Ribbon Synapse Markers

Alex Scharr

5/3/2018

One of the projects in our lab is looking at the ribbon synapses of hair cells, the cells of the inner ear that detect sound. These synapses have a specialized structure and are able to release neurotransmitters at high rates for sustained periods of time. This is important because each hair cell must always be able to communicate to the central nervous system that it has detected a sound, and it must also somehow communicate how loud that sound is. One way that it can communicate the loudness of a sound is by having different types of nerves connected to it. Some nerves "fire" only when they receive a large sustained signal, and some may "fire" at the slightest change. It has been observed in other species, like cats, that these nerves synapse onto the hair cells at different locations depending on whether or not they fire easily or not.

In this experiment, my labmate dissected the hair cells of mice, and used fluorescent tags (or markers) to mark the location of both ribbon synapses and the nerves connected to the hair cells. We then needed to explore the data to see if there were any patterns in the characteristics of either the ribbons, the nerves, or their particular combinations, that might correspond to the sensitive vs less sensitive synapses.

The data my labmate provided was the X, Y, and Z locations of either the marker for the ribbon synapse, CTBP2, and the post-synaptic nerve, GLUR, from multiple hair cells. At this stage, the markers were not separated by hair cell, but simply aggregated. Ribbon and synapse pairs were identified by locating the markers closest to each other. Other data included the volumes of each of the markers and/or pairs.

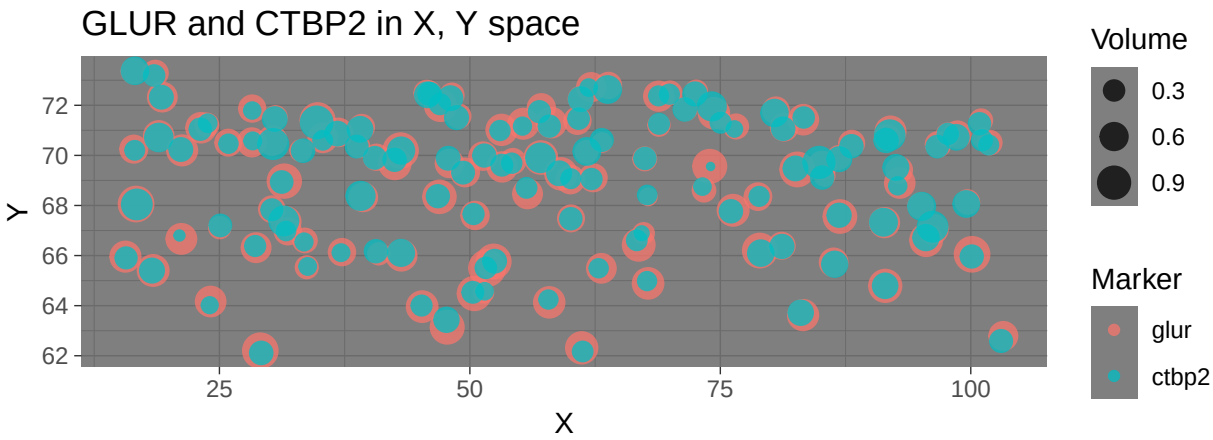
Read in data

First I read in the data and add some columns such as the ratio of the volumns between GLUR and CTBP2, as well as categorizing the locations of the markers as "high" or "low," and the volumes as "big" or "small."

Visualization of markers in space

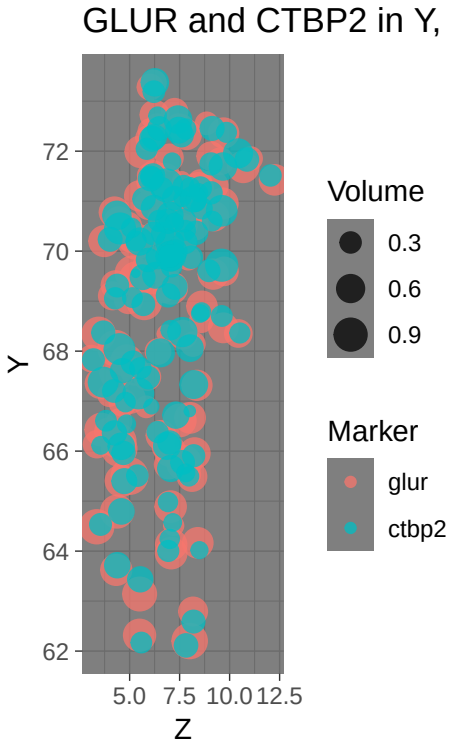
Here are the locations and the sizes of the markers, GLUR and CTBP2 in X and Y. These cover about ten cells in one horizontal row, but I did not have the data indicating which markers belonged to which cell, so they are presented here in aggregate. The plot covers about ten cells "wide" and one cell "deep."

```
p = ggplot(data = markers)+
  geom_point(aes(x = X,
                 y = Y,
                 color = Marker,
                 size = Volume),
             alpha = 0.75)+
  theme_dark()+
  coord_fixed(ratio = 2.5)+
  ggtitle("GLUR and CTBP2 in X, Y space")
p
```

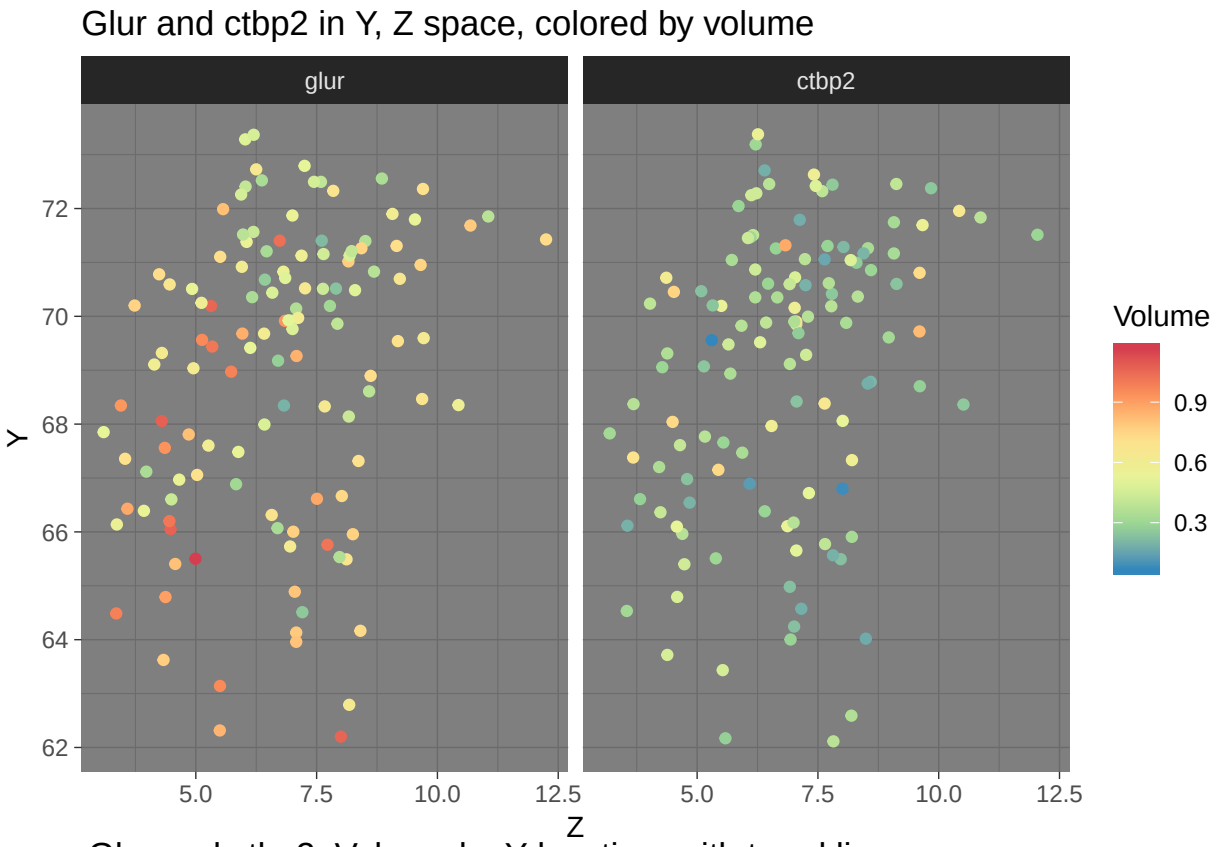


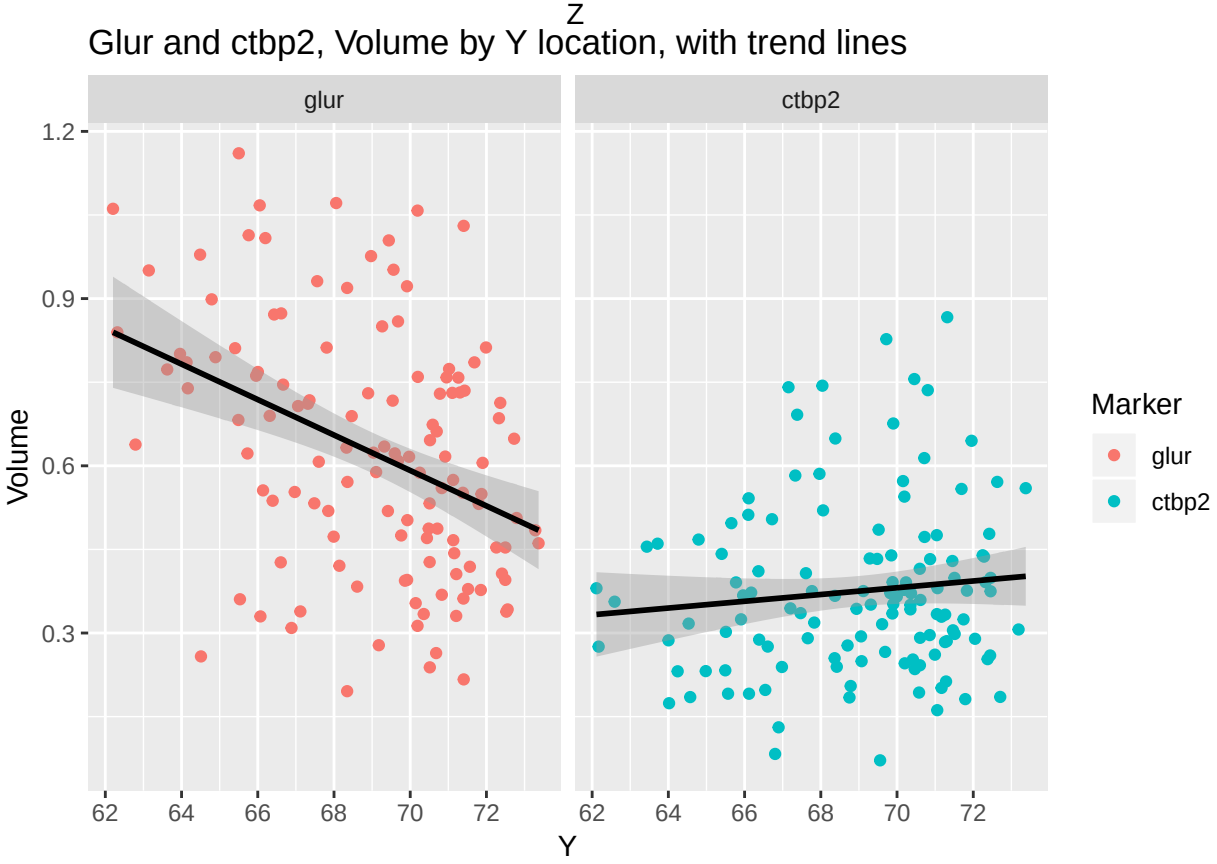
Here are the locations and the sizes of the markers, GLUR and CTBP2 in Z and Y. These cover about ten cells in one horizontal row, but now as viewed from the side, looking "down the row." So plot is one cell "wide" and ten cells "deep."

```
p = ggplot(data = markers)+
  geom_point(aes(x = Z,
                 y = Y,
                 color = Marker,
                 size = Volume),
             alpha = 0.75)+
  theme_dark()+
  coord_fixed(ratio = 2.5)+
  ggtitle("GLUR and CTBP2 in Y, Z space")
p
```



