

BOS-NY

October 23, 2023

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[2]: bos_swingtake = pd.read_csv("swing-take.csv")
ny_swingtake = pd.read_csv("swing-take (1).csv")
bos_pitchars = pd.read_csv("pitch-arsenal-stats.csv")
ny_pitchars = pd.read_csv("pitch-arsenal-stats (1).csv")
```

```
[3]: bos_swingtake
```

```
[3]:   year last_name first_name player_id team_id  pa  pitches  runs_all \
0  2023    Turner    Justin    457759     111  450    1935      17
1  2023     Duran    Jarren    680776     111  327    1270      16
2  2023    Devers    Rafael    646240     111  446    1754      18
3  2023   Yoshida  Masataka    807799     111  423    1594      15
4  2023   Verdugo     Alex    657077     111  432    1708       3
5  2023     Casas    Triston    671213     111  365    1471       5
6  2023      Wong    Connor    657136     111  285    1081     -10

   runs_heart  runs_shadow  runs_chase  runs_waste
0          -3           0         14          7
1           3           2          4          7
2           4          -8         12          9
3          -2           0         10          6
4           0         -16         12          7
5           1         -14         11          7
6           1         -20          5          5
```

```
[4]: bos_swingtake.sort_values(by="runs_all", ascending = False)
```

```
[4]:   year last_name first_name player_id team_id  pa  pitches  runs_all \
2  2023    Devers    Rafael    646240     111  446    1754      18
0  2023    Turner    Justin    457759     111  450    1935      17
1  2023     Duran    Jarren    680776     111  327    1270      16
3  2023   Yoshida  Masataka    807799     111  423    1594      15
```

5	2023	Casas	Triston	671213	111	365	1471	5
4	2023	Verdugo	Alex	657077	111	432	1708	3
6	2023	Wong	Connor	657136	111	285	1081	-10

	runs_heart	runs_shadow	runs_chase	runs_waste
2	4	-8	12	9
0	-3	0	14	7
1	3	2	4	7
3	-2	0	10	6
5	1	-14	11	7
4	0	-16	12	7
6	1	-20	5	5

```
[5]: bos_swingtake.sort_values(by="runs_heart", ascending = False)
```

```
[5]:   year last_name first_name player_id team_id pa pitches runs_all \
2  2023   Devers   Rafael    646240    111  446   1754    18
1  2023    Duran   Jarren    680776    111  327   1270    16
5  2023   Casas   Triston    671213    111  365   1471     5
6  2023    Wong   Connor    657136    111  285   1081   -10
4  2023  Verdugo    Alex    657077    111  432   1708     3
3  2023  Yoshida  Masataka    807799    111  423   1594    15
0  2023   Turner   Justin    457759    111  450   1935    17
```

	runs_heart	runs_shadow	runs_chase	runs_waste
2	4	-8	12	9
1	3	2	4	7
5	1	-14	11	7
6	1	-20	5	5
4	0	-16	12	7
3	-2	0	10	6
0	-3	0	14	7

```
[6]: bos_swingtake.sort_values(by="runs_shadow", ascending = False)
```

```
[6]:   year last_name first_name player_id team_id pa pitches runs_all \
1  2023    Duran   Jarren    680776    111  327   1270    16
0  2023   Turner   Justin    457759    111  450   1935    17
3  2023  Yoshida  Masataka    807799    111  423   1594    15
2  2023   Devers   Rafael    646240    111  446   1754    18
5  2023   Casas   Triston    671213    111  365   1471     5
4  2023  Verdugo    Alex    657077    111  432   1708     3
6  2023    Wong   Connor    657136    111  285   1081   -10
```

	runs_heart	runs_shadow	runs_chase	runs_waste
1	3	2	4	7
0	-3	0	14	7

3	-2	0	10	6
2	4	-8	12	9
5	1	-14	11	7
4	0	-16	12	7
6	1	-20	5	5

```
[7]: bos_swingtake.sort_values(by="runs_chase", ascending = False)
```

```
[7]:   year last_name first_name player_id team_id pa pitches runs_all \
0  2023    Turner    Justin   457759    111  450    1935    17
2  2023    Devers    Rafael   646240    111  446    1754    18
4  2023   Verdugo    Alex    657077    111  432    1708     3
5  2023    Casas    Triston   671213    111  365    1471     5
3  2023   Yoshida   Masataka   807799    111  423    1594    15
6  2023     Wong    Connor   657136    111  285    1081   -10
1  2023    Duran    Jarren   680776    111  327    1270    16
```

	runs_heart	runs_shadow	runs_chase	runs_waste
0	-3	0	14	7
2	4	-8	12	9
4	0	-16	12	7
5	1	-14	11	7
3	-2	0	10	6
6	1	-20	5	5
1	3	2	4	7

```
[8]: ny_pitchers
```

```
[8]:   last_name first_name player_id team_name_alt pitch_type pitch_name \
0      Cole    Gerrit   543037          NYY      FF    4-Seamer
1   Germán    Domingo   593334          NYY    CUKC  Curveball
2  Severino    Luis    622663          NYY      FF    4-Seamer
3    Cortes    Nestor   641482          NYY      FF    4-Seamer
4   Schmidt    Clarke   657376          NYY      FC    Cutter
..      ...      ...      ...      ...      ...
64  Vásquez    Randy   681190          NYY    CUKC  Curveball
65  Vásquez    Randy   681190          NYY      FC    Cutter
66 Weissert    Greg    669711          NYY      ST    Sweeper
67    Weber    Ryan    543901          NYY      CH  Changeup
68    Cortes    Nestor   641482          NYY      CH  Changeup
```

	run_value_per_100	run_value	pitches	pitch_usage	...	ba	slg	\
0	1.6	21	1300	54.4	...	0.201	0.333	
1	0.7	5	676	40.9	...	0.174	0.344	
2	-2.7	-15	553	48.7	...	0.370	0.719	
3	0.1	1	556	52.1	...	0.213	0.352	
4	-0.2	-1	496	26.9	...	0.288	0.525	

