USOpen2019

January 30, 2020

!!! NEED LINKS TO WEBSITE AND SEABORN THEME/PALLATTE !!!

1 US Open 2019 Analysis

May 17, 2019

This analysis aims to generate as comprehensive as possible summary of the event. Feel free to skip around and only view parts of the report that are interesting to you. Visualizations are presented in addtion to tables and narrations.

Table of Contents

Placements

Distributions

Trends

74 kilos 69 matches

Placements

Distributions

Trends

79 kilos 69 matches

Placements

Distributions

Trends

86 kilos 69 matches

Placements

Distributions

Trends

92 kilos 69 matches

Placements

Distributions

Trends

97 kilos 69 matches

Placements

Distributions

Trends

125 kilos 69 matches

Placements

Distributions

Trends

Comparison

Madness Matches

Match OEJV (Jesse Vasquez vs Brandon Wright)

Match QCY3 (Renaldo Rodriguez Spencer vs James Steerman)

Match WYEF (AJ Jaffe vs Drew Mandell)

Veritas Bites

Overall Event

High Value

Low Value

[5]:
```
%matplotlib inline
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import os
import django
os.chdir(os.getcwd())
```

[6]:
```
# read df
df = pd.read_csv('collection/stats/matchdata.csv')
df = df[df.Duration != 0.0]

# grouped by wrestler
name_grouped = df.groupby('Focus')

# grouped by weight
weight_grouped = df.groupby('Weight')

# filtered grouped by name (3 or more matches)
filt_name_grouped = df[df.Focus.isin(df['Focus']\
                                     .value_counts()[df['Focus'].value_counts()␣
 ↪>= 3].index)].groupby('Focus')
```

1.1 Individual Analysis

Top/bottom performances are mostly ranked on Action per Minute (APM), Neutral Pace Factor(NPF), and Results. Emphasis is placed on these metrics because APM and NPF have been shown to be highly correlated with long-term success (*r=0.92, r=0.88* respectively). The following section focuses on wrestlers' performances across the entire tournament and requires three or more matches to be included. Top performances for individual matches can be found further below.

1.1.1 Top Perfomers

Top performers for this event are (in order): 1. Justin Deangelis – 70kgs 2. Lavion Mayes – 70kgs 3. Yianni Diakomihalis – 65kgs

These wrestlers more than doubled the average APM for the event and competed for roughly 5 minutes (299 seconds) on average between them. They scored an average of 10.2 points and held their opponents to 5.8 for a 4.35 MoV. Interestingly, Justin and Yianni both had negative Passive differentials, but maintained a 4.2 and 6.6 MoV compared to Lavion's 2.3 MoV average.

```
[8]: print('Key perfomance metrics between all wrestlers and the top performers:')
top_perfs = filt_name_grouped.mean().sort_values(['APM', 'NPF', 'NumResult'],␣
 ↪ascending=False).head(3)

display(pd.concat([
    filt_name_grouped.mean()[['APM', 'NPF', 'NumResult']].mean().rename('All'),
    top_perfs.mean()[['APM', 'NPF', 'NumResult']].rename('Top')
    ], axis=1))

print('Additional metrics for Top Performers:')
display(top_perfs[['APM', 'NPF', 'NumResult', 'Duration', 'PassiveDiff',
                   'FocusPoints', 'OppPoints', 'MoV']])
```

```
Key perfomance metrics between all wrestlers and the top performers:

                 All        Top
APM        11.552341  22.200000
NPF         1.002481   1.212667
NumResult   0.911035   1.066667


Additional metrics for Top Performers:

                   APM    NPF  NumResult  Duration  PassiveDiff  FocusPoints  \
Focus
Justin Deangelis  23.0  1.296       1.18  269.0460        -0.60        10.00
Lavion Mayes      23.0  1.130       1.10  290.5375         0.25        10.75
Jesse Delgado     20.6  1.212       0.92  291.3860         0.00        11.20

                  OppPoints   MoV
Focus
Justin Deangelis        5.8  4.20
Lavion Mayes            8.5  2.25
Jesse Delgado          11.0  0.20
```

1.1.2 Bottom Performers

In this section it is important to remember that these wrestlers still competed three times, meaning they won at least one match and thus are **not** the *absolute* worst performers in the event. Bottom performers for this event are (in order): 1. Chance Goodman – 79kgs 2. Blake McNall – 65kgs 3. Fernando Serje – 92kgs

These wrestlers all averaged an APM of only 0.89 and a NPF score of 0.089. They were never turned and limited their opponents' NPF scores to only 1.02. All three of these wrestlers went 1-2 on the weekend with one Decision win and two Losses by Technical Superiority.

```
[10]: print('Key perfomance metrics between ALL wrestlers and the Bottom Performers:')
bottom_perfs = filt_name_grouped.mean().sort_values(['APM', 'NPF',␣
 ↪'NumResult']).head(3)
```

```
display(pd.concat([
    filt_name_grouped.mean()[['APM', 'NPF', 'NumResult']].mean().rename('All'),
    bottom_perfs.mean()[['APM', 'NPF', 'NumResult']].rename('Bottom')
    ], axis=1))

print('Additional metrics for Bottom Performers:')
display(bottom_perfs[['APM', 'NPF', 'NumResult', 'Duration', 'PassiveDiff',
                      'FocusPoints', 'OppPoints', 'MoV']])
```

```
Key perfomance metrics between ALL wrestlers and the Bottom Performers:

                 All    Bottom
APM        11.552341  3.666667
NPF         1.002481  0.794444
NumResult   0.911035  0.633333


Additional metrics for Bottom Performers:

                       APM       NPF  NumResult    Duration  PassiveDiff  \
Focus
Patrick Romero    3.333333  0.916667   0.583333  113.080000    -0.666667
Anthony Ashnault  3.666667  0.966667   0.733333  250.943333     0.666667
Kyle Kintz        4.000000  0.500000   0.583333   84.226667     0.000000

                  FocusPoints  OppPoints       MoV
Focus
Patrick Romero            1.0   7.666667 -6.666667
Anthony Ashnault          2.0   5.000000 -3.000000
Kyle Kintz                2.0   7.333333 -5.333333
```

1.1.3 Top Individual Match Performance

These sections focus on individual matches for all wrestlers who competed in the event.
Analysis will still favor the KPMs (Key Performance Metrics: APM and NPF).

There were 100 matches that resulted in a Win by Fall.
There were 88 matches with the maximum NPF of 2.0.
Of these, 11 matches resulted in an NPF of 2.0 **and** a Win by Fall.
There were 2 matches with an APM greater than 3 standard deviations from the event mean of 10.32.

[11]:
```
print('None of the 2 matches with an APM > 31 resulted in a Fall...')
display(df[df['APM'] > df['APM'].mean() * 3]\
        [['Result', 'MatchID', 'Focus', 'Opponent',
          'NPF', 'APM', 'Duration', 'FocusPoints', 'OppPoints']])
```

```
None of the 2 matches with an APM > 31 resulted in a Fall…
```

```
     Result MatchID             Focus         Opponent   NPF   APM  Duration  \
190    WinD    8020       Tyler Graff      Cody Brewer  1.15  40.0    360.04
350   WinTF    IFHT     Riley Lefever  Enock Francois  1.16  35.0    333.23
1011   WinD   FJ8O*  Justin Deangelis     Lavion Mayes  1.22  34.0    363.37

      FocusPoints  OppPoints
190            22         16
350            23         12
1011           16         11
```

[12]:

```python
print('11 matches with a WinF result and 2.0 NPF.')
display(df[(df['NPF'] == 2.0) & (df['Result'] == 'WinF')]\
        [['Result', 'MatchID', 'Focus', 'Opponent',
          'NPF', 'APM', 'Duration', 'FocusPoints', 'OppPoints']].
 ↪sort_values(['APM'], ascending=False))
```

```
11 matches with a WinF result and 2.0 NPF.

     Result MatchID                  Focus             Opponent  NPF   APM  \
843    WinF   HTMU*           Elroy Perkin       Joshua Kindig  2.0  12.0
669    WinF   DXNJ*            Conrad Cole      Ernesto Garcia  2.0  11.0
1598   WinF   PPH7*  Michael Macchiavello      Timothy Dudley  2.0  10.0
763    WinF    V8TQ           Joshua Asper   David Richardson  2.0   9.0
1400   WinF    SY70         Matthew Malcom     Austin Farbaugh  2.0   8.0
544    WinF   U72K*         Cameron Caffey          Jake Posey  2.0   7.0
955    WinF    QW3W           Geno Morelli      Kyle Matthews  2.0   6.0
524    WinF   DQW8*            Nikko Reyes  Nicholas Nottingham  2.0   5.0
652    WinF    KY05         Nathan Jackson        Ethan Vistro  2.0   5.0
997    WinF   M8Q4*            Noel Orozco       Kyle La Fritz  2.0   5.0
1165   WinF   5N6D*        Jarrod Hinrichs       Kyle Jennings  2.0   5.0

      Duration  FocusPoints  OppPoints
843      95.74            8          0
669      54.80            8          0
1598    263.87            9          0
763      36.90            8          0
1400     67.43            6          0
544      88.74            4          0
955      17.42            4          0
524      30.53            4          0
652      12.70            4          0
997      18.71            4          0
1165     21.41            4          0
```

From these two lists of matches it is clear that there are many great performances to choose from. The Perkin WinF over Kindig is tempting as it has the highest APM of any match with a 2.0 NPF and WinF result. However, the Cole WinF over Garcia seems a stronger contender as it has roughly the same APM and was finished in half the time (55 seconds versus 96 seconds).

Ultimately I will drop both of these options in favor of the 40.0 APM Graff WinD over Brewer. This match was a 38 point storm that resulted in the highest APM of the tournament despite a relateively even NPF (1.15 for Graff).
> Caveat: The second highest APM of 30.0 in the Glynn WinTF over Hessler was tempting as most metrics are the same (another 30 point match) or better (1.24 NPF for Glynn), but you don't get points for second place.

1.2 Weight Class Analysis

[13]:
```
base_cols = ['BinaryResult',
             'HIa', 'HIc2', 'HIc4', 'HOa', 'HOc2', 'HOc4', 'Da', 'Dc2', 'Dc4',
             'LSa', 'LSc2', 'LSc4', 'GBa', 'GBc2', 'Ta', 'Tc2', 'Tc4',
             'Exposure', 'Gut', 'LegLace',  'Turn', 'Pushout', 'Violation',␣
 ↪'Passive',
             'oHIa', 'oHIc2', 'oHIc4', 'oHOa', 'oHOc2', 'oHOc4', 'oDa', 'oDc2',␣
 ↪'oDc4',
             'oLSa', 'oLSc2', 'oLSc4', 'oGBa', 'oGBc2', 'oTa', 'oTc2', 'oTc4',
             'oExposure', 'oGut', 'oLegLace',  'oTurn', 'oPushout',␣
 ↪'oViolation', 'oPassive',
            ]
```

1.2.1 57 kilos 69 matches

Placements

1. Daton Fix
2. Thomas Gilman
3. Zane Richards
4. Vitali Arujau
5. Darian Cruz
6. Nathan Tomasello
7. Frank Perrelli
8. Zach Sanders

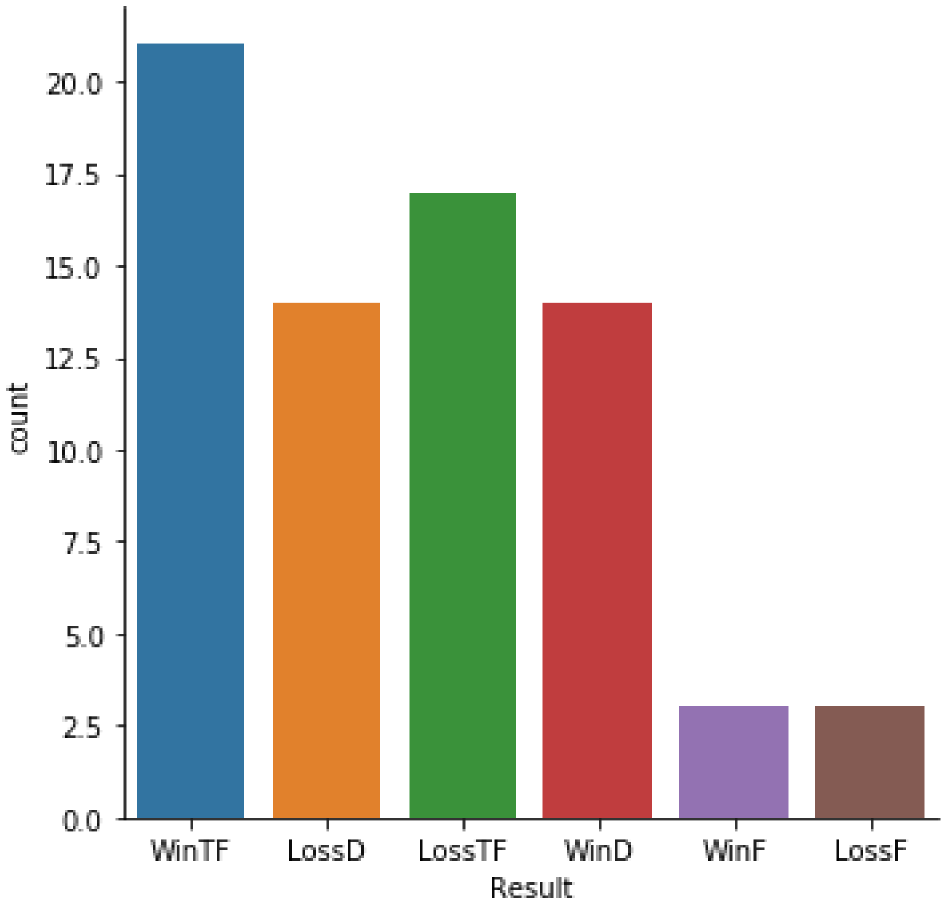
[20]:
```
kgs57 = df[df['Weight'] == 57]
```

Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

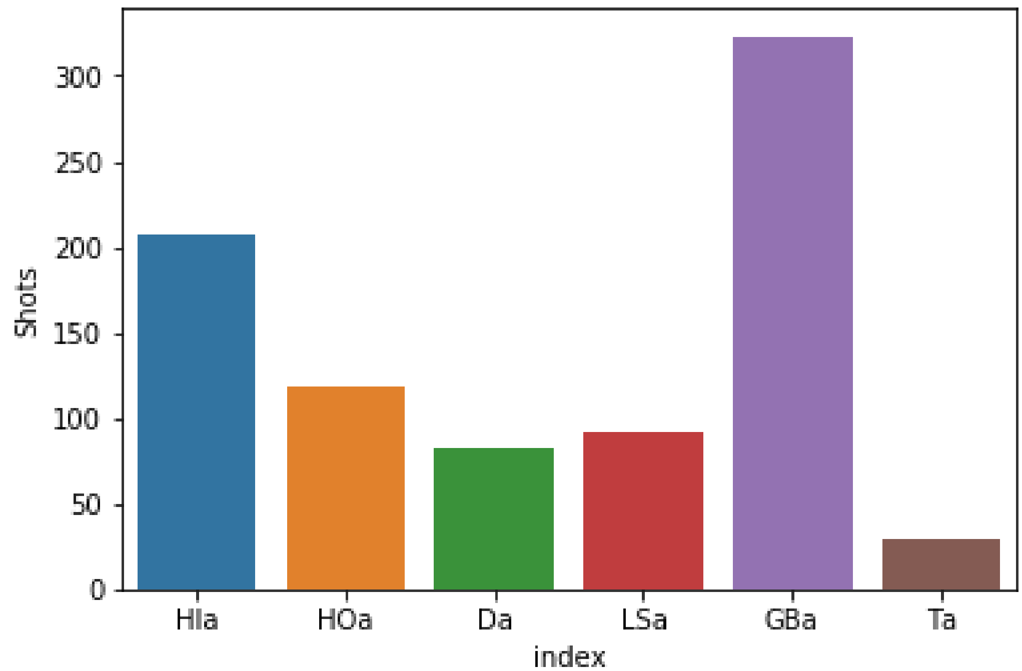
[21]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs57[kgs57['MatchID'].map(len) == 4], kind="count")
plt.show()