Exploration of volume:





Glur trends smaller when it is higher in Z. There is no strong trend for ctbp2.

Statistics on linear regression of glur volume as a function of Y

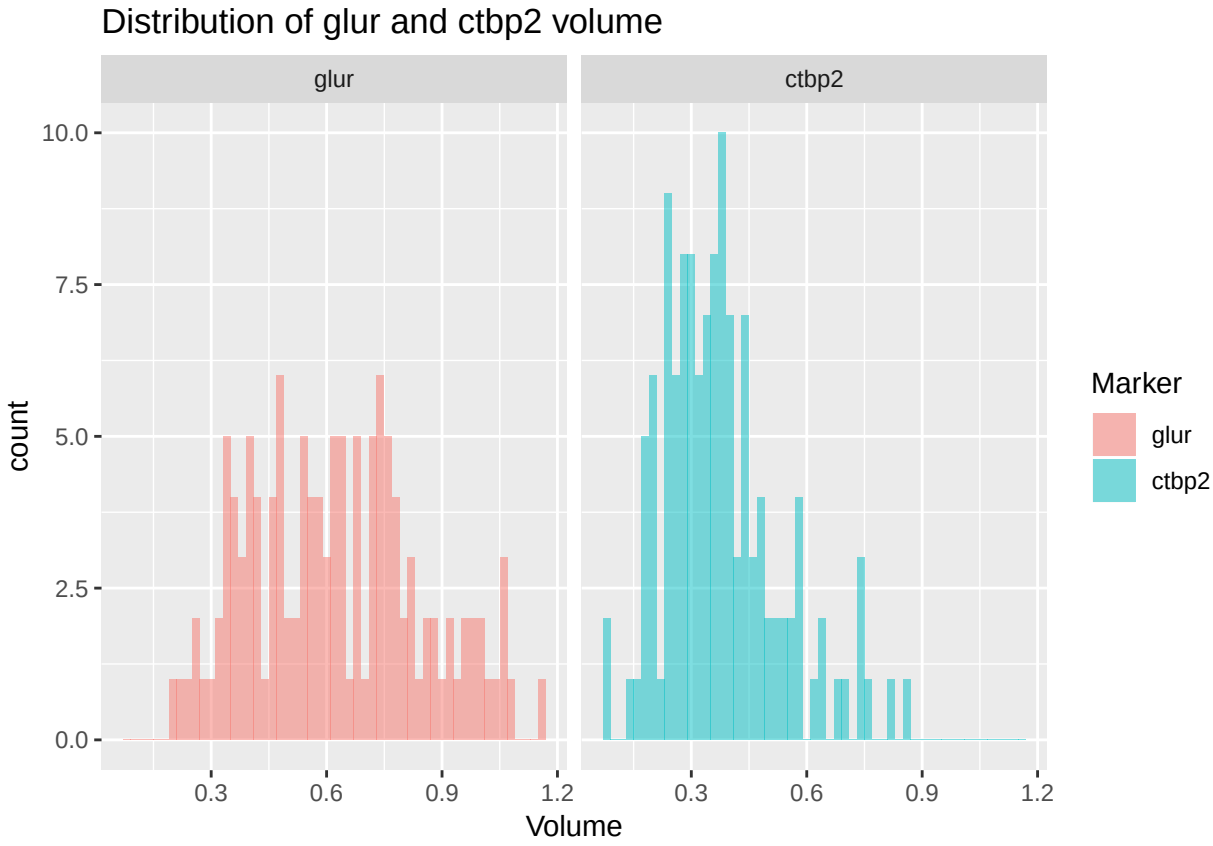
There is a statistically significant relationship between glur volume and Y position.

```
##
## Call:
## lm(formula = Volume ~ Y, data = markers[which(markers$Marker ==
##     "glur"), c("Volume", "Y")])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50820 -0.14095  0.00349  0.15076  0.48355
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.817550   0.476770   5.910 3.14e-08 ***
## Y           -0.031799   0.006905  -4.605 1.01e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2056 on 123 degrees of freedom
## Multiple R-squared:  0.1471, Adjusted R-squared:  0.1401
## F-statistic: 21.21 on 1 and 123 DF,  p-value: 1.013e-05
```

Statistics on linear regression of ctbp2 volume as a function of Y

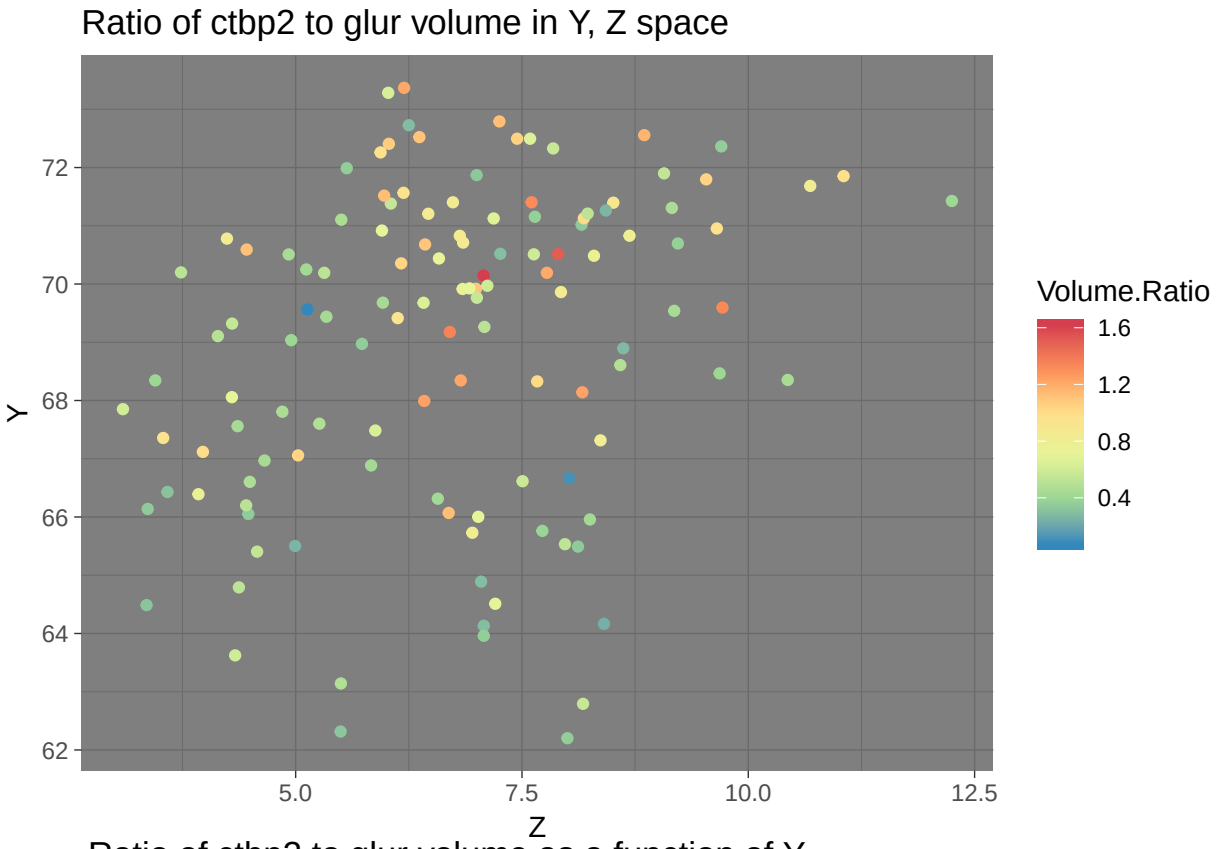
There is NOT a statistically significant relationship between ctbp2 volume and Y position.

