

A white golf ball sits on a lush green golf course. In the background, a tall flagpole with a white flag stands against a blue sky with scattered white clouds. A dense line of green trees forms the horizon. The scene is captured in a cinematic style with soft lighting.

# PrecisionPicksPGA

By: Matt Bonadies

# Agenda

1. Project Overview
2. Data Summary
3. Insights
4. Model Optimization
5. Streamlit Demo
6. Next Steps
7. Appendix





# Project Overview

## BACKGROUND

- A passionate golfer who enjoys exploring the influence of data on player performance & sports betting odds

## PROBLEM

- Many find golf to be a dull & difficult sport to understand, coupled with an increasing number of people placing uninformed bets across a range of sports

## OBJECTIVE

- Leverage machine learning to examine player statistics & predict future scores, thereby informing strategic betting decisions & showcasing the greatness of golf

## IMPACT

- In less than 5 years, the legal online sports betting market in the U.S. surpassed **\$10 billion** by 2023 with projections to mature to **\$45 billion** annually



# Data Summary

## DATA SOURCE

- All data was retrieved from datagolf.com, a partner of the PGA Tour, APIs

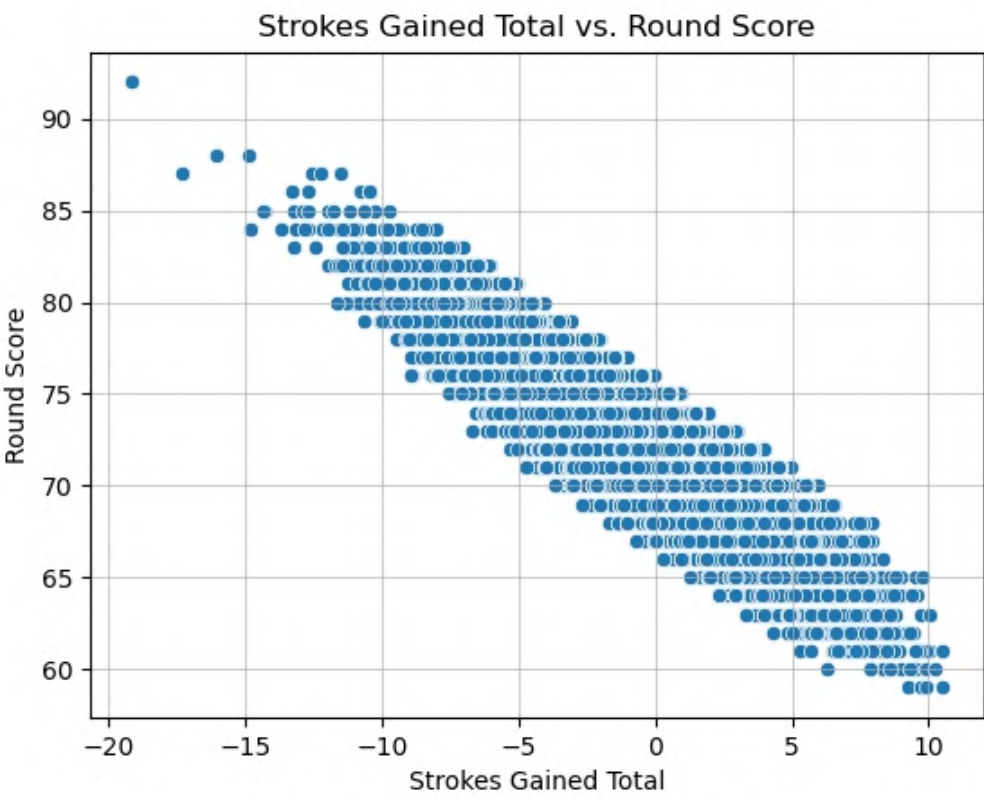
## DATASETS UTILIZED

- **Round Scoring & Stats:** Detailed player statistics since 2017
- **Player Information:** Data such as ID, Country, & Name
- **Tournament Field:** Up-to-date data regarding who is playing
- **Player Rankings:** Ranks of the Top 500 Players Globally

