Anxiety Metrics and Their Impact on Hippocampal Volume and Brain Connectivity

Daniel Meskill

November 30, 2023

Abstract

This report presents an investigation into the predictive relationship between anxiety metrics and measures of hippocampal volume and brain connectivity. Initial exploratory data analysis involved handling missing values through imputation, identifying and removing outliers, and visualizing relationships between variables using a correlation matrix. Subsequent predictive modeling using both multivariate linear and polynomial regression, however, did not yield significant results. The findings suggest that the relationship between anxiety metrics and hippocampal volume and brain connectivity may be complex and not easily captured by these forms of regression. Further research using more sophisticated modeling techniques may be warranted.

Introduction

Anxiety disorders are increasingly common and have been linked to changes in brain structures and functions. This study focuses on how anxiety levels correlate with hippocampal volume and brain connectivity. We investigate whether anxiety metrics can predict these measures. Understanding the relationship between anxiety and brain measures is important for advancing our knowledge of the neurological aspects of anxiety disorders.

Data Description

The data was provided to me in a post-study CSV file for me to analyze. There were 131 subjects in the study, each with up to 18 measures. These measures are split into 3 components: 2 of which are for response variables, and 1 for explanatory variables. The first response component is for hippocampal volume (mm^3), with 1 measure for the left hemisphere and 1 for the right hemisphere. The second response component is made up of 12 measures related to functional connectivity. Functional connectivity is between two regions, for this study those are combinations of left and right hippocampus, amygdala, dorso-lateral prefrontal cortex, and ventro-medial prefrontal cortex. Higher values suggest stronger connectivity between the regions under stress. Lastly, the explanatory component includes 4 measures related to anxiety. 3 are from the State-Trait Anxiety Inventory questionnaire, which consists of 40 self-report items on a 4-point Likert scale. The final measure is of MRI cortisol peak during an MRI session.

Data Exploration and Cleaning

In this section I explore the data and do any appropriate cleaning/processing.

The first thing I noticed was that there were a decent amount of missing values. To be exact, 242 missing values over 56 subjects, of which 42 subjects were missing their peak cortisol level. Instead of discarding over a third of the rows from our limited data, I chose to perform mean imputation to fill in the missing values.

Then I chose to check the histograms for normality. All of the explanatory variables looked good, for the most part. However, the cortisol peak histogram showed at least one outlier (see figure 1). For this I chose to remove the subject since it shouldn't impact the analysis much.