# somehow order so can use palatte for win/loss
```

```
Distributions of match results.
```

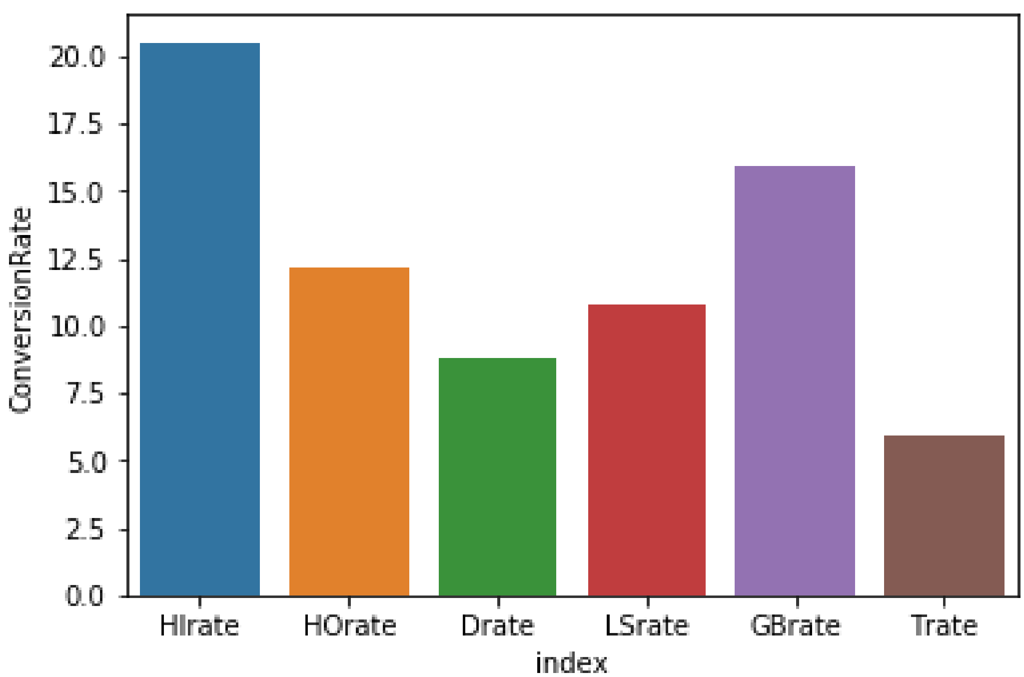


GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[22]:
```
sns.barplot(x='index', y='Shots',
            data=kgs57.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs57.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
 ↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



Shots
300
250
200
150
100
50
0
HIa
HOa
Da
LSa
GBa
Ta
index



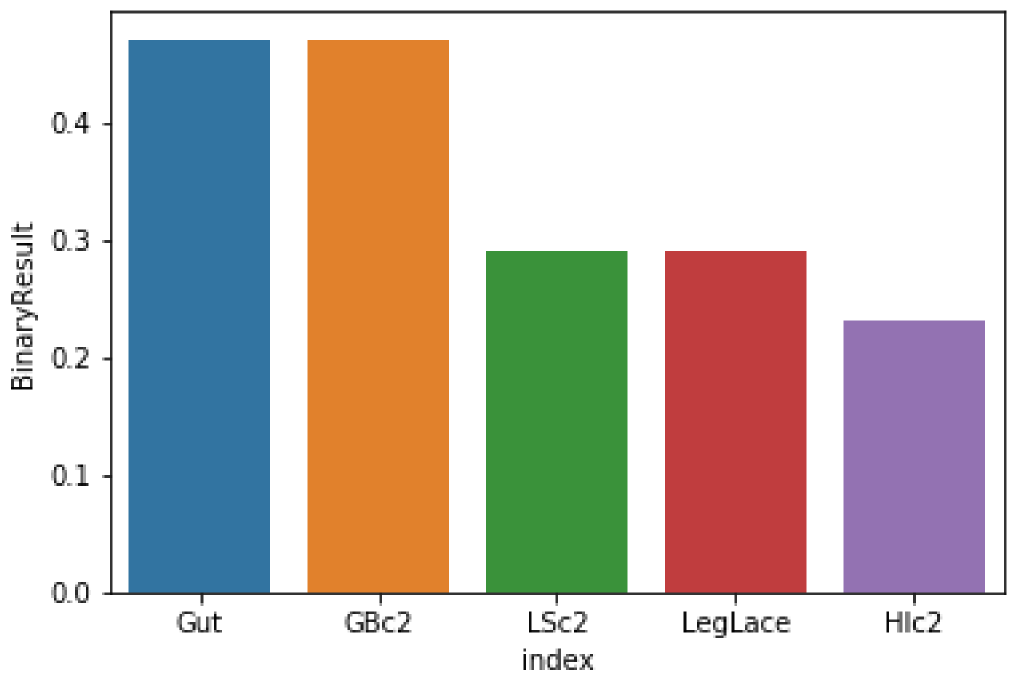
ConversionRate
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
HIrate
HOrate
Drate
LSrate
GBrate
Trate
index

Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[23]:
```
print('Top actions for 57kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs57[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 57kgs at this event.
```



1.2.2 61 kilos 61 matches

Placements

1. Cody Brewer
2. Nico Megaludis
3. Joey Palmer
4. Tyler Graff

5. Cory Clark
6. Anthony Ramos
7. Earl Hall
8. Beau Bartlett

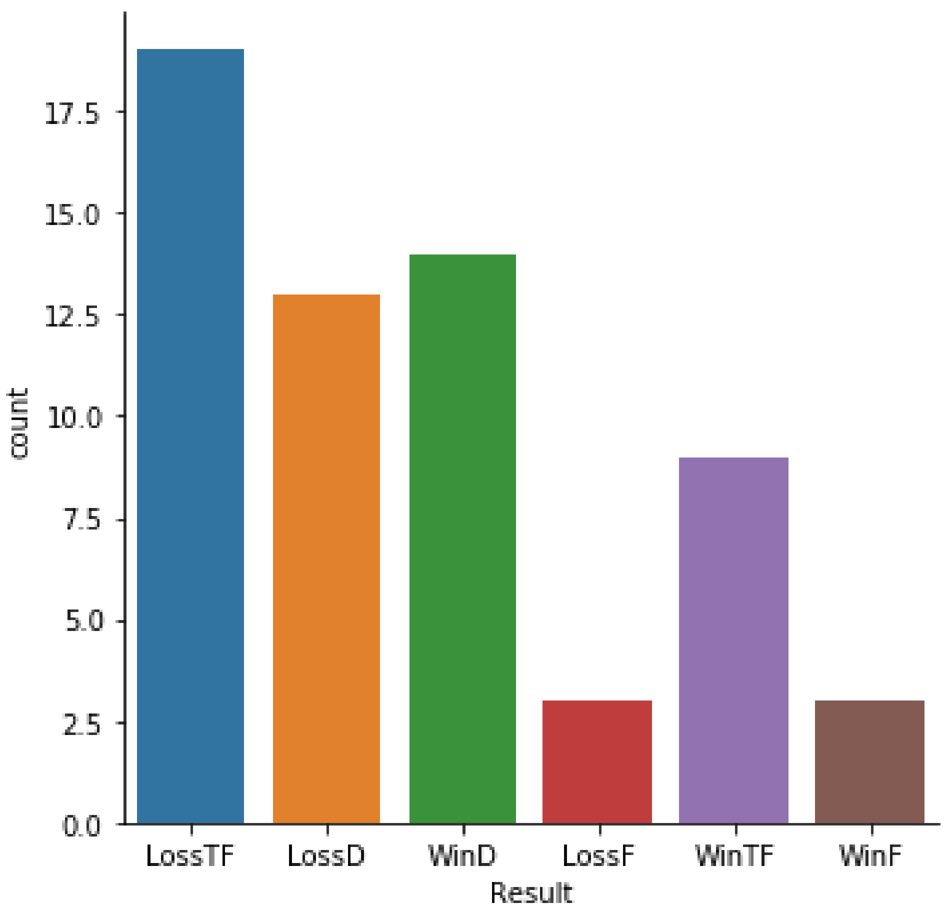
[684]:
```
kgs61 = df[df['Weight'] == 61]
```

Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

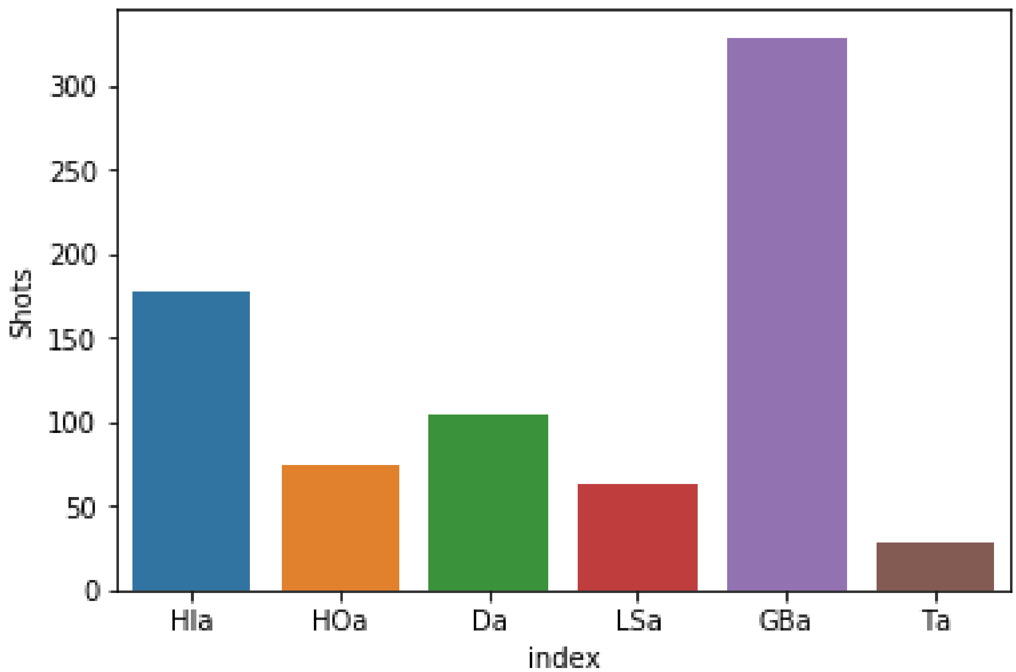
[691]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs61[kgs61['MatchID'].map(len) == 4], kind="count")
plt.show()

# somehow order so can use palatte for win/loss
```

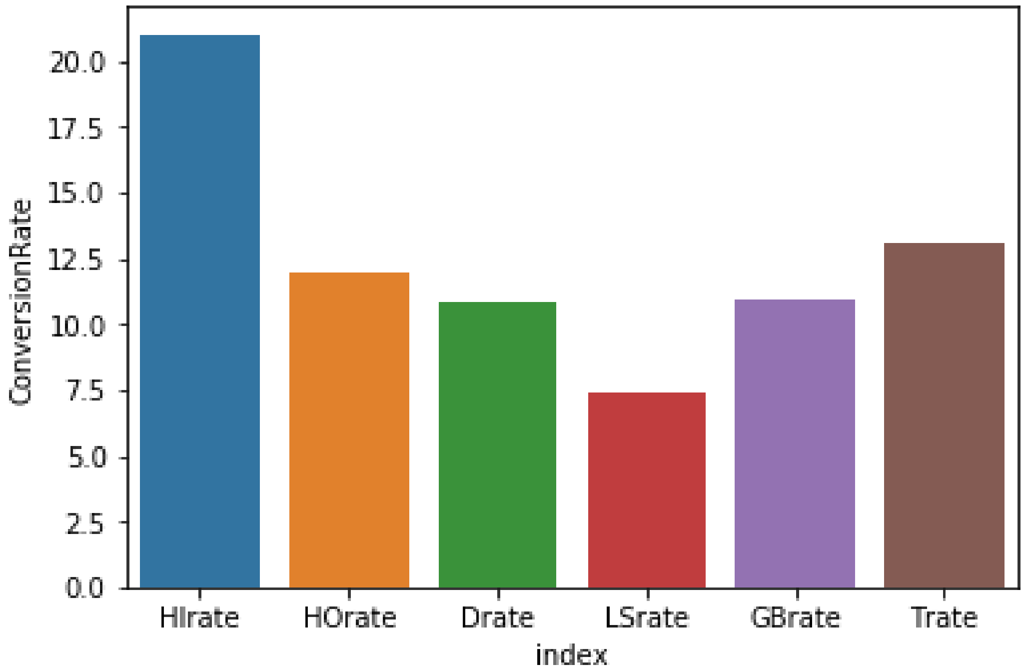
```
Distributions of match results.
```



GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[692]:
```
sns.barplot(x='index', y='Shots',
            data=kgs61.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs61.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
 ↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



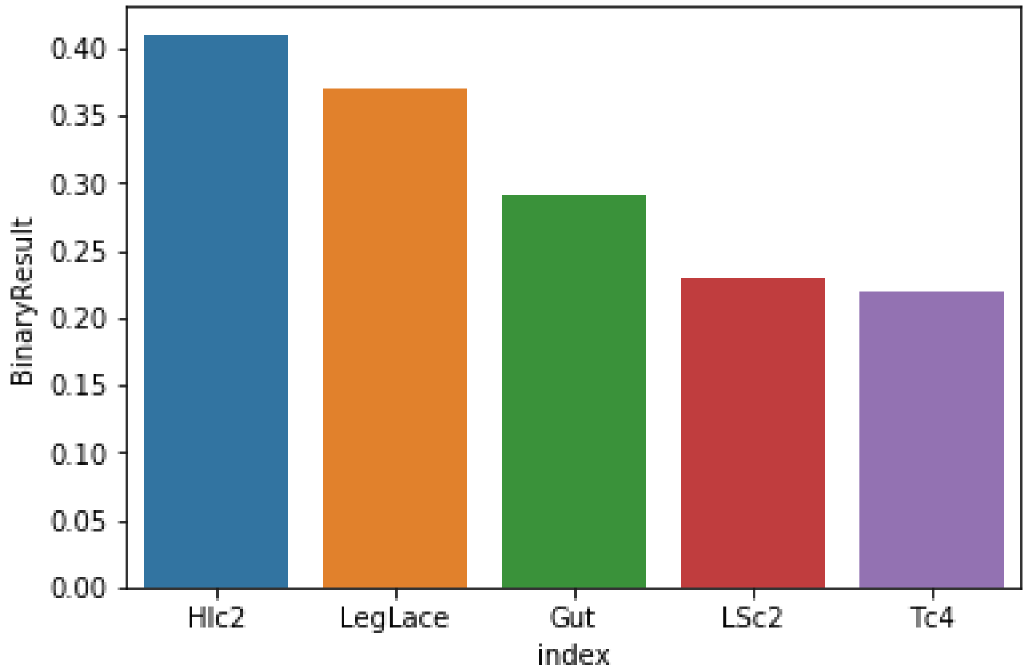


Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[693]:
```
print('Top actions for 61kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs61[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 61kgs at this event.
```



1.2.3 65 kilos 92 matches

Placements

1. Yianni Diakomihalis
2. Zain Retherford
3. Jason Ness
4. Frank Molinaro
5. Jaydin Eirman
6. Jordan Oliver
7. Kanen Storr
8. Bernard Futrell

