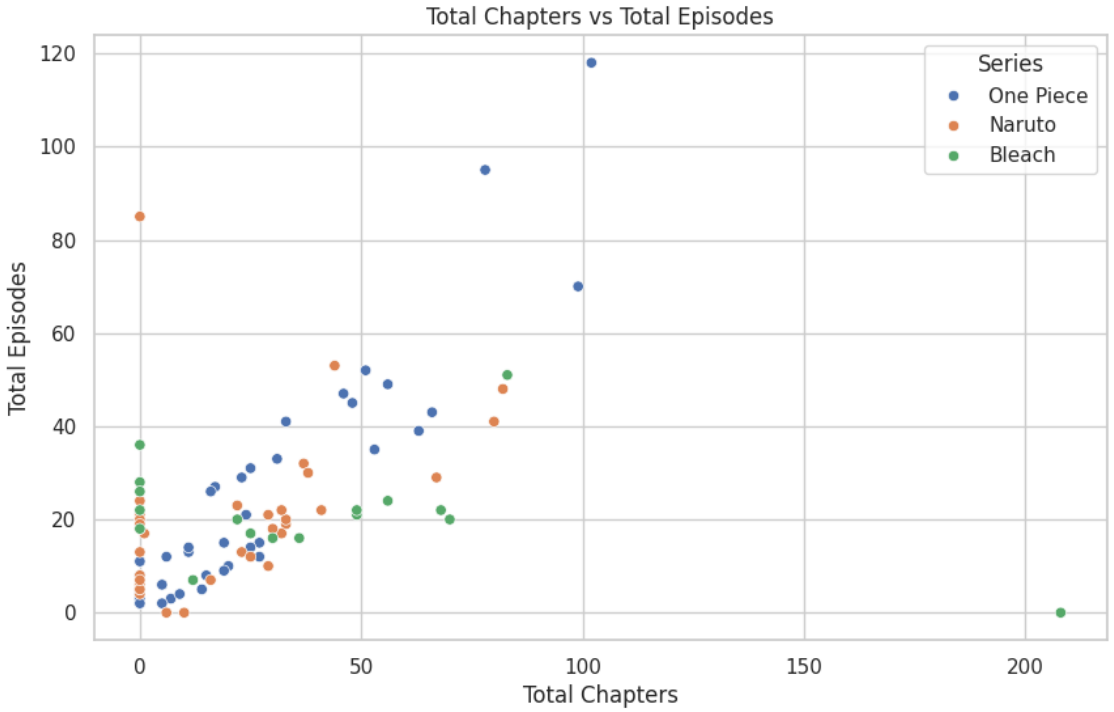
1.10 Plot 10: Total Minutes by Series

[16]:
```
plt.figure(figsize=(10, 6))
sns.boxplot(x='BigThree', y='TotalMinutes(avg 24)', data=df)
plt.title('Total Minutes by Series')
plt.xlabel('Series')
plt.ylabel('Total Minutes')
plt.show()
```



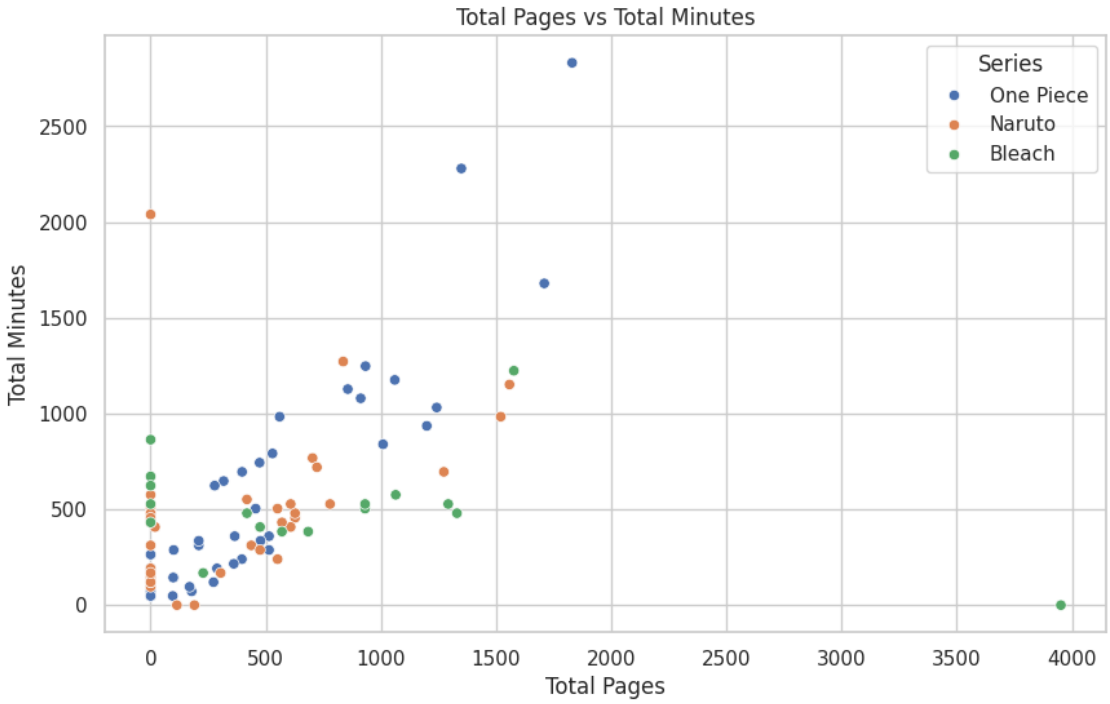
1.11 Plot 11: Relationship between Total Chapters and Total Episodes

[18]:
```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='TotalChapters', y='TotalEpisodes', hue='BigThree', data=df)
