



PM566 Final Project

2023 Post-Season Performance Review - Full Report

Physical and Wellness Metrics - FC566

Introduction

Player tracking is an ever-increasing practice in professional sports and is even trickling down to college and youth sports. Collecting GPS, accelerometer, HR, and self-reported wellness data can provide a means for coaches to adjust their training plans in an effort to minimize injuries and optimize athletic performance. These data can be utilized to answer questions regarding player readiness, training effort or load, potential cause/risk for injury, effectiveness of training, and more. Data from games can also be used to determine the desired load during training leading up to the game. It is best practice for the total training load throughout a week to be greater than the maximum potential load experienced during a game. A common practice is prescribing 180-200% of the game load, spread across training sessions throughout the week.

Understanding each metric is important to the proper prescription of training throughout the week. All physical metrics are obtained through GPS tracking devices. Distance is the total distance covered during an activity, acceleration is a bout of movement above the 3.5 m/s/s acceleration threshold, deceleration is a bout of movement below the -3.5 m/s/s acceleration threshold, and max velocity is the highest speed (kmh) achieved within an activity.

The goal of this data exploration is to identify any relationship between the team average of physical performance metrics during practice based on the number of days until the next game (this is going to be referred to as “training day” throughout this report.) I will also look at relationships between game data, position, and max velocity. Finally, I will assess player recovery by looking at the sleep patterns over time.

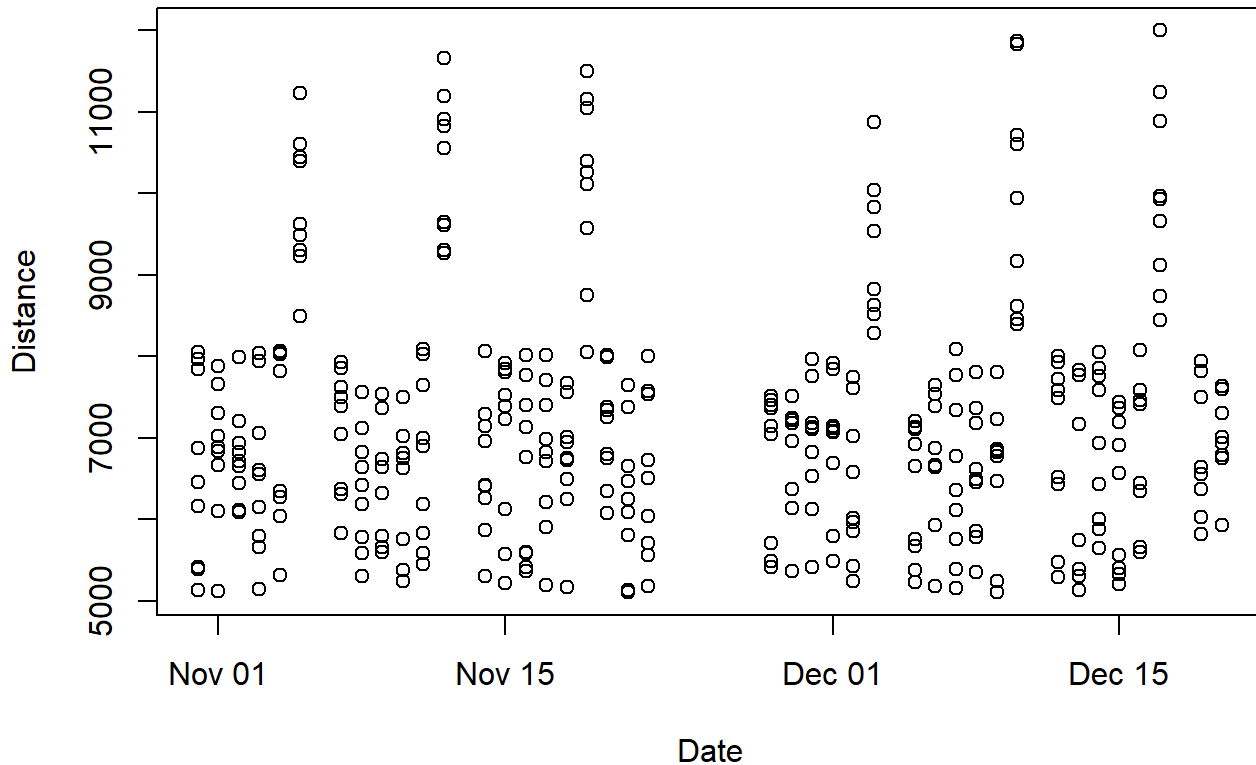
Methods

This data set was provided as a free resource for Sports Scientists to get exposure to working with these types of data without having to be working with a team, or without having to use potentially sensitive data from the team they are working with. This is training, wellness, and injury data for 10 hypothetical soccer players created by Micheal Hegedusich. These players now play on the fictitious team, Football Club 566!

