

Evan vs. Vegas

Who's Taking the Under?

September 8, 2017

Agenda

Data/Modeling

Exploratory Analysis

Concluding Thoughts

Future Studies

Objective:

**What Baseball Statistics are most
influential in determining the
Over/Under & Winning Team**

The Data

- Games from 2009 - 2016
- Common batting statistics
- Common pitching statistics
 - ◆ Starters
 - ◆ Bullpen
- Game Results
- Month of April removed from the dataset

Modeling

Over/Under

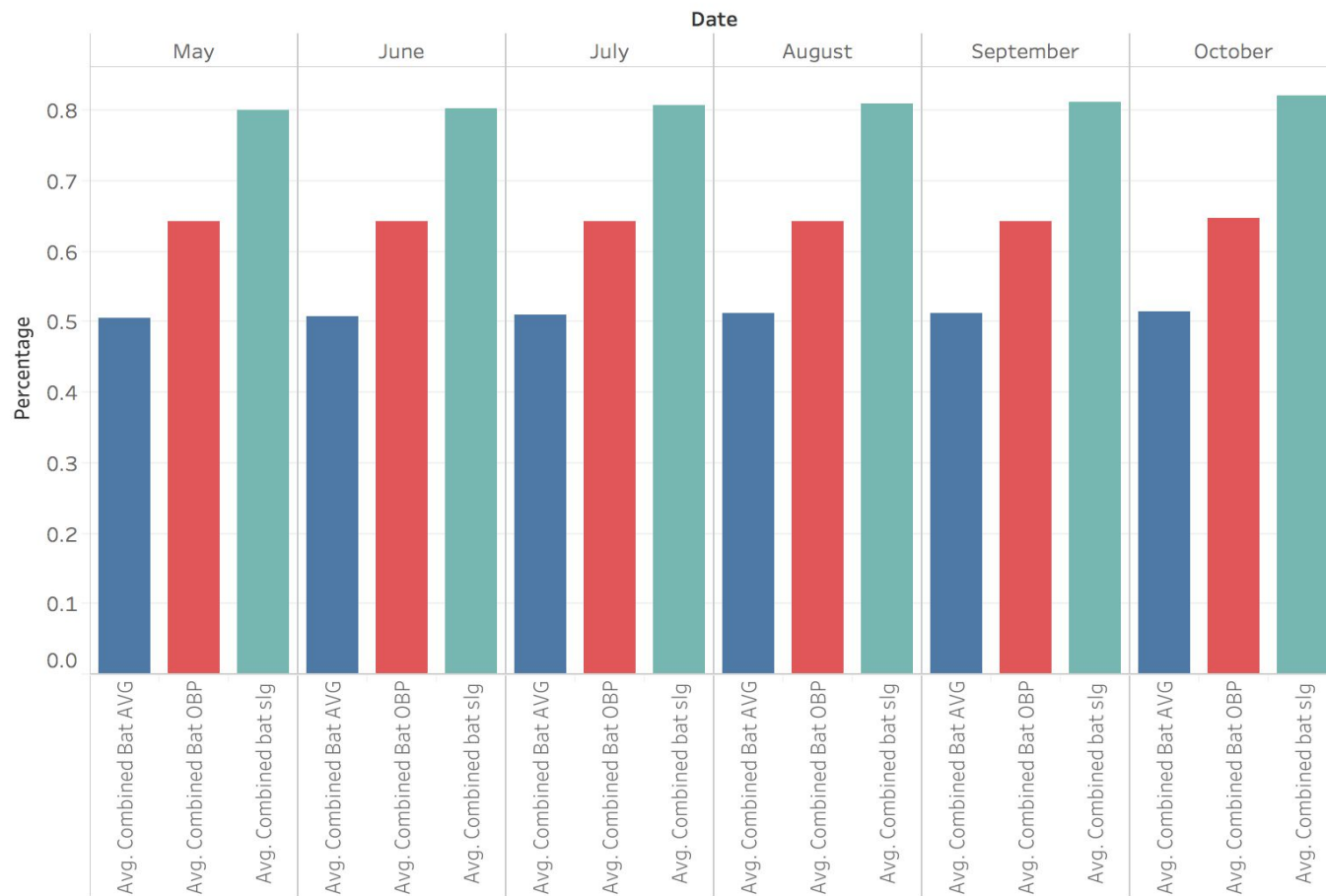
- Baseline
 - 48% Over
 - 52% Under
- Logistic Regression proved to be strongest predictor