plt.title('Total Chapters vs Total Episodes')
plt.xlabel('Total Chapters')
plt.ylabel('Total Episodes')
plt.legend(title='Series')
plt.show()
```



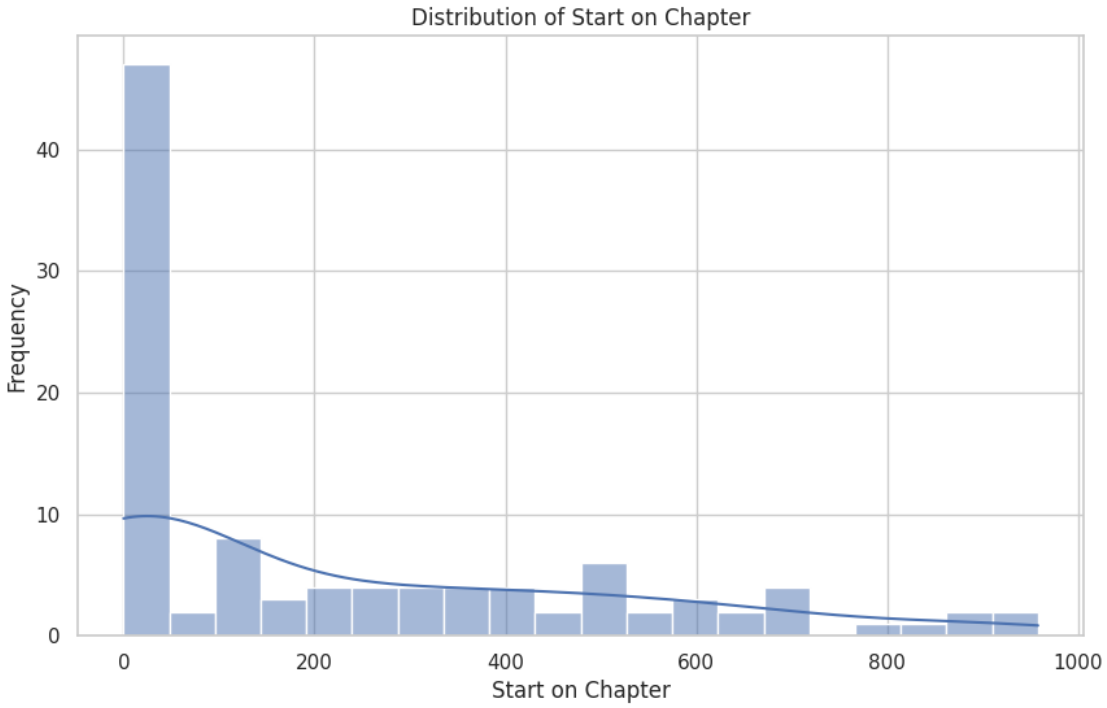
1.12 Plot 12: Relationship between Total Pages and Total Minutes

[19]:
```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='TotalPages', y='TotalMinutes(avg 24)', hue='BigThree',␣
↪data=df)
