

Notebook Contents

- Imports
 - Data
 - Index DataFrame to Get Pitch Types
 - Pitcher Overview
 - General Pitch Data
 - Pitch Usage
 - Pitch Usage by Batter Handedness
 - Velocity by Pitch Type
 - Pitch Velocity by Inning
 - Pitcher Stuff
 - Spin Rate by Pitch Type
 - Velocity & Spin Rate
 - Spin Axis
 - HB & VB Axis
 - HB & VB Due to Magnus Force
 - Release Position
 - Release Extension
 - Release Extension
 - Avrg. Bauer Units by Pitch Type
 - Count Breakdown
 - Pitch Usage by Count
 - Heatmaps
 - Pitch Location by Pitch Type
 - Quality Contact Heatmap - All Pitches
 - Slider Heatmaps
 - 4-Seam Heatmaps
 - Curveball Heatmaps
 - Changeup Heatmaps

Imports

```
In [1]: import math
import matplotlib.patches as mpatches
from matplotlib.patches import Rectangle
from matplotlib.ticker import PercentFormatter
import warnings
warnings.filterwarnings('ignore')
```

Data

```
In [2]: kershaw = pd.read_csv('../data/clayton-kershaw.csv')
kershaw.drop(columns = ['Unnamed: 0'], inplace = True)

# Font Dictionary
font_title = {
    'size': 14,
    'weight': 'bold',
    'verticalalignment': 'center_baseline',
    'horizontalalignment': 'center'
}

pd.set_option('max_columns', None)
print(kershaw.shape)
kershaw.head(2)
```

	pitch_type	game_date	release_speed	release_pos_x	release_pos_z	player_name	batter	pitcher	events	description	zone
0	SL	2021-05-30	87.5	1.6	6.13	Kershaw, Clayton	592332	477132	strikeout	called_strike	8
1	FF	2021-05-30	90.4	1.6	6.12	Kershaw, Clayton	592332	477132	N/A	called_strike	6

```
In [3]: gen_data = kershaw[['pitch_type', 'release_speed', 'release_spin_rate',
    'true_spin', 'spin_eff', 'phi', 'theta', 'pfx_e', 'pfx_s', 'pfx_x', 'pfx_z', 'is_strike', 'release_pos_x', 'release_pos_z', 'bauer_units']]

col_dict = {
    'release_speed': 'velo', 'release_spin_rate': 'spin', 'phi': 'spin_axis', 'theta': 'gyro_spin_deg',
    'pfx_e': 'ah', 'pfx_s': 'vb', 'is_strike': 'strike', 'release_pos_x': 'r_height',
    'release_pos_z': 'r_side'
}

gen_data.rename(columns = col_dict, inplace = True)
```

Index DataFrame to Get Pitch Types

```
In [4]: # kershaw.pitch_type.value_counts(normalize = True)
r_kershaw = kershaw.loc[kershaw['stand'] == 'R']
l_kershaw = kershaw.loc[kershaw['stand'] == 'L']

# Hitters
sl = kershaw.loc[kershaw['pitch_type'] == 'SL']
ff = kershaw.loc[kershaw['pitch_type'] == 'FF']
cu = kershaw.loc[kershaw['pitch_type'] == 'CU']
ch = kershaw.loc[kershaw['pitch_type'] == 'CH']

# Batters
r_sl = r_kershaw.loc[l_kershaw['pitch_type'] == 'SL']
r_ff = r_kershaw.loc[l_kershaw['pitch_type'] == 'FF']
r_cu = r_kershaw.loc[l_kershaw['pitch_type'] == 'CU']
r_ch = r_kershaw.loc[l_kershaw['pitch_type'] == 'CH']

# LHH
l_sl = l_kershaw.loc[l_kershaw['pitch_type'] == 'SL']
l_ff = l_kershaw.loc[l_kershaw['pitch_type'] == 'FF']
l_cu = l_kershaw.loc[l_kershaw['pitch_type'] == 'CU']
l_ch = l_kershaw.loc[l_kershaw['pitch_type'] == 'CH']
order = ['SL', 'FF', 'CU', 'CH']
```

Pitcher Overview

General Pitch Data

```
In [6]: gen_data.groupby(['pitch_type'], sort = False).mean()
```

pitch_type	velo	spin	true_spin	spin_eff	spin_axis	gyro_spin_deg	hb	vb	strike	r_side
SL	87.075000	2654.045402	501.376237	0.192000	223.276316	78.878947	8.182326	4.712372	0.734884	1.513756
FF	90.636707	2505.066566	1532.492784	0.624158	184.323024	51.099656	19.353897	0.525861	0.641994	1.525317
CU	74.239927	2535.468864	1372.876019	0.558889	346.722222	55.777778	-15.283916	4.782857	0.538462	1.178608
CH	87.045456	2163.272727	1684.480000	0.780000	137.363636	38.000000	15.665455	-12.992727	0.272727	1.790909

Slider: combo sweeping slider/gyro ball movement profile

Fastball: more of a Cutter/4-Seam profile

Curveball: True 12-12 banger

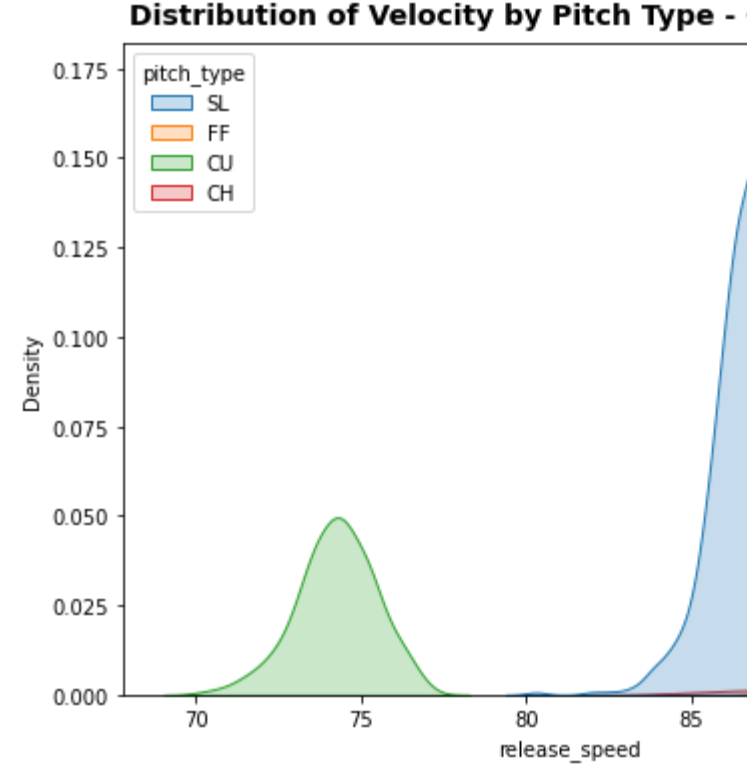
Change: straight change

Pitch Usage

```
In [7]: plt.figure(figsize = (8, 6))

dist = round(kershaw.pitch_type.value_counts(normalize = True), 2)
color = sns.color_palette('coolwarm_r')
plt.pie(dist, labels = order, colors = color, autopct = '%.0f%')
plt.title('Distribution of Pitch Types - Clayton Kershaw', fontdict = font_title, pad = 15);
```

