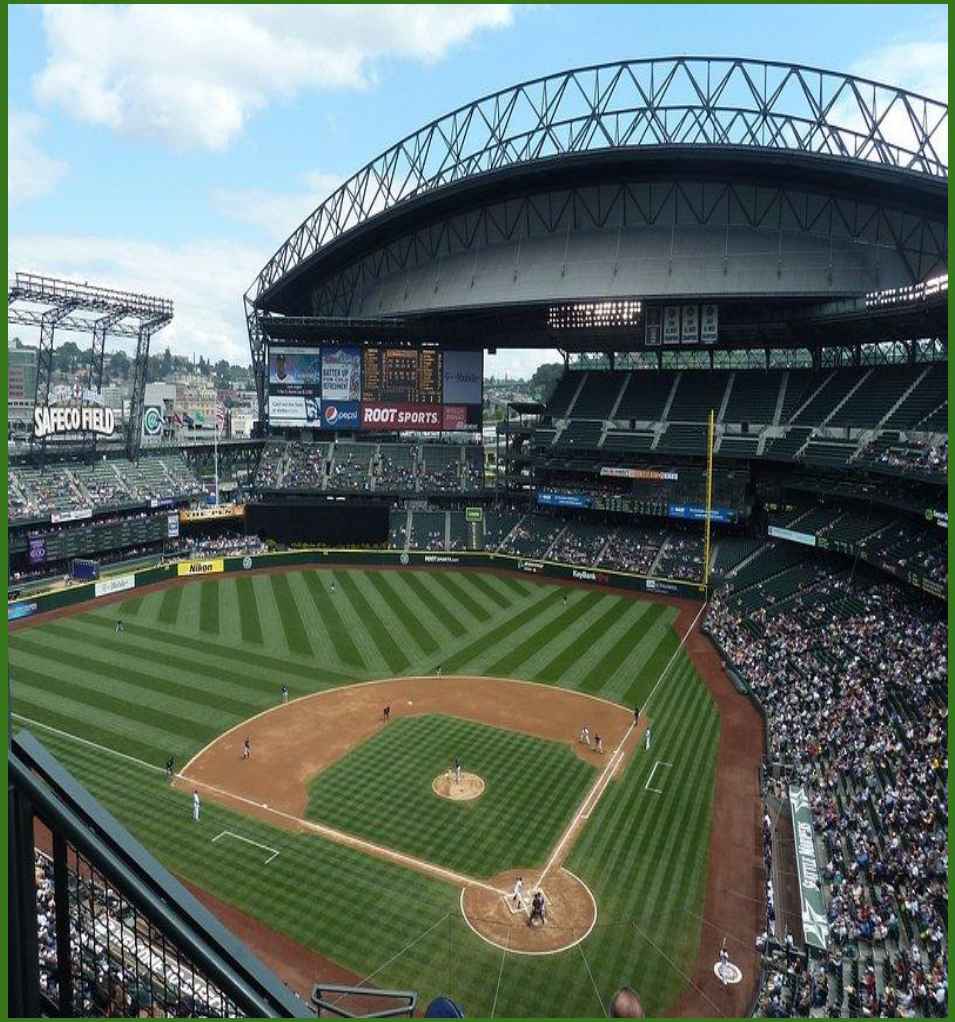
MLB Pitch Analysis

January 4, 2024



Dan Rosen

Github: dangrosen

Email: dan_rosen@outlook.com

Agenda

Business Problem

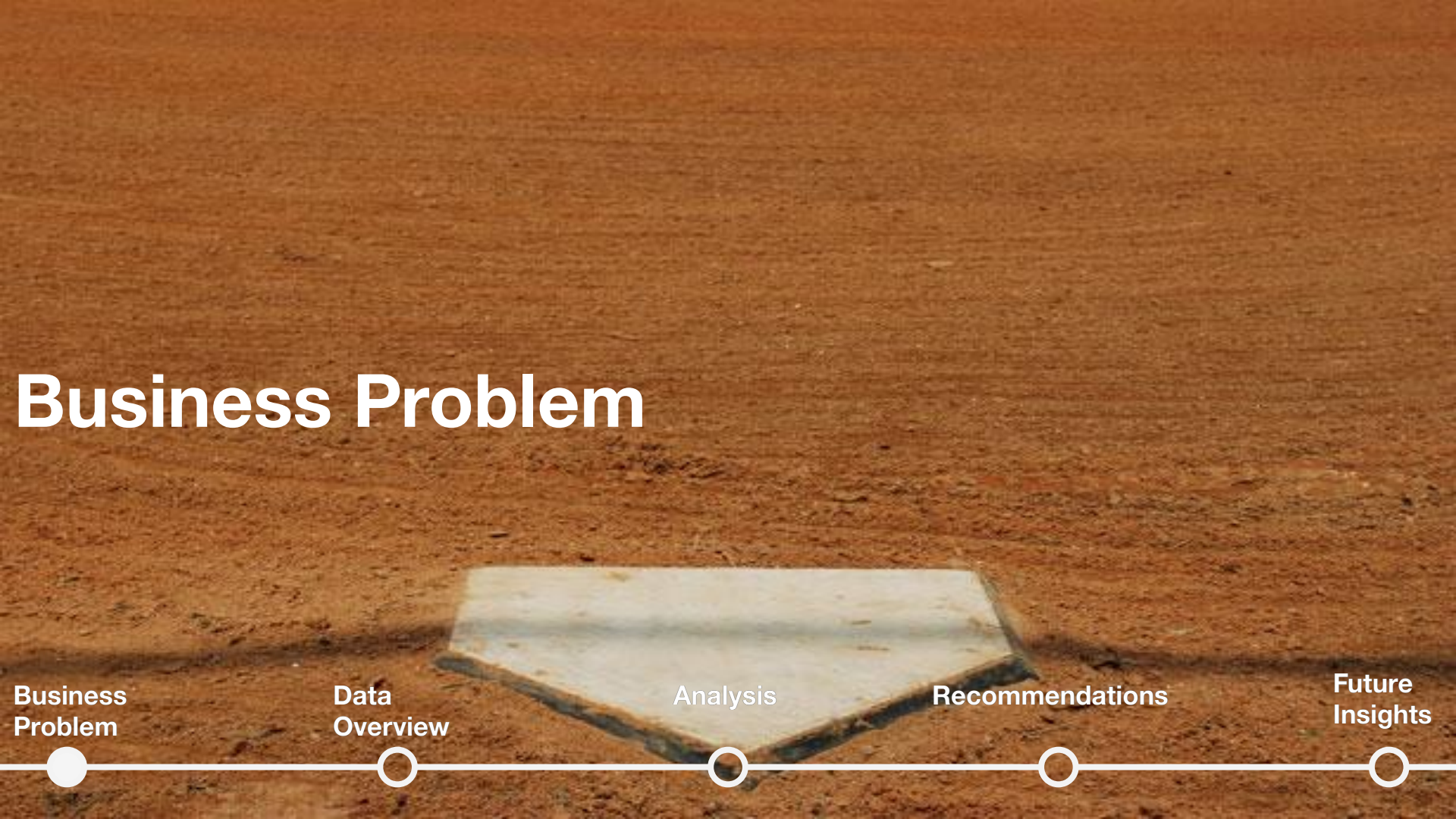
Data Overview

Analysis

Recommendations

Future Insights

Business Problem



Business Problem

- Look at different features of pitches
- Predict whether the batter makes an out or reaches base

Fastball

Finding

The model accurately predicts the outcome of an at bat 58% of the time.

Curveball

Data Overview

Business Problem | Data Overview | Analysis | Recommendations | Future Insights

Data Overview

Data
- ❏ 3.5 million pitches from 2015-2019
- ❏ Obtained from kaggle.com

Conditions
- ❏ Events grouped by out made or not
- ❏ Pitches with highest ratio of getting outs
- ❏ 283,887 records for modeling