```

..          ...          ...          ...          ...          ...          ...
64          3.7          1          24          9.8          ... 0.222 0.333
65         -0.5          0          50         20.4          ... 0.286 0.714
66         -5.9         -3          43         31.4          ... 0.500 0.875
67         -1.4          0          26         14.3          ... 0.300 0.600
68         -6.7         -2          30          2.8          ... 0.500 0.600

      woba  whiff_percent  k_percent  put_away  est_ba  est_slg  est_woba  \
0  0.282          24.2          30.2          23.7  0.221  0.397  0.306
1  0.270          39.3          36.8          32.5  0.180  0.320  0.251
2  0.490          18.0          17.9          16.7  0.332  0.633  0.437
3  0.284          24.9          30.6          19.7  0.185  0.276  0.244
4  0.379          24.4          20.5          17.4  0.248  0.445  0.333
..          ...          ...          ...          ...          ...          ...
64  0.323          33.3          27.3          25.0  0.182  0.234  0.273
65  0.500          22.7          20.0          25.0  0.297  0.780  0.520
66  0.610          21.4          10.0           7.7  0.356  0.774  0.519
67  0.380          35.3          30.0          37.5  0.225  0.545  0.319
68  0.485          29.4          40.0          36.4  0.355  0.408  0.332

      hard_hit_percent
0          46.9
1          35.9
2          54.1
3          47.6
4          45.2
..          ...
64          33.3
65          60.0
66          42.9
67          14.3
68          16.7

```

[69 rows x 21 columns]

```
[9]: bos_swingtake['shadowVsHeart'] = bos_swingtake['runs_shadow'] -
      ↪bos_swingtake['runs_heart']
```

```
[10]: bos_swingtake.sort_values(by = 'shadowVsHeart', ascending = True)
```

```
[10]:  year last_name  first_name  player_id  team_id  pa  pitches  runs_all  \
6  2023      Wong      Connor    657136     111   285    1081      -10
4  2023  Verdugo      Alex      657077     111   432    1708        3
5  2023     Casas  Triston    671213     111   365    1471        5
2  2023     Devers  Rafael    646240     111   446    1754       18
1  2023     Duran   Jarren    680776     111   327    1270       16
3  2023  Yoshida  Masataka    807799     111   423    1594       15

```

0	2023	Turner	Justin	457759	111	450	1935	17
		runs_heart	runs_shadow	runs_chase	runs_waste	shadowVsHeart		
6		1	-20	5	5	-21		
4		0	-16	12	7	-16		
5		1	-14	11	7	-15		
2		4	-8	12	9	-12		
1		3	2	4	7	-1		
3		-2	0	10	6	2		
0		-3	0	14	7	3		

```
[11]: bos_swingtake['shadowVsChase'] = bos_swingtake['runs_shadow'] -
↳ bos_swingtake['runs_chase']
```

```
[12]: bos_swingtake.sort_values(by = 'shadowVsChase', ascending = True)
```

```
[12]:   year last_name first_name player_id team_id pa pitches runs_all \
4  2023  Verdugo      Alex      657077      111  432    1708      3
5  2023   Casas    Triston      671213      111  365    1471      5
6  2023   Wong     Connor      657136      111  285    1081     -10
2  2023  Devers    Rafael      646240      111  446    1754     18
0  2023  Turner    Justin      457759      111  450    1935     17
3  2023  Yoshida  Masataka      807799      111  423    1594     15
1  2023   Duran    Jarren      680776      111  327    1270     16
```

	runs_heart	runs_shadow	runs_chase	runs_waste	shadowVsHeart	\
4	0	-16	12	7	-16	
5	1	-14	11	7	-15	
6	1	-20	5	5	-21	
2	4	-8	12	9	-12	
0	-3	0	14	7	3	
3	-2	0	10	6	2	
1	3	2	4	7	-1	

	shadowVsChase
4	-28
5	-25
6	-25
2	-20
0	-14
3	-10
1	-2

```
[13]: bos_swingtake['heartVsChase'] = bos_swingtake['runs_heart'] -
↳ bos_swingtake['runs_chase']
```

```
[14]: bos_swingtake.sort_values(by = 'heartVsChase', ascending = True)
```

```
[14]:   year last_name first_name player_id team_id pa pitches runs_all \
0  2023    Turner      Justin   457759    111  450   1935     17
3  2023   Yoshida   Masataka   807799    111  423   1594     15
4  2023   Verdugo     Alex    657077    111  432   1708      3
5  2023     Casas   Triston    671213    111  365   1471      5
2  2023    Devers   Rafael    646240    111  446   1754     18
6  2023     Wong    Connor    657136    111  285   1081    -10
1  2023     Duran   Jarren    680776    111  327   1270     16
```

```
   runs_heart runs_shadow runs_chase runs_waste shadowVsHeart \
0          -3           0         14           7           3
3          -2           0         10           6           2
4           0          -16         12           7          -16
5           1          -14         11           7          -15
2           4           -8         12           9          -12
6           1          -20          5           5          -21
1           3           2          4           7           -1
```

