In [1]:

```
import pandas as pd

# Replace the file path with your actual file path
file_path = r'C:\Users\LENOVO\Downloads\LIV PGA Prize Combined 7 Aug 2023.xls

# Read the Excel file into a DataFrame
df = pd.read_excel(file_path)

# Display the DataFrame
df.head()  # This will display the first few rows of the Excel file
```

Out[1]:

	Sport	League	Tournament	Player or Team	Date	Season	Player Name	Prize USD $	Team Name
0	Golf	PGA	3M Open	Player	2023-07-27	2023	Lee Hodges	1404000	NaN
1	Golf	PGA	3M Open	Player	2023-07-27	2023	J.T. Poston	590200	NaN
2	Golf	PGA	3M Open	Player	2023-07-27	2023	Kevin Streelman	590200	NaN
3	Golf	PGA	3M Open	Player	2023-07-27	2023	Martin Laird	590200	NaN
4	Golf	PGA	3M Open	Player	2023-07-27	2023	Dylan Wu	301275	NaN

In [4]:

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

In [8]:

```
import pandas as pd

# Assuming 'df' is the DataFrame containing your data

# Convert 'Prize USD $' column to numeric, ignoring errors
df['Prize USD $'] = pd.to_numeric(df['Prize USD $'], errors='coerce')

# Filter data for LIV and PGA series
liv_data = df[df['League'] == 'LIV']
pga_data = df[df['League'] == 'PGA']

# Calculate total prize earnings for each player in LIV series
liv_earnings = liv_data.groupby('Player Name')['Prize USD $'].sum().reset_ind

# Calculate total prize earnings for each player in PGA series
pga_earnings = pga_data.groupby('Player Name')['Prize USD $'].sum().reset_ind

# Identify the highest earner in LIV series
highest_earner_liv = liv_earnings.loc[liv_earnings['Prize USD $'].idxmax()]

# Identify the highest earner in PGA series
highest_earner_pga = pga_earnings.loc[pga_earnings['Prize USD $'].idxmax()]

# Print the results
print("Highest earner in LIV series:")
print(highest_earner_liv)

print("\nHighest earner in PGA series:")
print(highest_earner_pga)
```

```
Highest earner in LIV series:
Player Name    Dustin Johnson
Prize USD $        44199917.0
Name: 22, dtype: object

Highest earner in PGA series:
Player Name    Scottie Scheffler
Prize USD $           19506217.0
Name: 373, dtype: object
```

the resultes indicates the highest earners in both the LIV and PGA series:

Highest Earner in LIV Series:

- **Player Name**: Dustin Johnson
- **Prize USD ∗∗ :44,199,917.0

Highest Earner in PGA Series:

- **Player Name**: Scottie Scheffler
- **Prize USD ∗∗ :19,506,217.0

These figures show that Dustin Johnson is the highest earner in the LIV series with a