Distribution of Pitch Types - Clayton Kershaw

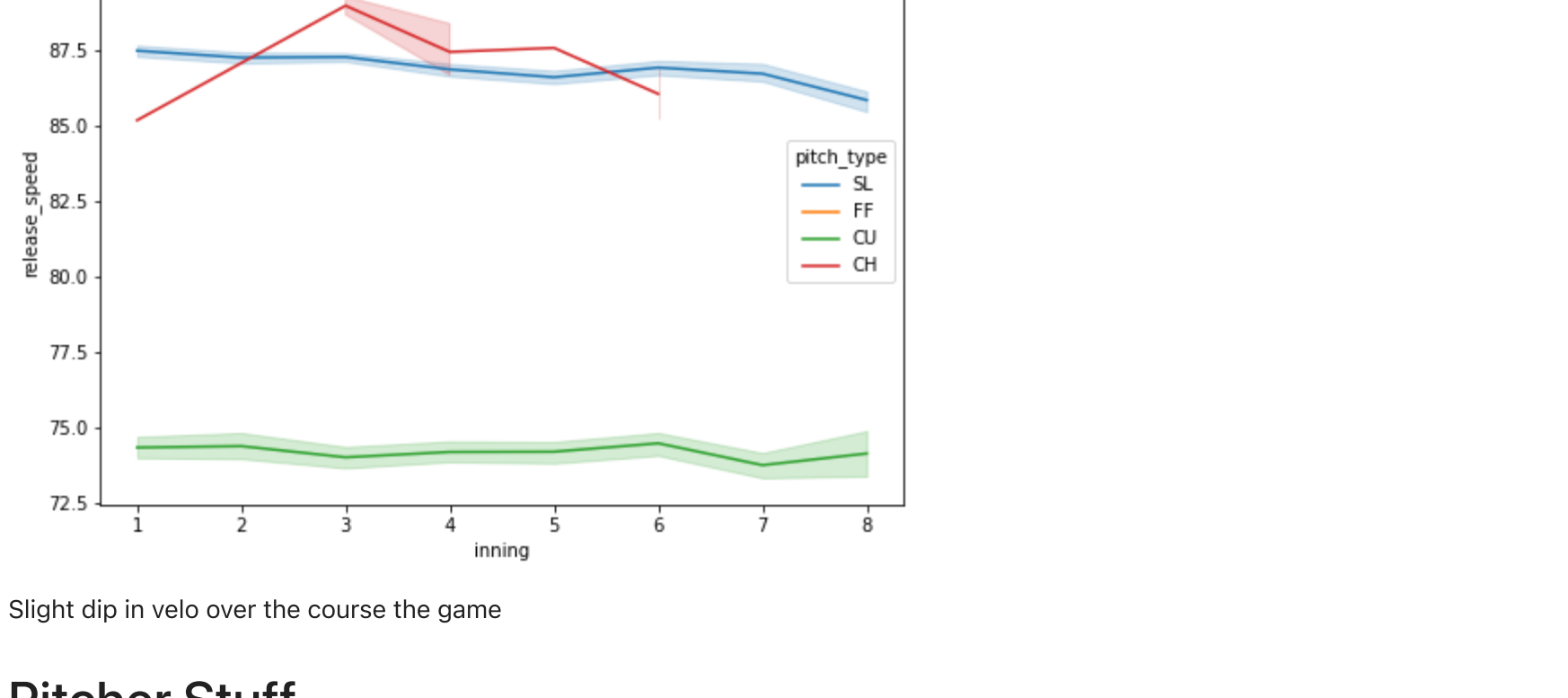


Pitch Usage by Batter Handedness

```
In [8]: blue = '#002072'
red = '#D60000'
fig, axs = plt.subplots(1, 2, figsize = (20, 6))
fig.suptitle('Pitch Usage by Batter Handedness', fontdict = font_title, pad = 15)

dist_r = r_kershaw.pitch_type.sort_values(ascending = False)
axs[0].hist(dist_r, weights = np.ones(len(dist_r)) / len(dist_r), color = blue)
axs[0].set_title('Distribution of Pitch Types - RHH', fontdict = font_title, pad = 15)

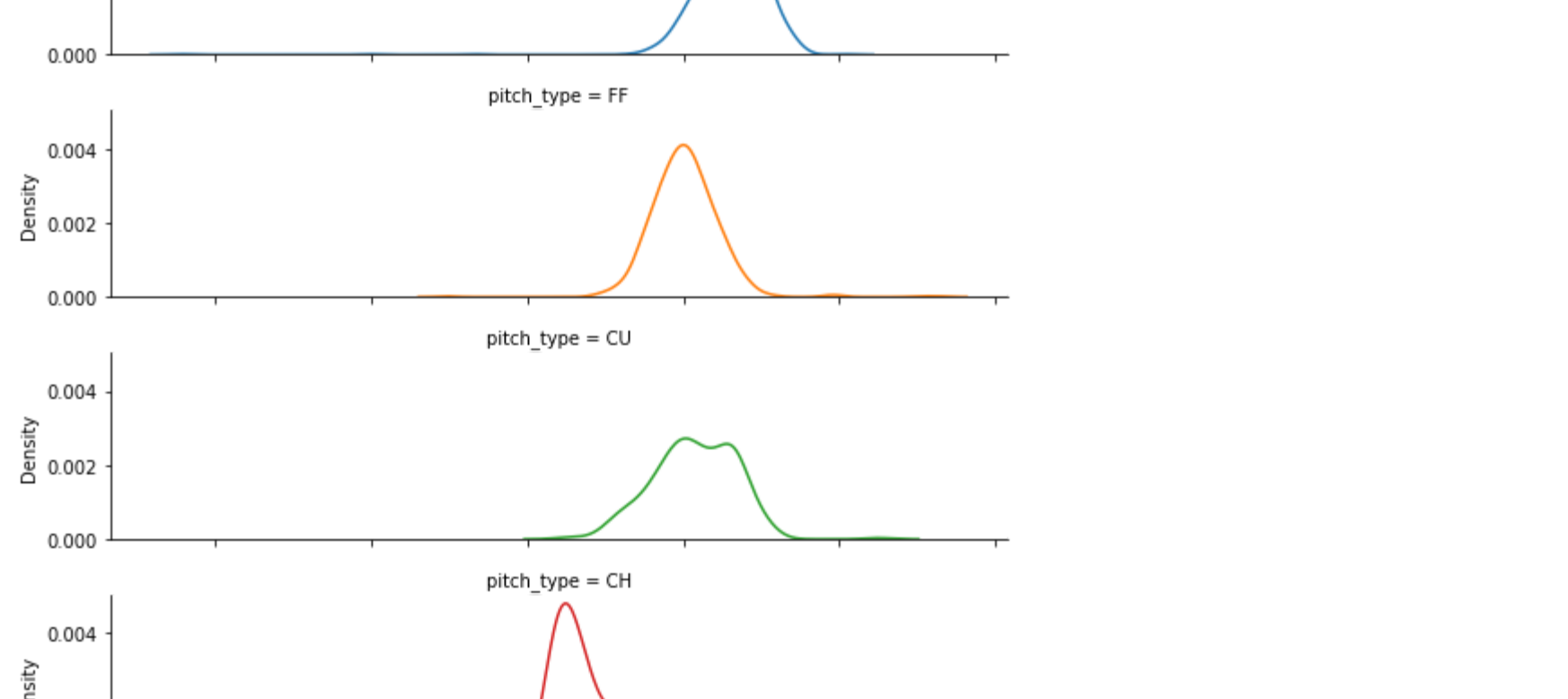
dist_l = l_kershaw.pitch_type.sort_values(ascending = False)
axs[1].hist(dist_l, weights = np.ones(len(dist_l)) / len(dist_l), color = red)
axs[1].set_title('Distribution of Pitch Types - LHH', fontdict = font_title, pad = 15);
```



