

MLB Pitch Analysis

January 3, 2024





Dan Rosen

Github: **dangrosen**

Email: **dan_rosen@outlook.com**

**Business
Problem**

**Data
Overview**

Analysis

Recommendations

**Future
Insights**



Findings

The model improves accuracy in predicting the outcome of an at bat by 9%.



Business Problem



Business Problem

- Look at features of pitches
- Out recorded or not
- Meaningful predictions?



Data Overview



Data Overview

Data

- ❑ 3.5 million pitches from 2015-2019
- ❑ Data from kaggle.com

Conditions

- ❑ Grouped events by out made or not
- ❑ Pitches with highest ratio of getting outs
- ❑ 283,887 records for modeling

Limitations

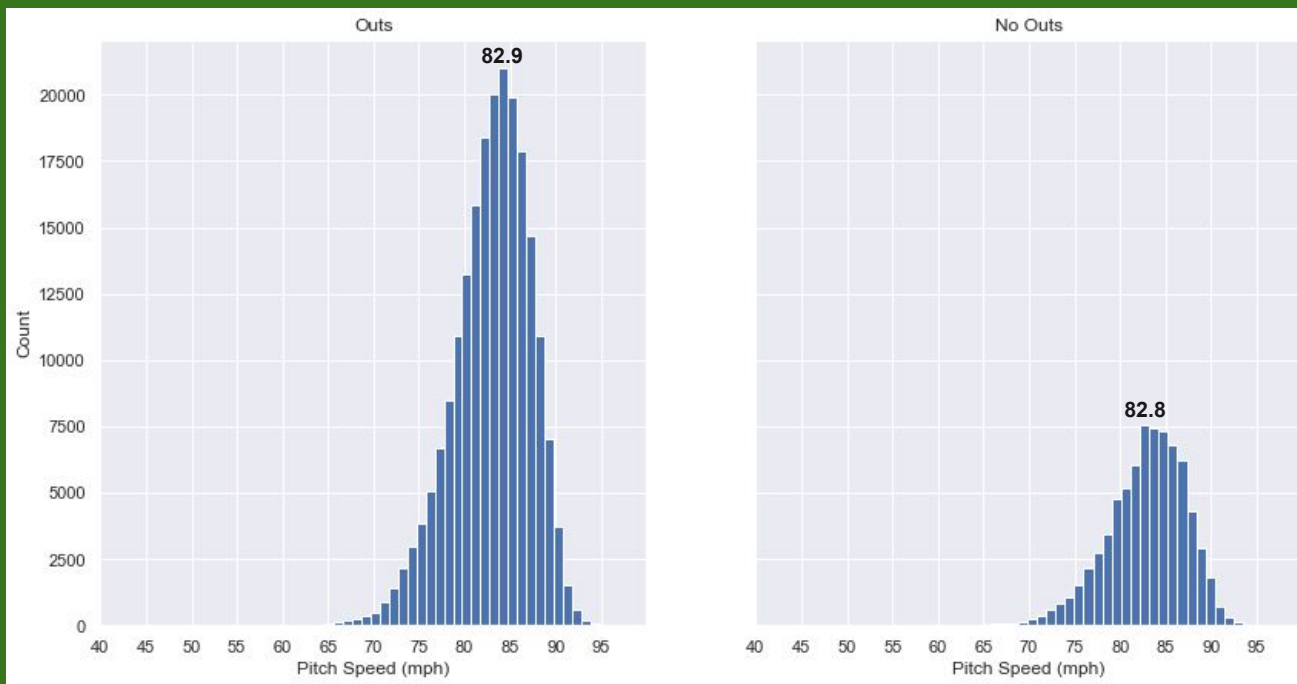
- ❑ Explanation of some columns lacking
- ❑ 2019 data not recorded for some variables