In [12]:

```
import pandas as pd

# Assuming 'df' is the DataFrame containing your data

# Convert 'Prize USD $' column to numeric, ignoring errors
df['Prize USD $'] = pd.to_numeric(df['Prize USD $'], errors='coerce')

# Filter data for LIV and PGA series
liv_data = df[df['League'] == 'LIV']
pga_data = df[df['League'] == 'PGA']

# Calculate total prize earnings for each player in LIV series
liv_earnings = liv_data.groupby('Player Name')['Prize USD $'].sum().reset_ind

# Calculate total prize earnings for each player in PGA series
pga_earnings = pga_data.groupby('Player Name')['Prize USD $'].sum().reset_ind

# Identify the highest earner in LIV series
highest_earner_liv = liv_earnings.loc[liv_earnings['Prize USD $'].idxmax()]

# Identify the highest earner in PGA series
highest_earner_pga = pga_earnings.loc[pga_earnings['Prize USD $'].idxmax()]

# Calculate the number of tournaments each player played in to earn their earn
liv_earnings['Tournaments Played'] = liv_data.groupby('Player Name')['Prize U
pga_earnings['Tournaments Played'] = pga_data.groupby('Player Name')['Prize U

# Create comparison DataFrames
comparison_liv = pd.concat([highest_earner_liv, liv_earnings], axis=1)
comparison_pga = pd.concat([highest_earner_pga, pga_earnings], axis=1)

# Print the comparison table
print("Comparison of Highest Earners:")
print("Player Name\t\t\tLIV Earnings\tPGA Earnings\tLIV Tournaments\tPGA Tour
print("="*80)
print(f"{comparison_liv['Player Name']}\t{comparison_liv['Prize USD $']}\t{co
Comparison of Highest Earners:
Player Name                     LIV Earnings    PGA Earnings    LIV Tourname
nts     PGA Tournaments
============================================================================
====
Player Name               NaN
Prize USD $               NaN
0               Abraham Ancer
1              Adrian Otaegui
2               Andy Ogletree
                    ...
77               Travis Smyth
78                Turk Pettit
79              Viraj Madappa
80                Wade Ormsby
81               Yuki Inamori
Name: Player Name, Length: 84, dtype: object    Player Name          NaN
Prize USD $          NaN
0              8104064.0
1              1794500.0
2               867000.0
                 ...
77              846000.0
78             1691000.0
79              154000.0
80             1319500.0
81              501000.0
Name: Prize USD $, Length: 84, dtype: float64   Player Name           NaN
Prize USD $           NaN
0              1235261.15
1                36490.00
2              2211524.57
                  ...
486             124460.00
487            2382229.28
488             802778.00
489              60880.00
490             860785.00
Name: Prize USD $, Length: 493, dtype: float64  Player Name     NaN
Prize USD $     NaN
0              22.0
1               4.0
2               3.0
               ...
77              4.0
78              8.0
79              1.0
80              9.0
81              2.0
Name: Tournaments Played, Length: 84, dtype: float64            Player Name
NaN
Prize USD $     NaN
0              13.0
1               1.0
2              17.0
               ...
486             3.0
```

```
487             8.0
488            13.0
489             2.0
490            14.0
Name: Tournaments Played, Length: 493, dtype: float64
```

