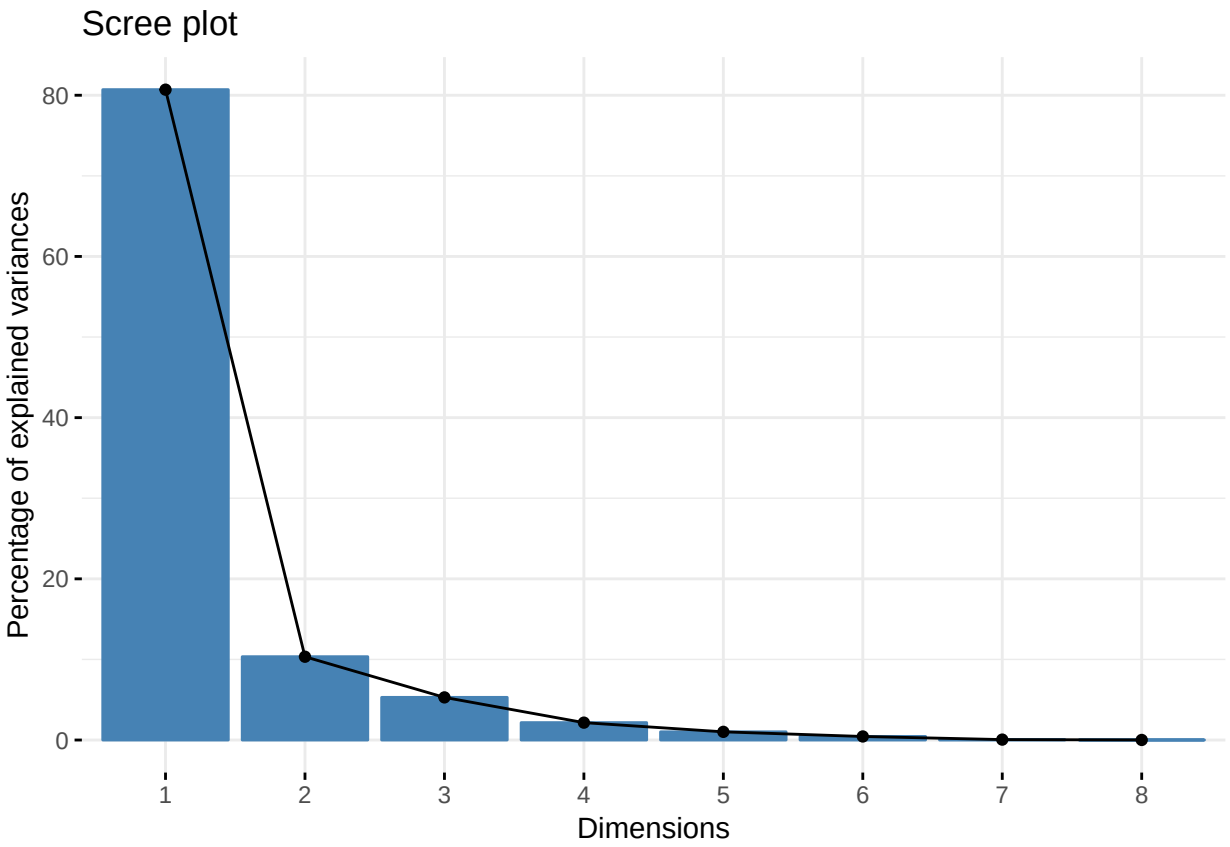
```r
#data preprocessing
library(tidyverse)
library(MRIcloudT1volumetrics)

roiDir = "C:/Users/lcqi/OneDrive/Desktop/bcaffo/data_analysis_project/data"
fileList = dir(roiDir)


dat = list()
for (i in 1:length(fileList)){
  fullPath = paste(roiDir, fileList[i], sep = "/")
  raw_dat = readSubject(fullPath) %>% subject2df()
  #type 1
  dat_type1 = raw_dat %>% filter(type ==1)
  #diff level
  #for level 5, same roi have different volumn.