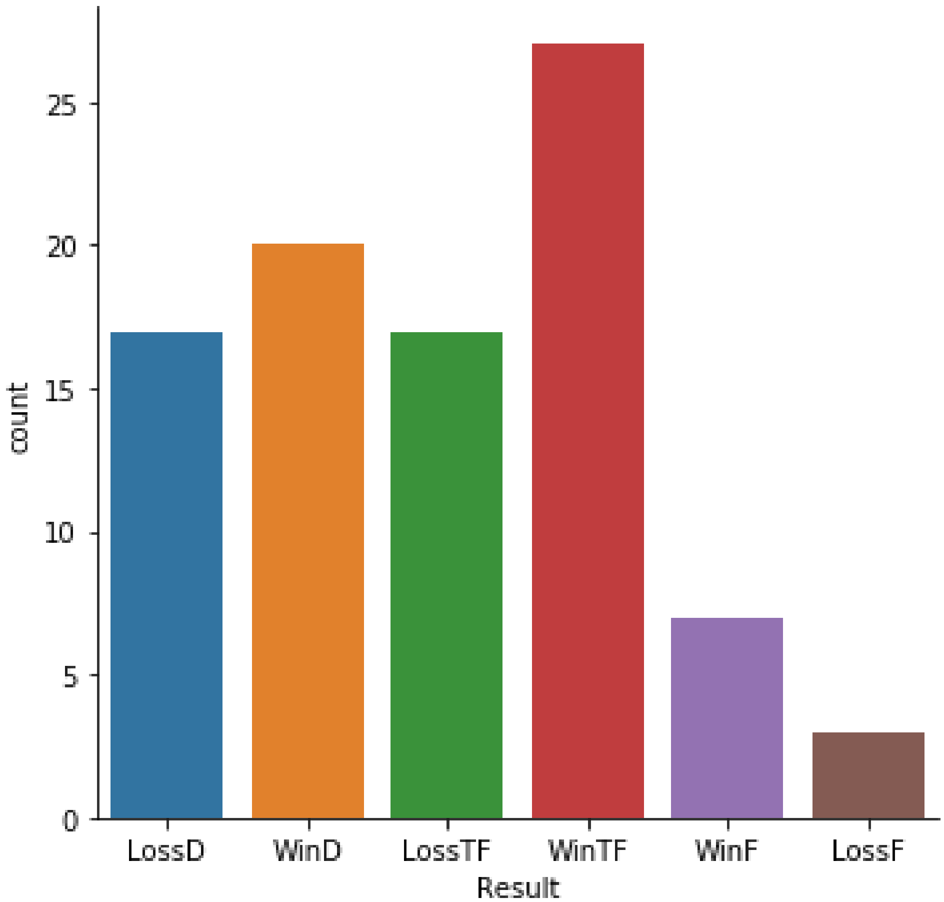
[696]:
```
kgs65 = df[df['Weight'] == 65]
```

Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

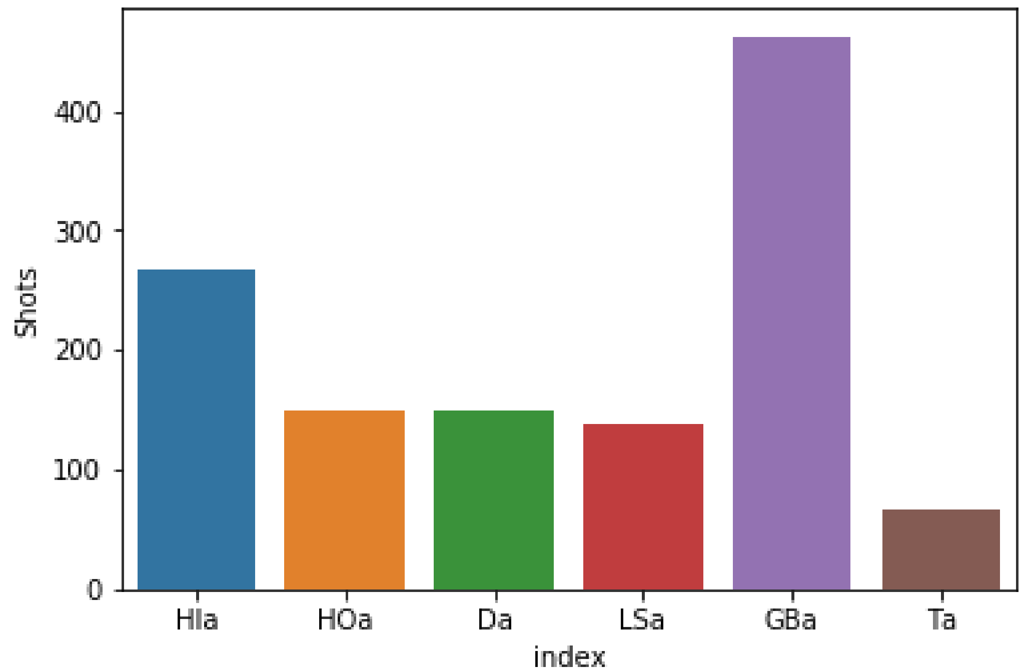
[697]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs65[kgs65['MatchID'].map(len) == 4], kind="count")
plt.show()

# somehow order so can use palatte for win/loss
```

```
Distributions of match results.
```

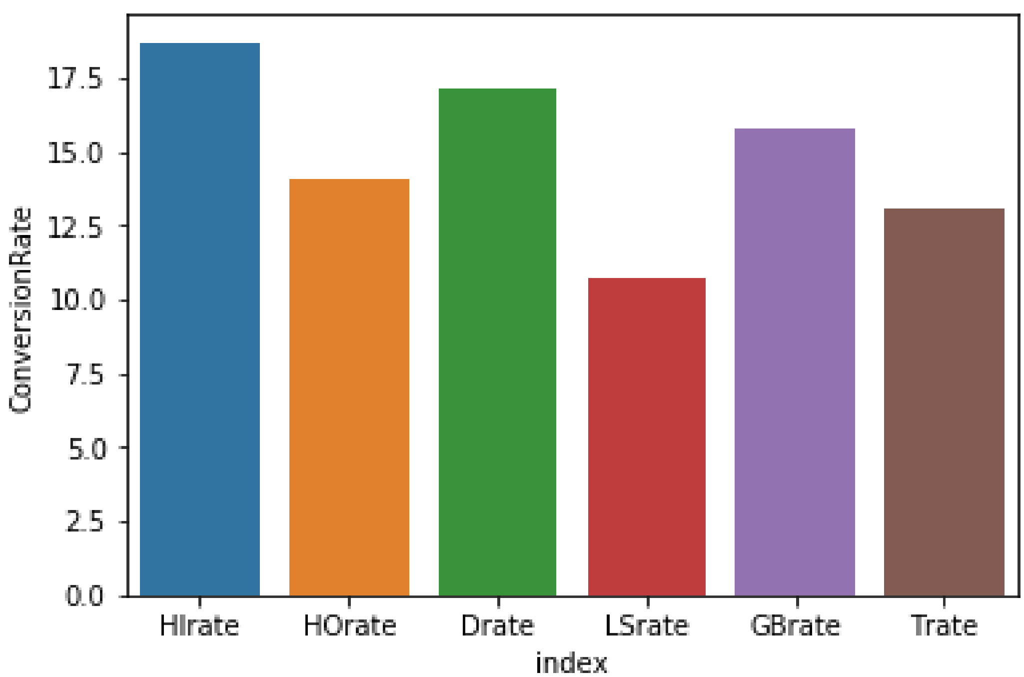


GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[698]:
```
sns.barplot(x='index', y='Shots',
            data=kgs65.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs65.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
 ↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



Shots
400
300
200
100
0
HIa
HOa
Da
LSa
GBa
Ta
index



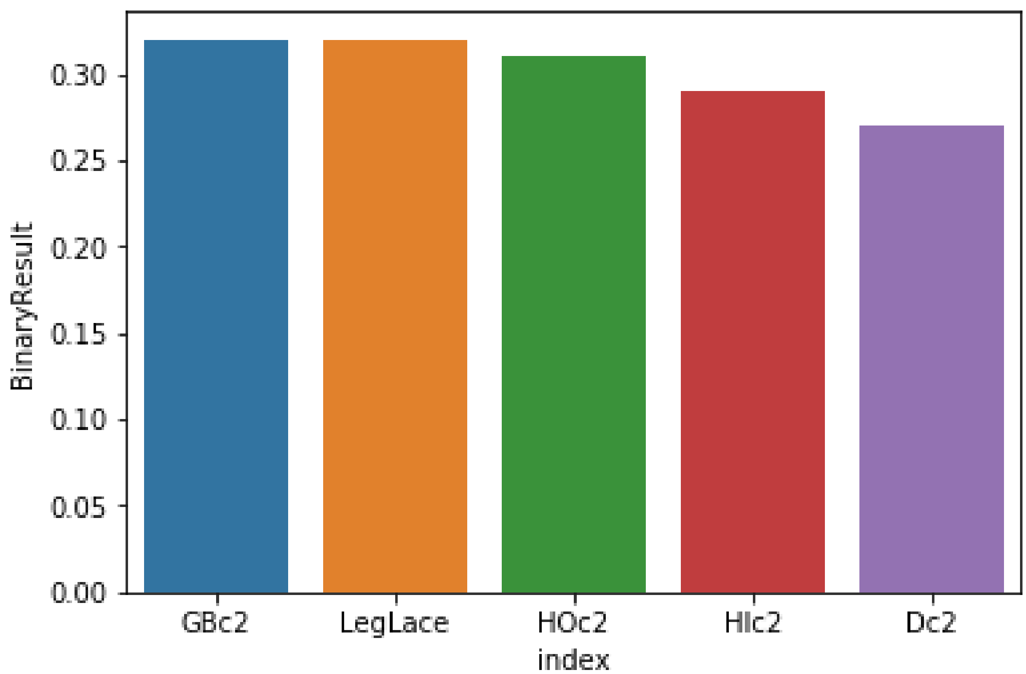
ConversionRate
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
HIrate
HOrate
Drate
LSrate
GBrate
Trate
index

Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[699]:
```
print('Top actions for 65kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs65[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 65kgs at this event.
```



1.2.4 70 kilos 96 matches

Placements

1. Ryan Deakin
2. James Green
3. Jason Nolf
4. Alec Pantaleo

5. Brandon Sorensen
6. Nazar Kulchytskyy

7. Jason Chamberlain
8. Anthony Collica

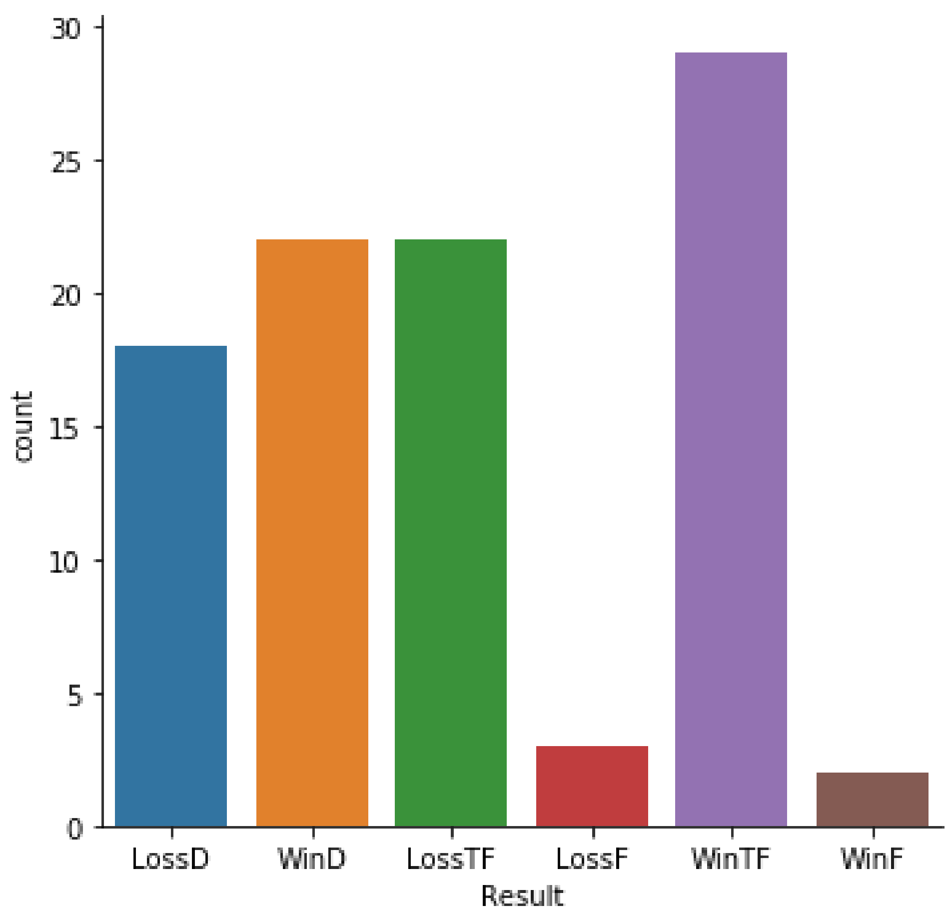
[702]:
```
kgs70 = df[df['Weight'] == 70]
```

Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

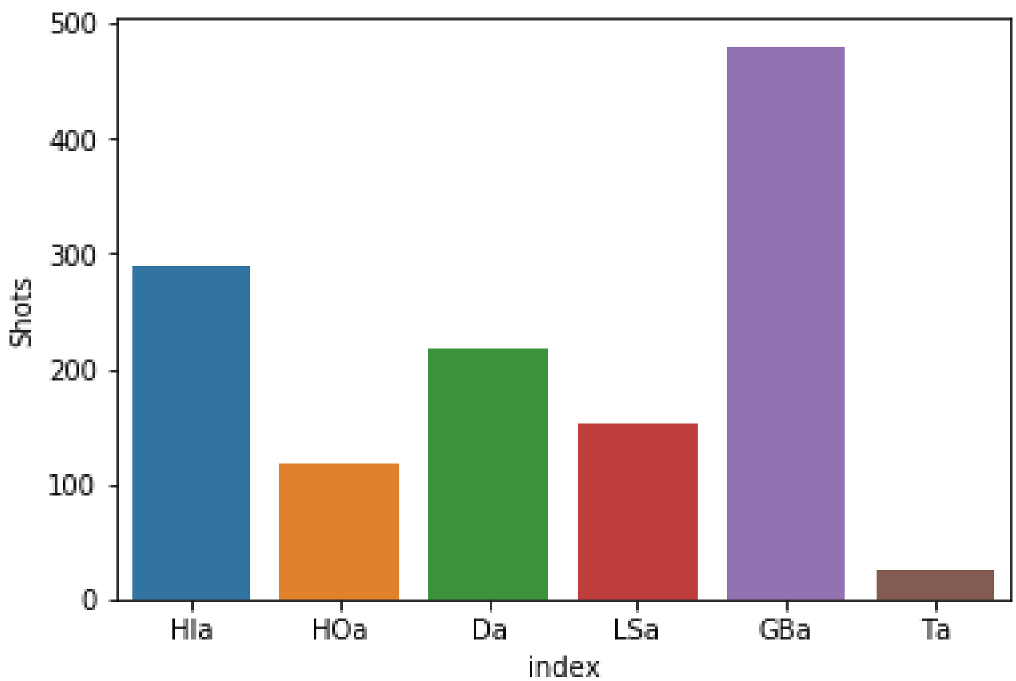
[703]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs70[kgs70['MatchID'].map(len) == 4], kind="count")
plt.show()

