# Is the Previous Week's Load Related to RSI?

First, we need to define what "load" means and what it means in the context of our datasets.

- From Blake:
    - Use Acceleration Load as the primary load variable
    - Could also be interesting to look at Jump.Load and Session.Mechanical.Load
        - Jump.Load could have a strong relationship with RSI since both involve jump mechanics
            - Can also use Jump.Load.Per.Mass, which standardizes Jump Load based on player mass (not important for basketball)

RSI testing dates during 2024-25 season:

```
"2025-03-07" "2025-03-04"
"2025-02-14" "2025-02-07" "2025-02-04"
"2025-01-27" "2025-01-20" "2025-01-14" "2025-01-11" "2025-01-03"
"2024-12-29" "2024-12-20" "2024-12-12" "2024-12-06" "2024-12-01"
"2024-11-21" "2024-11-16" "2024-11-12" "2024-11-07" "2024-11-03"
"2024-10-18" "2024-10-11"
```

Acceleration Load dates:

```
 [1] "2025-04-18" "2025-04-17" "2025-04-14" "2025-04-12" "2025-04-11"
"2025-04-10"
  [7] "2025-04-08" "2025-04-05" "2025-04-04" "2025-04-03" "2025-04-01"
"2025-03-31"
 [13] "2025-03-29" "2025-03-28" "2025-03-21" "2025-03-20" "2025-03-19"
"2025-03-18"
 [19] "2025-03-14" "2025-03-13" "2025-03-12" "2025-03-11" "2025-03-10"
"2025-03-08"
 [25] "2025-03-07" "2025-03-06" "2025-03-05" "2025-03-04" "2025-03-03"
"2025-03-02"
 [31] "2025-03-01" "2025-02-28" "2025-02-27" "2025-02-26" "2025-02-25"
"2025-02-24"
 [37] "2025-02-23" "2025-02-22" "2025-02-21" "2025-02-20" "2025-02-18"
"2025-02-17"
 [43] "2025-02-15" "2025-02-14" "2025-02-13" "2025-02-11" "2025-02-10"
"2025-02-08"
 [49] "2025-02-07" "2025-02-05" "2025-02-04" "2025-02-03" "2025-02-02"
"2025-02-01"
 [55] "2025-01-31" "2025-01-30" "2025-01-28" "2025-01-27" "2025-01-26"
"2025-01-25"
 [61] "2025-01-23" "2025-01-21" "2025-01-20" "2025-01-19" "2025-01-18"
"2025-01-17"
 [67] "2025-01-16" "2025-01-15" "2025-01-14" "2025-01-12" "2025-01-11"
"2025-01-10"
```

```
 [73] "2025-01-08" "2025-01-07" "2025-01-06" "2025-01-04" "2025-01-03"
"2025-01-02"
 [79] "2025-01-01" "2024-12-30" "2024-12-29" "2024-12-27" "2024-12-26"
"2024-12-21"
 [85] "2024-12-20" "2024-12-19" "2024-12-18" "2024-12-16" "2024-12-15"
"2024-12-13"
 [91] "2024-12-12" "2024-12-11" "2024-12-10" "2024-12-07" "2024-12-06"
"2024-12-05"
 [97] "2024-12-04" "2024-12-02" "2024-11-30" "2024-11-27" "2024-11-26"
"2024-11-25"
[103] "2024-11-24" "2024-11-23" "2024-11-21" "2024-11-20" "2024-11-19"
"2024-11-17"
[109] "2024-11-16" "2024-11-15" "2024-11-13" "2024-11-12" "2024-11-11"
"2024-11-10"
[115] "2024-11-08" "2024-11-07" "2024-11-06" "2024-11-04" "2024-11-03"
"2024-11-02"
[121] "2024-11-01" "2024-10-30" "2024-10-29" "2024-10-27" "2024-10-25"
"2024-10-22"
[127] "2024-10-21" "2024-10-19" "2024-10-18" "2024-10-17" "2024-10-15"
"2024-10-14"
[133] "2024-10-12" "2024-10-11" "2024-10-10"
```

- Created data frame RSI
    - Filtered the VALD ForceDecks dataset to only include data from this season (starting 10/10/2024).
    - Selected columns: anon_id, Date, RSI..m.per.s., and Trial.
    - Then I created a new column Daily.Avg.RSI..m.per.s that takes the mean of the RSI scores recorded for an athlete for each date. They had multiple trials of RSI tests for a day so I averaged them.
    - This data frame makes it easier to examine RSI score by date.

- Using the RSI data frame, I created a scatter plot of Daily.Avg.RSI..m.per.s by date and colored the points by anon_id. I also added a regression line for each athlete.
    - Overall trend seems like the RSI scores for each athlete increase over the course of the season.

- Created load data frame
    - Filtered the Kinexon Session dataset to only include data from 10/10/2024 - present.
    - Selected anon_id, Date. Daily.Total.Accel.Load.Accum, and Jump.Load. Then filtered out the NAs.
    - This data frame makes it easier to explore load data by date

- Using the load data frame, I created a scatter plot of Daily.Total.Accel.Load.Accum by date and colored the points by anon_id. I also added a regression line for each athlete as well.
    - Overall trend seems like acceleration load goes down throughout the course of the season.

- Using the load data frame I calculated new columns of rolling 7 day acceleration and jumpload.
    - One of the columns is the average acceleration load in the last 7 days
    - Another is average jump load in last 7 days
    - Another is total jump load in the last 7 days
    - The other column is the total acceleration load in the last 7 days.
    - Note: Last 7 calendar days, not last 7 sessions
    - This is super helpful for answering the question of prior week's load affecting RSI

- Then I ran some simple plots and put color=anon_id so you can see each athlete
    - One of them is RSI scores over course of the season
        - Trends generally increase as season goes on
    - One of them is daily acceleration total throughout season
        - Trends generally decrease
    - Another is daily jump load
        - Stayed about the same

- Then I merged the load and RSI dataframes
- Then I did some correlations of avg and total of both jump and acceleration load to RSI. Did pearson and spearman
    - Found that Jump Load had higher correlations to RSI
- Then I made some per athlete correlations graphs where you can see how much an athlete's RSI scores are  affected by load (acceleration or jump)
- Then did some more simple scatter plots of RSI vs. Avg/Total Jump/Acceleration Load

- Compared a mixed effects model and a GAM model using CV, also created a fixed effects model
    - The mixed effects and GAM models showed very similar predictive accuracy, with mean RMSEs of about 0.412 and 0.406, respectively, indicating that both approaches capture the relationship between previous week's load and RSI well.
    - The fixed effects model revealed that higher average acceleration load in the last week is significantly associated with a slight decrease in RSI, while higher average jump load is associated with a slight increase. Together with athlete-specific differences, these factors explain about 70% of the variation in RSI.