Visualization Codes

1.

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

# Load your DataFrame
# df = pd.read_csv('your_file.csv')  # Uncomment and specify your file
path

# List of artists
artists = ['Common', 'Kendrick Lamar', 'J. Cole', 'Drake', 'Nas', 'Lil
Wayne', 'Cardi B', 'Meagan Thee Stallion', 'Travis Scott']

# Create a dictionary to hold the correct column names for revenue
average_revenue_columns = {
    'Common': 'Average Revenue per Artist playlist: Common',
    'Kendrick Lamar': 'Average Revenue per Artist playlist: Kendrick
Lamar',
    'J. Cole': 'Average Revenue per Artist playlist: J. Cole',
    'Drake': 'Average Revenue per Artist playlist: Drake',
    'Nas': 'Average Revenue per Artist playlist:  Nas',  # Note the space
before 'Nas'
    'Lil Wayne': 'Average Revenue per Artist playlist: Lil Wayne',  #
Ensure the casing is consistent
    'Cardi B': 'Average Revenue per Artist playlist: Cardi B',
    'Meagan Thee Stallion': 'Average Revenue per Artist playlist: Meagan
Thee Stallion',
    'Travis Scott': 'Average Revenue per Artist playlist: Travis Scott'
}

# 1. Calculate Average Revenue
average_revenue = []
for artist in artists:
    col_name = average_revenue_columns[artist]
    if col_name in df.columns:
        average_revenue.append(df[col_name].mean())
    else:
        print(f"Column '{col_name}' not found.")
        average_revenue.append(0)  # or any default value you prefer
```

```
# 2. Calculate Total Revenue
total_revenue = []
for artist in artists:
    col_name = f'Total Revenue: {artist}'
    if col_name in df.columns:
        total_revenue.append(df[col_name].sum())
    else:
        print(f"Column '{col_name}' not found.")
        total_revenue.append(0)  # or any default value you prefer

# 3. Calculate Average Revenue Percentage of Total Revenue
average_revenue_percentage = [(avg / total) * 100 if total > 0 else 0 for
avg, total in zip(average_revenue, total_revenue)]

# 4. Calculate Cumulative Revenue
cumulative_revenue = pd.Series(total_revenue).cumsum()

# 5. Box Plot of Average Revenue
plt.figure(figsize=(12, 6))
plt.boxplot(average_revenue, labels=['Average Revenue'])
plt.title('Box Plot of Average Revenue')
plt.ylabel('Revenue')
plt.show()

# Visualization 1: Average Revenue as Percentage of Total Revenue per
Artist
plt.figure(figsize=(12, 6))
plt.bar(artists, average_revenue_percentage, color='teal')
plt.title('Average Revenue as Percentage of Total Revenue per Artist')
plt.xlabel('Artists')
plt.ylabel('Average Revenue Percentage (%)')
plt.xticks(rotation=45)
plt.show()

# Visualization 2: Cumulative Revenue per Artist
plt.figure(figsize=(12, 6))
plt.plot(artists, cumulative_revenue, marker='o', color='magenta')
plt.title('Cumulative Revenue per Artist')
plt.xlabel('Artists')
plt.ylabel('Cumulative Revenue')
```

```
plt.xticks(rotation=45)
plt.grid()
plt.show()