Velocity by Pitch Type

```
In [9]: plt.figure(figsize = (8, 6))

ax = sns.kdeplot(data = kershaw, x = 'release_speed', shade = 'fill', hue = 'pitch_type',
    hue_order = order, palette = 'tab10')
sns.move_legend(ax, 'upper left')
plt.title('Distribution of Velocity by Pitch Type - Clayton Kershaw', fontdict = font_title, pad = 15);
```



Slider: 86 - 88 mph (league avg: 85 mph)

Cutter / 4-Seam: Sits 90 - 91 mph (league avg: 92.3 mph)

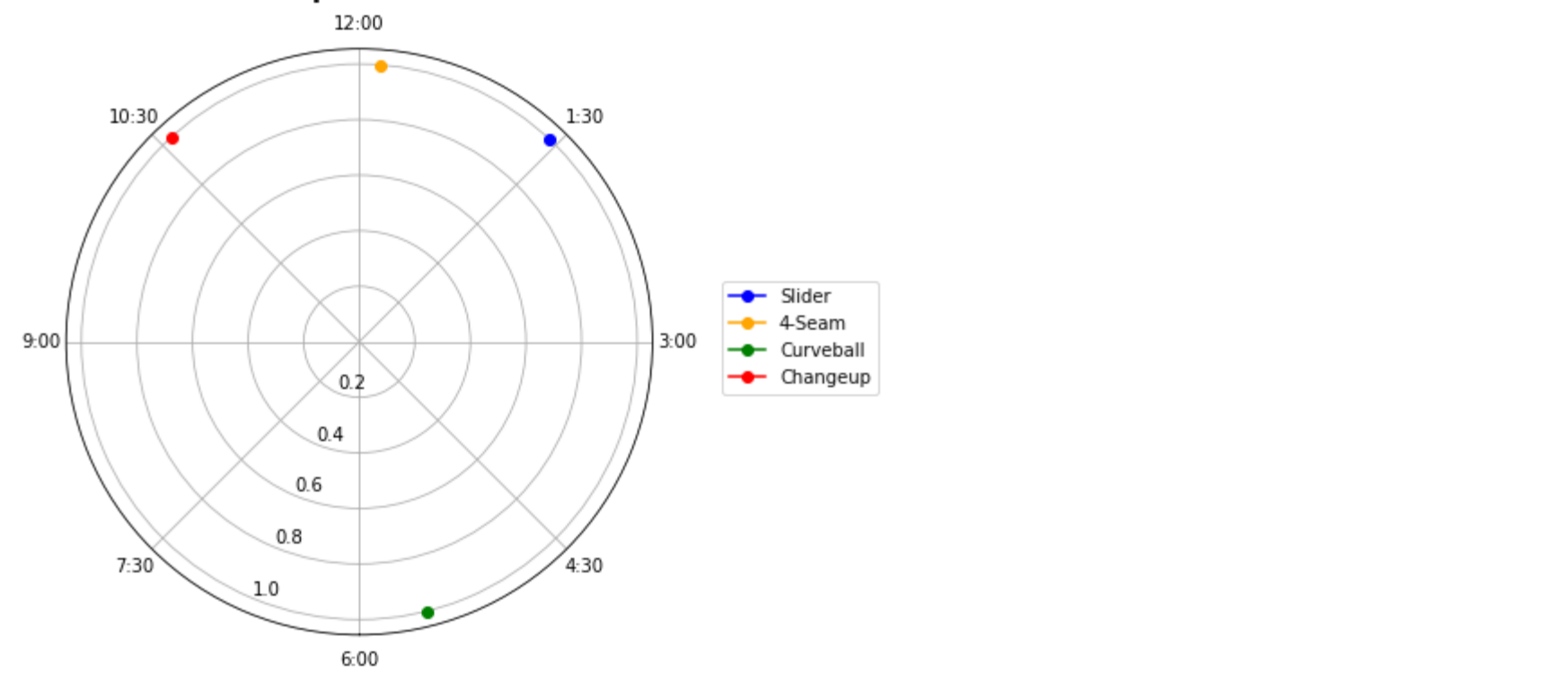
Curveball: Sits 74 - 75 mph (league avg: 77 mph (11 - 18 mph slower compared to FB))

Changeup: Sits 86 - 88 mph (league avg: 85 mph (8 - 12 mph difference from FB))

Pitch Velocity by Inning

```
In [10]: plt.figure(figsize = (8, 6))

sns.lineplot(data = kershaw, x = 'inning', y = 'release_speed', hue = 'pitch_type',
    hue_order = order, palette = 'tab10')
plt.title('Pitch Velocity by Inning', fontdict = font_title, pad = 15);
```

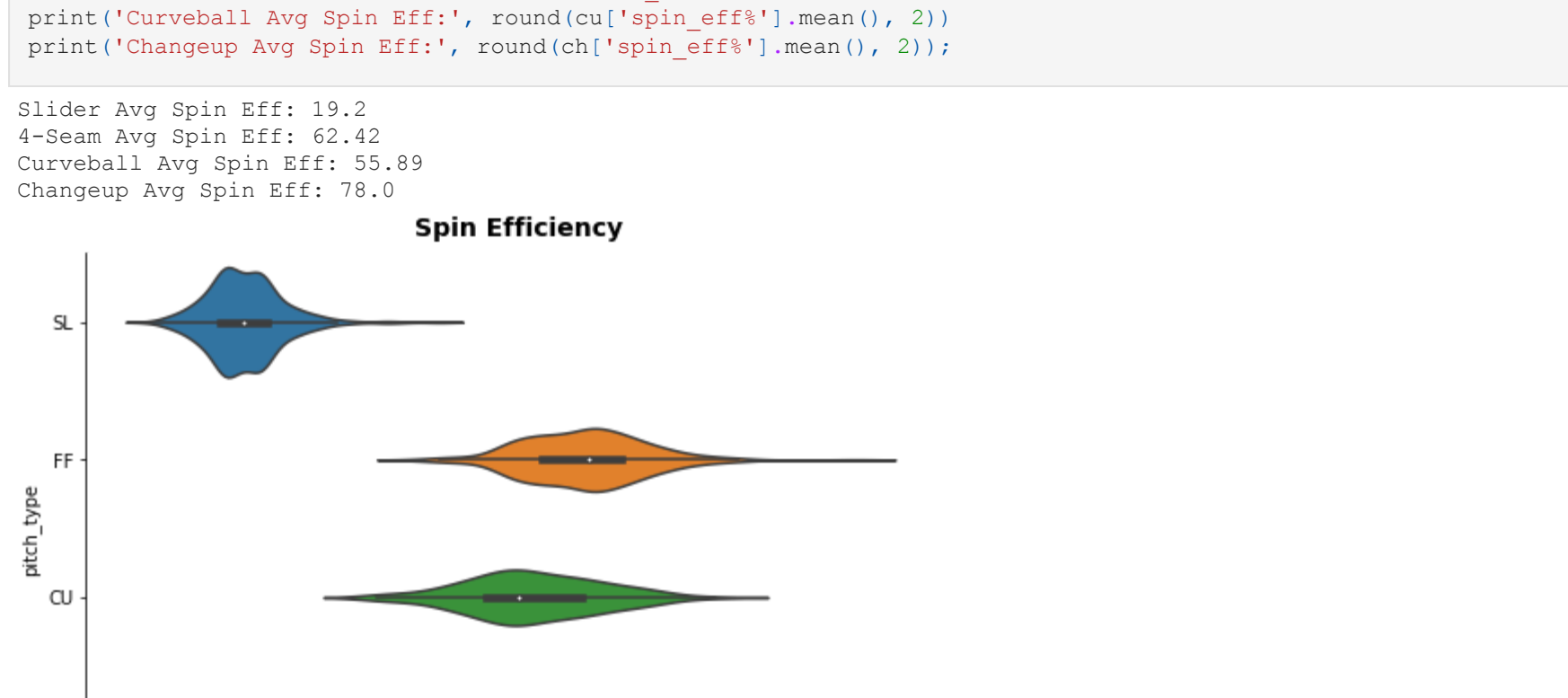


Slight dip in velo over the course of the game

Pitcher Stuff

Spin Rate by Pitch Type

```
In [11]: g = sns.FacetGrid(kershaw, row = 'pitch_type', hue = 'pitch_type', height = 2, aspect = 4, )
g.map(sns.kdeplot, 'release_spin_rate', palette = 'tab10')
```