```
   shadowVsChase heartVsChase
0          -14          -17
3          -10          -12
4          -28          -12
5          -25          -10
2          -20           -8
6          -25           -4
1           -2           -1
```

```
[15]: bos_swingtake.iloc[:,7:15].describe()
```

```
[15]:   runs_all runs_heart runs_shadow runs_chase runs_waste \
count    7.000000    7.000000    7.000000    7.000000    7.000000
mean     9.142857    0.571429   -8.000000    9.714286    6.857143
std     10.350983    2.507133    8.869423    3.773340    1.214986
min    -10.000000   -3.000000  -20.000000    4.000000    5.000000
25%      4.000000   -1.000000  -15.000000    7.500000    6.500000
50%     15.000000    1.000000   -8.000000   11.000000    7.000000
75%     16.500000    2.000000    0.000000   12.000000    7.000000
max     18.000000    4.000000    2.000000   14.000000    9.000000
```

```
   shadowVsHeart shadowVsChase heartVsChase
count    7.000000    7.000000    7.000000
mean     -8.571429   -17.714286   -9.142857
std      9.710083    9.464218    5.367450
min     -21.000000  -28.000000  -17.000000
25%     -15.500000  -25.000000  -12.000000
50%     -12.000000  -20.000000  -10.000000
75%      0.500000  -12.000000   -6.000000
```

max 3.000000 -2.000000 -1.000000

```
[16]: nybat_pitchers = pd.read_csv('pitch-arsenal-stats-ny-bat.csv')
      bosbat_pitchers = pd.read_csv('pitch-arsenal-stats-rs-bat.csv')
      bospt_swingtake = pd.read_csv('swing-take-bs-pit.csv')
      nypit_swingtake = pd.read_csv('swing-take-ny-pit.csv')
```

```
[17]: nypit_swingtake
```

```
[17]:   year  last_name  first_name  player_id  team_id  pa  pitches  runs_all  \
0  2023      Cole      Gerrit    543037     147  596    2389      28
1  2023  Severino      Luis    622663     147  298    1135     -28
2  2023   Holmes      Clay    605280     147  179     720     13
3  2023   Brito      Jhony    666745     147  243     976     -6
4  2023    King  Michael    650633     147  249     952      8
5  2023  Marinaccio      Ron    676760     147  199     861      2
6  2023   Germán  Domingo    593334     147  444    1653      1
7  2023   Abreu   Albert    656061     147  205     830      0
8  2023   Cortes   Nestor    641482     147  265    1068     -5
9  2023   Schmidt   Clarke    657376     147  495    1933      2

      runs_heart  runs_shadow  runs_chase  runs_waste
0             26           22         -12          -7
1            -11           -4          -7          -6
2             12            7          -4          -2
3              3            2          -6          -5
4              4           10          -4          -3
5              5           13          -8          -7
6              4            6          -3          -5
7              1            8          -4          -5
8            -10           14          -6          -3
9              1           14          -7          -7
```

```
[18]: nypit_swingtake['shadowVsHeart'] = nypit_swingtake['runs_shadow'] -_
      ↪nypit_swingtake['runs_heart']
      nypit_swingtake['shadowVsChase'] = nypit_swingtake['runs_shadow'] -_
      ↪nypit_swingtake['runs_chase']
      nypit_swingtake['heartVsChase'] = nypit_swingtake['runs_heart'] -_
      ↪nypit_swingtake['runs_chase']
```

```
[19]: nypit_swingtake.iloc[:,7:15].describe()
```

```
[19]:   runs_all  runs_heart  runs_shadow  runs_chase  runs_waste  \
count  10.000000   10.000000    10.00000    10.000000   10.000000
mean     1.500000    3.500000     9.20000    -6.100000   -5.000000
std     14.331395   10.469533     7.20802     2.643651    1.825742
min    -28.000000  -11.000000    -4.00000   -12.000000   -7.000000
```

25%	-3.750000	1.000000	6.25000	-7.000000	-6.750000
50%	1.500000	3.500000	9.00000	-6.000000	-5.000000
75%	6.500000	4.750000	13.75000	-4.000000	-3.500000
max	28.000000	26.000000	22.00000	-3.000000	-2.000000

	shadowVsHeart	shadowVsChase	heartVsChase
count	10.000000	10.000000	10.000000
mean	5.700000	15.300000	9.600000
std	8.590046	8.895067	11.843423
min	-5.000000	3.000000	-4.000000
25%	-0.250000	9.500000	5.500000
50%	6.500000	13.000000	8.000000
75%	7.750000	20.750000	12.000000
max	24.000000	34.000000	38.000000

```
[20]: nypit_swingtake
```

```
[20]:   year  last_name  first_name  player_id  team_id  pa  pitches  runs_all  \
0  2023      Cole      Gerrit    543037    147    596    2389      28
1  2023  Severino      Luis    622663    147    298    1135     -28
2  2023    Holmes      Clay    605280    147    179     720     13
3  2023     Brito     Jhony    666745    147    243     976     -6
4  2023      King   Michael    650633    147    249     952      8
5  2023  Marinaccio      Ron    676760    147    199     861      2
6  2023   Germán   Domingo    593334    147    444    1653      1
7  2023     Abreu   Albert    656061    147    205     830      0
8  2023    Cortes   Nestor    641482    147    265    1068     -5
9  2023   Schmidt   Clarke    657376    147    495    1933      2
```

	runs_heart	runs_shadow	runs_chase	runs_waste	shadowVsHeart	\
0	26	22	-12	-7	-4	
1	-11	-4	-7	-6	7	
2	12	7	-4	-2	-5	
3	3	2	-6	-5	-1	
4	4	10	-4	-3	6	
5	5	13	-8	-7	8	
6	4	6	-3	-5	2	
7	1	8	-4	-5	7	
8	-10	14	-6	-3	24	
9	1	14	-7	-7	13	

	shadowVsChase	heartVsChase
0	34	38
1	3	-4
2	11	16
3	8	9
4	14	8

5	21	13
6	9	7
7	12	5
8	20	-4
9	21	8