Winning Team

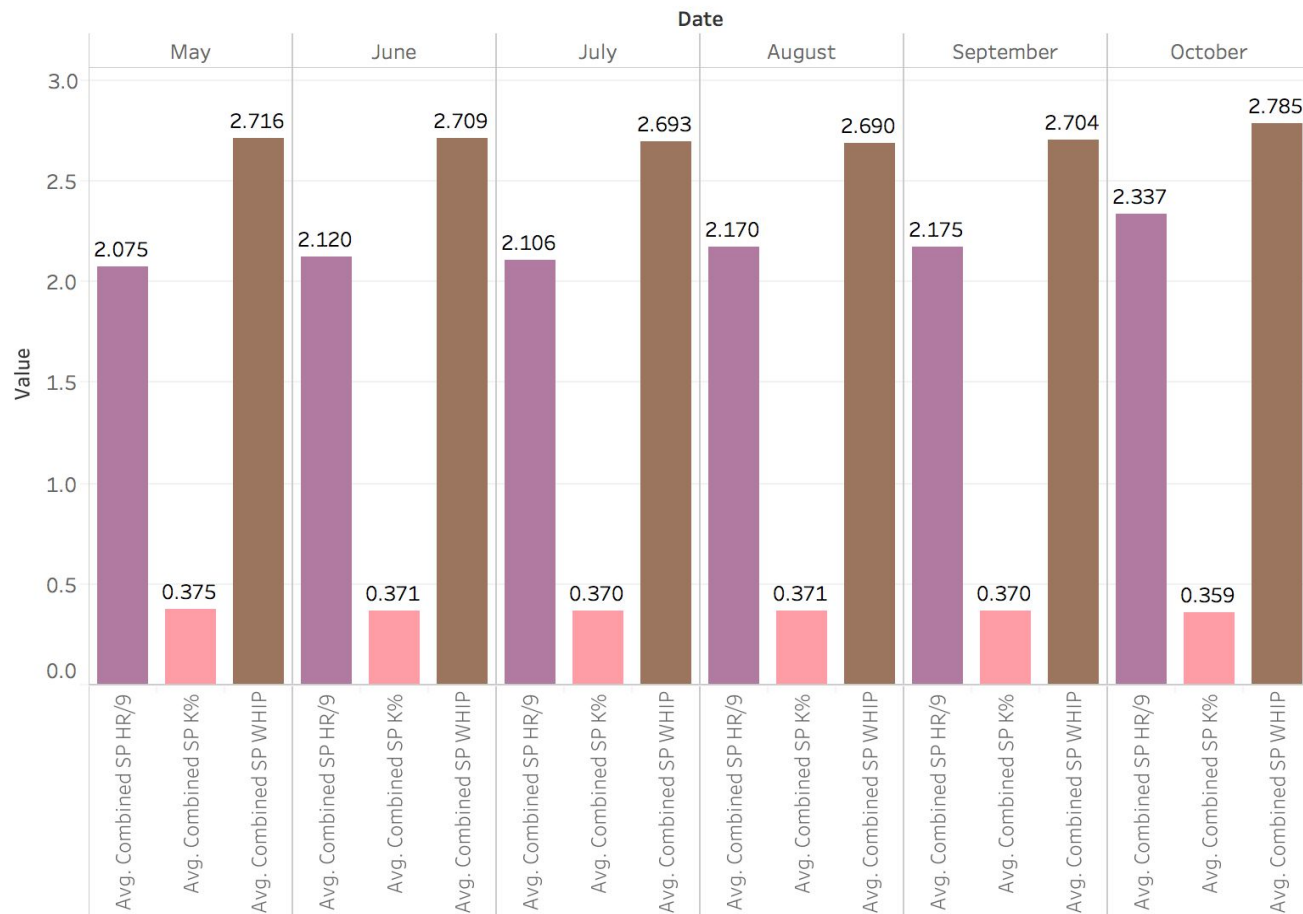
- Baseline score
 - 52% Home Team
 - 48% Away Team
- Logistic Regression proved to be strongest predictor