# somehow order so can use palatte for win/loss
```

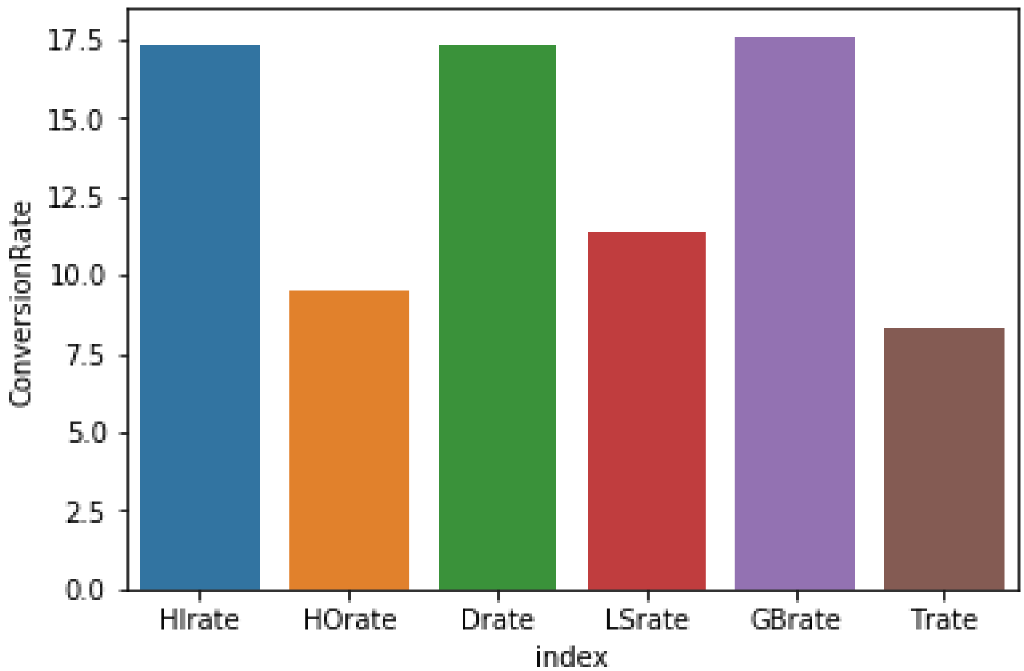
```
Distributions of match results.
```



GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[704]:
```
sns.barplot(x='index', y='Shots',
            data=kgs70.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs70.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
 ↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



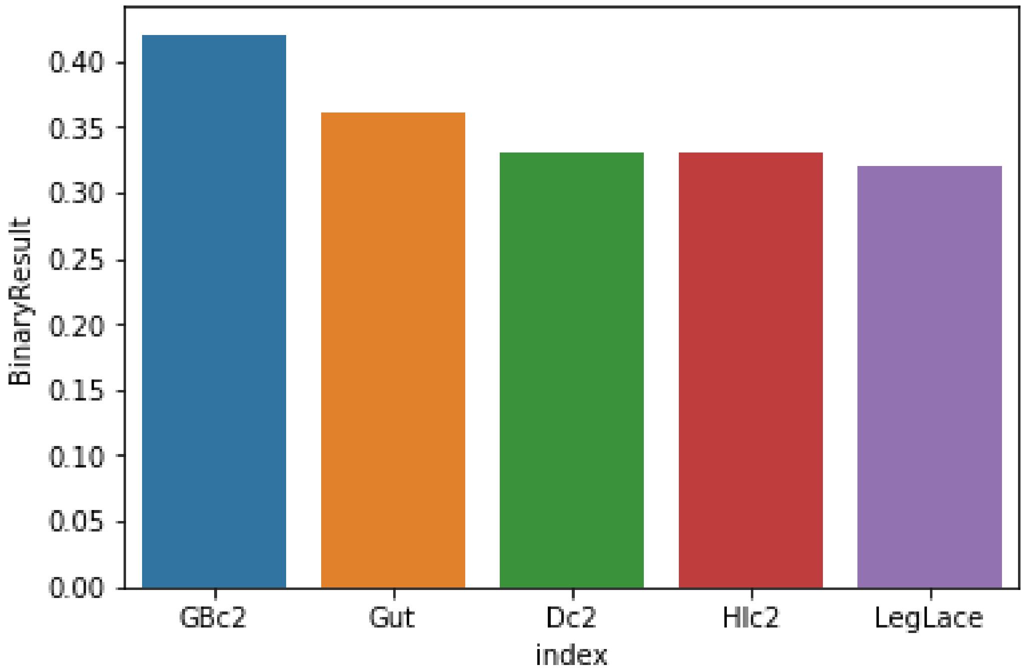


Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[705]:
```
print('Top actions for 70kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs70[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 70kgs at this event.
```



1.2.5 74 kilos 69 matches

Placements

1. Daton Fix
2. Thomas Gilman
3. Zane Richards
4. Vitali Arujau
5. Darian Cruz
6. Nathan Tomasello
7. Frank Perrelli
8. Zach Sanders

[321]:
```
kgs57 = df[df['Weight'] == 57]
```

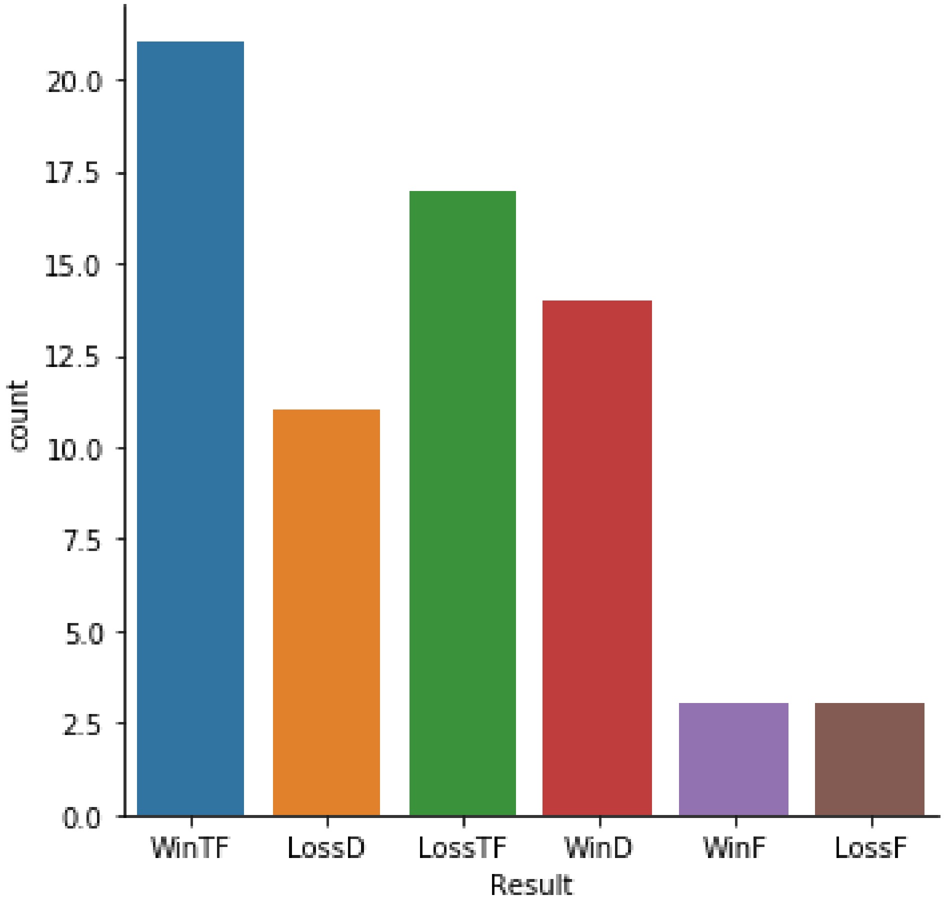
Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

[499]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs57[kgs57['MatchID'].map(len) == 4], kind="count")
g.set_xticklabels(rotation=30)
plt.show()
```

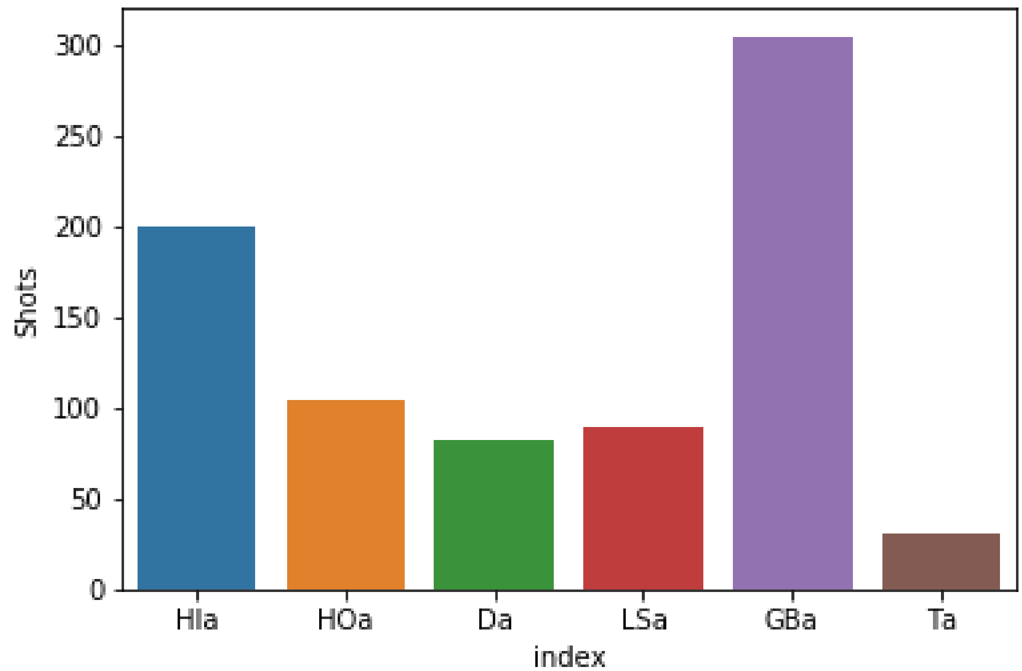
```
# somehow order so can use palatte for win/loss
```

```
Distributions of match results.
```

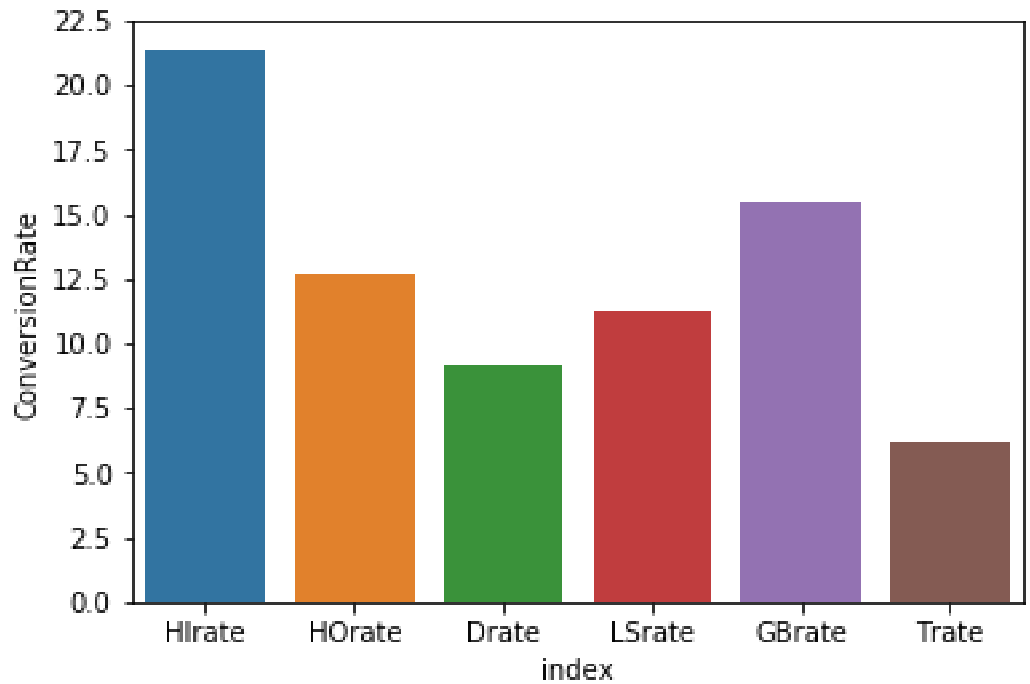


GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[405]:
```
sns.barplot(x='index', y='Shots',
            data=kgs57.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs57.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
 ↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



Shots
300
250
200
150
100
50
0
HIa
HOa
Da
LSa
GBa
Ta
index



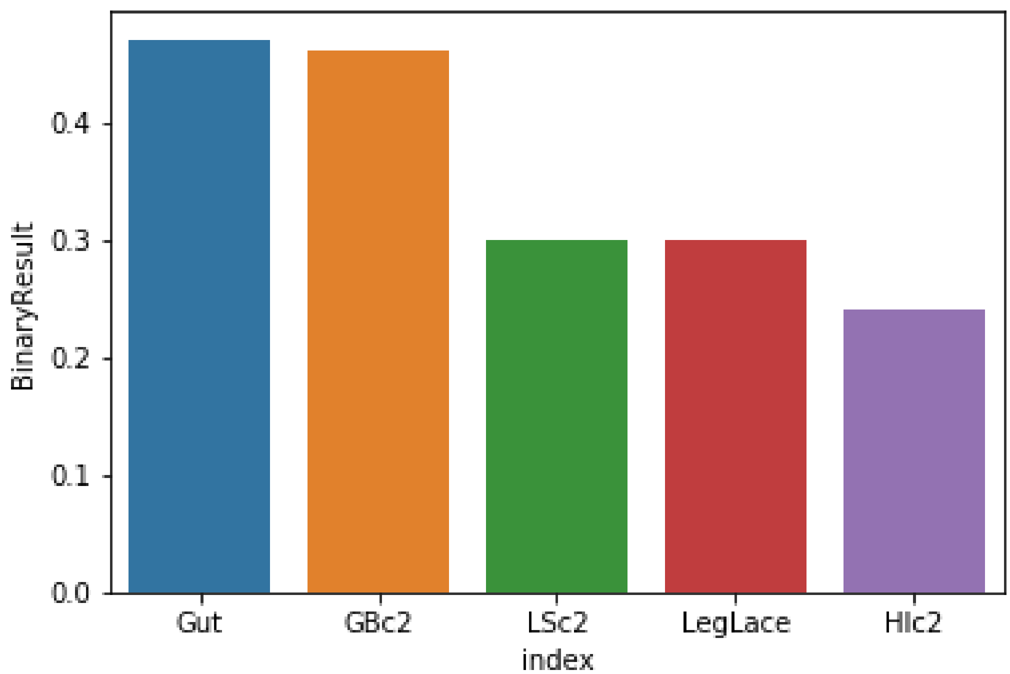
ConversionRate
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
HIrate
HOrate
Drate
LSrate
GBrate
Trate
index

Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[627]:
```
print('Top actions for 57kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs57[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 57kgs at this event.
```



