CS 6140 Spring 2025 Project Presentation

INSIDE OUT ML: CLASSIFY FMRI BY EMOTION

Predicting Emotion States and Priming Conditions

NAME OF PROJECT:
NEURO-CV-EMOTION

PRESENTED BY:
Anissa Vaughn

PRESENTED TO:
Dr. Uzair Ahmad

æ anissa vaughn

About This Dataset

- **Source**: Northeastern University's Affective and Brain Sciences Lab
- **Participants**: 30 individuals underwent fMRI scans ()
  - Each underwent 9 trials: one for each emotion × priming combination
- **Task**: Participants were shown emotionally evocative images after hearing a word.
  - **Emotions:** anger, fear, disgust
  - **Priming Conditions**:
    - Incongruent: Emotion word does not match the image
      - fearful word → disgusting image
    - Congruent: Emotion word matches the image
      - angry word → angry image
    - Neutral: Emotionally neutral word before the image

- **Dataset**:
  - Preprocessed 3D beta maps from fMRI
  - Each map reflects voxelwise brain activation levels
  - Map Dimensions: 91 × 109 × 91 voxels per map
  - Sample Size: 270 total images
  - 30 subjects × 3 emotions × 3 priming conditions

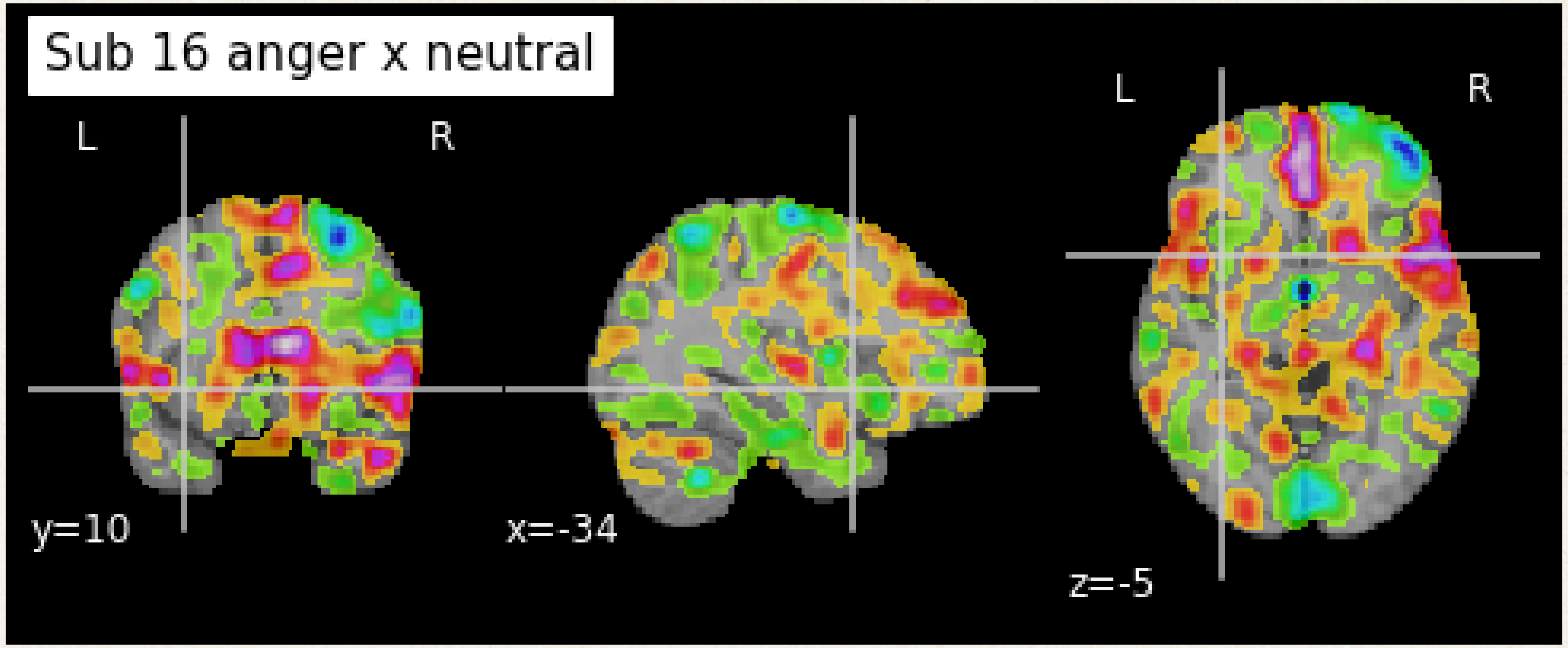


Figure 1: fmri image at coordinates (-34, 10, -5) represented as a ROI mask for Subject 16. S hown angry image after neutral word.

Research Objectives

Overall: Investigate how statistical and domain-specific feature extraction (parcellation and projection) techniques impact multi-class fMRI classification of emotional states

Goal # 1

Identify optimal feature extraction technique for classification

Goal # 2

Identify relationships between features, classes and regions of the brain

Goal # 3

Identify how to manage autocorrelation and bias-variance

Data Preprocessing

How was the raw .nii.gz images prepared for analysis?