There are seven cases in which performance data was recorded, while wellness and RPE was left blank by the athlete resulting in an NA. For RPE, I will use the mean of the RPE on the date of the missing value, as this will be the most accurate for the training session. However, for other wellness data, I will use the mean for each player to

best fill in the missing data. Finally, I am going to remove all goalkeeper data, as this position has very different demands of a field player and will interfere with the interpretation of our data.

For my first question, I want to explore the relationship between the training day (number of days before/after a game) and the physical output metrics.

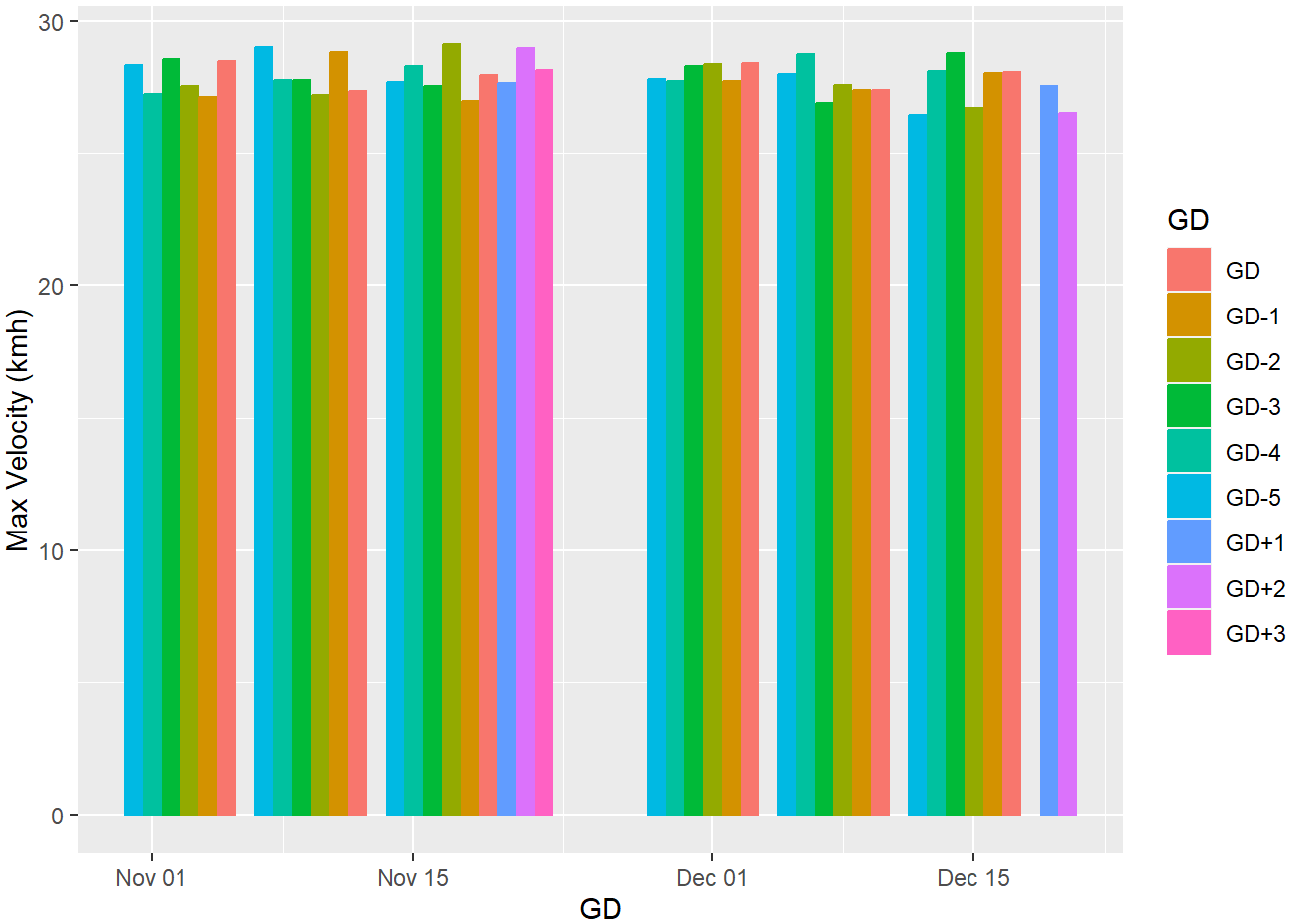
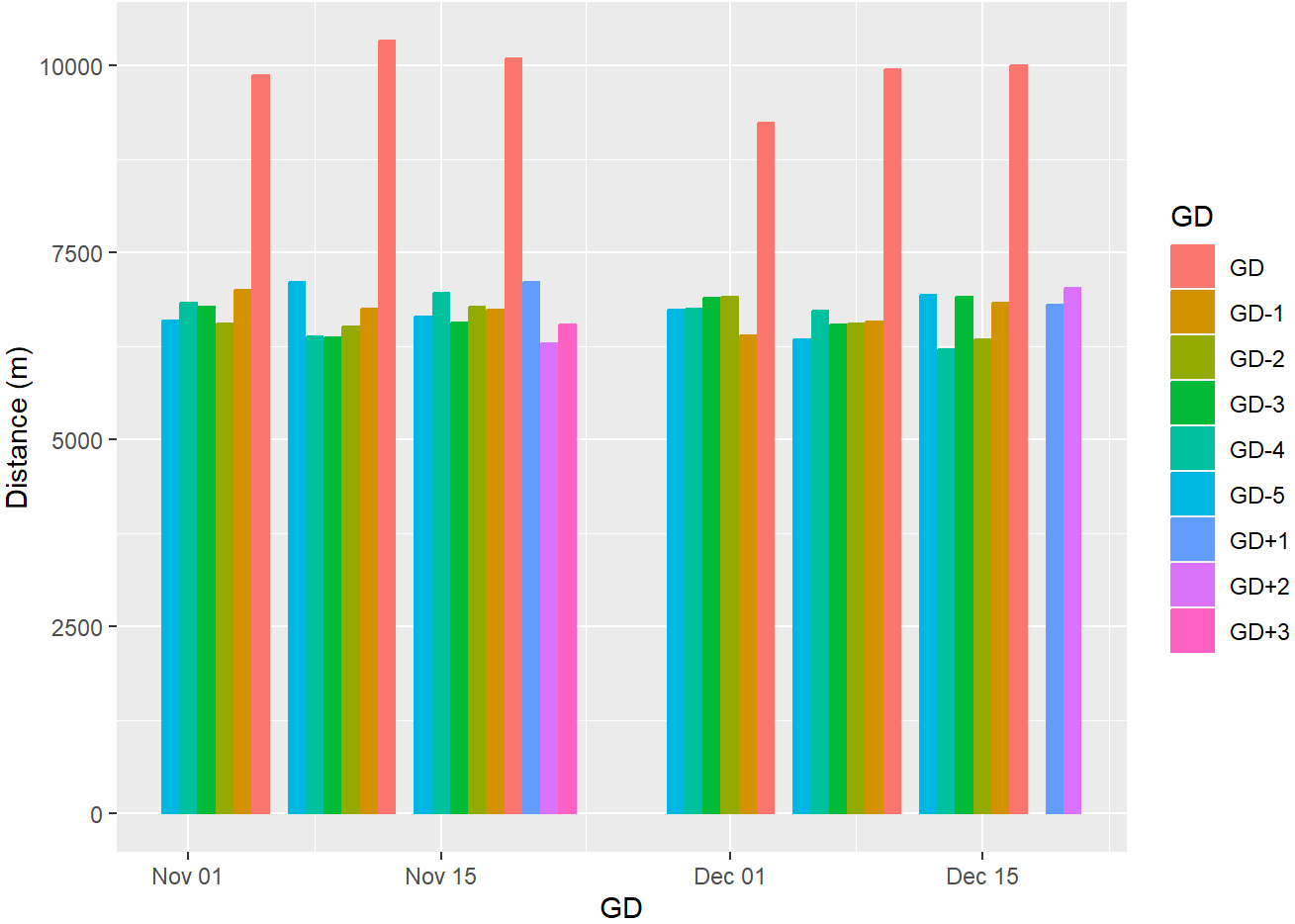


When plotting all points for distance, there does not seem to be a lot of meaningful information to gather.

