Clinical data hydra results (primarily using Hydra_k2)

Erica Baller

3/16/2018

This script goes through demographics, clinical scores, health, and psych summaries, adds clustering information, runs statistics and makes graphs from results.

Part 1 : Read in csv.s -This script reads in demographics, clinical_scores, health and psych summaries, merges them, removes NAs, codes and separates by depression.

Part 2 : merge with hydra -It then merges these documents with hydra output (made in cbica), adding Hydra_k1 through Hydra_k10 columns (which represent the number of clusters) -The script reads in 3 different types of groups (matched, unmatched, and residualized unmatched groups), and also does all gender together as well as separating them by gender.

Part 3 : Demographics tables - Demographics tables for each group (matched, unmatched, resid) were produced

Part 4 : Graphing - Graphs were then made.
For continuous variables(age, medu1), the graphs represent means, with SEM as error bars For categorical variables (race, sex) the graphs are percentages (caucasian, male) per group, with chisq used to calculate significance

Part 5 : LM -The script then runs LM on each cognitive score (clinical_measure ~ hydra_group).
-There is a test option that does this for all clinical measures and all hydra groups, but for the remainder of the analysis, Hydra_k2 was the only classification more deeply explored.

Part 6: Visreg : Look at results of linear model graphically -Allows you to visualize each cluster by cognitive measure

Part 7 : Anova -Anovas were also run on the results of the LM of each clinical value by cluster.

Part 8 : FDR Correction -FDR correction was calculated for each clinical measure ANOVA output -A table of the results was extracted

Part 1-2: Prep (read in csvs and merge with hydra)