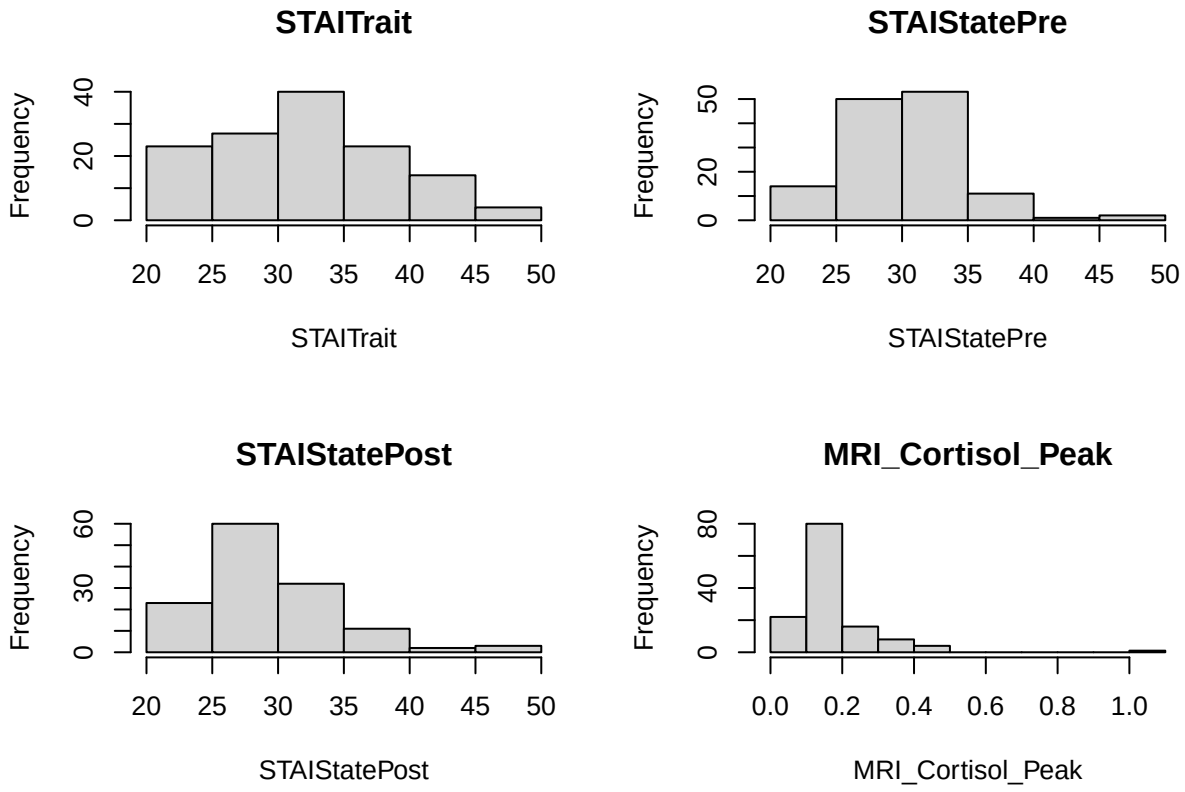

Figure 1: Histograms of the 4 Predictor Variables

Next I explored the data through a correlation heat map (see figure 2). This shows whether two variables are positively correlated (darker blue) or negatively correlated (darker red). From this figure we can conclude that although there are no signs of high collinearity, there also isn't much correlation between the explanatory and response variables. This foreshadows our results later.

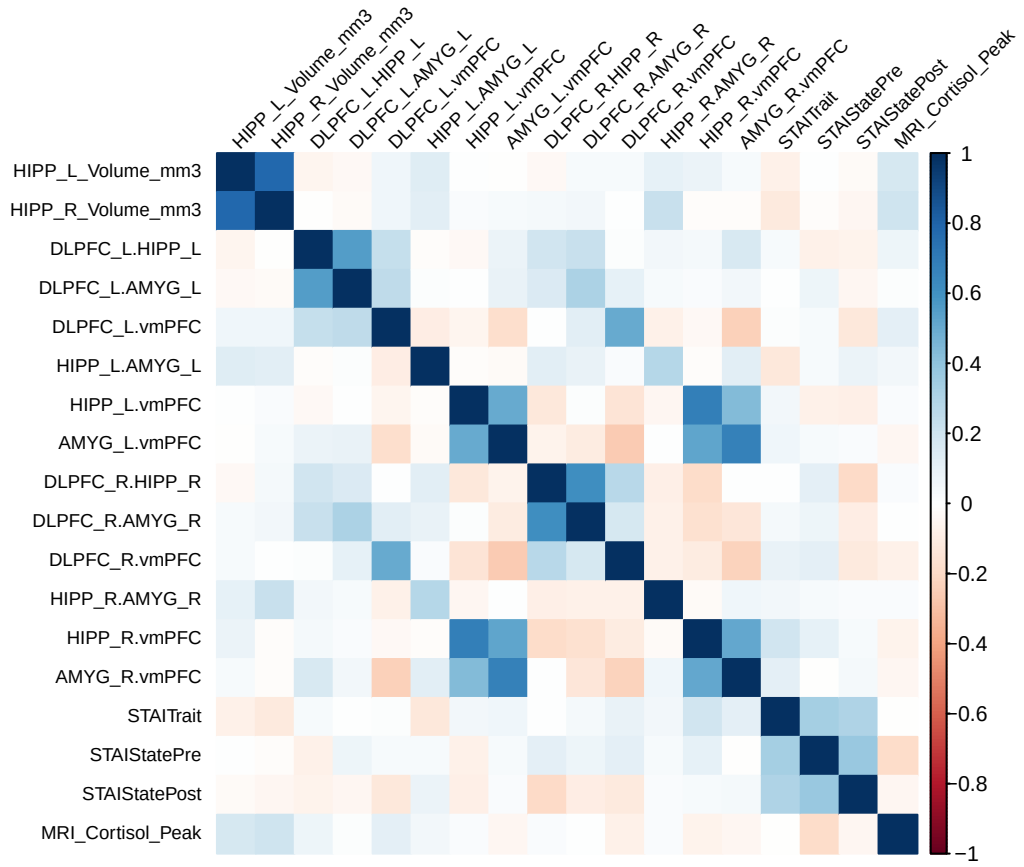

Figure 2: Heat Map of Every Variable

Regression

Then I moved on to modeling our data. I first applied a simple multivariate linear regression model for each response variable and found the following R-squared values (see Table 1).

Response	R_Squared
HIPP_L_Volume_mm3	0.0406
HIPP_R_Volume_mm3	0.0581
DLPFC_L.HIPP_L	0.0149
DLPFC_L.AMYG_L	0.0129
DLPFC_L.vmPFC	0.0362
HIPP_L.AMYG_L	0.0380
HIPP_L.vmPFC	0.0180
AMYG_L.vmPFC	0.0059
DLPFC_R.HIPP_R	0.0809
DLPFC_R.AMYG_R	0.0256
DLPFC_R.vmPFC	0.0580
HIPP_R.AMYG_R	0.0040
HIPP_R.vmPFC	0.0437
AMYG_R.vmPFC	0.0178

Table 1: R-squared Values for Every Model in Linear Regression

R-squared is a measure of how much variability in the response variable can be explained by the predictor variables. Of the models the highest R-squared was about .08 while the lowest was .004. Quite honestly all of these models performed very poorly, suggesting there is not a linear model that fits the data well.

I then went on to plot the residuals (see Figure 3) and conducted a Shapiro-Wilk test for normality. From this we can conclude they are approximately normally distributed.