However, there are a few things to point out. Most of the data falls between 5000 and 8000, with 6 sets of “outliers” with values in the 10000s. Based, on my topic knowledge, these are most likely the distance measures from the games included in this data set. Within the data set, there are 41 unique dates, each representing either a practice or a game. Of those, 6 unique dates are game days, which aligns with my assumption based on the distance plot.

To help establish a better understanding of the load, I am going to take the average across players for each date. This will allow me to look deeper into trends that may exist based on the training day. I am also going to take the maximum values of each metric for every game, which will be used in further analysis.

Preliminary Results



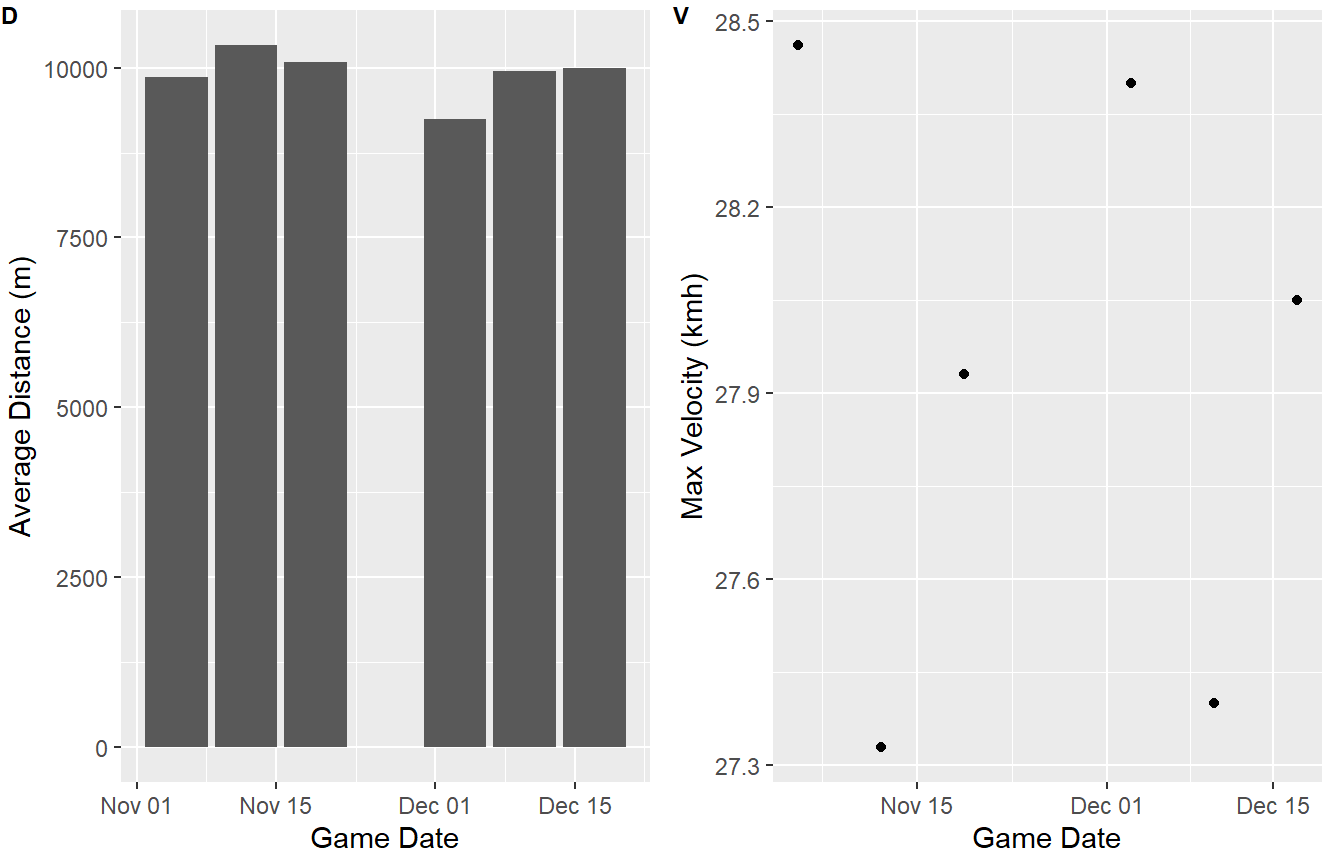
After looking at the mean values for distance covered and max velocity, it is clear there was no real pattern across the training days. For distance, the averages remain fairly constant until GD, when they peak. It is expected that the values for distance are significantly greater during games than in training, so that appears normal. However, in an ideal training structure, there would be a clear delineation between the training days. Distances should look similar to a bell-shaped curve, starting the week with lower distances to ensure proper recovery from the game, peaking midweek when players are fresh and still have time to recover before the game, and taper off before that game to ensure players are not too fatigued. With speed, you want players to hit their peak within 1 to 2 days of a game. This ensures they are sharp, and prepared to hit extreme speeds if needed in a game situation. With this team, players occasionally peak on the GD-1 or GD-2, but it is not consistent throughout each week.

