How to Measure Fatigue and What We Can Do with the Existing Measurements

Team Babies
Younghoo Cho
Irene Li
Murphy Lu
Lynn Zhu
Zoe Zhou

Substitute for the Self-reported Fatigue

- Median speed
- Reasons why we choose this as a substitute:

1. High intensity

2. Visualization

Identification of Explanatory Variables

- Duration (Frame ID)
- Load (Median load detected by the accelerometer)
- Impulse (Median absolute value of change in speed divided by change in time)
- Jump (Median acceleration in vertical direction)
- Interaction terms

Models and Results

- Multiple Linear Regression Model
    - R-squared
    - Statistical significance

Reflections and Further Investigations

- How to improve individual's performances during games

  - Monitor jump height and control for it

  - Help to decide which player to play the second half of the game/game later

- More inclusion of training data and identification of difference between baseline and competition phase

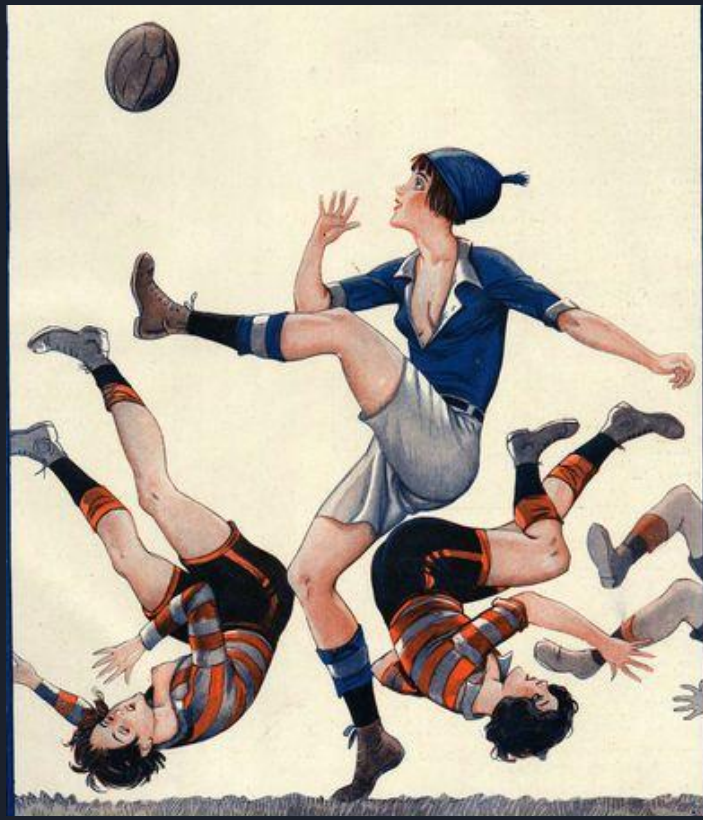
- Application of more models with higher accuracy or more direct insights demo video

Wellness and Performance in Rugby

SupeR squared, Smith College
Carol Liu, Ruby Ru, Abby Cui, Lily Jin, Rebecca Wang

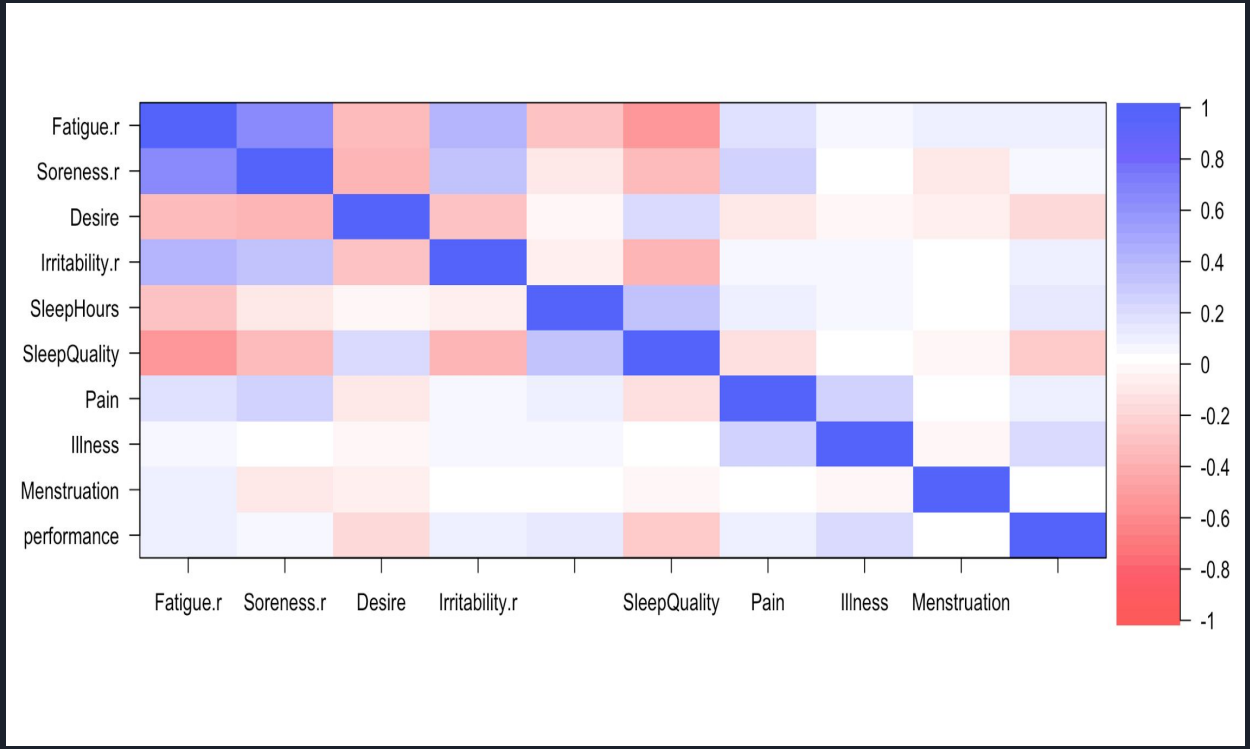
Understanding of the DATA

- Performance: Weighted average of four variables in GPS dataset: Mean(speed), Max(speed), Mean(accell), Max(accell)
- Factors considered: Fatigue, Soreness, Desire, Irritability, Sleep Hours, Sleep Quality, Pain, Illness, Menstruation, Performance.
- Special Note: Reverse coded fatigue, soreness and irritability.

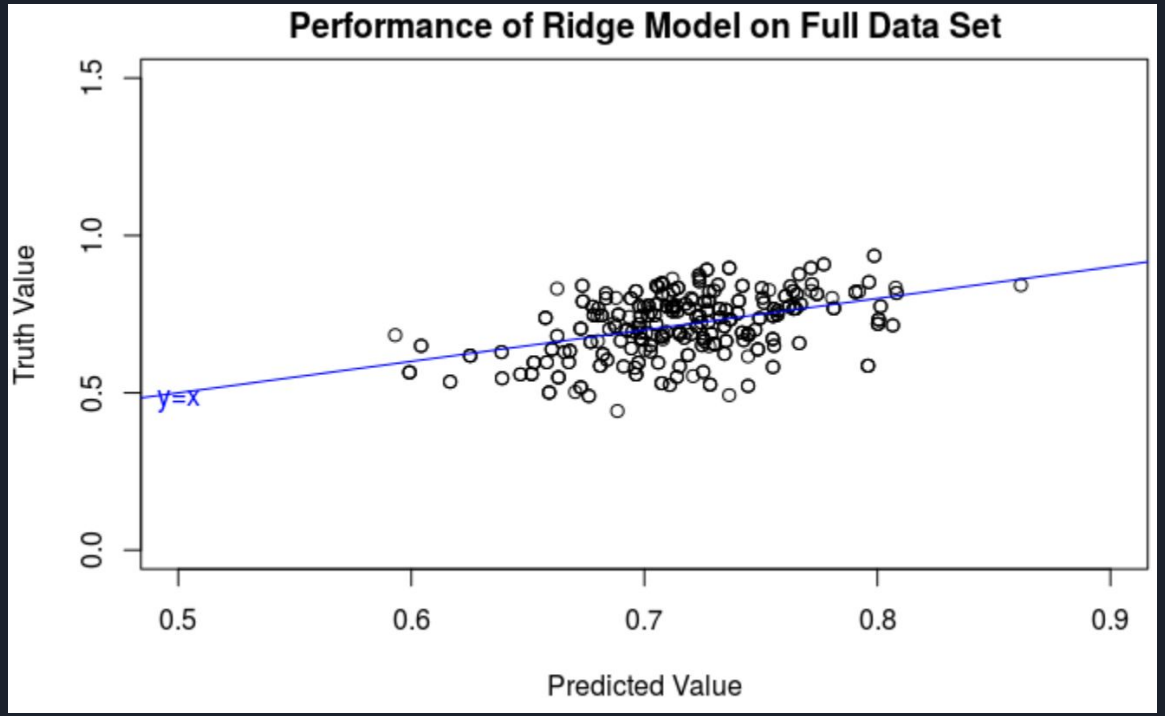


What are the correlations?

- Fatigue is positively strongly correlated with soreness (r = 0.64) and negatively strongly correlated sleep quality (r = -0.54).
- Both fatigue and soreness are weakly correlated with pain, illness, and menstruation.