Part 3: Demographics

```
##                            Stratified by Cluster
##                             level           -1              1
##   n                                         711            376
##   Race (%)               Caucasian          393 ( 55.3)    155 ( 41.2)
##                          Non-caucasian      318 ( 44.7)    221 ( 58.8)
##   Sex (%)                Female             476 ( 66.9)    270 ( 71.8)
##                          Male               235 ( 33.1)    106 ( 28.2)
##   Maternal Ed (mean (sd))                 14.14 (2.26)   13.75 (2.23)
##   Age (mean (sd))                         16.11 (3.00)   15.66 (3.14)
##   Depression (%)         Depressed            0 (  0.0)    376 (100.0)
##                          Non-depressed      711 (100.0)      0 (  0.0)
##   Cluster (%)            -1                  711 (100.0)      0 (  0.0)
##                          1                     0 (  0.0)    376 (100.0)
```
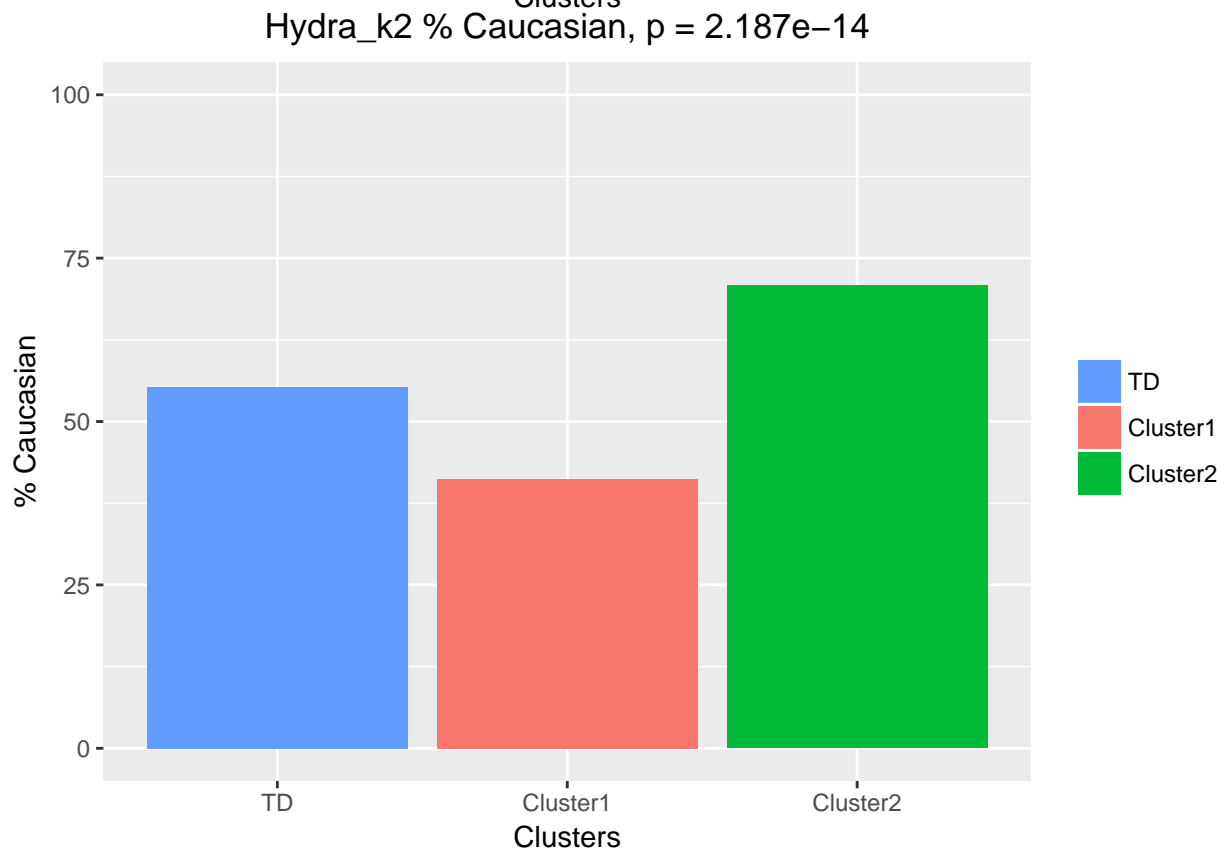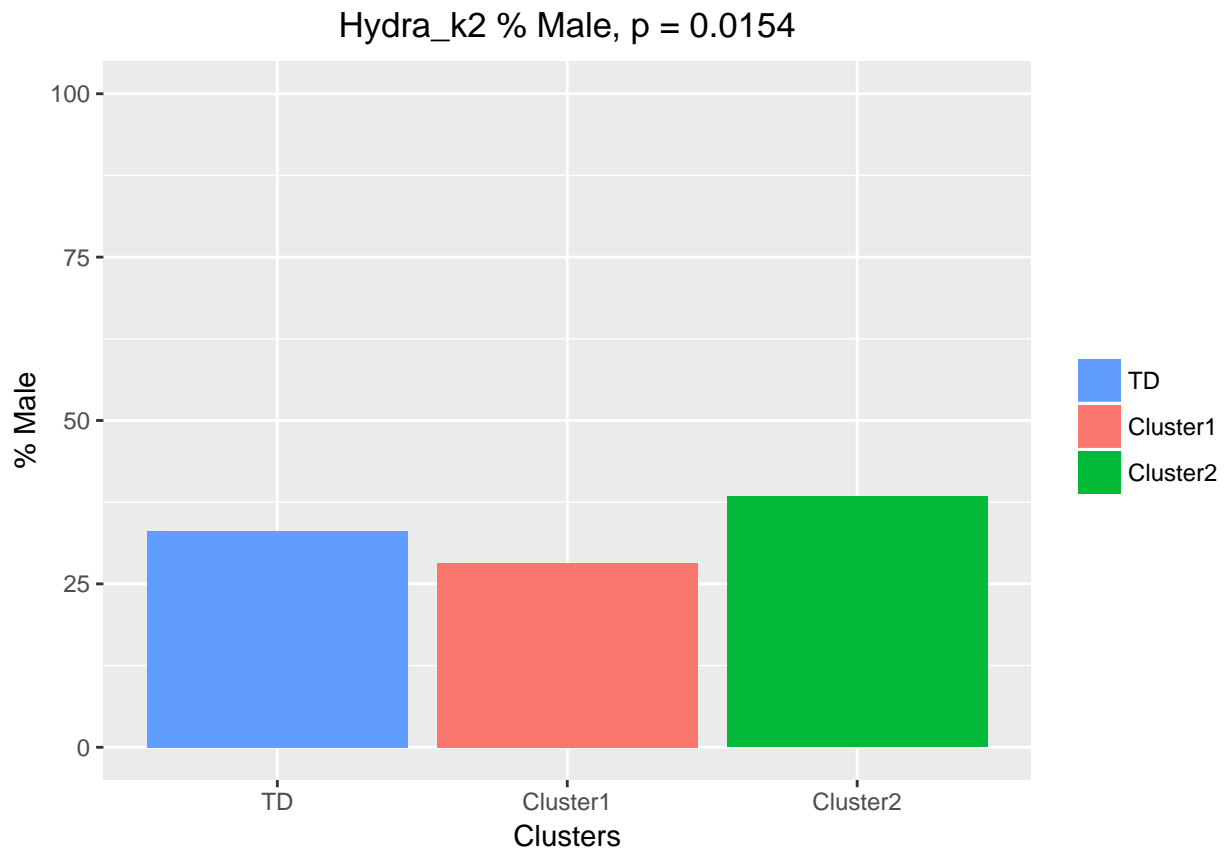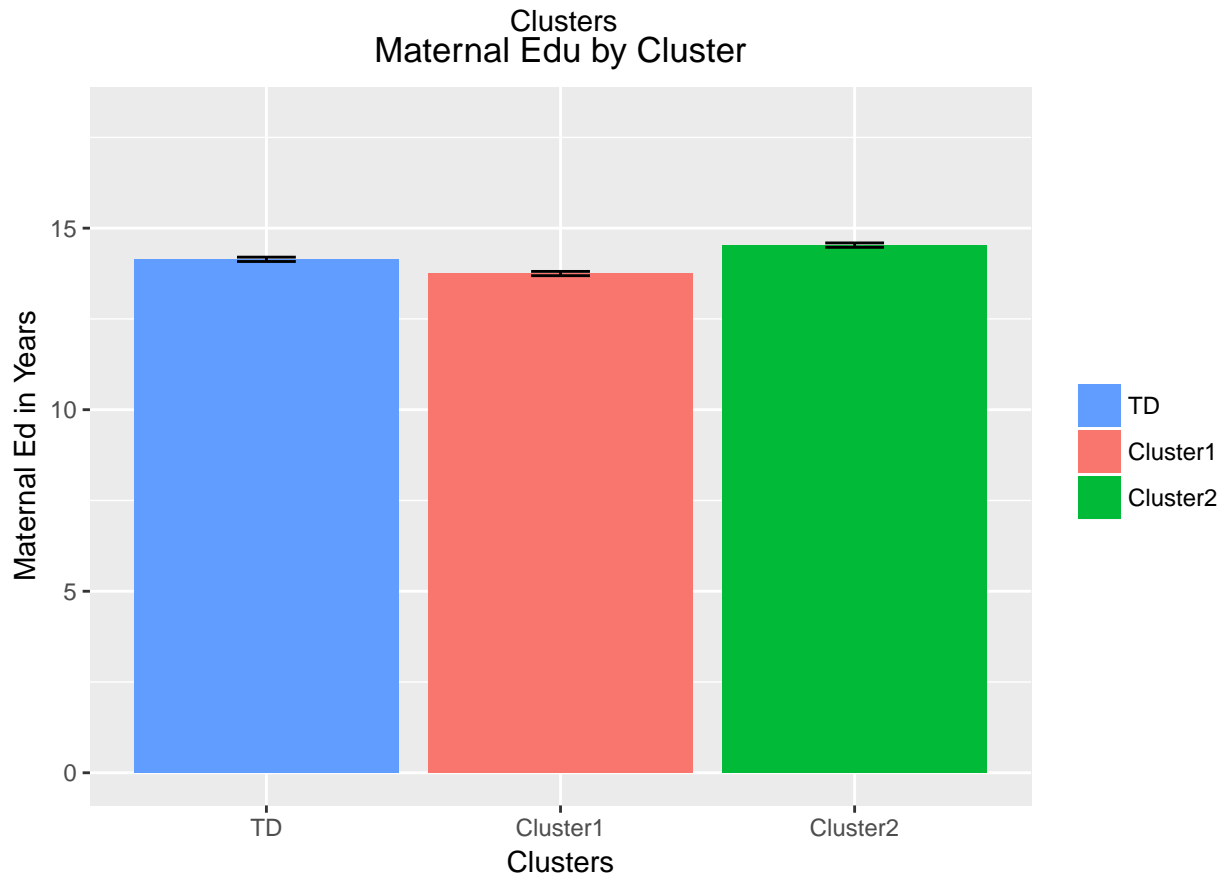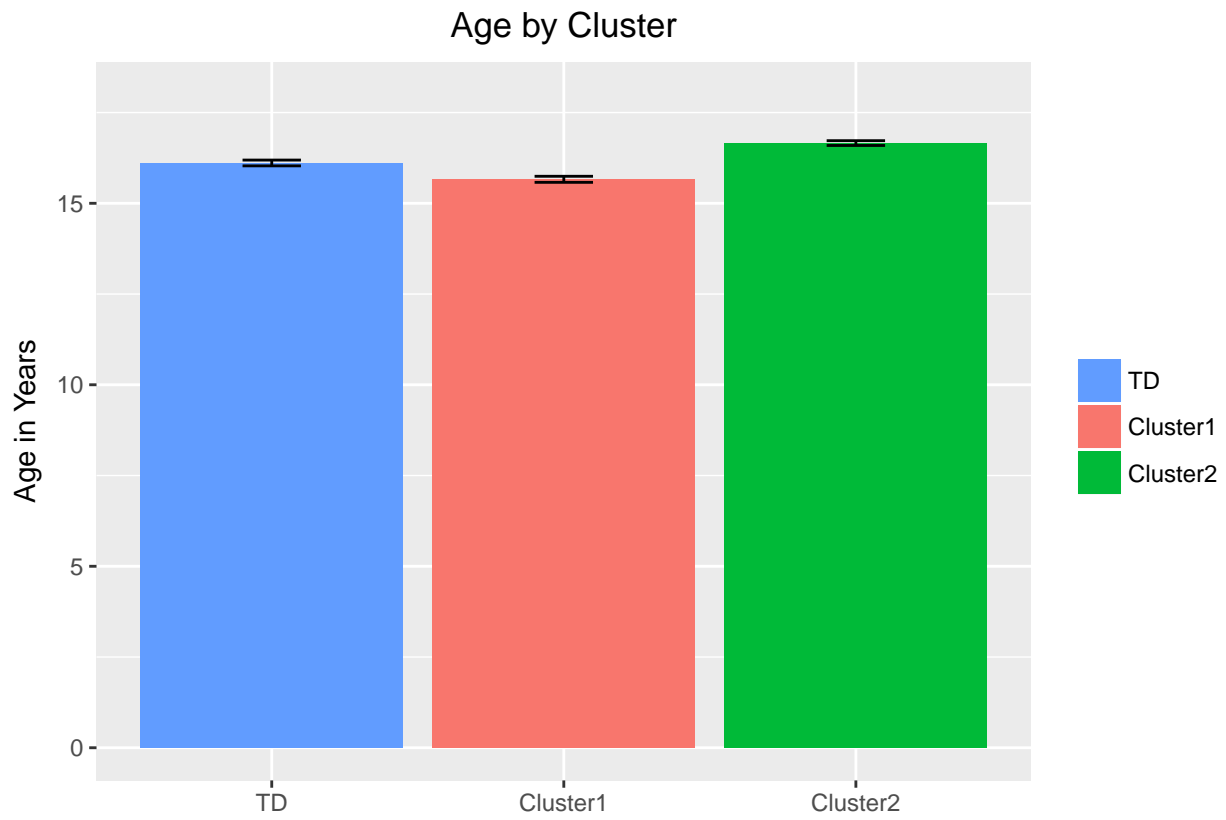
```
##                                  2               0 (  0.0)      0 (  0.0)
##                       Stratified by Cluster
##                                  2              p      test
##   n                            336
##   Race (%)                     238 ( 70.8)  <0.001
##                                 98 ( 29.2)
##   Sex (%)                      207 ( 61.6)   0.015
##                                129 ( 38.4)
##   Maternal Ed (mean (sd)) 14.53 (2.29)   <0.001
##   Age (mean (sd))         16.66 (2.49)   <0.001
##   Depression (%)               336 (100.0)  <0.001
##                                  0 (  0.0)
##   Cluster (%)                    0 (  0.0)  <0.001
##                                  0 (  0.0)
##                                336 (100.0)

##                       Stratified by Cluster
##                       level           -1              1
##   n                                  2297            376
##   Race (%)            Caucasian      1469 ( 64.0)    177 ( 47.1)
##                       Non-caucasian   828 ( 36.0)    199 ( 52.9)
##   Sex (%)             Female         1103 ( 48.0)    264 ( 70.2)
##                       Male           1194 ( 52.0)    112 ( 29.8)
##   Maternal Ed (mean (sd))          14.93 (2.45)   13.89 (2.24)
##   Age (mean (sd))                  13.83 (3.72)   16.27 (2.84)
##   Depression (%)      Depressed         0 (  0.0)    376 (100.0)
##                       Non-depressed  2297 (100.0)      0 (  0.0)
##   Cluster (%)         -1             2297 (100.0)      0 (  0.0)
##                       1                 0 (  0.0)    376 (100.0)
##                       2                 0 (  0.0)      0 (  0.0)
##                       Stratified by Cluster
##                                  2              p      test
##   n                            341
##   Race (%)                     218 ( 63.9)  <0.001
##                                123 ( 36.1)
##   Sex (%)                      217 ( 63.6)  <0.001
##                                124 ( 36.4)
##   Maternal Ed (mean (sd)) 14.37 (2.31)   <0.001
##   Age (mean (sd))         15.96 (2.95)   <0.001
##   Depression (%)               341 (100.0)  <0.001
##                                  0 (  0.0)
##   Cluster (%)                    0 (  0.0)  <0.001
##                                  0 (  0.0)
##                                