plt.title('Total Pages vs Total Minutes')
plt.xlabel('Total Pages')
plt.ylabel('Total Minutes')
plt.legend(title='Series')
plt.show()
```



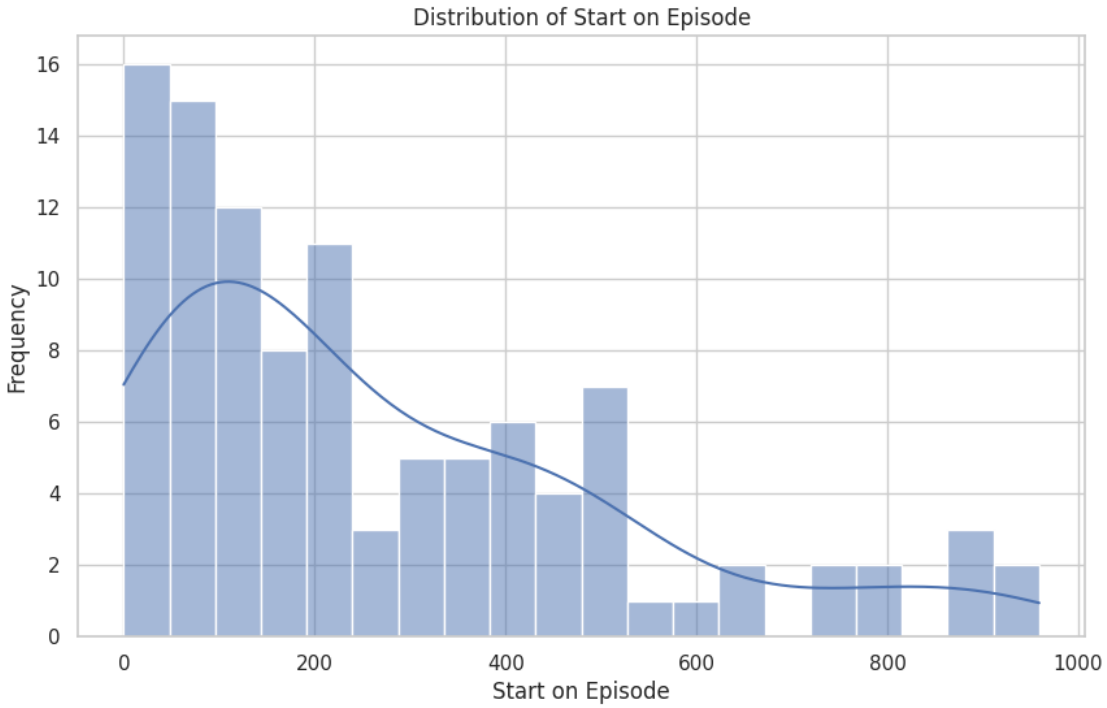
2 Plot 13: Distribution of Start on Chapter

[20]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['Start onChapter'], bins=20, kde=True)
plt.title('Distribution of Start on Chapter')
plt.xlabel('Start on Chapter')
plt.ylabel('Frequency')
plt.show()
```



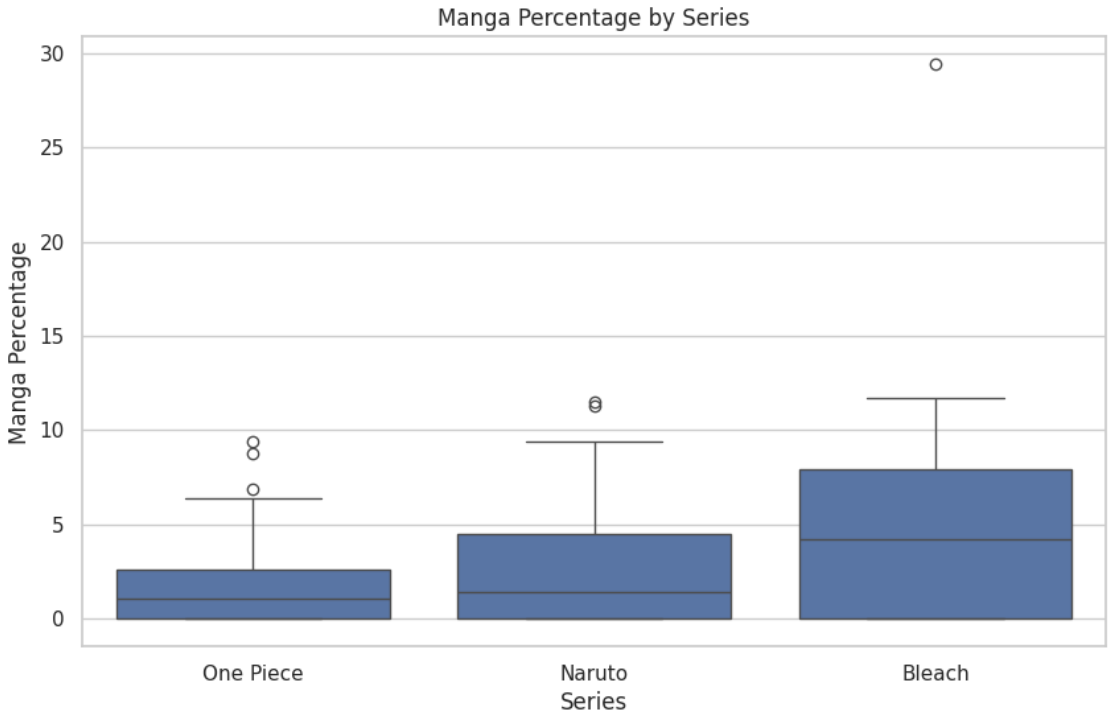
