SIR Report:

This week I reviewed some a past thesis paper related to my project, and the code related to the thesis. This helps me understand my goal of understanding the past projects related to mine, and what I may need to do in order to expand on that project. Getting the code to run and understanding the math and basic ideas of the paper was a challenge, but I was able to discuss with my mentor and get the code to run. I spent 1 hour with my mentor to discuss of my problems and to understand the main idea of the past papers. I also got a guide on the things I should study. My mentor stated how creating a project timeline and plan would be good in order to stay on track. I plan to review the code, the thesis, and some literature related to the project to just get a better understanding of what my project will be about.

Look at the paper <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2942750/>

***F ratio*** meaning: ratio between group variance and within group variance.

People were asked pains and predictors weights and variance were nonted and calculated through multivariate regressions.

(*r* = −0.24, *p* < 0.01). P is significance, and r is Pearson correlation coefficient.

Beta\* meaning.

α = 0.72 (Cronbach alpha, is a reliability coefficient)

Use when:

1. The data is [normally distributed](https://en.wikipedia.org/wiki/Normal_distribution) and [linear](https://en.wikipedia.org/wiki/Linearity)[[ii]](https://en.wikipedia.org/wiki/Cronbach%27s_alpha#cite_note-Footnote2-20);
2. The compared tests or measures are essentially tau-equivalent;
3. Errors in the measurements are [independent](https://en.wikipedia.org/wiki/Independence_(statistics)).

-Looked at the thesis

47.40% pPMSEr calculated. Simulated gives 40.42% for the same sample size, indicating

“This discrepancy may be attributed to the potential introduction of random error

due to the inclusion of insignificant predictors, which can diminish the amount

of variance explained and increase the e↵ective sample size n⇤ needed.”

Increasaing sample size is needed to offset extra error from insignificant predictors

N=600 results in a closer 46.2% to calclaulte 49.33

Choosing predictors is important

Correlation increases with samplesize but decreases with some useless predictors.

Future studies: “general linear models, additional outcome measures,

and a more comprehensive set of predictors, potentially utilizing machine learning

or other statistical techniques to identify complex relationships among predictors”

Correlation matrix was used instead of actual data