Load data: Nibabel

Store NIfTI files into np.arrays of voxel data

Assess Feature Extraction Methods

Use cost functions, run time and ROI count to select best methods

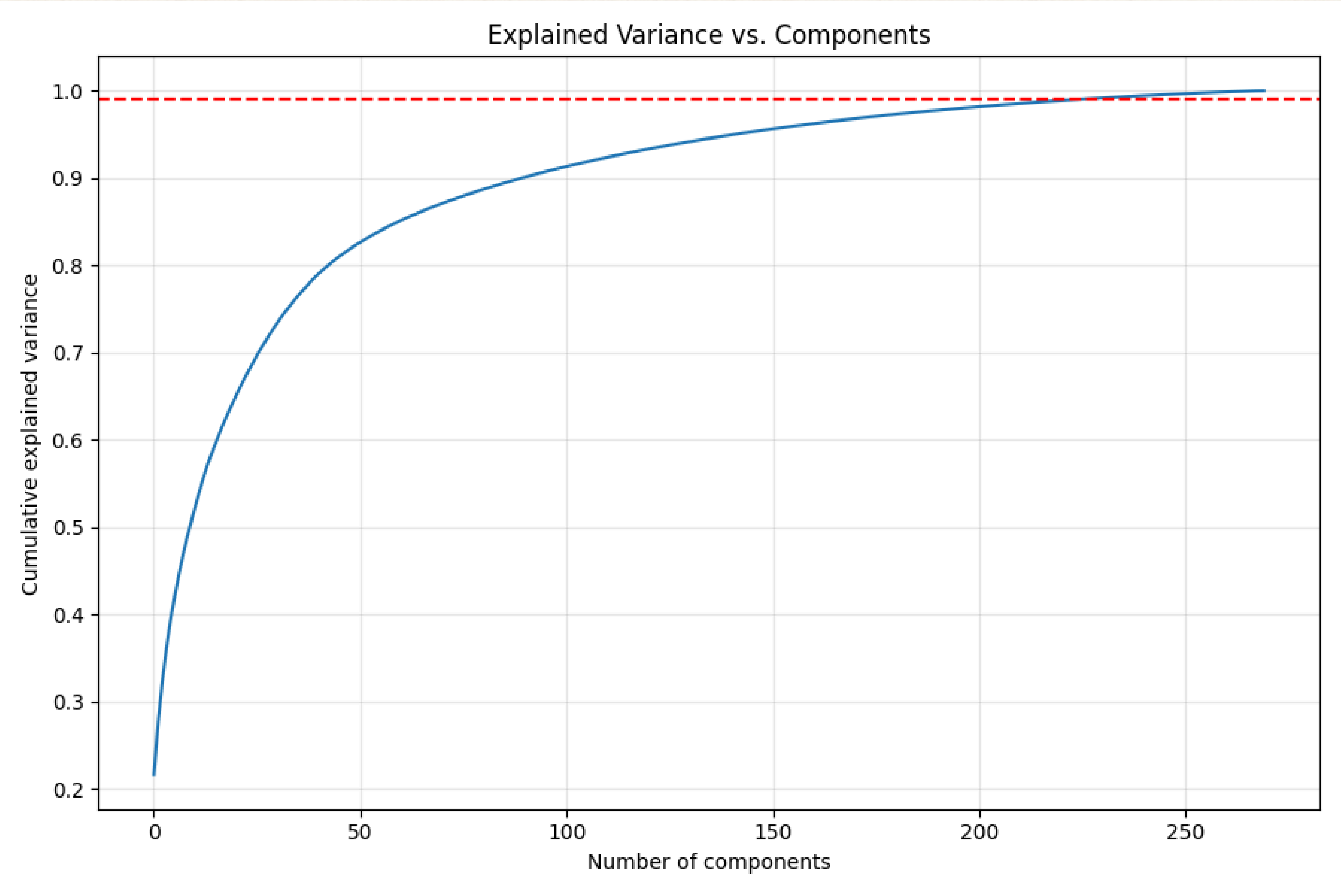
Parcellation Maps: Nilearn

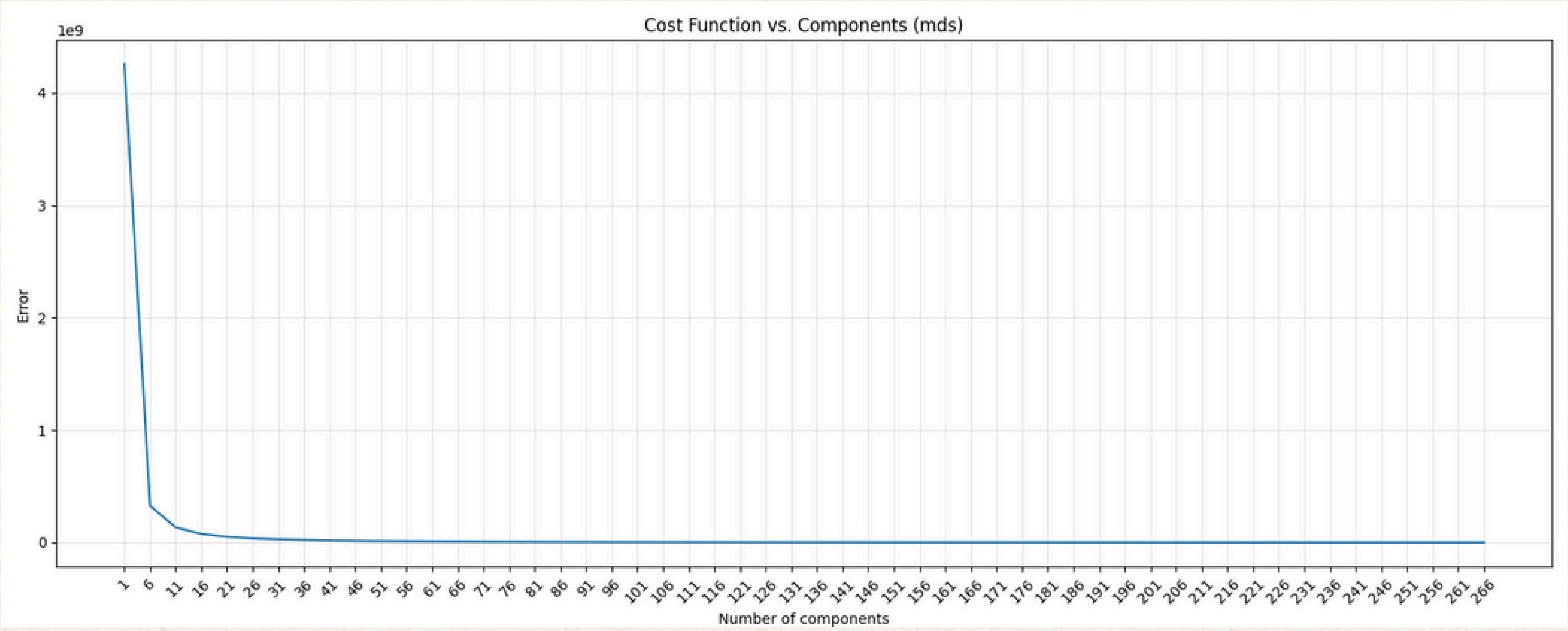
Reduce data to one value per brain region of interest (ROI)

Dimension Reduction: Sci-kit Learn

Reduce scaled data to best components for explained variance

Preprocessing Step: Cost Functions





Investigation

Homogeneity of Variance

Identify features that variance changes across target classes

Multicollinearity

Identify features with high VIF values and correlation coefficients to other input feautres

Autocorrelation

Identify input features with high correlation to subject ID due to relatedness

Logistic Regression

Conduct Multi-class classification on emotion and priming classes. leverging Logistic Regression

Model Training

01 **Problem:** Dependent samples (same subject)

Solution: Hierarchical K-Fold Cross Validation: prevent data leaks of same subject in test and train splits

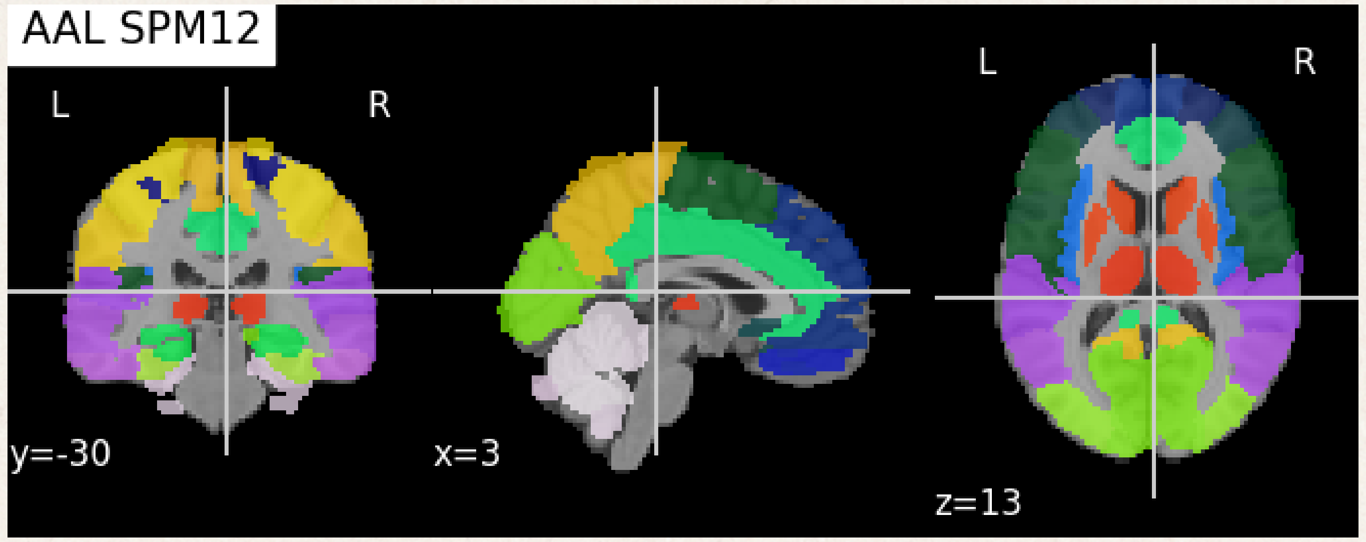
02 **Problem**: Bias-Variance Tradeoff

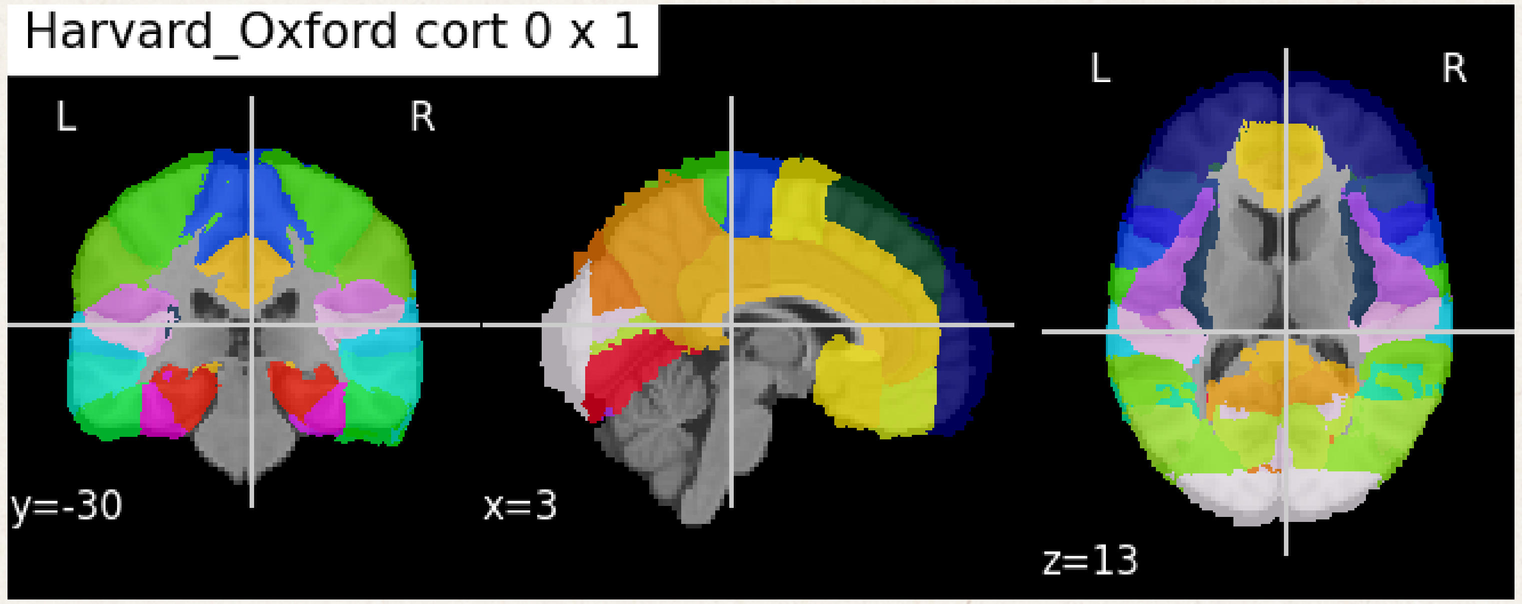
Solution: Calculate test and train metrics to identify overfitting

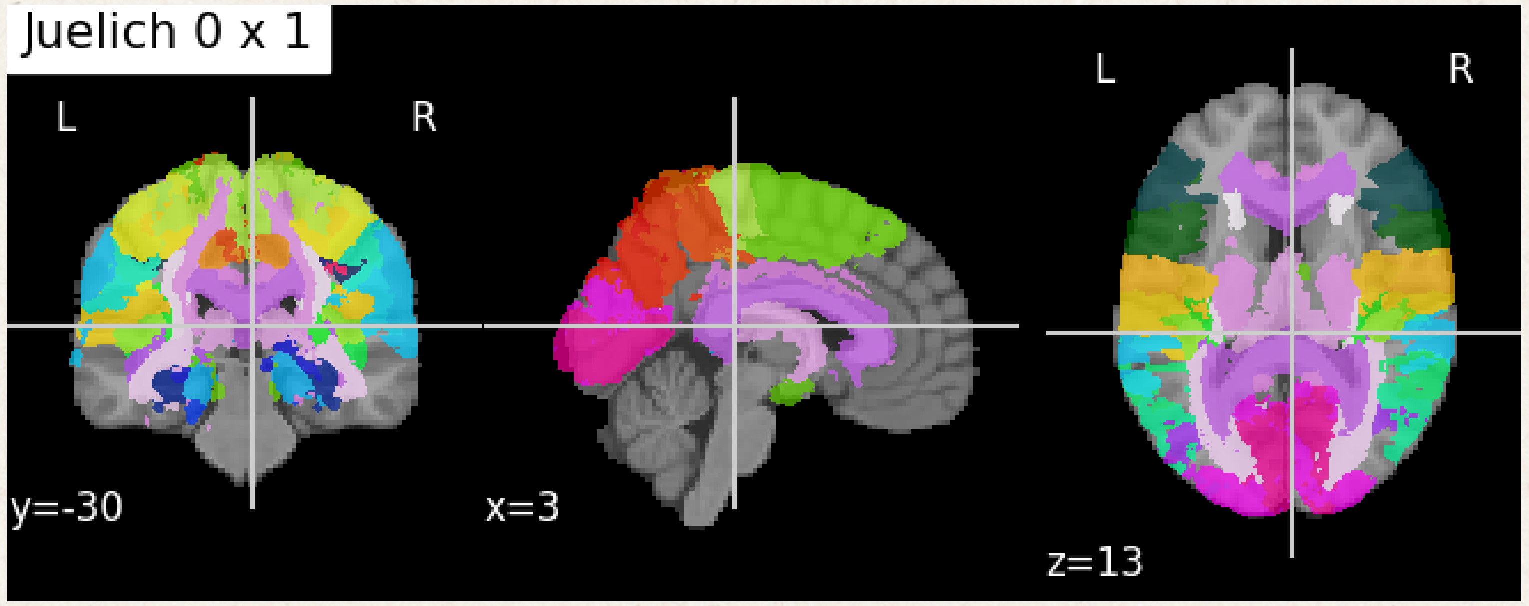
03 **Problem**: Explainability and Feature Importance