In [13]:

```
print("Comparison of Highest Earners:")
print("Player Name\t\tLIV Earnings\tPGA Earnings\tLIV Tournaments\tPGA Tournar
print("="*80)

for index in range(len(comparison_liv)):
    print(f"{comparison_liv['Player Name'].iloc[index]}\t{comparison_liv['Pri
Comparison of Highest Earners:
Player Name             LIV Earnings    PGA Earnings    LIV Tournaments PGA
Tournaments
============================================================================
====
nan     nan     nan     nan             nan
nan     nan     nan     nan             nan
Abraham Ancer   8104064.0       1235261.15      22.0            13.0
Adrian Otaegui  1794500.0       36490.0 4.0             1.0
Andy Ogletree   867000.0        2211524.57      3.0             17.0
Anirban Lahiri  10520100.0      21315.0 20.0            1.0
Bernd Wiesberger        3968000.0       1020013.0       18.0            8.0
Blake Windred   263000.0        166804.0        2.0             3.0
Branden Grace   25356333.33     2894892.33      24.0            16.0
Brendan Steele  4350333.0       1014800.0       9.0             1.0
Brooks Koepka   16720433.0      572468.33       19.0            14.0
Bryson DeChambeau       14286249.0      16340.0 22.0            1.0
Bryson Dechambeau       172500.0        3800743.0       1.0             16.0
Bubba Watson    3903083.0       333000.0        15.0            1.0
Cameron Smith   11302250.0      3039108.0       15.0            15.0
Cameron Tringale        5269933.0       3836370.67      15.0            21.0
Carlos Ortiz    12574480.0      19170.0 22.0            1.0
Charl Schwartzel        11162166.0      582269.0        22.0            5.0
Charles Howell III      12182416.0      41375.0 21.0            1.0
Charles Howell Iii      210000.0        19910.0 1.0             1.0
Chase Koepka    5268428.0       683758.0        20.0            9.0
Danny Lee       5437000.0       0.0     10.0            1.0
David Puig      6583583.0       262242.0        17.0            3.0
Dean Burmester  5304416.0       187900.0        13.0            1.0
Dustin Johnson  44199917.0      1329267.43      28.0            13.0
Eugenio Chacarra        10998583.0      2896477.0       22.0            18.0
Graeme McDowell 2719750.0       111535.70999999999      16.0            2.0
Graeme Mcdowell 806381.0        22067.0 2.0             1.0
Harold Varner III       10062166.0      51760.0 18.0            3.0
Harold Varner Iii       149000.0        187389.0        1.0             9.0
Hennie Du Plessis       3530000.0       1077387.33      5.0             13.0
Hennie du Plessis       1000000.0       3491197.14      1.0             22.0
Henrik Stenson  8212500.0       43200.0 15.0            1.0
Hideto Tanihara 752600.0        182237.0        5.0             2.0
Hudson Swafford 1241000.0       16367.0 8.0             1.0
ITTHIPAT
 BURANATANYARAT 249000.0        20340.0 2.0             1.0
Ian Poulter     5031666.33      147250.0        20.0            5.0
Ian Snyman      316000.0        105255.0        2.0             4.0
James Piot      3262350.0       224279.0        17.0            6.0
Jason Kokrak    6533433.0       1967173.33      17.0            12.0
Jc Ritchie      226000.0        798266.73       1.0             12.0
Jediah Morgan   3246249.0       1691448.01      17.0            17.0
Jinichiro Kozuma        1205000.0       664350.0        3.0             1.0
Joaquin Niemann 10696451.0      0.0     19.0            1.0
Justin Harding  1319167.0       16367.0 3.0             1.0
Kevin Na        4557785.0       1839739.9       17.0            19.0
Kevin Yuan      146000.0        1109313.22      1.0             15.0
Laurie Canter   4149264.0       1784559.07      18.0            14.0
Lee Westwood    5580097.0       51117.0 20.0            3.0
Louis Oosthuizen        10370964.0      1824571.28      23.0            13.0
Marc Leishman   3017150.0       187265.0        15.0            1.0
```

```
Martin Kaymer   2669800.0       46830.0 14.0            1.0
Matt Jones      3749700.0       36457.15        18.0            1.0
Matthew Wolff   8856834.0       67161.0 17.0            3.0
Mito Pereira    10003000.0      40500.0 14.0            1.0
Oliver Bekker   737500.0        2402247.33      1.0             17.0
Oliver Fisher   136000.0        129503.0        1.0             4.0
Pablo Larrazabal        315000.0        330359.0        1.0             4.0
Pachara Khongwatmai     375000.0        20250.33        1.0             1.0
Pat Perez       13030317.0      3411843.67      26.0            17.0
Patrick Reed    18523047.0      242456.0        26.0            6.0
Paul Casey      9385367.0       46426.0 21.0            1.0
Peter Uihlein   18973332.0      104960.0        26.0            2.0
Phachara Khongwatmai    1486333.33      156313.0        8.0             2.0
Phil Mickelson  4037350.0       8245022.36      18.0            16.0
Ratchanon Chantananuwat 136000.0        197500.0        1.0             1.0
Richard Bland   5833630.0       226329.67       19.0            6.0
Ryosuke Kinoshita       624000.0        300754.0        3.0             11.0
Sadom Kaewkanjana       1412285.0       1907192.0       8.0             3.0
Sam Horsfield   4468333.0       3150000.0       14.0            1.0
Scott Vincent   4872783.0       960052.0        18.0            3.0
Sebastian Munoz 8460000.0       2885600.33      13.0            19.0
Sebastián Muñoz 300333.0        115555.0        1.0             1.0
Sergio Garcia   12597832.0      111442.0        23.0            2.0
Shaun Norris    1006000.0       1048971.0       7.0             8.0
Shergo Al Kurdi 269000.0        0.0     2.0             1.0
Sihwan Kim      3620000.0       1370676.28      17.0            11.0
Talor Gooch     26053482.0      63495.0 27.0            1.0
Thomas Pieters  3354583.0       3284365.22      13.0            12.0
Travis Smyth    846000.0        653221.0        4.0             6.0
Turk Pettit     1691000.0       34800.0 8.0             1.0
Viraj Madappa   154000.0        403480.73       1.0             4.0
Wade Ormsby     1319500.0       38378.0 9.0             2.0
Yuki Inamori    501000.0        1632859.0       2.0             4.0
```