5 Day Build Breakdown

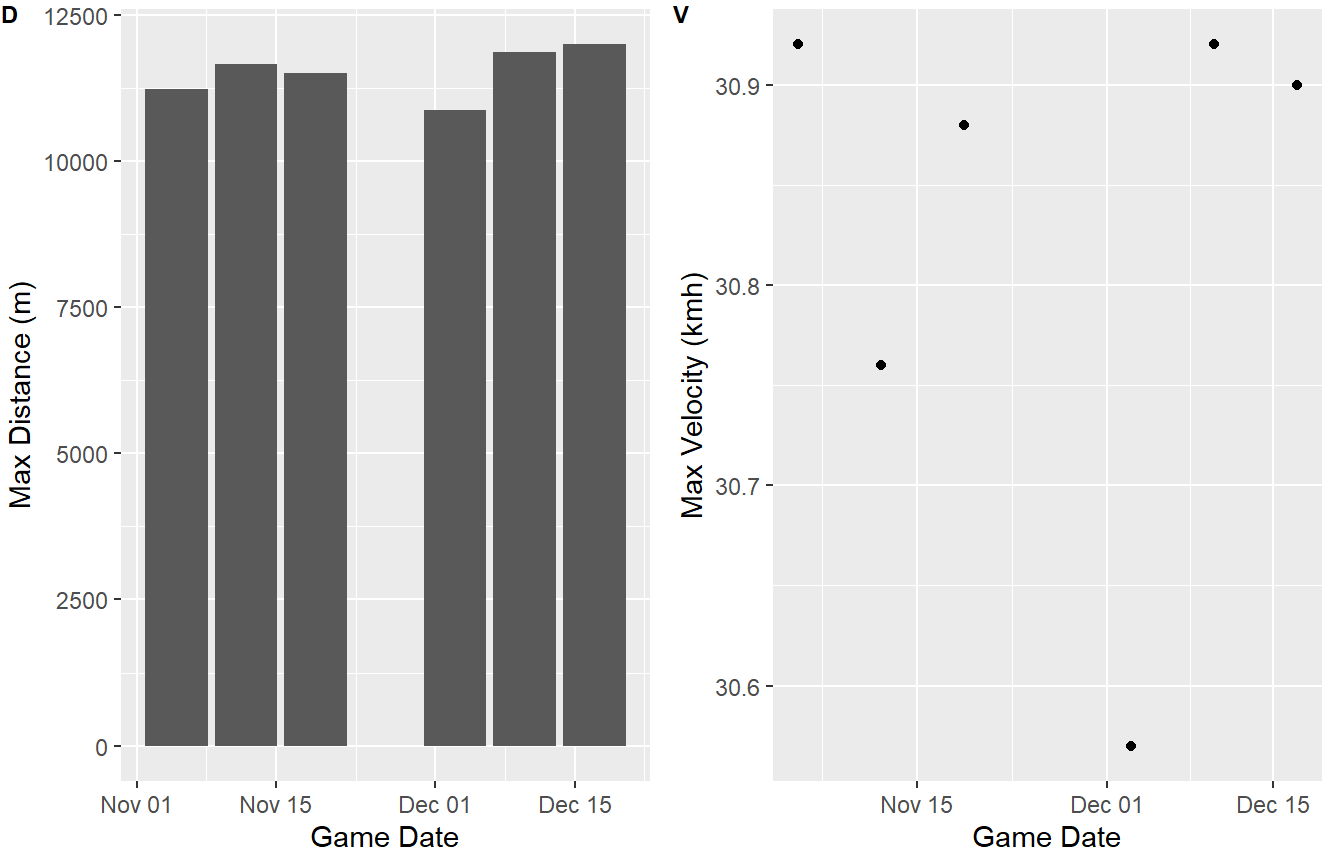
GD	RPE	Distance	Accelerations	Decelerations	Max.Speed
GD-5	5.62	6722	61	54	27.86
GD-4	5.52	6642	61	52	27.96
GD-3	5.56	6674	62	55	27.96
GD-2	5.73	6604	66	55	27.75
GD-1	5.70	6711	60	52	27.66
GD+3	5.37	6539	58	53	28.14
GD+2	5.66	6658	62	50	27.73
GD+1	6.08	6956	60	54	27.60
GD	7.46	9913	56	48	27.93

