Exploring Golf Analytics From Trackman System: Consistency and Clustering Analysis

Jackson Meehan, Emily Feng, Rohan Patel
Faculty Advisor: Ron Yurko Client: Daniel Rodgers

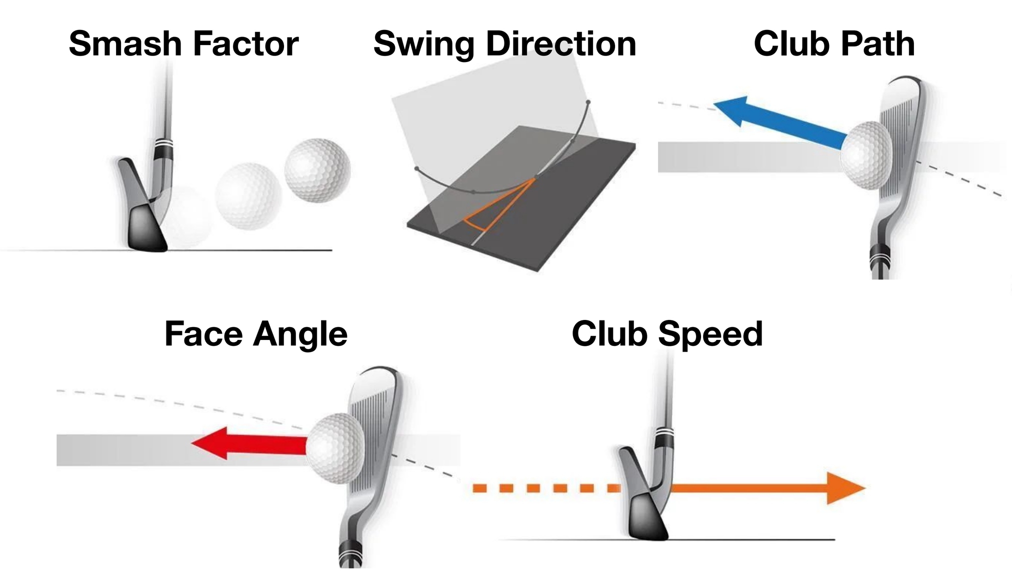
Background & Introduction

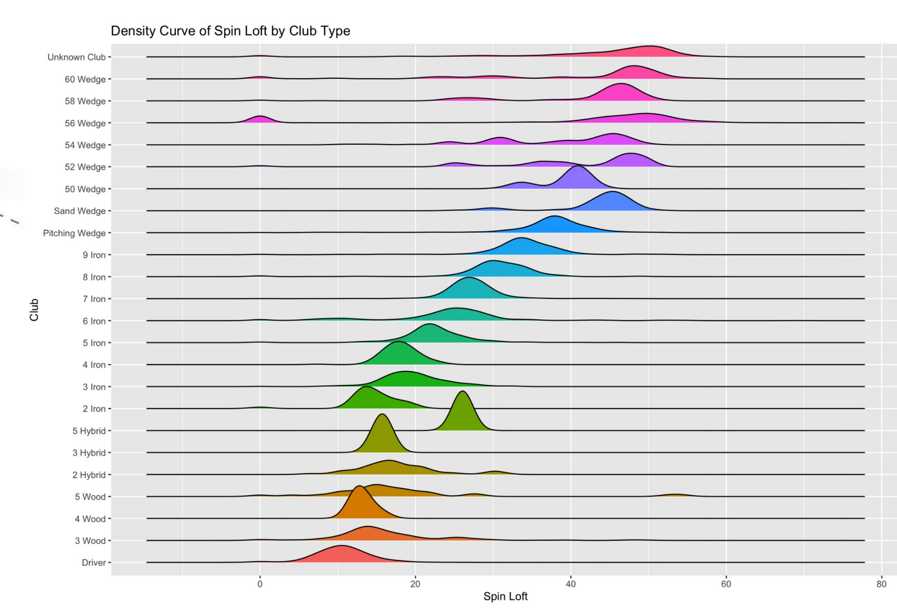
Carnegie Mellon University's Varsity Golf Team practices with the Trackman system, which utilizes radar to record information about each swing and uses it to infer the flight path of the golf ball. Goals:

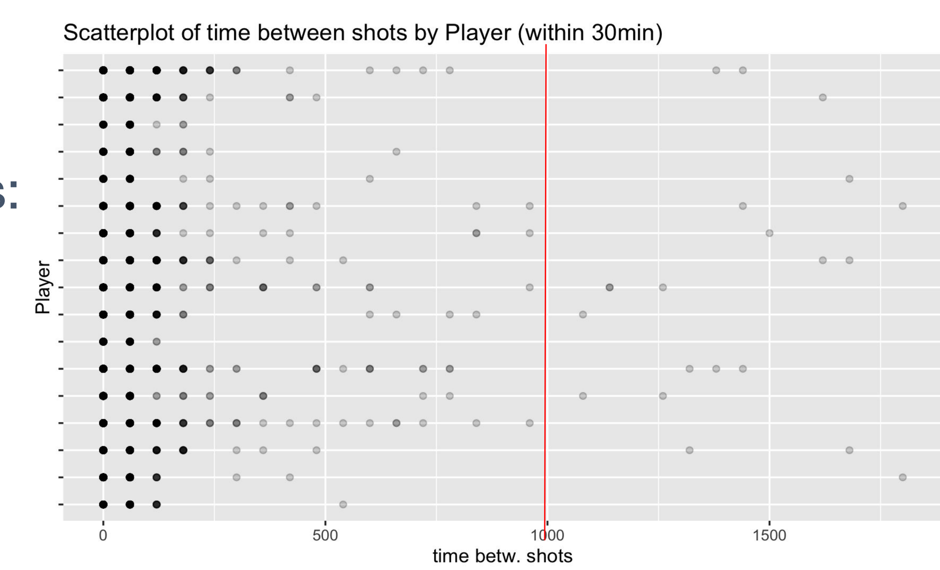
1.) **Explore relationships between Trackman variables**
2.) **Quantify and compare players' consistency**
3.) **Visualize deviation across and within player shot sessions**

Data & Feature Engineering

Dataset consists of **11,924** shots by **16** players with **40** metrics, such as:







Feature Engineering
Clustered club types into 4 categories: Woods, Irons, Hybrids, and Wedges
New "session" begins when:

- Next shot by different player OR
- > 17 mins after the previous shot

Methods

Player-Agnostic:

- Principal Component Analysis (PCA) identifies largest sources of variance
- Gaussian Mixture Models (GMM) used for classification, exploring how shot attributes could characterize athletes

$$p(x) = \sum_{i=1}^{k} \pi_i N(x|\mu_k, \Sigma_k)$$

- Adjusted Rand Index (ARI) measures partition similarity, where <0 is worse than random, 0 is random and 1 represents perfect correspondence

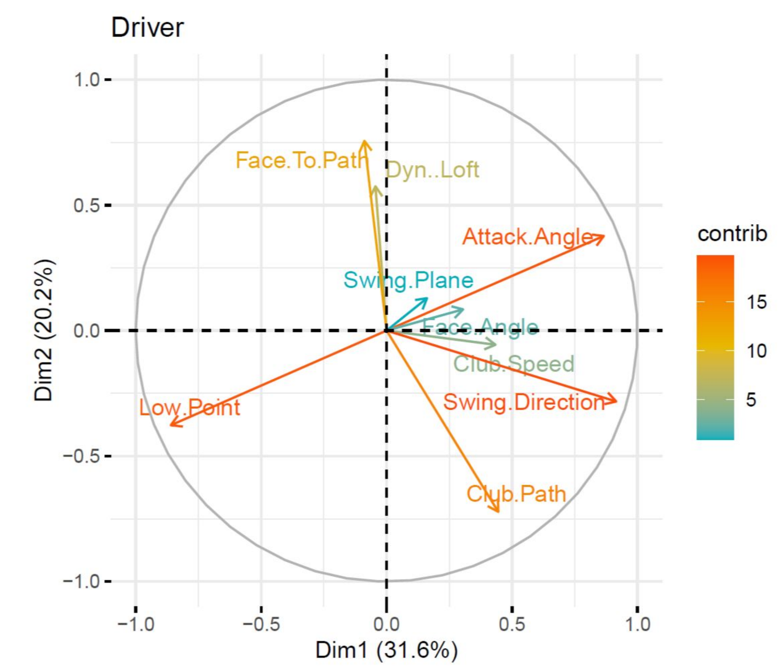
- **Player-Specific:**
- Standardized player metrics by club for each session