3 Plot 14: Distribution of Start on Episode

[21]:
```
plt.figure(figsize=(10, 6))
sns.histplot(df['Start onEpisode'], bins=20, kde=True)
plt.title('Distribution of Start on Episode')
plt.xlabel('Start on Episode')
plt.ylabel('Frequency')
plt.show()
```



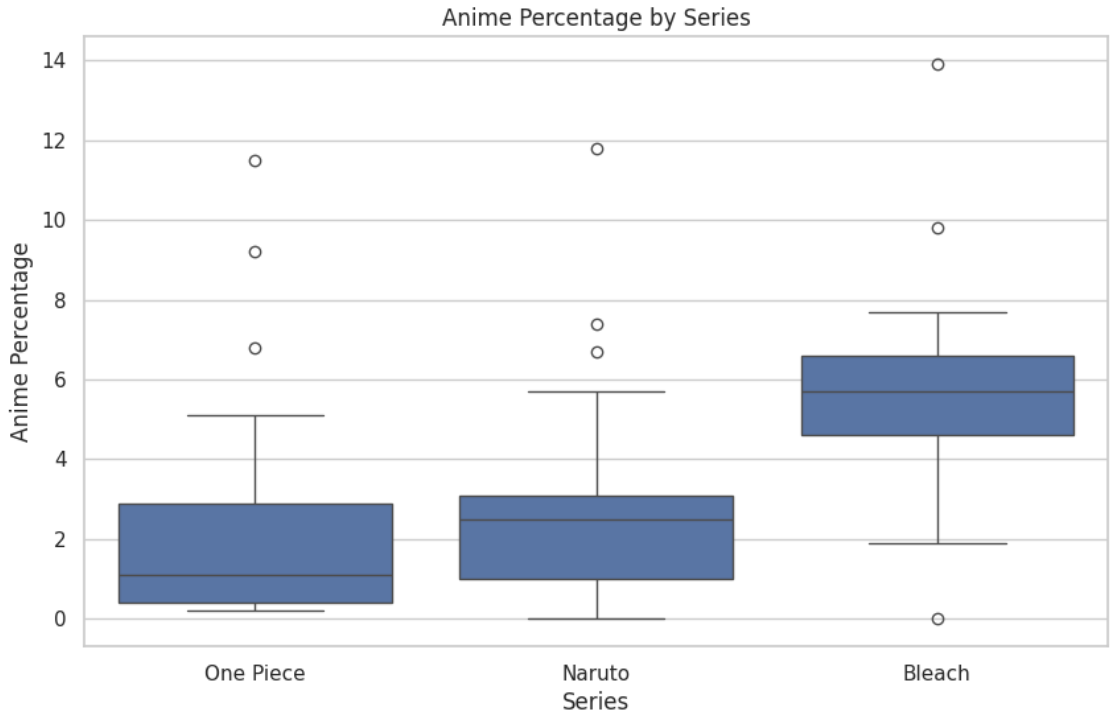
4 Plot 15: Boxplot of Manga Percentage by Series

[22]:
```
plt.figure(figsize=(10, 6))
sns.boxplot(x='BigThree', y='Manga%', data=df)