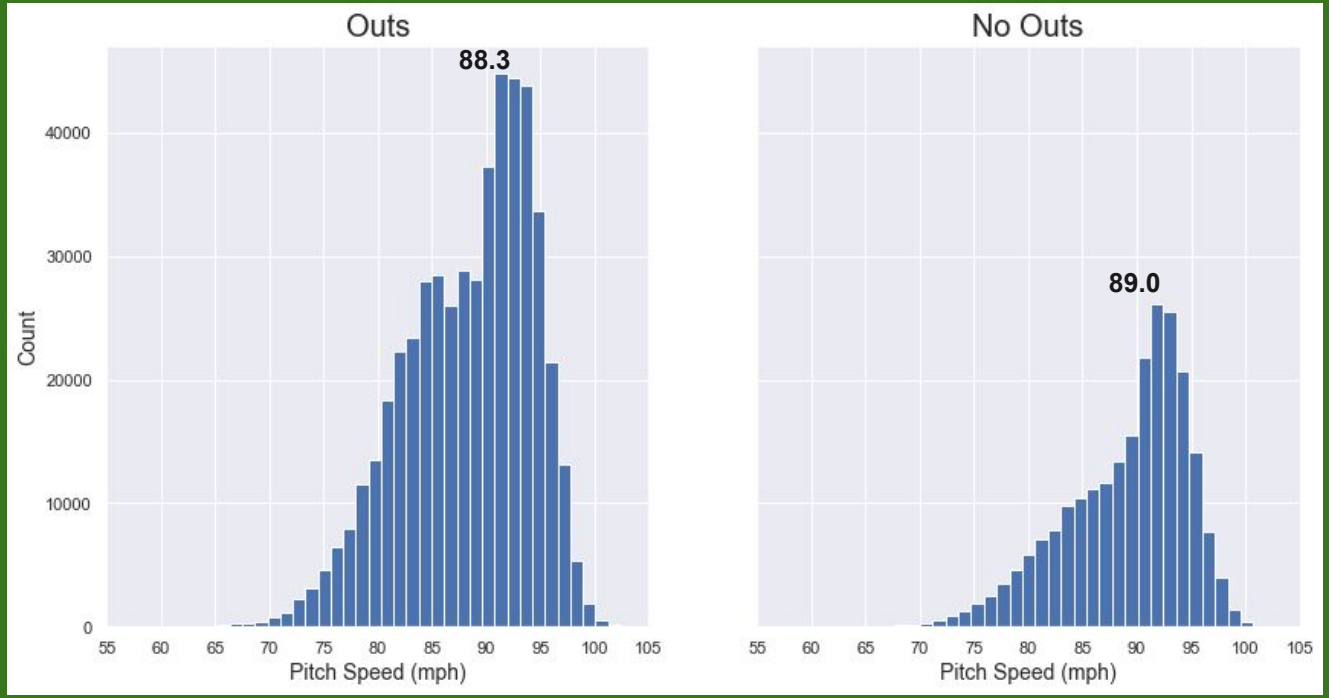
Limitations
- ❏ Explanation of some columns lacking
- ❏ 2019 data not recorded for some variables