Machine Learning Model:
Use Ridge Regression to Predict Performance



- Use Wellness data to predict performance
- Training set & Test set
- Test Mean Square Error: 0.0085
- Best result comparing to Lasso and Principal Component Regression

Results, Interpretations & Suggestions

(Intercept)	Fatigue	Soreness	Desire	Irritability	SleepHours	SleepQuality
0.728373482	-0.008867016	-0.046153185	-0.074925384	-0.058935602	0.019916013	-0.176153082
Pain	Illness	Menstruation				
-0.004157314	0.130421044	-0.003436111				

Future Research:

- Include number of assists, tackle number in Performance
- Response Bias

Suggestions:

- Sleep!
- Modify questionnaire
- Further Investigation on unexpected relationships

ggsquad2

Paige Patrick, Elaona Lemoto, Mathena Abramson, Ana Porro, Jordan Moody

Data : Canadian Women's 7s Rugby

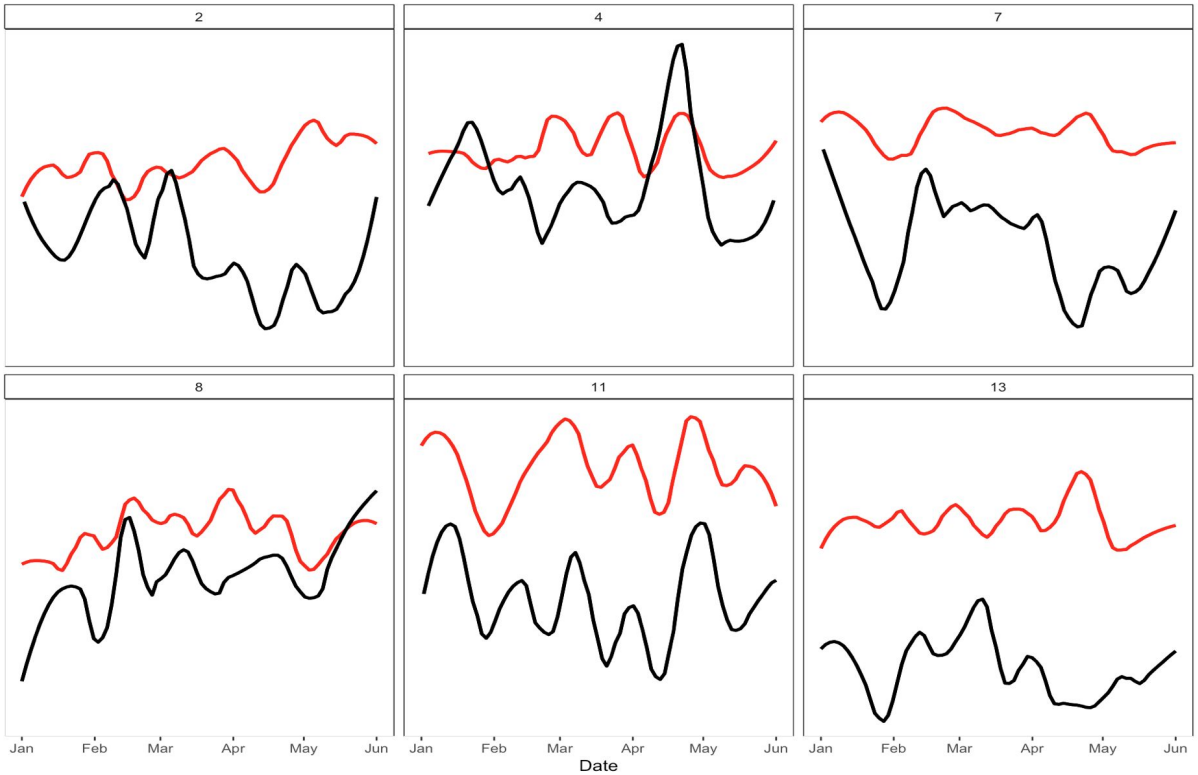
Goals:

- Understand relationship between training and fatigue
  - Quantify fatigue (rating vs numerical exertion)
- Is fluctuation relative to tournament outcomes?
  - Training synchronization

Variables:
Daily, for each player

- Session Load
  - Rate of Perceived Exertion (RPE)
  - Training duration (minutes)
- Fatigue
  - Self Reported (1 hour after waking)

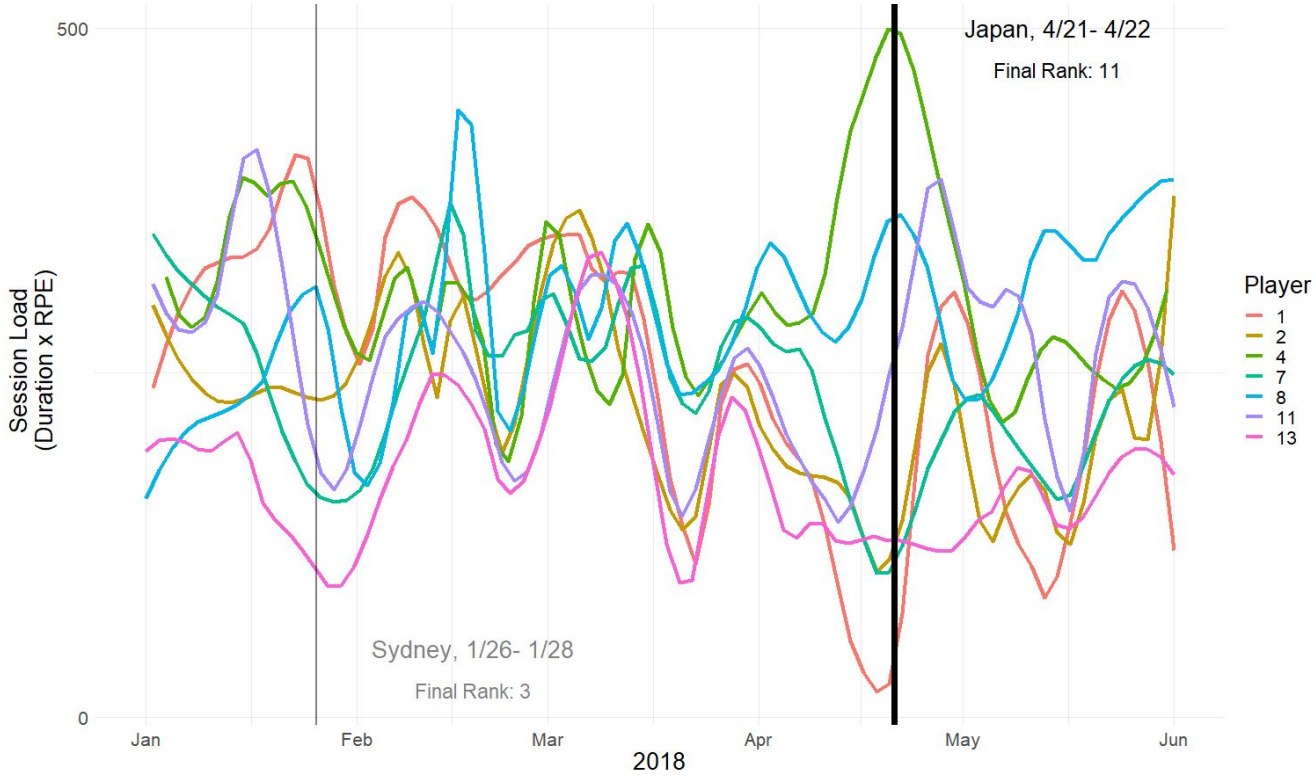
Fatigue and Session Load Trends for Top Six Players



Fatigue

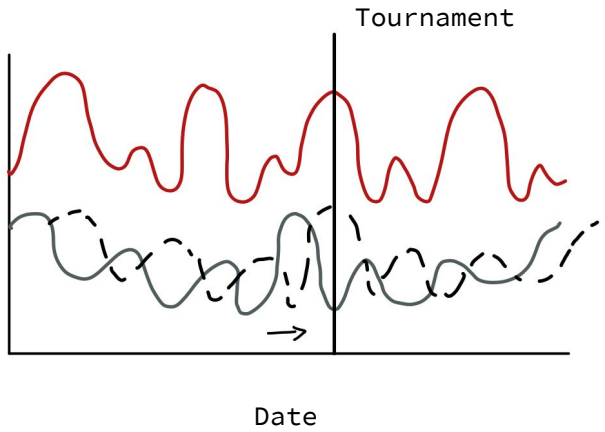
Session Load

Canadian Rugby 7s Top Seven Player's Session Load Trends

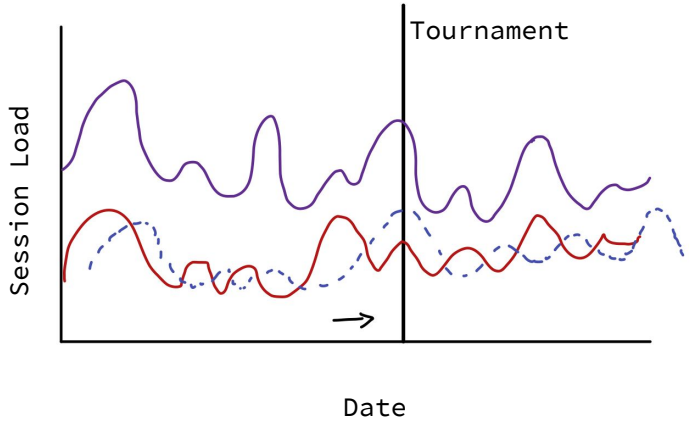


Actionable Steps

Individual: Maximize Corr.



Team: Sync Session Peaks



DataFest 2019

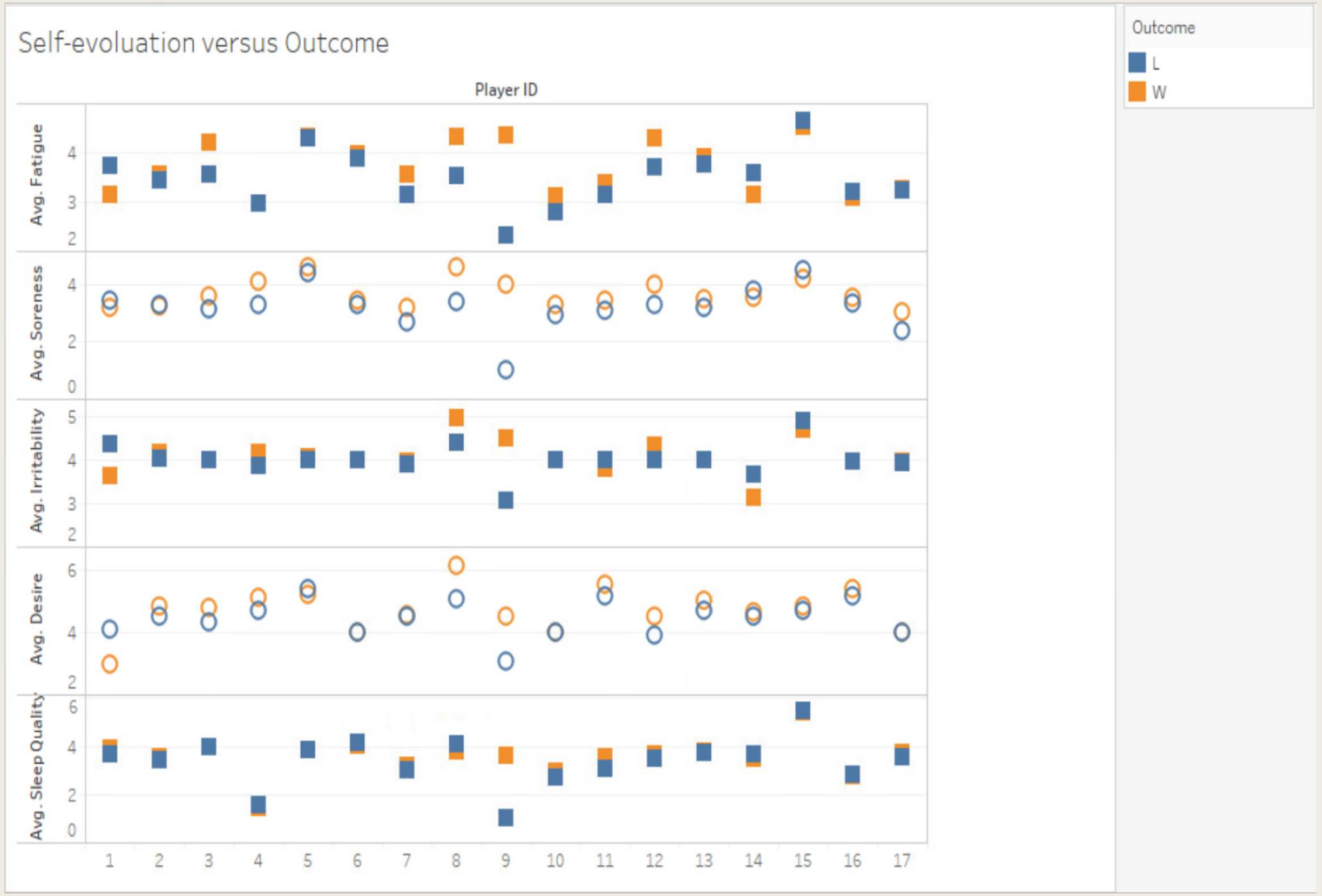
Fusheng Yang
Hao Gao
Ray Huang
Candice Zheng
Iris Li