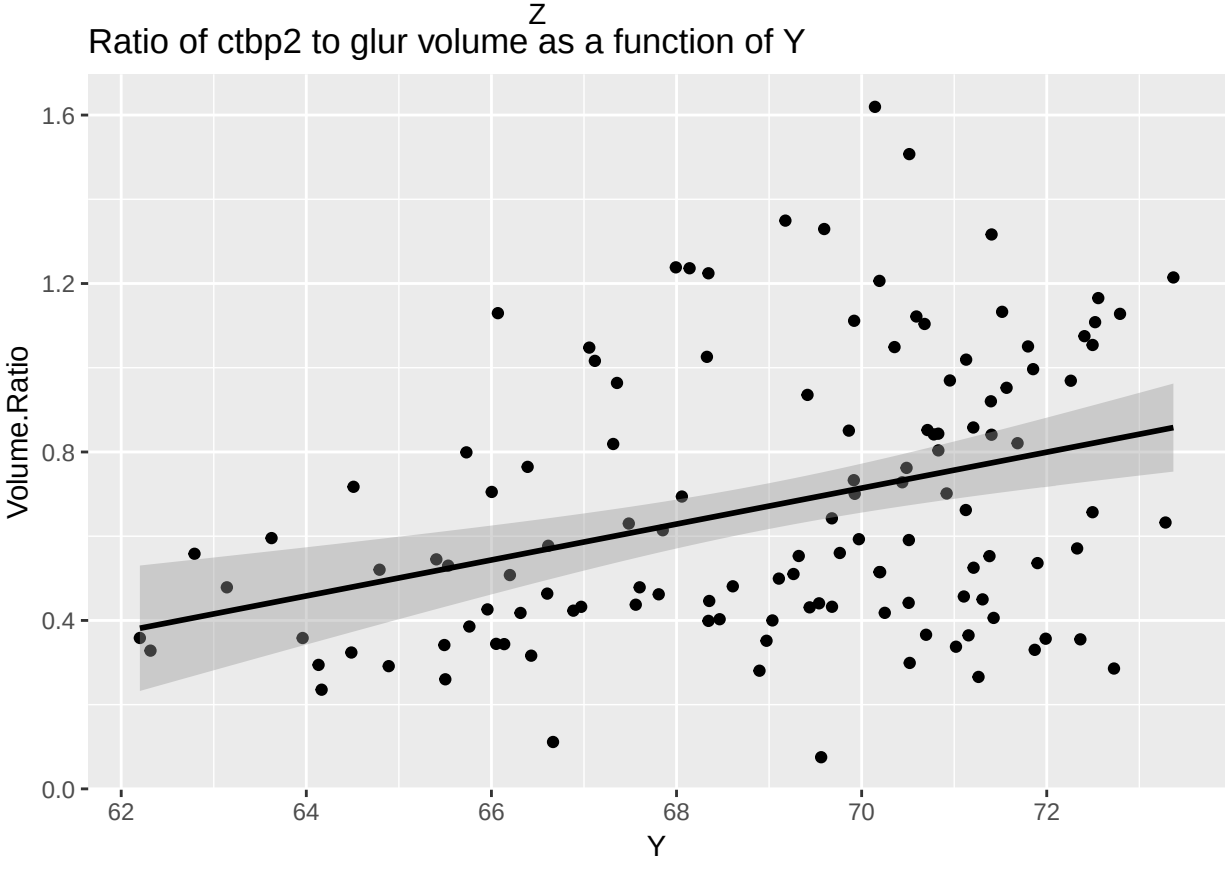
```
##
## Call:
## lm(formula = Volume ~ Y, data = markers[which(markers$Marker ==
##     "ctbp2"), c("Volume", "Y")])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30697 -0.10512 -0.02562  0.05897  0.47741
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.043749   0.358739  -0.122    0.903
## Y            0.006071   0.005196   1.168    0.245
##
## Residual standard error: 0.1535 on 123 degrees of freedom
## Multiple R-squared:  0.01097,    Adjusted R-squared:  0.002934
## F-statistic: 1.365 on 1 and 123 DF,  p-value: 0.245
```



Volumes of glur and ctbp2 are not distributed the same. Glur has a larger more even range of volumes, whereas ctbp2 is generally smaller.

Exploring volume ratio





The ratio of ctbp2:glur volume trends upward as Y increases.

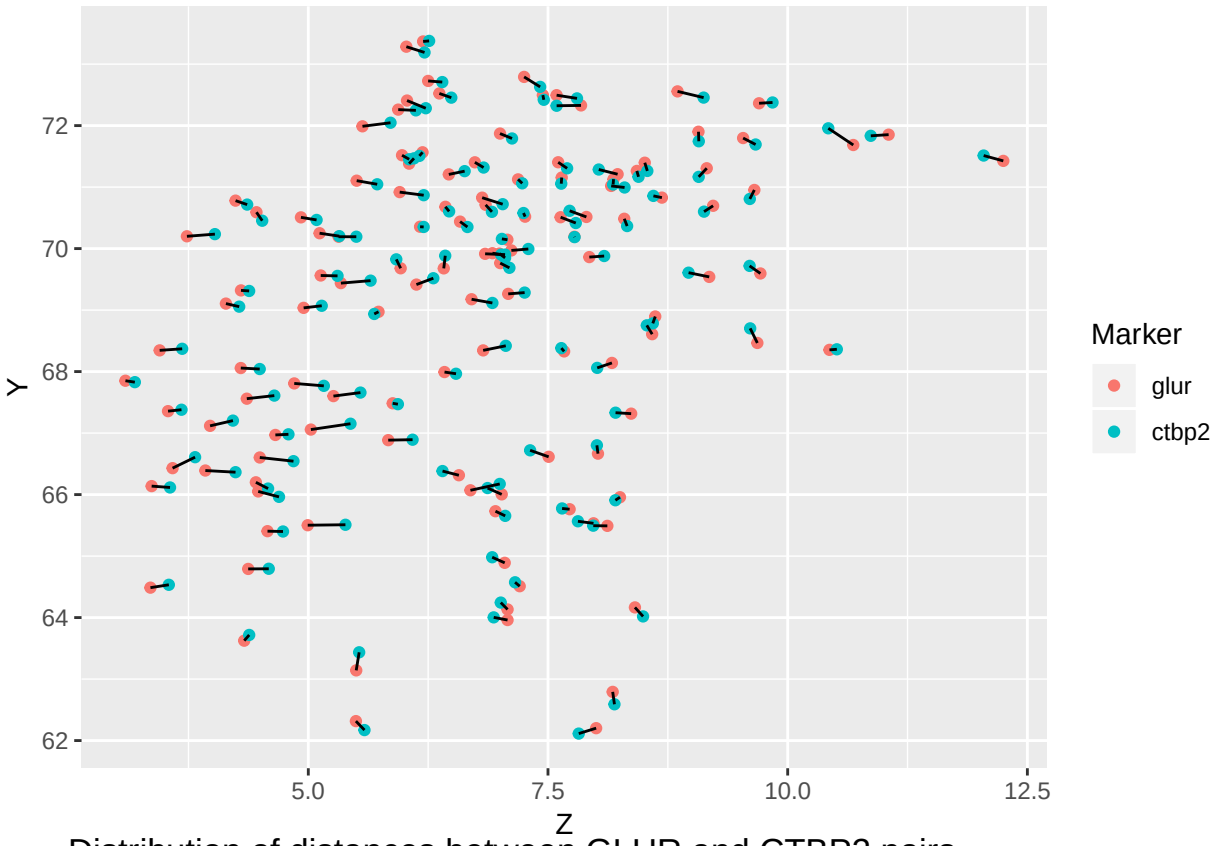
Statistics on linear regression of volume ratio as a function of Y

There is a statistically significant relationship between volume ratio and Y position.

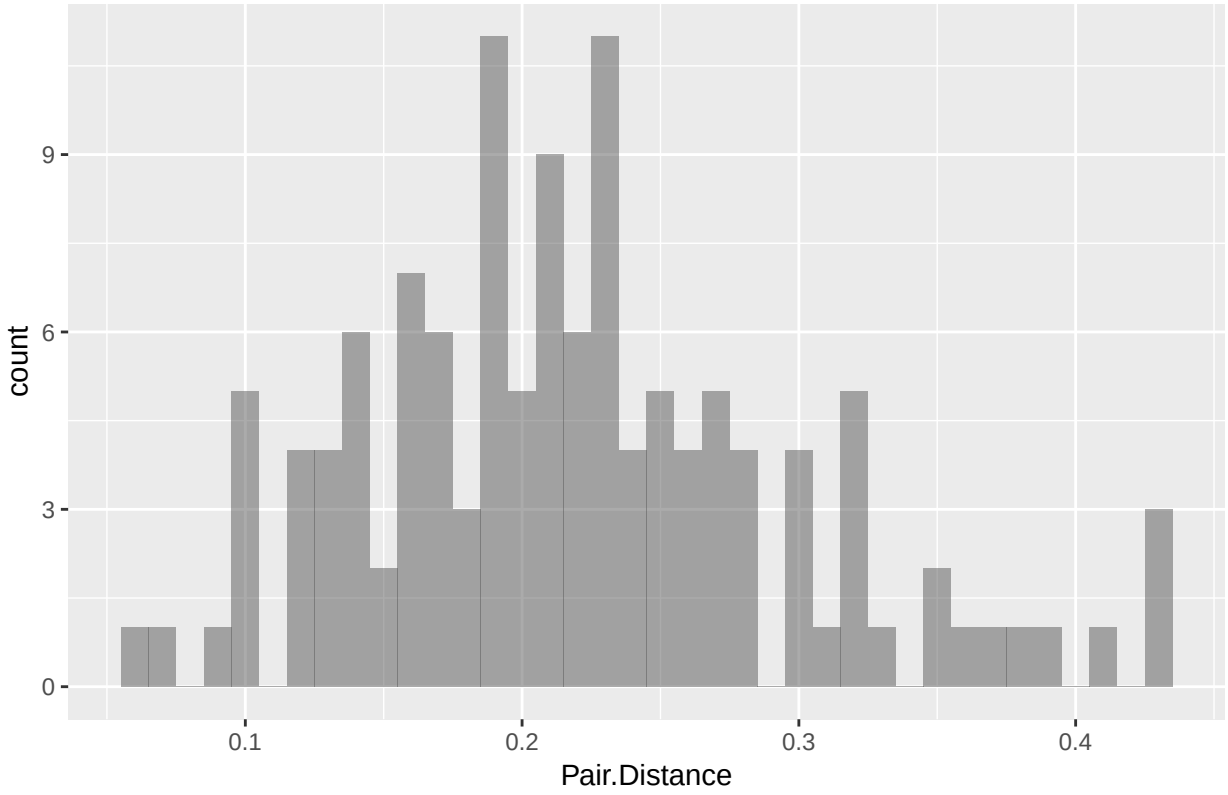
```
##
## Call:
## lm(formula = Volume.Ratio ~ Y, data = markers[which(markers$Marker ==
##     "glur"), c("Volume.Ratio", "Y")])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.62036 -0.22044 -0.05187  0.21493  0.89919
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.27279    0.71203  -3.192  0.00179 **
## Y            0.04267    0.01031   4.138 6.45e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.307 on 123 degrees of freedom
## Multiple R-squared:  0.1222, Adjusted R-squared:  0.115
## F-statistic: 17.12 on 1 and 123 DF,  p-value: 6.452e-05
```