This table provides a breakdown of the average output for each metric on any given training day or a game. Looking at this table, you can see values for RPE are almost always around 5. This level of consistency may suggest there is not enough variation in the difficulty of training throughout the week, an assumption that is supported by the lack of variation in distance values as well. Another key observation is that the acceleration and deceleration values on any given day are fairly close to each other. The similarity between the two is important for preventing injuries, as an imbalance between the number of accelerations and decelerations performed within a session can indicate some level of neuromuscular fatigue and increased injury risk. Finally, when looking at the distance values for the week, this team completes almost 300% of what they may do in a game throughout the course of the week. It is possible this players are being overexposed and may be at greater risk for overuse injuries.

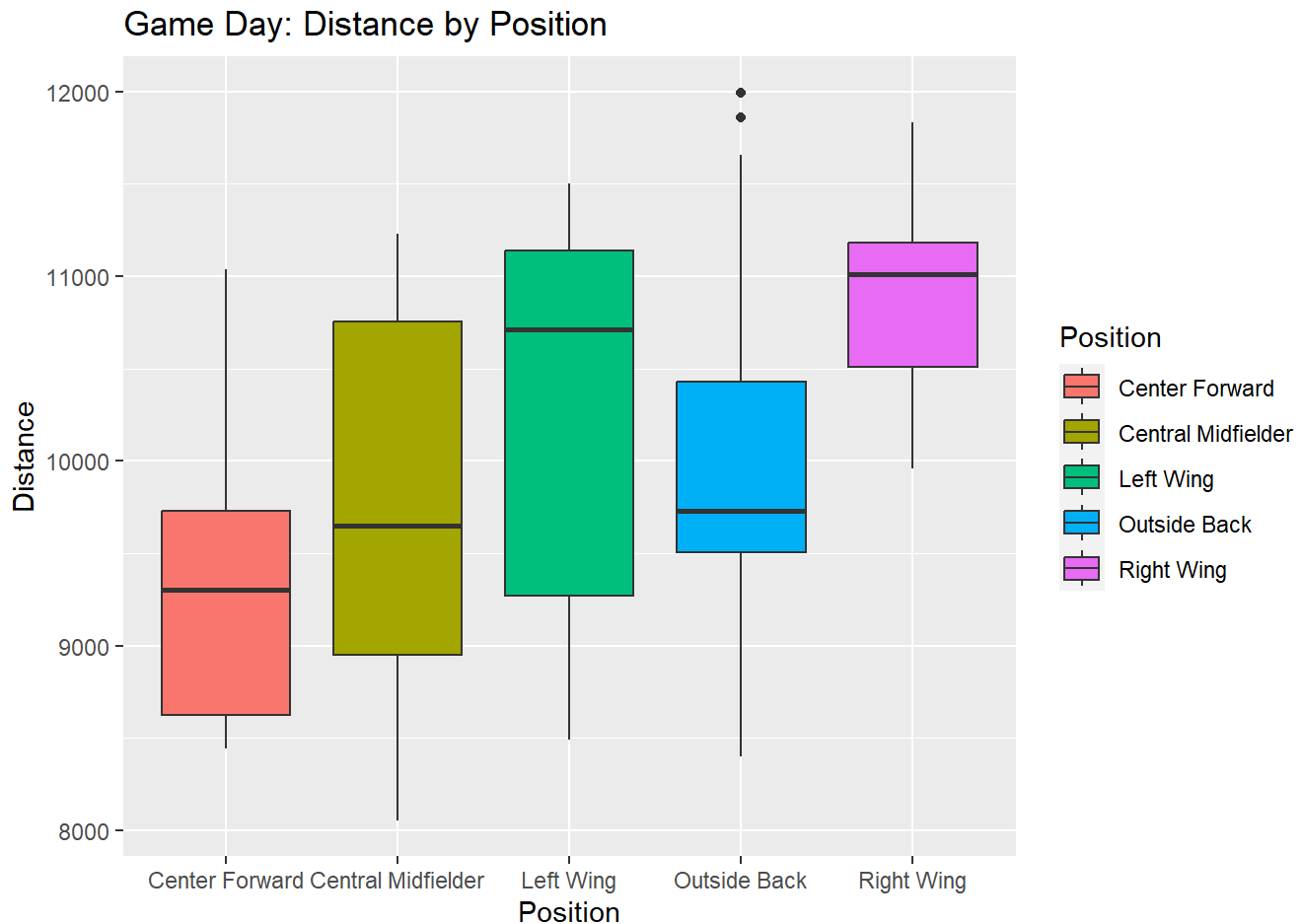
Game Day Average: Distance (m) and Max Velocity (kmh)