Exploratory Findings

No Seasonal Changes in OBP, SLG, or AVG



No Seasonal Changes in Pitchers WHIP, K%, or HR/9



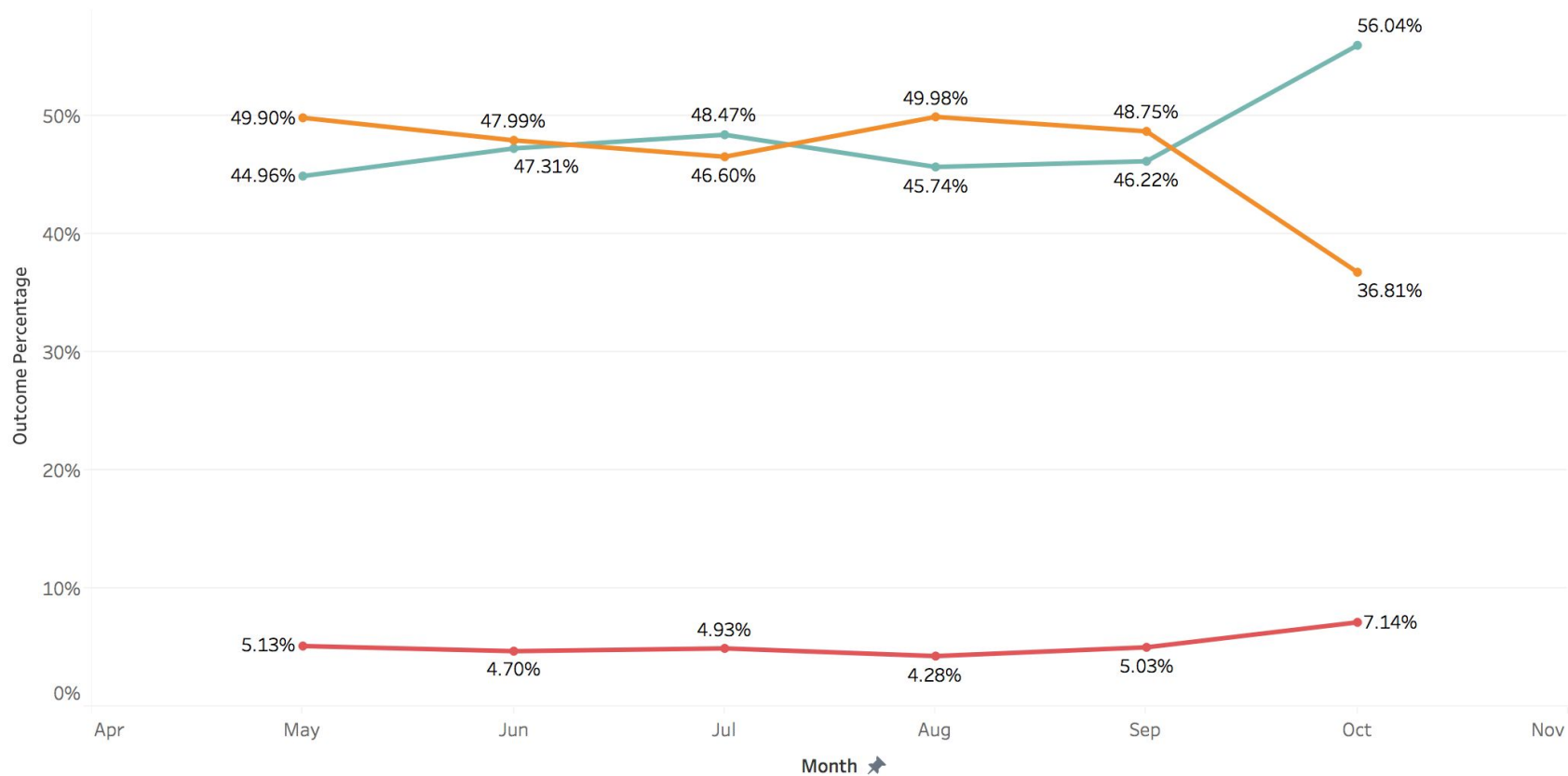
O/U Game Result

Over

Push

Under

Breakdown of Over/Under By Month

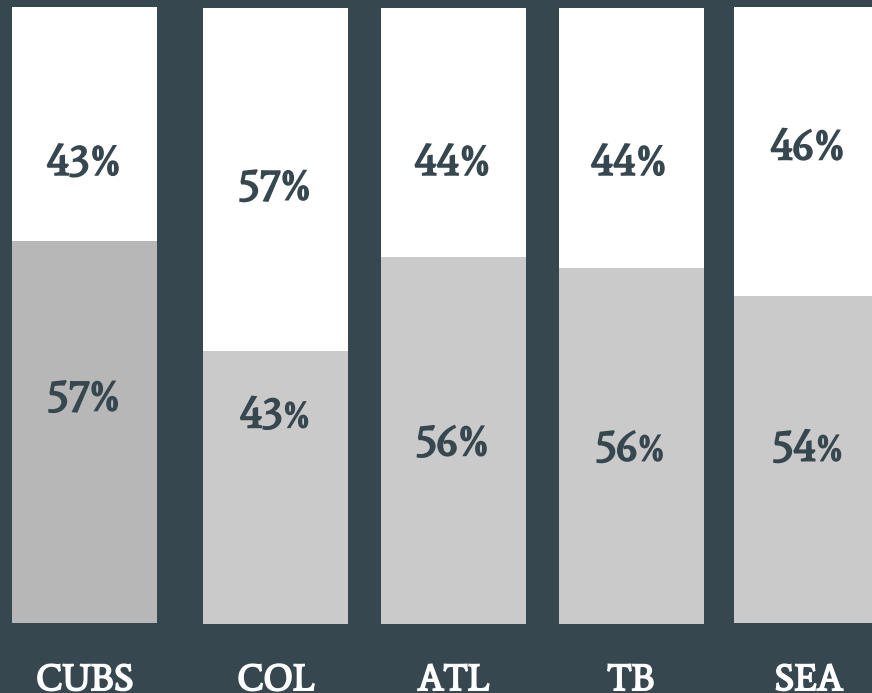


Largest Over/Under Differentials By Ballpark

Over 
Under 

Findings

- Hammer the Over in games played in Colorado
- Hammer the Under in games played in Wrigley Field (Cubs Stadium)



Concluding Thoughts

When it comes to betting baseball...

- Subject Matter Knowledge is of utmost importance
 - ◆ Stadium
 - ◆ Team's current position
 - ◆ Managerial Decisions
 - Monthly data can guide educated decision making
-

Future Studies

Future Studies

- Data on Data on Data
- Feature Engineering
 - ◆ Dummy Variables for Stadium
 - ◆ Weather data
 - ◆ Manager Win %