$$Z = \frac{X - \mu}{\sigma}$$

- Trajectory clustering:
  a. computes 24 measures of each trajectory
  b. selects subset of measures that describe main features
  c. clusters using cubic clustering criteria
- Trajectory clustering approach is agnostic to session length

Results

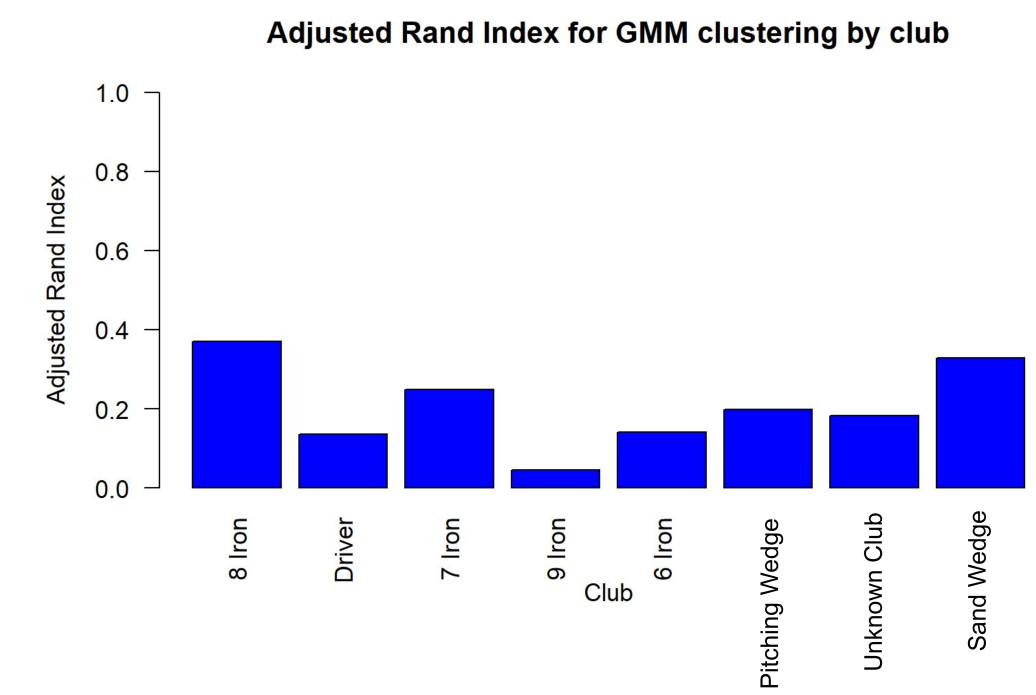
PCA identifies four key metrics driving variances



Feature selection for GMM corroborates relevance of key metrics across club types