plt.title('Manga Percentage by Series')
plt.xlabel('Series')
plt.ylabel('Manga Percentage')
plt.show()
```



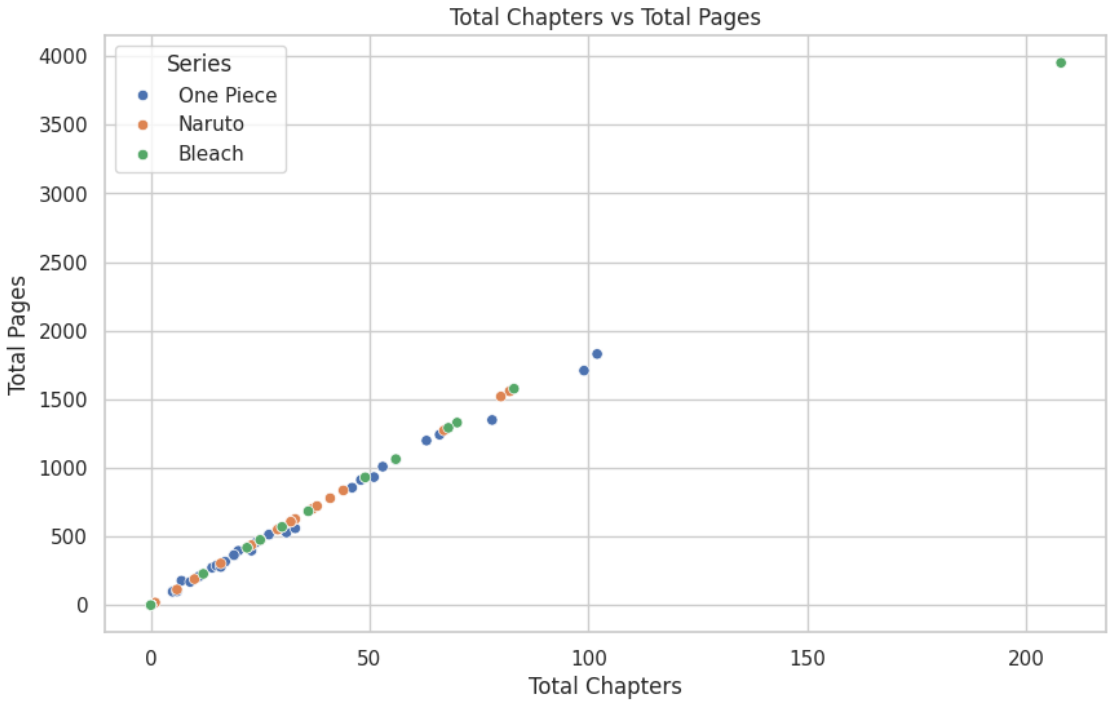
5 Plot 16: Boxplot of Anime Percentage by Series

[23]:
```
plt.figure(figsize=(10, 6))
sns.boxplot(x='BigThree', y='Anime%', data=df)
plt.title('Anime Percentage by Series')
plt.xlabel('Series')
plt.ylabel('Anime Percentage')
plt.show()
```



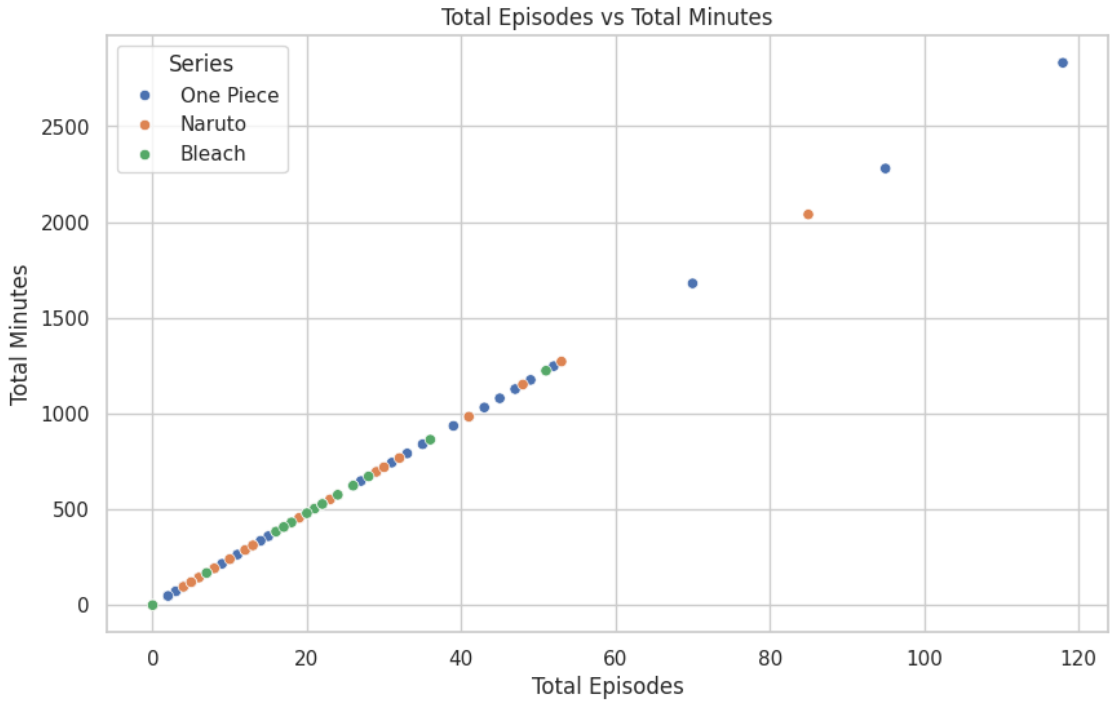
6 Plot 17: Relationship between Total Chapters and Total Pages