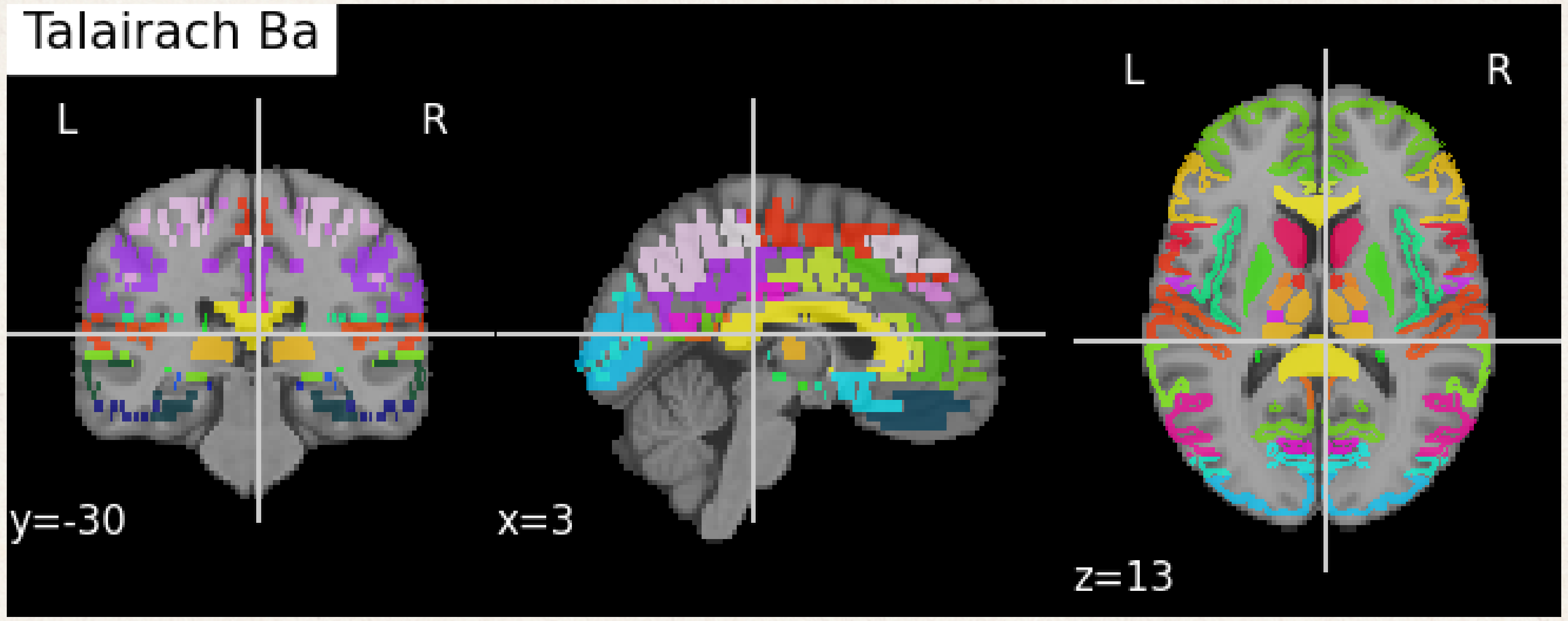
Solution: Choose appropriate model for classification task and output coefficients