341 (100.0)

##                       Stratified by Cluster
##                       level           -1              1
##   n                                  2297            346
##   Race (%)            Caucasian      1469 ( 64.0)    211 ( 61.0)
##                       Non-caucasian   828 ( 36.0)    135 ( 39.0)
##   Sex (%)             Female         1103 ( 48.0)    219 ( 63.3)
##                       Male           1194 ( 52.0)    127 ( 36.7)
##   Maternal Ed (mean (sd))          14.93 (2.45)   14.34 (2.31)
##   Age (mean (sd))                  13.83 (3.72)   16.02 (3.03)
##   Depression (%)      Depressed         0 (  0.0)    346 (100.0)
```
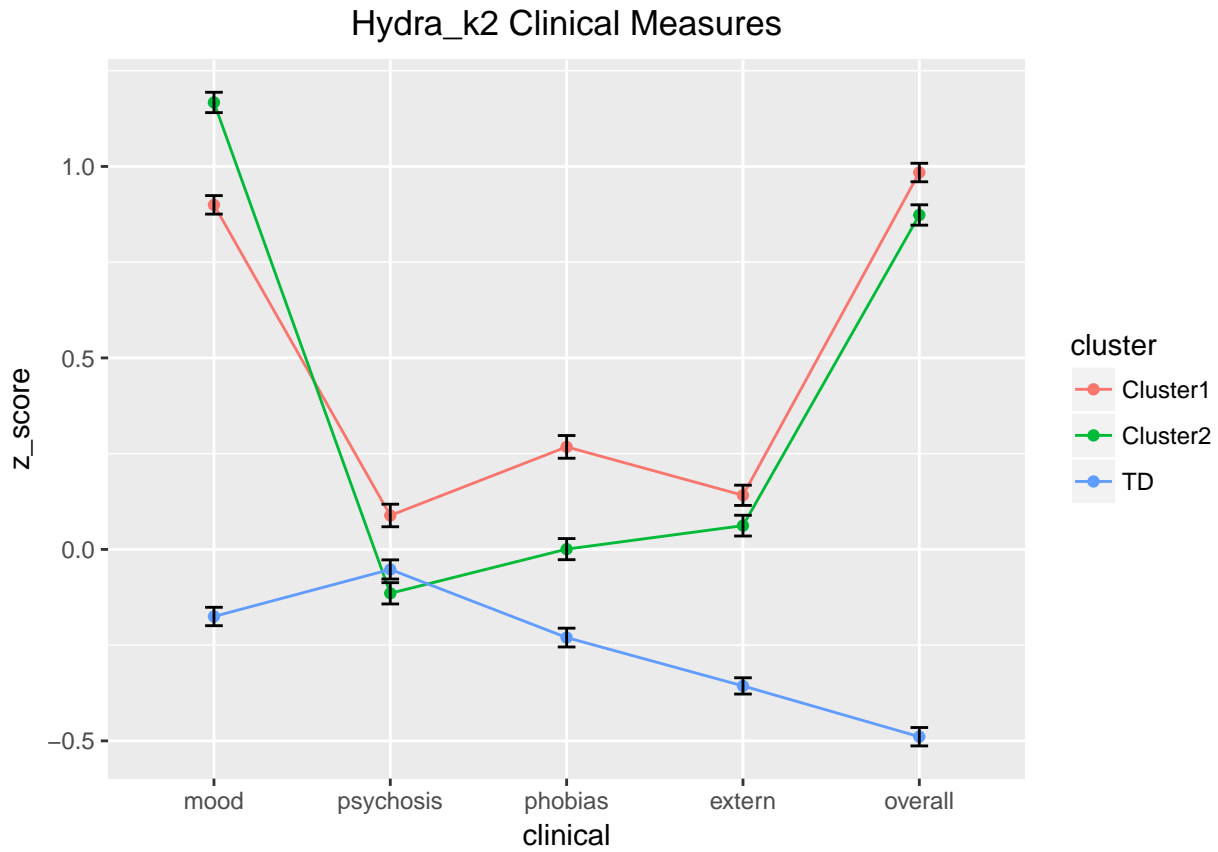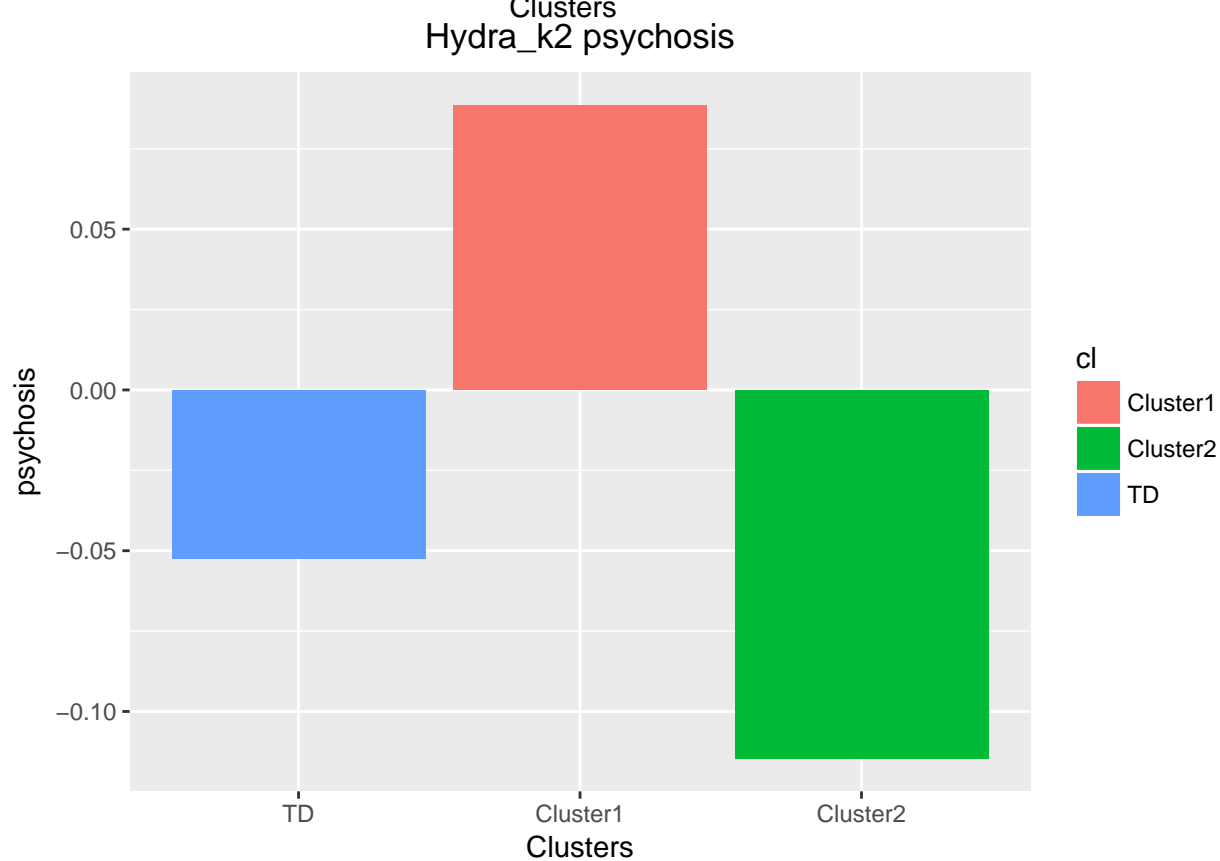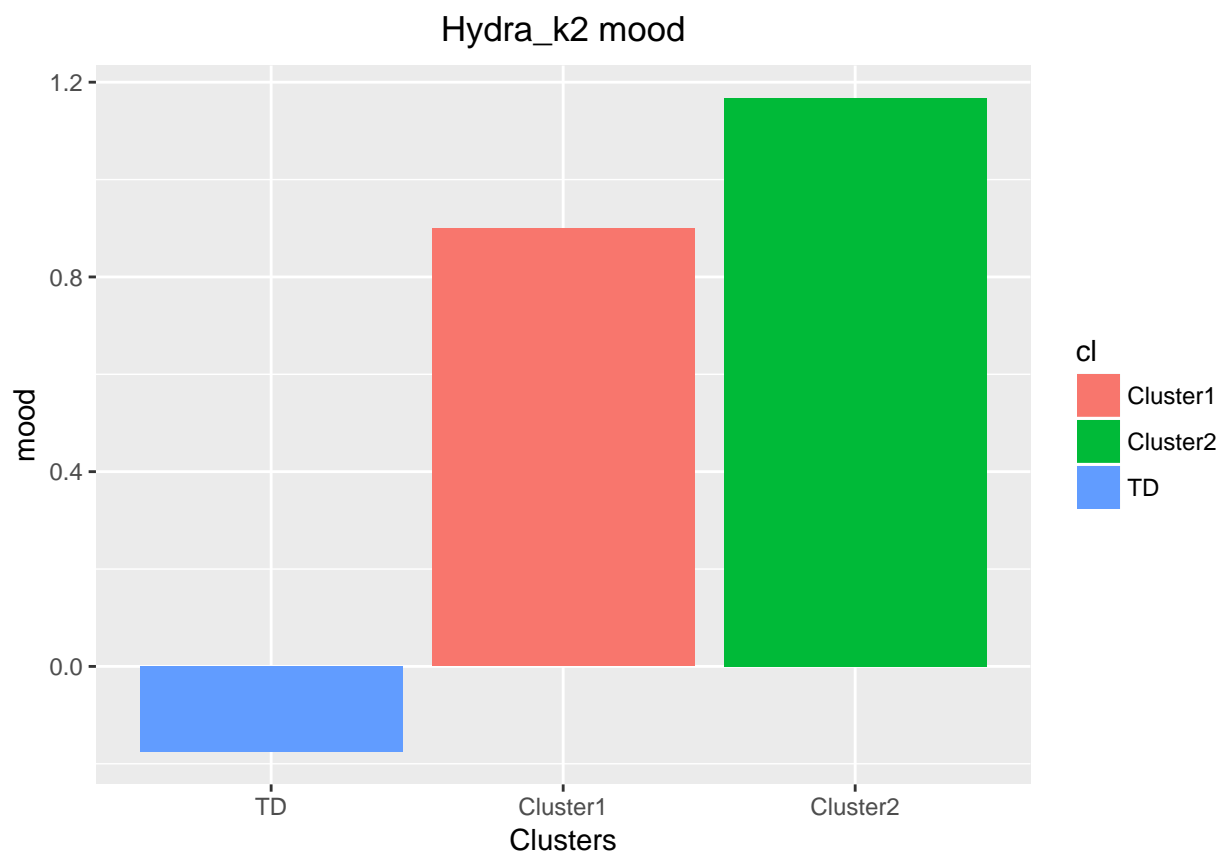
```
##                         Non-depressed  2297 (100.0)       0 (   0.0)
##    Cluster (%)          -1             2297 (100.0)       0 (   0.0)
##                         1                 0 (   0.0)    346 (100.0)
##                         2                 0 (   0.0)      0 (   0.0)
##                         Stratified by Cluster
##                         2              p       test
##    n                       371
##    Race (%)                184 ( 49.6)  <0.001
##                            187 ( 50.4)
##    Sex (%)                 262 ( 70.6)  <0.001
##                            109 ( 29.4)
##    Maternal Ed (mean (sd)) 13.92 (2.25)  <0.001
##    Age (mean (sd))         16.22 (2.77)  <0.001
##    Depression (%)          371 (100.0)  <0.001
##                              0 (   0.0)
##    Cluster (%)               0 (   0.0)  <0.001
##                              0 (   0.0)
##                            371 (100.0)
```