  #e.g. BasalForebrain_L have 5 diff volumn.here choose level 1:4
  for (j in 1:4){
    dat_each = dat_type1 %>% filter(level == j) %>% select(rawid,roi,volume) %>% spread(roi,volume)
    dat[[paste0('level',j)]] = rbind(dat[[paste0('level',j)]],dat_each)
  }

}


#transform rawid to integer for further work
for (j in 1:4){
  id = dat[[paste0('level',j)]]$rawid
  dat[[paste0('level',j)]]$rawid = as.numeric(sapply(strsplit(id,"_"),function(x) x[1]))
}
```

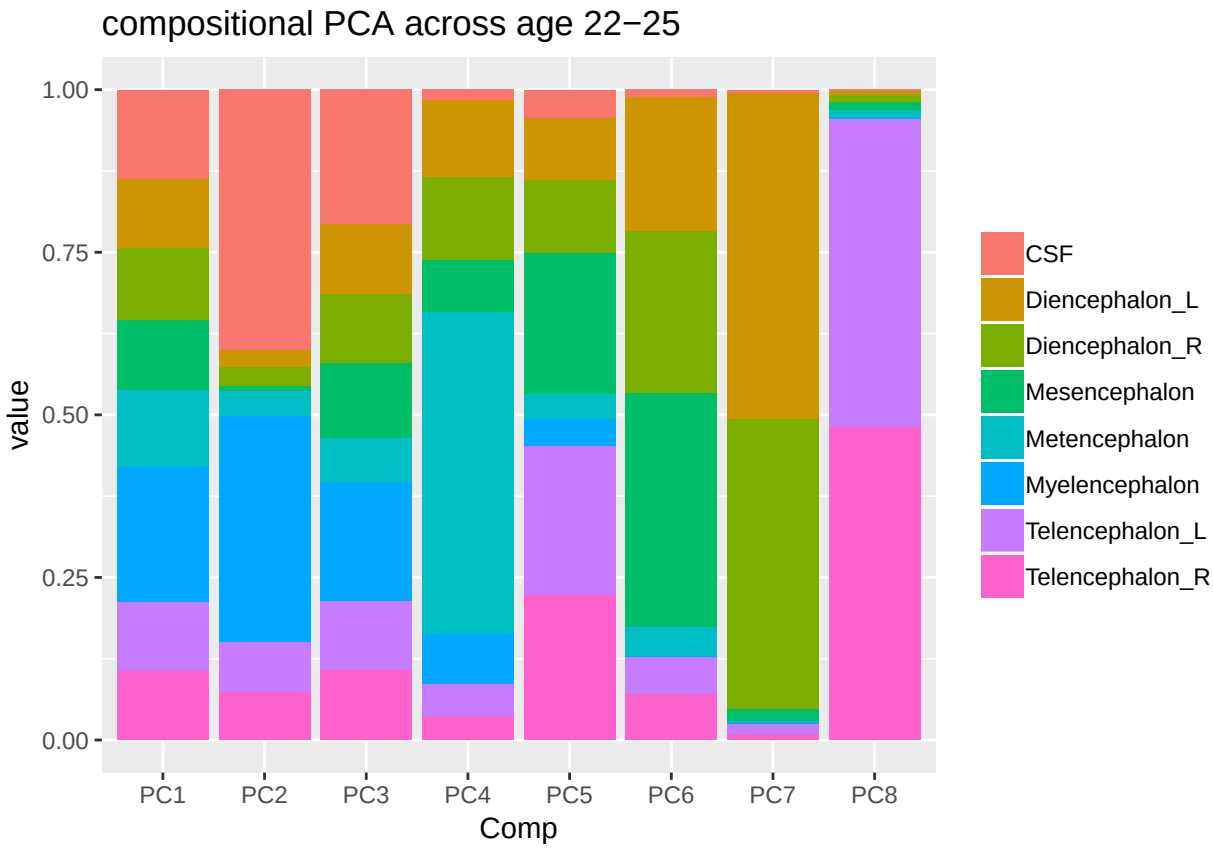
```
## Warning: NAs introduced by coercion
```

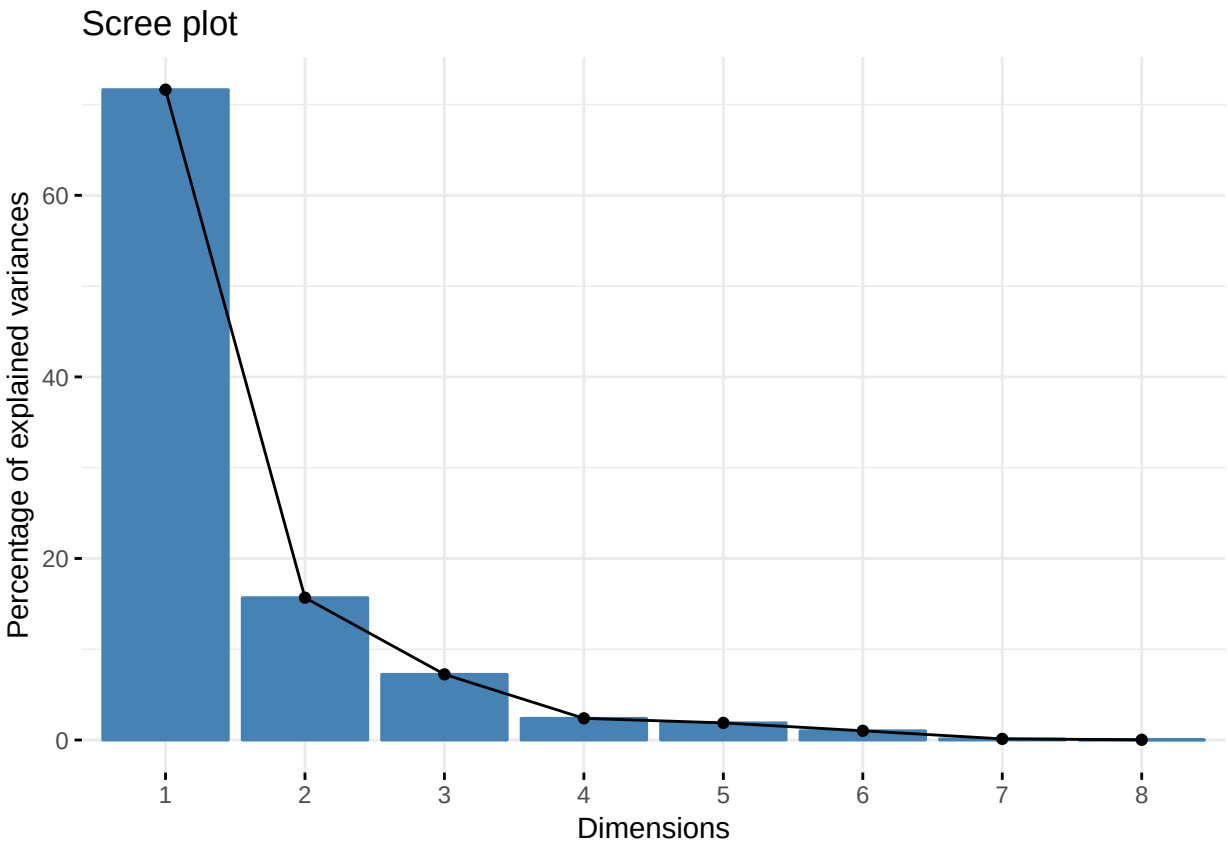
```
## Warning: NAs introduced by coercion
```

```r
#compositional pca analysis across age
library(factoextra)
library(magrittr)
library(compositions)
library(data.table)


#data processing
library(tidyverse)
ref = read.csv("C:/Users/lcqi/OneDrive/Desktop/bcaffo/data_analysis_project/unrestricted_bcaffo_12_12_2