Brain Parcellation Visualizations

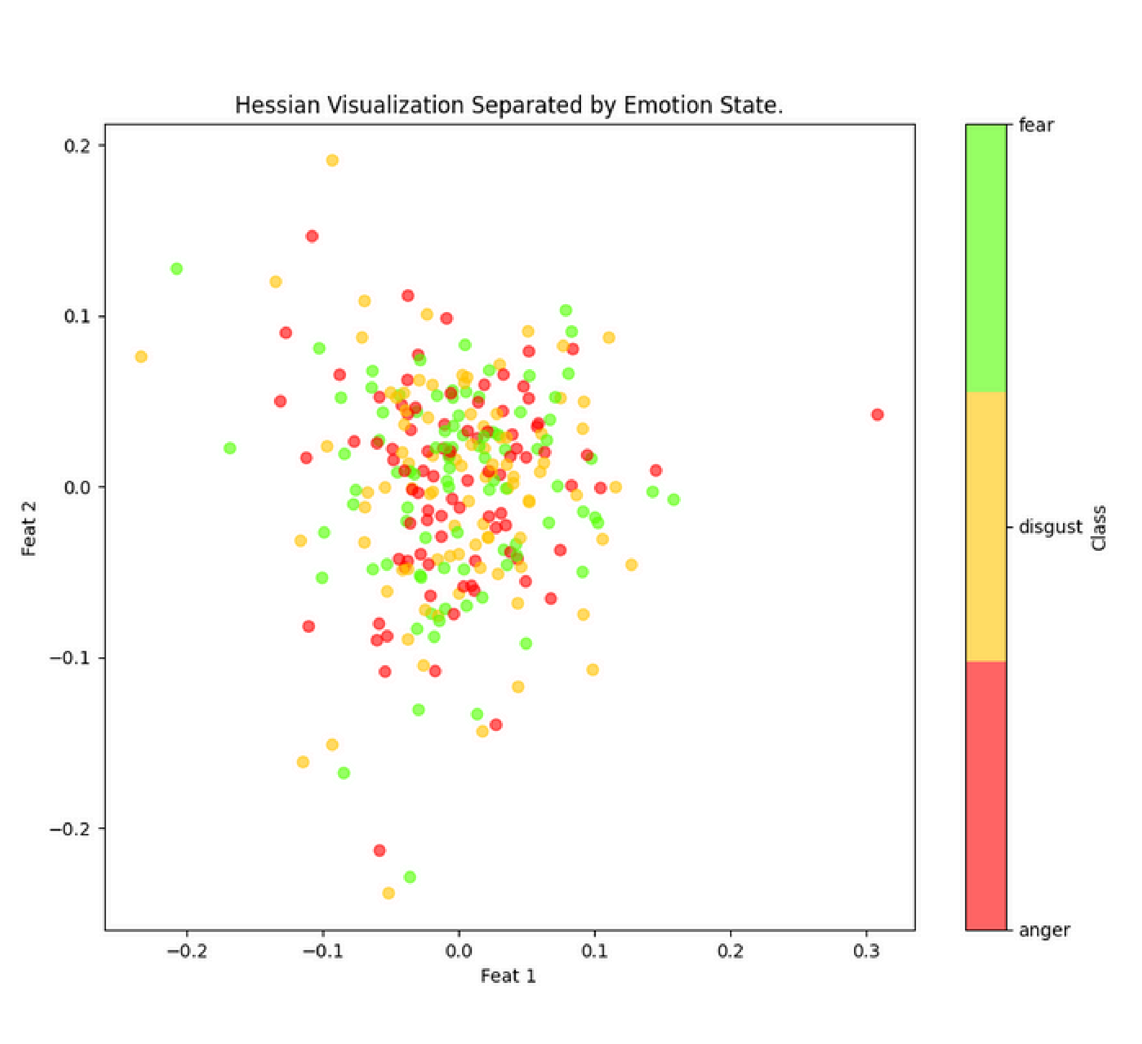


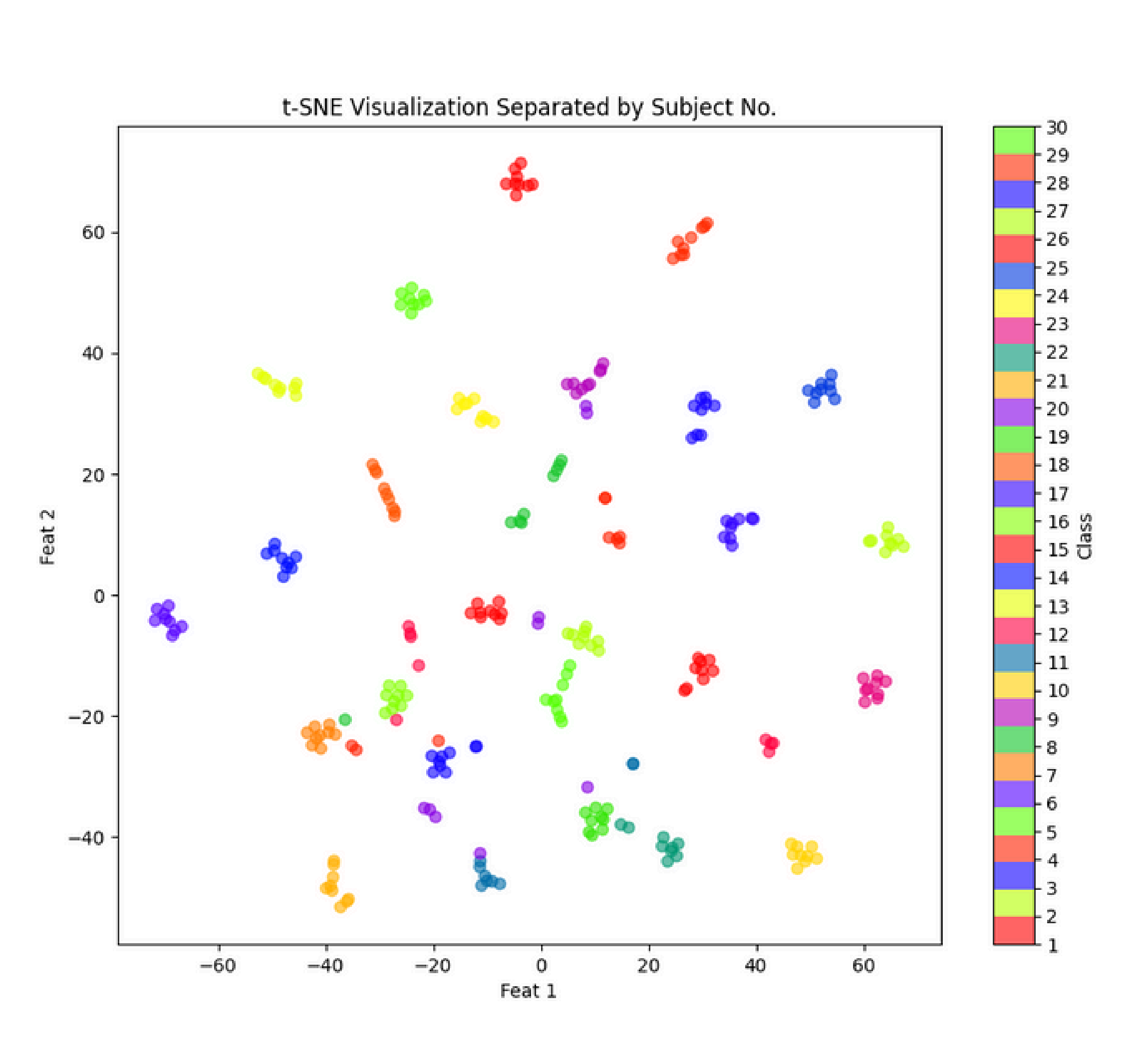


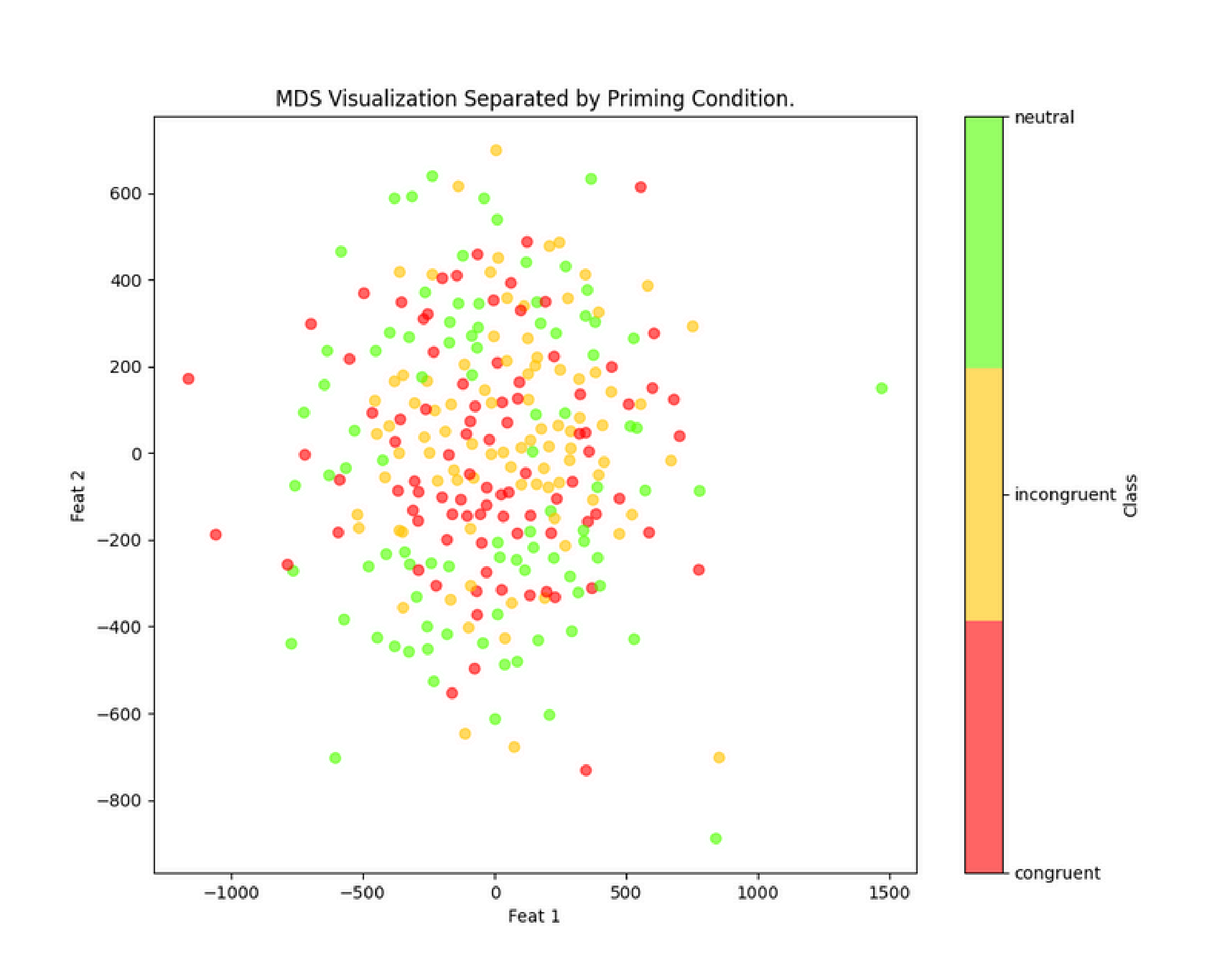