## GOAL

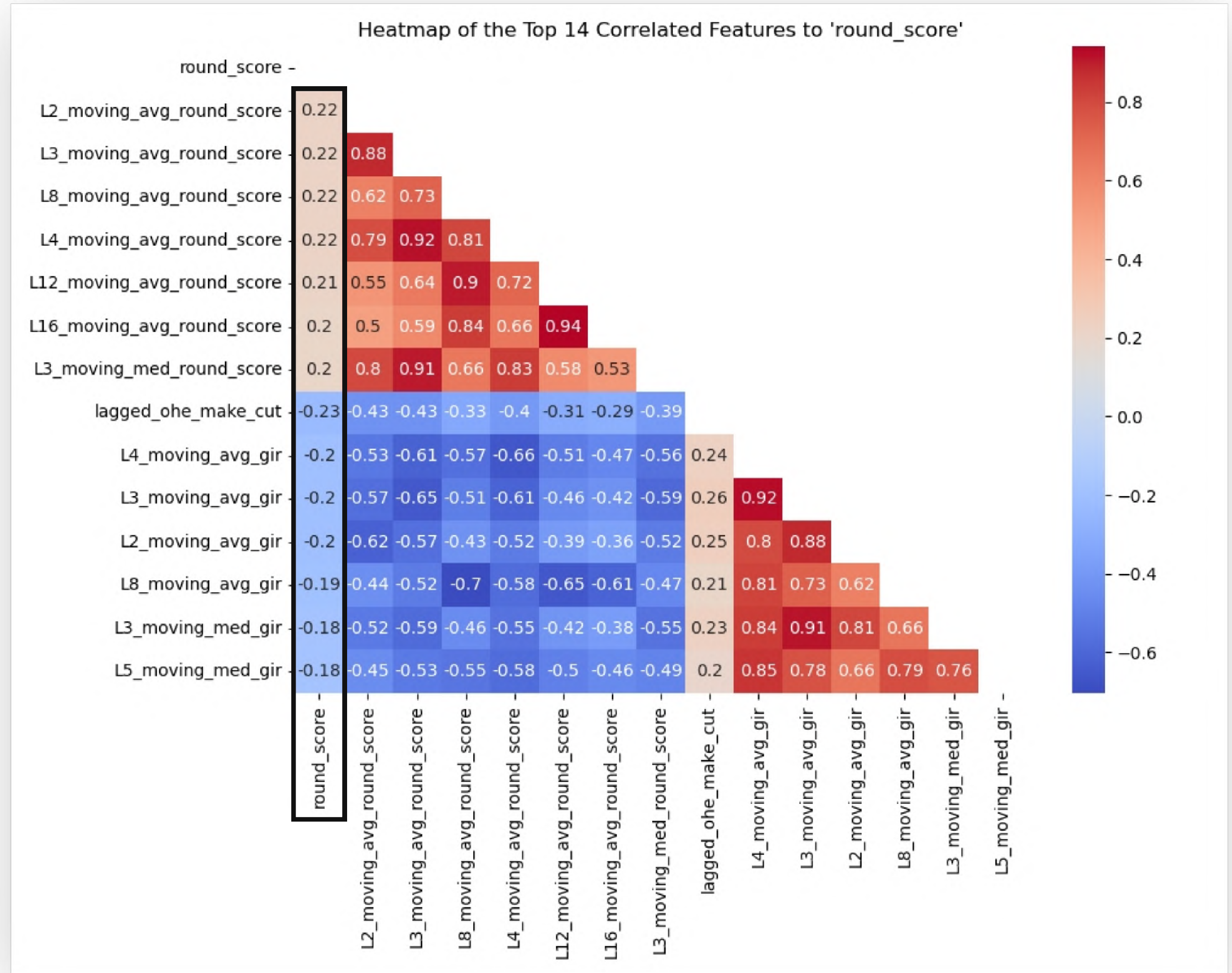
- Develop robust & detailed data to assess its predictive influence on future outcomes