colnames(ref)[1] = 'rawid'
dat_cor = left_join(dat$level1, ref[,c('rawid','Age','Gender')], by="rawid")


#across age

Age = as.character(unique(dat_cor$Age))
Age = sort(Age[!is.na(Age)])
```

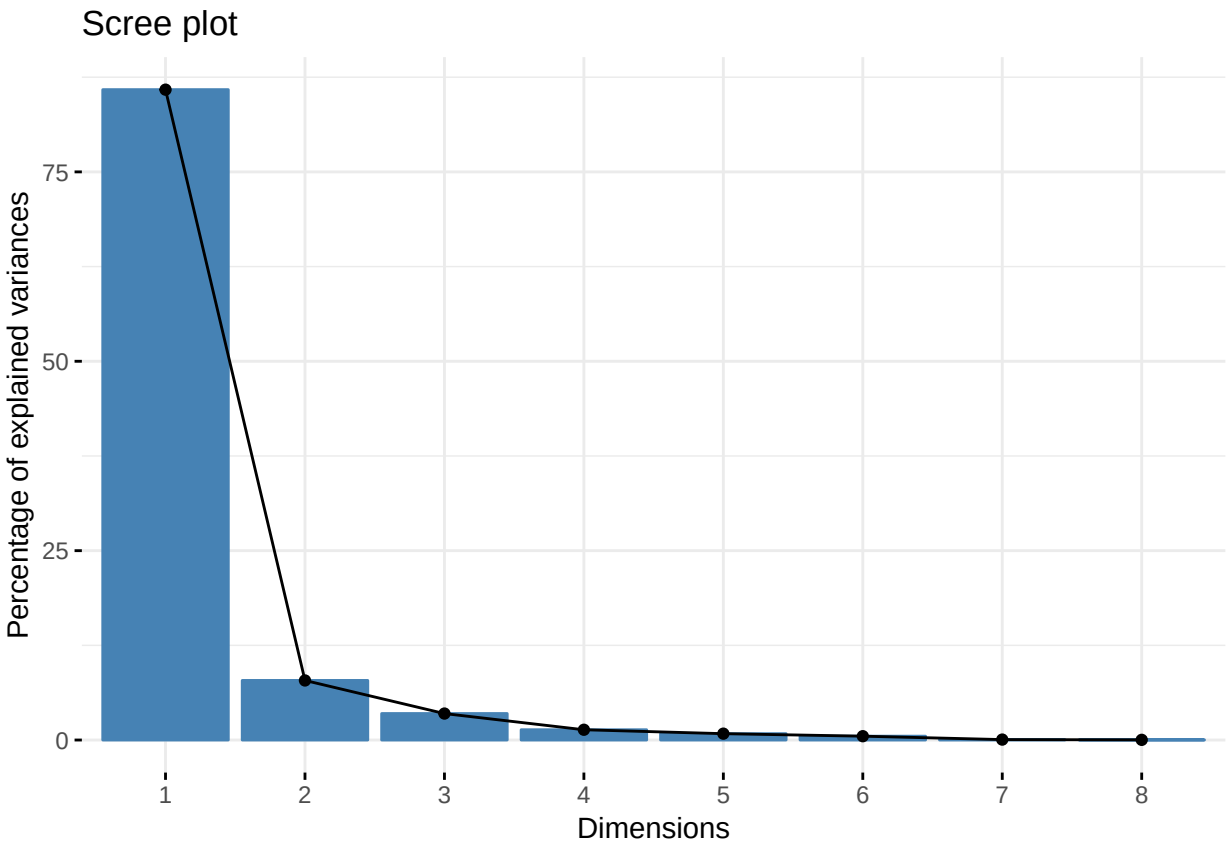
```
dat_age = list()

for (i in Age){
  a = dat_cor %>% filter(Age == i)
  dat_age[[i]] =  a[, !colnames(a) %in% c('rawid','Age','Gender')]
}


res.rotations = list()

for (i in names(dat_age)){