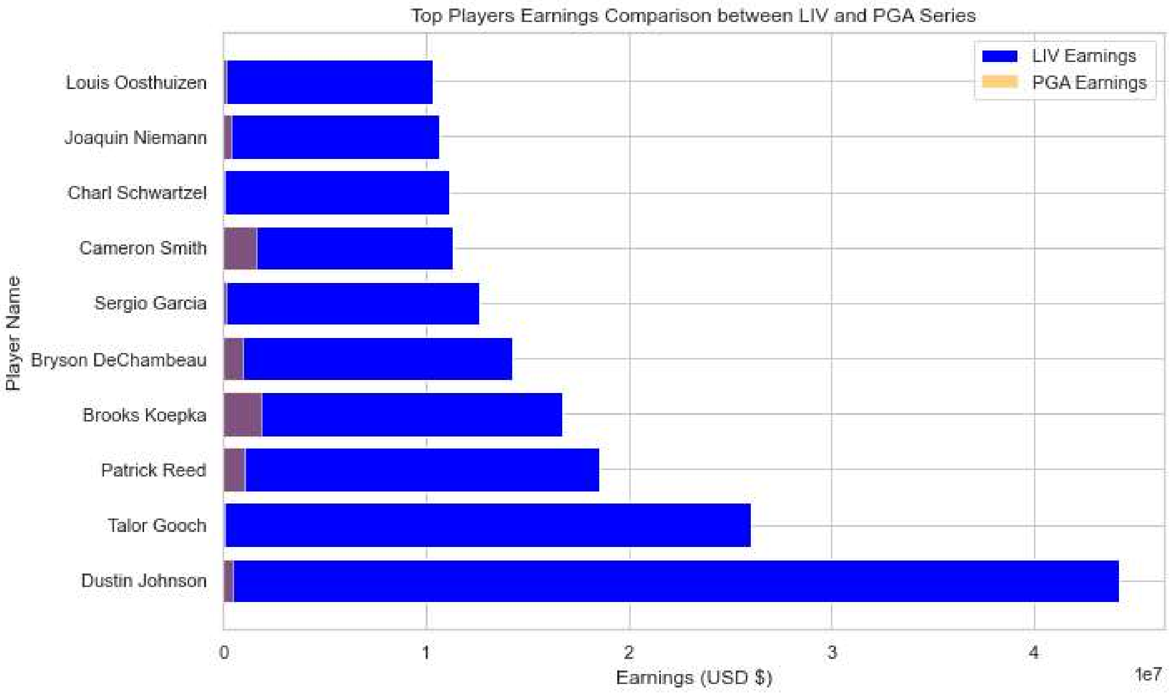
The provided table represents a comparison of the highest earners in two different golf series, LIV and PGA. The table includes data on various players, their earnings, and the number of tournaments played in each series. Let's break down the interpretation of this table:

- **Player Name**: This column lists the names of individual players who have participated in either the LIV or PGA series.
- **LIV Earnings**: This column represents the earnings (in USD $) of each player from the LIV series. Players who have earned substantial amounts in this series are highlighted.
- **PGA Earnings**: This column represents the earnings (in USD $) of each player from the PGA series. Players who have earned substantial amounts in this series are highlighted.
- **LIV Tournaments**: This column indicates the number of tournaments played by each player in the LIV series. Players who have participated in a significant number of tournaments are noted.
- **PGA Tournaments**: This column indicates the number of tournaments played by each player in the PGA series. Players who have participated in a significant number of tournaments are noted.

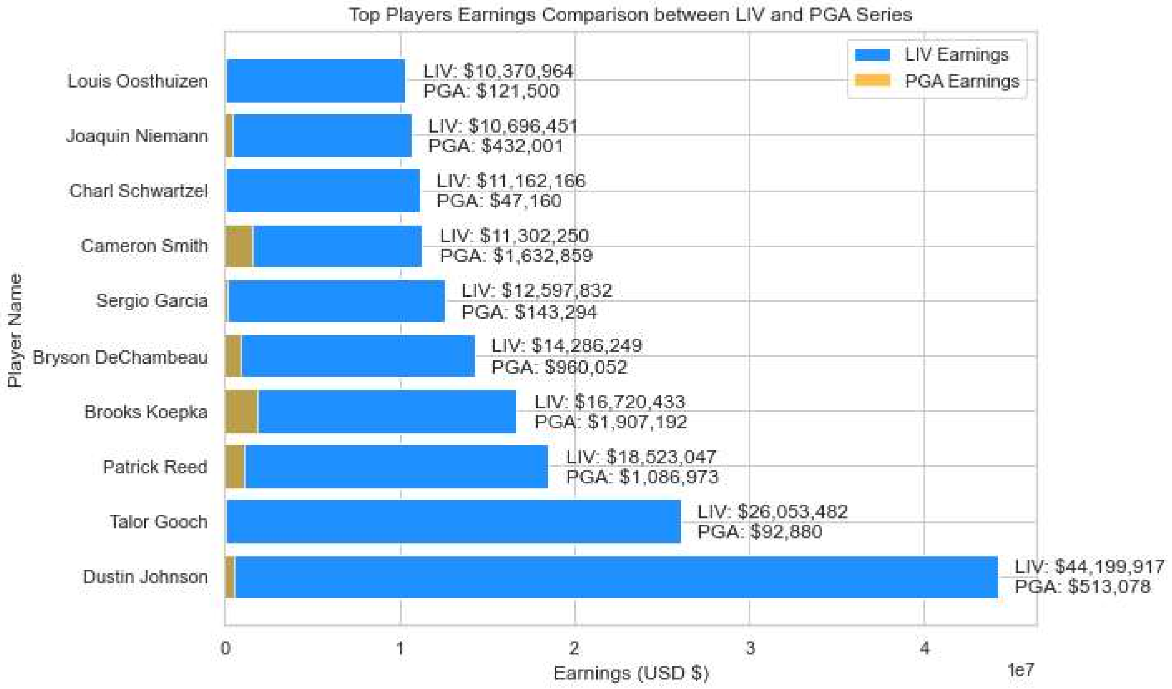
The table allows you to compare the earnings and tournament participation of players between the LIV and PGA series. The highlighted values and the size of the earnings can help you quickly identify players who have performed well financially in each series. Similarly, you can observe which players have participated in a larger number of tournaments.