1.2.6 79 kilos 69 matches

Placements

1. Daton Fix
2. Thomas Gilman
3. Zane Richards
4. Vitali Arujau

5. Darian Cruz
6. Nathan Tomasello
7. Frank Perrelli
8. Zach Sanders

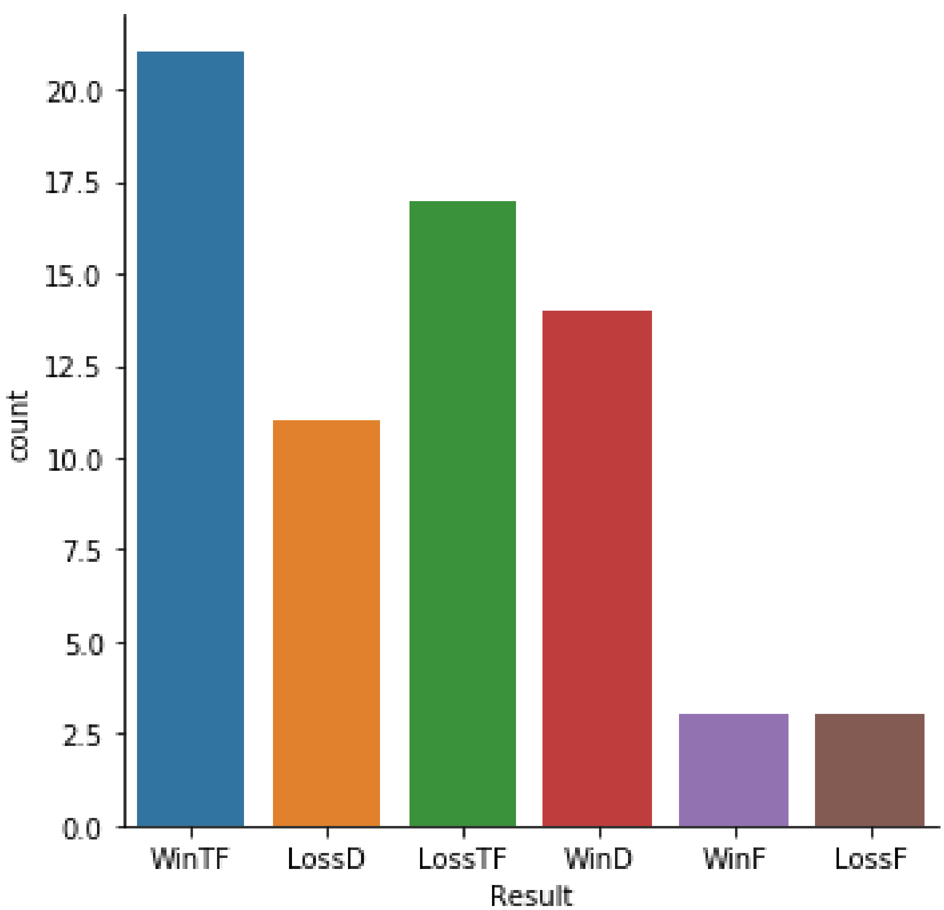
[321]:
```
kgs57 = df[df['Weight'] == 57]
```

Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

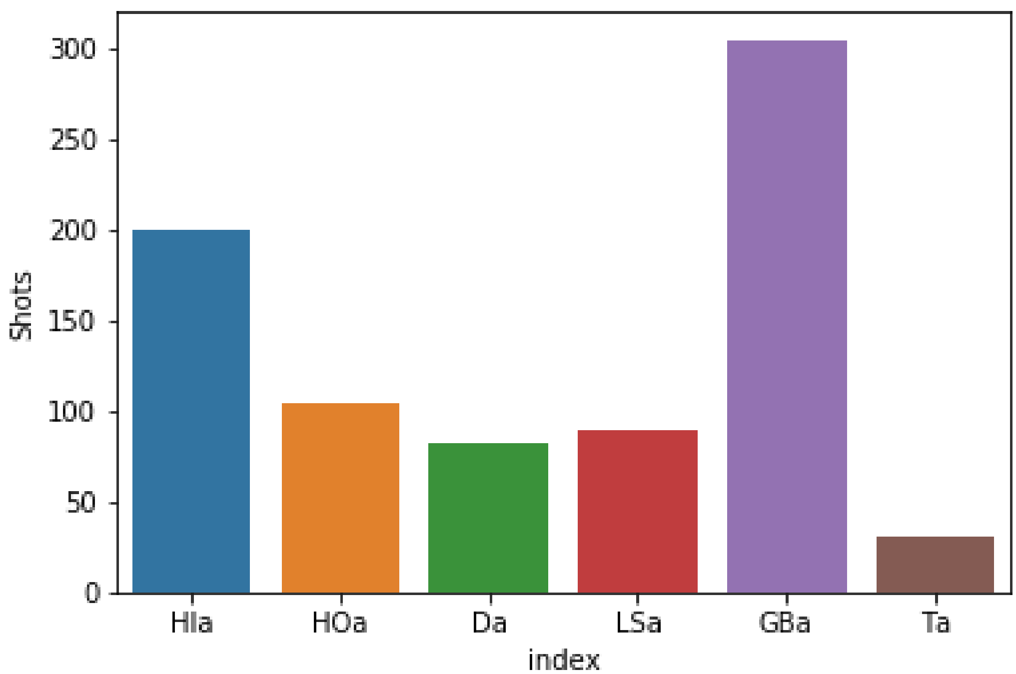
[499]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs57[kgs57['MatchID'].map(len) == 4], kind="count")
g.set_xticklabels(rotation=30)
plt.show()

# somehow order so can use palatte for win/loss
```

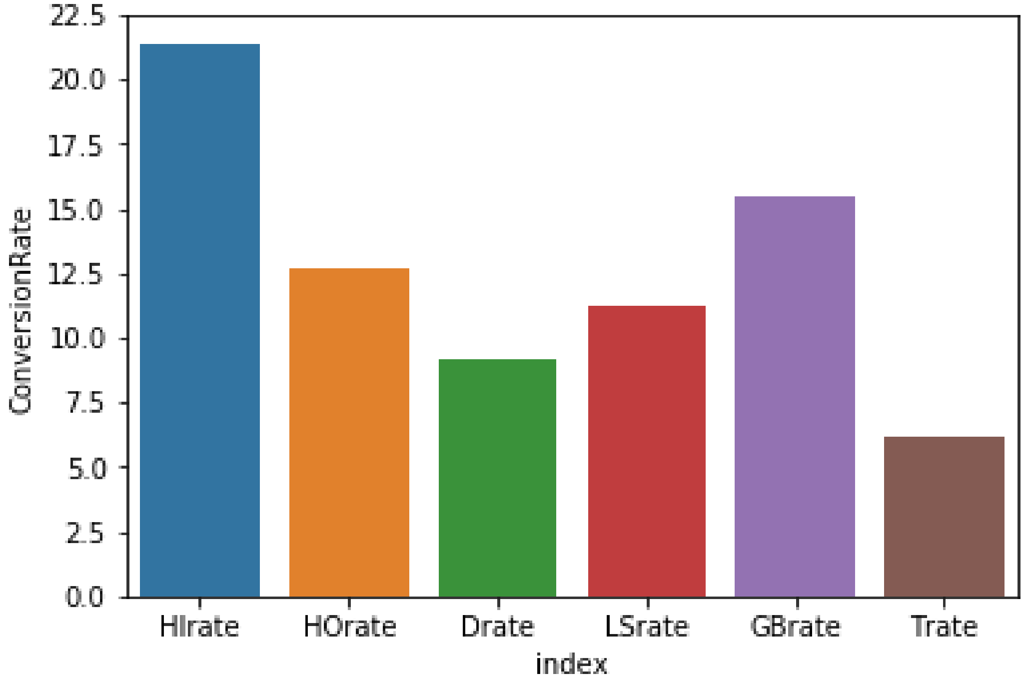
```
Distributions of match results.
```



GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[405]:
```
sns.barplot(x='index', y='Shots',
            data=kgs57.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs57.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



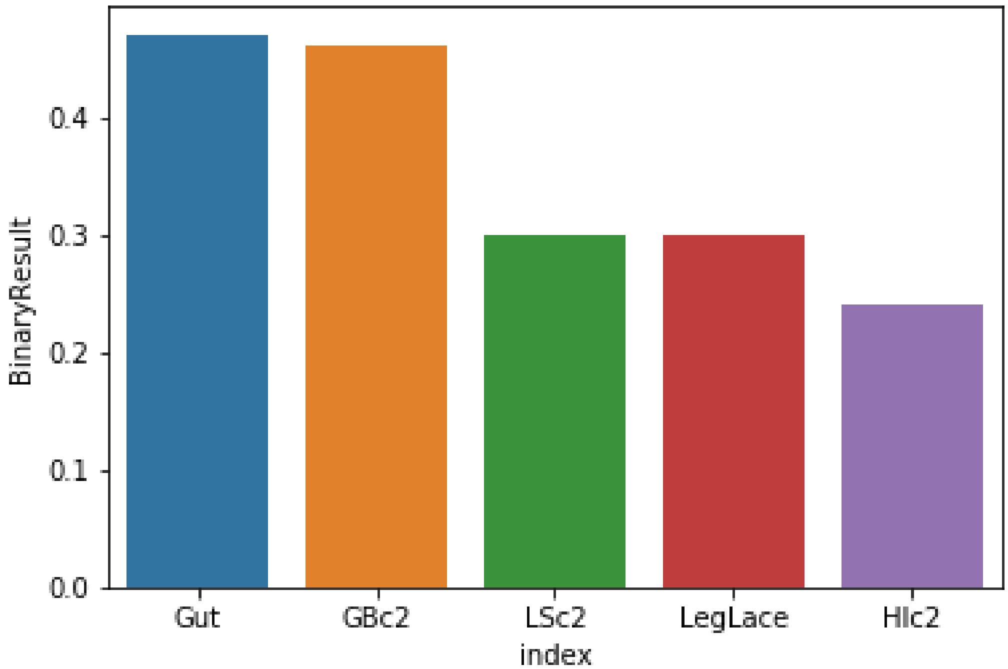


Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[627]:
```
print('Top actions for 57kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs57[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 57kgs at this event.
```



1.2.7 86 kilos 69 matches

Placements

1. Daton Fix
2. Thomas Gilman
3. Zane Richards
4. Vitali Arujau
5. Darian Cruz
6. Nathan Tomasello
7. Frank Perrelli
8. Zach Sanders

[321]:
```
kgs57 = df[df['Weight'] == 57]
```

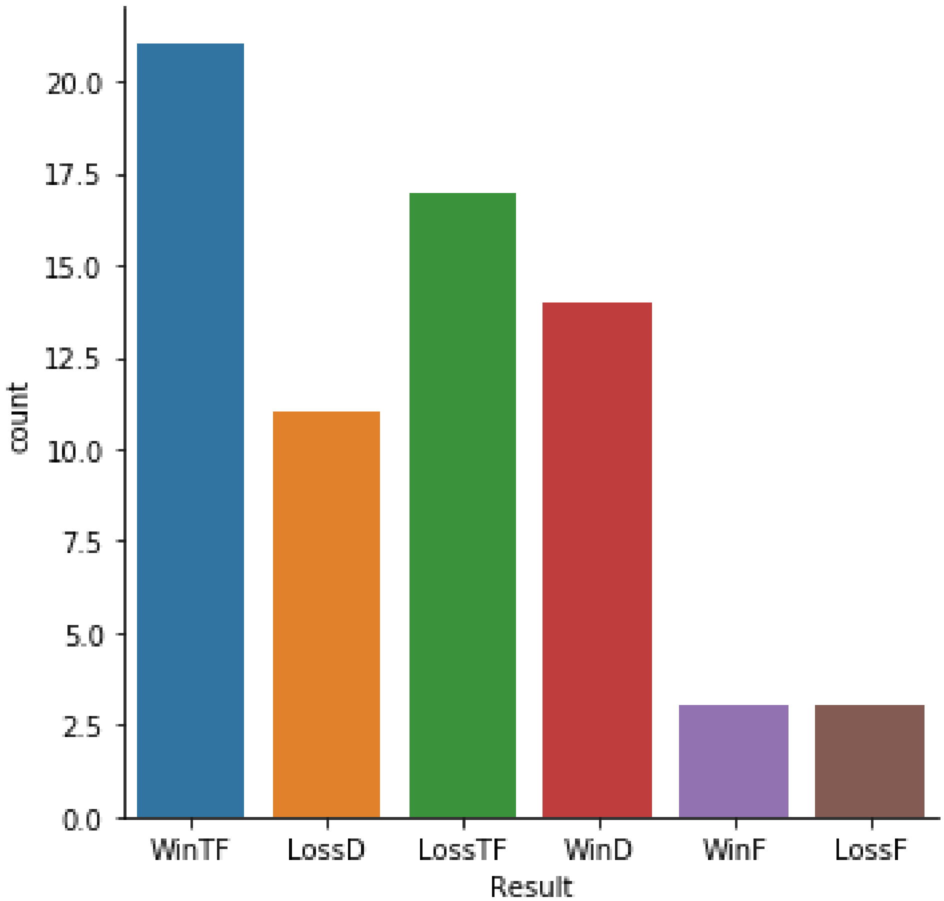
Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

[499]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs57[kgs57['MatchID'].map(len) == 4], kind="count")
g.set_xticklabels(rotation=30)
plt.show()
```

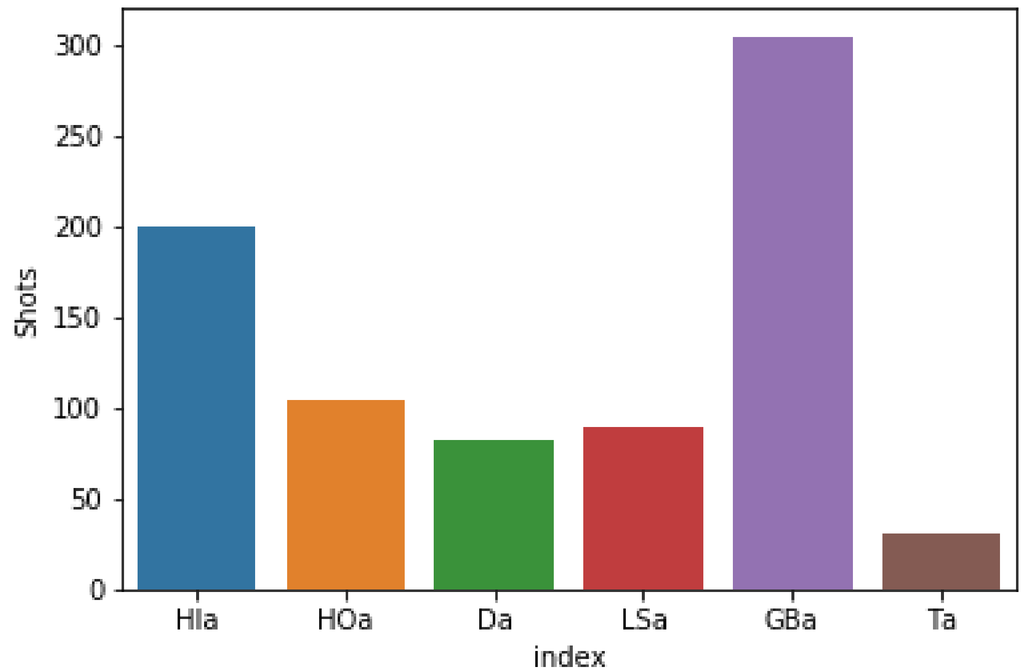
```
# somehow order so can use palatte for win/loss
```

Distributions of match results.

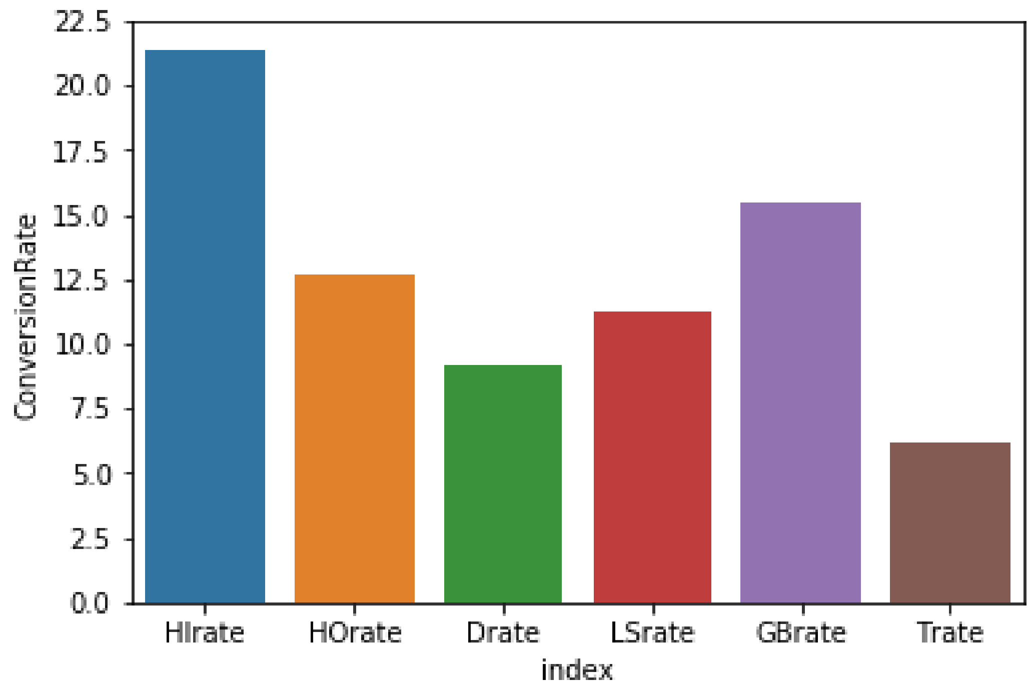


GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[405]:
```
sns.barplot(x='index', y='Shots',
            data=kgs57.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs57.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