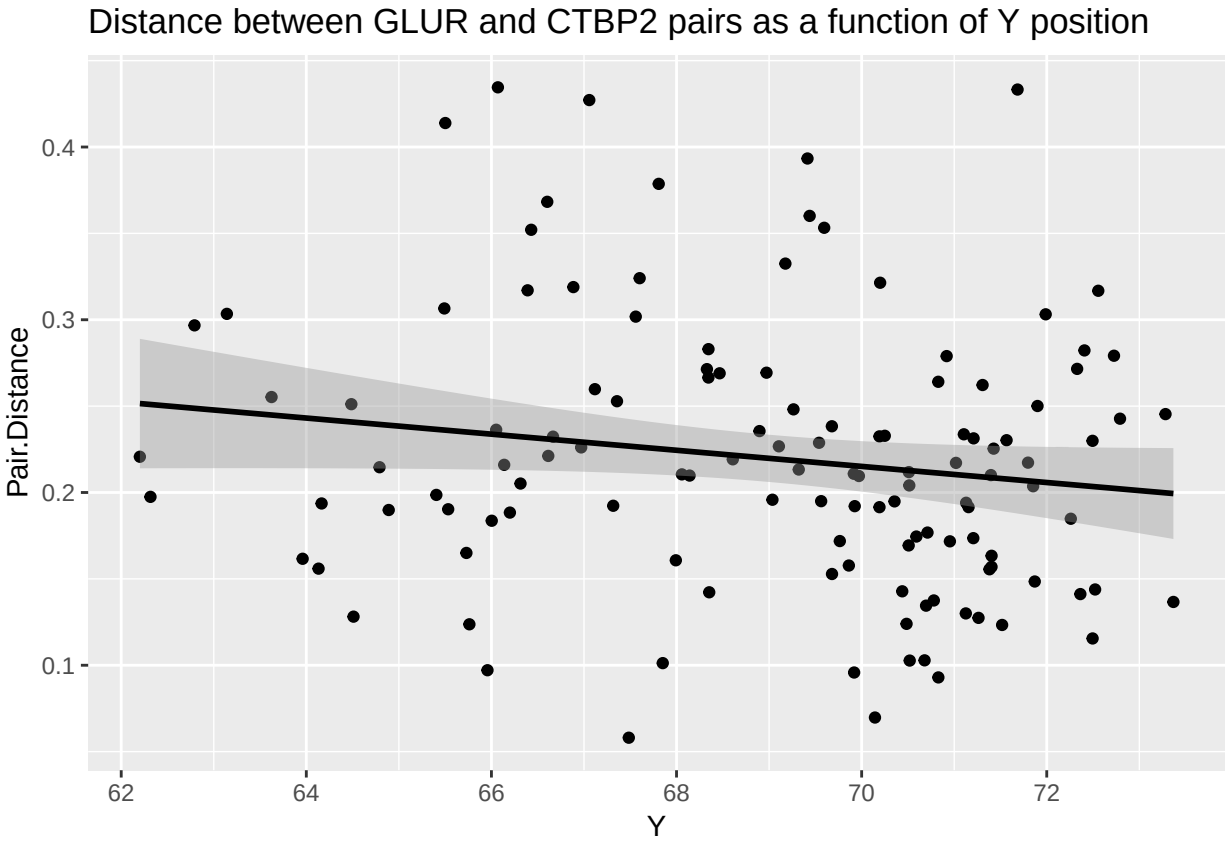
Exploration of distance between glur and ctbp2 centers

```
## $title
## [1] "Z, Y centroid locations of GLUR and CTBP2 pairs"
##
## $subtitle
## NULL
##
## attr(,"class")
## [1] "labels"
```



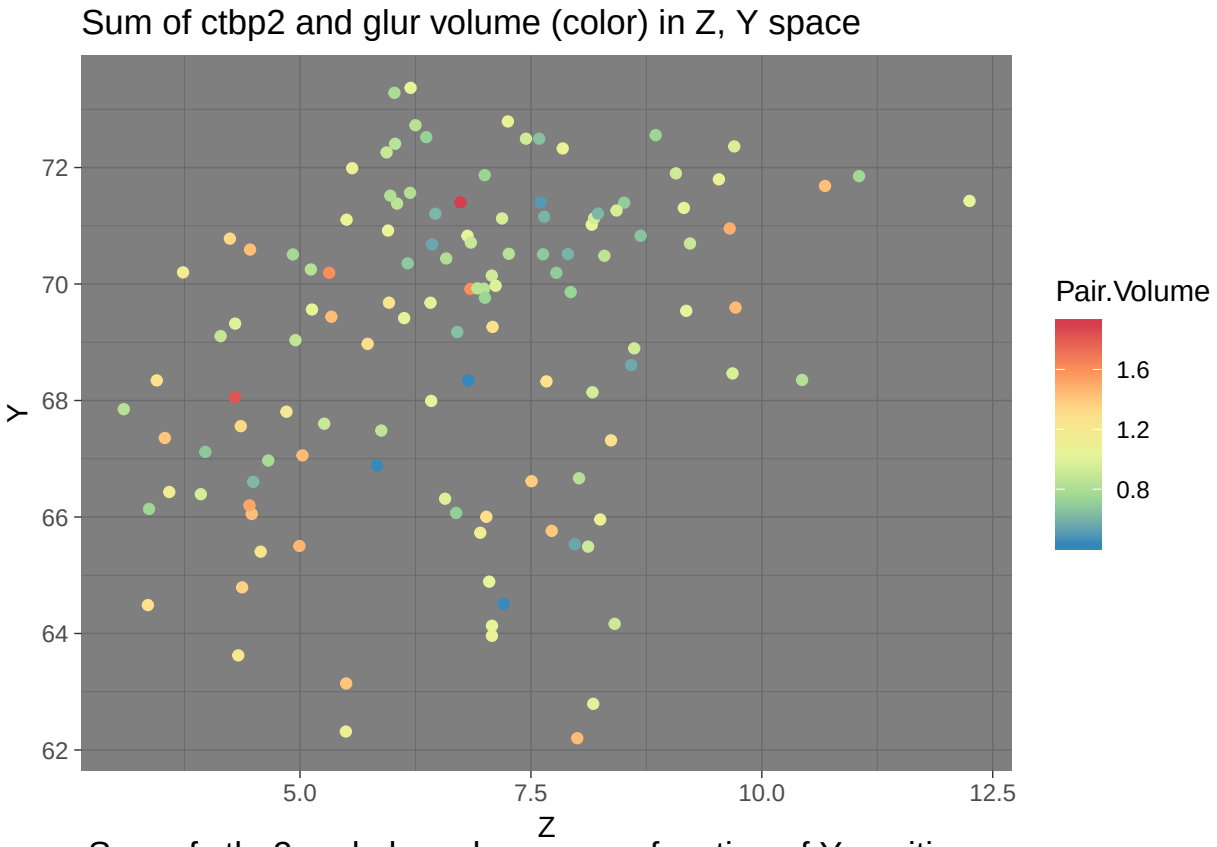
Distribution of distances between GLUR and CTBP2 pairs





The distance between the centroids of glur and ctbp2 does not seem to change with changing Y location (or Z or X, not shown).

Exploring sum of volumes of glur and ctbp2



The sum of the volumes of glur and ctbp2 pairs does not change much, but may trend downward with increasing Y.

Statistics on linear regression of glur volume as a function of Y

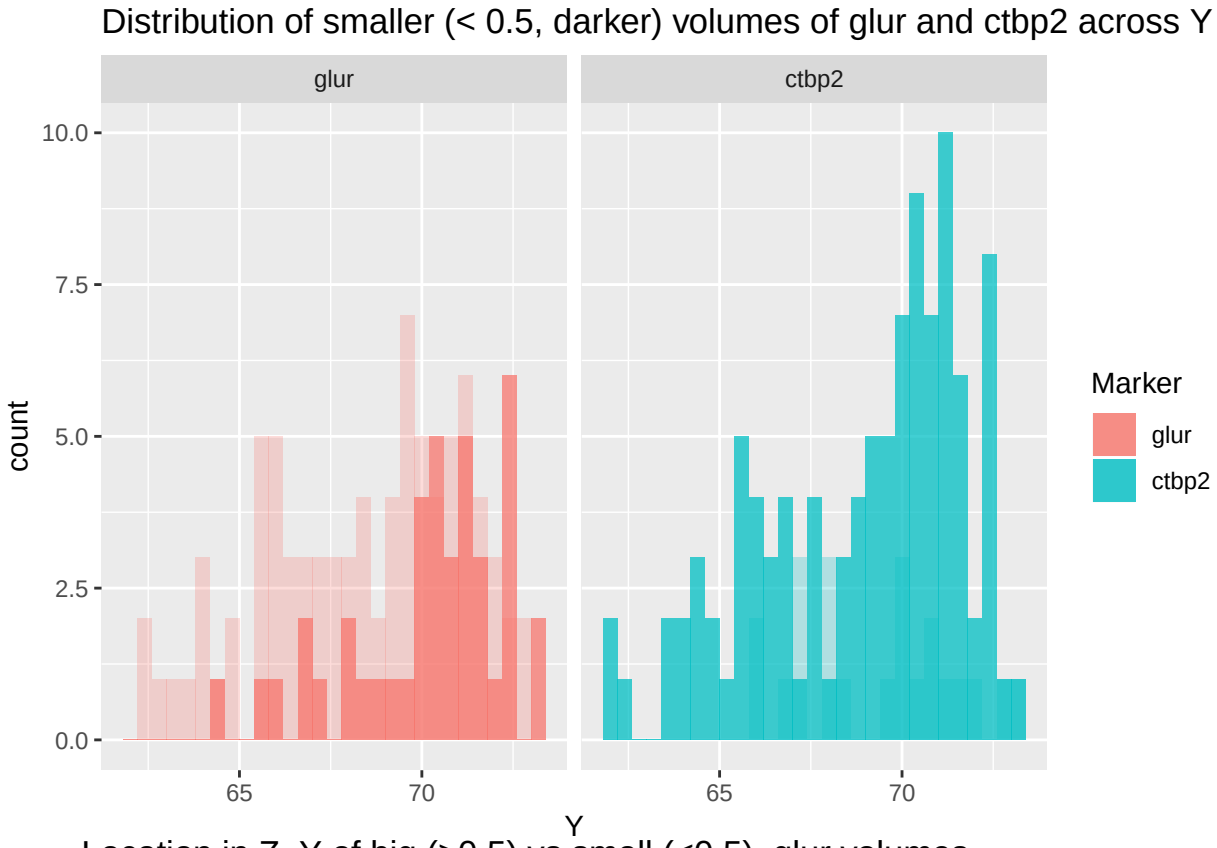
There is a statistically significant relationship between pair volume and Y position.