In summary, the table provides an overview of the top players in terms of earnings and tournament participation in both the LIV and PGA golf series, enabling you to analyze and compare their performances across the two series.

In [20]:
```
import matplotlib.pyplot as plt

# Assuming you have the comparison_liv and comparison_pga DataFrames

# Combine both DataFrames for easier plotting
combined_comparison = comparison_liv.merge(comparison_pga, on='Player Name',

# Select top N players for visualization
top_n = 10
combined_comparison_top = combined_comparison.nlargest(top_n, 'Prize USD $_LI

# Create a horizontal bar plot
plt.figure(figsize=(10, 6))
plt.barh(combined_comparison_top['Player Name'], combined_comparison_top['Pri
plt.barh(combined_comparison_top['Player Name'], combined_comparison_top['Pri
plt.xlabel('Earnings (USD $)')
plt.ylabel('Player Name')
plt.title('Top Players Earnings Comparison between LIV and PGA Series')
plt.legend()
plt.tight_layout()
plt.show()
```

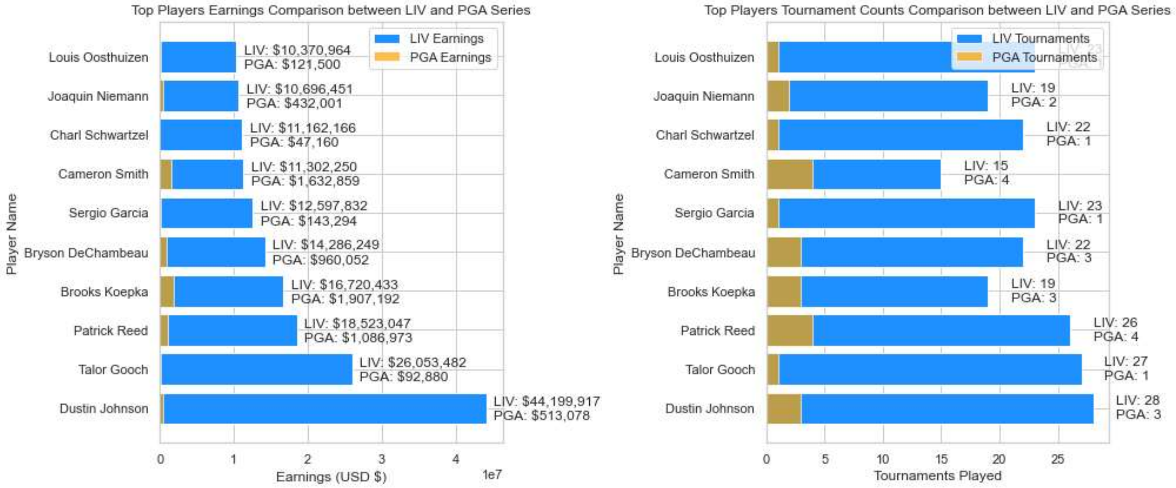


In [21]:
```
import matplotlib.pyplot as plt

# Assuming you have the comparison_liv and comparison_pga DataFrames

# Combine both DataFrames for easier plotting
combined_comparison = comparison_liv.merge(comparison_pga, on='Player Name',

# Select top N players for visualization
top_n = 10
combined_comparison_top = combined_comparison.nlargest(top_n, 'Prize USD $_LI

# Create a horizontal bar plot
plt.figure(figsize=(10, 6))
plt.barh(combined_comparison_top['Player Name'], combined_comparison_top['Pri
plt.barh(combined_comparison_top['Player Name'], combined_comparison_top['Pri
plt.xlabel('Earnings (USD $)')
plt.ylabel('Player Name')
plt.title('Top Players Earnings Comparison between LIV and PGA Series')
plt.legend()