#	Column	Non-Null Count		Dtype
0	event_completed	131333	non-null	object
1	event_name	131333	non-null	object
2	player_name	131333	non-null	object
3	fin_text	131333	non-null	object
4	course_name	131333	non-null	object
5	teetime	131333	non-null	object
6	sg_putt	104164	non-null	float64
7	sg_arg	104164	non-null	float64
8	sg_app	104164	non-null	float64
9	sg_ott	104164	non-null	float64
10	sg_t2g	104164	non-null	float64
11	sg_total	131333	non-null	float64
12	driving_dist	114688	non-null	float64
13	driving_acc	118779	non-null	float64
14	gir	116134	non-null	float64
15	scrambling	102436	non-null	float64
16	prox_rgh	99588	non-null	float64
17	prox_fw	102480	non-null	float64
18	round_score	130873	non-null	float64



## SOLUTION

- ### Heatmap of the Top 14 Correlated Features to 'round\_score'
- 
- | Feature                    | round_score | L2_moving_avg_round_score | L3_moving_avg_round_score | L8_moving_avg_round_score | L4_moving_avg_round_score | L12_moving_avg_round_score | L16_moving_avg_round_score | L3_moving_med_round_score | lagged_ohe_make_cut | L4_moving_avg_gir | L3_moving_avg_gir | L2_moving_avg_gir | L8_moving_avg_gir | L3_moving_med_gir | L5_moving_med_gir |
|----------------------------|-------------|---------------------------|---------------------------|---------------------------|---------------------------|----------------------------|----------------------------|---------------------------|---------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| round_score                | 1.0         | 0.22                      | 0.22                      | 0.22                      | 0.22                      | 0.21                       | 0.2                        | 0.2                       | -0.23               | -0.2              | -0.2              | -0.2              | -0.19             | -0.18             | -0.18             |
| L2_moving_avg_round_score  | 0.22        | 1.0                       | 0.88                      | 0.62                      | 0.79                      | 0.55                       | 0.5                        | 0.8                       | -0.43               | -0.53             | -0.57             | -0.62             | -0.44             | -0.52             | -0.45             |
| L3_moving_avg_round_score  | 0.22        | 0.88                      | 1.0                       | 0.73                      | 0.92                      | 0.64                       | 0.59                       | 0.91                      | -0.43               | -0.61             | -0.65             | -0.57             | -0.52             | -0.59             | -0.53             |
| L8_moving_avg_round_score  | 0.22        | 0.62                      | 0.73                      | 1.0                       | 0.81                      | 0.9                        | 0.84                       | 0.66                      | -0.33               | -0.57             | -0.51             | -0.43             | -0.7              | -0.46             | -0.55             |
| L4_moving_avg_round_score  | 0.22        | 0.79                      | 0.92                      | 0.81                      | 1.0                       | 0.72                       | 0.66                       | 0.83                      | -0.4                | -0.66             | -0.61             | -0.52             | -0.58             | -0.55             | -0.58             |
| L12_moving_avg_round_score | 0.21        | 0.55                      | 0.64                      | 0.9                       | 0.72                      | 1.0                        | 0.94                       | 0.58                      | -0.31               | -0.51             | -0.46             | -0.39             | -0.65             | -0.42             | -0.5              |
| L16_moving_avg_round_score | 0.2         | 0.5                       | 0.59                      | 0.84                      | 0.66                      | 0.94                       | 1.0                        | 0.53                      | -0.29               | -0.47             | -0.42             | -0.36             | -0.61             | -0.38             | -0.46             |
| L3_moving_med_round_score  | 0.2         | 0.8                       | 0.91                      | 0.66                      | 0.83                      | 0.58                       | 0.53                       | 1.0                       | -0.39               | -0.56             | -0.59             | -0.52             | -0.47             | -0.55             | -0.49             |
| lagged_ohe_make_cut        | -0.23       | -0.43                     | -0.43                     | -0.33                     | -0.4                      | -0.31                      | -0.29                      | -0.39                     | 1.0                 | 0.24              | 0.26              | 0.25              | 0.21              | 0.23              | 0.2               |
| L4_moving_avg_gir          | -0.2        | -0.53                     | -0.61                     | -0.57                     | -0.66                     | -0.51                      | -0.47                      | -0.56                     | 0.24                | 1.0               | 0.92              | 0.8               | 0.88              | 0.81              | 0.78              |
| L3_moving_avg_gir          | -0.2        | -0.57                     | -0.65                     | -0.51                     | -0.61                     | -0.46                      | -0.42                      | -0.59                     | 0.26                | 0.92              | 1.0               | 0.88              | 0.73              | 0.91              | 0.79              |
| L2_moving_avg_gir          | -0.2        | -0.62                     | -0.57                     | -0.43                     | -0.52                     | -0.39                      | -0.36                      | -0.52                     | 0.25                | 0.8               | 0.88              | 1.0               | 0.62              | 0.81              | 0.66              |
| L8_moving_avg_gir          | -0.19       | -0.44                     | -0.52                     | -0.7                      | -0.58                     | -0.65                      | -0.61                      | -0.47                     | 0.21                | 0.81              | 0.73              | 0.62              | 1.0               | 0.84              | 0.76              |
| L3_moving_med_gir          | -0.18       | -0.52                     | -0.59                     | -0.46                     | -0.55                     | -0.42                      | -0.38                      | -0.55                     | 0.23                | 0.84              | 0.91              | 0.81              | 0.66              | 1.0               | 0.79              |
| L5_moving_med_gir          | -0.18       | -0.45                     | -0.53                     | -0.55                     | -0.58                     | -0.5                       | -0.46                      | -0.49                     | 0.2                 | 0.85              | 0.78              | 0.66              | 0.79              | 0.76              | 1.0               |