300
250
200
150
100
50
0
Shots
HIa
HOa
Da
LSa
GBa
Ta
index



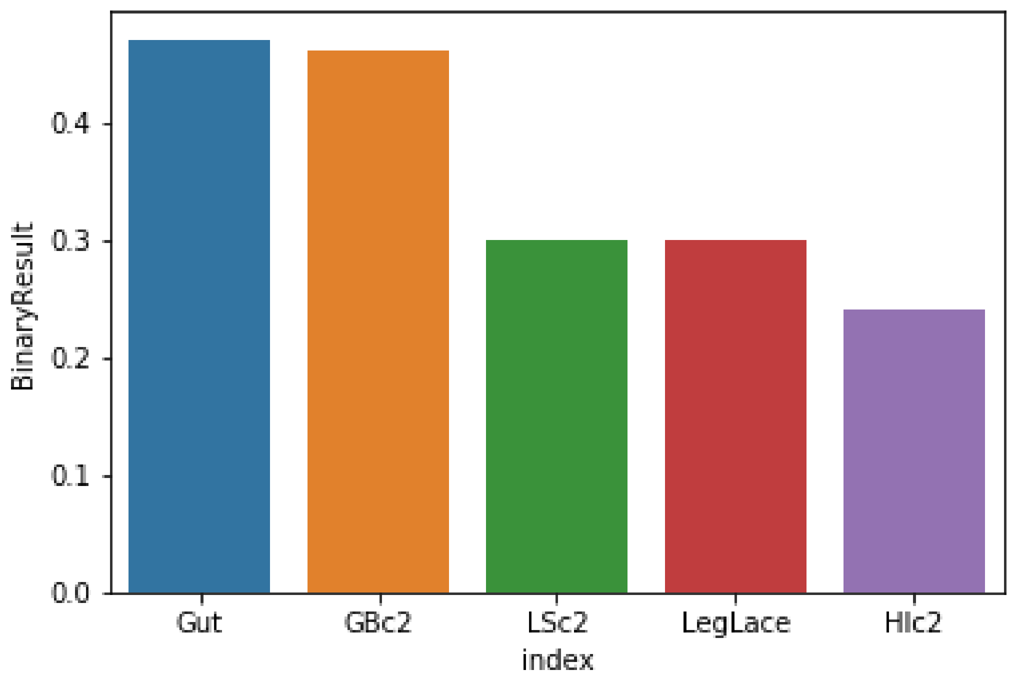
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
ConversionRate
HIrate
HOrate
Drate
LSrate
GBrate
Trate
index

Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[627]:
```
print('Top actions for 57kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs57[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 57kgs at this event.
```



1.2.8 92 kilos 69 matches

Placements

1. Daton Fix
2. Thomas Gilman
3. Zane Richards
4. Vitali Arujau

5. Darian Cruz
6. Nathan Tomasello
7. Frank Perrelli
8. Zach Sanders

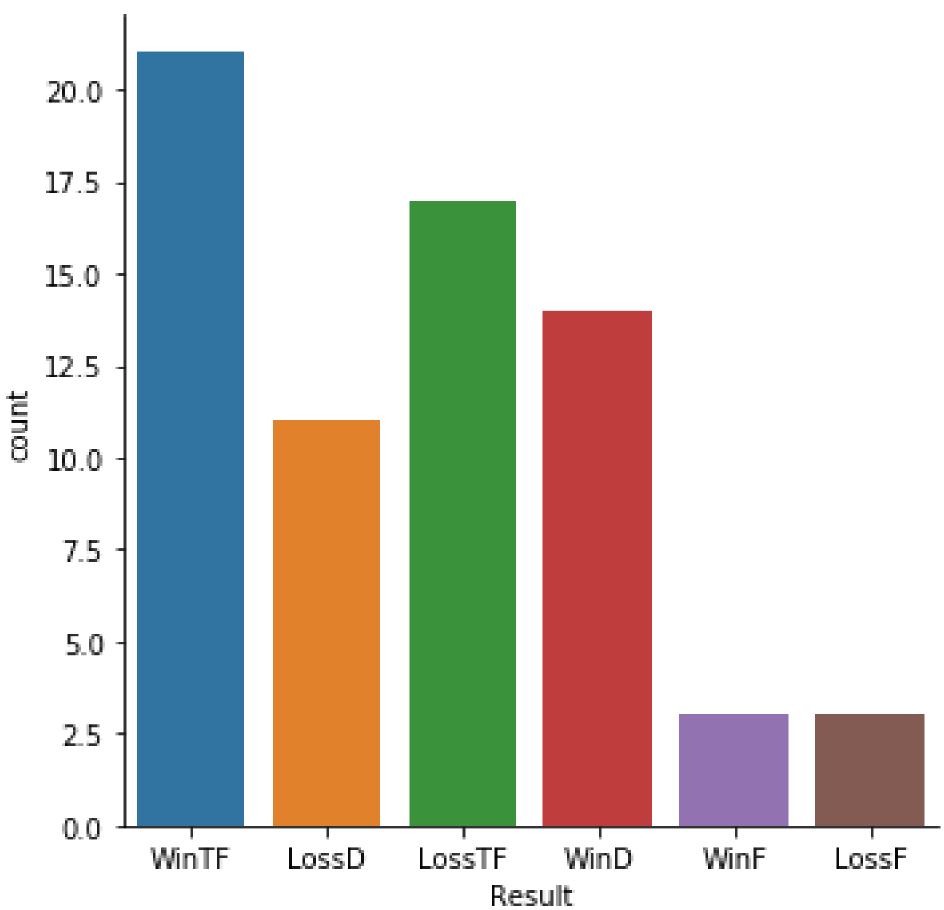
[321]:
```
kgs57 = df[df['Weight'] == 57]
```

Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

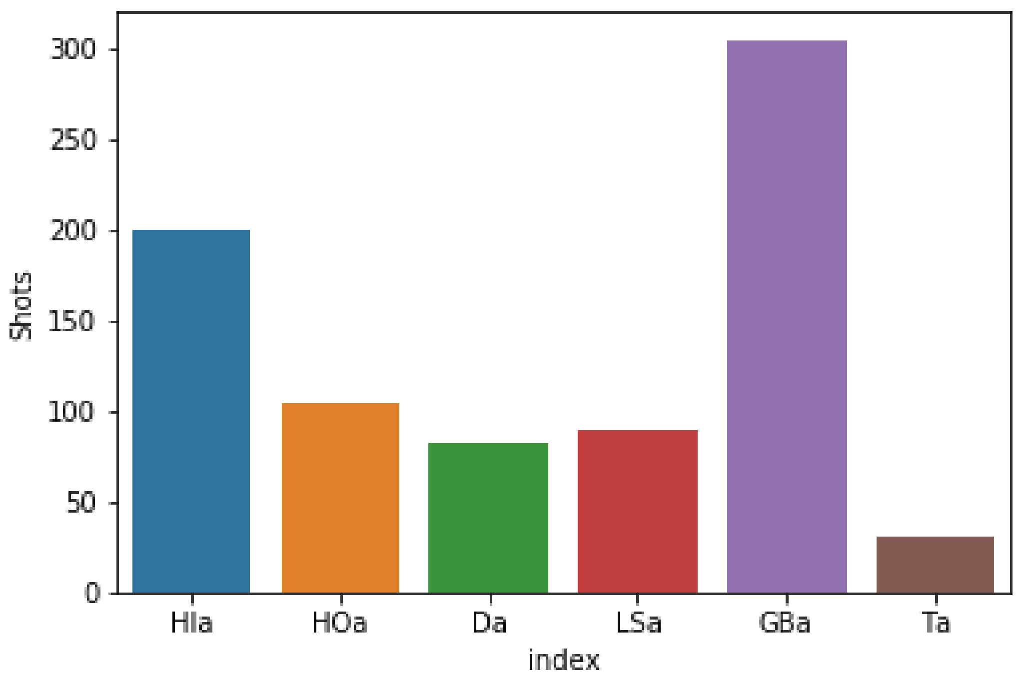
[499]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs57[kgs57['MatchID'].map(len) == 4], kind="count")
g.set_xticklabels(rotation=30)
plt.show()

# somehow order so can use palatte for win/loss
```

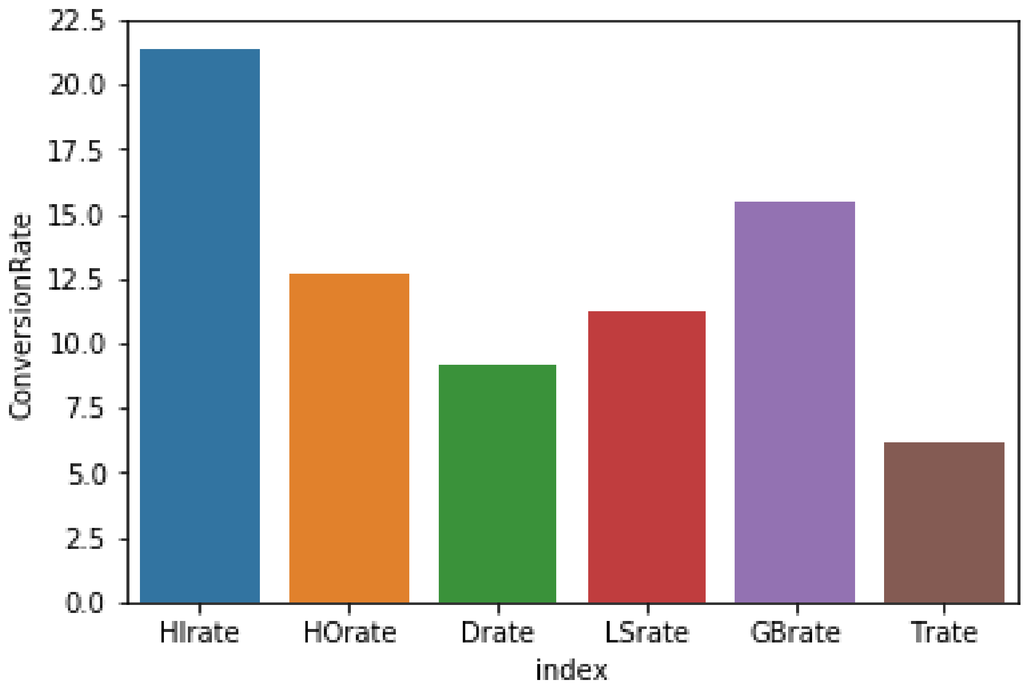
```
Distributions of match results.
```



GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[405]:
```
sns.barplot(x='index', y='Shots',
            data=kgs57.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs57.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



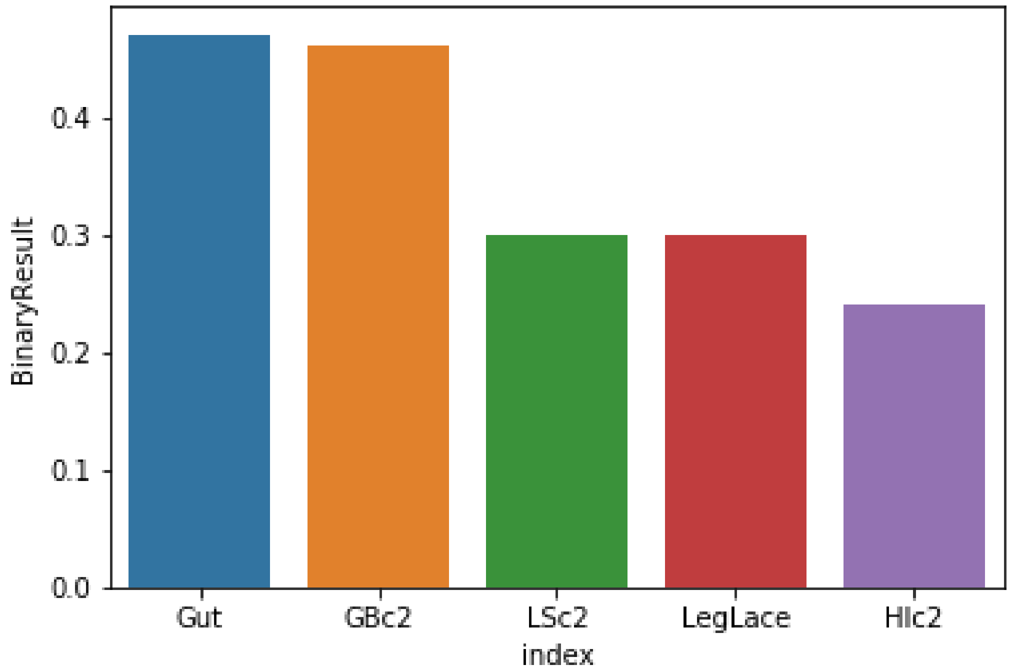


Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[627]:
```
print('Top actions for 57kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs57[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 57kgs at this event.
```



1.2.9 97 kilos 69 matches

Placements

1. Daton Fix
2. Thomas Gilman
3. Zane Richards
4. Vitali Arujau
5. Darian Cruz
6. Nathan Tomasello
7. Frank Perrelli
8. Zach Sanders