Part 4: Including Graphs

Age by Cluster

Maternal Edu by Cluster

Part 5-8: Stats: LM with visreg, anova, fdr correction
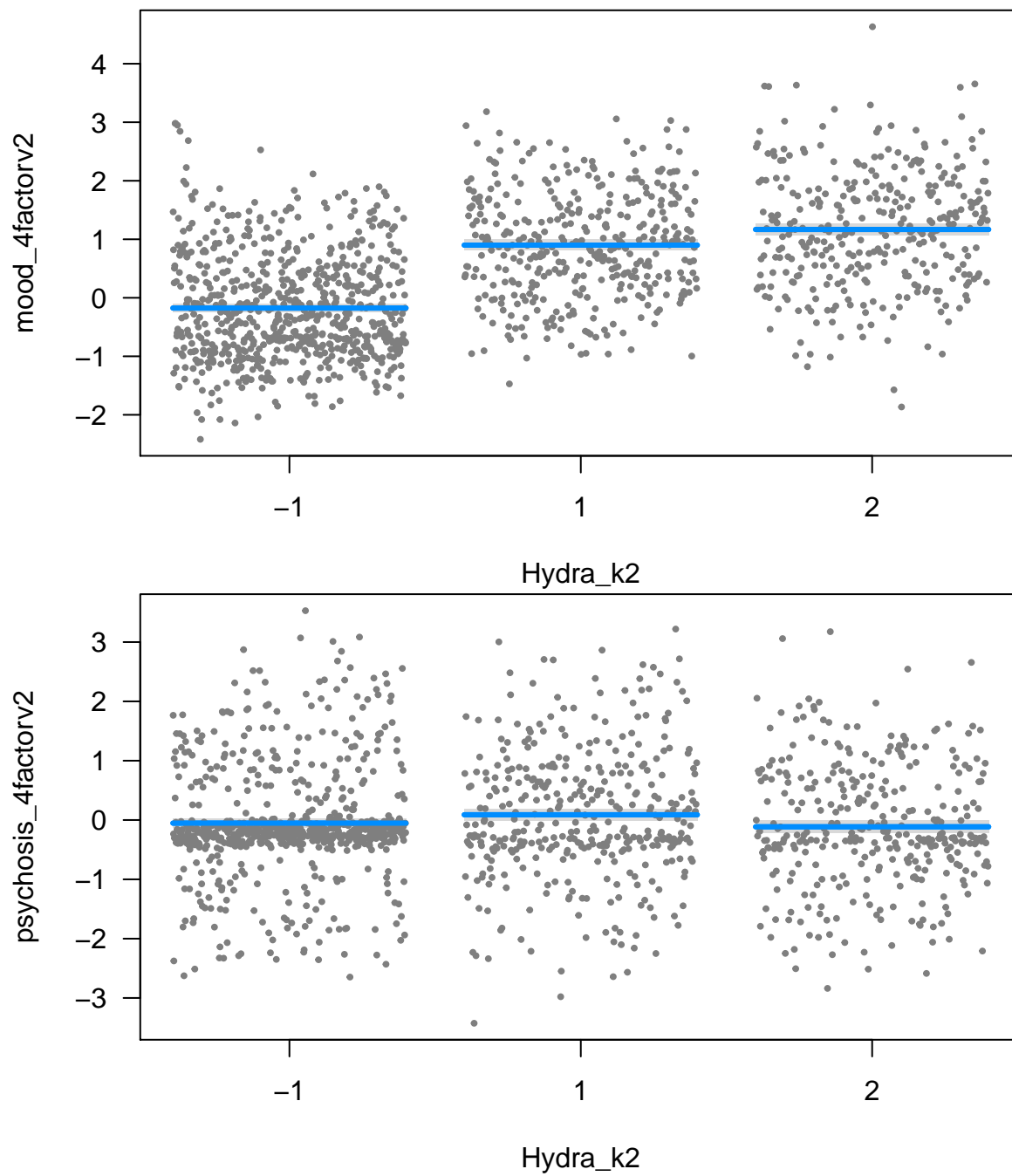
Using cl as id variables