Analysis



Analysis - Pitch Speed



Business
Problem

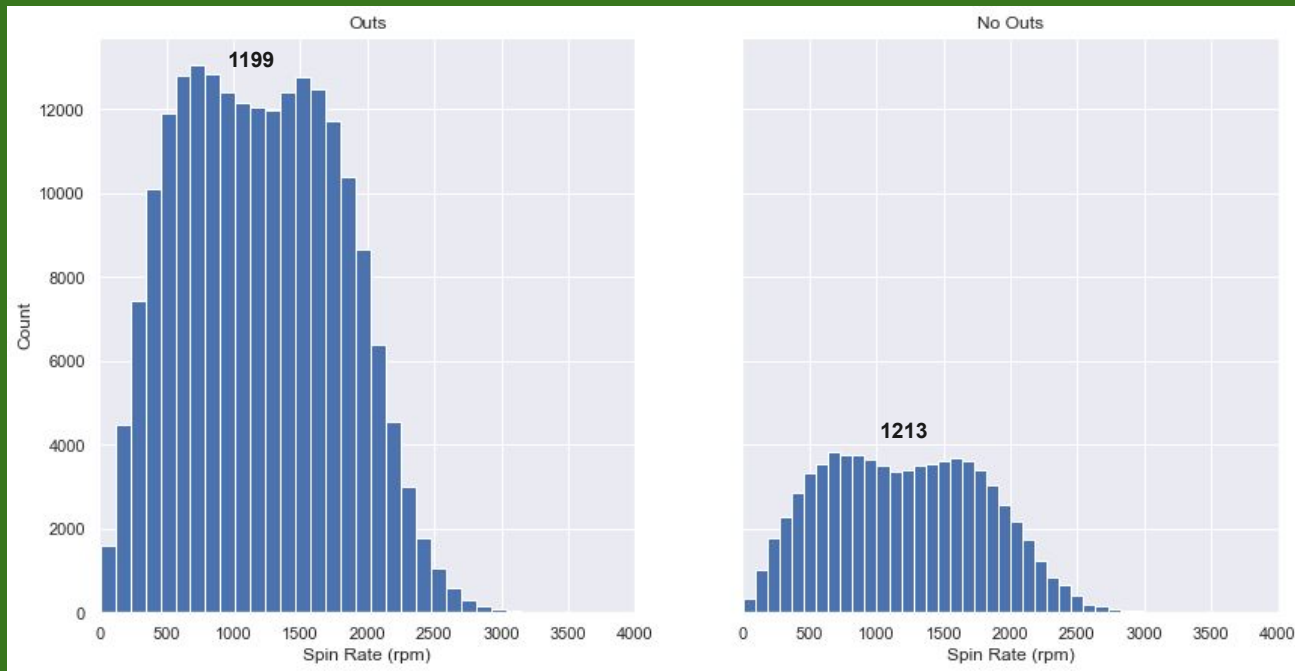
Data
Overview

Analysis

Recommendations

Future
Insights

Analysis - Spin Rate



Business
Problem

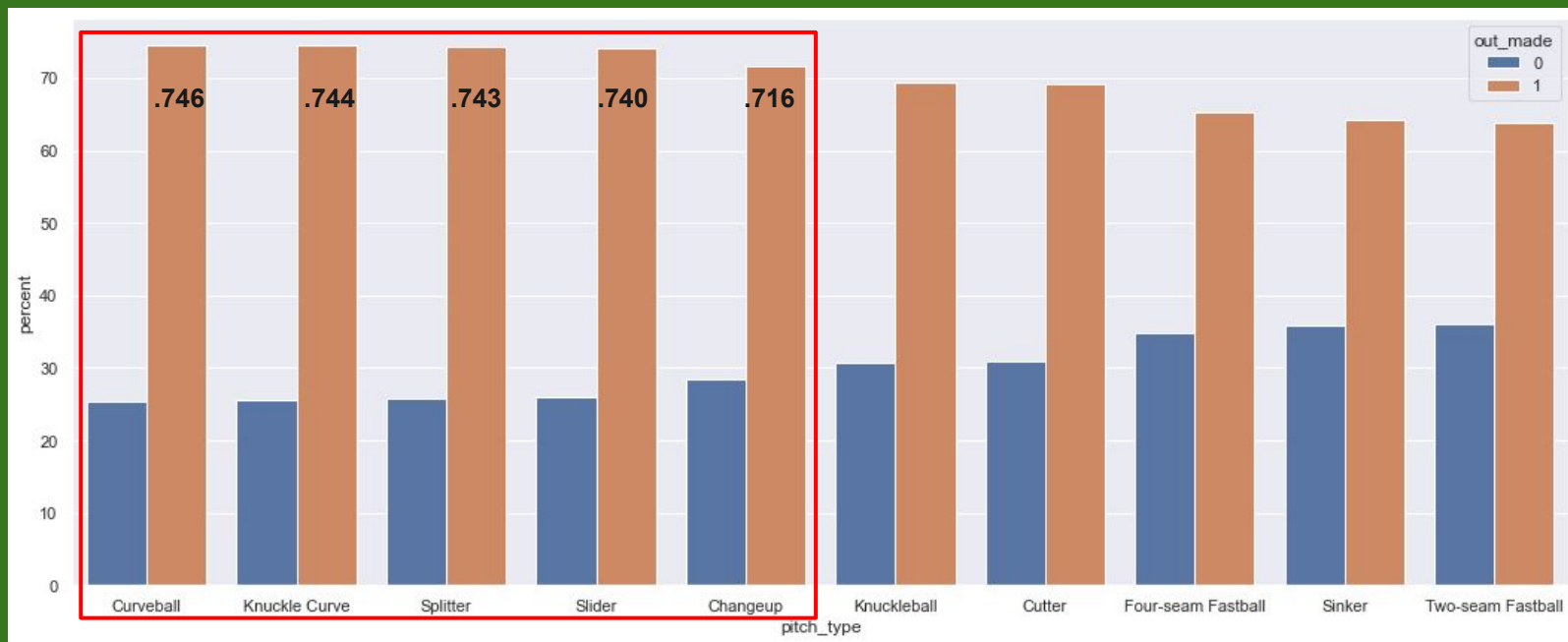
Data
Overview

Analysis

Recommendations

Future
Insights