[321]:
```
kgs57 = df[df['Weight'] == 57]
```

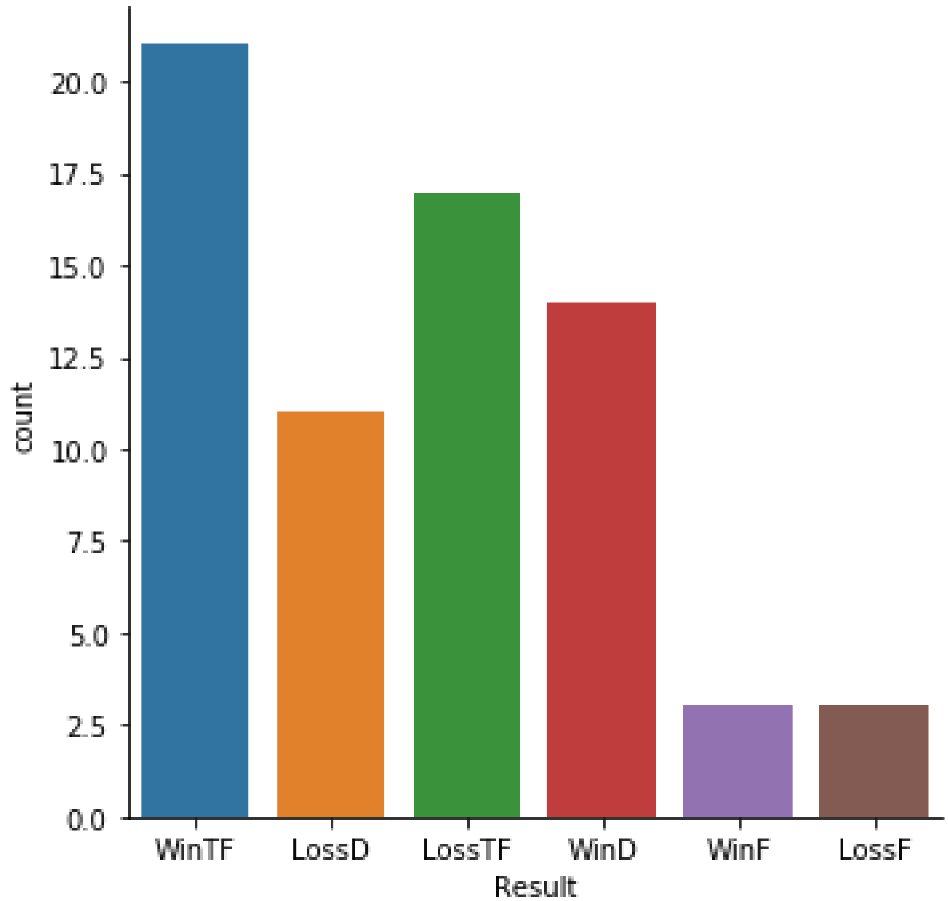
Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

[499]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs57[kgs57['MatchID'].map(len) == 4], kind="count")
g.set_xticklabels(rotation=30)
plt.show()
```

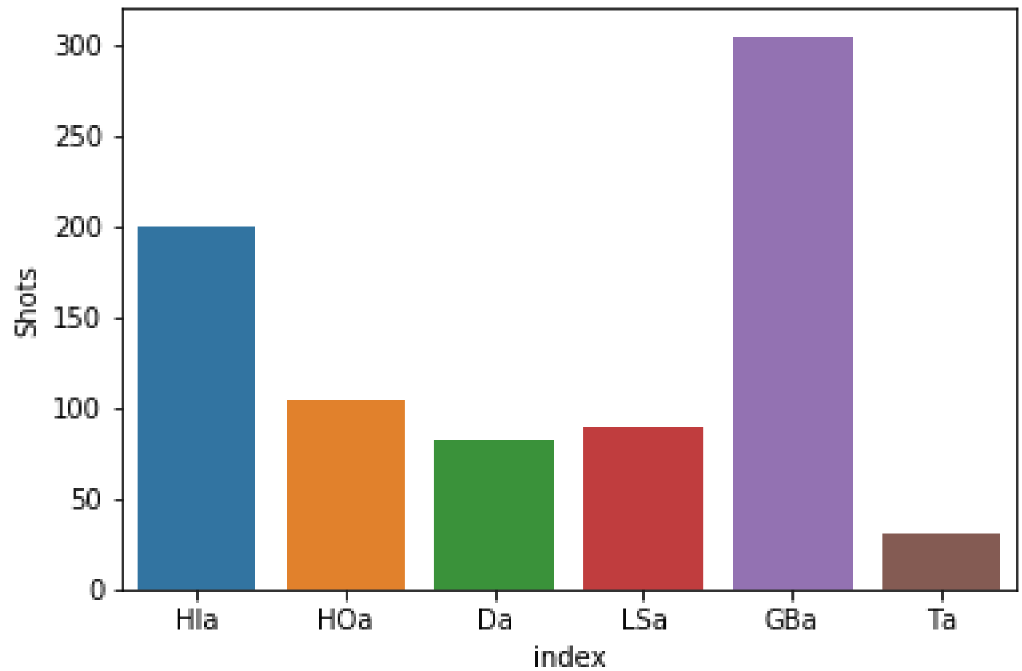
```
# somehow order so can use palatte for win/loss
```

```
Distributions of match results.
```

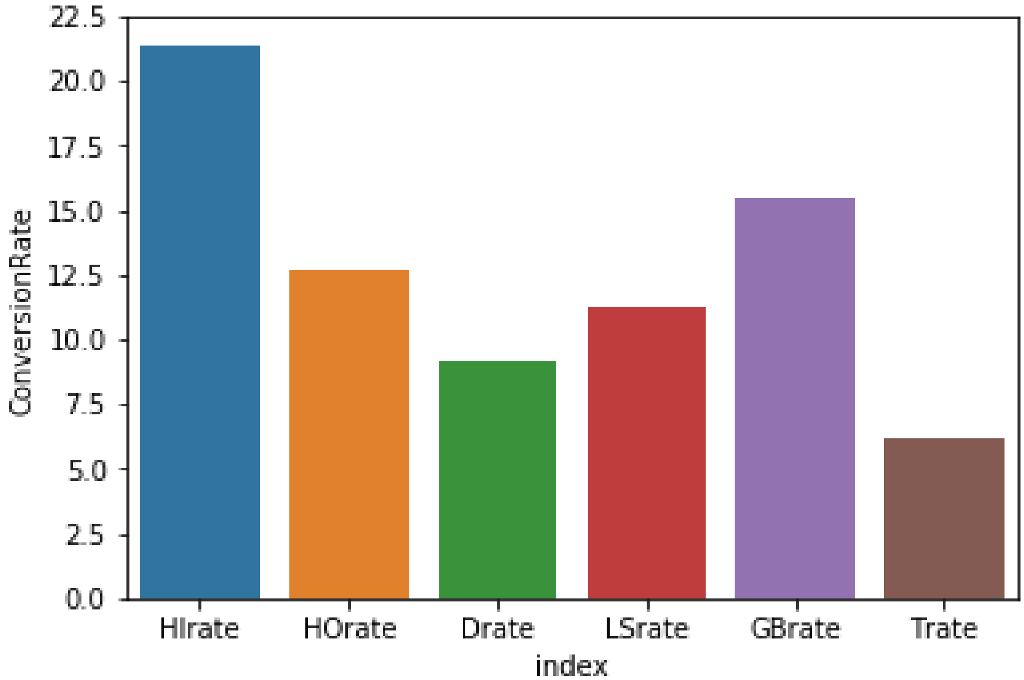


GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[405]:
```
sns.barplot(x='index', y='Shots',
            data=kgs57.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs57.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
 ↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



300
250
200
150
100
50
0
Shots
HIa
HOa
Da
LSa
GBa
Ta
index



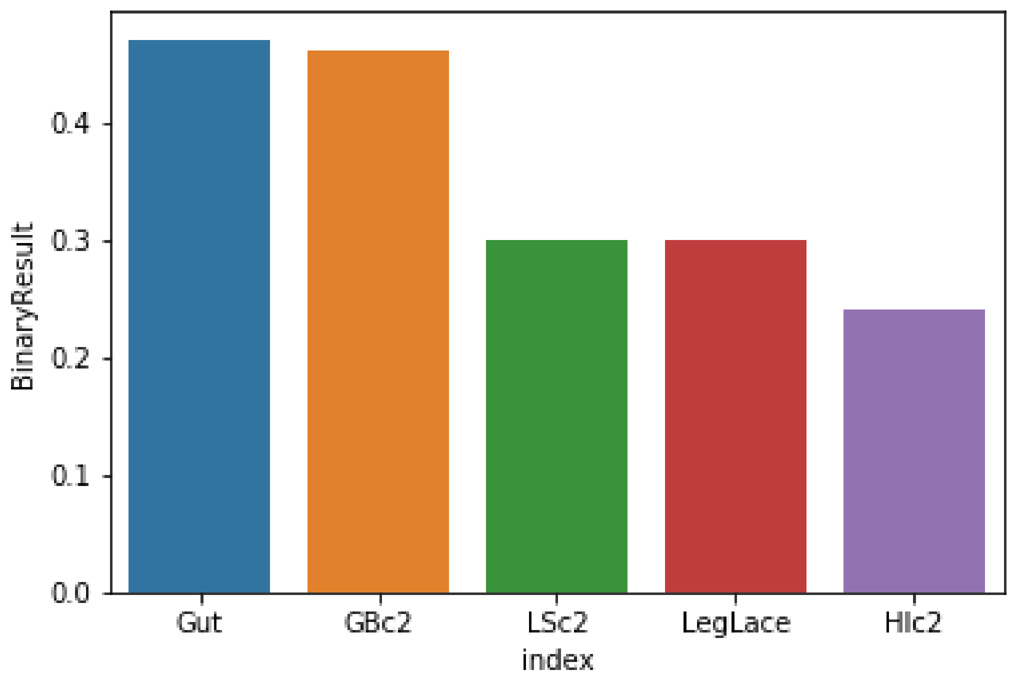
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
ConversionRate
HIrate
HOrate
Drate
LSrate
GBrate
Trate
index

Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[627]:
```
print('Top actions for 57kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs57[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 57kgs at this event.
```



1.2.10 125 kilos 69 matches

Placements

1. Daton Fix
2. Thomas Gilman
3. Zane Richards
4. Vitali Arujau

5. Darian Cruz
6. Nathan Tomasello
7. Frank Perrelli
8. Zach Sanders

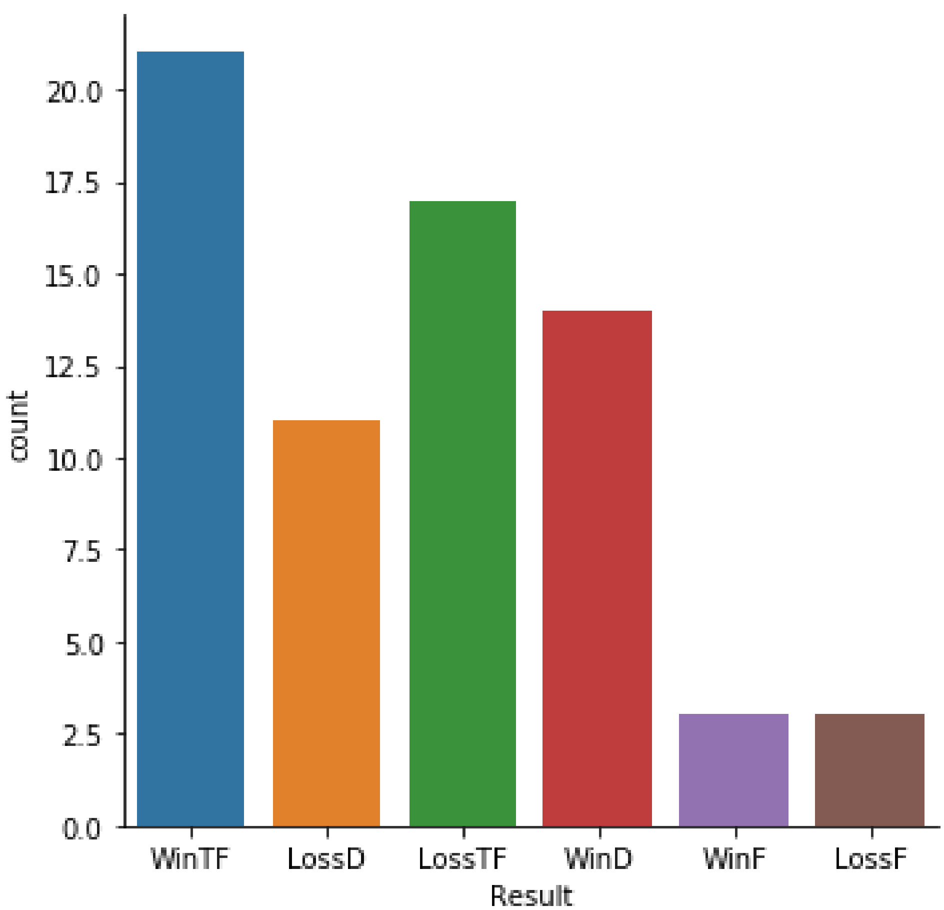
[321]:
```
kgs57 = df[df['Weight'] == 57]
```

Distributions This section focuses on illustrating the distribution of Results and ShotTypes for this event.

Techs are the most common result followed by Decisions and then Falls as is typical with a Freestyle event.

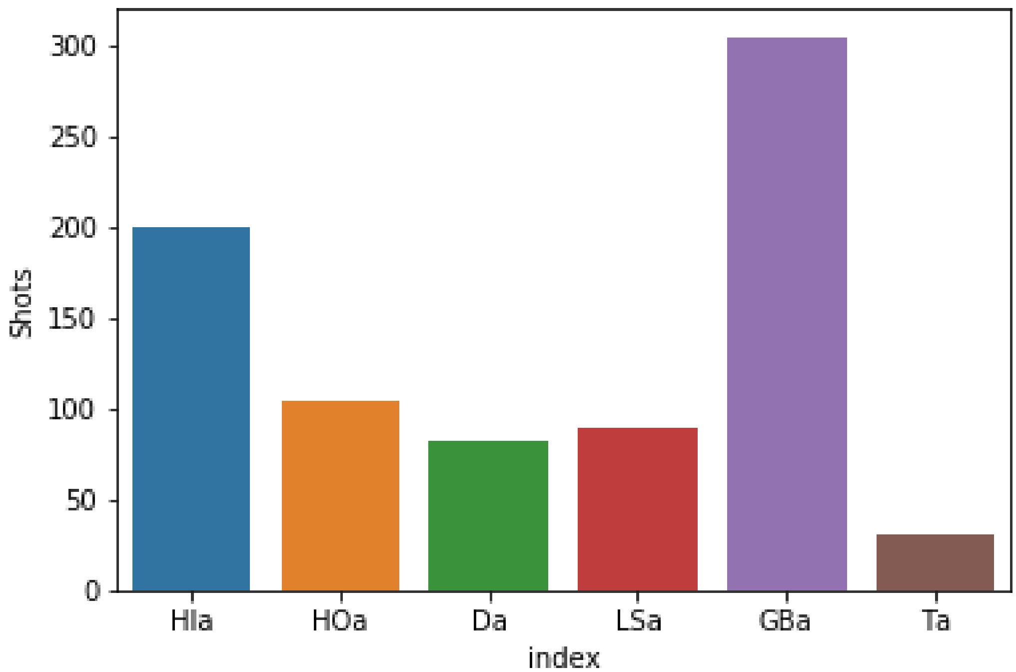
[499]:
```
print('Distributions of match results.')
sns.catplot("Result", data=kgs57[kgs57['MatchID'].map(len) == 4], kind="count")
g.set_xticklabels(rotation=30)
plt.show()

# somehow order so can use palatte for win/loss
```

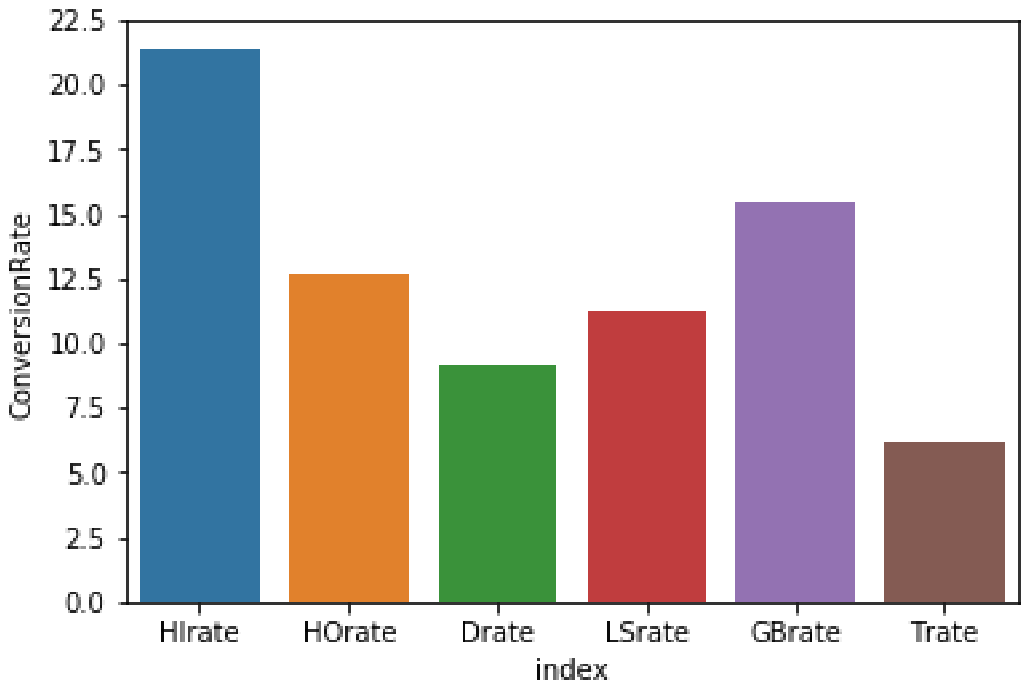
```
Distributions of match results.
```



GBa will almost exclusively be #1 on this list due to the large variety of conditions under which it is recorded. More interesting is the high GBrate and even higher HIrate at this weight. Other ShotTypes follow their corresponding attempts patterns.