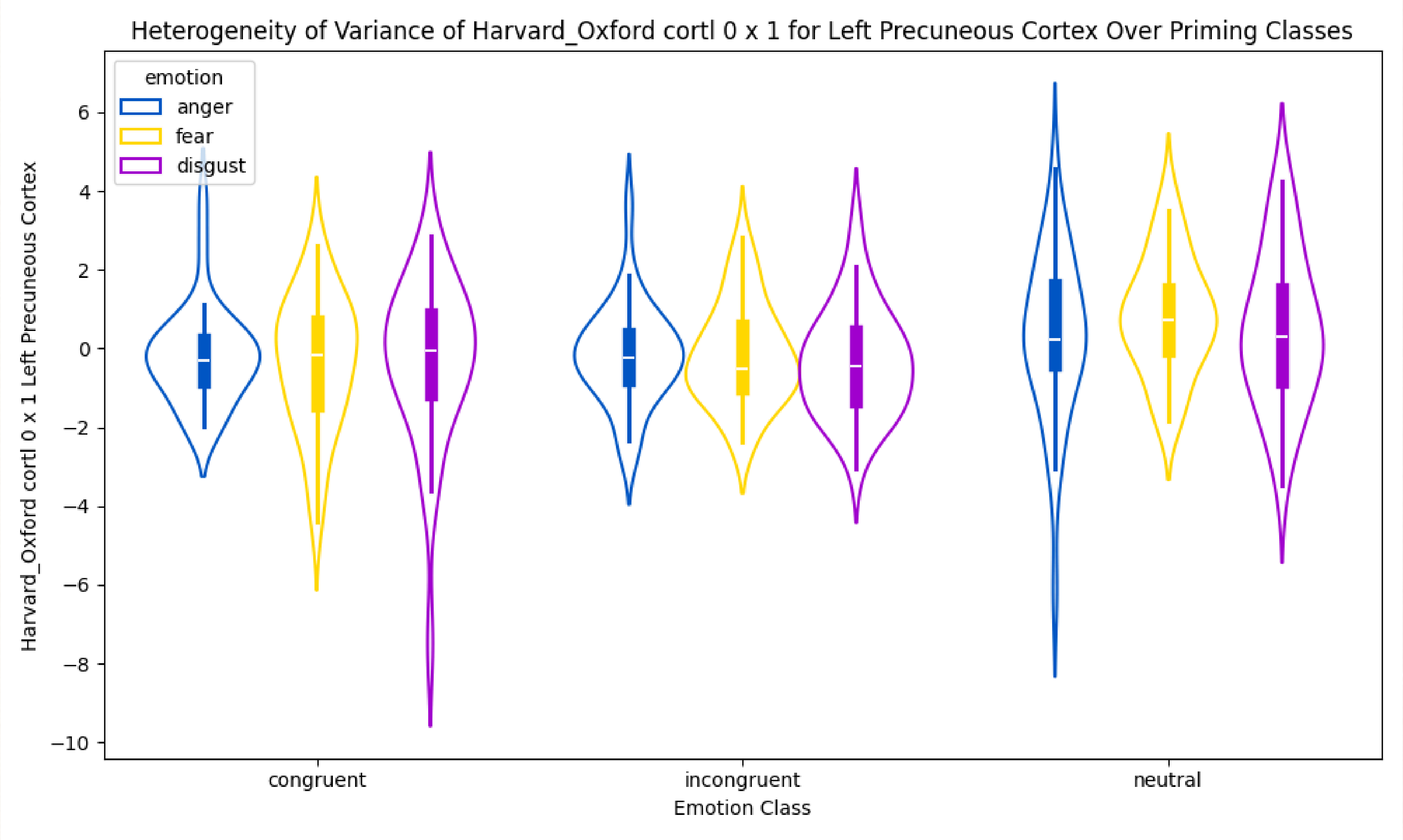
Geometric Projection Visualizations

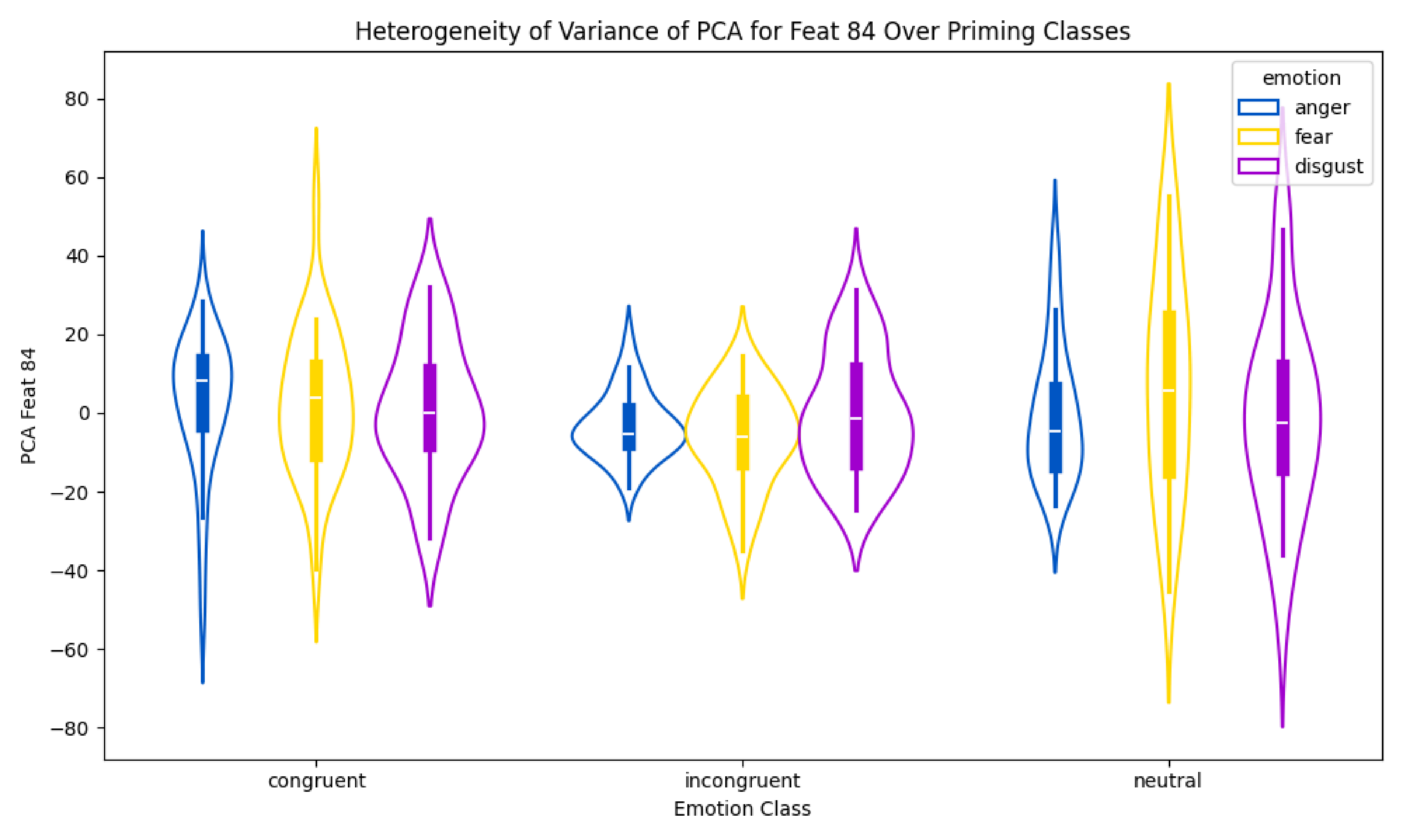




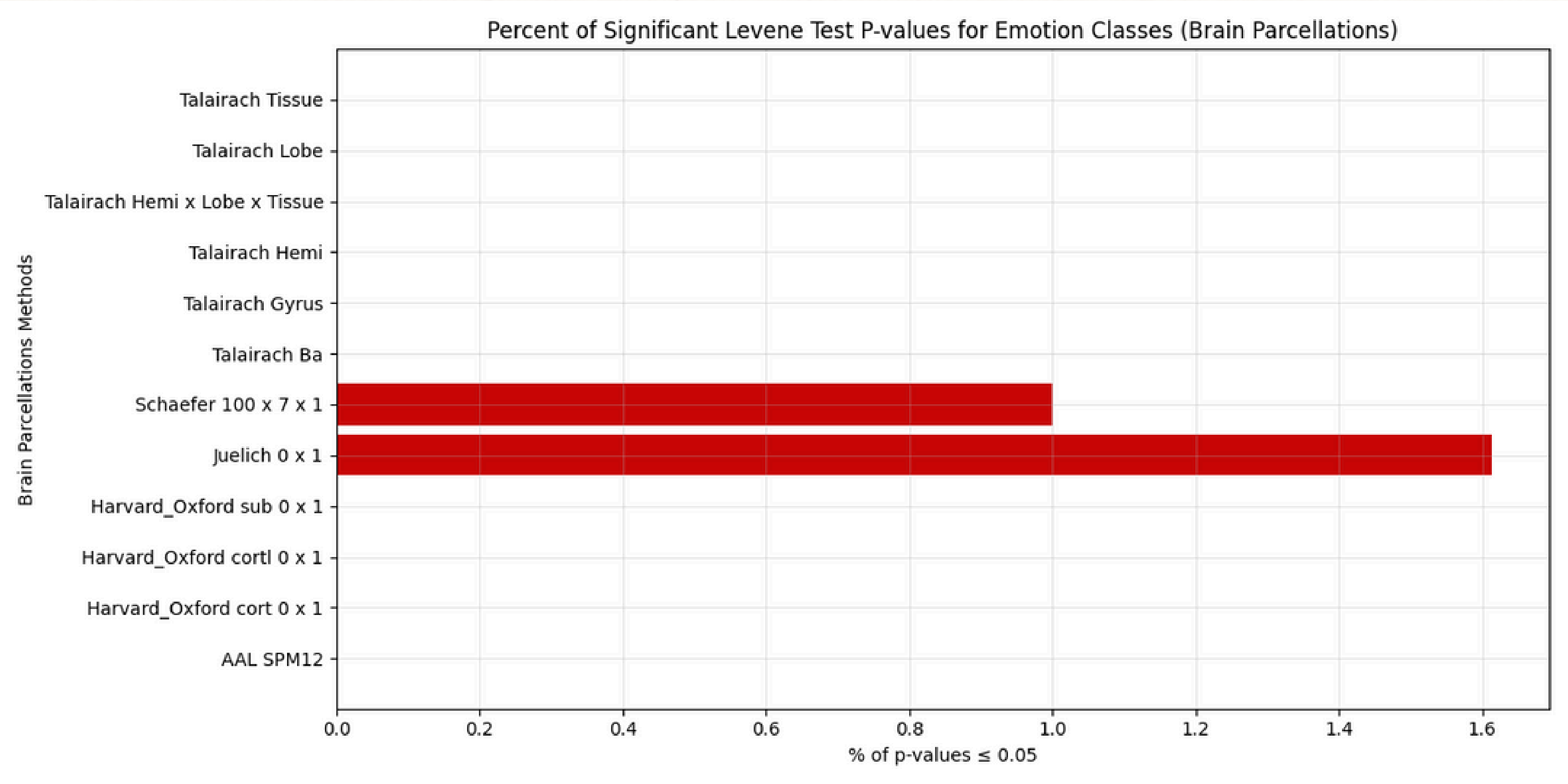


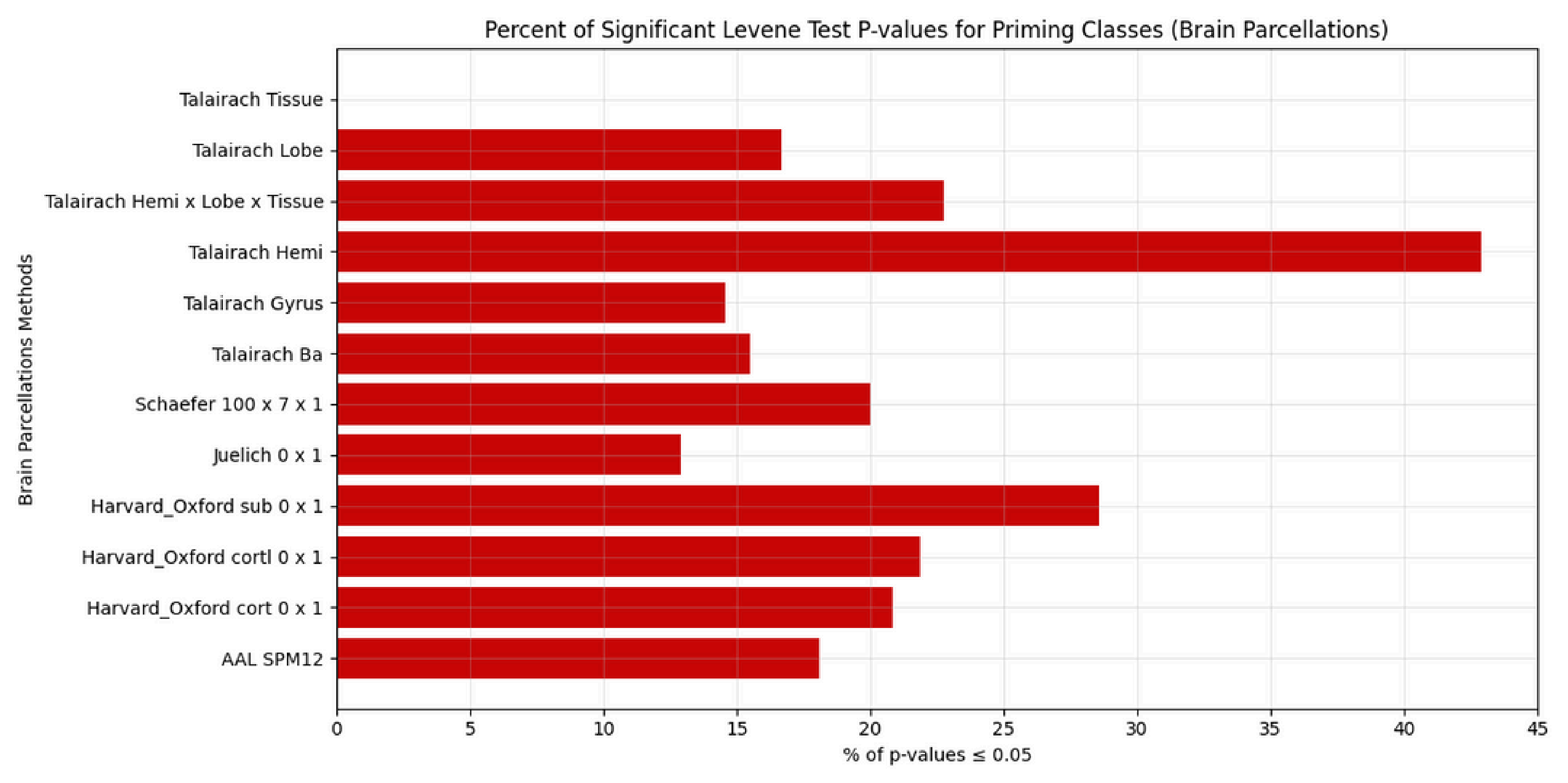