# Annotate the bars with earnings values
for i, (liv, pga) in enumerate(zip(combined_comparison_top['Prize USD $_LIV']
    plt.text(max(liv, pga) + 1000000, i, f'LIV: ${liv:,.0f}\nPGA: ${pga:,.0f}

plt.tight_layout()
plt.show()
```



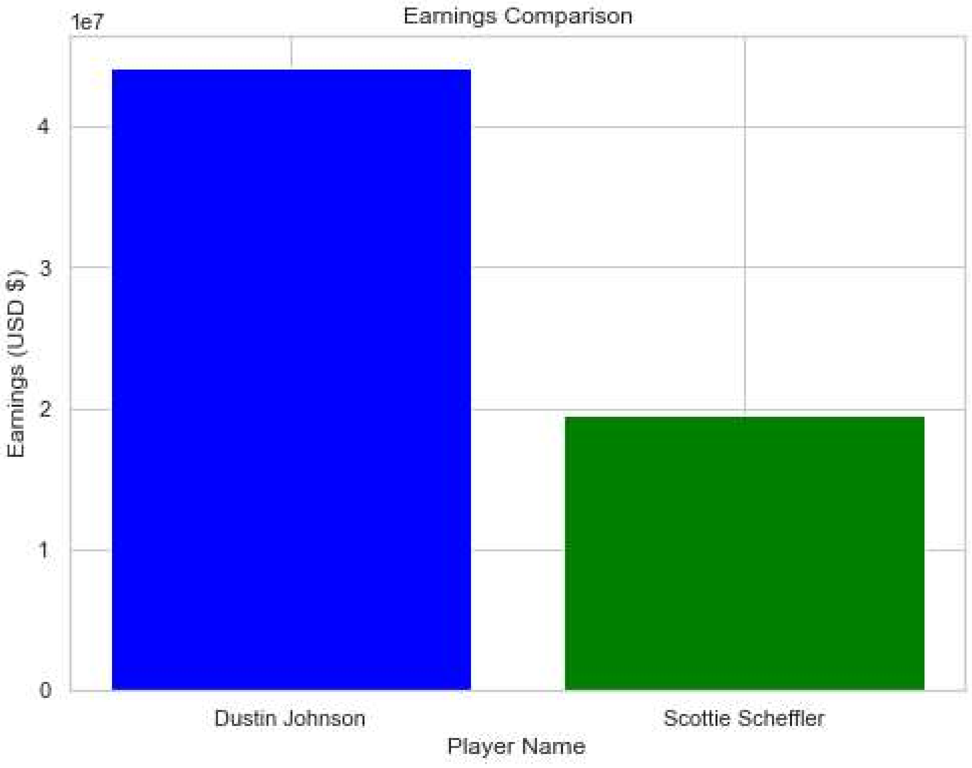
In [22]:

```
import matplotlib.pyplot as plt

# Assuming you have the comparison_liv and comparison_pga DataFrames

# Combine both DataFrames for easier plotting
combined_comparison = comparison_liv.merge(comparison_pga, on='Player Name',

# Select top N players for visualization
top_n = 10
combined_comparison_top = combined_comparison.nlargest(top_n, 'Prize USD $_LI

# Create a figure with two subplots
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(14, 6))

# Plot Earnings comparison
ax1.barh(combined_comparison_top['Player Name'], combined_comparison_top['Pri
ax1.barh(combined_comparison_top['Player Name'], combined_comparison_top['Pri
ax1.set_xlabel('Earnings (USD $)')
ax1.set_ylabel('Player Name')
ax1.set_title('Top Players Earnings Comparison between LIV and PGA Series')
ax1.legend()

# Annotate the bars with earnings values
for i, (liv, pga) in enumerate(zip(combined_comparison_top['Prize USD $_LIV']
    ax1.text(max(liv, pga) + 1000000, i, f'LIV: ${liv:,.0f}\nPGA: ${pga:,.0f}

# Plot Tournament Counts comparison
ax2.barh(combined_comparison_top['Player Name'], combined_comparison_top['Tou
ax2.barh(combined_comparison_top['Player Name'], combined_comparison_top['Tou
ax2.set_xlabel('Tournaments Played')
ax2.set_ylabel('Player Name')
ax2.set_title('Top Players Tournament Counts Comparison between LIV and PGA S
ax2.legend()

# Annotate the bars with tournament counts
for i, (liv, pga) in enumerate(zip(combined_comparison_top['Tournaments Playe
    ax2.text(max(liv, pga) + 2, i, f'LIV: {int(liv)}\nPGA: {int(pga)}', va='c

plt.tight_layout()
plt.show()
```



