Morgan Fissel

# Final Report:

## How Sleep Rates and Marijuana Usage Vary By Age

### Executive Summary

Personal health means many things to different people. To me, I am curious what I may be able to learn about health trends relating to age, sleep, and marijuana. My primary goal in this project was to practice using some of the methods we learned about in class. The NHANES dataset is well established and relatively legible. I investigated trends relating to age using hierarchical modeling. I also attempted regression, a binary/binomial model, and a poisson/gamma model. My best results were with my hierarchical modeling suggesting some potential for that type of structure in the data. It seems that health indicators do indeed vary over time within different age groups in the U.S.

### Description of Data

National Health and Nutrition Examination Survey. The NHANES program is designed to assess the health and nutritional status of adults and children in the United States. There are 76 variables but the ones of interest are: Age, AgeFirstMarij, AgeRegMarij, EducationNumeric, SleepTroubleNumeric, and SleepHrsNight.

Morgan Fissel

## Graphs and Summaries

Figure 1:


Figure 2:

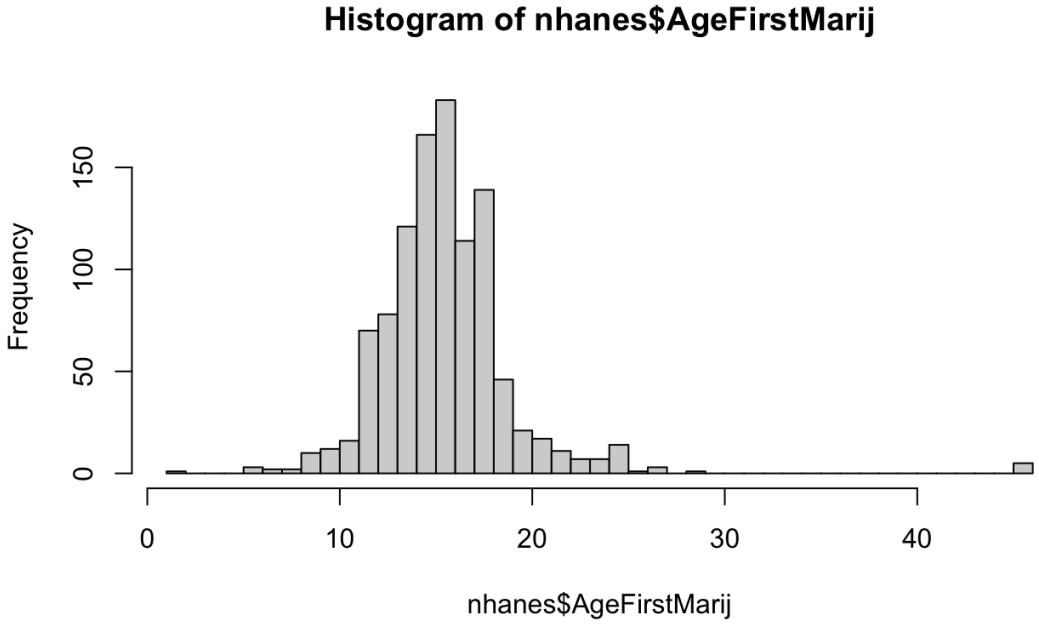


Figure 3:

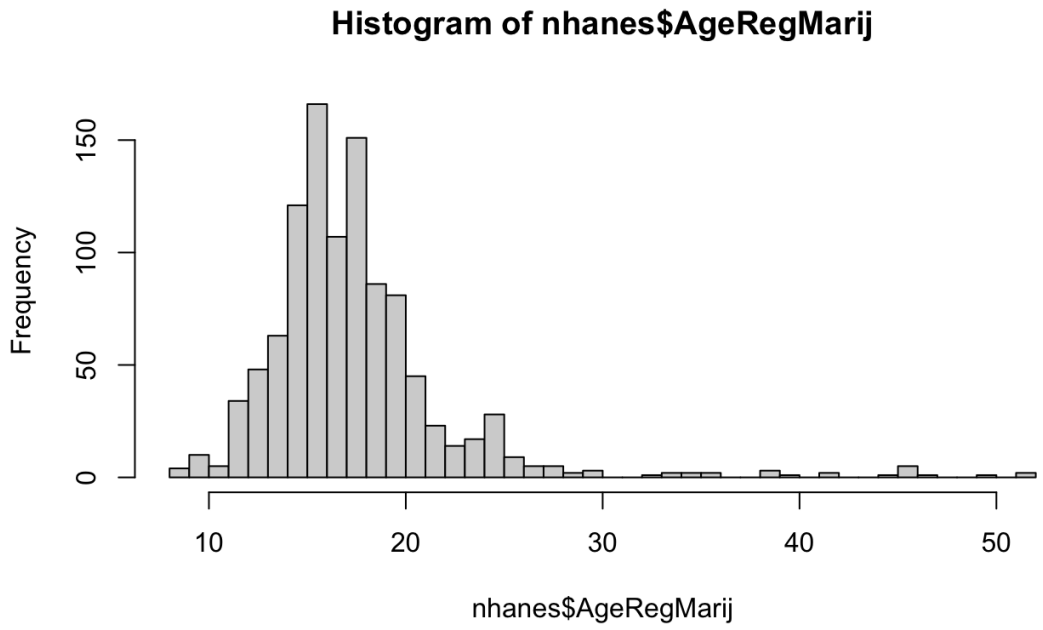
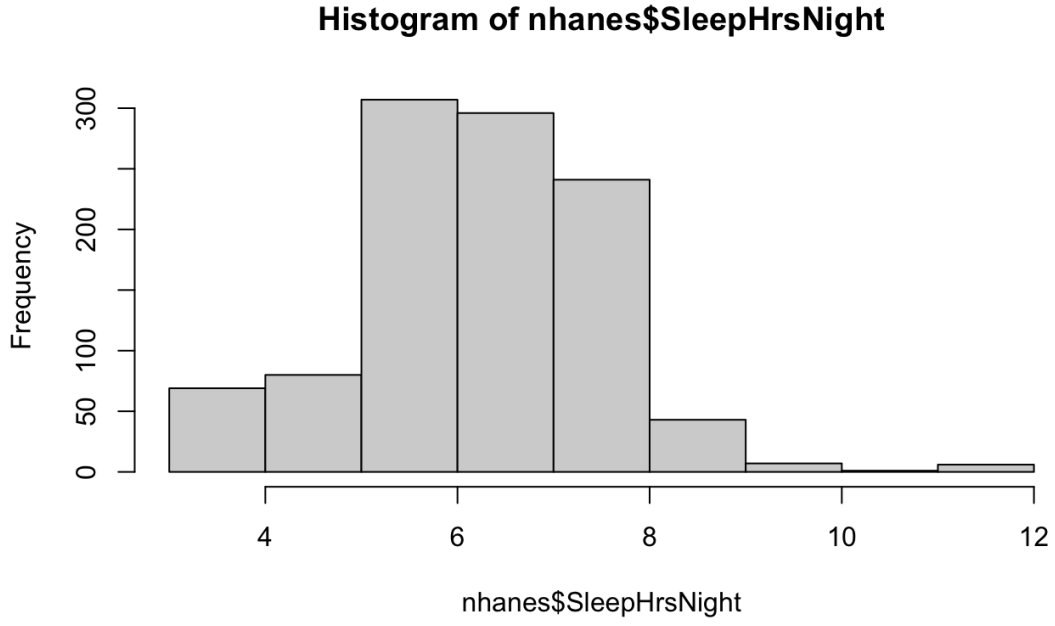
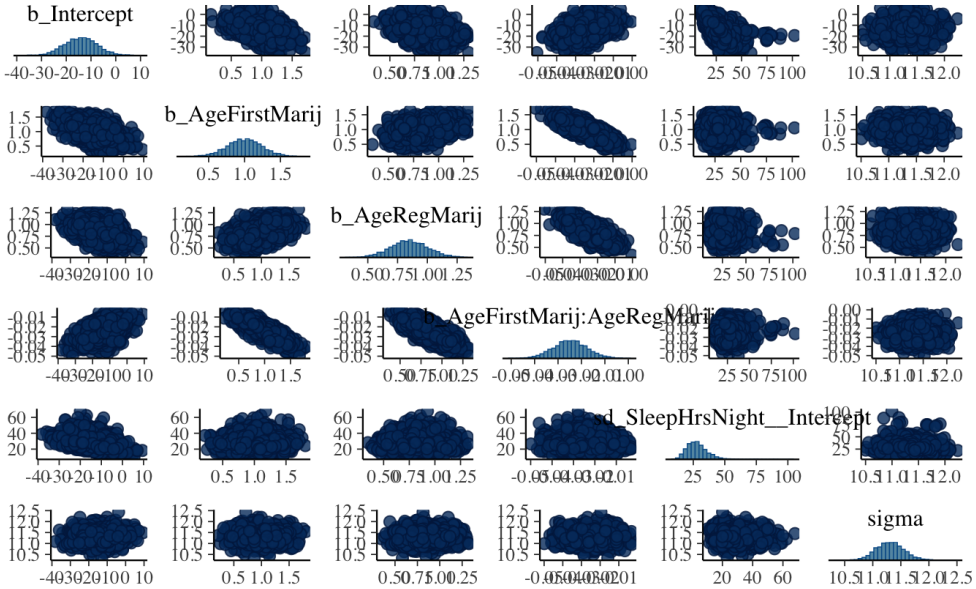



Figure 4:



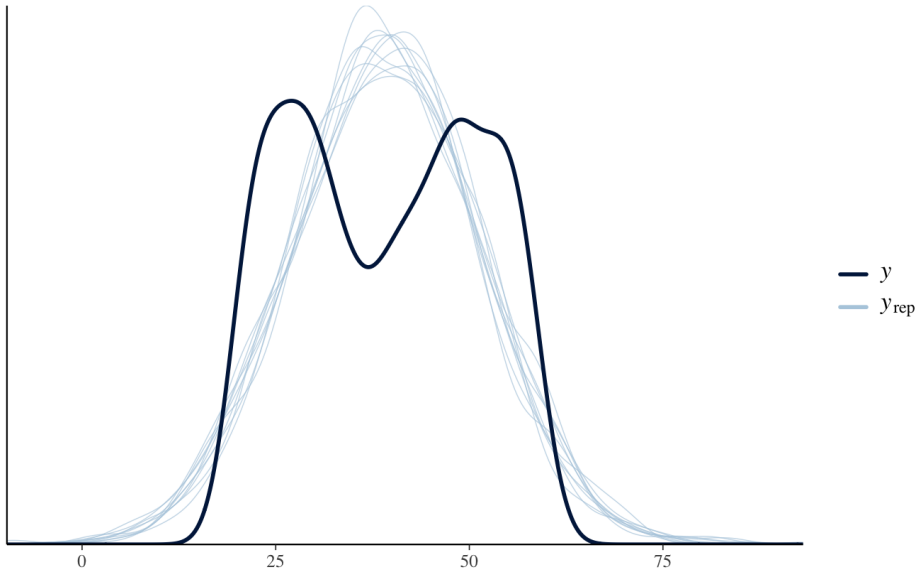
## Figure 5:

This is the pairs plot for the hierarchical interaction model: Age ~ (AgeFirstMarij) * (AgeRegMarij) + (1|SleepHrsNight)



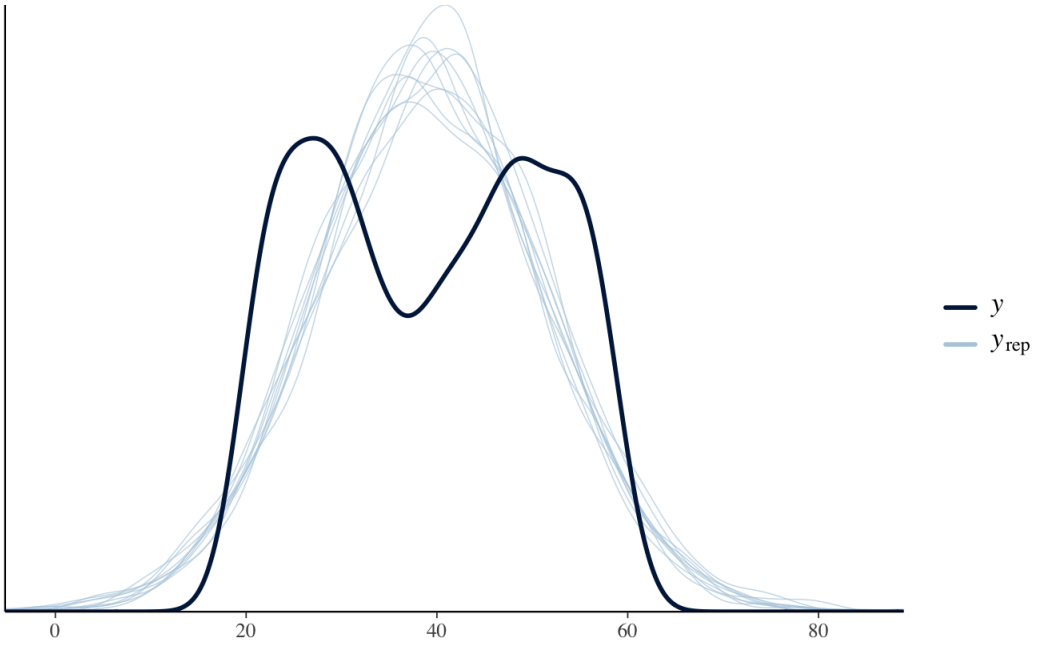

## Figure 6:

This is the pp_check for the hierarchical interaction model: Age ~ (AgeFirstMarij) * (AgeRegMarij) + (1|SleepHrsNight)



## Figure 7:

This is the pp_check for the hierarchical model: Age ~ AgeFirstMarij + AgeRegMarij + (1|SleepHrsNight)




It is quite apparent that for both **Figure 6** and **Figure 7** our y_rep does not do a perfect job explaining the 2 humps. However, even though it's not the same amount of humps, the model does a pretty good job smoothing out the overall curve of y. y_rep is valuable because it is simple and still manages to capture the data decently. Both models are very close to being the same, but differ slightly. The better model is the one with an interaction term (Figure 7).

## Subject Matter Implications

Contribute information about the relationships between age, marijuana use, and sleep duration based on the NHANES dataset which could potentially be used to predict on a much larger scale. I won't attempt to say this has some widely impactful usage, but the analysis just interested me. I don't often get to work with those variables or things that are related to those often. Just good practice for understanding things we discussed in class as well.

## Further Questions

The next steps of this investigation would be to try out some new models. I have a general outline for a regression model, a binary model, and a gamma model. I wasn't able to get them running perfectly but they may offer more insights into the data and trends given more time. In all honesty I wasn't able to get them working primarily because I don't understand the brms package incredibly well and have struggled with even being able to use the package half the time. See the code file for the skeleton that I created for these models (lines 93-167).