Slider: Above avg spin (league avg: ~2450 rpm)

4-Seam: Above avg spin (league avg: ~2200 rpm)

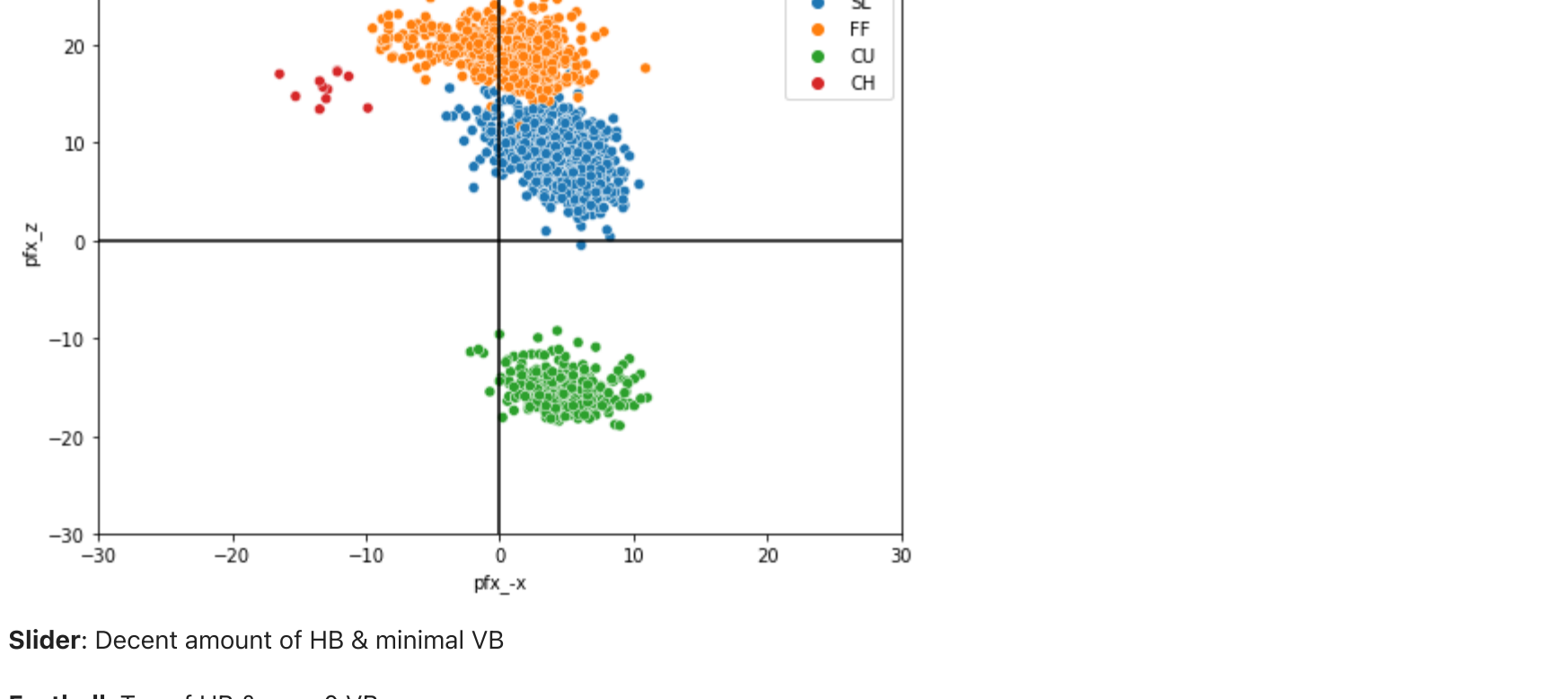
Curveball: Above average spin (league avg: ~2500 rpm)

Changeup: Above average spin (league avg range: ~1200 2400 rpm)

Spin Axis

```
In [12]: ax = plt.figure(figsize = (8, 6))

ax = plt.subplot(polar = True, theta_direction = 'l')
ax.plot(math.radians(sl_tilt), 1, color = 'blue', marker = 'o', label = 'Slider')
ax.plot(math.radians(ff_tilt), 1, color = 'orange', marker = 'o', label = '4-Seam')
ax.plot(math.radians(cu_tilt), 1, color = 'green', marker = 'o', label = 'Curveball')
ax.plot(math.radians(ch_tilt), 1, color = 'red', marker = 'o', label = 'Changeup')
ticks = ['6:00', '7:30', '9:00', '10:30', '12:00', '1:30', '3:00', '4:30']
ax.set_xticklabels(ticks, ax.legend(bbox_to_anchor=(1.4, 62)), ax.set_theta_zero_location='S')
ax.set_title('Spin Axis', fontdict = font_title, pad = 15);
```



Pairs: 4-Seam / Curveball pair extremely well. 4-Seam / Slider pair well

Slider: combo sweeping slider/gyro ball movement profile Sweeping slider (side spin): 2:00 - 4:00 L spin tilt Decent amount of HB and near 0 VB Sweeping slider: 2:0 - 4:0% Gyro ball: 0 - 20%