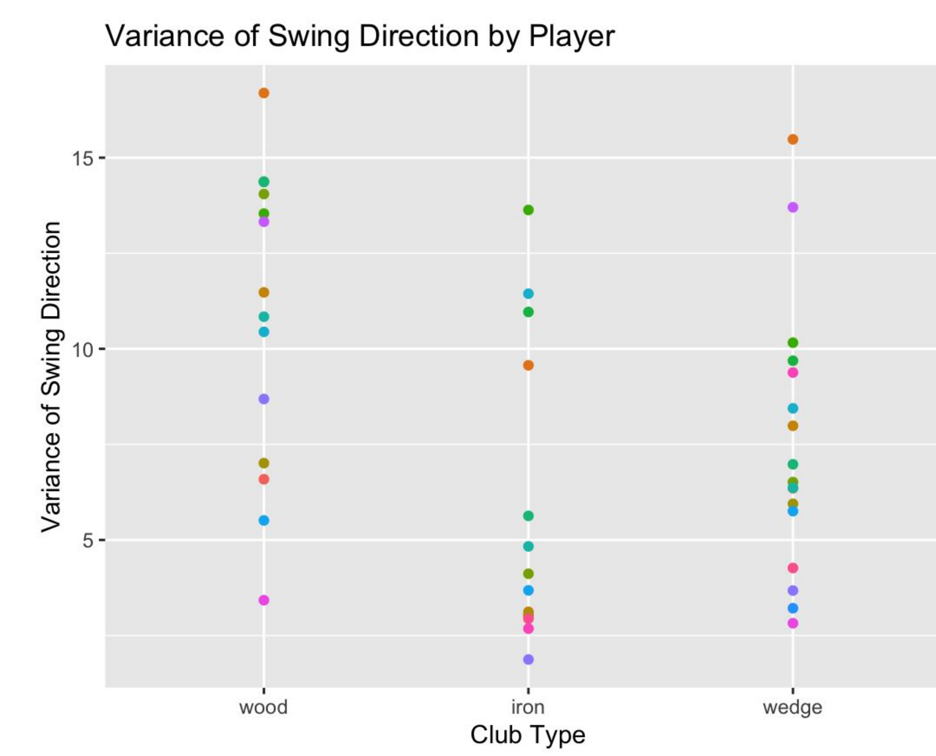
	Club Types			
Metrics	IRON	WOOD	HYBRID	WEDGE
Club Speed	✔	✔	✔	✔
Face Angle	✔	✔		✔
Club Path	✔	✔	✔	✔
Swing Direction	✔	✔		✔
Smash Factor	✔	✔	✔	✔