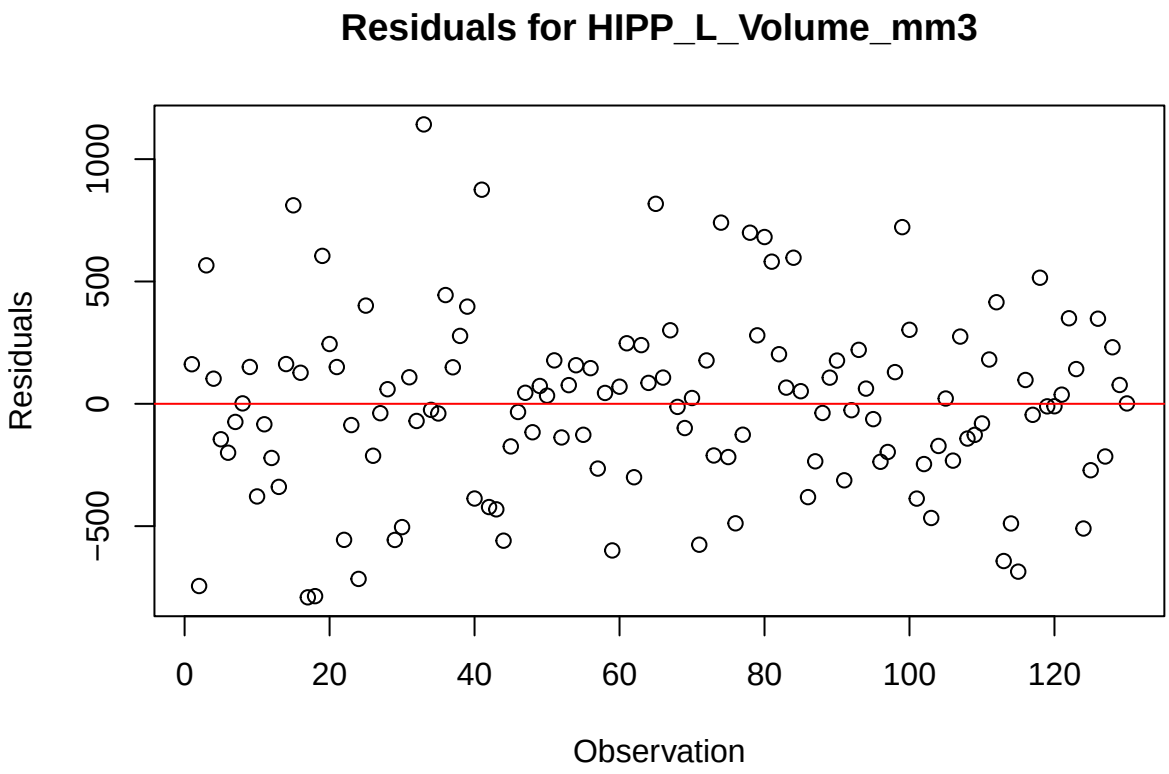

Figure 3: Residuals for the Left Hippocampal Volume Model

```
##
##  Shapiro-Wilk normality test
##
## data:  results[[1]]$residuals
## W = 0.98334, p-value = 0.1123
```

Since the linear model did not perform well, I applied a polynomial regression model (of degree 2) to see if it would yield better results The following table summarizes those R-squared results (see Table 2).

Response	R_Squared
HIPP_L_Volume_mm3	0.0945
HIPP_R_Volume_mm3	0.0944
DLPFC_L.HIPP_L	0.0953
DLPFC_L.AMYG_L	0.0641
DLPFC_L.vmPFC	0.1466

Response	R_Squared
HIPP_L.AMYG_L	0.1036
HIPP_L.vmPFC	0.1343
AMYG_L.vmPFC	0.1101
DLPFC_R.HIPP_R	0.1653
DLPFC_R.AMYG_R	0.0892
DLPFC_R.vmPFC	0.2284
HIPP_R.AMYG_R	0.1099
HIPP_R.vmPFC	0.1739
AMYG_R.vmPFC	0.1218

Table 2: R-squared Values for Every Model in Polynomial Regression (degree 2)

This table shows significant improvement over the linear models, but still nothing too impressive for predicting each response.

Conclusion

This study investigated the relationship between anxiety metrics and brain-related measures, specifically hippocampal volume and brain connectivity. Despite the data exploration and cleaning processes, and the application of both multivariate linear and polynomial regression analyses, the results did not reveal a significant predictive relationship. The low R-squared values in linear regression models, along with only marginal improvements in polynomial regression, indicate that the relationship between anxiety metrics and the chosen brain measures might be more intricate than initially hypothesized.

The complexity of the brain's response to anxiety, coupled with the limitations of the modeling techniques used, suggests that this field warrants further investigation with more sophisticated models. Future research might benefit from integrating more comprehensive data, considering non-linear models, or employing machine learning techniques that can capture more nuanced relationships within such complex biological data. Additionally, exploring other potential mediating or moderating variables could provide deeper insights into the intricate dynamics between anxiety and brain structure and function.