Slider

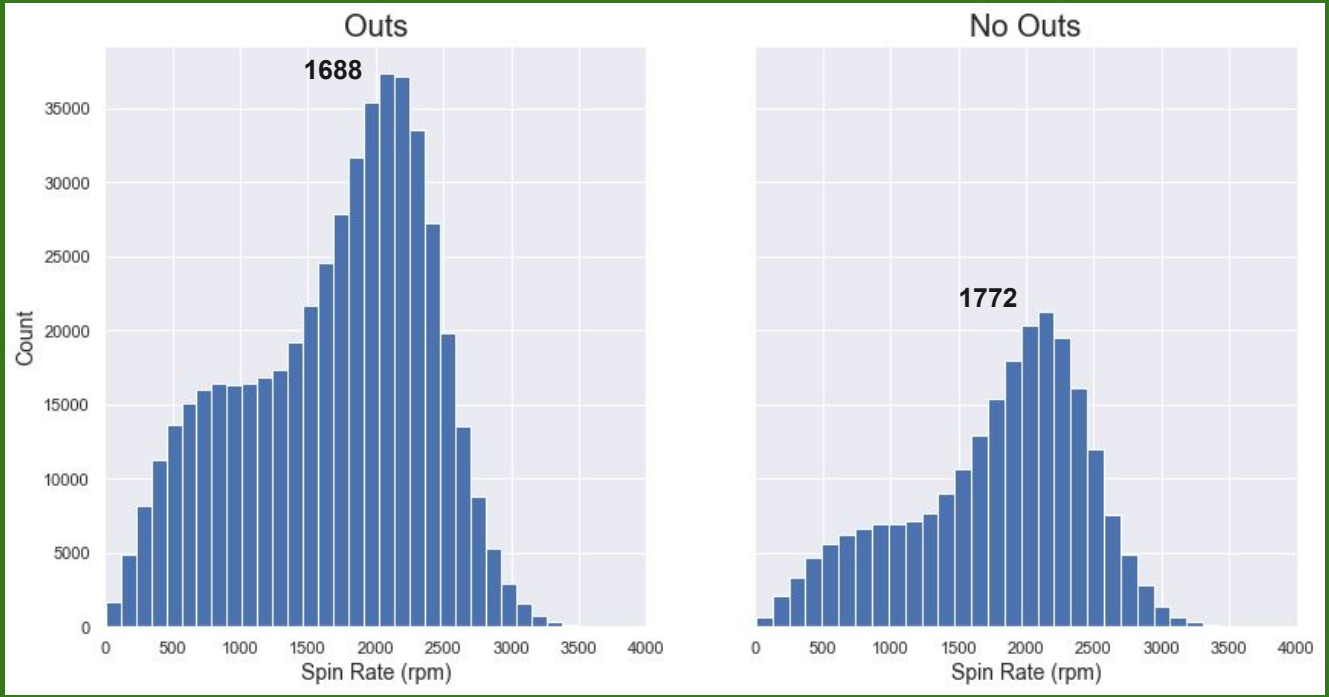
Analysis

Business Problem
Data Overview
Analysis
Recommendations
Future Insights

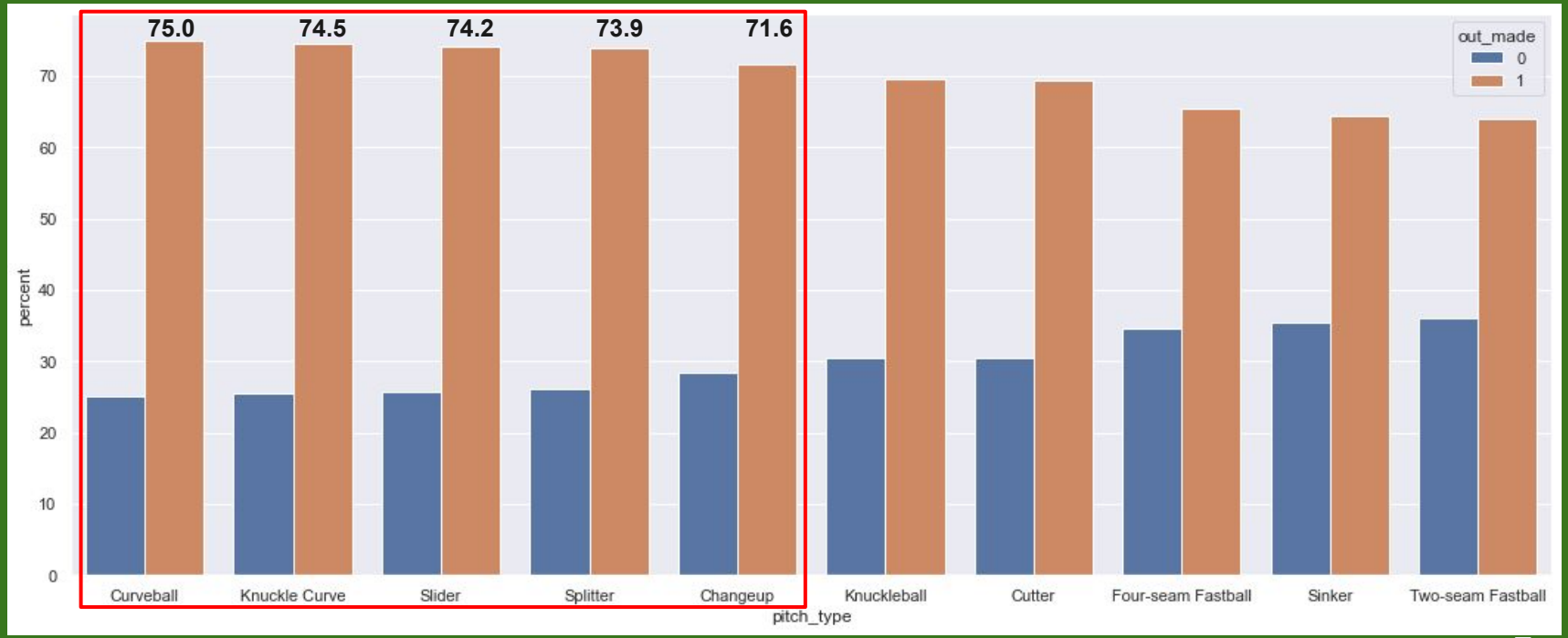
Pitch Speed not predictive on its own



Spin Rate not predictive on its own



Pitch Types that get the most outs



Business Problem

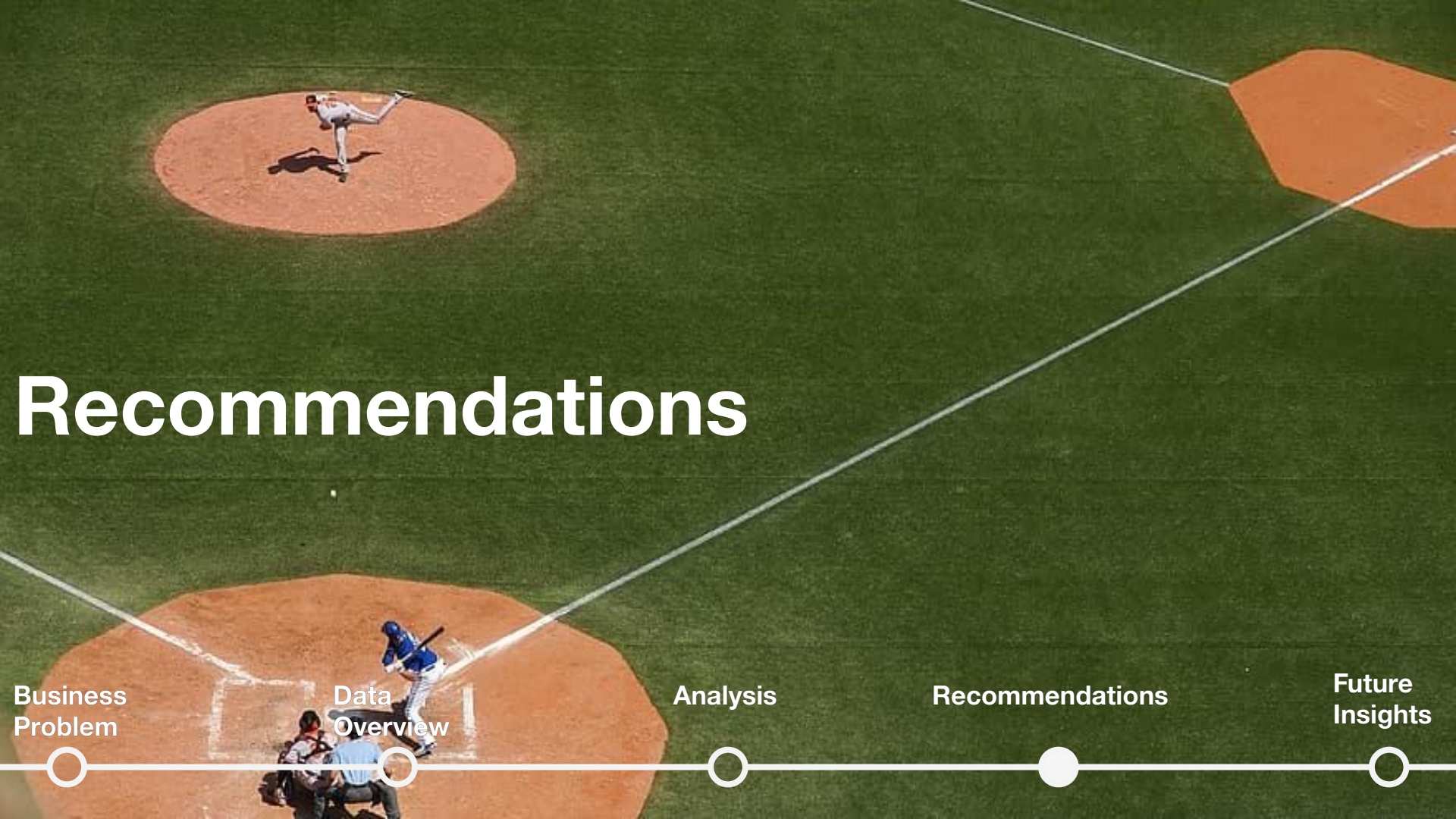
Data Overview

Analysis

Recommendations

Future Insights

Model Analysis