ARI indicates that GMM classification performance varies strongly by club

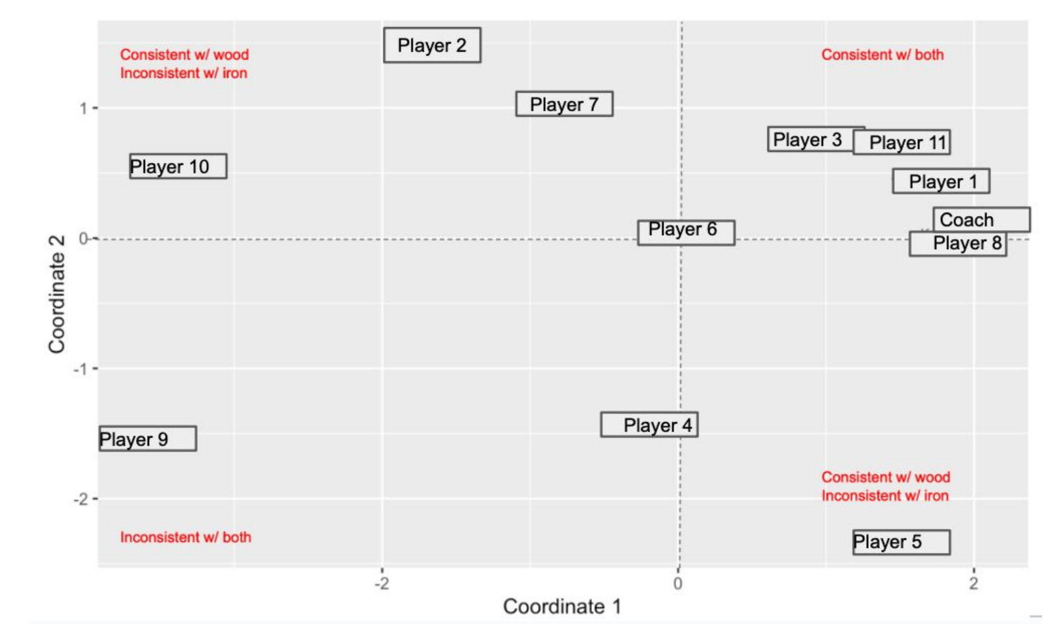


Measure player consistency by club type with variance of 4 key measures

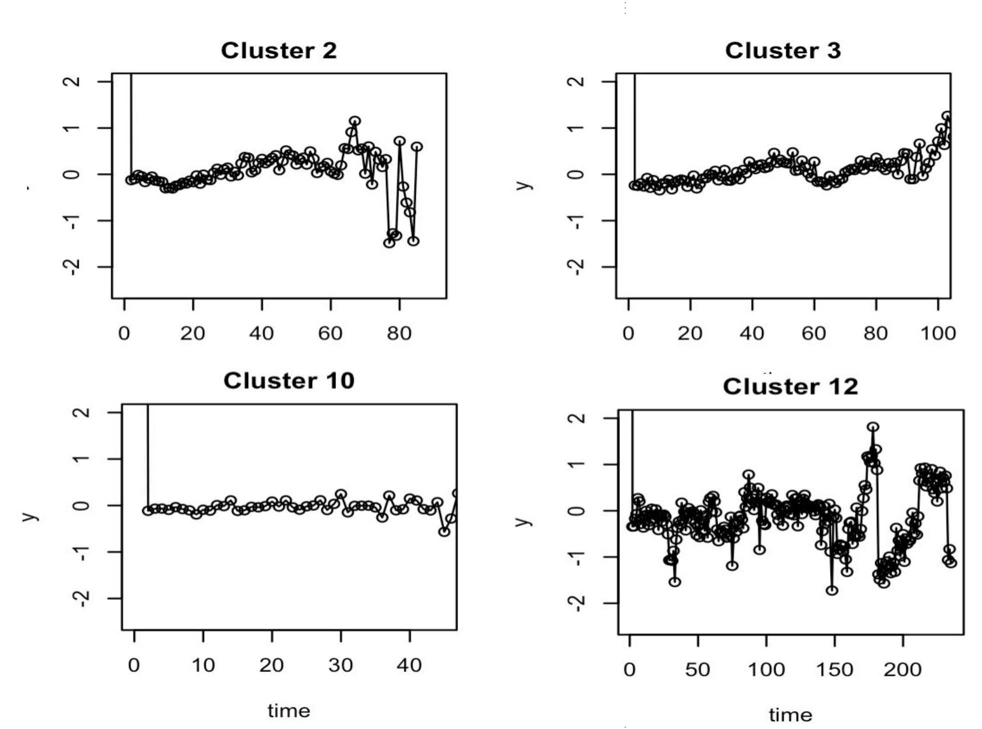
Player	wood_Club.Speed	iron_Club.Speed
Player 1	94.984272	39.223081
Player 2	14.816921	106.679144



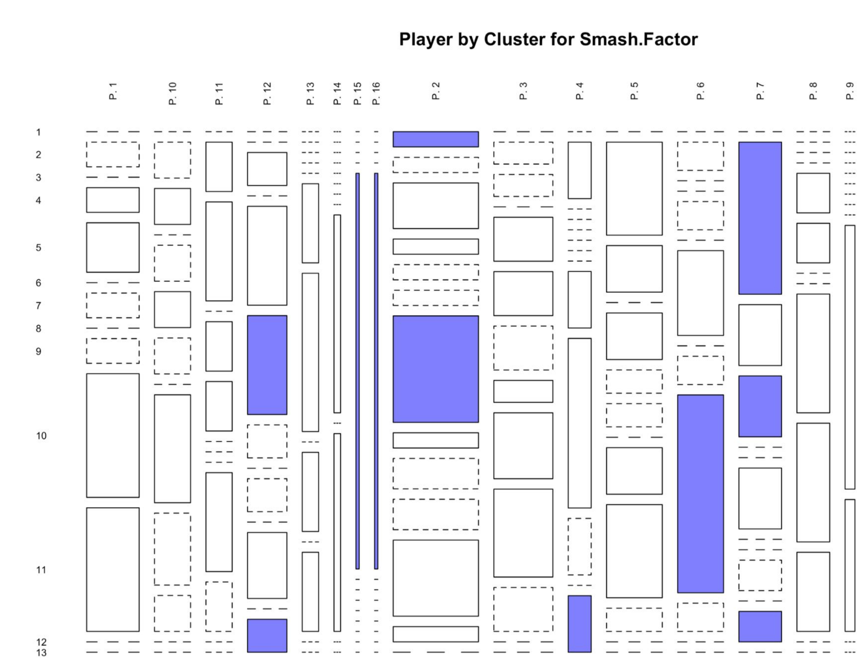
Visualize distance between players with multidimensional scaling, revealing players with similar and dissimilar consistency profiles



