

DataFest 2019

Fusheng Yang
Hao Gao
Ray Huang
Candice Zheng
Iris Li

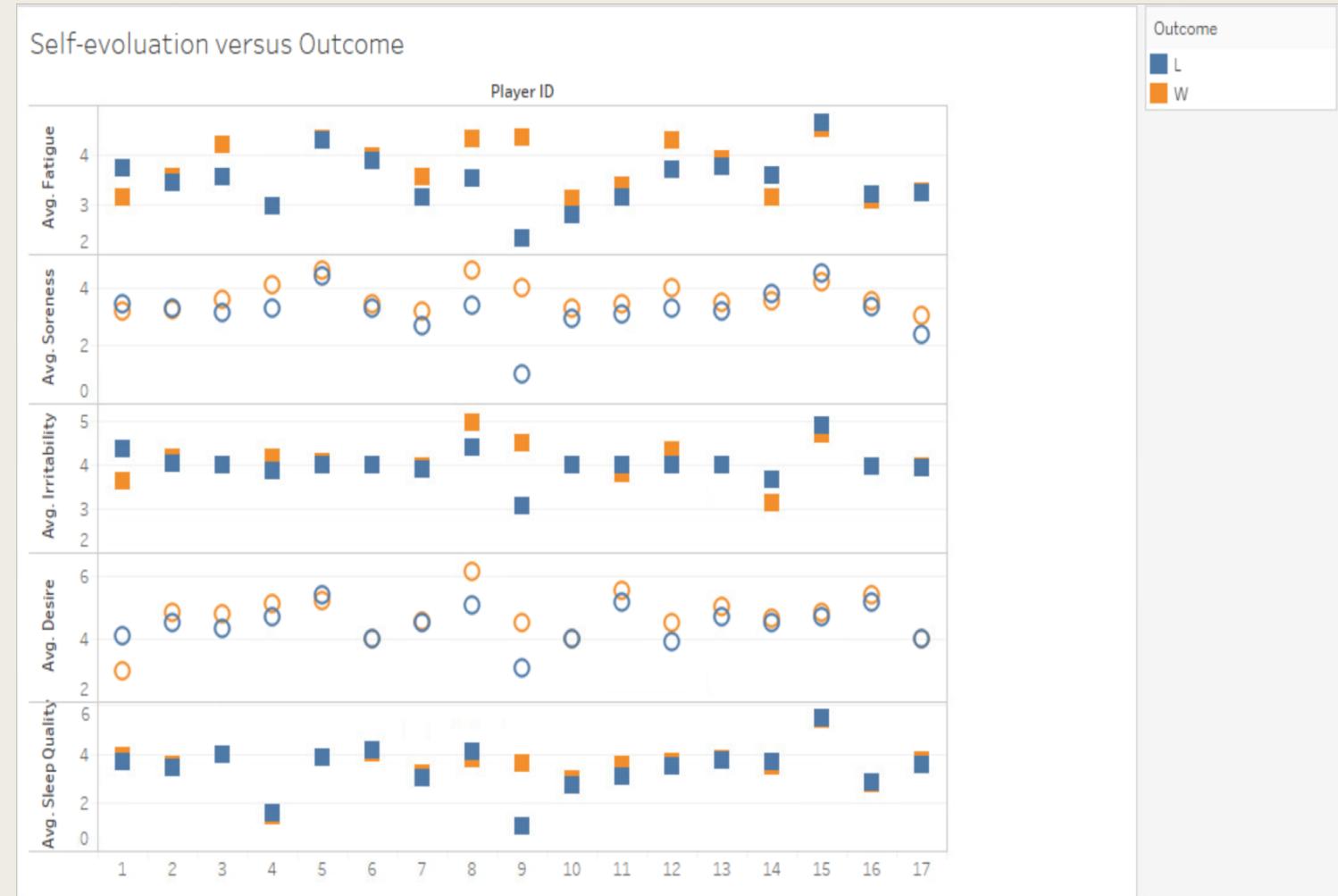
Umass-Amherst

568 negative

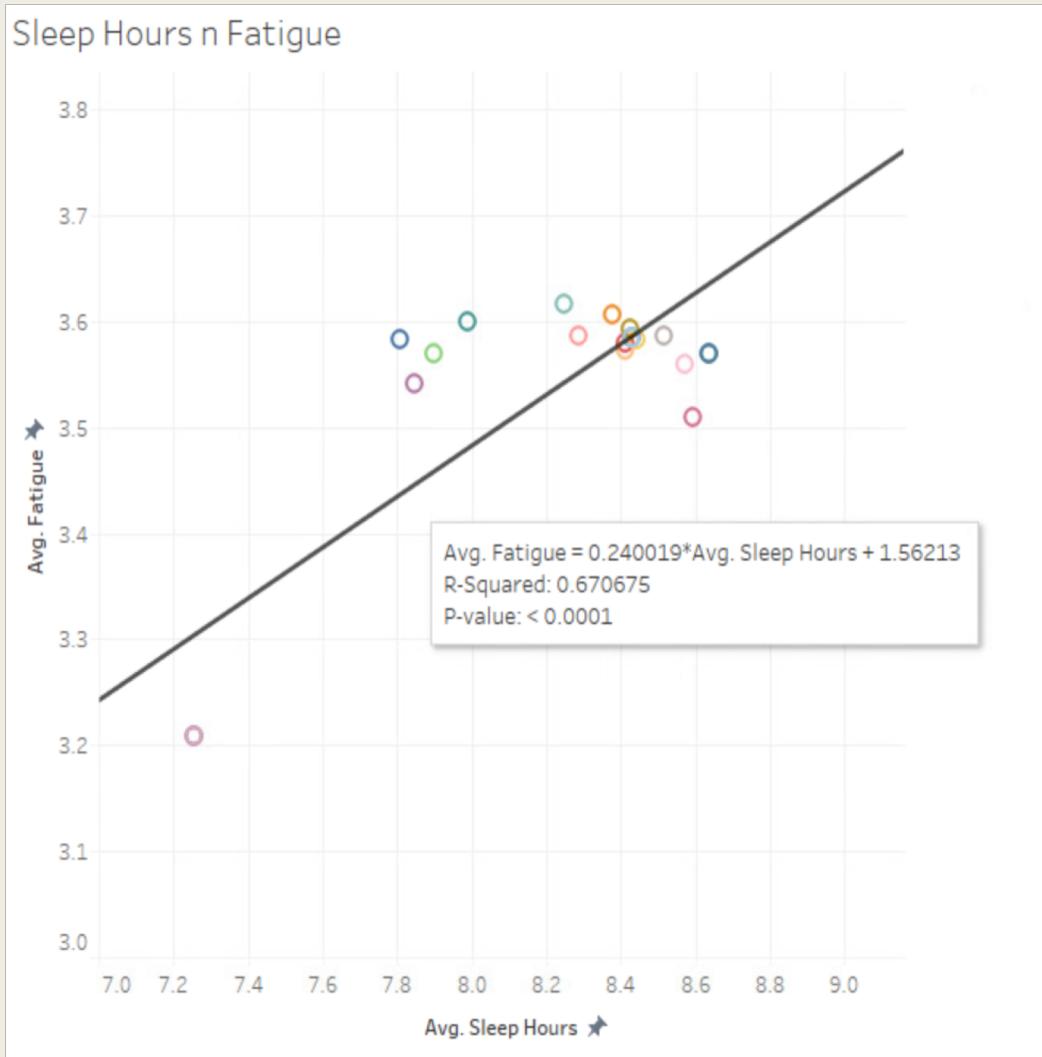
Initial Attempt

- Predict:
 - Fatigue
 - Outcome
- Variable include:
 - Subjective
(Soreness, Desire,
Irritability,
SleepQuantity, ...)
 - Objective
(SleepHours,
Speed,
SessionType,
Duration, ...)
- Also tried:
 - Neutral network
 - Cluster analysis

ALL FAILED!



Fun Facts of Sleep Hours



First Attempt at modeling

Model the “Performance Score”



f_{θ_1} : Model(NN etc.) that learn a function to estimate the performance score for each player

g_{θ_2} : A Classifier that predict win/lose base on the performance score

Assumption: The higher P score implies higher probability to be classified as “win”

$$cost(\theta_1, \theta_2) = \frac{1}{2n} \sum_i y_i - g_{\theta_2}(f_{\theta_1}(x_i))^2 - \sum_{i,j} \sigma K(x_i, x_j) + \lambda(|\theta|^2 + |\theta|^2)$$

$$K(x_1, x_2) = (|g_{\theta_2}(f_{\theta_1}(x_1))|^2 - |g_{\theta_2}(f_{\theta_1}(x_2))|^2) * (|f_{\theta_1}(x_1)|^2 - |f_{\theta_1}(x_2)|^2)$$

Random Forest Regression