Analysis - Pitch Type



Business
Problem

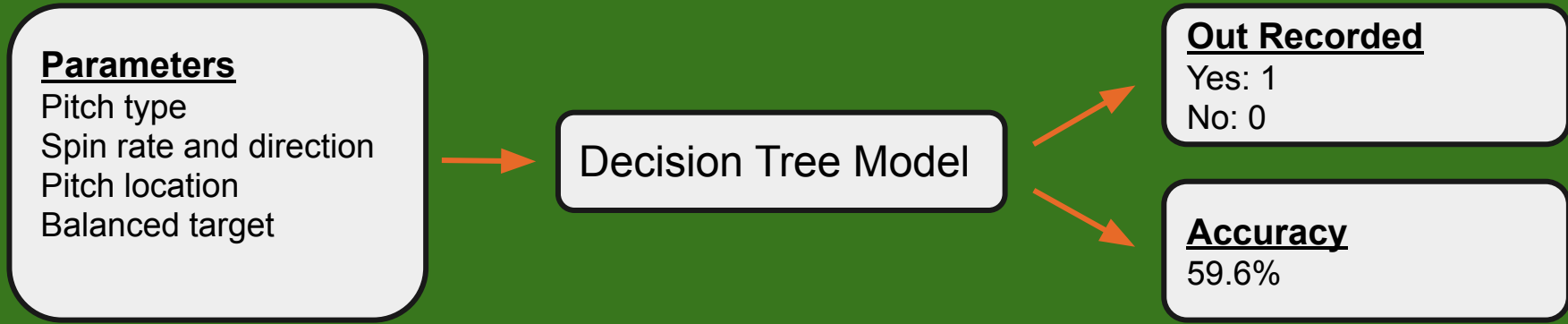
Data
Overview

Analysis

Recommendations

Future
Insights

Analysis - Modeling



Recommendations



Recommendations



Future Insights



Future Insights





Questions

Github: [dangrosen](#)

Email:
dan_rosen@outlook.com

Appendix