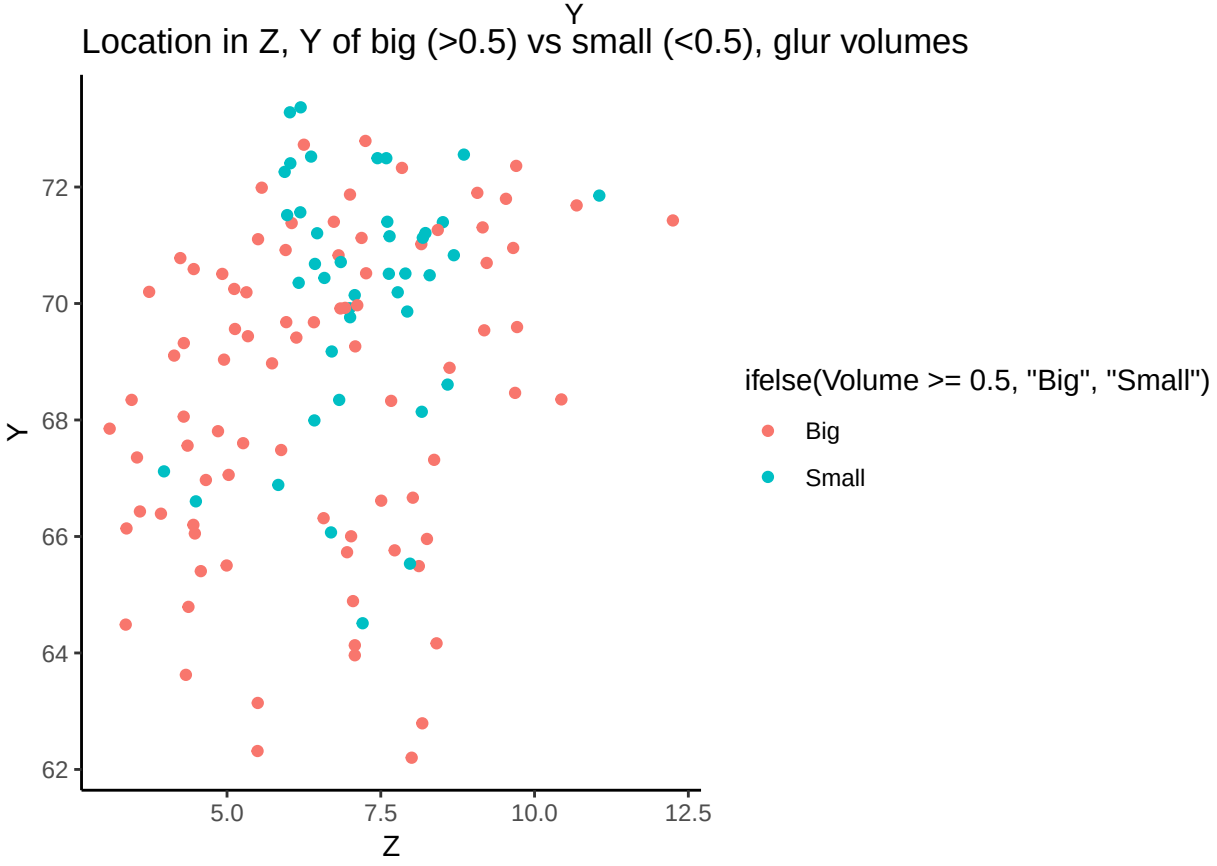
```
fit = lm(Pair.Volume~Y, data = markers[which(markers$Marker == "glur"),c("Pair.Volume","Y")])
summary(fit)
```

```
##
## Call:
## lm(formula = Pair.Volume ~ Y, data = markers[which(markers$Marker ==
##     "glur"), c("Pair.Volume", "Y")])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.67118 -0.18928 -0.04289  0.17664  0.96059
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.776790   0.669676   4.146 6.24e-05 ***
## Y           -0.025772   0.009699  -2.657  0.00893 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2887 on 123 degrees of freedom
## Multiple R-squared:  0.05429,    Adjusted R-squared:  0.0466
## F-statistic: 7.061 on 1 and 123 DF,  p-value: 0.008926
```

Suggestion for possible grouping: big vs small glur volumes

You can split glur volumes into two categories "big" >= .5 and "small" < .5, and see that they are distributed differently along the Y axis.

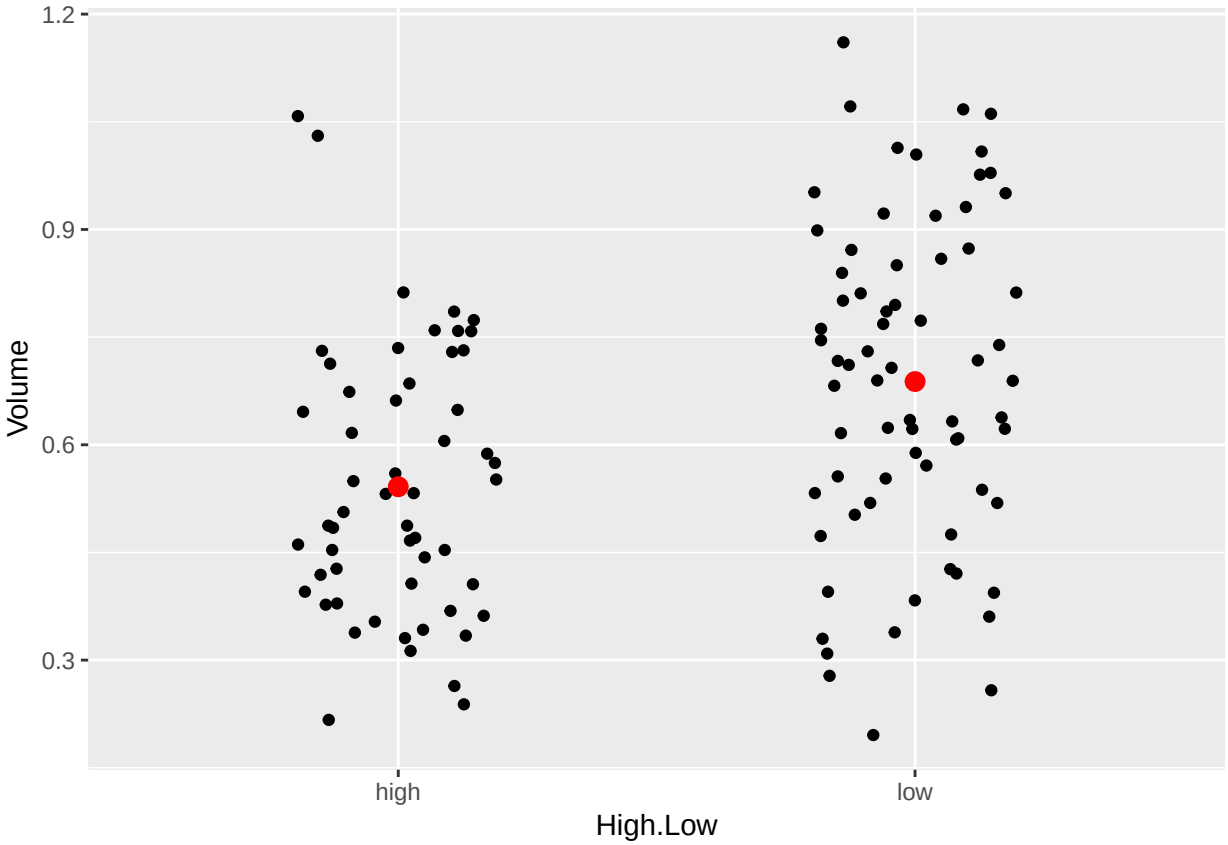

Distribution of smaller (< 0.5, darker) volumes of glur and ctbp2 across Y
glur
ctbp2
count
10.0
7.5
5.0
2.5
0.0
65
70
65
70
Y
Marker
glur
ctbp2


Location in Z, Y of big (>0.5) vs small (<0.5), glur volumes
Y
72
70
68
66
64
62
5.0
7.5
10.0
12.5
Z
ifelse(Volume >= 0.5, "Big", "Small")
Big
Small

Suggestion for population comparisons: Y >= 70, vs Y < 70

I think that doing a regression for Volume against Y location gives more information, but you can also find a statistically significant difference in the mean volumes of glur above and below Y = 70.

```
## No summary function supplied, defaulting to `mean_se()
```




```
## [1] "mean +/- SD glur volume high: 0.542 +/- 0.187"
```

```
## [1] "mean +/- SD glur volume low: 0.688 +/- 0.227"
```

```
##
##  Welch Two Sample t-test
##
## data:  high and low
## t = -3.9618, df = 122.72, p-value = 0.0001253
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.21980547 -0.07333934
## sample estimates:
## mean of x mean of y
## 0.5415386 0.6881110
```

Principle components analysis

Trying PCA to see if variance can be better described by principle components.

```
library(class)
```

```
## Warning: package 'class' was built under R version 3.5.2
```