```
[21]: nypit_swingtake.sort_values(by="runs_all", ascending = False)
```

```
[21]:   year  last_name  first_name  player_id  team_id  pa  pitches  runs_all  \
0  2023      Cole      Gerrit    543037    147    596    2389      28
2  2023     Holmes      Clay    605280    147    179     720      13
4  2023      King    Michael    650633    147    249     952       8
5  2023  Marinaccio      Ron    676760    147    199     861       2
9  2023     Schmidt    Clarke    657376    147    495    1933       2
6  2023     Germán    Domingo    593334    147    444    1653       1
7  2023      Abreu    Albert    656061    147    205     830       0
8  2023     Cortes    Nestor    641482    147    265    1068      -5
3  2023      Brito    Jhony    666745    147    243     976      -6
1  2023    Severino      Luis    622663    147    298    1135     -28
```

	runs_heart	runs_shadow	runs_chase	runs_waste	shadowVsHeart	\
0	26	22	-12	-7	-4	
2	12	7	-4	-2	-5	
4	4	10	-4	-3	6	
5	5	13	-8	-7	8	
9	1	14	-7	-7	13	
6	4	6	-3	-5	2	
7	1	8	-4	-5	7	
8	-10	14	-6	-3	24	
3	3	2	-6	-5	-1	
1	-11	-4	-7	-6	7	

	shadowVsChase	heartVsChase
0	34	38
2	11	16
4	14	8
5	21	13
9	21	8
6	9	7
7	12	5
8	20	-4
3	8	9
1	3	-4

```
[22]: nypit_swingtake.sort_values(by="runs_heart", ascending = False)
```

```
[22]:
```

	year	last_name	first_name	player_id	team_id	pa	pitches	runs_all	\
0	2023	Cole	Gerrit	543037	147	596	2389	28	
2	2023	Holmes	Clay	605280	147	179	720	13	
5	2023	Marinaccio	Ron	676760	147	199	861	2	
4	2023	King	Michael	650633	147	249	952	8	
6	2023	Germán	Domingo	593334	147	444	1653	1	
3	2023	Brito	Jhony	666745	147	243	976	-6	
7	2023	Abreu	Albert	656061	147	205	830	0	
9	2023	Schmidt	Clarke	657376	147	495	1933	2	
8	2023	Cortes	Nestor	641482	147	265	1068	-5	
1	2023	Severino	Luis	622663	147	298	1135	-28	

	runs_heart	runs_shadow	runs_chase	runs_waste	shadowVsHeart	\
0	26	22	-12	-7	-4	
2	12	7	-4	-2	-5	
5	5	13	-8	-7	8	
4	4	10	-4	-3	6	
6	4	6	-3	-5	2	
3	3	2	-6	-5	-1	
7	1	8	-4	-5	7	
9	1	14	-7	-7	13	
8	-10	14	-6	-3	24	
1	-11	-4	-7	-6	7	

	shadowVsChase	heartVsChase
0	34	38
2	11	16
5	21	13
4	14	8
6	9	7
3	8	9
7	12	5
9	21	8
8	20	-4
1	3	-4

```
[23]: nypit_swingtake.sort_values(by="runs_shadow", ascending = False)
```

```
[23]:
```

	year	last_name	first_name	player_id	team_id	pa	pitches	runs_all	\
0	2023	Cole	Gerrit	543037	147	596	2389	28	
8	2023	Cortes	Nestor	641482	147	265	1068	-5	
9	2023	Schmidt	Clarke	657376	147	495	1933	2	
5	2023	Marinaccio	Ron	676760	147	199	861	2	
4	2023	King	Michael	650633	147	249	952	8	
7	2023	Abreu	Albert	656061	147	205	830	0	
2	2023	Holmes	Clay	605280	147	179	720	13	
6	2023	Germán	Domingo	593334	147	444	1653	1	

3	2023	Brito	Jhony	666745	147	243	976	-6
1	2023	Severino	Luis	622663	147	298	1135	-28

	runs_heart	runs_shadow	runs_chase	runs_waste	shadowVsHeart	\
0	26	22	-12	-7	-4	
8	-10	14	-6	-3	24	
9	1	14	-7	-7	13	
5	5	13	-8	-7	8	
4	4	10	-4	-3	6	
7	1	8	-4	-5	7	
2	12	7	-4	-2	-5	
6	4	6	-3	-5	2	
3	3	2	-6	-5	-1	
1	-11	-4	-7	-6	7	

	shadowVsChase	heartVsChase
0	34	38
8	20	-4
9	21	8
5	21	13
4	14	8
7	12	5
2	11	16
6	9	7
3	8	9
1	3	-4