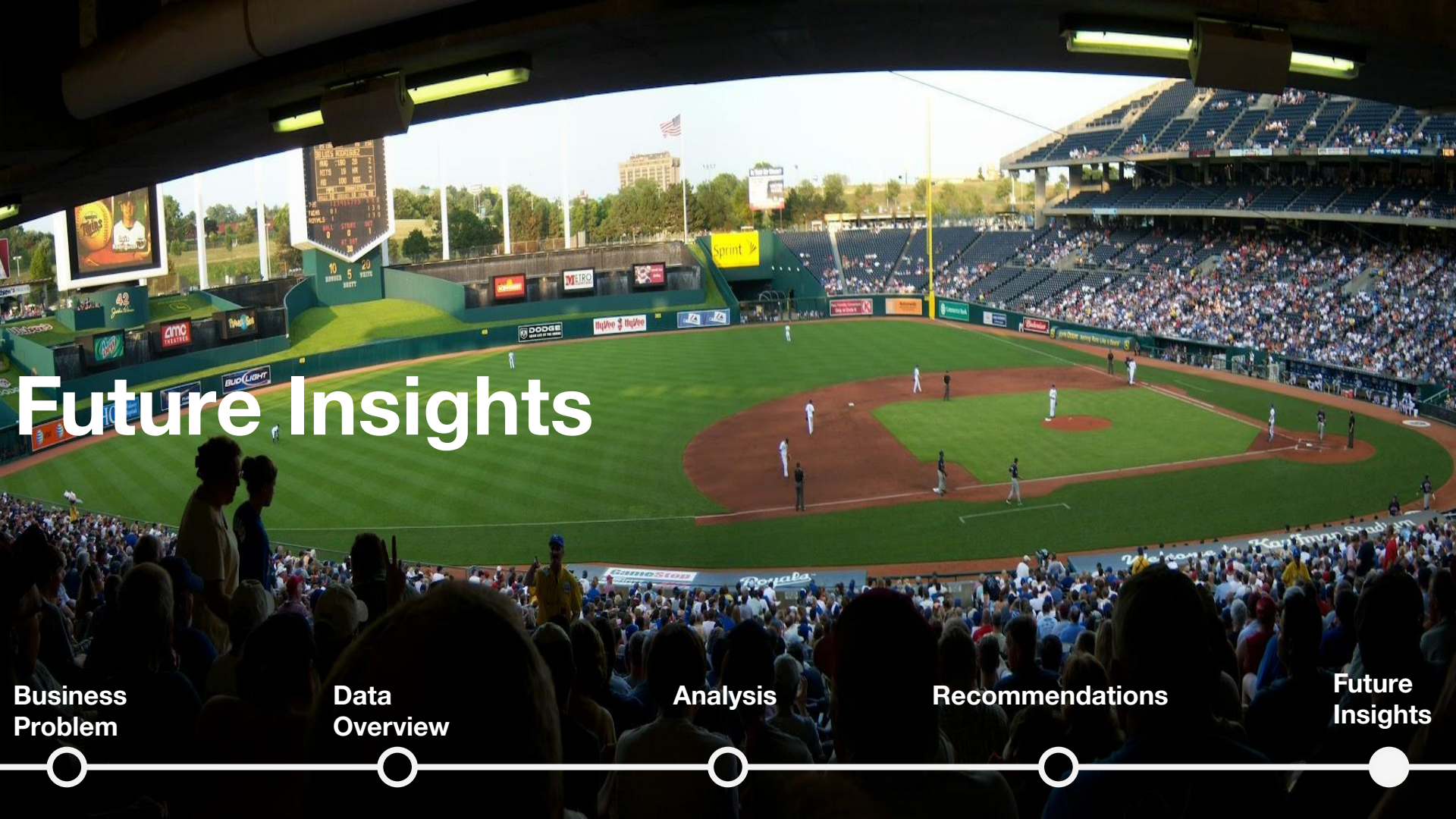
Recommendations

Business Problem
Data Overview
Analysis
Recommendations
Future Insights

Recommendations

- ★ Teach young players the pitches getting outs most often
  - ○ Curve, knuckle curve, slider, splitter, changeup
- ★ Individualize which pitch or pitches work with each pitcher
- ★ Use additional pitch data (including minor league data) to enhance results and further coaching

Changeup



Future Insights

Business Problem

Data Overview

Analysis

Recommendations

Future Insights

Future Insights

- Study pitch sequencing
  - How the order of pitches affects results
- Measure other targets
  - Exit velocity
  - Launch angle
- Monitor trends in the game
  - Rule changes
    - Pitch clock
    - Automated balls and strikes

Thank you

Email:
dan_rosen@outlook.com

Github: dangrosen