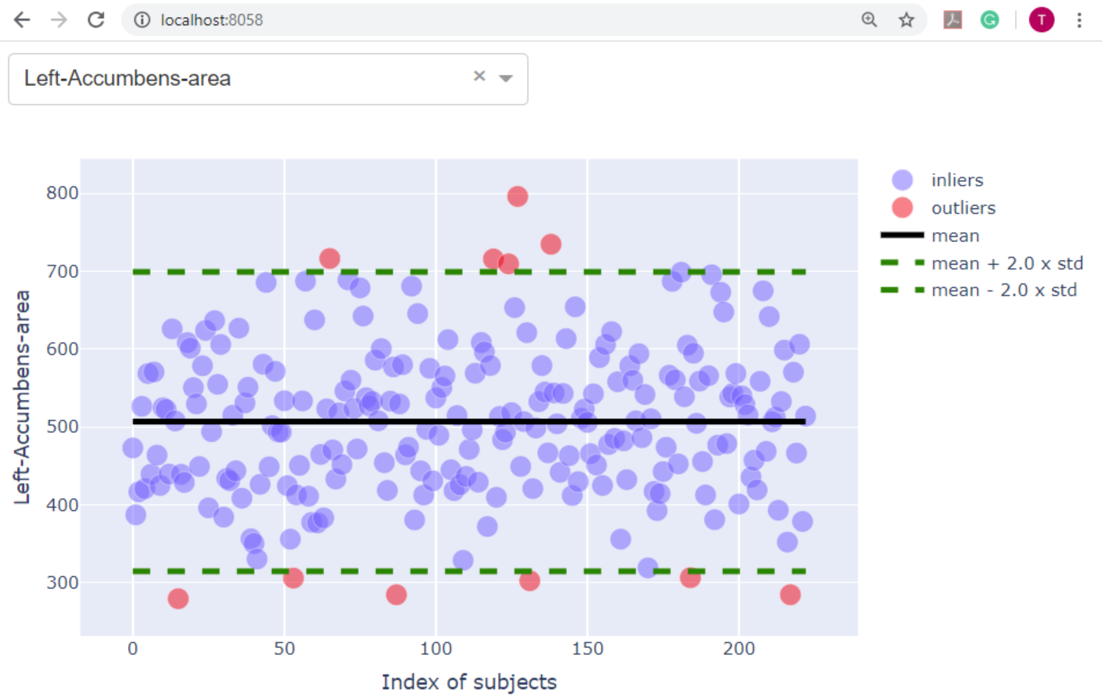
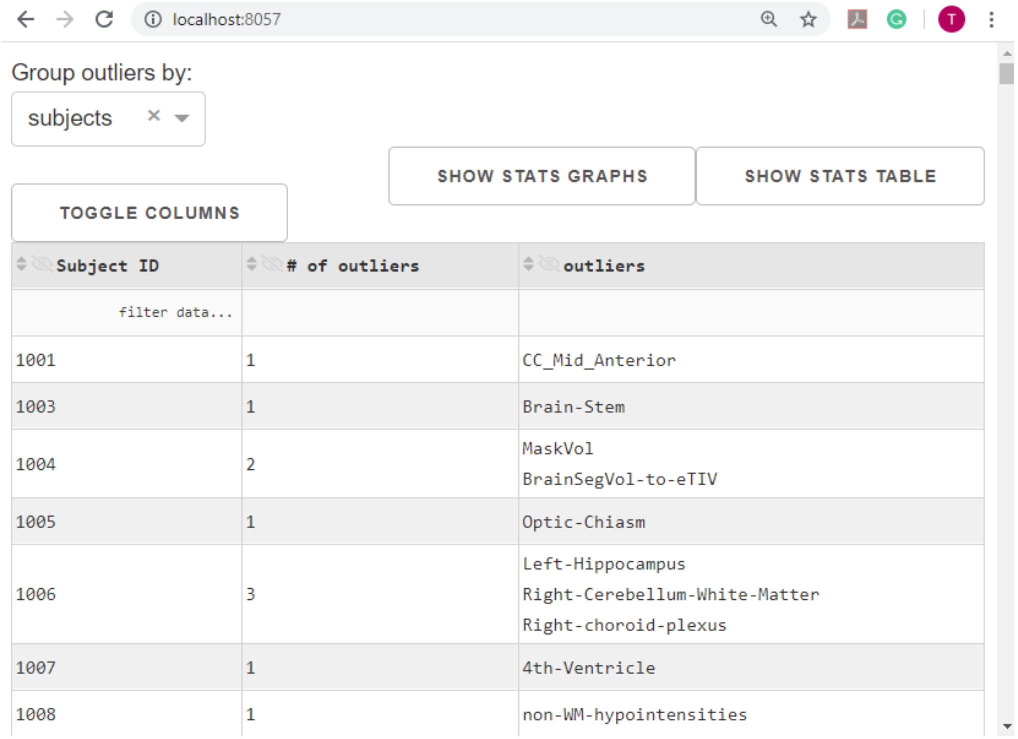
We developed a tool for performing statistical analysis on brain regions and displaying results on a web server. Input to the tool is a summary table with rows for subjects and columns for regions obtained from FreeSurfer statistics of a set of subjects. Although the tool is developed for analyzing FreeSurfer statistics, it can be readily employed with other statistics having a summary table such as those obtained from Tract-Based Spatial Statistics (TBSS) study.