```
[24]: nypit_swingtake.sort_values(by="runs_chase", ascending = False)
```

```
[24]:
```

	year	last_name	first_name	player_id	team_id	pa	pitches	runs_all	\
6	2023	Germán	Domingo	593334	147	444	1653	1	
2	2023	Holmes	Clay	605280	147	179	720	13	
4	2023	King	Michael	650633	147	249	952	8	
7	2023	Abreu	Albert	656061	147	205	830	0	
3	2023	Brito	Jhony	666745	147	243	976	-6	
8	2023	Cortes	Nestor	641482	147	265	1068	-5	
1	2023	Severino	Luis	622663	147	298	1135	-28	
9	2023	Schmidt	Clarke	657376	147	495	1933	2	
5	2023	Marinaccio	Ron	676760	147	199	861	2	
0	2023	Cole	Gerrit	543037	147	596	2389	28	

	runs_heart	runs_shadow	runs_chase	runs_waste	shadowVsHeart	\
6	4	6	-3	-5	2	
2	12	7	-4	-2	-5	
4	4	10	-4	-3	6	
7	1	8	-4	-5	7	
3	3	2	-6	-5	-1	

8	-10	14	-6	-3	24
1	-11	-4	-7	-6	7
9	1	14	-7	-7	13
5	5	13	-8	-7	8
0	26	22	-12	-7	-4

	shadowVsChase	heartVsChase
6	9	7
2	11	16
4	14	8
7	12	5
3	8	9
8	20	-4
1	3	-4
9	21	8
5	21	13
0	34	38

```
[25]: bos_swingtake_vs_nypit = bos_swingtake.merge(nypit_swingtake, how='cross',
↪suffixes=("_BOShitter", "_NYpitcher"))
```

```
[26]: bos_swingtake_vs_nypit.query("last_name_BOShitter == 'Turner'")
```

```
[26]:   year_BOShitter last_name_BOShitter first_name_BOShitter \
0          2023          Turner          Justin
1          2023          Turner          Justin
2          2023          Turner          Justin
3          2023          Turner          Justin
4          2023          Turner          Justin
5          2023          Turner          Justin
6          2023          Turner          Justin
7          2023          Turner          Justin
8          2023          Turner          Justin
9          2023          Turner          Justin
```

	player_id_BOShitter	team_id_BOShitter	pa_BOShitter	pitches_BOShitter	\
0	457759	111	450	1935	
1	457759	111	450	1935	
2	457759	111	450	1935	
3	457759	111	450	1935	
4	457759	111	450	1935	
5	457759	111	450	1935	
6	457759	111	450	1935	
7	457759	111	450	1935	
8	457759	111	450	1935	
9	457759	111	450	1935	

	runs_all_BOShitter	runs_heart_BOShitter	runs_shadow_BOShitter	...	\
0	17	-3	0	...	
1	17	-3	0	...	
2	17	-3	0	...	
3	17	-3	0	...	
4	17	-3	0	...	
5	17	-3	0	...	
6	17	-3	0	...	
7	17	-3	0	...	
8	17	-3	0	...	
9	17	-3	0	...	

	pa_NYpitcher	pitches_NYpitcher	runs_all_NYpitcher	runs_heart_NYpitcher	\
0	596	2389	28	26	
1	298	1135	-28	-11	
2	179	720	13	12	
3	243	976	-6	3	
4	249	952	8	4	
5	199	861	2	5	
6	444	1653	1	4	
7	205	830	0	1	
8	265	1068	-5	-10	
9	495	1933	2	1	

	runs_shadow_NYpitcher	runs_chase_NYpitcher	runs_waste_NYpitcher	\
0	22	-12	-7	
1	-4	-7	-6	
2	7	-4	-2	
3	2	-6	-5	
4	10	-4	-3	
5	13	-8	-7	
6	6	-3	-5	
7	8	-4	-5	
8	14	-6	-3	
9	14	-7	-7	

	shadowVsHeart_NYpitcher	shadowVsChase_NYpitcher	heartVsChase_NYpitcher
0	-4	34	38
1	7	3	-4
2	-5	11	16
3	-1	8	9
4	6	14	8
5	8	21	13
6	2	9	7
7	7	12	5
8	24	20	-4
9	13	21	8

[10 rows x 30 columns]

```
[27]: bos_swingtake_vs_nypit['overallDiff'] =  
      ↪bos_swingtake_vs_nypit['runs_all_BOShitter'] -  
      ↪bos_swingtake_vs_nypit['runs_all_NYpitcher']  
      bos_swingtake_vs_nypit['heartDiff'] =  
      ↪bos_swingtake_vs_nypit['runs_heart_BOShitter'] -  
      ↪bos_swingtake_vs_nypit['runs_heart_NYpitcher']  
      bos_swingtake_vs_nypit['shadowDiff'] =  
      ↪bos_swingtake_vs_nypit['runs_shadow_BOShitter'] -  
      ↪bos_swingtake_vs_nypit['runs_shadow_NYpitcher']  
      bos_swingtake_vs_nypit['chaseDiff'] =  
      ↪bos_swingtake_vs_nypit['runs_chase_BOShitter'] -  
      ↪bos_swingtake_vs_nypit['runs_chase_NYpitcher']
```

```
[28]: bos_swingtake_vs_nypit.groupby('last_name_NYpitcher')['overallDiff'].agg(np.  
      ↪mean).sort_values(ascending=True)
```

```
[28]: last_name_NYpitcher  
Cole          -18.857143  
Holmes         -3.857143  
King           1.142857  
Marinaccio     7.142857  
Schmidt        7.142857  
Germán        8.142857  
Abreu          9.142857  
Cortes        14.142857  
Brito         15.142857  
Severino      37.142857  
Name: overallDiff, dtype: float64
```