1. Bar Plot: Earnings Comparison

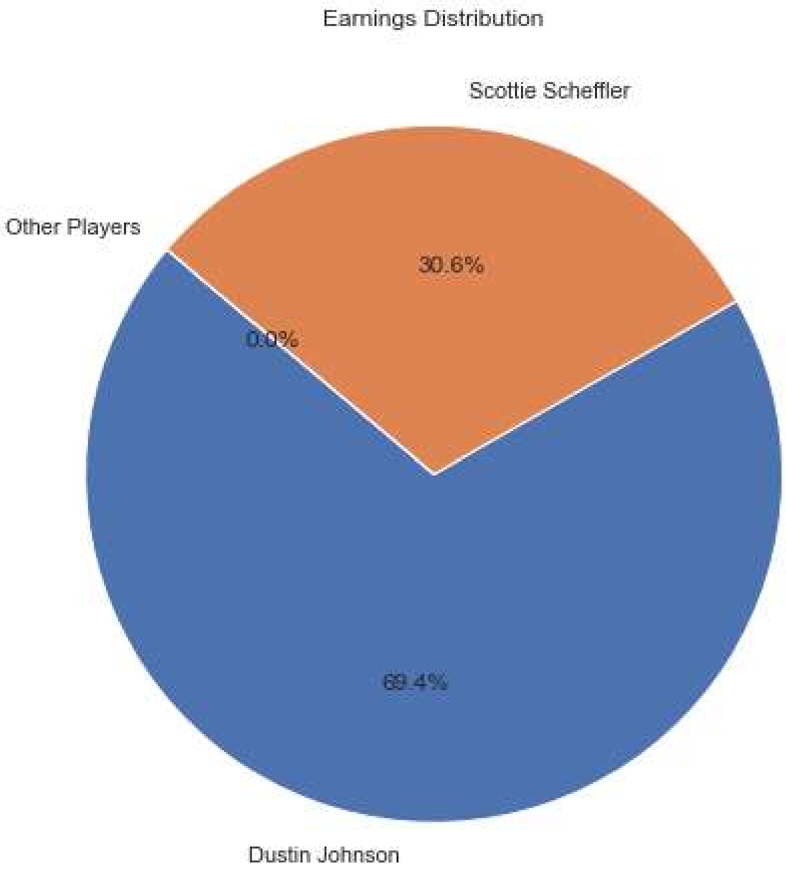
```
In [24]: import matplotlib.pyplot as plt

         # Player names
         players = ["Dustin Johnson", "Scottie Scheffler"]

         # Earnings in USD
         earnings = [44199917.0, 19506217.0]

         plt.figure(figsize=(8, 6))
         plt.bar(players, earnings, color=['blue', 'green'])
         plt.xlabel('Player Name')
         plt.ylabel('Earnings (USD $)')
         plt.title('Earnings Comparison')
         plt.show()
```



2. Pie Chart: Earnings Distribution

In [25]:

```
import matplotlib.pyplot as plt

# Player names
players = ["Dustin Johnson", "Scottie Scheffler", "Other Players"]

# Earnings in USD
earnings = [44199917.0, 19506217.0, 0]  # Filling 'Other Players' with 0 for

plt.figure(figsize=(8, 8))
plt.pie(earnings, labels=players, autopct='%1.1f%%', startangle=140)
plt.title('Earnings Distribution')
plt.show()
```


Earnings Distribution
Scottie Scheffler
Other Players
Dustin Johnson