To find outliers pertinent to each region, we standardize the statistics of each column. Subjects with standard scores beyond ±2 range (can be varied) are classified as outliers for that region. Using Dash, a Python framework for building web applications, we demonstrate the outlying subjects in a graph. Summary of outlying subjects for all regions are also put together in another webpage.

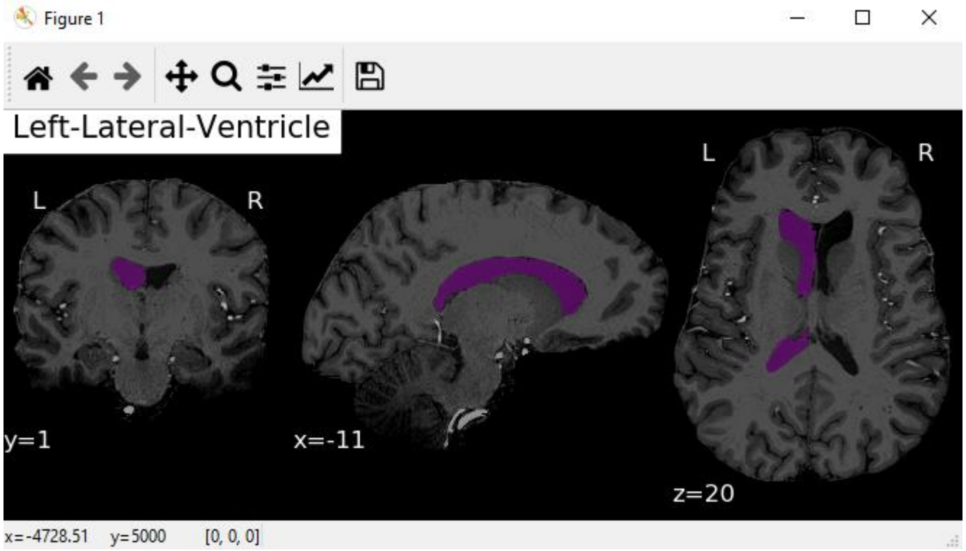




Group outliers by: subjects

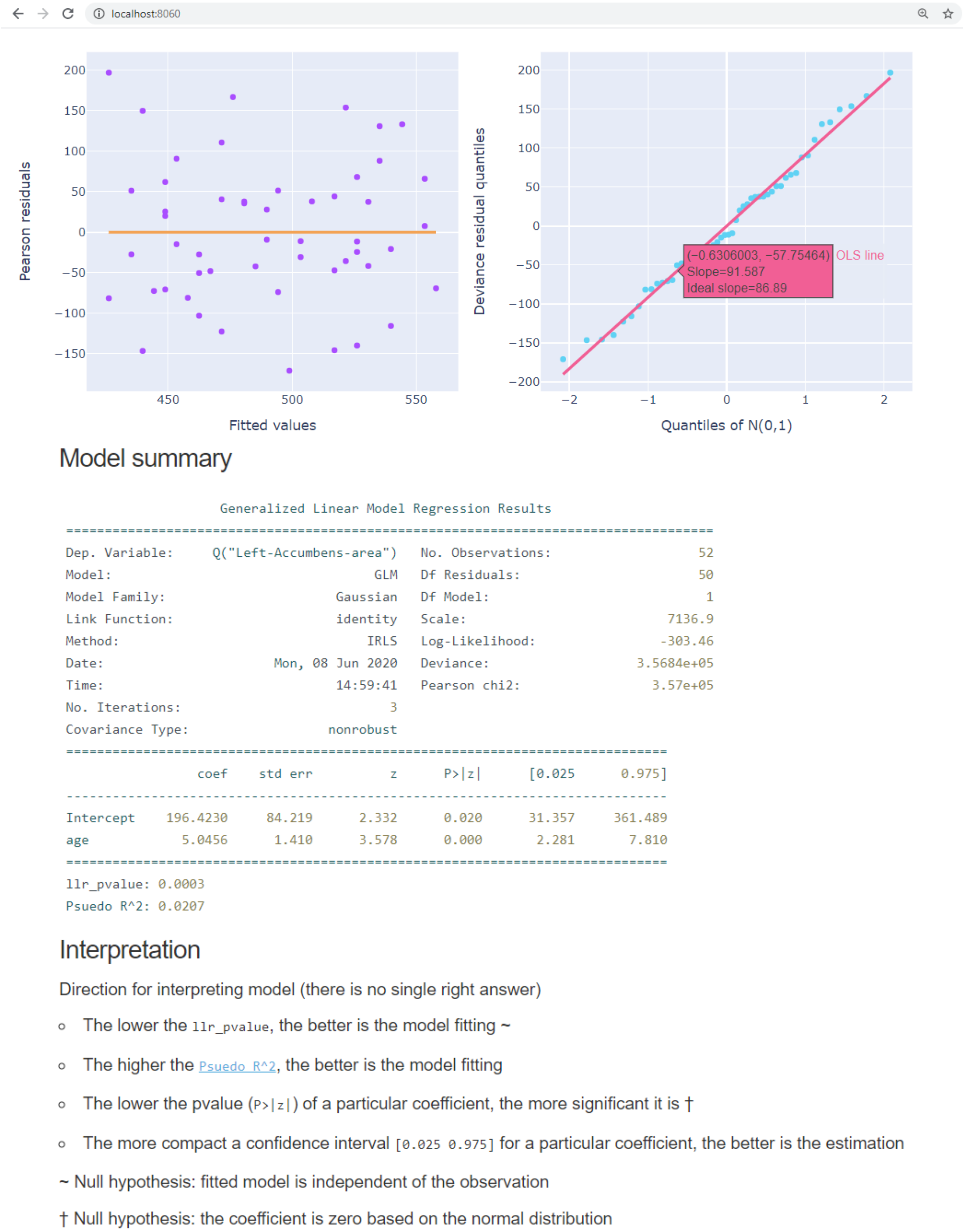
Subject ID	# of outliers	outliers
1001	1	CC_Mid_Anterior
1003	1	Brain-Stem
1004	2	MaskVol BrainSegVol-to-eTIV
1005	1	Optic-Chiasm
1006	3	Left-Hippocampus Right-Cerebellum-White-Matter Right-choroid-plexus
1007	1	4th-Ventricle
1008	1	non-WM-hypointensities

Moreover, the user can click on a cell of the summary table to view static (nilearn) or dynamic (freeview) rendering of the region of interest overlaid upon the subject's brain.



We also developed a means for outlier analysis accounting for demographics— age, weight, years of education etc. In this feature, we fit a Generalized Linear Model (GLM) over the statistics of each column i.e. region of a user-specified control group of subjects. The region statistics are response variables in the model while demographics are independent variables. We assume the responses are normally distributed and an identity link is present between the mean of response variables and the linear predictor over independent variables. Python package statsmodels is used to perform Maximum Likelihood Estimation (MLE) of the coefficients associated with independent variables (can be more than one, user specified). Finally, we declare each subject an outlier based on residuals between the predicted responses and given responses. To reach that decision, residuals are standardized in the same way we did for detecting outliers without considering the effect of demographics.

