1. The team is using force plates to evaluate the neuromuscular characteristics of the players using a CMJ. Please answer the following questions

a. Provide a definition for each of the metrics collected. As much as possible use simple terms like if you were explaining it to a coach.

Concentric Impulse:
Concentric impulse is the amount of force applied during the time spent in the concentric phase of the jump. We filter out any outliers by looking for tests greater than 3 std's away from mean, and then take the average of the remaining jumps.

Concentric Impulse - 100ms:
This metric is the amount of force applied during the first 100ms of the concentric phase of a jump. We filter out the outlier jumps using the criteria from above and then take the average of the metric between remaining jumps.

Concentric Mean Power / Body Weight:
This metric is how much power was produced during the concentric phase of a jump relative to an athlete's body weight. This metric describes how powerful concentrically an athlete is relative to their size.

Eccentric Mean Power / Body Weight:
This metric is how much power was produced during the eccentric phase of a jump relative to an athlete's body weight. This metric describes how powerful eccentrically an athlete is relative to their size.

RSI - Modified:
This metric describes how 'explosive' an athlete is, that is how high they jump relative to how fast it takes them to get off the ground.

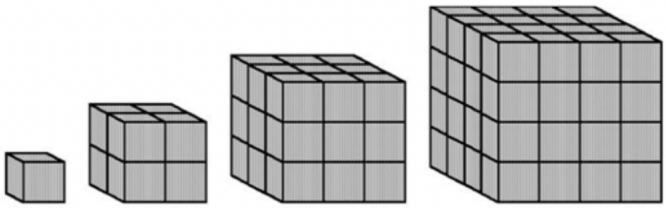
Vertical Velocity @ Take off:
This metric describes how rapidly an athlete displaces their center of mass vertically at takeoff.

b. What are your thoughts on the overall metrics collected within the context of this team (amateur football)? Would recommend removing or adding any metrics moving forward? Please explain the rationale behind your answer

I think the metrics listed are pretty good, but while at Driveline I found peak power in the CMJ and Squat jump to be the most predictive of sport success and used it in models. I also think it's valuable to track countermovement depth to flag if a significant change in jump mechanics may be a reason for increased jump height. In addition, the relative metrics are nice but I would prefer it to be allometrically scaled as not to be biased towards lighter athletes.

Allometric scaling

Let's represent a muscle (or the force generator) with a cube with the side length L



Length	1	2	3	4
Surface	1	4	9	16
Volume	1	8	27	64
Surface / Volume	1	0.50	0.33	0.25

Figure 4.8. Small World model of the muscle using cube with side lengths L.

The surface of the cube (one side surface) is proportional to the cross-sectional area of the muscle, and hence directly proportional to the maximal strength. The volume of the cube is proportional to the weight of the muscle.

As can be seen, the ratio of surface to volume, or strength to weight, is not linear, because weight increase much quicker than surface area. That's why simple strength ratio is biased against heavier individuals.

This excerpt from Mladen Jovanovic's book "Strength Training Manual" explains the importance of this. Since muscle CSA doesn't increase linearly with muscle weight we must consider this when exploring relative metrics.

2. Some of the assessments in the data contain multiple trials for each test. For example, the linear running speed or the CMJ test have multiple trials. Please develop a solution to report the results from tests with multiple trials. How do you calculate the results, how do you decide when to remove or keep a trial? Work with the dataset and develop a solution to report only one result per assessment. The new data set must be used to complete questions 3 and 4.

a) We can say that any trials that fall +/- 3 standard deviations away from the mean of trials should be removed as an outlier and then the average of the remaining should be taken to represent an athlete's score for the day. The solution was implemented starting in line 36 of the blue_crow_project.R file.

3. The director of strength & conditioning for Blue Crow Sports is interested to know whether 30Y sprint times can be predicted with the data that we have available and he/she is asking if you can help. How do you approach this question analytically? Please elaborate a solution and provide both the code and process that you follow to come up with an answer as well as an explanation of your results.

They might, first thing I would do is see what metrics correlate strongest to 30Y sprint times and whether or not that relationship is significant. Then I run predictive power scores to see if maybe there is a significant non-linear relationship that exists. If a strong linear relationship exists, I'd build a simple linear model. I looked into it and we probably can create a model to predict 30y sprint times, as these metrics look to share a strong relationship with 0-30y sprint time. The issue is we don't have enough data to create a model I'd be confident in. Only 216 athletes had 0-30y times and of those only 4 also had concentric impulse and vertical velocity at takeoff recorded the same day. A solution to potentially increase the sample is saying any assessments within 1 week of the 0-30y sprint test could be paired. My issue is the more time between test types, the more different the athlete is and those jump metrics could be impacted by something like fatigue/training load causing odd relationships between data.

	x0_30y
height_in	1.00
x0_30y	1.00
x10_30y	0.84
concentric_impulse	0.81
x0_10y	0.78
body_weight_lbs	0.61
concentric_impulse_100ms	0.29
lateral_assymetry_percent	-0.05
left_hamstring_strength	-0.13
right_hamstring_strength	-0.17
eccentric_mean_power_bm	-0.65
rsi_modified	-0.93
concentric_mean_power_bm	-0.94
vertical_velocity_at_takeoff	-0.97

4. Blue Crow would like to have an in-house application to report this information and ensure that coaches are able to access it on a daily basis. Please create an interactive application that contains the following functionalities

a. A tab that enables users to select an athlete and see the most recent results for all tests, shown as a percentile. Provide some context for users to understand what is a "poor", "good", "average", etc, results

b. A second tab that enables users to filter by position, test and metric and display a table or a chart to show the reference values from all the historical data

c. A third tab that lets users filter by player, test, metric, date and display a longitudinal trendline to track changes over time. Try to provide some context, for example, how does the coach know when there is a meaningful improvement?

An application developed programmatically would be preferred, such as R Shiny, Python Dash, etc. Please provide a link to the live app as well as the code.

Link to app: https://aosnacz.shinyapps.io/blue_crow_dash/