Heterogeneity of Variance



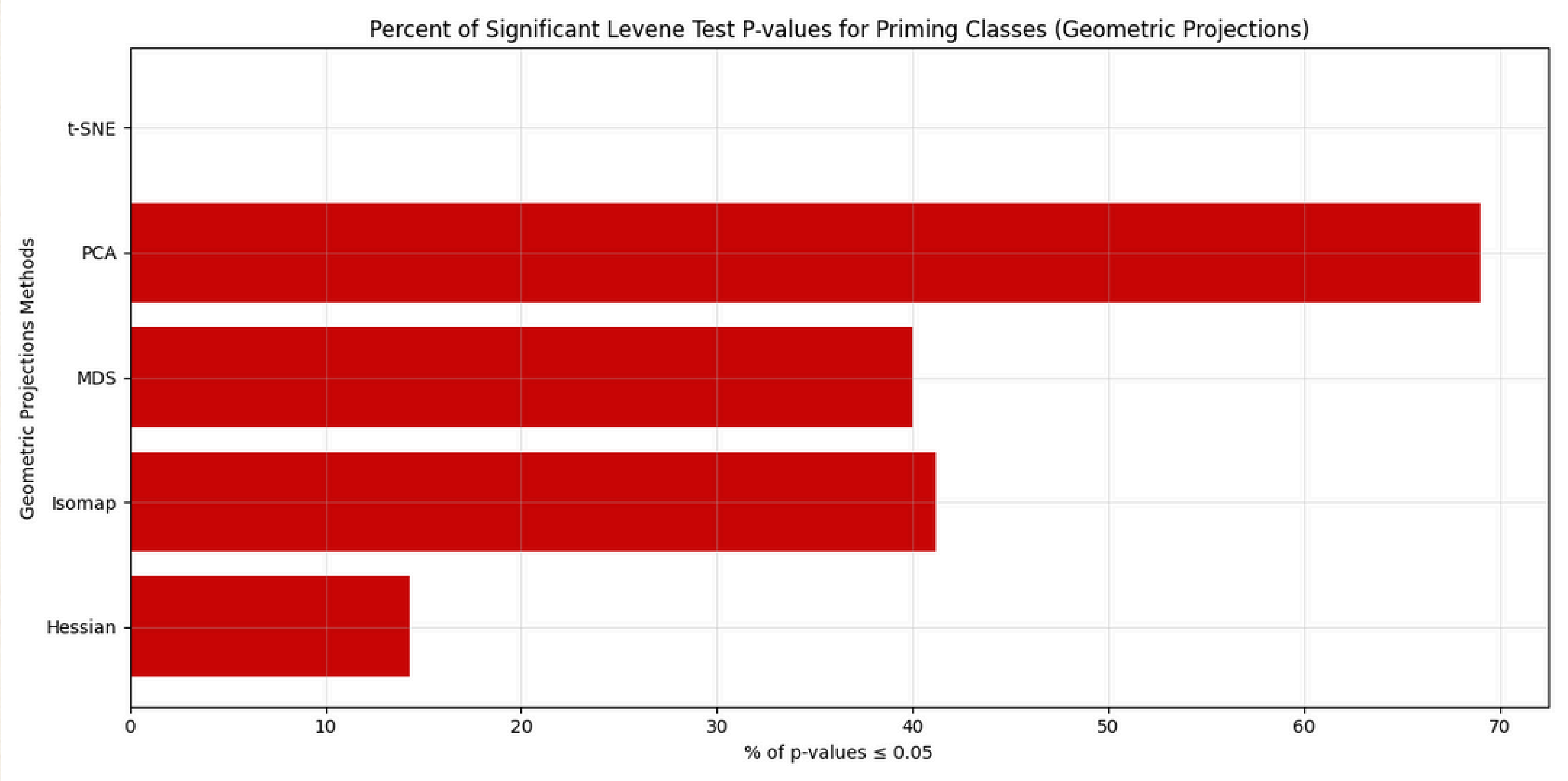


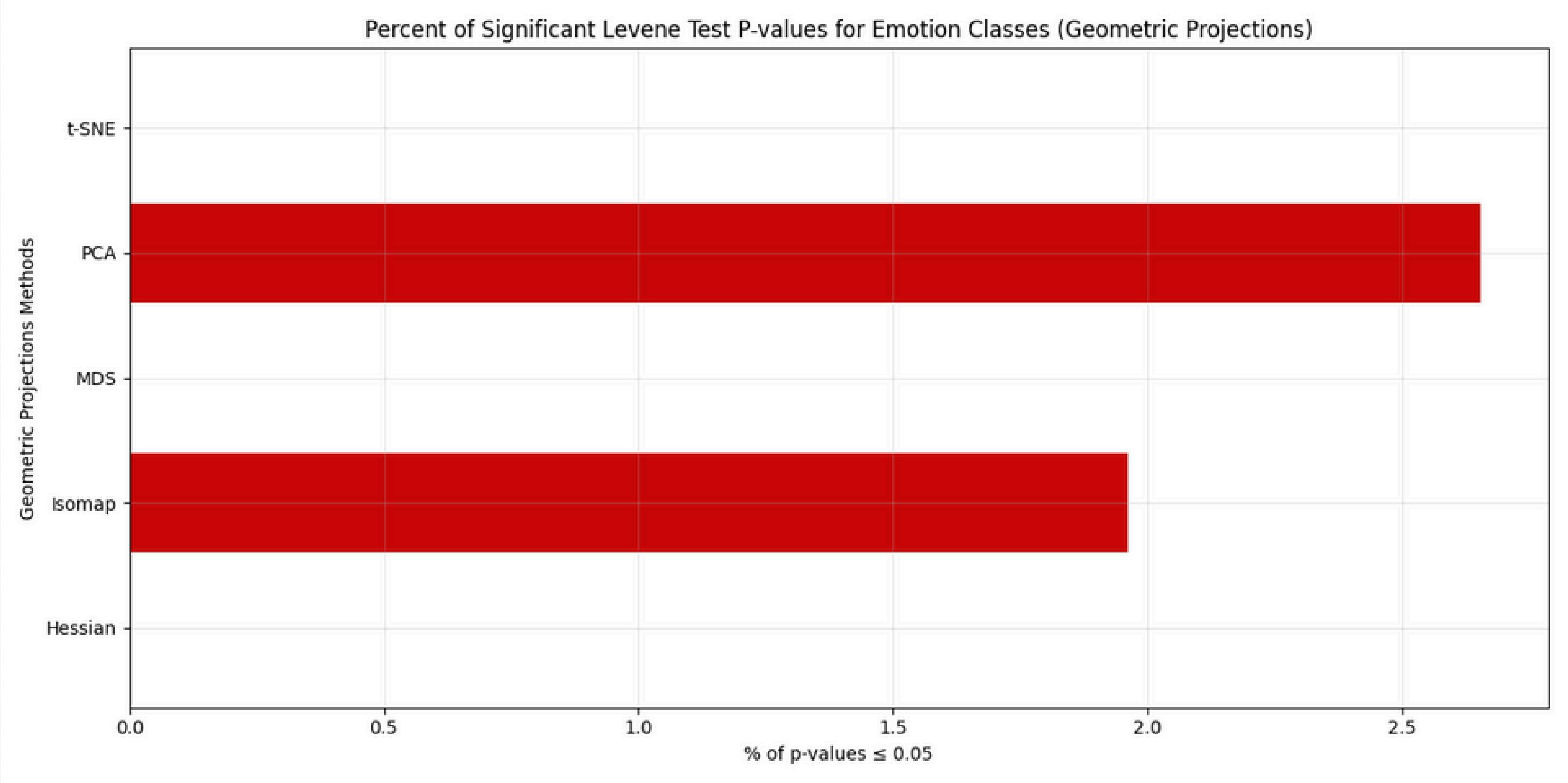
Heterogeneity of Variance pt. 2