Umass-Amherst

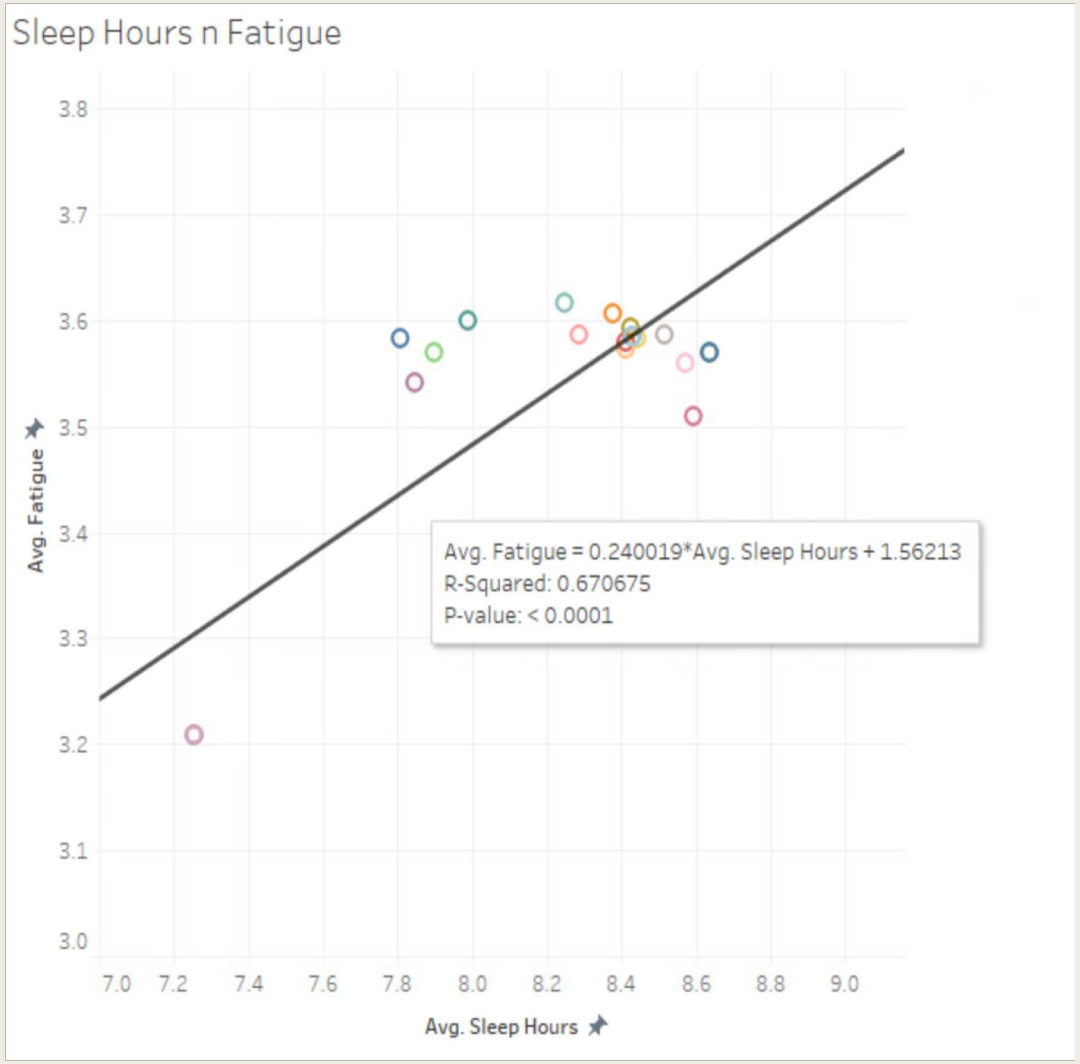
Initial Attempt

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

ALL FAILED!

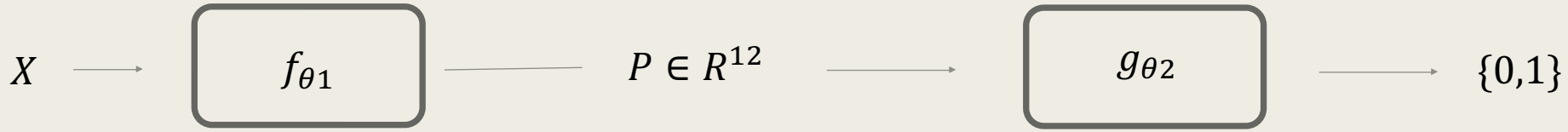


Fun Facts of Sleep Hours



First Attempt at modeling

Model the “Performance Score”



$f_{\theta 1}$: Model(NN etc.) that learn a function to estimate the performance score for each player

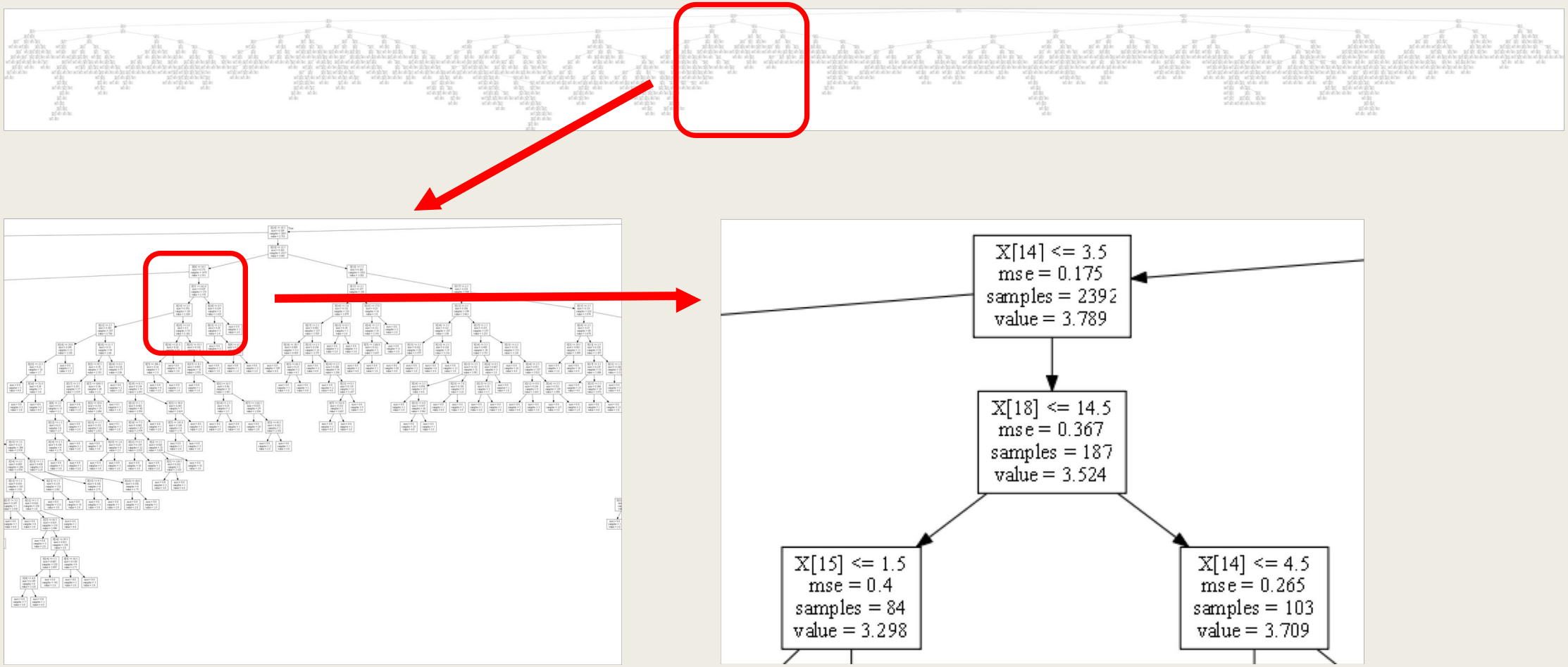
$g_{\theta 2}$: A Classifier that predict win/lose base on the performance score

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

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

$$K(x_1, x_2) = \left(\left|g_{\theta 2}\left(f_{\theta 1}(x_1)\right)\right|^2 - \left|g_{\theta 2}\left(f_{\theta 1}(x_2)\right)\right|^2\right) * \left(|f_{\theta 1}(x_1)|^2 - |f_{\theta 1}(x_2)|^2\right)$$