  #compositional analysis
  cdata = acomp(dat_age[[i]])
  #pca analysis
  cdata.pca = prcomp(cdata,scale. = T)
  dat_visualization = prop.table(abs(cdata.pca$rotation),margin = 2) %>% melt()
  colnames(dat_visualization) = c('roi','Comp','value')

  res.rotations[[i]] = prop.table(abs(cdata.pca$rotation),margin = 2) %>% melt()
  # %>% filter(Var2 == 'PC1')
  colnames(res.rotations[[i]]) = c('roi','PC','value')

print(
  fviz_eig(cdata.pca)
)

print(
  ggplot(dat_visualization, aes(x = as.factor(Comp),
                    y = value,
                    fill = as.factor(roi))
           ) + geom_col() +theme(legend.title = element_blank()) +
    labs(x = 'Comp', title = paste('compositional PCA across age',i))
)
}
```

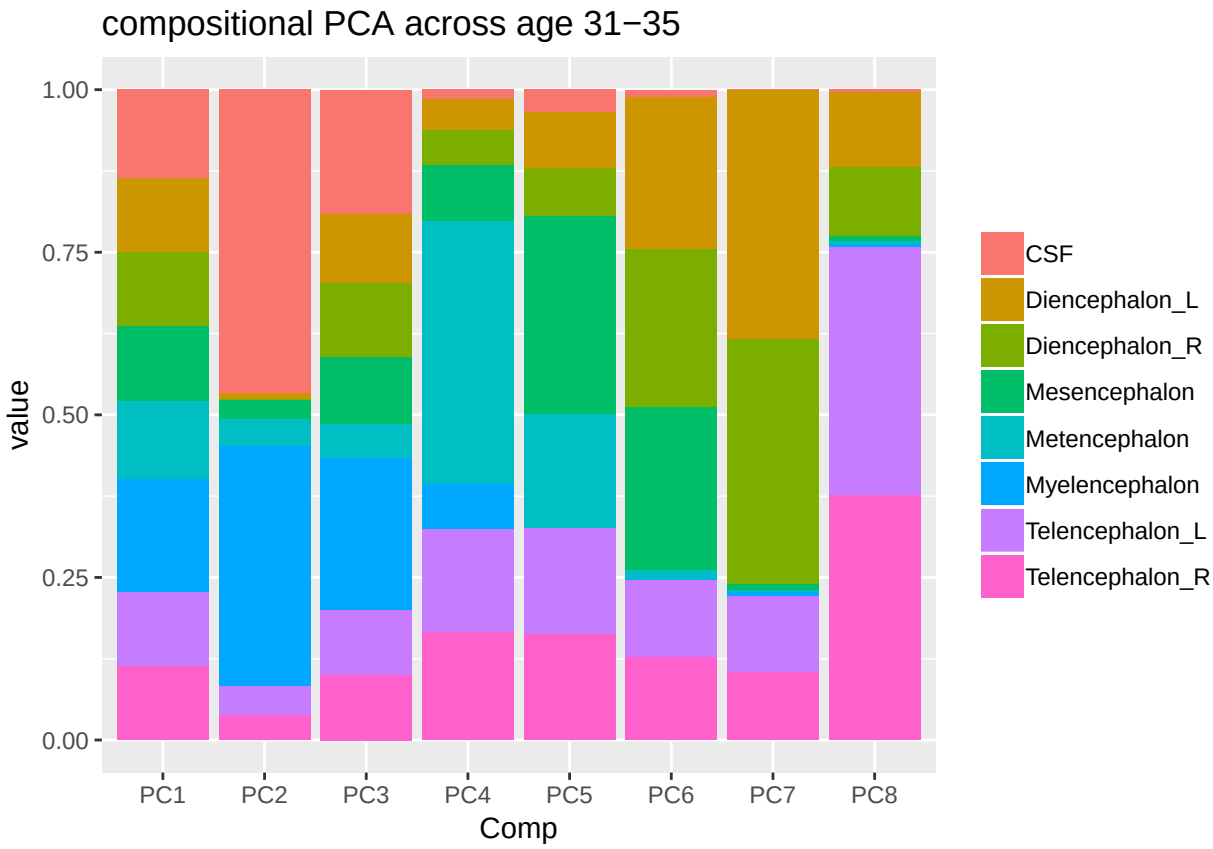

Scree plot
Percentage of explained variances
80
60
40
20
0
1
2
3
4
5
6
7
8
Dimensions

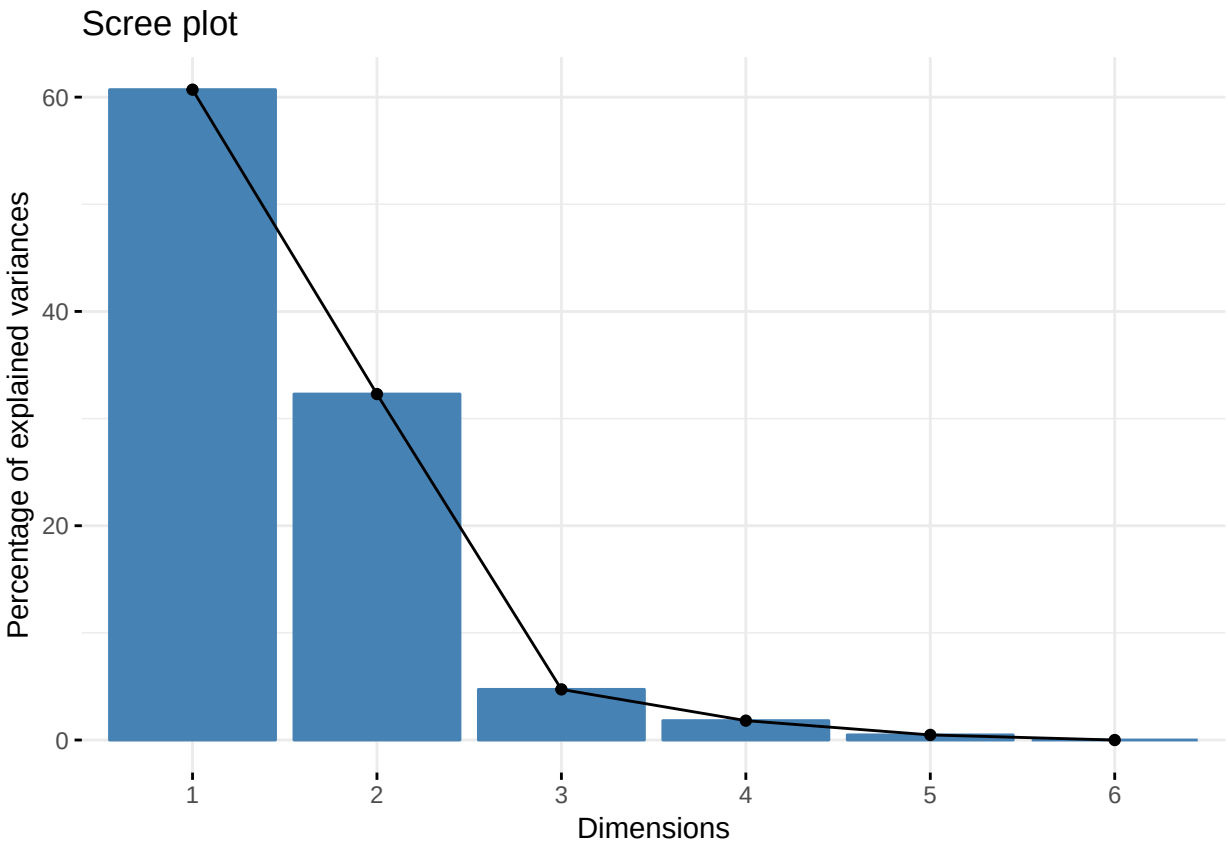