The goodness of fit of the GLM can be realized through graphical and parametrical interpretation that we provide in the summary webpage. As outlined in Gill 2001, graphical interpretation includes observing a Quantile-Quantile (Q-Q) plot where residuals follow a straight line with a slope almost equal to the standard deviation of the residuals for a good model. On the other hand, parametrical interpretation includes observing the value of an R^2 statistics ~ [0,1], the greater it is, the better is the model.



Model summary

```
                 Generalized Linear Model Regression Results
==================================================================================
Dep. Variable:     Q("Left-Accumbens-area")   No. Observations:                 52
Model:                                  GLM   Df Residuals:                     50
Model Family:                      Gaussian   Df Model:                          1
Link Function:                     identity   Scale:                        7136.9
Method:                                IRLS   Log-Likelihood:              -303.46
Date:                      Mon, 08 Jun 2020   Deviance:                 3.5684e+05
Time:                              14:59:41   Pearson chi2:               3.57e+05
No. Iterations:                           3
Covariance Type:                  nonrobust
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept    196.4230     84.219      2.332      0.020      31.357     361.489
age            5.0456      1.410      3.578      0.000       2.281       7.810
==============================================================================
llr_pvalue: 0.0003
Psuedo R^2: 0.0207
```

Interpretation

Direction for interpreting model (there is no single right answer)

- The lower the `llr_pvalue`, the better is the model fitting ~
- The higher the `Psuedo R^2`, the better is the model fitting
- The lower the pvalue (`P>|z|`) of a particular coefficient, the more significant it is †
- The more compact a confidence interval `[0.025 0.975]` for a particular coefficient, the better is the estimation

~ Null hypothesis: fitted model is independent of the observation

† Null hypothesis: the coefficient is zero based on the normal distribution