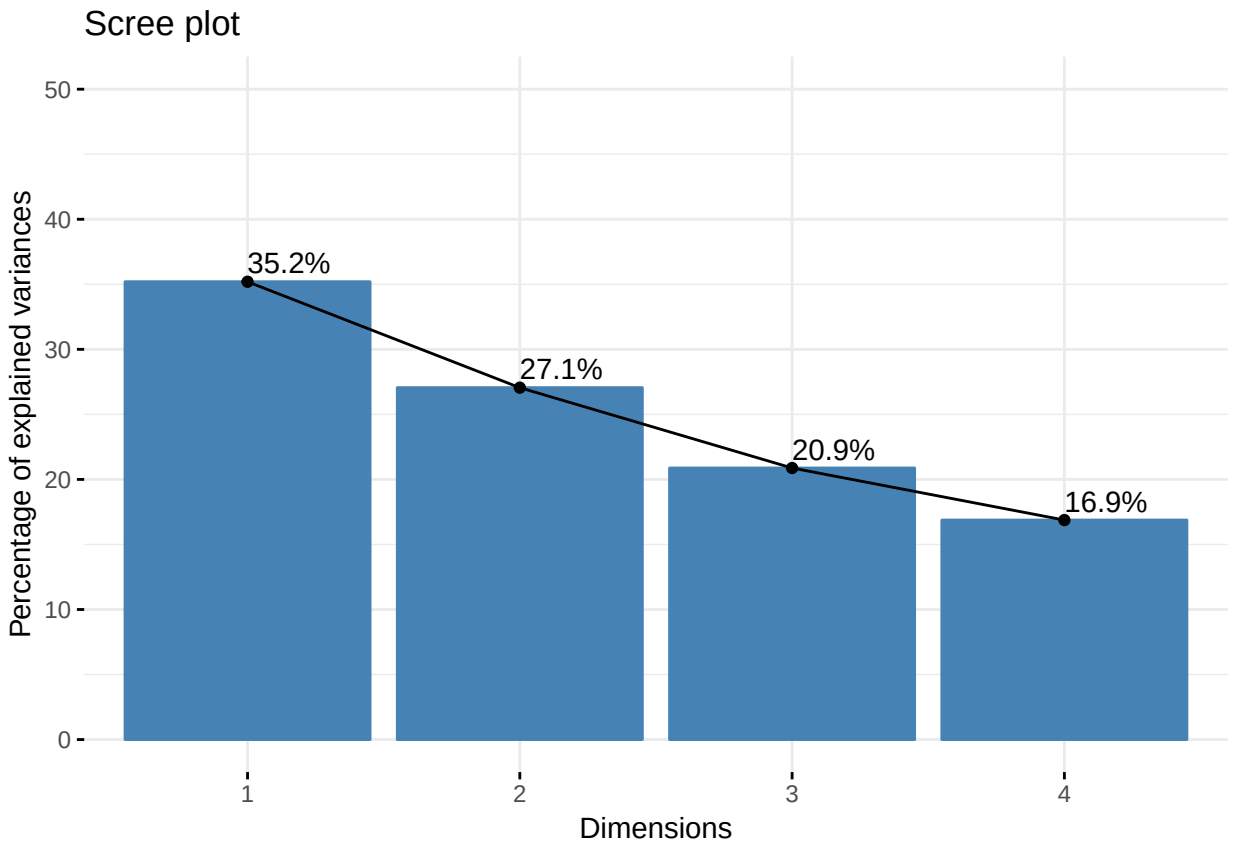
```
glur.pc = prcomp(markers[which(markers$Marker == "glur"),c("Y","Z","Partner.Volume","Pair.Distance")], 
print(glur.pc)
```

```
## Standard deviations (1, .., p=4):
## [1] 1.1865166 1.0402533 0.9138210 0.8215732
##
## Rotation (n x k) = (4 x 4):
##                         PC1         PC2         PC3        PC4
## Y              -0.61047426  0.35354557 -0.07954882 -0.7042718
## Z              -0.61932996 -0.02083207 -0.52274732  0.5854329
## Partner.Volume  0.03518468  0.86717349  0.33477715  0.3670101
## Pair.Distance   0.49245672  0.35011670 -0.77995608 -0.1630129
```

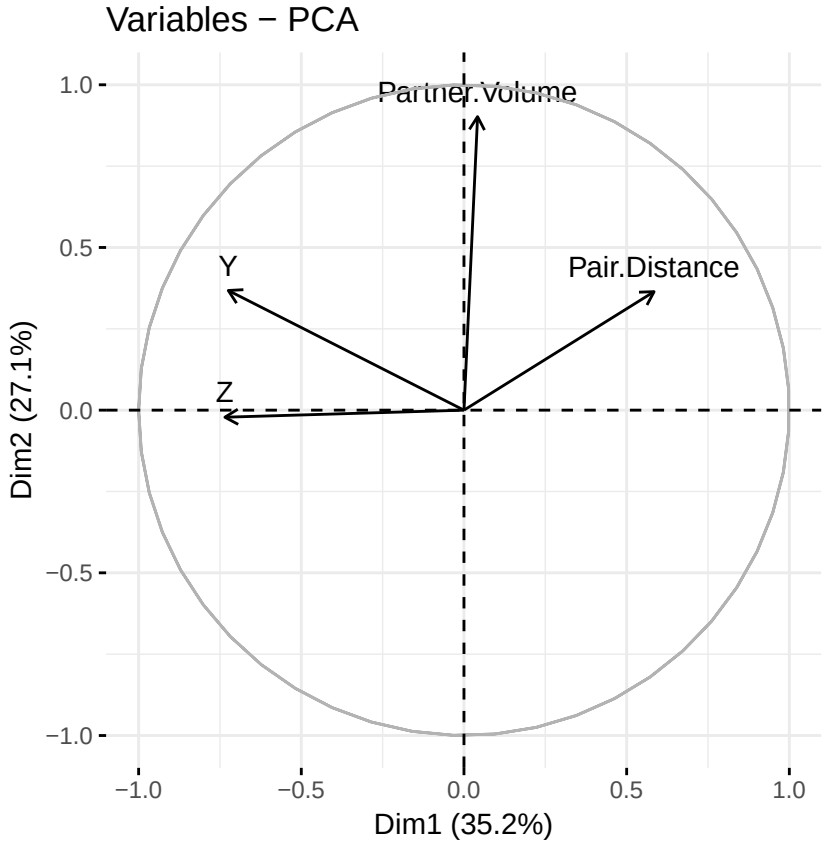
```
eig.val <- get_eigenvalue(glur.pc)
print(eig.val)
```

```
##       eigenvalue variance.percent cumulative.variance.percent
## Dim.1  1.4078217         35.19554                    35.19554
## Dim.2  1.0821269         27.05317                    62.24871
## Dim.3  0.8350688         20.87672                    83.12544
## Dim.4  0.6749826         16.87456                   100.00000
```

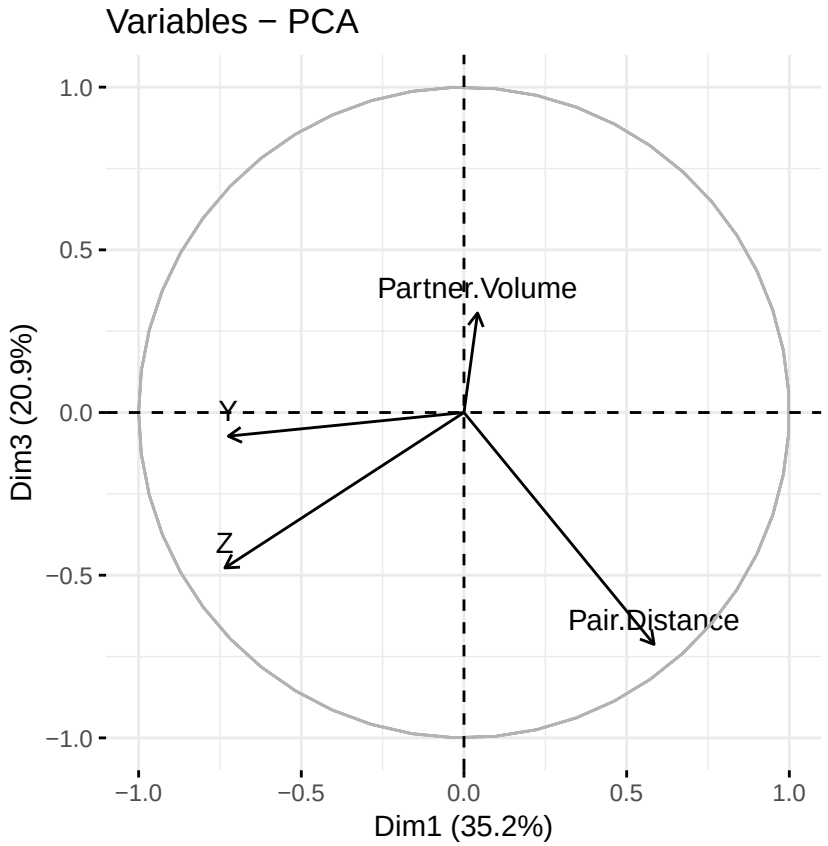
```
fviz_eig(glur.pc, addlabels = TRUE, ylim = c(0, 50))
```



```
## code below does same thing as fviz_eig()
# glur.pc.var = glur.pc$sdev^2
# prop.glur.pc.var = glur.pc.var/sum(glur.pc.var)
# p = plot(prop.glur.pc.var, xlab = "Principal Component",
#              ylab = "Proportion of Variance Explained",
#              type = "b")
# p
var <- get_pca_var(glur.pc)
fviz_pca_var(glur.pc,
             axes = c(1,2),
             col.var = "black")
```



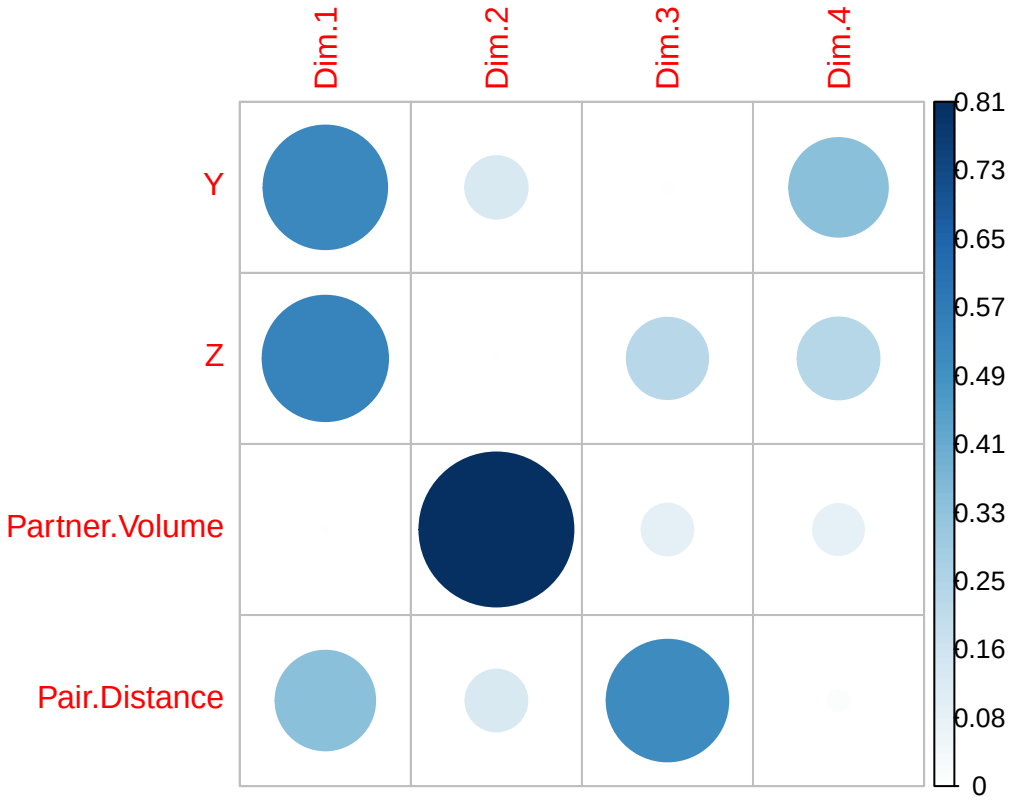

```
fviz_pca_var(glur.pc,
             axes = c(1,3),
             col.var = "black")
```