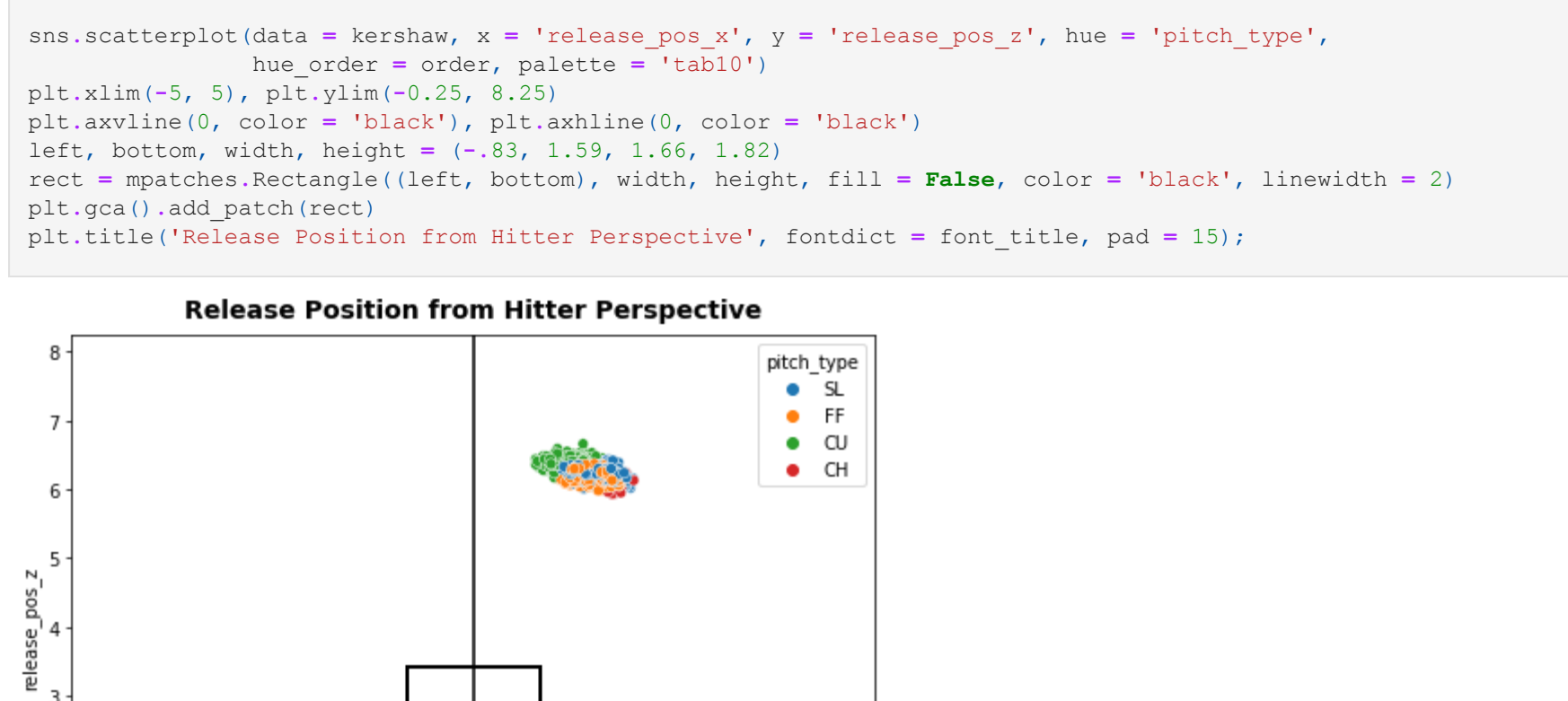
Fastball: more of a Cutter/4-Seam profile Cutter (side + back spin): 12:00 - 1:30 L spin tilt Positive VB & Some HB Hybrid between slider and FB, typically harder than slider

Curveball: True 12-12 banger True 12 - 6 CB: 5:00 - 6:00 L spin tilt Much more negative VB and minimal HB Curveballs thrown with low spin efficiency (below 70% efficiency) Movement profile gets closer to (0, 0) Higher spin efficiency Movement profile gets further from (0, 0)

Change: straight change Straight change: 10:30 - 12:00 L spin tilt More VB & less HB Spin rate: 1700 - 1800 rpm Spin efficiency: 90% Arm slot: overhand

Spin Efficiency

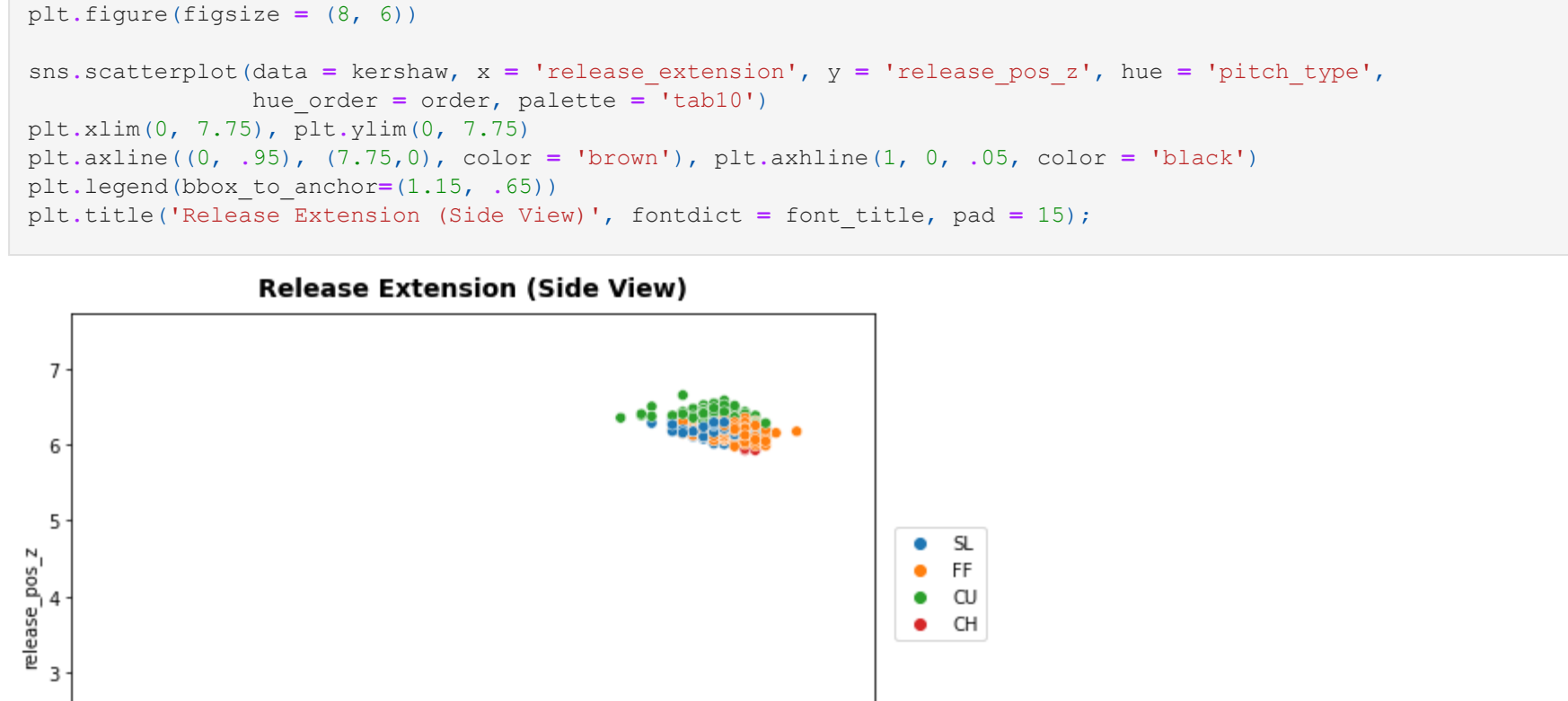
```
In [13]: sns.catplot(data = kershaw, x = 'spin_eff', y = 'pitch_type', kind = 'violin', aspect = 1.5, palette = 'tab10')
sns.catplot(ax = axs[0][0].data, ax.set_title('Slider Avg Spin Eff: 19.2')
sns.catplot(ax = axs[1][0].data, ax.set_title('4-Seam Avg Spin Eff: 62.42')
sns.catplot(ax = axs[2][0].data, ax.set_title('Curveball Avg Spin Eff: 55.89')
sns.catplot(ax = axs[3][0].data, ax.set_title('Changeup Avg Spin Eff: 78.0')
```