Hydra_k2 Clinical Measures

Hydra_k2 mood

Hydra_k2 psychosis

Hydra_k2 phobias

Hydra_k2 externalizing

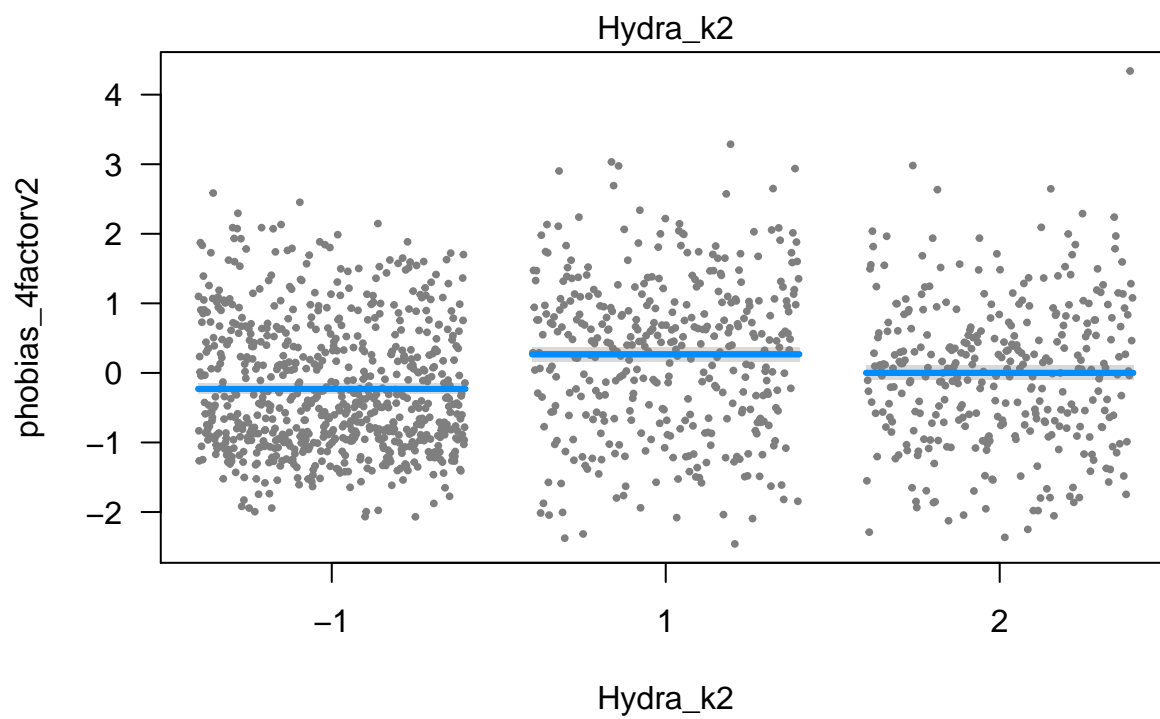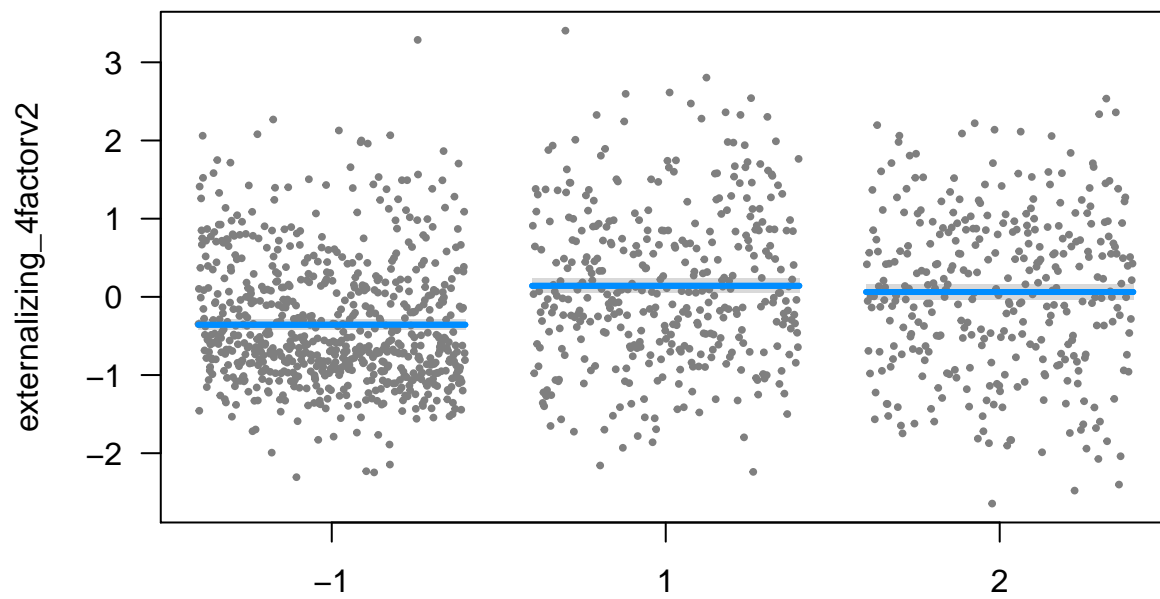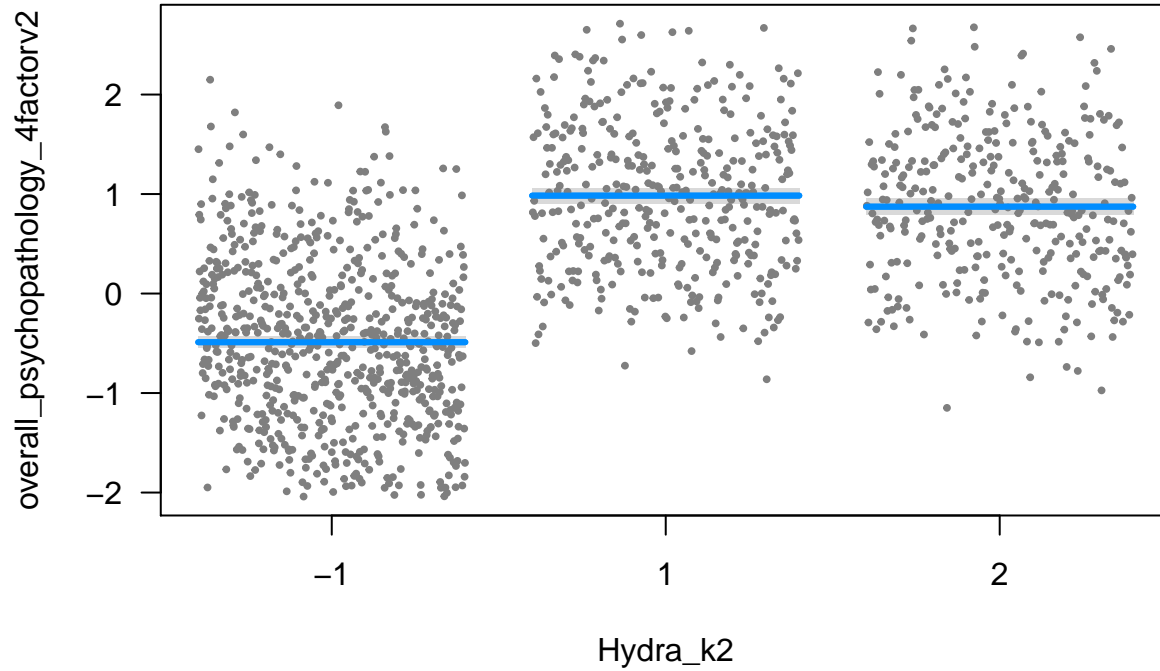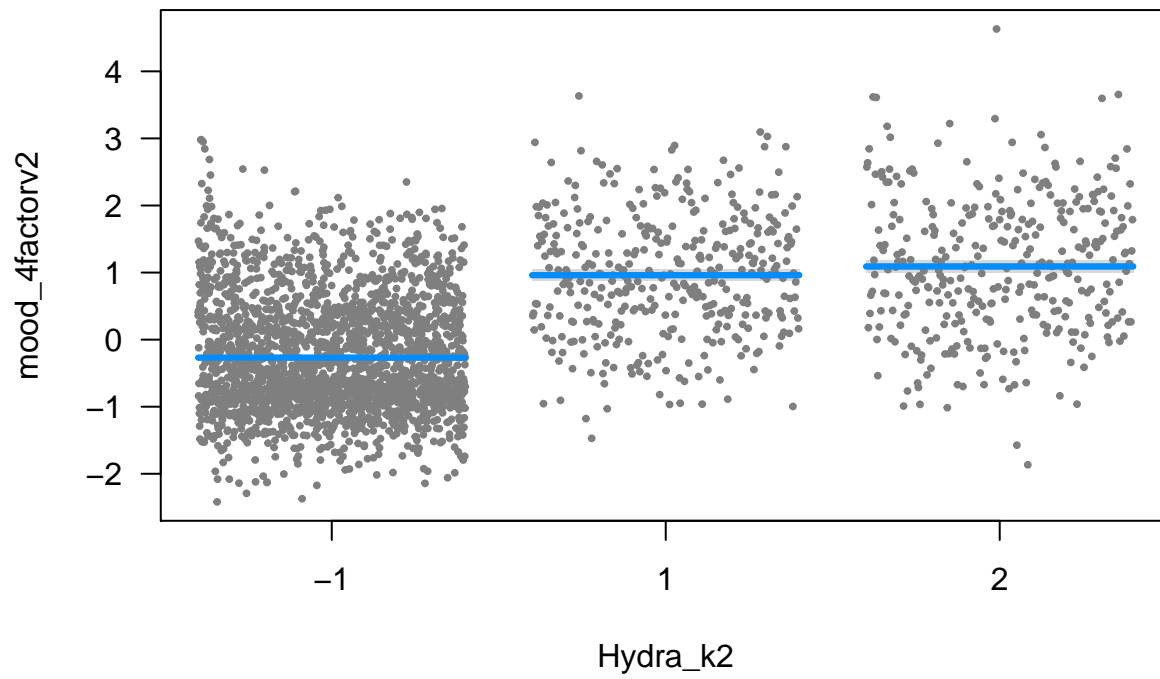## Hydra_k2 overall