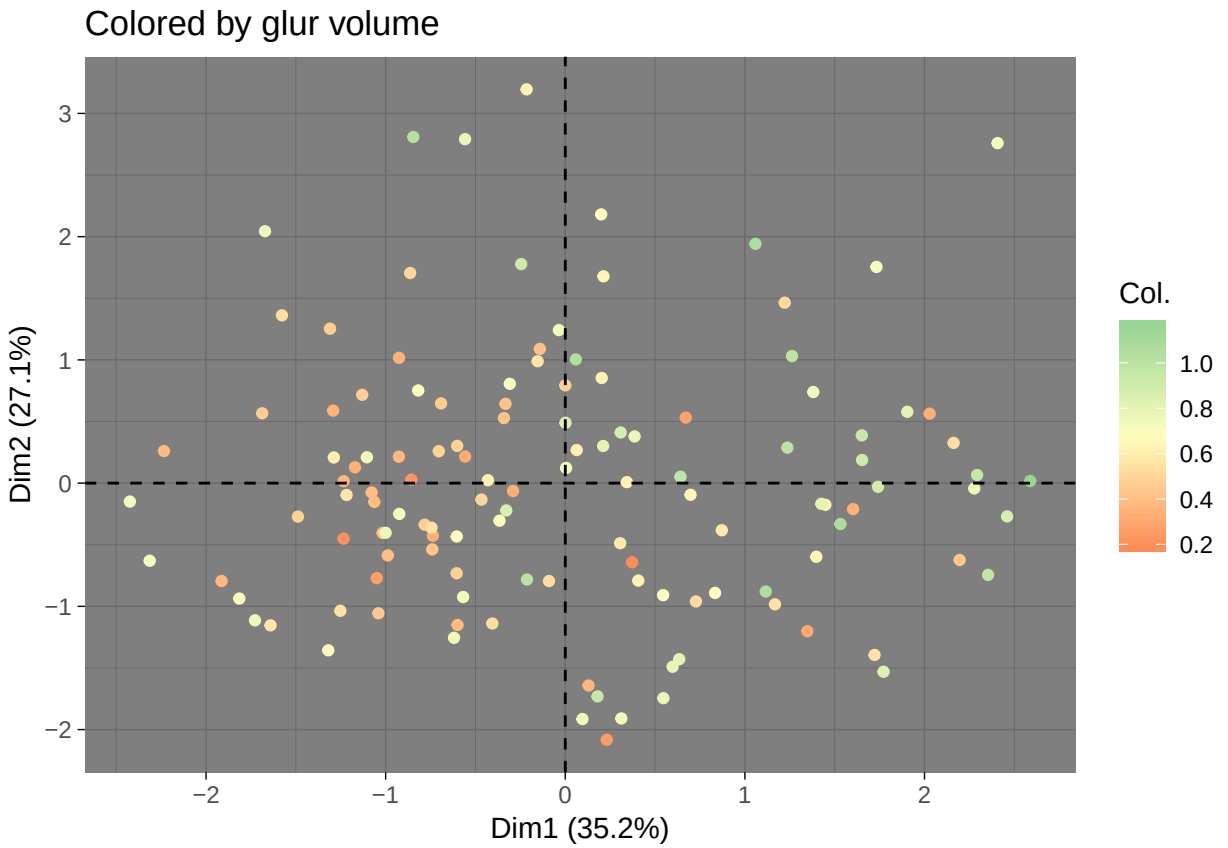

```
library("corrplot")
```

```
## corrplot 0.84 loaded
```

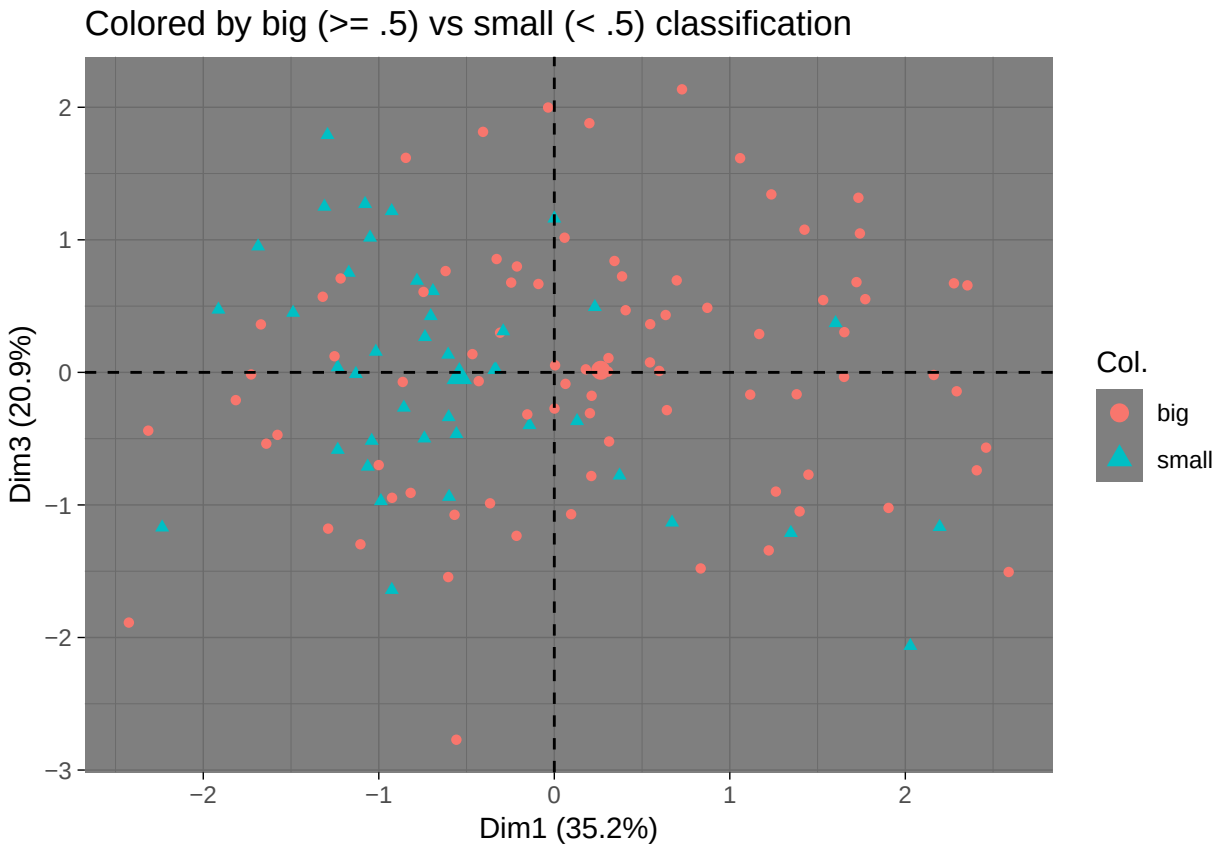
```
corrplot(var$cos2, is.corr=FALSE)
```



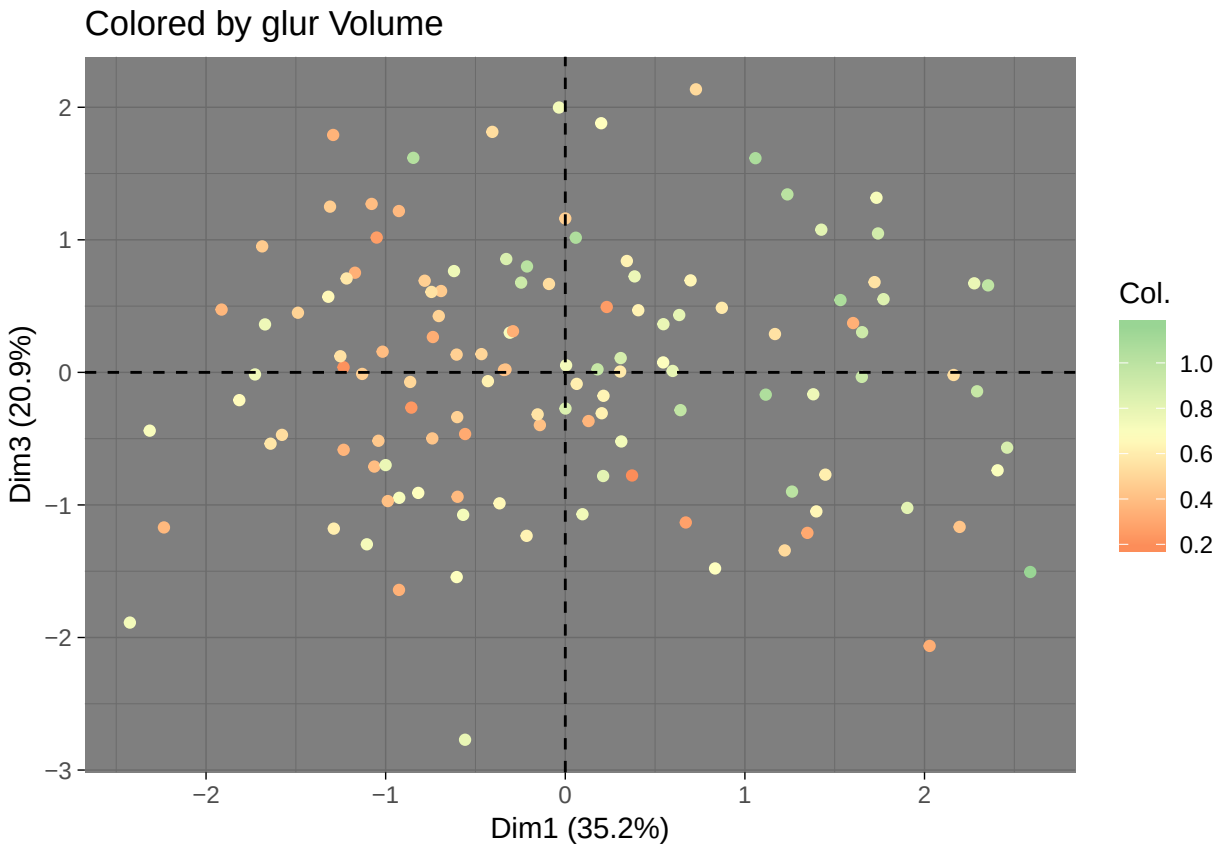

```
# Create a grouping variable using kmeans
# Create 3 groups of variables (centers = 3)
# set.seed(123)
# res.km <- kmeans(var$coord, centers = 3, nstart = 25)
# grp <- as.factor(res.km$cluster)
# # Color variables by groups
# fviz_pca_var(glur.pc, col.var = grp,
#              palette = c("#0073C2FF", "#EFC000FF", "#868686FF"),
#              legend.title = "Cluster")

ind <- get_pca_ind(glur.pc)
fviz_pca_ind(glur.pc,
             axes = c(1,2),
             geom.ind = "point",
             #habillage=markers[which(markers$Marker == "glur"),c("Big.Small")],
             col.ind = markers[which(markers$Marker == "glur"),c("Volume")],
             gradient.cols = "Spectral",
             title = "Colored by glur volume",
             ggtheme = theme_dark())
```