Random Forest Regression

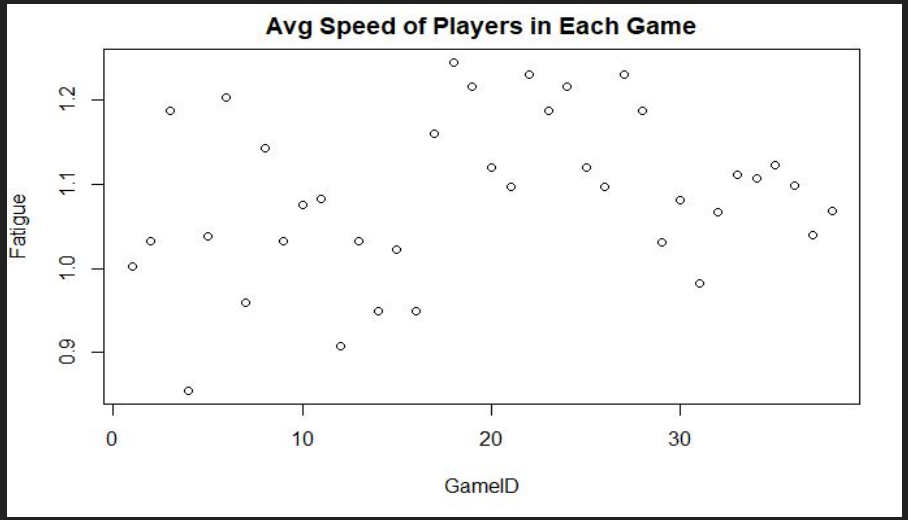


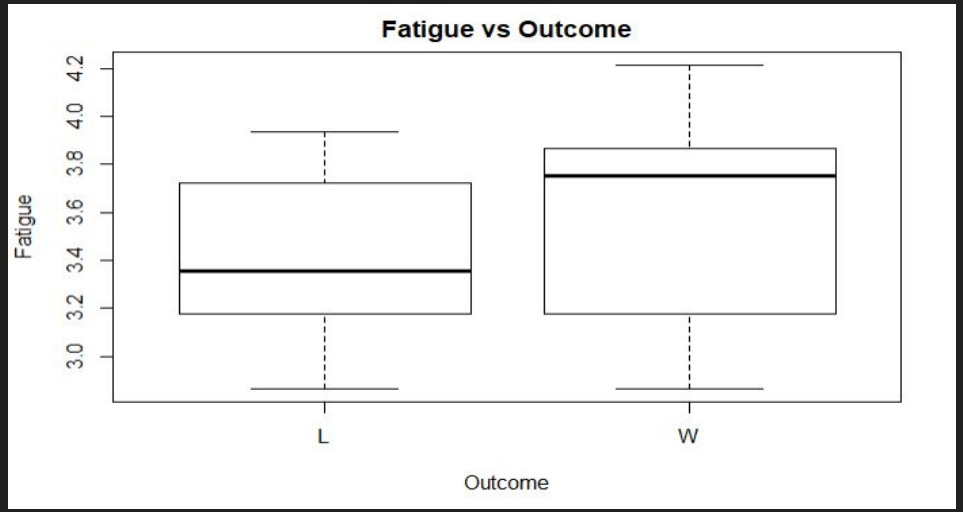
Effects of Wellness and Trainings on Fatigue

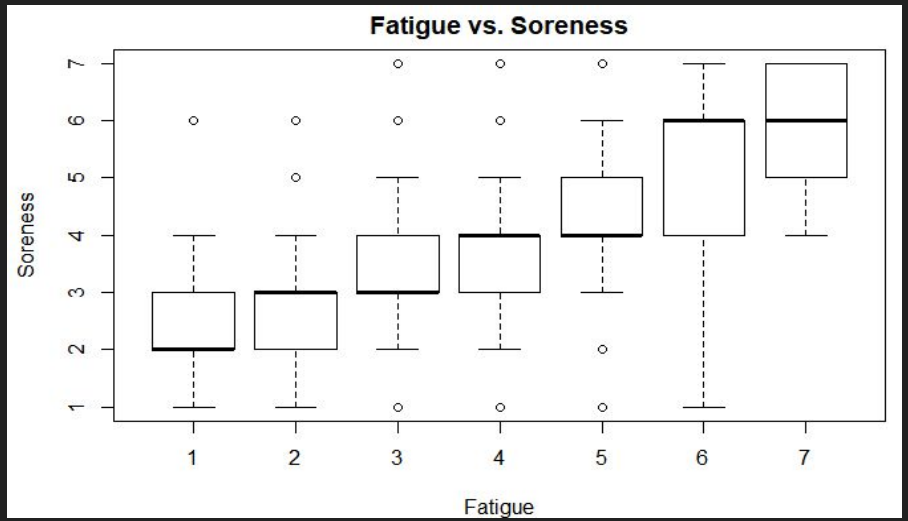
By
Team Significant
(Anthony Woo, Alexander Yang, Che-An Lin)

Explore the Data...

Finding relationship of fatigue and different variables ...







Linear Regression: Results

Looking at the significance

on each type of training ...

```
lm(formula = Fatigue ~ Desire + SleepQuality + Nutrition + Menstruation +
    SessionType + DailyLoad, data = merge.df)

Residuals:
     Min       1Q   Median       3Q      Max
-2.91997 -0.46974  0.07145  0.47586  2.60972

Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                 4.175e-01  8.424e-02   4.957 7.65e-07 ***
Desire                      4.817e-01  1.607e-02  29.981  < 2e-16 ***
SleepQuality                3.459e-01  1.355e-02  25.522  < 2e-16 ***
NutritionOkay               5.548e-02  3.094e-02   1.793  0.07309 .
NutritionPoor              -8.942e-02  1.379e-01  -0.648  0.51680
MenstruationYes            -8.528e-03  4.083e-02  -0.209  0.83456
SessionTypeGame            -2.410e-01  7.694e-02  -3.132  0.00176 **
SessionTypeMobility/Recovery -1.745e-01  5.549e-02  -3.144  0.00169 **
SessionTypeSkills          -9.321e-02  4.890e-02  -1.906  0.05672 .
SessionTypeSpeed           -3.422e-02  4.484e-02  -0.763  0.44547
SessionTypeStrength         1.059e-01  5.729e-02   1.848  0.06471 .
DailyLoad                  -2.584e-04  4.343e-05  -5.949 3.07e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6963 on 2507 degrees of freedom
  (4789 observations deleted due to missingness)
Multiple R-squared:  0.5123,    Adjusted R-squared:  0.5101
F-statistic: 239.4 on 11 and 2507 DF,  p-value: < 2.2e-16
```

Linear Regression: Results Cont.