# Visualization 3: Scatter Plot of Average vs. Total Revenue
plt.figure(figsize=(12, 6))
plt.scatter(average_revenue, total_revenue, color='orange')
plt.title('Scatter Plot of Average vs. Total Revenue')
plt.xlabel('Average Revenue')
plt.ylabel('Total Revenue')
plt.axhline(y=sum(total_revenue)/len(total_revenue), color='r',
linestyle='--', label='Overall Average Total Revenue')
plt.axvline(x=sum(average_revenue)/len(average_revenue), color='b',
linestyle='--', label='Overall Average Revenue')
plt.legend()
plt.show()
```

2.

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

# Load your DataFrame
# df = pd.read_csv('your_file.csv')  # Uncomment and specify your file
path

# List of artists
artists = ['Common', 'Kendrick Lamar', 'J. Cole', 'Drake', 'Nas', 'Lil
Wayne', 'Cardi B', 'Meagan Thee Stallion', 'Travis Scott']

# Create a dictionary to hold the correct column names for revenue
average_revenue_columns = {
```

```
    'Common': 'Average Revenue per Artist playlist: Common',
    'Kendrick Lamar': 'Average Revenue per Artist playlist: Kendrick
Lamar',
    'J. Cole': 'Average Revenue per Artist playlist: J. Cole',
    'Drake': 'Average Revenue per Artist playlist: Drake',
    'Nas': 'Average Revenue per Artist playlist:  Nas',  # Note the space
before 'Nas'
    'Lil Wayne': 'Average Revenue per Artist playlist: Lil Wayne',  #
Ensure the casing is consistent
    'Cardi B': 'Average Revenue per Artist playlist: Cardi B',
    'Meagan Thee Stallion': 'Average Revenue per Artist playlist: Meagan
Thee Stallion',
    'Travis Scott': 'Average Revenue per Artist playlist: Travis Scott'
}

# 1. Average Revenue per Artist
average_revenue = []
for artist in artists:
    col_name = average_revenue_columns[artist]
    if col_name in df.columns:
        average_revenue.append(df[col_name].mean())
    else:
        print(f"Column '{col_name}' not found.")
        average_revenue.append(0)  # or any default value you prefer

# Plotting the average revenue
plt.figure(figsize=(12, 6))
plt.bar(artists, average_revenue, color='green')
plt.title('Average Revenue per Artist')
plt.xlabel('Artists')
plt.ylabel('Average Revenue')
plt.xticks(rotation=45)
plt.show()

# 2. Total Revenue per Artist
total_revenue = []
for artist in artists:
    col_name = f'Total Revenue: {artist}'
    if col_name in df.columns:
        total_revenue.append(df[col_name].sum())
```

```
    else:
        print(f"Column '{col_name}' not found.")
        total_revenue.append(0)  # or any default value you prefer

# Plotting the total revenue
plt.figure(figsize=(12, 6))
plt.bar(artists, total_revenue, color='blue')
plt.title('Total Revenue per Artist')
plt.xlabel('Artists')
plt.ylabel('Total Revenue')
plt.xticks(rotation=45)
plt.show()

# 3. Count of Revenue Greater than Average
revenue_greater_than_average = []
overall_average = sum(average_revenue) / len(average_revenue)

for artist in artists:
    col_name = average_revenue_columns[artist]
    if col_name in df.columns:
        avg_revenue = df[col_name].mean()
        revenue_greater_than_average.append(1 if avg_revenue >
overall_average else 0)
    else:
        print(f"Column '{col_name}' not found.")
        revenue_greater_than_average.append(0)

# Plotting the count of revenue greater than average
plt.figure(figsize=(12, 6))
plt.bar(artists, revenue_greater_than_average, color='orange')
plt.title('Revenue Greater than Average')
plt.xlabel('Artists')
plt.ylabel('Count (1 if above average, 0 if not)')
plt.xticks(rotation=45)
plt.show()

# 4. Average Revenue Comparison Among Artists (Horizontal Bar Chart)
plt.figure(figsize=(12, 8))
plt.barh(artists, average_revenue, color='purple')
plt.title('Average Revenue Comparison Among Artists')
```

```
plt.xlabel('Average Revenue')
plt.ylabel('Artists')
plt.show()

# 5. Distribution of Revenue
plt.figure(figsize=(12, 6))
plt.hist(df['Revenue'], bins=30, color='gray', alpha=0.7)
plt.title('Distribution of Revenue')
plt.xlabel('Revenue')
plt.ylabel('Frequency')
plt.show()
```