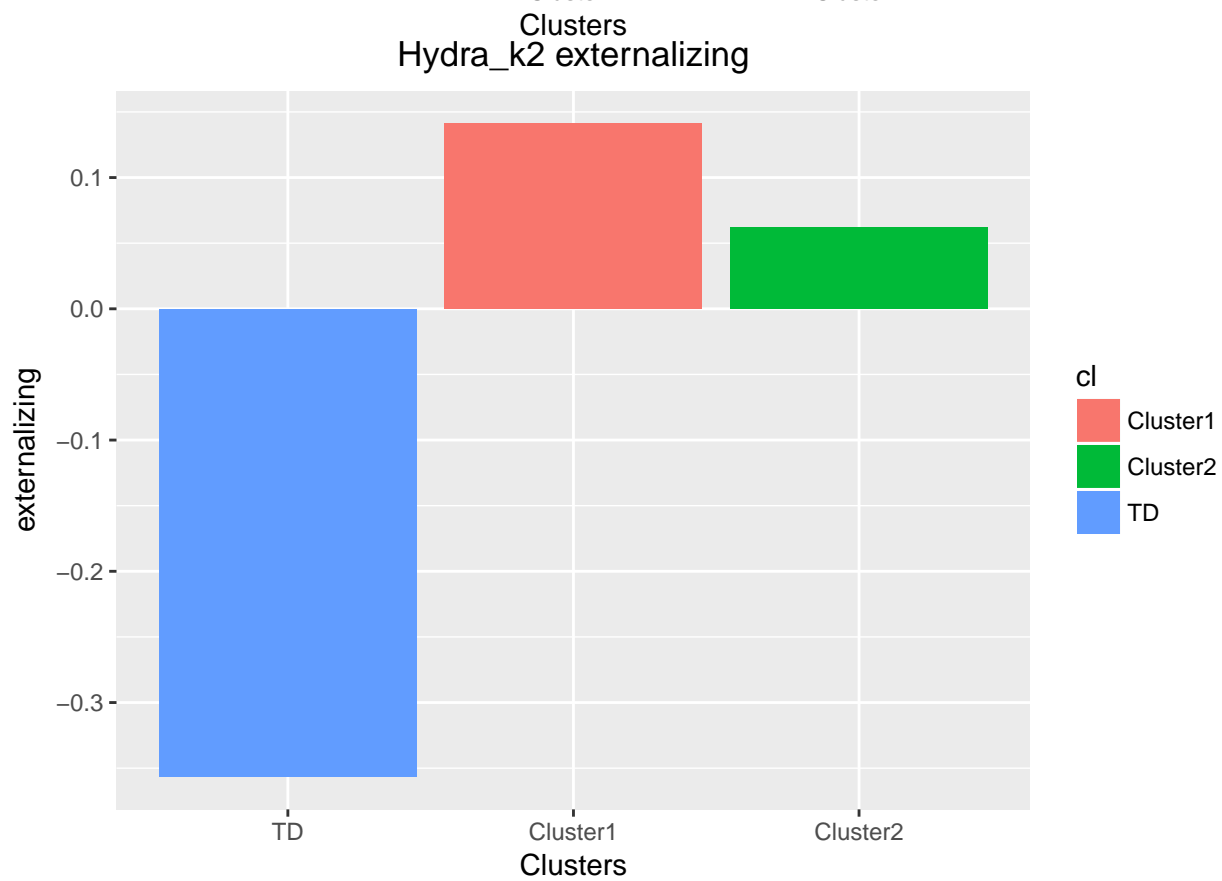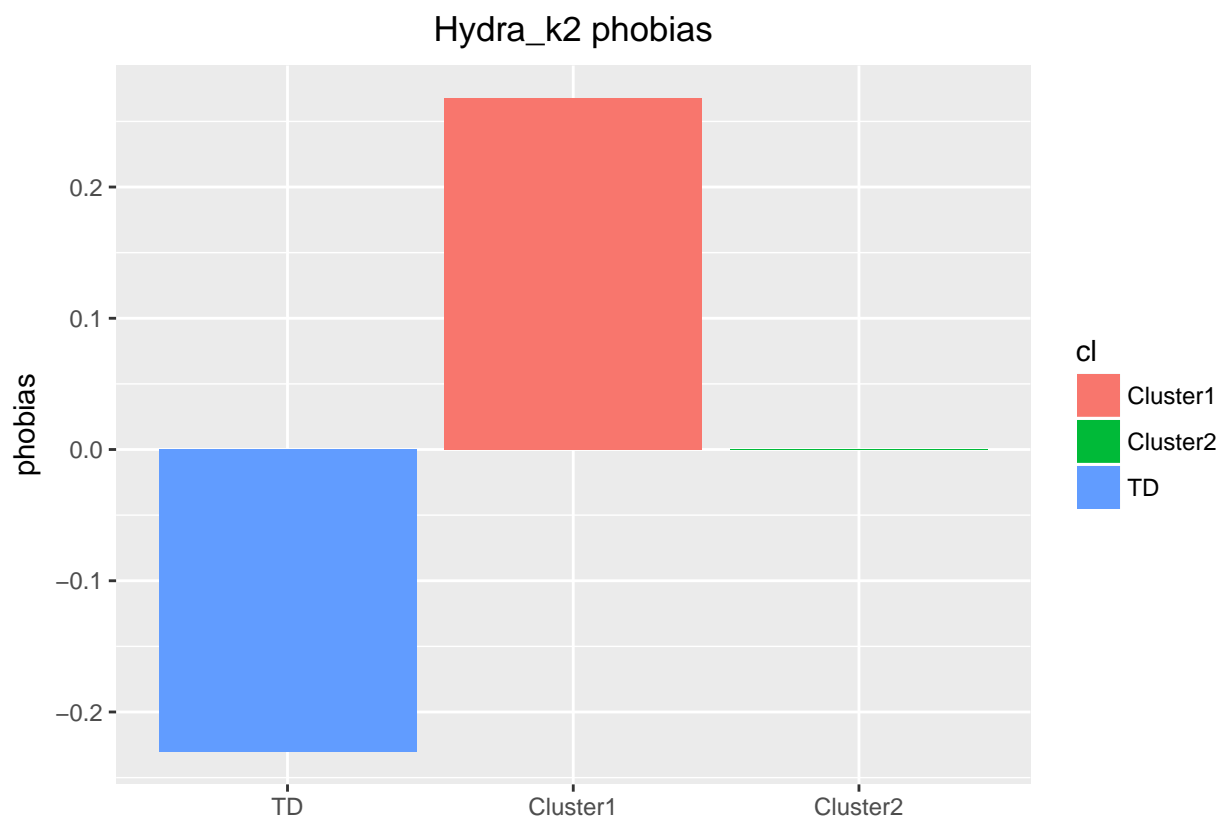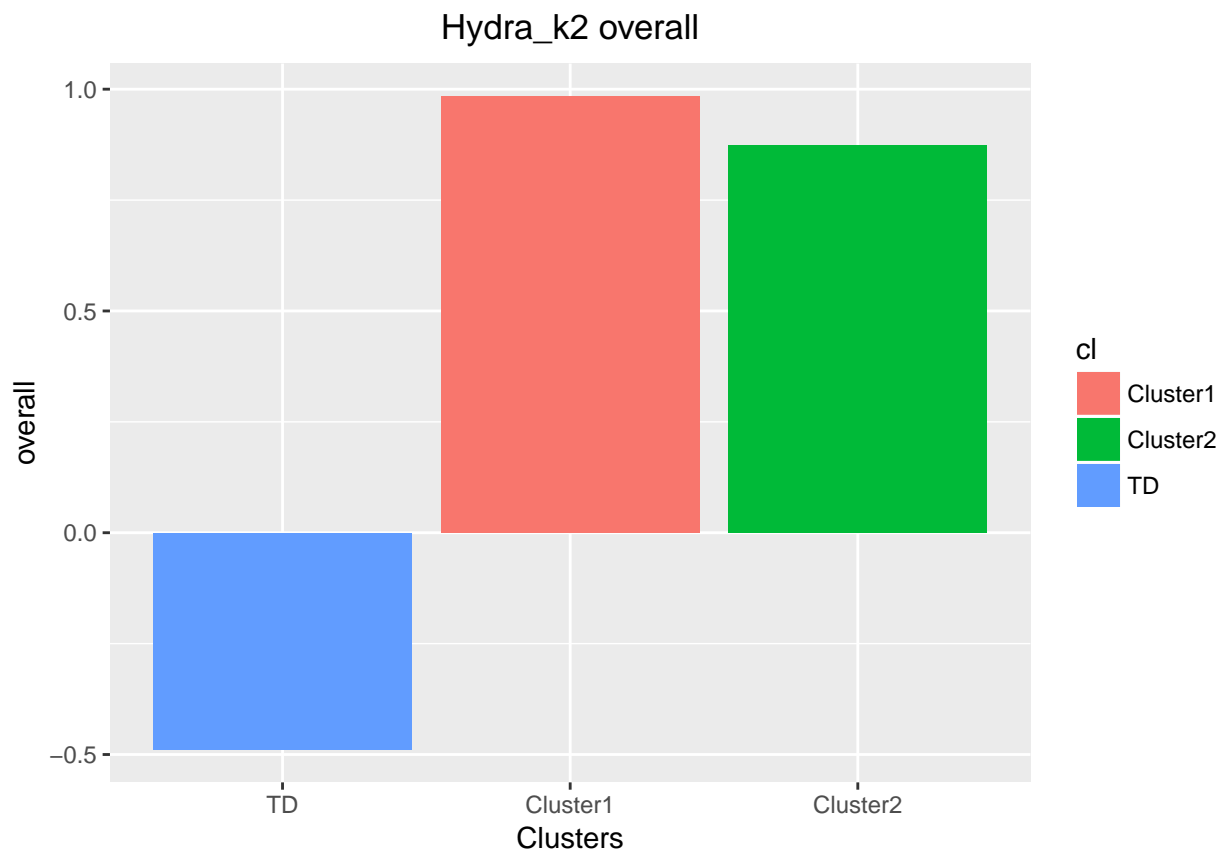

```
##                      clinical_measure p_FDR_corr
## 1                      mood_4factorv2          0
## 2                 psychosis_4factorv2       0.02
## 3             