Notebook 3 - Final Visual and Resources

sxm4bz

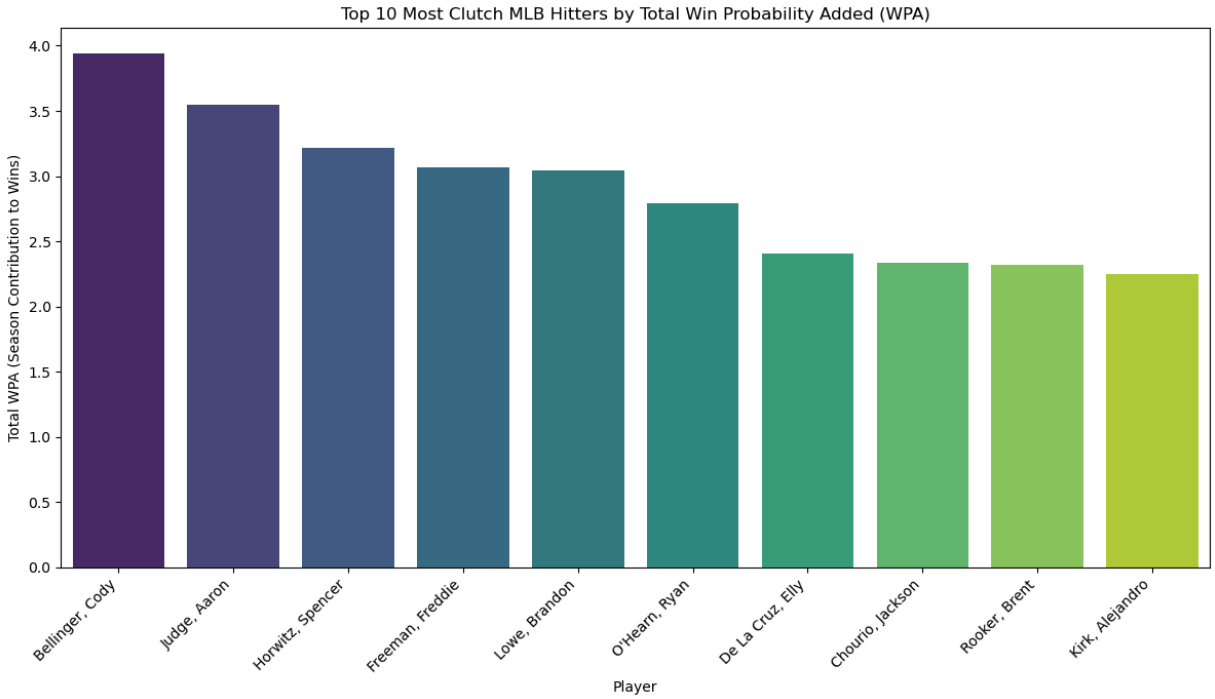
Ryan Colavita

In this notebook, I create my final visualization putting all the data that was loaded in notebook 1 and calculated in notebook 2, into one visual. In the visual, I decided to display the top 10 players since it would be easier to see a ton of players. I have also added a resources section that gives a summary and description of where all of my data came from.

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

#Making the Visual

try:
    clutch_stats_ranked = pd.read_csv('clutch_stats_ranked.csv')
except FileNotFoundError:
    print("ERROR: 'clutch_stats_ranked.csv' not found. Please run Notebook 2 first.
    exit()

clutch_stats_ranked['Total WPA'] = clutch_stats_ranked['total_wpa'].round(3)
clutch_stats_ranked.rename(columns={'player_name': 'Player', 'plate_appearances': '
top_10 = clutch_stats_ranked[['Player', 'Total WPA', 'PA']].head(10)


plt.figure(figsize=(12, 7))
sns.barplot(
    x='Player',
    y='Total WPA',
    data=top_10,
    palette=sns.color_palette("viridis", len(top_10))
)
plt.title('Top 10 Most Clutch MLB Hitters by Total Win Probability Added (WPA)')
plt.xlabel('Player')
plt.ylabel('Total WPA (Season Contribution to Wins)')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()

final_product_image_path = 'most_clutch_player_final_graph.png'
plt.savefig(final_product_image_path)
plt.show()
#

print(f"Final graph saved to: {final_product_image_path}")

print("\n" + "="*70)
print("Final Ranking: Most Clutch Hitters by Total WPA")
```

```
print("="*70)
print(top_10.to_markdown(index=False))
```

```
/tmp/ipykernel_740/3174579584.py:19: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.1
4.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

  sns.barplot(
```




```
Final graph saved to: most_clutch_player_final_graph.png

======================================================================
Final Ranking: Most Clutch Hitters by Total WPA
======================================================================
| Player           |   Total WPA |   PA |
|:-----------------|------------:|-----:|
| Bellinger, Cody  |       3.938 |  300 |
| Judge, Aaron     |       3.548 |  260 |
| Horwitz, Spencer |       3.214 |  224 |
| Freeman, Freddie |       3.07  |  264 |
| Lowe, Brandon    |       3.044 |  234 |
| O'Hearn, Ryan    |       2.794 |  246 |
| De La Cruz, Elly |       2.41  |  308 |
| Chourio, Jackson |       2.332 |  246 |
| Rooker, Brent    |       2.322 |  272 |
| Kirk, Alejandro  |       2.25  |  190 |
```

In [4]:
```
#Resources
manifest = pd.DataFrame({
    'Resource Name': [
        'Raw Statcast Data',
        'Notebook 1: Data Establishment',
        'Notebook 2: Exploratory Analysis',
        'Notebook 3: Final Product',
        'Calculated Clutch Stats (CSV)',
```

```
        'External Resource: Baseball Savant'
    ],
    'Brief Description': [
        'Pitch-by-pitch data for the season (must contain WPA/LI columns).',
        'Acquires, reads, and verifies the raw data.',
        'Calculates WPA, filters, and generates exploratory graphs.',
        'Presents final ranking table and bar chart.',
        'Intermediate file containing aggregated WPA metrics.',
        'https://baseballsavant.mlb.com/statcast_search'
    ],
    'Link/File Location': [
        'savant_data.csv',
        'data_establishment.ipynb',
        'exploration.ipynb',
        'final_product.ipynb',
        'clutch_stats_ranked.csv',
        'External Link'
    ]
})

print("\n\nResources")
print(manifest.to_markdown(index=False))
```

```
Resources
| Resource Name                      | Brief Description
| Link/File Location       |
|:-----------------------------------|:-----------------------------------------
---------------------|:-------------------------|
| Raw Statcast Data                  | Pitch-by-pitch data for the season (must cont
ain WPA/LI columns). | savant_data.csv          |
| Notebook 1: Data Establishment     | Acquires, reads, and verifies the raw data.
| data_establishment.ipynb |
| Notebook 2: Exploratory Analysis   | Calculates WPA, filters, and generates explor
atory graphs.        | exploration.ipynb        |
| Notebook 3: Final Product          | Presents final ranking table and bar chart.
| final_product.ipynb      |
| Calculated Clutch Stats (CSV)      | Intermediate file containing aggregated WPA m
etrics.              | clutch_stats_ranked.csv  |
| External Resource: Baseball Savant | https://baseballsavant.mlb.com/statcast_searc
h                    | External Link            |
```