Horizontal & Vertical Break Axis

```
In [14]: plt.figure(figsize = (8, 6))

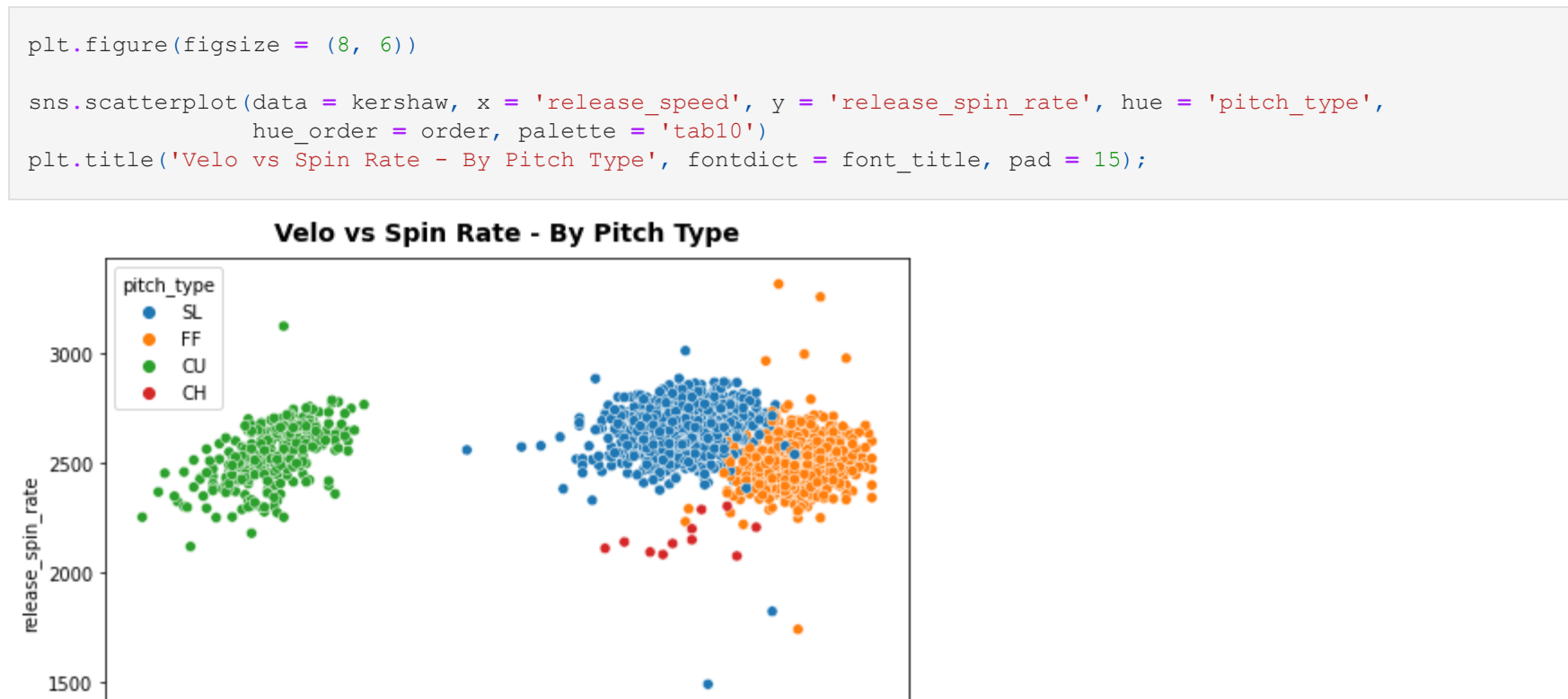
sns.scatterplot(data = kershaw, x = 'pfx_x', y = 'pfx_z', hue = 'pitch_type',
    hue_order = order, palette = 'tab10')
plt.xlim(-30, 30), plt.ylim(-30, 30)
plt.axhline(0, color = 'black'), plt.axhline(0, color = 'black')
plt.title('HB & VB Axis', fontdict = font_title, pad = 15);
```



Horizontal & Vertical Break Due to Magnus Force Axis

```
In [15]: plt.figure(figsize = (8, 6))

sns.scatterplot(data = kershaw, x = 'Mx', y = 'My', hue = 'pitch_type',
    hue_order = order, palette = 'tab10')
plt.xlim(-30, 30), plt.ylim(-30, 30)
plt.axhline(0, color = 'black'), plt.axhline(0, color = 'black')
plt.title('HB & VB from Magnus Force Axis', fontdict = font_title, pad = 15);
```

