ASA | DATAFEST 2019

GoodGame

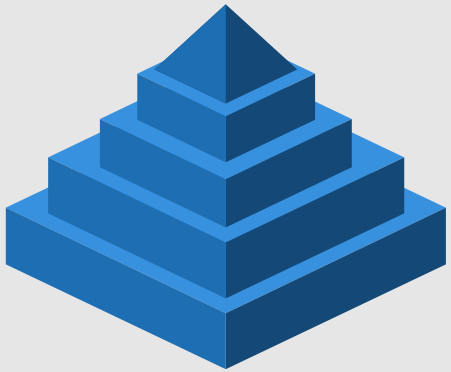
Daksh Jain & Yash Tyagi.

01 | Document Parsing

Due to large data set, we decided to focus on data from wellness.csv to form a expression for it from the self reported values.

02 | Cleaning

A major challenge was cleaning up the data. In wellness.csv, 5 columns had more than 50% of their values listed as NA. And quite a few of them were binary string or arbitrary variables which need to be scaled to some numeric value.

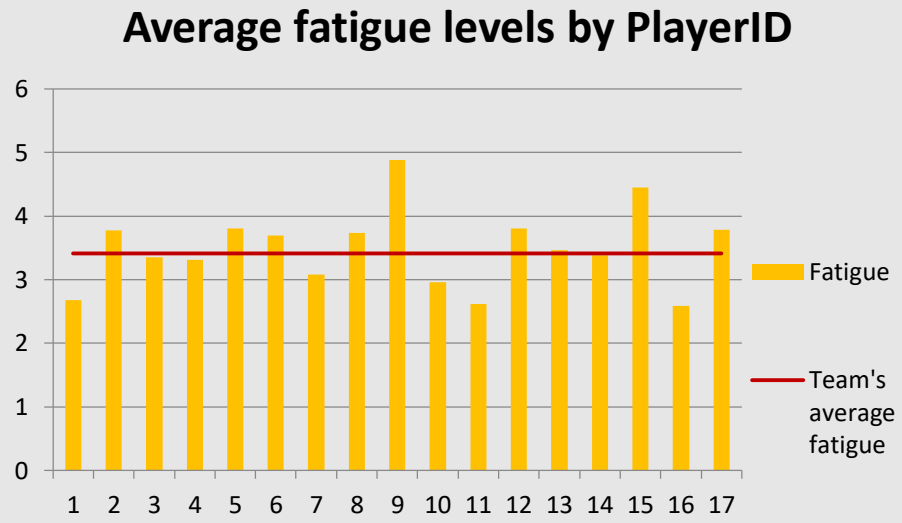


03 | Testing

We came up with an agreeable system to convert the variables to numeric values. Then we carried out testing to check their viability. And finally decide to sort the data by Player ID to process it further.

04 | Averages

From wellness.csv. We found individual averages and team average of all the players in the team. This gave an indication of the average condition of the players. Fatigue levels can be seen on the right.



05 | Further exploration

After averaging we decided to create a comprehensive equation to predict fatigue levels based on the provided data points.

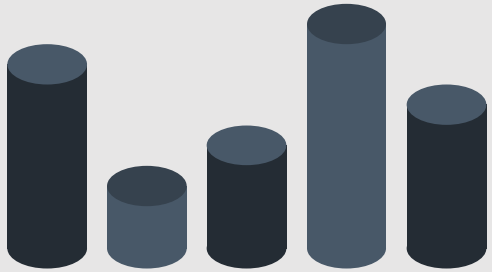
Correlogram							
	Fatigue	Soreness	Desire	Irritability	SleepHours	SleepQuality	MonitoringScore
Fatigue	1						
Soreness	0.5199	1					
Desire	0.5802	0.4259	1				
Irritability	0.4815	0.3292	0.4511	1			
SleepHours	0.1858	0.0245	0.0675	0.0888	1		
SleepQuality	0.5254	0.2818	0.3118	0.4028	0.2818	1	
MonitoringScore	0.8463	0.6943	0.7455	0.6886	0.1845	0.71	1

06 | Correlation

We made a correlogram which included all the variables in the welleness.csv. And based on the their correlation with fatigue we decided to regress the latter against them.* (The above shows the variables we decided on to use.)

*Data with more values available on request.

Fatigue	Coef.
Soreness	-1
Desire	-1
Irritability	-1
SleepQuality	-1
MonitoringScore	1
_constant	2.71E-13



07 | Regression

Various regression models were run and tested. Most of them very close to each other however we ultimately selected one with the following variables: Soreness, Desire, Irritability, SleepQuality and MonitoringScore. *

*Data available on request.

08 | Quantification

Therefore, fatigue can be quantified as:

Fatigue = MonitoringScore – Soreness – Desire – Irritability - SleepQuality