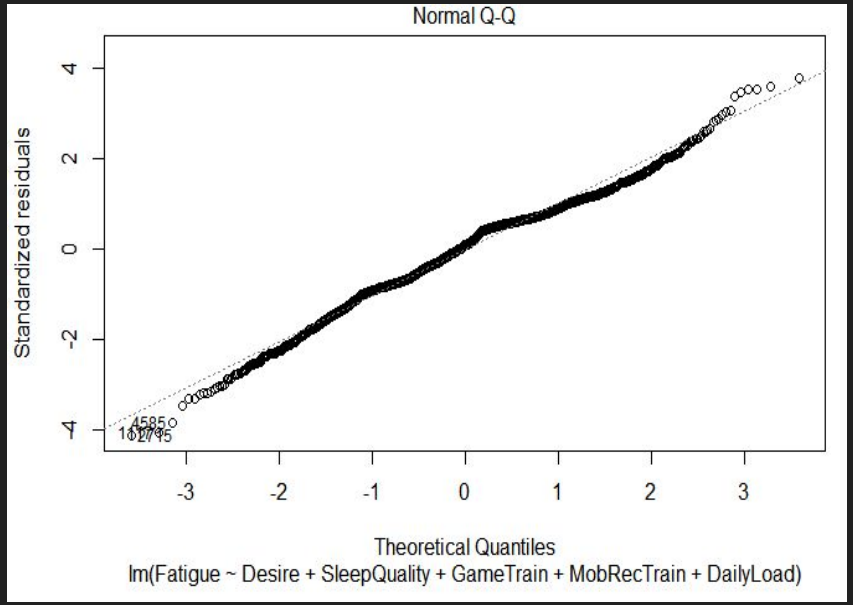
Remove the trainings…

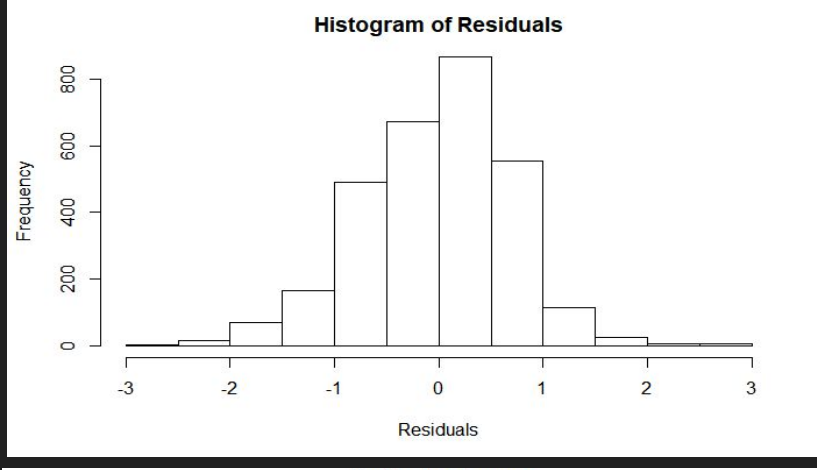
Now focusing at wellness

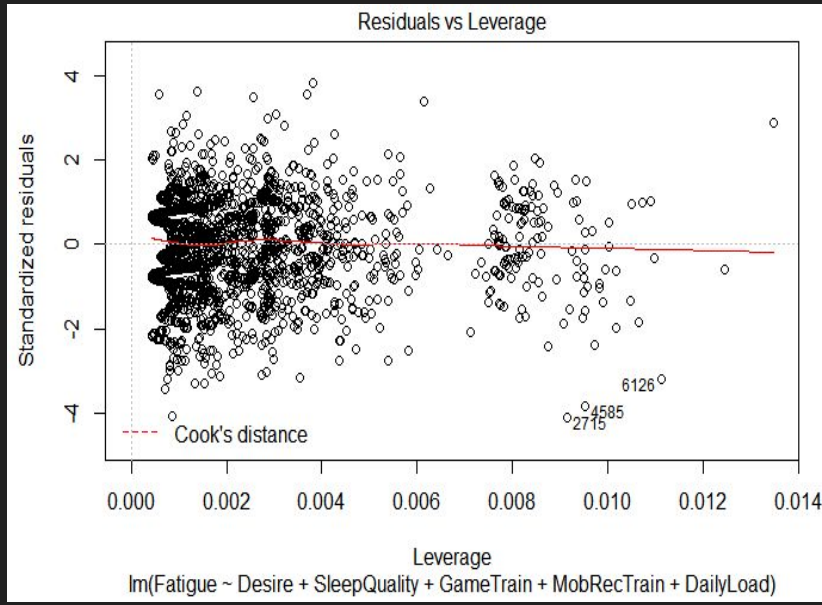
```
Call:
lm(formula = Fatigue ~ Desire + SleepQuality + GameTrain + MobRecTrain +
    DailyLoad, data = change.df)

Residuals:
     Min       1Q   Median       3Q      Max
-2.91539 -0.49879  0.07276  0.48237  2.70873

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     5.141e-01  7.066e-02   7.275 4.39e-13 ***
Desire          4.742e-01  1.474e-02  32.175  < 2e-16 ***
SleepQuality    3.310e-01  1.245e-02  26.581  < 2e-16 ***
GameTrainYes   -2.759e-01  6.491e-02  -4.251 2.19e-05 ***
MobRecTrainYes -1.399e-01  4.278e-02  -3.270  0.00109 **
DailyLoad      -2.725e-04  3.704e-05  -7.357 2.42e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7132 on 2981 degrees of freedom
  (4321 observations deleted due to missingness)
Multiple R-squared:  0.4763,    Adjusted R-squared:  0.4754
F-statistic: 542.2 on 5 and 2981 DF,  p-value: < 2.2e-16
```

Conclusion…

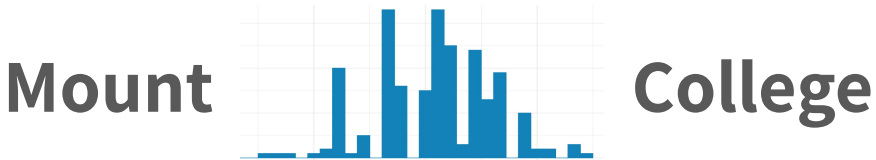
Training doesn't affect fatigue



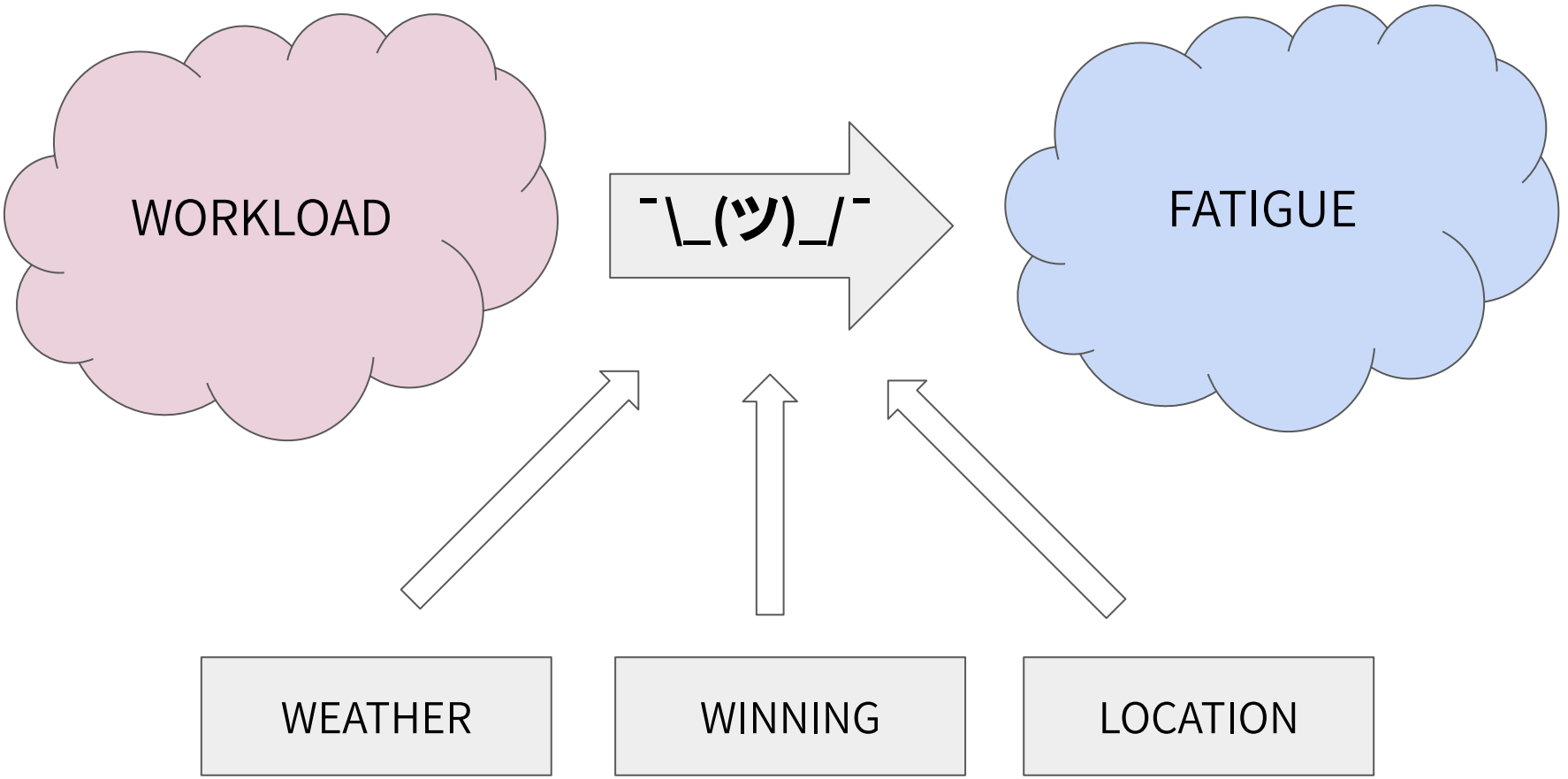




Effect of External Variables on Workload-Fatigue Relationship



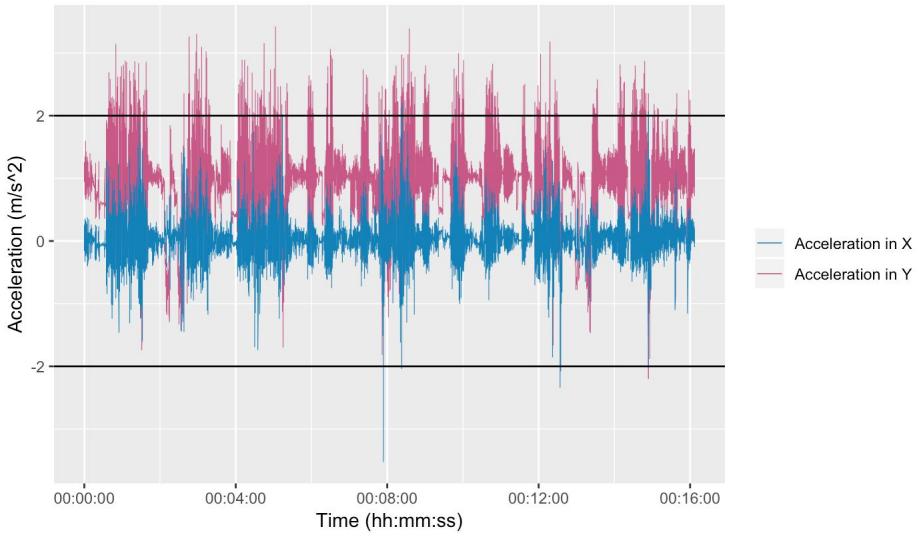
Emma Grotto, Eleanor Harris, Sara Pradhan


WORKLOAD
¯_(ツ)_/¯
FATIGUE
WEATHER
WINNING
LOCATION

Improving the Physical Preparation and Development of Women's Rugby Sevens Players

sprint distance, and mean sprint duration (s). Distance covered above 3.5 $m.s^{-1}$ are potentially a better measure of 'high intensity' for women than the commonly used 5 $m.s^{-1}$ threshold (25). Sprint distance was determined as the distance covered while accelerating >2.0 $m.s^{-2}$ for longer than one sec (90). The use of GPS-derived impacts >10g has not yet

Acceleration vs Game time a player



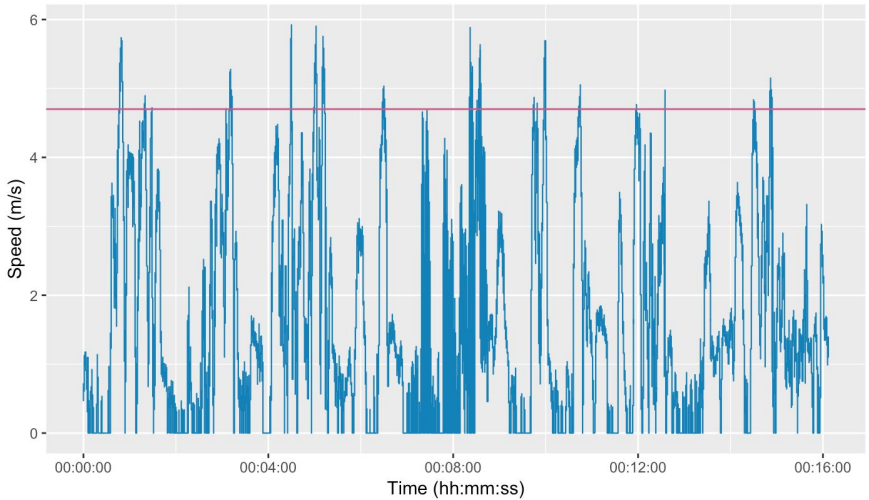
*http://www.canberra.edu.au/researchrepository/file/58674629-ecdc-4bcc-9d5b-1c55a90fea7a/1/full_text.pdf