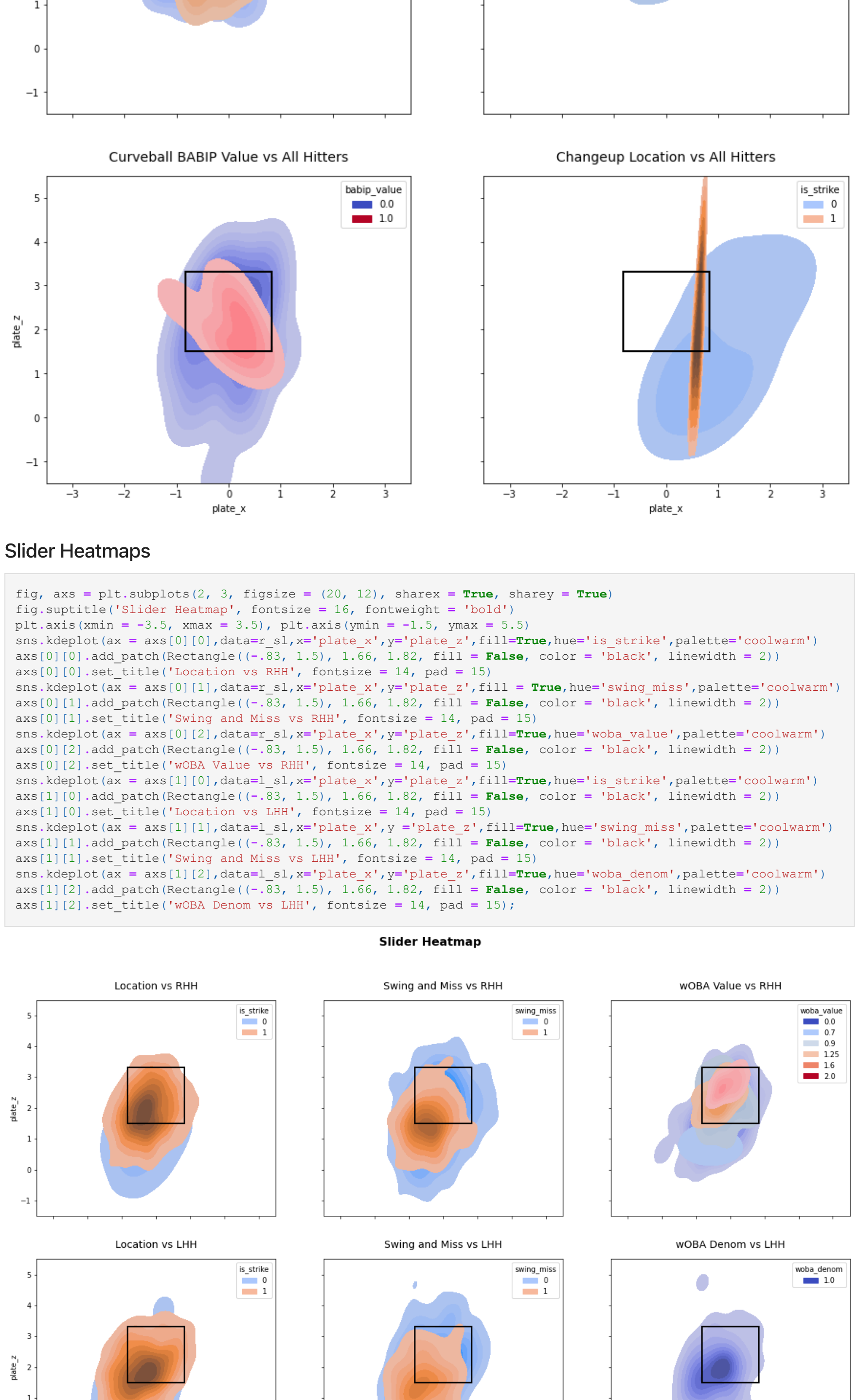


Launch Speed Angle

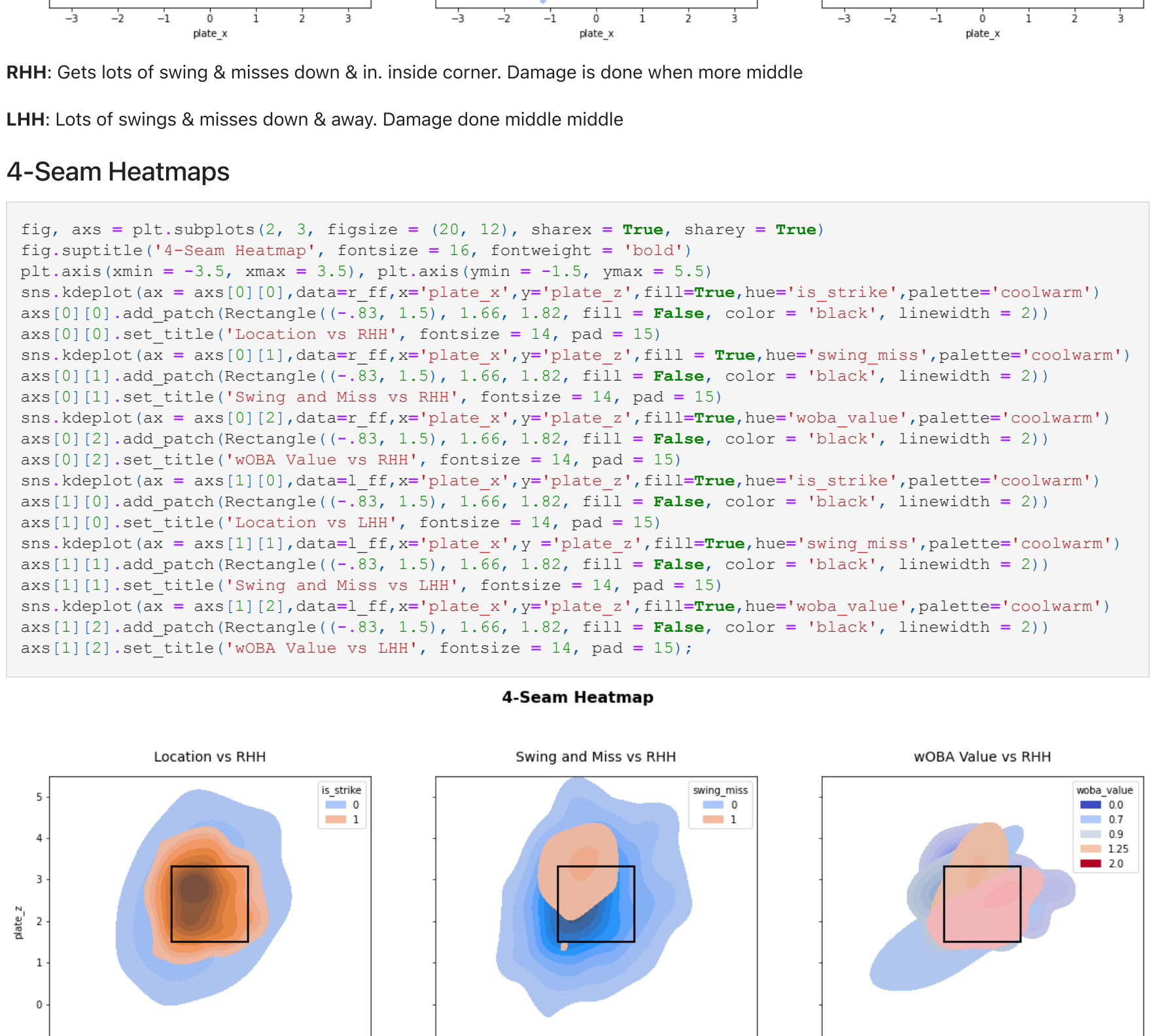
1: weak 2: topped 3: under 4: flare/burner 5: solid contact 6: barrel


```
fig, axs = plt.subplots(2, 2, figsize = (15, 15), sharex = True, sharey = True)
fig.suptitle('Quality Contact by Pitch Type', fontsize = 16, fontweight = 'bold')
plt.axis(xmin = -3.5, xmax = 3.5), plt.axis(ymin = -1.5, ymax = 5.5)
sns.kdeplot(ax=axs[0][0], data=df, x='plate_x', y='plate_z', fill=True, hue='launch_speed_angle', palette='coolwarm')
axs[0][0].add_patch(Rectangle((-85, 1.5), 1.66, 1.82, fill = False, color = 'black', linewidth = 2))
axs[0][0].set_title('Slider Quality Contact vs All Hitters', fontsize = 14, pad = 15)
sns.kdeplot(ax = axs[0][1], data=df, x='plate_x', y='plate_z', fill = True, hue='launch_speed_angle', palette='coolwarm')
axs[0][1].add_patch(Rectangle((-85, 1.5), 1.66, 1.82, fill = False, color = 'black', linewidth = 2))
axs[0][1].set_title('Location vs RHH', fontsize = 14, pad = 15)
sns.kdeplot(ax = axs[0][2], data=df, x='plate_x', y='plate_z', fill=True, hue='woba_value', palette='coolwarm')
axs[0][2].add_patch(Rectangle((-85, 1.5), 1.66, 1.82, fill = False, color = 'black', linewidth = 2))
axs[0][2].set_title('Woba Value vs RHH', fontsize = 14, pad = 15)
sns.kdeplot(ax = axs[1][0], data=df, x='plate_x', y='plate_z', fill=True, hue='is_strike', palette='coolwarm')
axs[1][0].add_patch(Rectangle((-85, 1.5), 1.66, 1.82, fill = False, color = 'black', linewidth = 2))
axs[1][0].set_title('Swing and Miss vs RHH', fontsize = 14, pad = 15)
sns.kdeplot(ax = axs[1][1], data=df, x='plate_x', y='plate_z', fill=True, hue='babip_value', palette='coolwarm')
axs[1][1].add_patch(Rectangle((-85, 1.5), 1.66, 1.82, fill = False, color = 'black', linewidth = 2))
axs[1][1].set_title('Curveball BABIP Value vs All Hitters', fontsize = 14, pad = 15)
sns.kdeplot(ax=axs[1][2], data=df, x='plate_x', y='plate_z', fill=True, hue='is_strike', palette='coolwarm')
axs[1][2].add_patch(Rectangle((-85, 1.5), 1.66, 1.82, fill = False, color = 'black', linewidth = 2))
axs[1][2].set_title('Changeup Location vs All Hitters', fontsize = 14, pad = 15);
```