```
[29]: grouped_df = bos_swingtake_vs_nypit.groupby(['last_name_NYpitcher',  
      ↪'last_name_BOShitter'])['overallDiff', 'heartDiff',  
      ↪'shadowDiff', 'chaseDiff'].agg(np.mean).reset_index()
```

```
-----  
ValueError                                Traceback (most recent call last)
```

```
Cell In[29], line 1
```

```
----> 1 grouped_df =
```

```
      ↪bos_swingtake_vs_nypit.groupby(['last_name_NYpitcher', 'last_name_BOShitter'])  
      ↪agg(np.mean).reset_index()
```

```
File /srv/conda/envs/notebook/lib/python3.11/site-packages/pandas/core/groupby/
```

```
      ↪generic.py:1767, in DataFrameGroupBy.__getitem__(self, key)
```

```
    1763 # per GH 23566
```

```
    1764 if isinstance(key, tuple) and len(key) > 1:
```

```

1765     # if len == 1, then it becomes a SeriesGroupBy and this is actually
1766     # valid syntax, so don't raise
-> 1767     raise ValueError(
1768         "Cannot subset columns with a tuple with more than one element. "
1769         "Use a list instead."
1770     )
1771     return super().__getitem__(key)

```

ValueError: Cannot subset columns with a tuple with more than one element. Use a list instead.

```
[ ]:
```

```

[30]: bos_swingtake_vs_nypit.groupby(['last_name_NYpitcher',
    ↪ 'last_name_BOShitter'])['shadowDiff'].agg(np.mean).
    ↪ sort_values(ascending=False).to_frame()

```

```

[30]:

```

last_name_NYpitcher	last_name_BOShitter	shadowDiff
Severino	Duran	6.0
	Yoshida	4.0
	Turner	4.0
Brito	Duran	0.0
	Turner	-2.0
...
Cortes	Wong	-34.0
Schmidt	Wong	-34.0
Cole	Casas	-36.0
	Verdugo	-38.0
	Wong	-42.0

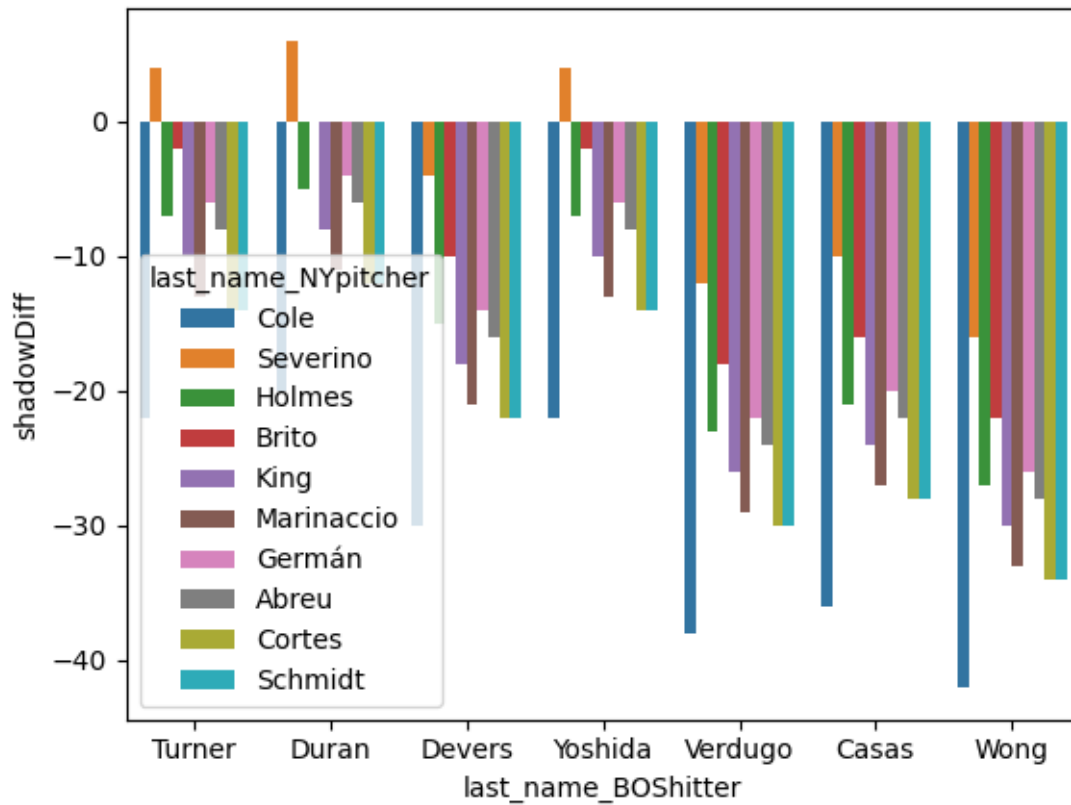
[70 rows x 1 columns]

```

[31]: sns.barplot(x="last_name_BOShitter",
    y="shadowDiff",
    hue="last_name_NYpitcher",
    data=bos_swingtake_vs_nypit)