Speed Demands of Women's Rugby Sevens Match Play

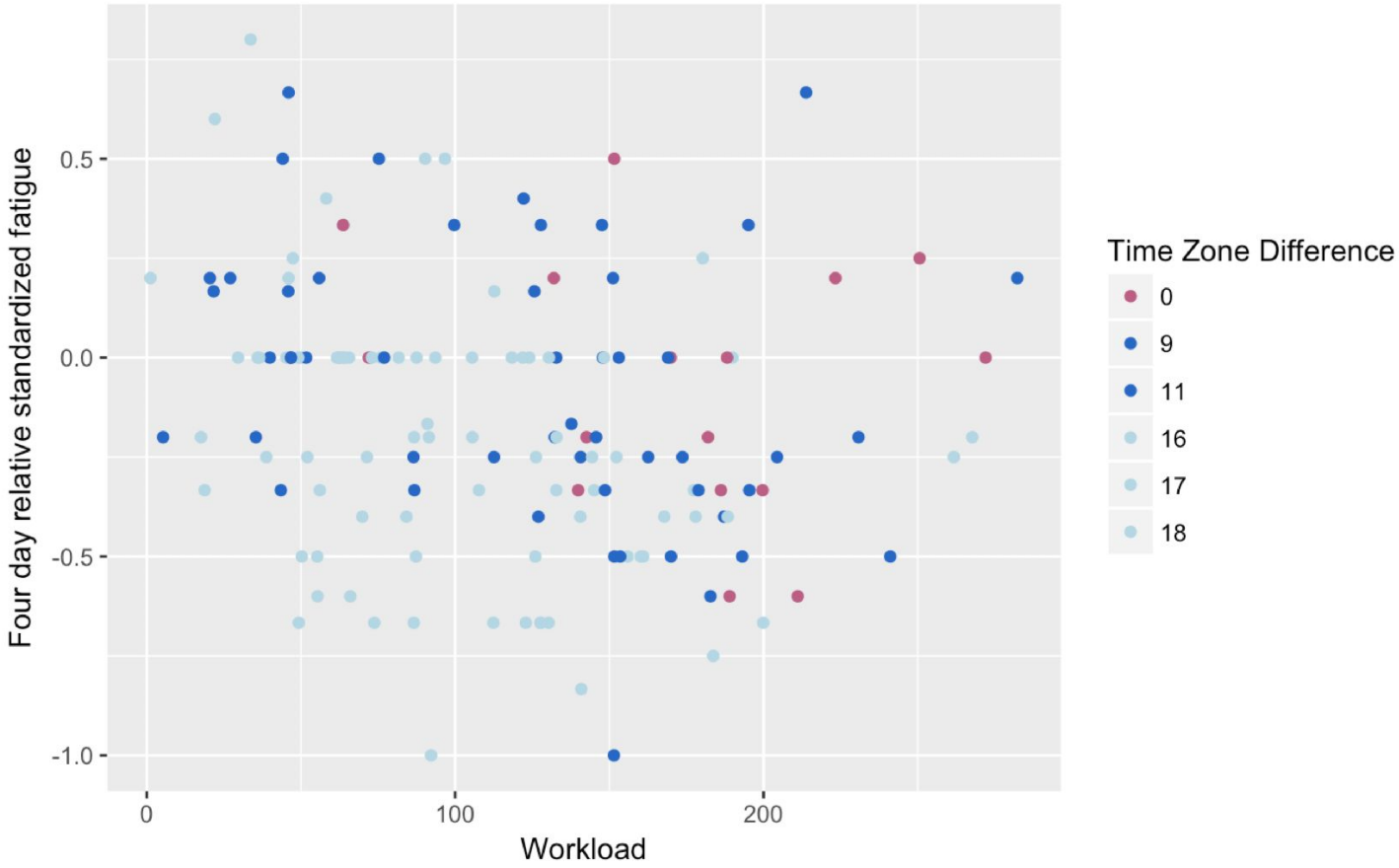
Nicole D. Misseldine,[1] Richard C. Blagrove,[2,3] and Jon E. Goodwin[1]

Players covered 1,556 ± 233 m per game, with "sprinting" representing 6 ± 4% of this total distance using the typical-standard "sprint" threshold (5.6 $m \cdot s^{-1}$), but a significantly ($p < 0.001$) greater 12 ± 4% using the female-adjusted threshold (4.7 $m \cdot s^{-1}$). Despite similar total distances,

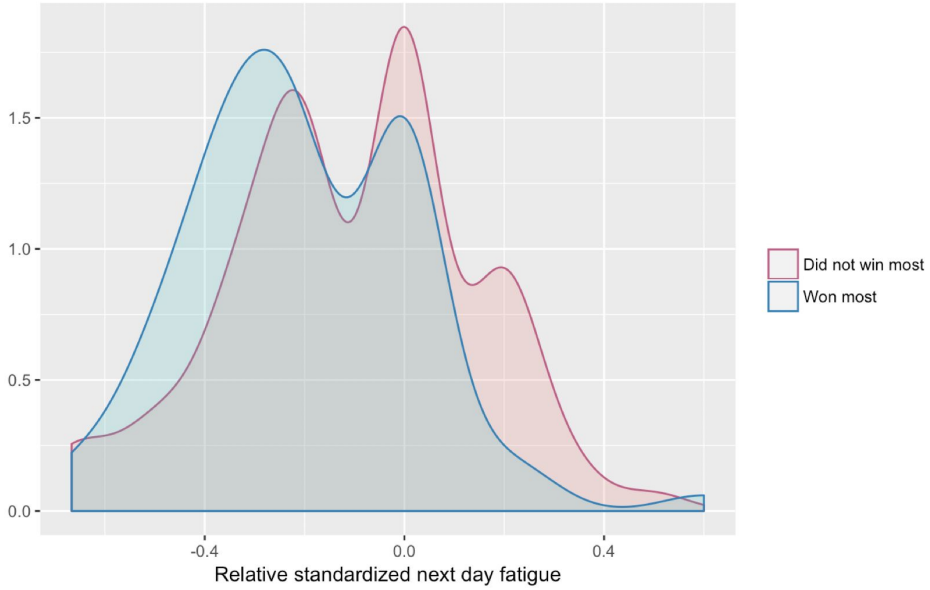
Speed vs Game time of a player

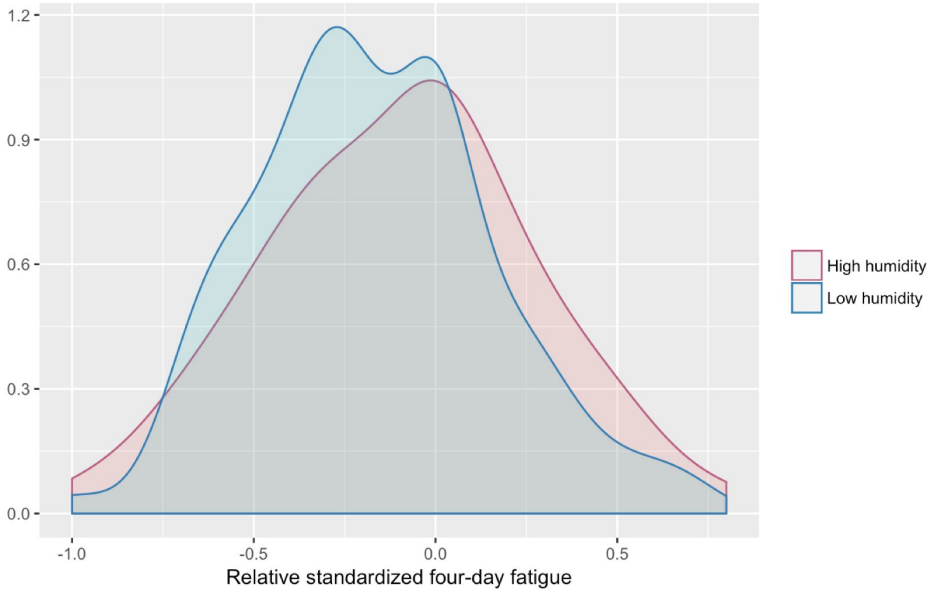


*https://journals.lww.com/nsca-jscr/Abstract/publishahead/Speed_Demands_of_Women_s_Rugby_Sevens_Match_Play.95342.aspx



How does time zone affect workload and fatigue?





Does winning change the workload-fatigue relationship?

Higher margin of winning predicts a faster recovery ($p < 0.05$).

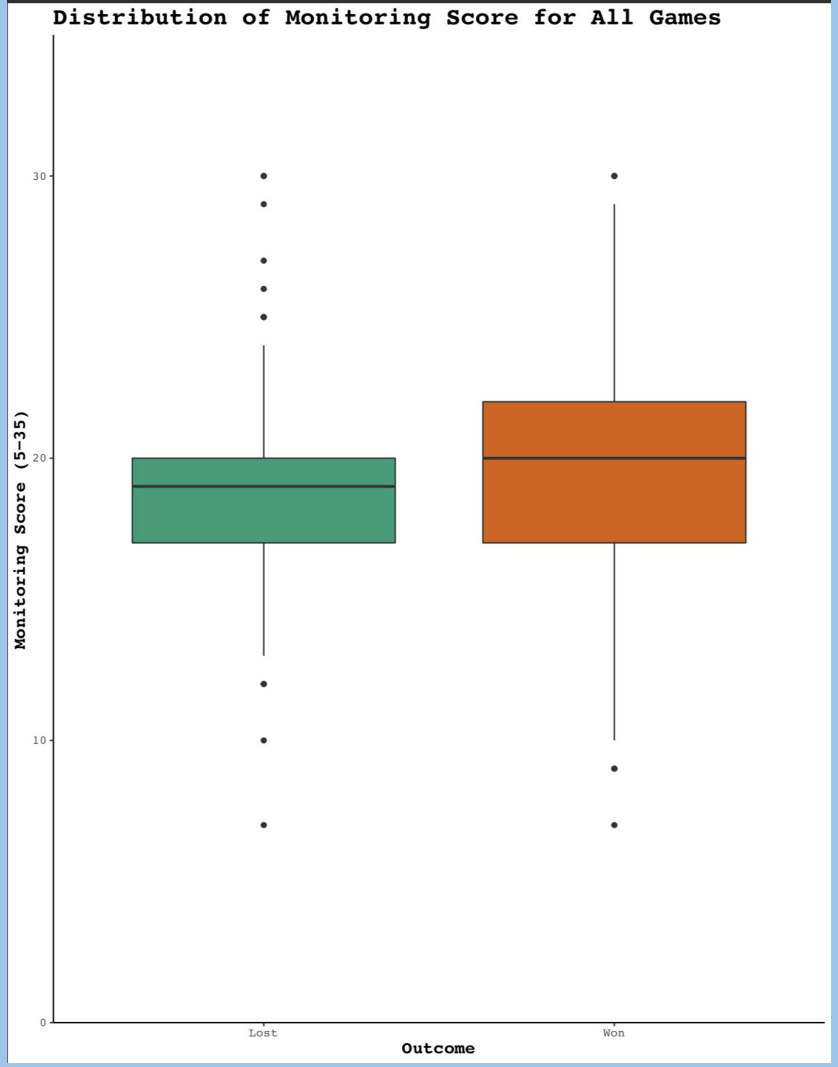
Does weather change the workload-fatigue relationship?

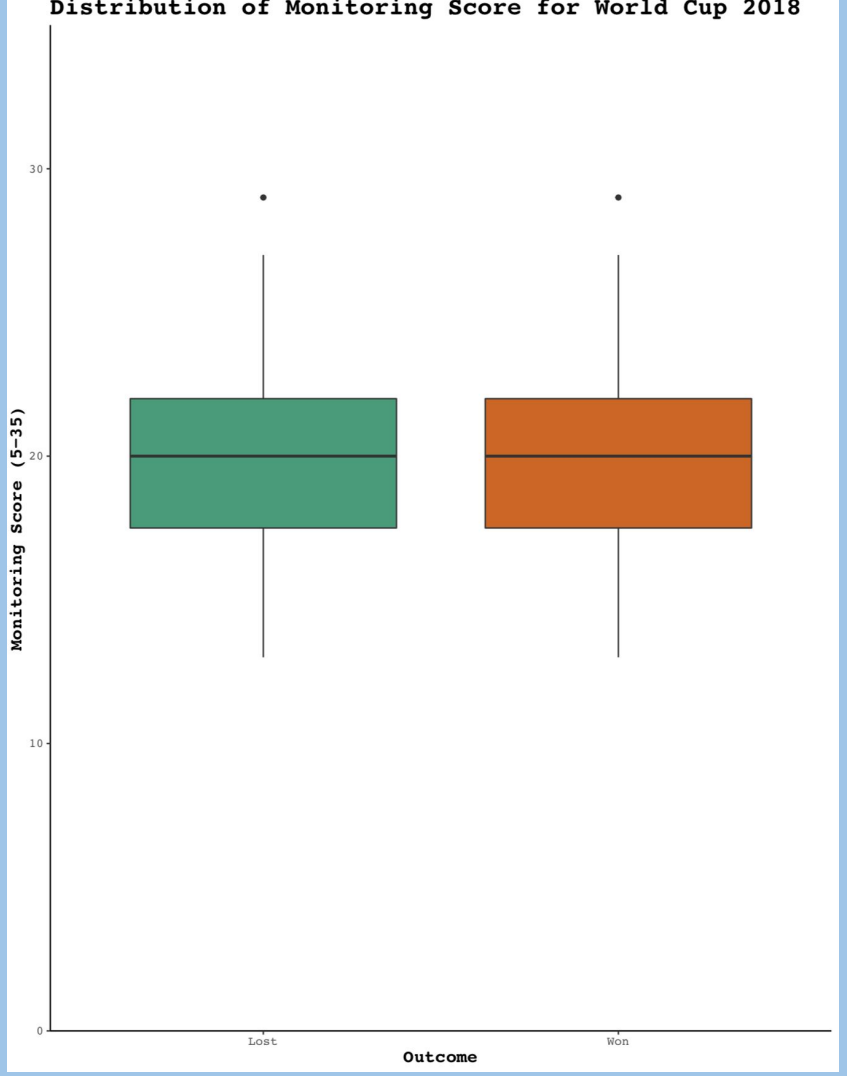
Using Riem package to get weather data for location/times.

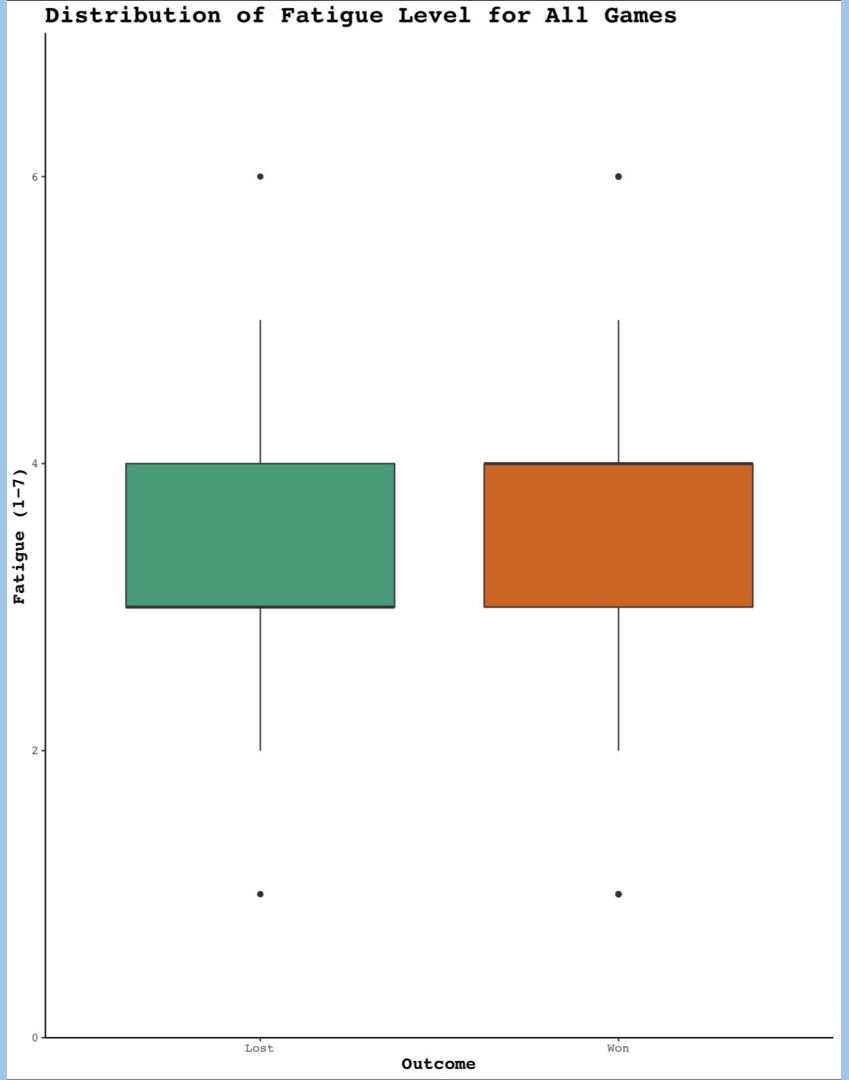
How Wellness Relates to Outcome in Rugby

Empty Environment:

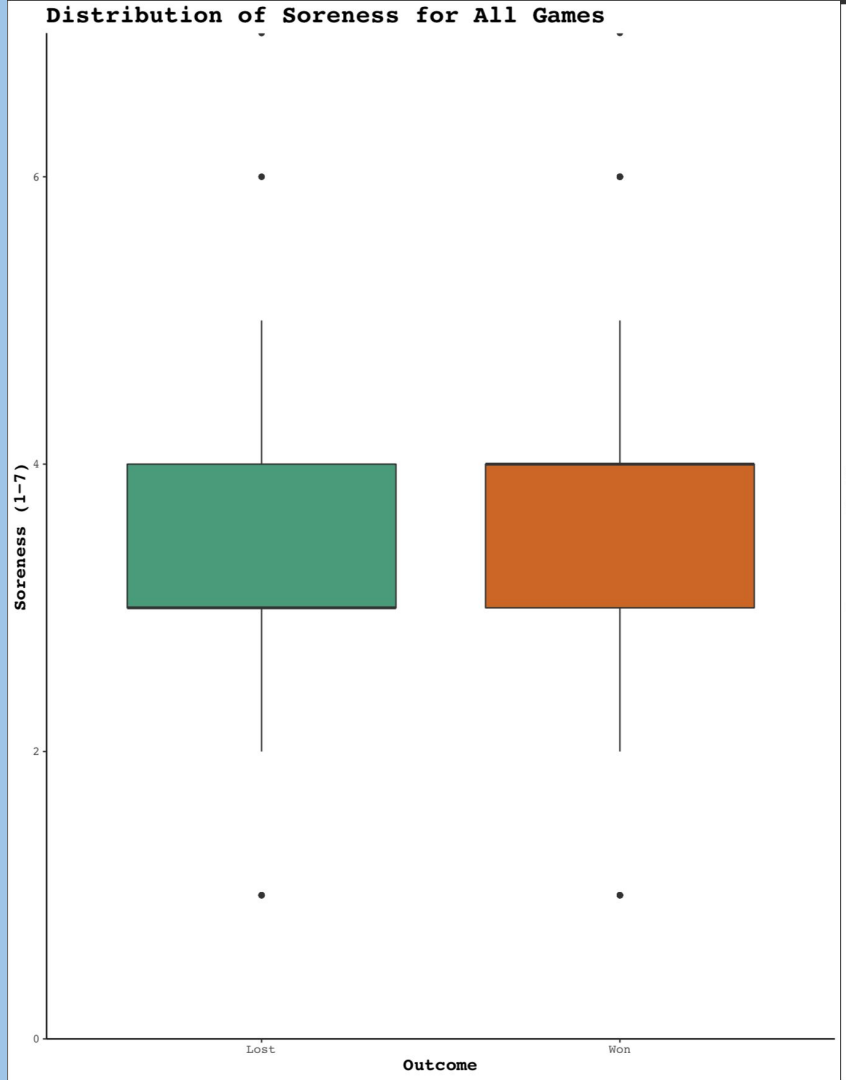
Chhiring Lama, Marium Tapal, and Grace Hartley

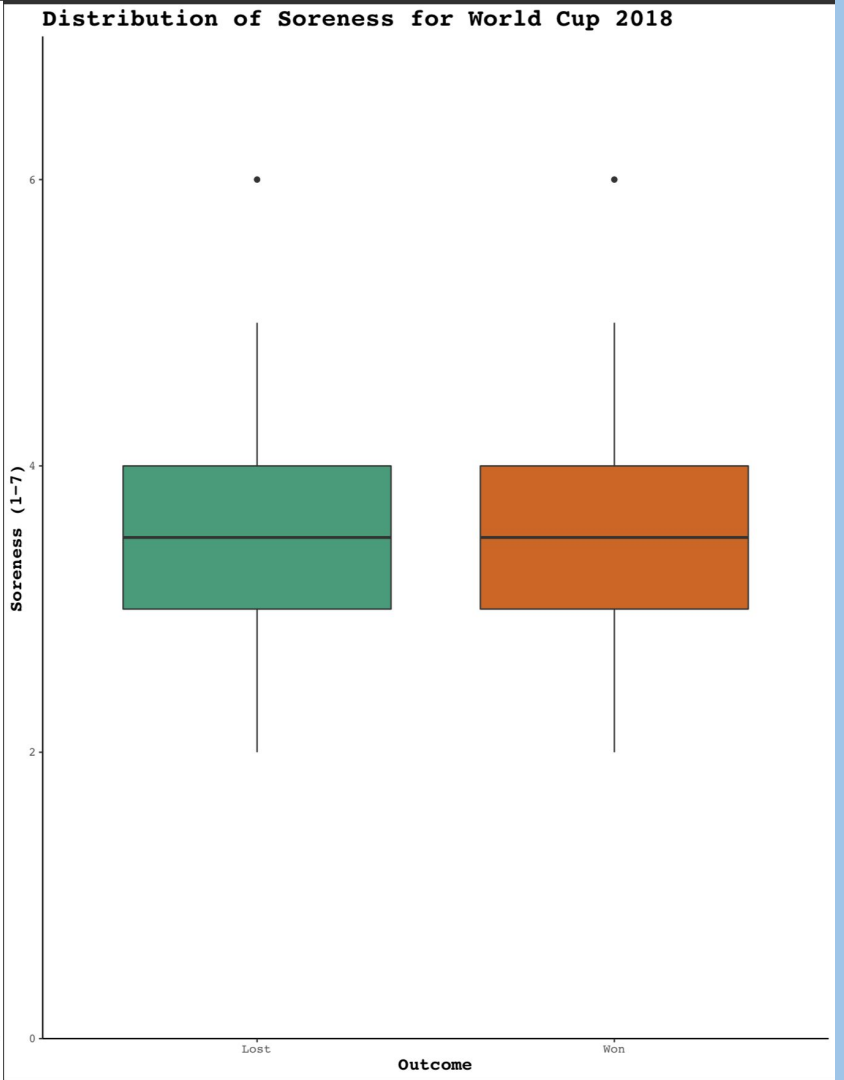

Distribution of Monitoring Score for All Games
Monitoring Score (5-35)
30
20
10
0
Lost
Won
Outcome


Distribution of Monitoring Score for World Cup 2018
Monitoring Score (5-35)
30
20
10
0
Lost
Won
Outcome


Distribution of Fatigue Level for All Games
Fatigue (1-7)
0
2
4
6
Lost
Won
Outcome


Distribution of Fatigue Level for World Cup 2018
Fatigue (1-7)
0
2
4
6
Lost
Won
Outcome


Distribution of Soreness for All Games
Soreness (1-7)
0
2
4
6
Lost
Won
Outcome


Distribution of Soreness for World Cup 2018
Soreness (1-7)
0
2
4
6
Lost
Won
Outcome

Conclusions

- More observations
  - Across more years
  - Frequency during games
  - Observations of specific players' performance
- More precise measures of wellness for higher-level performance
- Our limitations

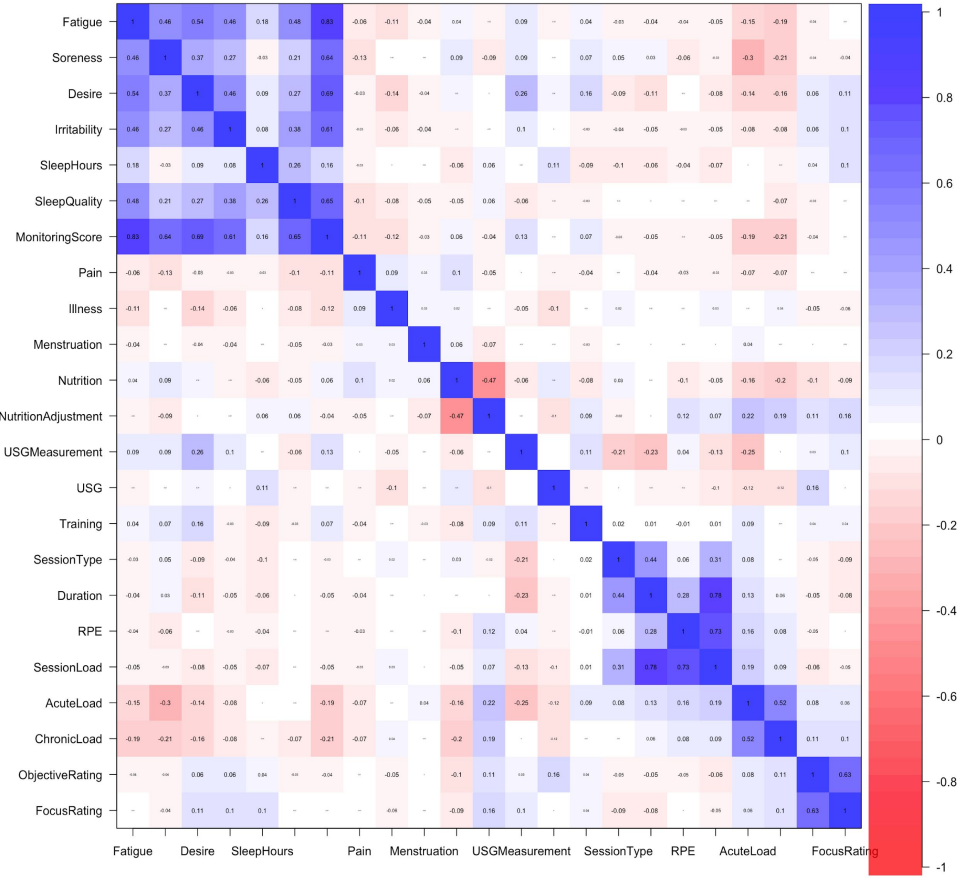
Data Fest 2019

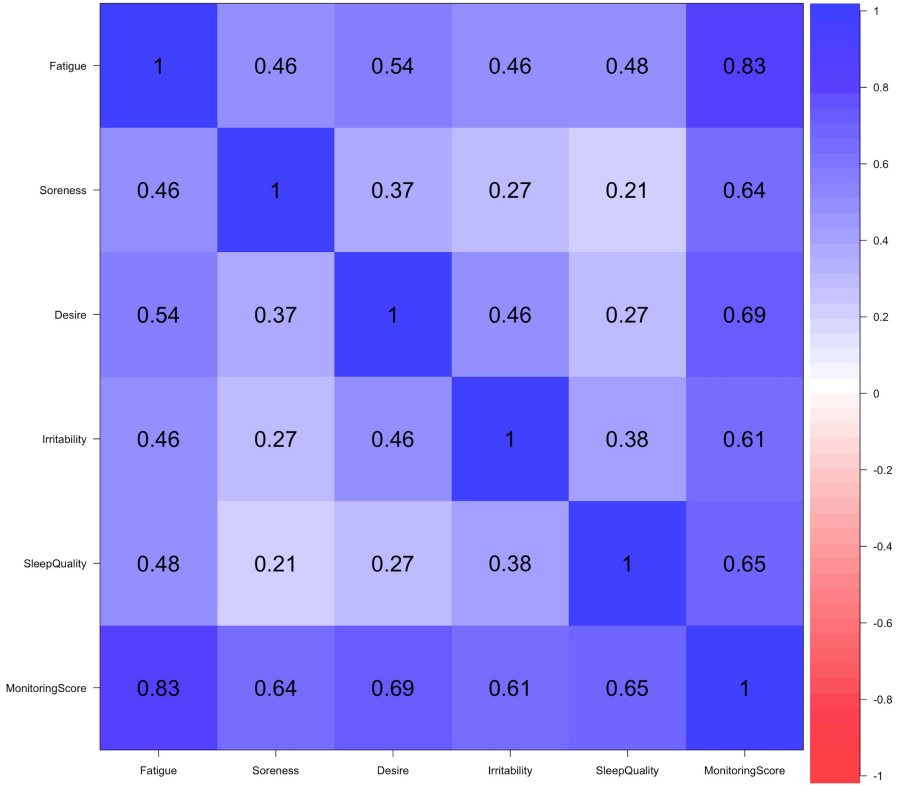
Ugonna Ezeaka, Rachael Abbott, Samad Abbasi,
Morgan White, Peter Cassels
-UMass Amherst-

Objectives

- Currently, training load is measured through combinations of subjective and objective measurements.
- We looked at these subjective measurements more in-depth.
  - Sport is as much psychological as physical
- Fatigue plays a critical role in a player’s performance but how do we define it?

Correlation Matrix





10-Fold Cross Validation

Naive Bayesian Algorithm

```
6345 samples
   5 predictor
   3 classes: 'Disagree', 'Average', 'Fresher than usual'

No pre-processing
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 5711, 5710, 5711, 5709, 5711, 5711, ...
Resampling results:

  Accuracy   Kappa
  0.7380657  0.5446714

Tuning parameter 'fL' was held constant at a value of 0
Tuning parameter 'usekernel' was held constant
 at a value of FALSE
Tuning parameter 'adjust' was held constant at a value of 1
```

Naive Bayes tries to classify instances based on the probabilities of previously seen attributes/instances

Confusion Matrix and Statistics

prediction_training_sample	Disagree	Average	Fresher than usual
Disagree	2430	333	11
Average	993	1904	138
Fresher than usual	39	144	353

Overall Statistics

Accuracy	0.7387
95% CI	(0.7277, 0.7495)
No Information Rate	0.5456
P-Value [Acc > NIR]	< 2.2e-16
Kappa	0.5458
Mcnemar's Test P-Value	< 2.2e-16

Statistics by Class:

	Class: Disagree	Class: Average	Class: Fresher than usual
Sensitivity	0.7019	0.7997	0.70319
Specificity	0.8807	0.7147	0.96868
Pos Pred Value	0.8760	0.6273	0.65858
Neg Pred Value	0.7110	0.8559	0.97435
Prevalence	0.5456	0.3753	0.07912
Detection Rate	0.3830	0.3001	0.05563
Detection Prevalence	0.4372	0.4783	0.08448
Balanced Accuracy	0.7913	0.7572	0.83593

Conclusion

- Our Model does a good job of predicting fatigue.
- Moving forward, we could perform some more exploratory and regression analyses to further study fatigue levels.
- For now, it would be advisable to implement methods that decrease fatigue, focusing on these model parameters, to improve the player's overall health and increase performance in games.
- Also we would need consistent data collection for other omitted variables to improve analyses.