[24]:
```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='TotalChapters', y='TotalPages', hue='BigThree', data=df)
plt.title('Total Chapters vs Total Pages')
plt.xlabel('Total Chapters')
plt.ylabel('Total Pages')
plt.legend(title='Series')
plt.show()
```



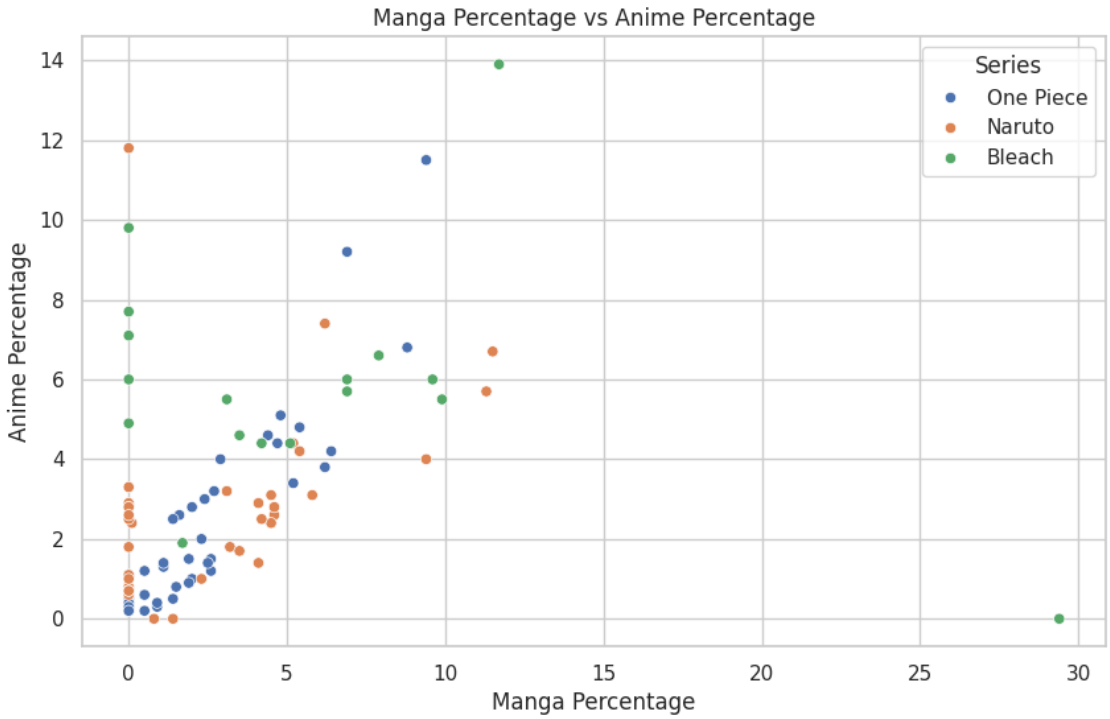
7 Plot 18: Relationship between Total Episodes and Total Minutes

[25]:
```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='TotalEpisodes', y='TotalMinutes(avg 24)', hue='BigThree',␣
↪data=df)
plt.title('Total Episodes vs Total Minutes')
plt.xlabel('Total Episodes')
plt.ylabel('Total Minutes')
plt.legend(title='Series')
plt.show()
```



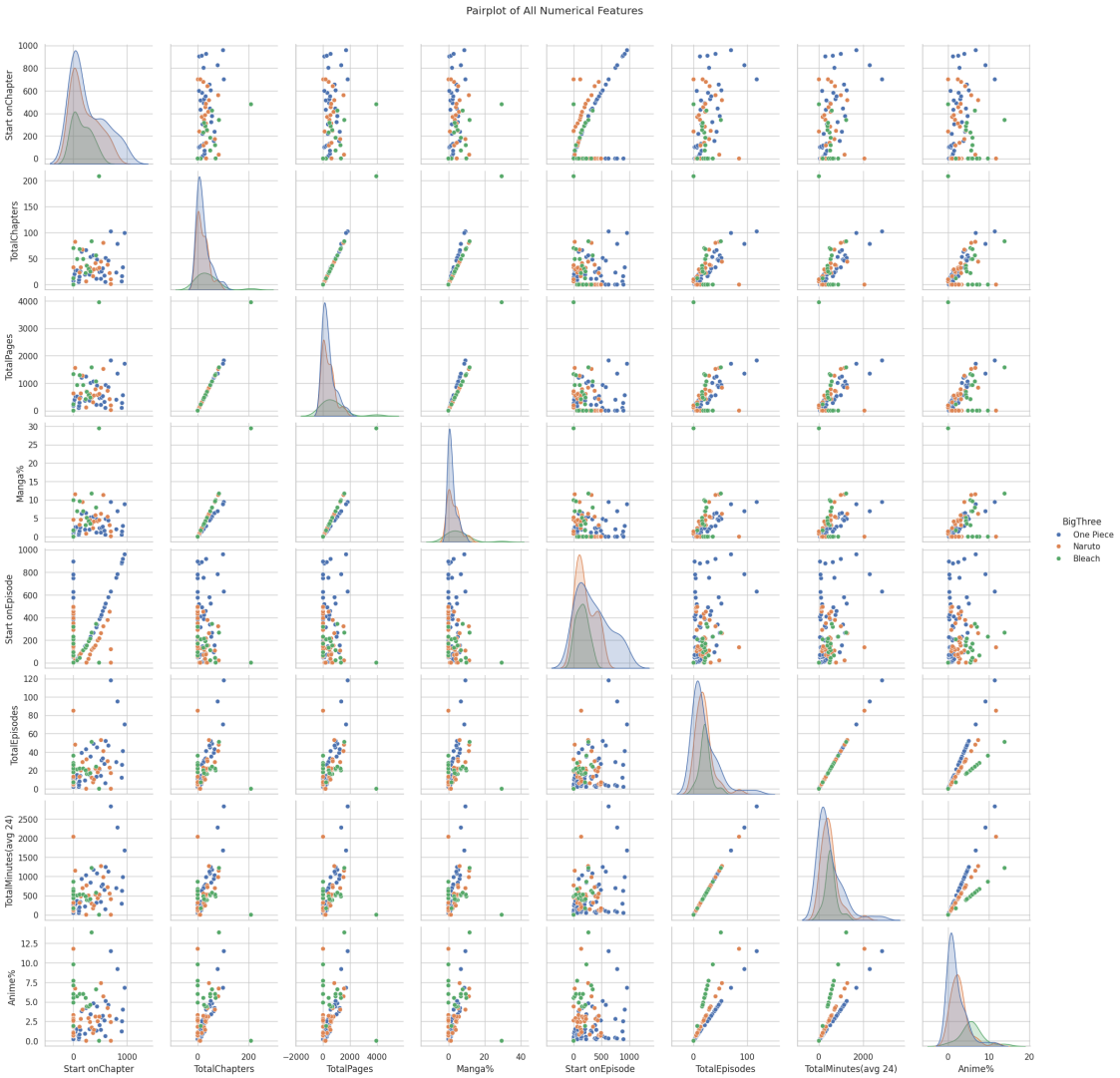
8 Plot 19: Manga Percentage vs Anime Percentage

[26]:
```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Manga%', y='Anime%', hue='BigThree', data=df)
plt.title('Manga Percentage vs Anime Percentage')
plt.xlabel('Manga Percentage')
plt.ylabel('Anime Percentage')
plt.legend(title='Series')
plt.show()
```



9 Plot 20: Pairplot of all numerical features

[27]:
```
sns.pairplot(df, hue='BigThree')
plt.suptitle('Pairplot of All Numerical Features', y=1.02)
plt.show()
```


Pairplot of All Numerical Features
Start onChapter
TotalChapters
TotalPages
Manga%
Start onEpisode
TotalEpisodes
TotalMinutes(avg 24)
Anime%
BigThree
One Piece
Naruto
Bleach

[ ]: