**The Dataset:**

Please find attached the dataset to be used in this project. This data contains simulated results from various performance assessments collected over various years. For more information about each of the metrics please refer to the following table:

|  |  |  |
| --- | --- | --- |
| **Assessment** | **Metrics** | **Explanation** |
| 30-15 IFT | Final Speed (km/h) | The speed of the final stage completed before being removed from the test. |
| V02max | An estimation of the VO2max value for the athlete based on the final speed of the test, age, body weight and gender. |
| 0-30Y Linear Speed | 0-10Y Split | Time to complete the first 10Y in seconds |
| 10-30Y Split | Time to complete the last 20Y in seconds |
| 0-30Y Split | Time to complete the total distance from 0 to 30Y in seconds |
| Anthropometry | 7 Site Fat % | Overall body % percentage taken using skinfold methods from seven sites |
| Weight (lb) | Body weight in pounds |
| Height (in) | Player's height in inches |
| CMJ (Force Plates) | Concentric Impulse |  |
| Concentric Impulse - 100ms |  |
| Concentric Mean Power / Body Weight |  |
| Eccentric Mean Power / Body Weight |  |
| RSI - Modified |  |
| Vertical Velocity @ Take off |  |
| Max Strength (Trapbar Deadlift) | Relative Max Strength | The value of the estimated 1RM divided by the body weight of the athlete |
| Estimated 1RM | The estimated 1RM based on weight and number of repetitions using Brzycki's method |

**Tasks**

**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.

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

**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.

**3.** The director of strength & conditioning 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.

**4.** Design an in-house flexdashboard 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 such as "good", "average", "poor", "excellent", etc.

**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?

Upload to your tool to your Github page. Additionally, please load the .rmd file to D2L. Please provide a link to the Github site within your .rmd file.