[405]:
```
sns.barplot(x='index', y='Shots',
            data=kgs57.sum()[['HIa', 'HOa', 'Da', 'LSa', 'GBa', 'Ta']]\
            .rename('Shots').reset_index())
plt.show()
sns.barplot(x='index', y='ConversionRate',
            data=kgs57.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate', 'GBrate',␣
 ↪'Trate']]\
            .rename('ConversionRate').reset_index())
plt.show()
```



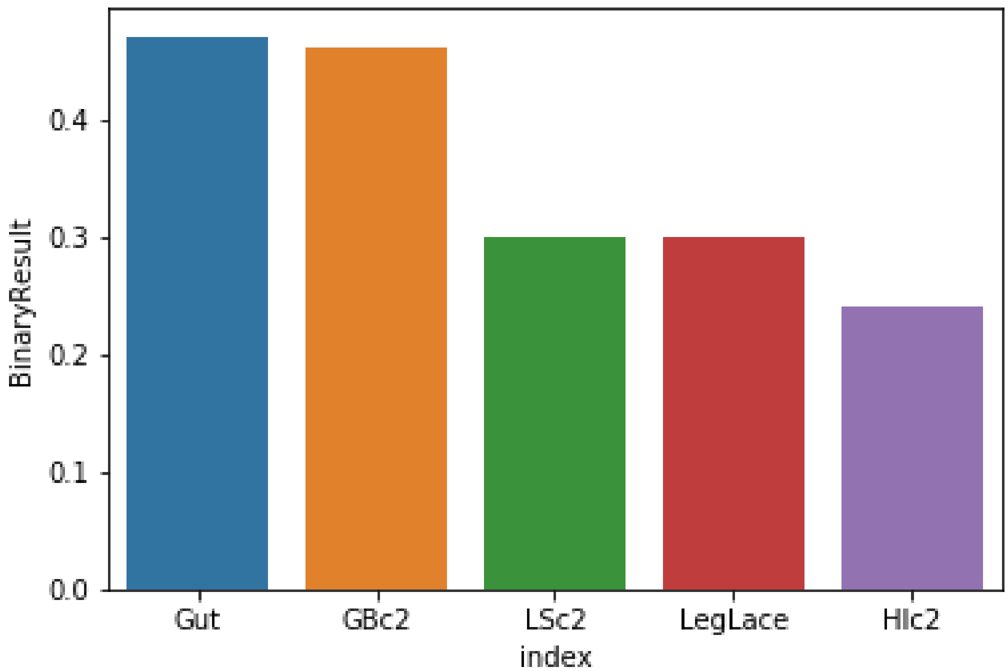


Trends This section focuses on identifying key metrics for success at this weight (at this event). The analysis does not take into consideration the value of more powerful results. In otherwords, correlations are done only on binary win/loss results and **not** on the degree of the result. So a strong correlation indicates a higher relationship with winning but **not** the degree of winning.

Remember correlations range from -1 to +1 the former representating a strong relationship with a negative outcome (in this case losing the match) and the latter representation a strong relationship with a positive outcome (i.e. winning a match). Correlations closer to 0 represent weak relationships and correlations around 0.5 represent moderate relationships.

[627]:
```
print('Top actions for 57kgs at this event.')
sns.barplot(x='index', y='BinaryResult',
    data=kgs57[base_cols].corr().round(2).sort_values(by=['BinaryResult'],␣
 ↪ascending=False)['BinaryResult'][1:6].reset_index())
plt.show()
```

```
Top actions for 57kgs at this event.
```



1.2.11 Comparison

The comparison between weight classes will focus mostly on Shot Types and Point Distributions as opposed to the result-based distribution above.

Below is a table illustrating different Shot Type conversion rates for the various weight classes. This table can be interpreted multiple ways. For instance, is a low shot conversion rate due to wrestlers at that weight being incapable/inefficient at finishing, or is it a sign that opponents at that weight are excellent defenders of that shot? The following discussion attempts to take both of these perspectives into consideration.

Firstly, it is important to state that top scores for a particular column (or shot type) are highlighted in blue, while bottom scores are highlighted in gold. This table should be read across (left to right), checking the weight class and then identifying its highest/lowest values. Knowing this, we can clearly identify the best/worst weight class for any given shot type. We can see, for example, that 74kgs has the highest Throw-rate. Interestingly, 125kgs has the highest GBrate… perhaps because of their size.

What may appear less obvious are some trends. For example, that 61kgs is the worst in two categories (LS and GB) as is 79kgs (HO and D and the lowest overall average) but **no** weight is the best in two categories. This showcases the diversity of competing at different weights. 65kgs has the highest overall average, but most scores are in the teens.

Further, some athletes or coaches may find this chart interesting to compare the dynamics of different weight classes in the neutral position. However, these numbers do not represent correlations to success but only raw averages, so bear that in mind.

[645]:
```
def highlight_max(s):
    '''
    highlight the maximum in a Series blue.
    '''
    is_max = s == s.max()
    return ['background-color: lightblue' if v else '' for v in is_max]


def highlight_min(s):
    '''
    highlight the minimum in a Series gold.
    '''
    is_max = s == s.min()
    return ['background-color: gold' if v else '' for v in is_max]


weight_shots = weight_grouped.mean()[['HIrate', 'HOrate', 'Drate', 'LSrate',␣
↪'GBrate', 'Trate']]
weight_shots['AverageRate'] = weight_shots.mean(axis=1)
display(weight_shots.style.apply(highlight_max).apply(highlight_min))
```

```
<pandas.io.formats.style.Styler at 0x7f39704a57d0>
```

The next table illustrates the various ways to earn points in a match (with the exception of opponent violations) and their highest/lowest values. This table has the same highlighting as the previous one, but should be read vertically (top to bottom) first identifying the weight class of interest then locating its highest/lowest scores.

The scores in this table represent the average points earned per match by each action/technique for each weight class. Similar to the table above this table has multiple possible interpretations. For example, does an action/technique have a low value because it is executed poorly, defended well, ***or*** because it is rare and does not occur often?

Guts and GBc2 are clearly popular ways of scoring and account for almost identical (0.83 and 0.82 respectively) points per match. Tc2 is obviously rare as you would expect most Throws to go for 4. The same logic applies to all the leg attacks (HI, HO, D, LS) where it is less common to see a 4-point leg attack finish.

[683]:
```
def apply_points(row):
    """Multiples by 1, 2, or 4 based on the action to change numeric␣
↪representation to points from action."""
    row['Pushout'] *= 1
    row['HIc2'] *= 2
    row['HOc2'] *= 2
    row['Dc2'] *= 2
    row['LSc2'] *= 2
    row['GBc2'] *= 2
    row['Tc2'] *= 2
```

```
    row['Exposure'] *= 2
    row['Gut'] *= 2
    row['LegLace'] *= 2
    row['Turn'] *= 2
    row['HIc4'] *= 2
    row['HOc4'] *= 2
    row['Dc4'] *= 2
    row['LSc4'] *= 2
    row['Tc4'] *= 2
    return row


weight_actions = weight_grouped.mean()[[
    'HIc2', 'HIc4', 'HOc2', 'HOc4', 'Dc2', 'Dc4', 'LSc2', 'LSc4', 'GBc2',
    'Tc2', 'Tc4', 'Exposure', 'Gut', 'LegLace', 'Turn', 'Pushout']].T.
 ↪apply(lambda row: apply_points(row))
weight_actions['AvgPoints'] = weight_actions.mean(axis=1)
display(weight_actions.round(2).style.apply(highlight_min).apply(highlight_max))
```

```
<pandas.io.formats.style.Styler at 0x7f397150b690>
```

1.3 Madness Matches

This section focuses on matches with abnormal outcomes. This is sometimes statistically determined (i.e. matches with various metric values outside their third standard deviation values, but sometimes can be determined visually using common sense.
inter

1.3.1 Match OEJV (Jesse Vasquez vs Brandon Wright)

This match went the distance as Vasquez and Wright put up 35 combined points.
A back and forth battle, the MoV never exceeded +/- 4 for either wrestler. The deciding factor came when Wright took the lead with a go-behind and 4 succesive leg laces with less than a minute to go. Vasquez managed a takedown and two gut wrenches to come within 1 point but was unable to take the lead in the last 5 seconds.

[563]:
```
display(df[df['Focus'] == 'Jesse Vasquez'])
```

```
       APM  Drate  Da        Date  Dc2  Dc4  Duration  Exposure  \
1215  30.0  100.0   1  2019-04-26    1    0    360.45         2

              Focus  FocusPoints  ... Tc4  Turn  Violation    VS  Weight  \
1215  Jesse Vasquez           17  ...   0     0          0  45.9      65

      PassiveDiff  NumResult  BinaryResult  BinaryResultText  ResultType
1215            0        0.9             0              Loss    Decision
```

```
[1 rows x 85 columns]
```

1.3.2 Match QCY3 (Renaldo Rodriguez Spencer vs James Steerman)

You might've missed this match if you blinked… it lasted only 10.37 seconds.
Spencer came out strong with a 4-point takedown followed by three succesive gut wrenches to put the match away.

[579]:
```
display(df[df['FocusPoints'] > 0].sort_values('Duration').head(1).to_html())
```

```
'<table border="1" class="dataframe">\n  <thead>\n    <tr style="text-align: right;">\n      <
```

1.3.3 Match WYEF (AJ Jaffe vs Drew Mandell)

Mandell came out strong with an exposure early in the match. However Jaffe stayed calm and slowly gained an increasing lead despite trading points throughout the bout. Eventually he ended it around the 5:30 mark with a WinTF at a whooping 24-13!

All of those points lead to a high APM score of 31 and 46.65 VS for Jaffe after the win. Interestingly this was an even bout on their feet as Jaffe registered only a 1.11 NPF score.

[580]:
```
df[df['FocusPoints'] == df['FocusPoints'].max()]
```

[580]:
```
      APM  Drate  Da        Date  Dc2  Dc4  Duration  Exposure     Focus  \
448  31.0  100.0   2  2019-04-26    1    1    320.96         2  AJ Jaffe

     FocusPoints  … Tc4  Turn  Violation     VS  Weight  PassiveDiff  \
448           24  …   1     2          0  46.65      70            0

     NumResult  BinaryResult  BinaryResultText  ResultType
448        1.5             1               Win        Tech

[1 rows x 85 columns]
```

1.4 Veritas Bites

This section emphasizes Veritas Score which accounts for the match result, NPF, and APM and puts all of these KPMs plus over 70 other data points from each match to into a formula that outputs a numeric representation of overall match performance. Veritas Score scales based on the degree of victory.

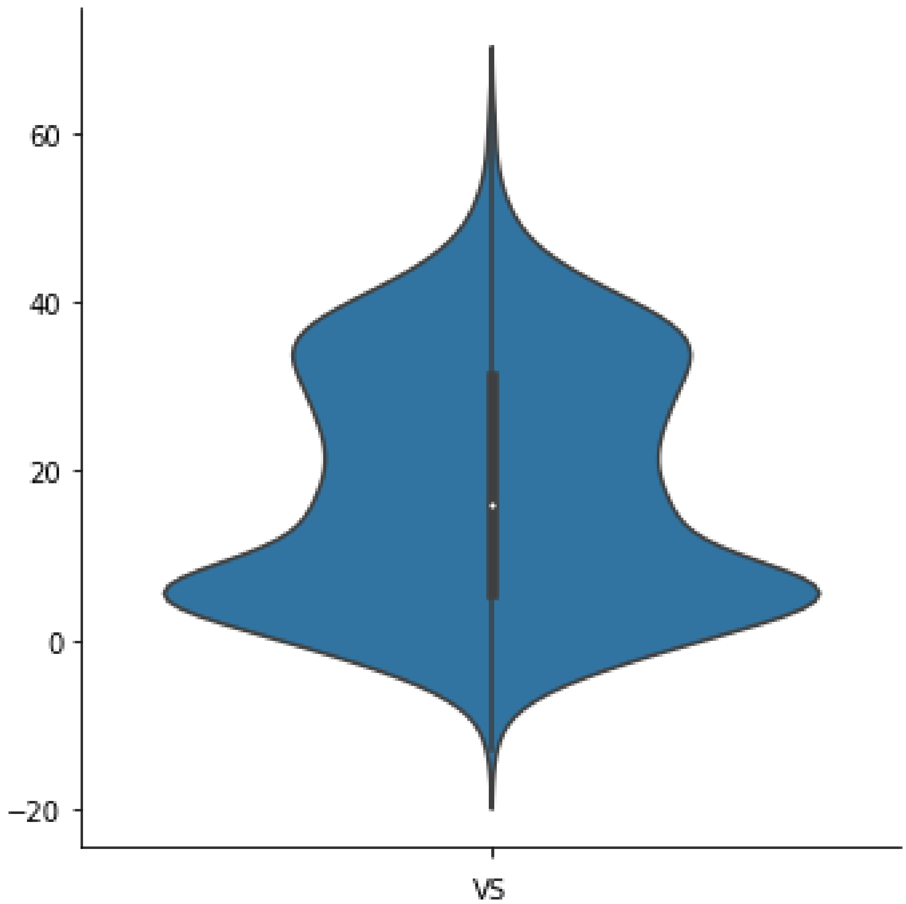
1.4.1 Overall Event

The event average a VS of 18.3 with a standard deviation of 14.6.
Most matches fell between 5.5 - 31.1 VS.

[582]:
```
display(df['VS'].describe().to_frame().to_html())
```

```python
sns.catplot(x='VS', kind='violin', orient='v', data=df)
plt.show()
```

```
'<table border="1" class="dataframe">\n  <thead>\n    <tr style="text-align: right;">\n      <
```



1.4.2 High Value

X4HH had the highest VS (63.5) in a Jason Ness WinTF over Dom Demas.

[524]:
```python
display(df[df['VS'] == df['VS'].max()].to_html())
```

```
'<table border="1" class="dataframe">\n  <thead>\n    <tr style="text-align: right;">\n      <
```

1.4.3 Low Value

Match 9AOY* had the lowest VS (-13) in a LossTF Donald Mcneil fell to Wynn Michalak.

[528]:

```
display(df[df['VS'] == df['VS'].min()].to_html())
```

```
'<table border="1" class="dataframe">\n  <thead>\n    <tr style="text-align: right;">\n      <t
```

[ ]:

```

```

[ ]:

```

```