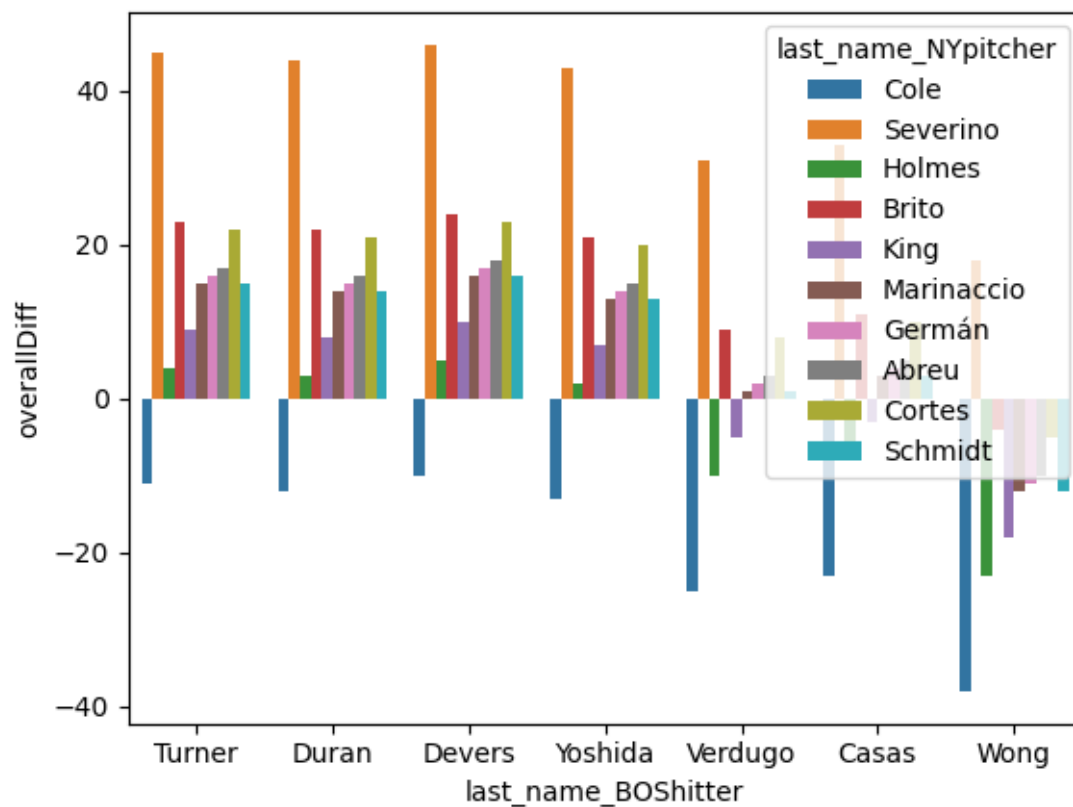
```

```
[31]: <Axes: xlabel='last_name_BOShitter', ylabel='shadowDiff'>
```



```
[32]: sns.barplot(x="last_name_BOShitter",
                  y="overallDiff",
                  hue="last_name_NYpitcher",
                  data=bos_swingtake_vs_nypit)
```

```
[32]: <Axes: xlabel='last_name_BOShitter', ylabel='overallDiff'>
```

```
[33]: bos_swingtake_vs_nypit.groupby(['last_name_NYpitcher',
    ↪ 'last_name_BOShitter'])['heartDiff'].agg(np.mean).
    ↪sort_values(ascending=False).to_frame()
```

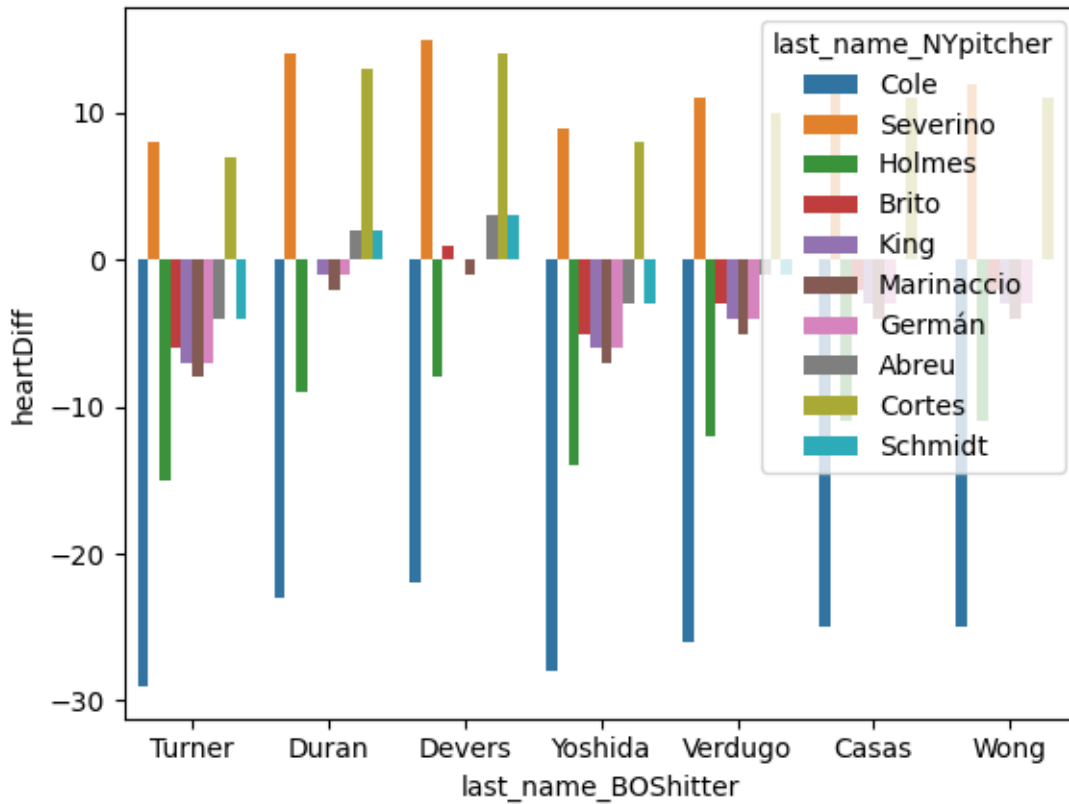
```
[33]:
```

last_name_NYpitcher	last_name_BOShitter	heartDiff
Severino	Devers	15.0
Cortes	Devers	14.0
Severino	Duran	14.0
Cortes	Duran	13.0
Severino	Wong	12.0
...
Cole	Wong	-25.0
	Casas	-25.0
	Verdugo	-26.0
	Yoshida	-28.0
	Turner	-29.0

[70 rows x 1 columns]