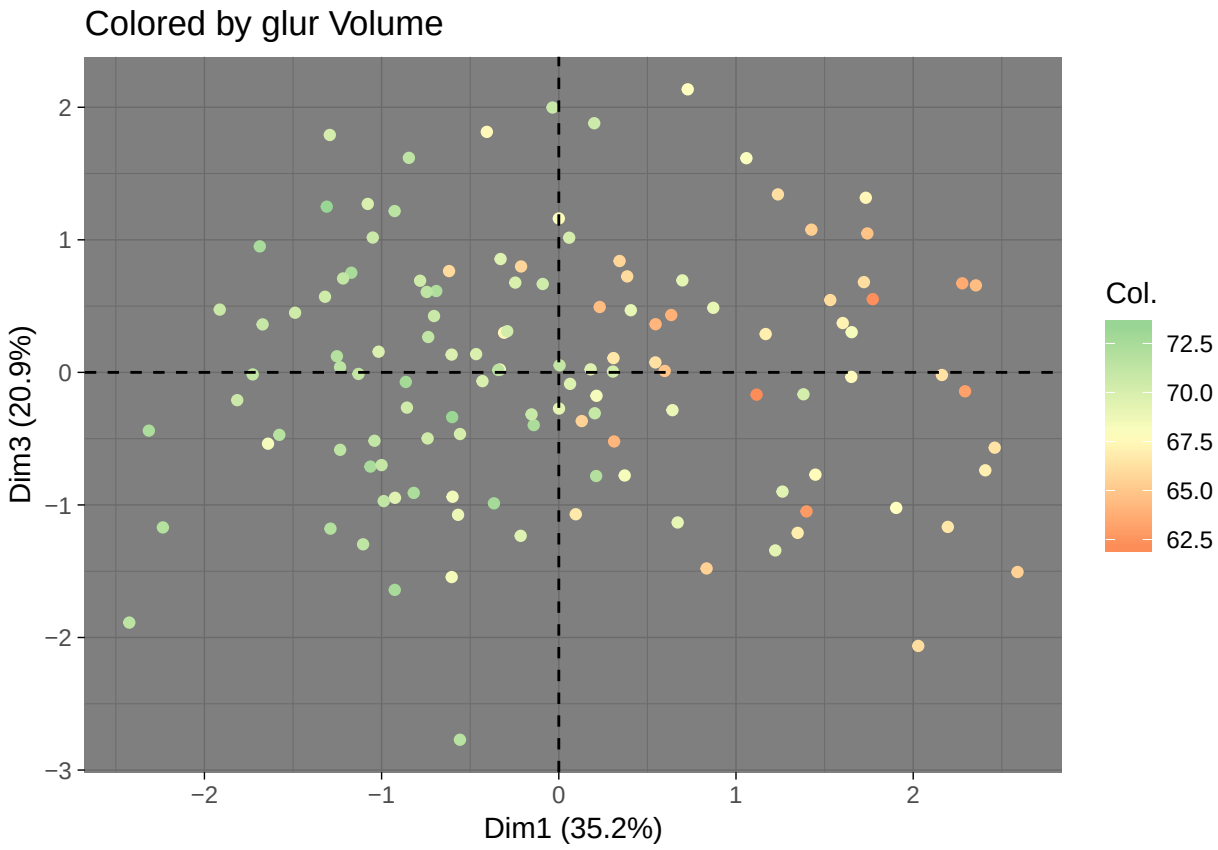

```
test = ind$coord[,1:2]
test.knn = knn(train = test, test = test, cl = as.factor(markers[which(markers$Marker == "glur"),"Big.Sm

fviz_pca_ind(glur.pc,
             axes = c(1,3),
             geom.ind = "point",
             #habillage=test.knn,
             col.ind = as.factor(markers[which(markers$Marker == "glur"),"Big.Small"]),
             title = "Colored by big (>= .5) vs small (< .5) classification",
             ggtheme = theme_dark())
```



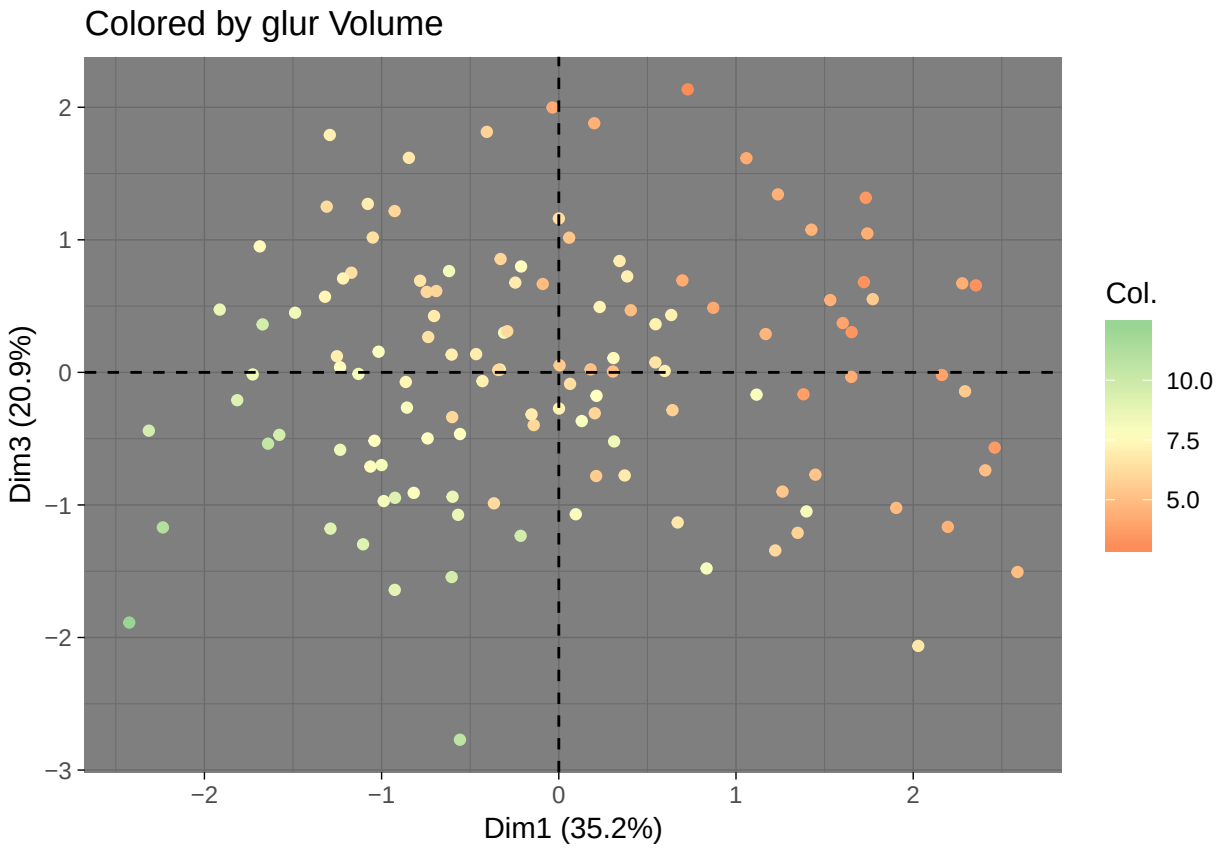

```
fviz_pca_ind(glur.pc,
             axes = c(1,3),
             geom.ind = "point",
             col.ind = markers[which(markers$Marker == "glur"),c("Volume")],
             gradient.cols = "Spectral",
             title = "Colored by glur Volume",
             ggtheme = theme_dark())
```




```
fviz_pca_ind(glur.pc,
             axes = c(1,3),
             geom.ind = "point",
             col.ind = markers[which(markers$Marker == "glur"),c("Y")],
             gradient.cols = "Spectral",
             title = "Colored by glur Volume",
             ggtheme = theme_dark())
```




```
fviz_pca_ind(glur.pc,
             axes = c(1,3),
             geom.ind = "point",
             col.ind = markers[which(markers$Marker == "glur"),c("Z")],
             gradient.cols = "Spectral",
             title = "Colored by glur Volume",
             ggtheme = theme_dark())
```



After exploring the data, it seems like separating the data into groups that may correspond to sensitive vs non-sensitive synapses based on criteria like volume, volume ratio, total volume, etc in ZY space might not be possible, because the markers were spread across multiple cells that were not closely aligned enough in X. That's not to say there were no trends. The strongest trend seemed to be GLUR volume along the Y axis. GLUR gets smaller further up in Y. The distributions of the volumes of GLUR and CTBP2 were also different, with CTBP2 being generally smaller than GLUR.