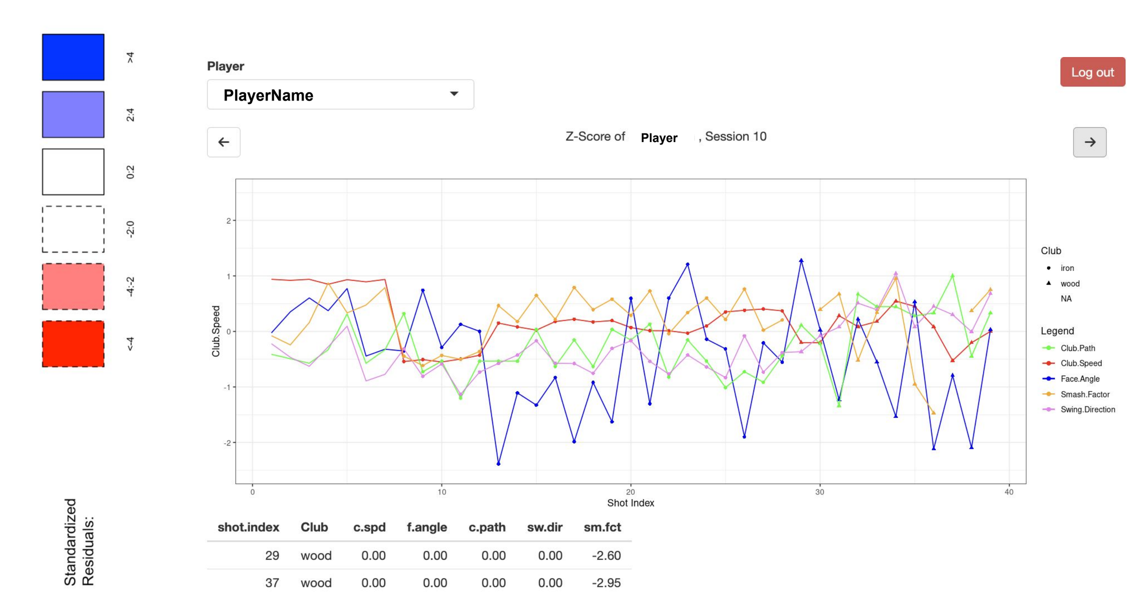
Trajectory Clustering
Chose 14 clusters based on CCC index [3], example mean of 4 clusters



Observe an association between player metrics and consistency



R-Shiny App to zoom in on each player's session performance by metrics



Discussion

- Few key metrics found to be important when attempting to classify shots
- Visualized variance of metrics for players to compare overall consistency
- R Shiny App that visualizes how players change within a session
- Using app and trajectory clustering, we identify trends for CMU golfers that provide insight on the results from their practices

References

1. Dreger, Bill. "Understanding Your TrackmanTM Analysis Data." NonstopGOLF, https://nonstopgolf.ca/wp-content/uploads/2011/09/TrackMan-Data-Primer.pdf.
2. Hahn, Christian. "Club Data Definitions." TrackMan Golf, 23 Aug. 2017, https://blog.trackmangolf.com.
3. Charrad, M., . N. Ghazzali, V. Boiteau, and A. Niknafs. "NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set". Journal of Statistical Software, vol. 61, no. 6, Nov. 2014, pp. 1-36, doi:10.18637/jss.v061.i06.