# Model Optimization

Model Type	Train R^2	Test R^2
Linear Regression	26.29%	26.35%
Lasso Regression	25.81%	26.41%
Neural Netork	42.98%	31.84%

*\*For additional scoring metrics please refer to Jupyter Notebooks*



An aerial photograph of a golf course. In the center is a large, light-green putting green. To its upper right is an irregularly shaped sand trap. To its lower left is another sand trap of a similar shape. The surrounding area is covered in lush green grass. Several trees are scattered around the course, including a large, dense tree on the left and a cluster of trees on the right. Long, dark shadows are cast across the grass, suggesting the sun is low in the sky. The text "Streamlit Demo" is written in a white, serif font, centered over the green.

# Streamlit Demo



## **Next Steps:**

1. Enhance Models with New Features
2. Deploy Streamlit Application
3. License GitHub Repository



# Appendix





# What is Golf?

## SUMMARY OF THE SPORT

- Played on an 18 hole course spanning thousands of yards
- Objective is to hit the ball into each hole in the fewest strokes
- Baseline score for most courses is 72 strokes
- Various different shots to hit such as driving, chipping, etc.

## PGA TOUR TOURNAMENT

- 4 rounds of 18 holes each, played over 4 days (Thu-Sun)
- Approx. 120 players competing against each other
- Player with the lowest total score wins
- New competition begins every week in a different location





# What is Sports Betting?

## SUMMARY OF SPORTS BETTING

- Individuals can bet money on certain outcomes occurring
- Every bet has odds representing the likelihood of an event occurring which determines how much you win
- Bets can be placed online or in-person for most sports

## GOLF BETTING

- Golf bets come in many forms:
  - Guessing the winner
  - Predicting an individual to finish in the top 5, 10, or 20
  - Wagering on the winner out of 3 players in a group
  - Betting on one player to finish better than another



# Impact & Vision

## INDUSTRY IMPACT

- **In 2018**, the Supreme Court reversed a federal law banning sports betting
  - Approx. **38 states** have legalized sports betting, both in-person and online
- In less than 5 years, the legal online sports betting market in the U.S. surpassed **\$10 billion** by 2023
- The market is expected to reach **\$45 billion** annually at market maturity

## VISION

- **Growing Trend:** As the number of sports bettors increases, navigating the complexities of betting becomes challenging.
- **Reality:** Many are unaware that the break-even point for betting odds exceeds a 50% probability, as set by casinos.
- **Objective:** Analyze extensive individual player statistics using machine learning to forecast future scores, thereby guiding the most strategic bets for tournament winners, finishing positions, and head-to-head matchups.