Quality Contact by Pitch Type



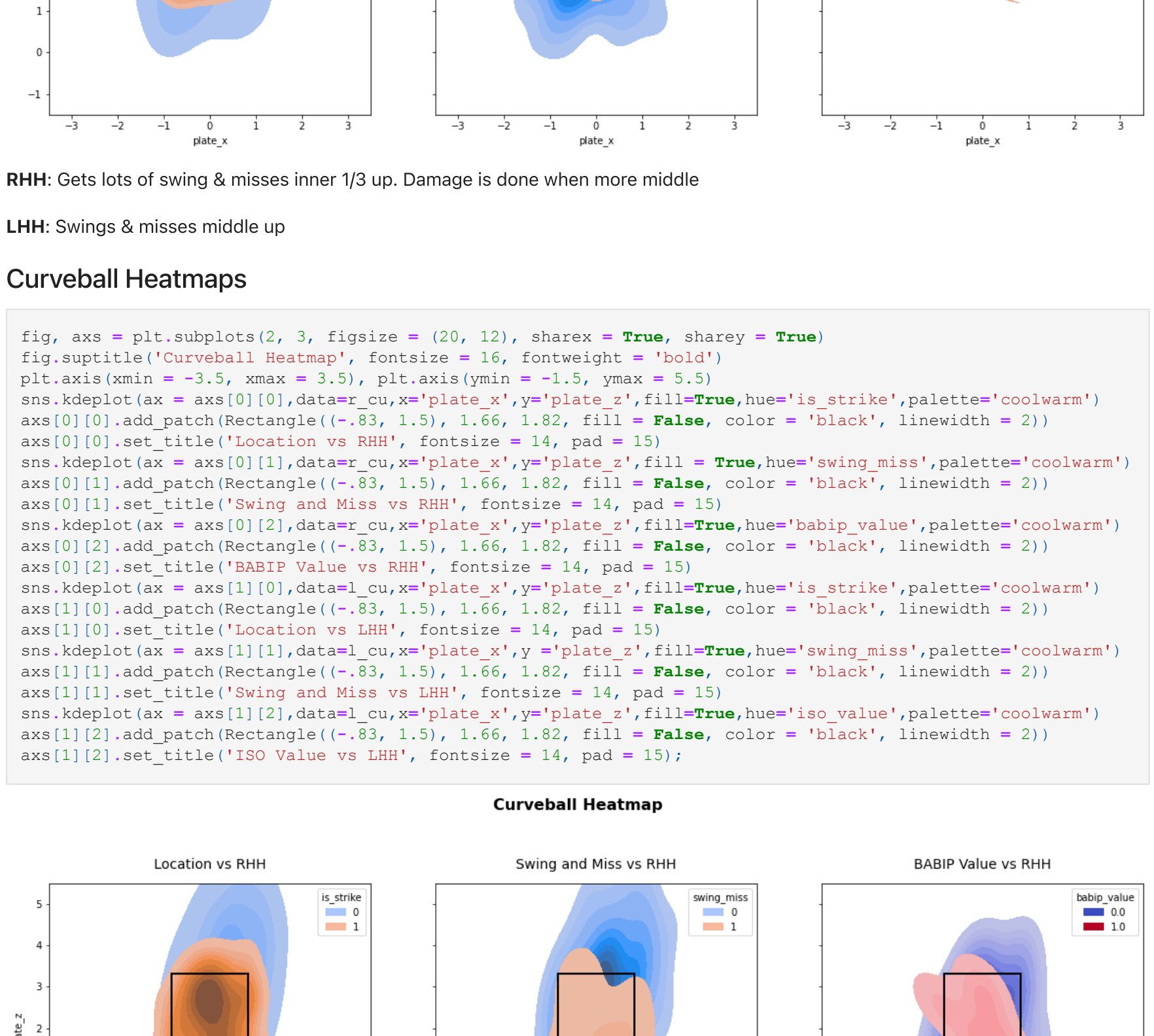
Slider Heatmaps



RHH: Gets lots of swing & misses down & in, inside corner. Damage is done when more middle

LHH: Lots of swings & misses down & away. Damage done middle middle

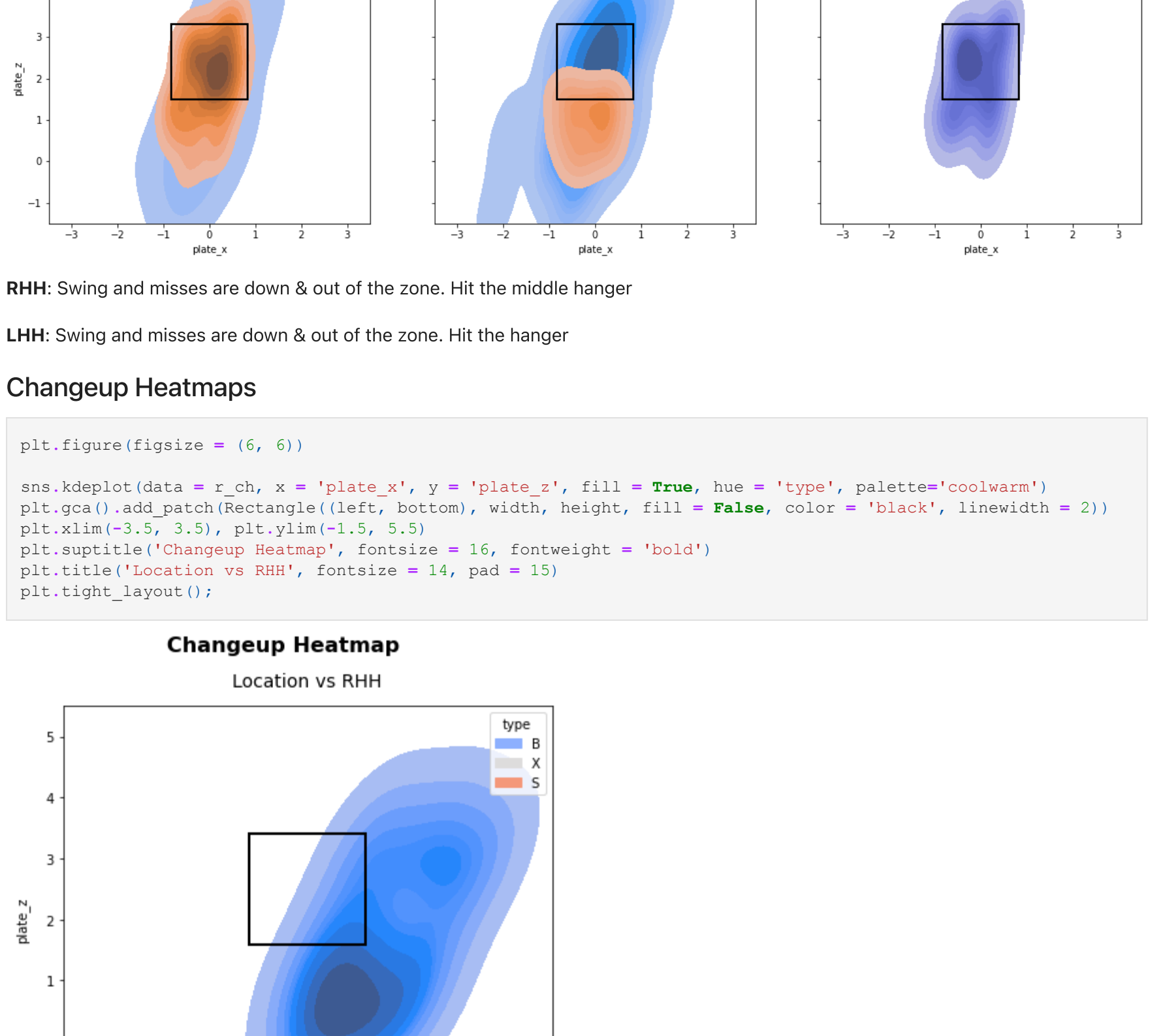
4-Seam Heatmaps



RHH: Gets lots of swing & misses inner 1/3 up. Damage is done when more middle

LHH: Swings & misses middle up

Curveball Heatmaps



RHH: Swing and misses are down & out of the zone. Hit the middle hanger

LHH: Swing and misses are down & out of the zone. Hit the hanger

Changeup Heatmaps



RHH: Barely thrown for strike. Sits down and out of the zone