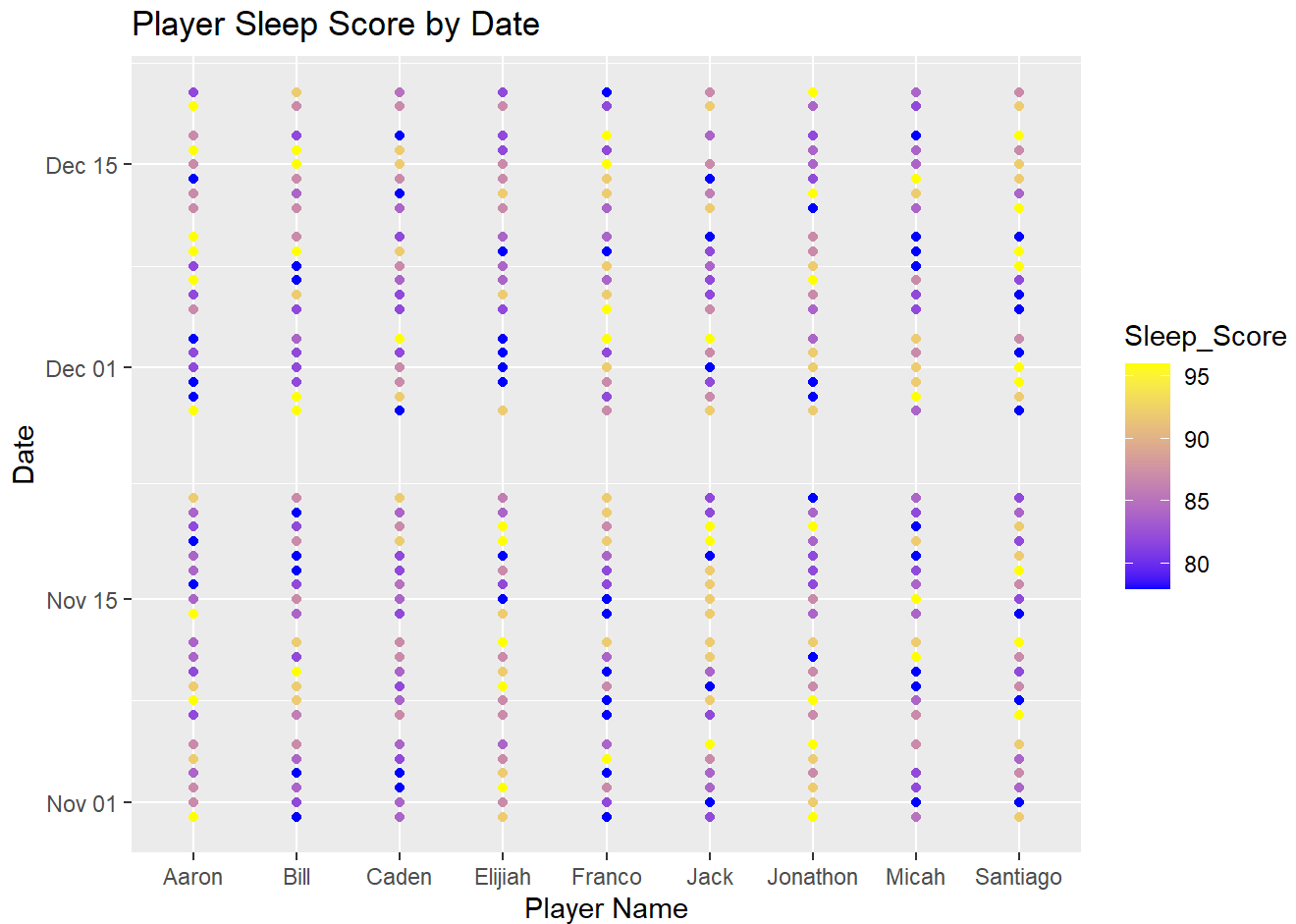
Game Day Maximum: Distance (m) and Max Velocity (kmh)