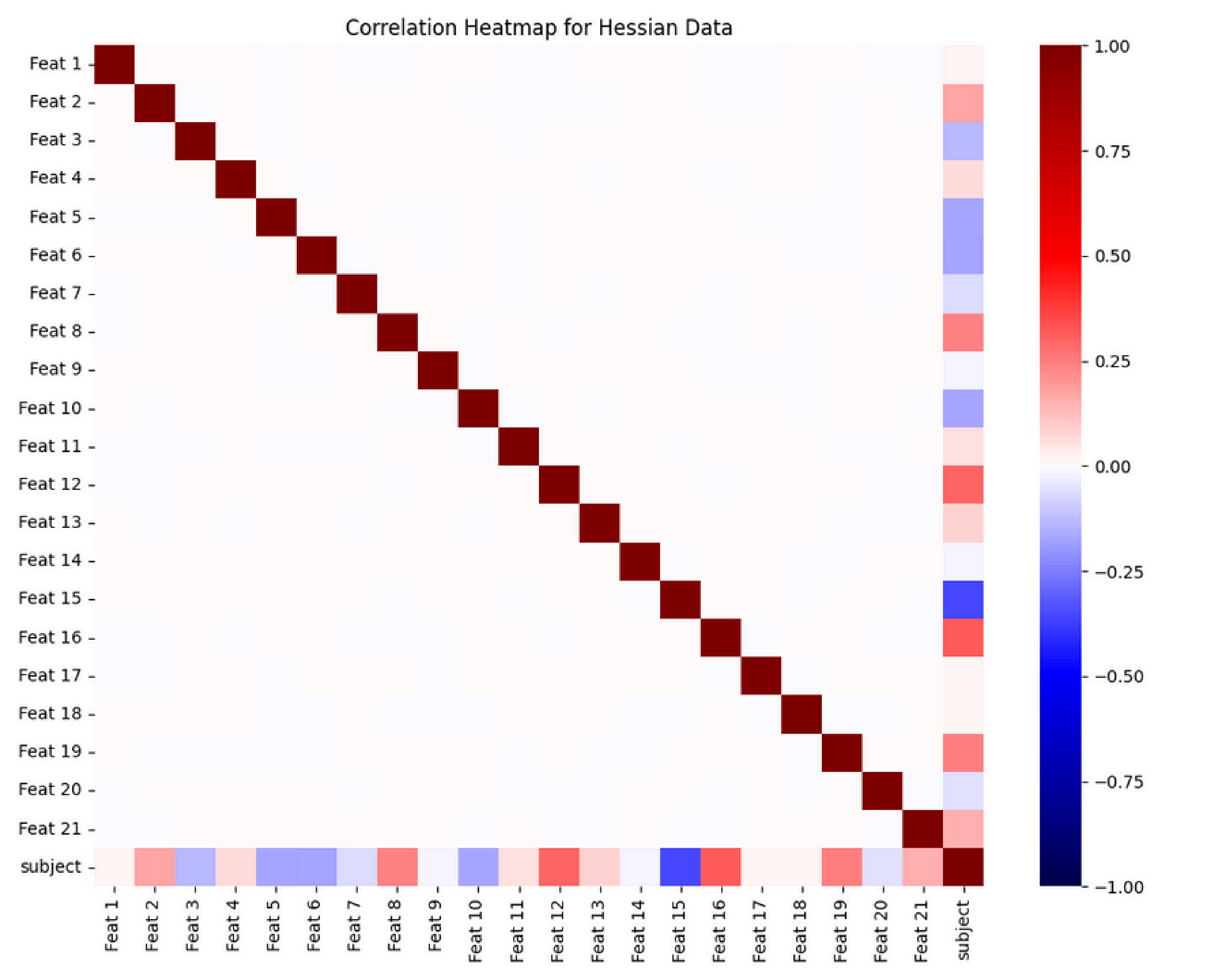


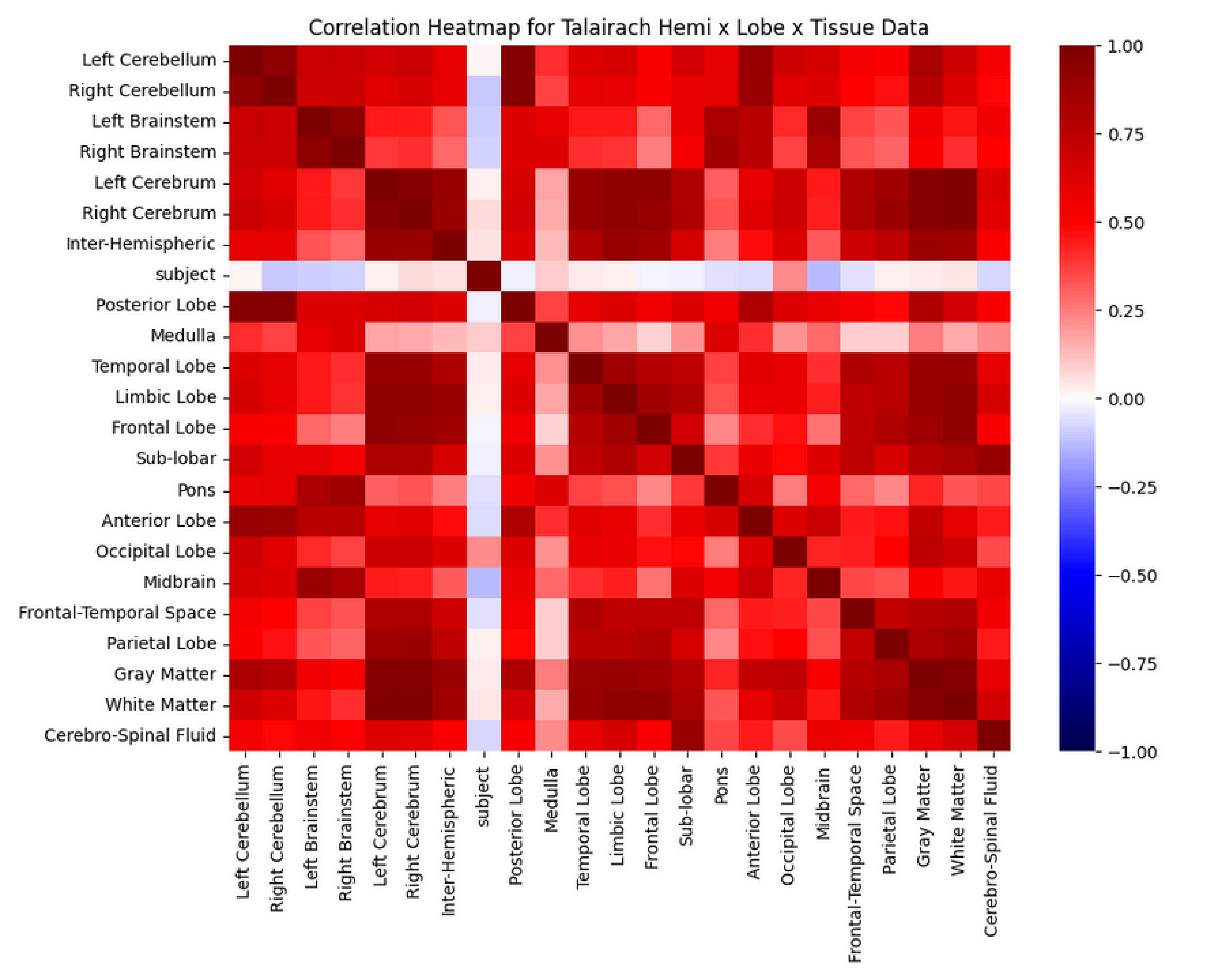
Heterogeneity of Variance pt. 2



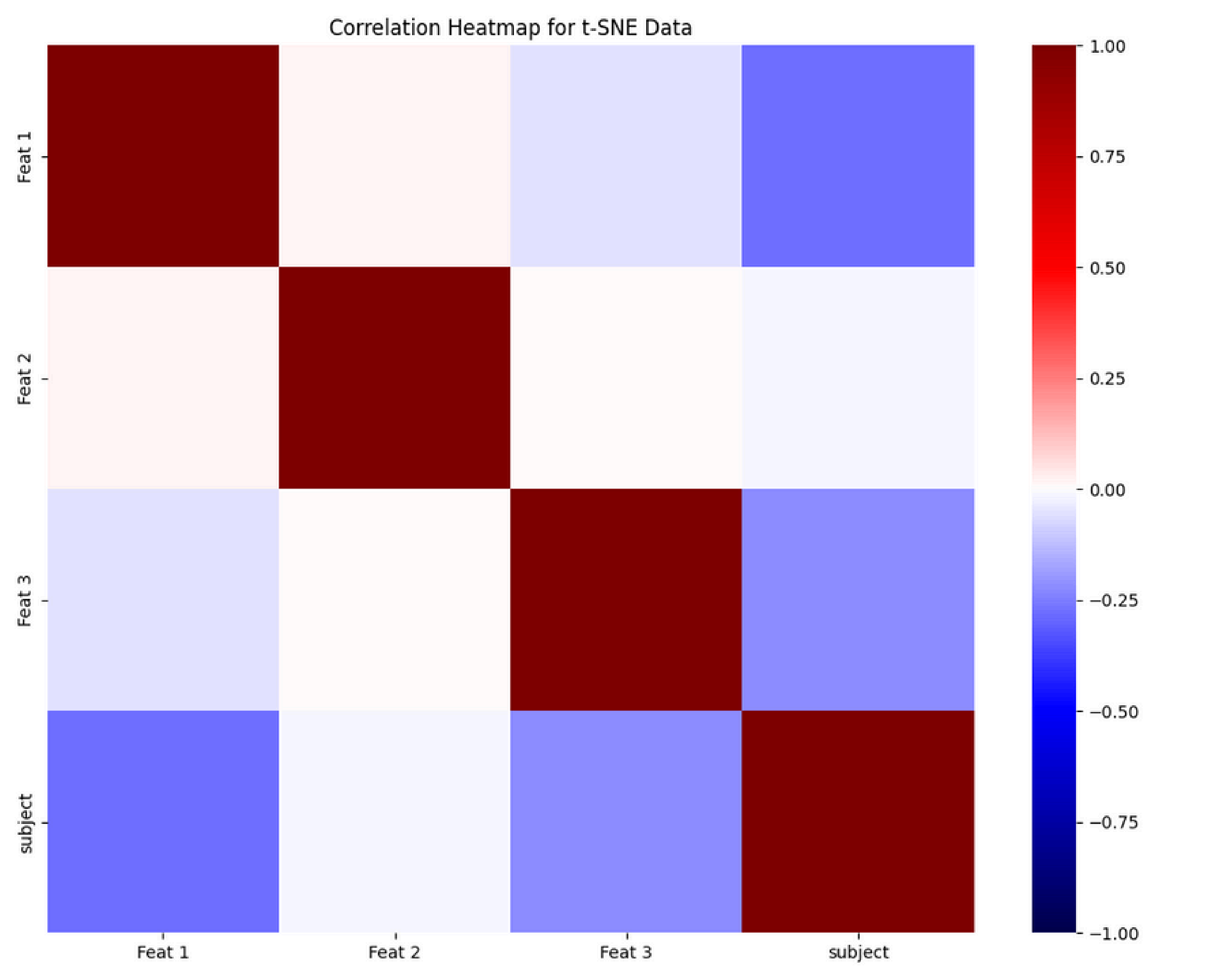


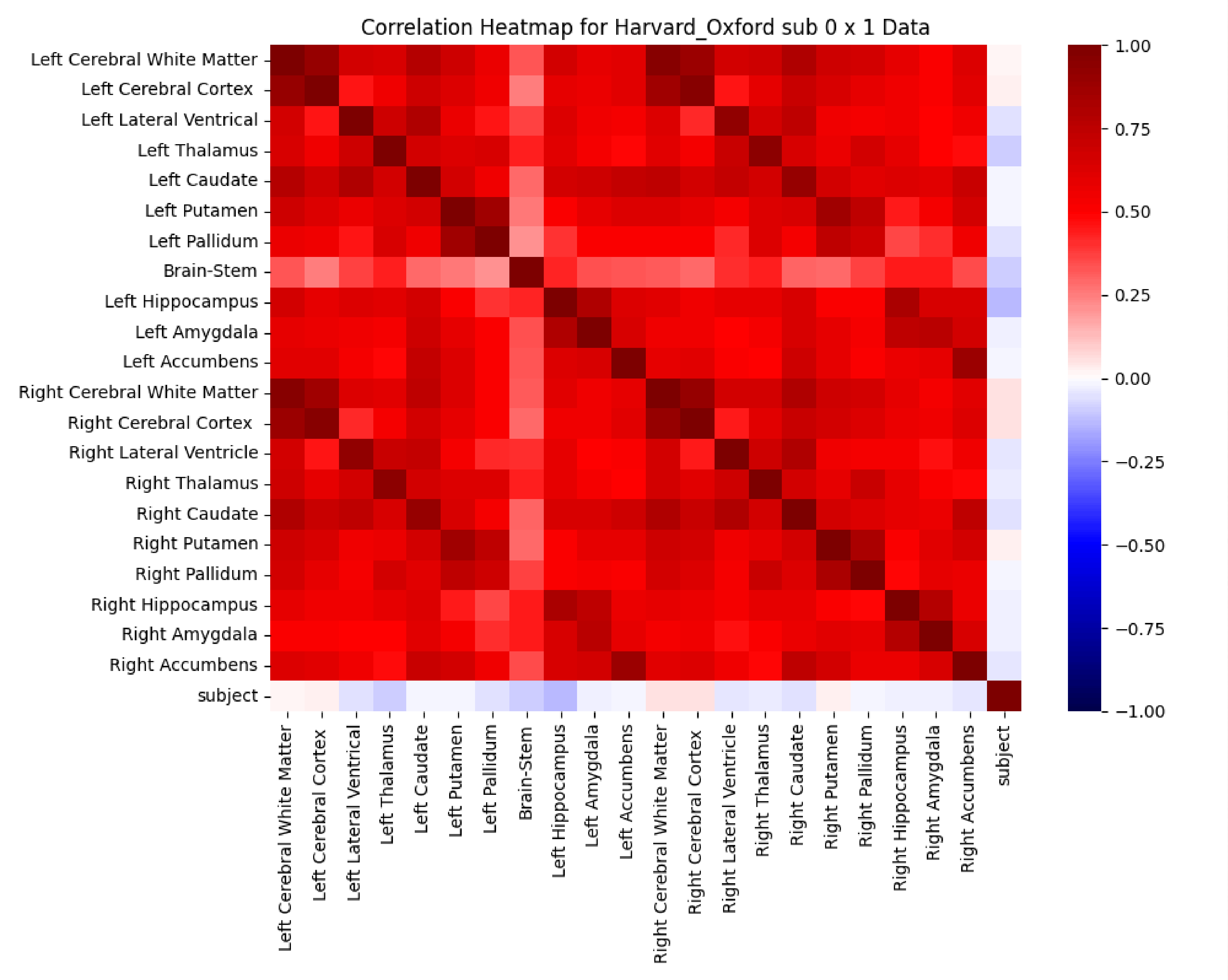