externalizing_4factorv2          0
## 4                   phobias_4factorv2          0
## 5 overall_psychopathology_4factorv2             0

## [1] "MATCHED"
```
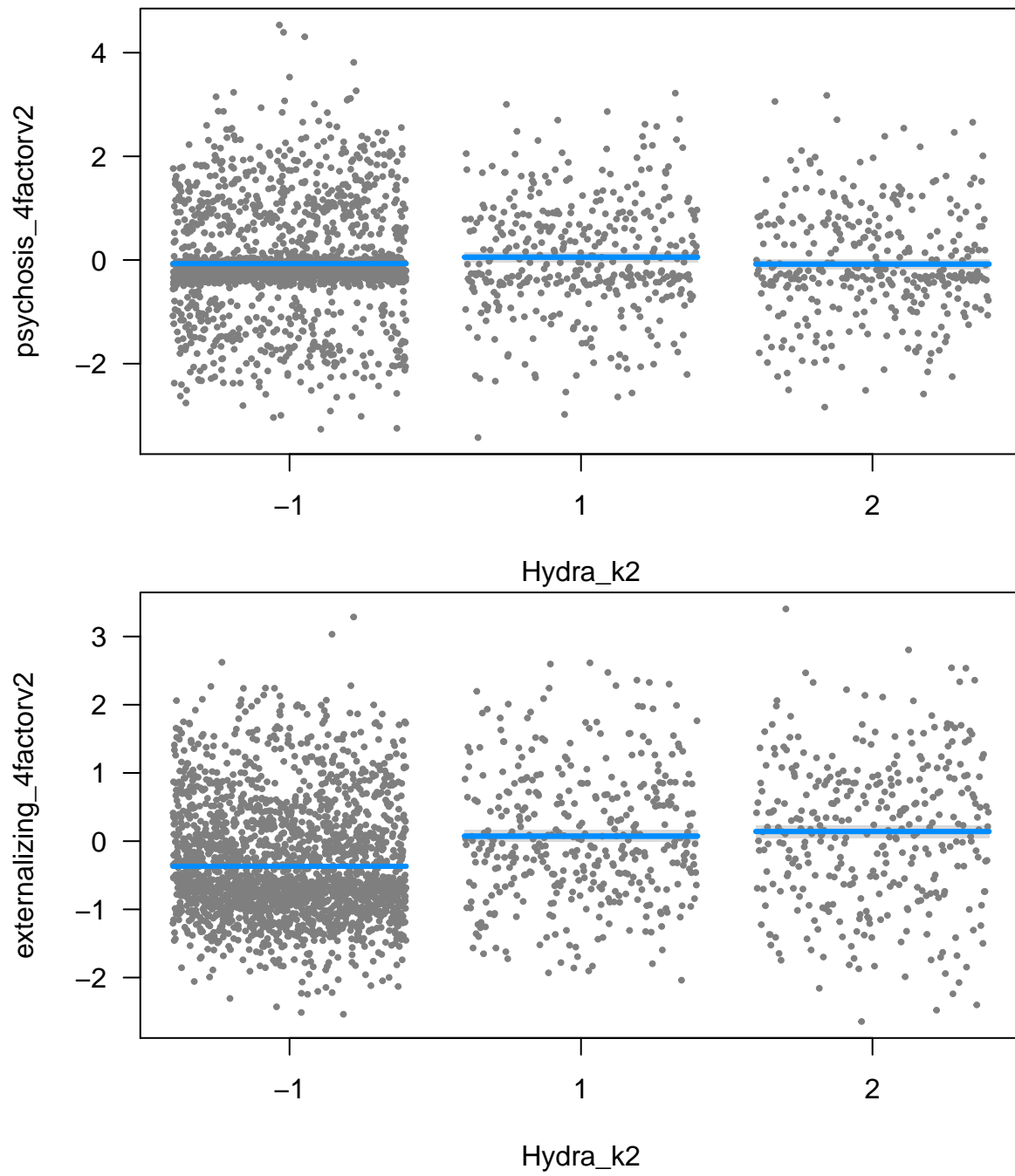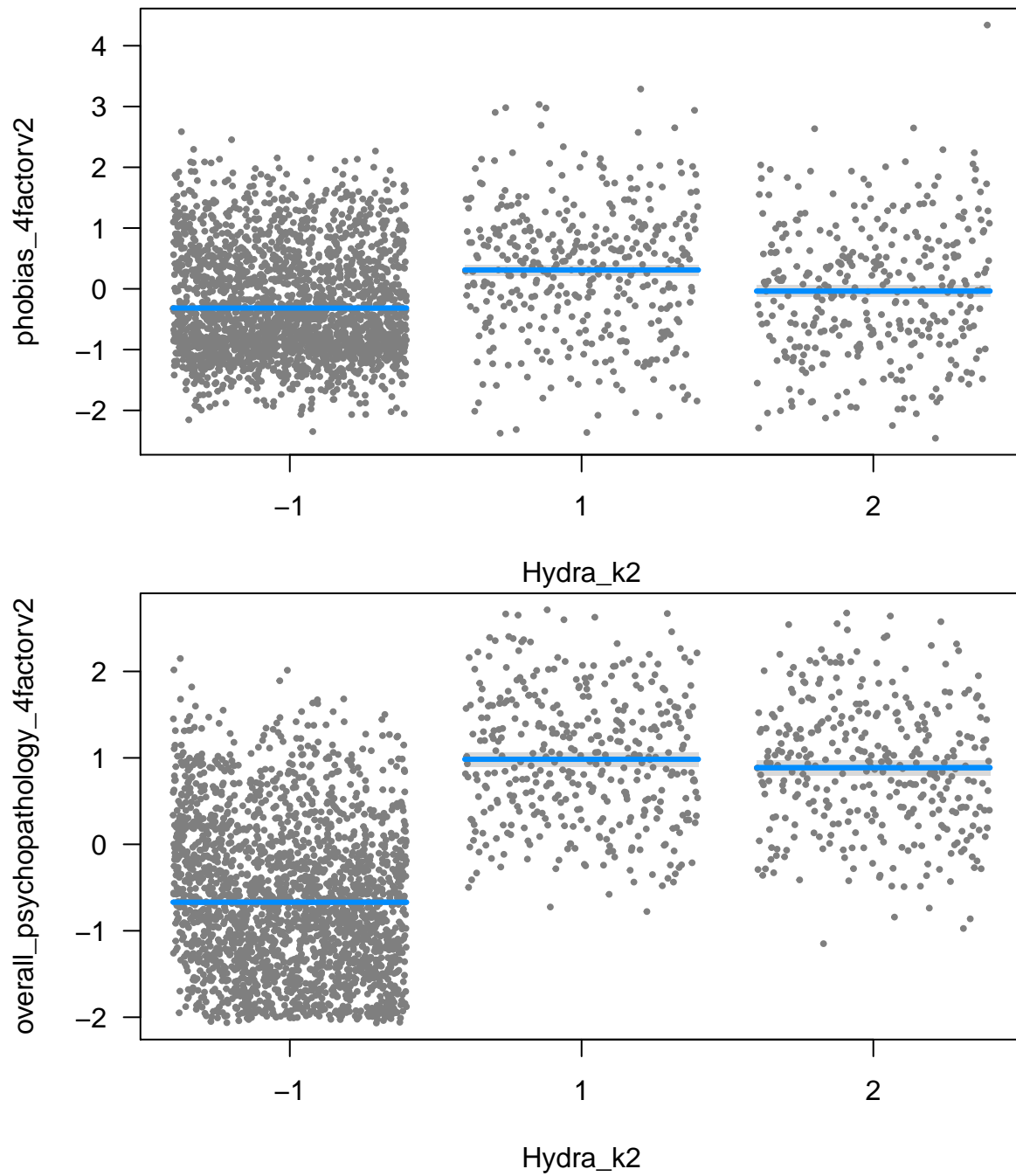
[1] "UNMATCHED"

12

[1] "RESIDUALIZED unmatched"