```
[34]: sns.barplot(x="last_name_BOShitter",
                  y="heartDiff",
                  hue="last_name_NYpitcher",
                  data=bos_swingtake_vs_nypit)
```

```
[34]: <Axes: xlabel='last_name_BOShitter', ylabel='heartDiff'>
```



```
[35]: bos_swingtake_vs_nypit.groupby('last_name_BOShitter')['overallDiff'].agg(np.
      ↳mean).sort_values(ascending=False)
```

```
[35]: last_name_BOShitter
Devers      16.5
Turner      15.5
Duran       14.5
Yoshida     13.5
Casas        3.5
Verdugo      1.5
Wong       -11.5
Name: overallDiff, dtype: float64
```

```
[36]: bos_swingtake_vs_nypit.groupby('last_name_BOShitter')['heartDiff'].agg(np.mean).  
      ↪sort_values(ascending=False)
```

```
[36]: last_name_BOShitter  
      Devers      0.5  
      Duran     -0.5  
      Casas     -2.5  
      Wong      -2.5  
      Verdugo   -3.5  
      Yoshida   -5.5  
      Turner    -6.5  
      Name: heartDiff, dtype: float64
```

```
[37]: bos_swingtake_vs_nypit.groupby('last_name_BOShitter')['shadowDiff'].agg(np.  
      ↪mean).sort_values(ascending=False)
```

```
[37]: last_name_BOShitter  
      Duran      -7.2  
      Turner     -9.2  
      Yoshida    -9.2  
      Devers    -17.2  
      Casas    -23.2  
      Verdugo   -25.2  
      Wong     -29.2  
      Name: shadowDiff, dtype: float64
```

```
[38]: bos_swingtake_vs_nypit.groupby('last_name_BOShitter')['chaseDiff'].agg(np.mean).  
      ↪sort_values(ascending=False)
```

```
[38]: last_name_BOShitter  
      Turner     20.1  
      Devers     18.1  
      Verdugo    18.1  
      Casas     17.1  
      Yoshida    16.1  
      Wong       11.1  
      Duran      10.1  
      Name: chaseDiff, dtype: float64
```

```
[39]: boshIT_pitch_ars = pd.read_csv('pitch-arsenal-stats-rs-bat.csv')
```

```
[40]: bos_ny_pitchars = boshIT_pitch_ars.merge(ny_pitchars, how='cross',  
      ↪suffixes=("_BOShitter", "_NYpitcher"))
```

```
[41]: bos_ny_pitchars['RVDiff'] = bos_ny_pitchars['run_value_BOShitter'] -  
      ↪bos_ny_pitchars['run_value_NYpitcher']
```

```
[42]: grouped_pitchers = bos_ny_pitchers.groupby(['last_name_BOShitter',
↳ 'last_name_NYpitcher', 'pitch_name_NYpitcher',
↳ 'pitch_name_BOShitter'])[['RVDiff']].agg(np.mean).reset_index()
```

```
[43]: for i in range(0, len(grouped_pitchers)):
    if grouped_pitchers.loc[i, 'pitch_name_BOShitter'] != grouped_pitchers.
↳ loc[i, 'pitch_name_NYpitcher']:
        grouped_pitchers.drop(i, inplace=True)
```

```
[45]: len(grouped_pitchers)
```

```
[45]: 912
```

```
[48]: grouped_pitchers = bos_ny_pitchers.groupby(['last_name_BOShitter',
↳ 'last_name_NYpitcher', 'pitch_name_NYpitcher'])[['RVDiff']].agg(np.mean).
↳ reset_index()
```

```
[49]: grouped_pitchers
```

```
[49]:
```

	last_name_BOShitter	last_name_NYpitcher	pitch_name_NYpitcher	RVDiff
0	Alfaro	Abreu	Changeup	0.00
1	Alfaro	Abreu	Sinker	-2.00
2	Alfaro	Abreu	Slider	-4.00
3	Alfaro	Brewer	Cutter	-4.00
4	Alfaro	Brito	4-Seamer	4.00
...
1151	Yoshida	Vásquez	Sinker	-0.25
1152	Yoshida	Weber	Changeup	1.75
1153	Yoshida	Weber	Sinker	-3.25
1154	Yoshida	Weber	Slider	3.75
1155	Yoshida	Weissert	Sweeper	4.75

```
[1156 rows x 4 columns]
```

```
[50]: grouped_pitchers.to_csv('grouped_bos_ny_pitchers.csv')
```

```
[ ]:
```