Multicollinearity



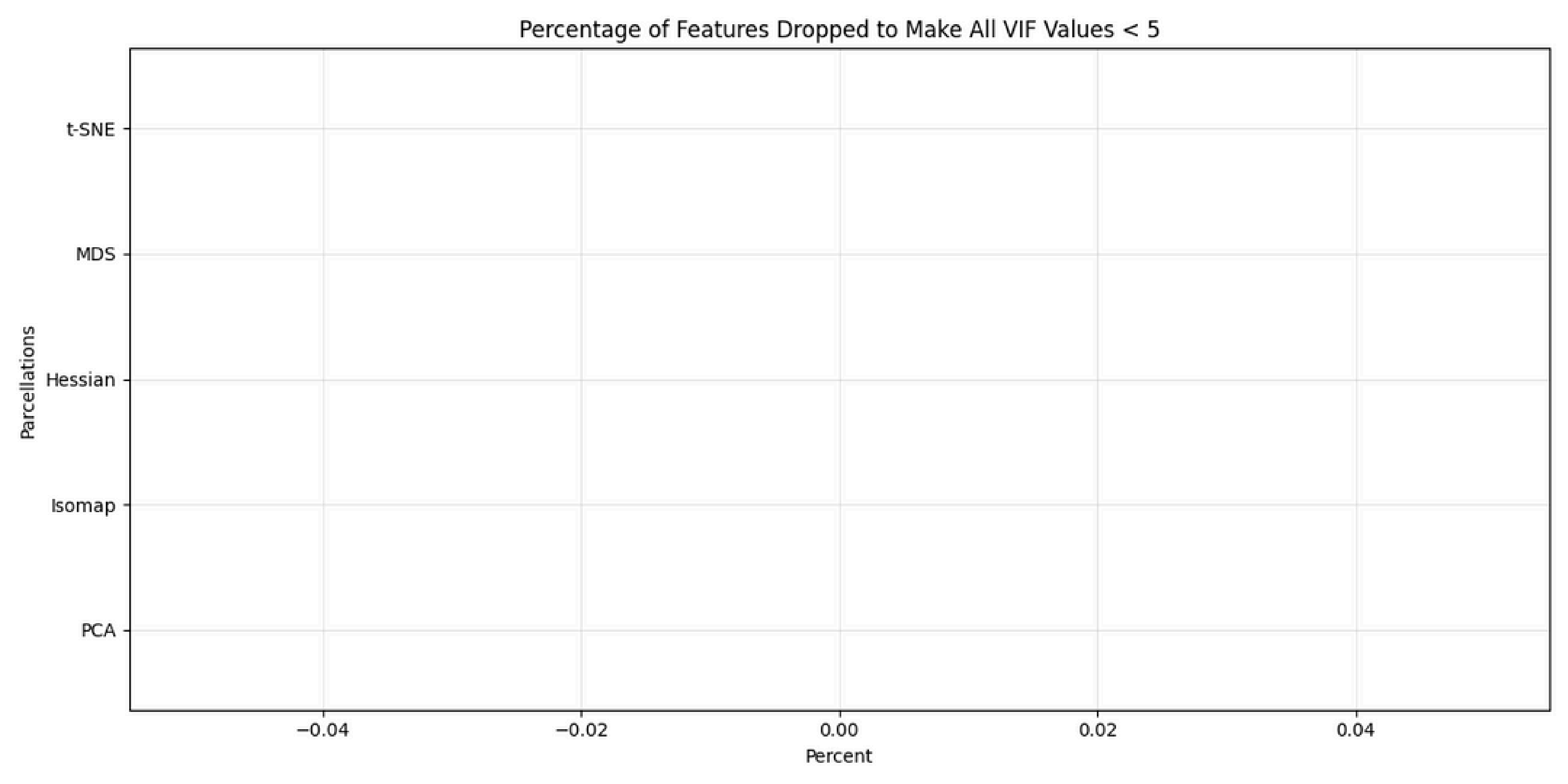


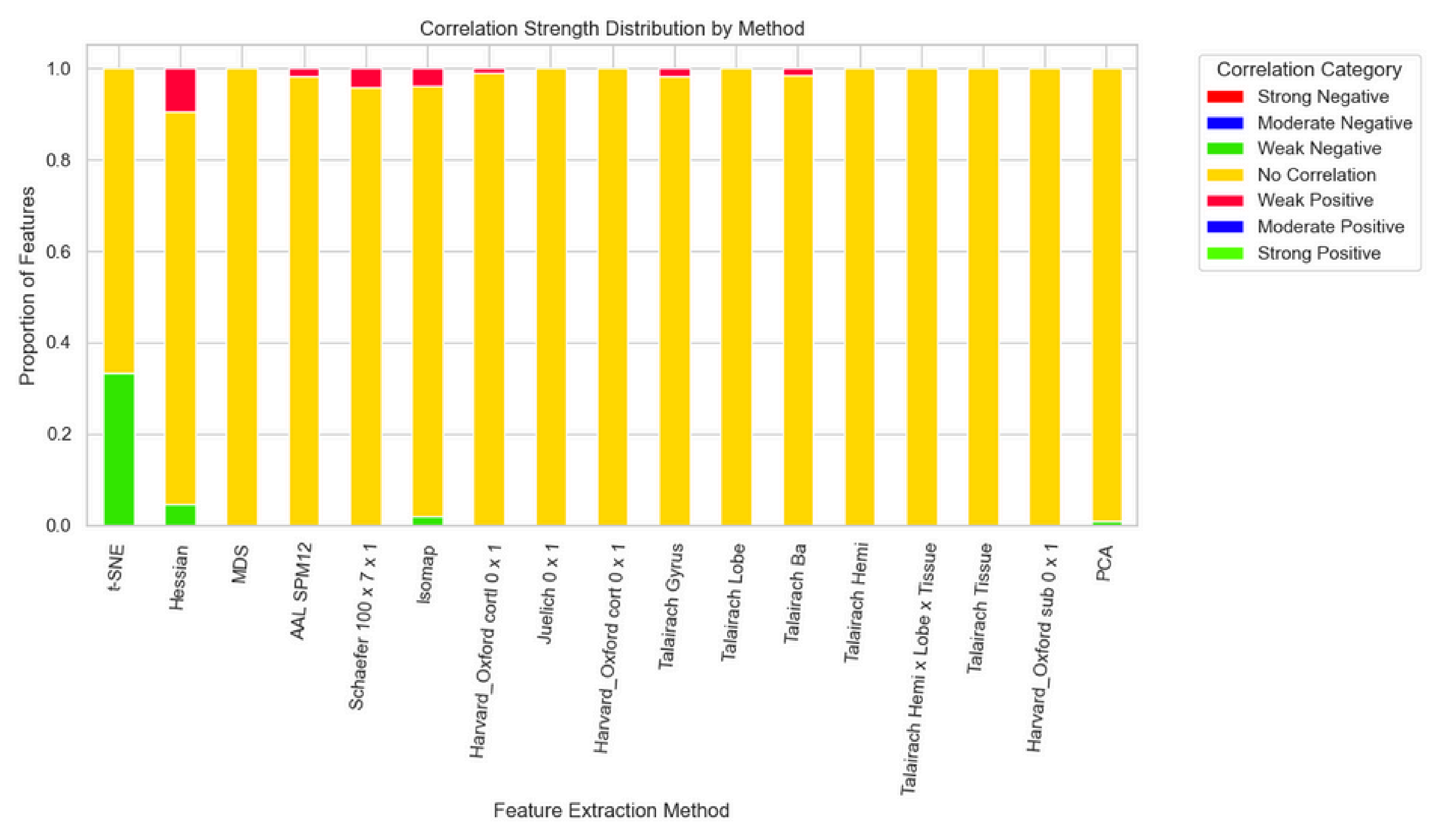
Multicollinearity





Multicollinearity





Thank you