In addition to distance, another metric to help indicate load and demands of a game is max velocity. Because there does not appear to be a clear pattern between max velo and distance in these charts, it is possible the demands for each game were different. For example, the second game has a high distance and a low max velocity. This could be the team was in control of the game: covering large distances in slow, methodical patterns. The first game has a low distance and a high max velocity. In this case, players may have done more standing before having to make rapid movements to defend or attack the goal.



During a game, the demands for each position are unique, as is displayed in this boxplot. Generally, the right wing appears to cover the most distance. A winger plays on the edge of the field and has to constantly make runs to create chances for crosses and shots. Therefore, it makes sense that they would cover the most distance. The left wing does have values that appear similar to those of the right wing, but it has a much larger range. This may indicate the left wing does not get as much of the play as the right wing and doesn't always have to make as many runs. The center forward seems to generally cover the least distance, which is logical as they often stay towards the top of the field as a target and don't help as much on defense.



Sleep is a good indicator of recovery, and is key to being at peak performance. Here we can see the players' sleep scores, with higher being better sleep. There are a few players that may need to focus on creating better sleep habits to improve overall health and performance. Players like Bill, Caden, and Micah have more purple and blue dots than yellow and seem to have lower scores more often than others. This was during November and December, so it is possible the holidays may have played a role in the poor sleep, but they should still look for places within their routines to create an environment for improved sleep.

Conclusion

After reviewing the data from part of this season, my main conclusion is there is much room for improvement. When looking at the build-up to a game, there is no distinction between days. To optimize preparation, recovery, and performance there should be a gradual introduction to load at the start of the week, the maximum load at the middle of the week, and a taper off of load right before the game. This will ensure players are adequately prepared, but not overexposed. The change in difficulty can be complemented by focusing on position-specific loads throughout the training week. These data clearly indicate that each position has its own demands that should be introduced within the controlled training environment. Along with the physical aspects, sleep should also be improved upon. While the data shows there are many times the athletes are getting quality sleep, the goal should always be to eliminate as many nights of poor sleep as possible.

One positive shown by the data is the similarity between acceleration and deceleration counts across the week. Soccer is a sport that requires a lot of speeding up and slowing down rapidly. If your body stops doing one of these, there would be some negative consequences. An imbalance could indicate the start of an injury or

increasing fatigue, two things that you want to avoid. Therefore, the athletes should continue being exposed to both accel and decel in similar number.

Overall, the data displayed no pattern between training day and physical metrics. Position appears to be a strong determinant of distance covered in games. Finally, while not entirely concerning, there are some low sleep scores that could be improved to aid in better recovery.