compositional PCA across age 22−25
1.00
0.75
0.50
0.25
0.00
value
PC1
PC2
PC3
PC4
PC5
PC6
PC7
PC8
Comp
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R

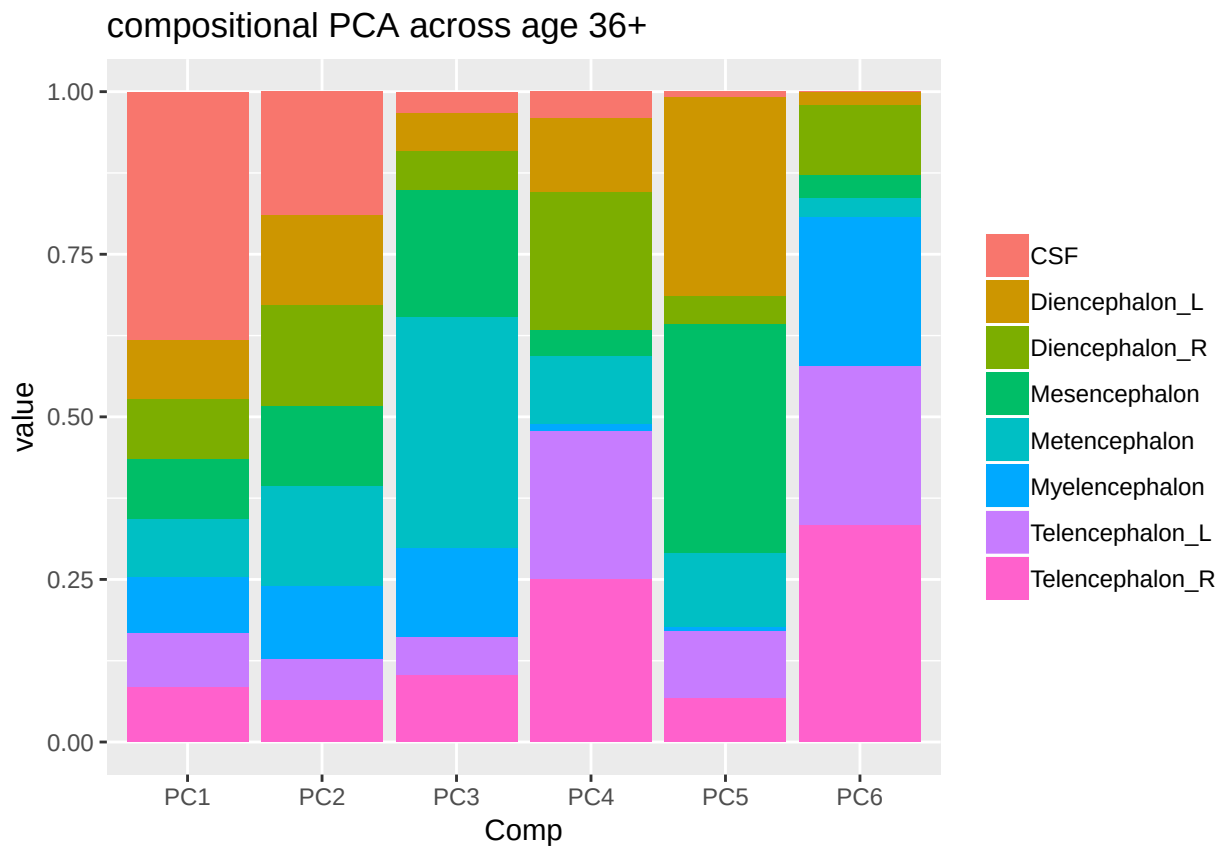

Scree plot
Percentage of explained variances
60
40
20
0
1
2
3
4
5
6
7
8
Dimensions


compositional PCA across age 26−30
value
1.00
0.75
0.50
0.25
0.00
PC1
PC2
PC3
PC4
PC5
PC6
PC7
PC8
Comp
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R

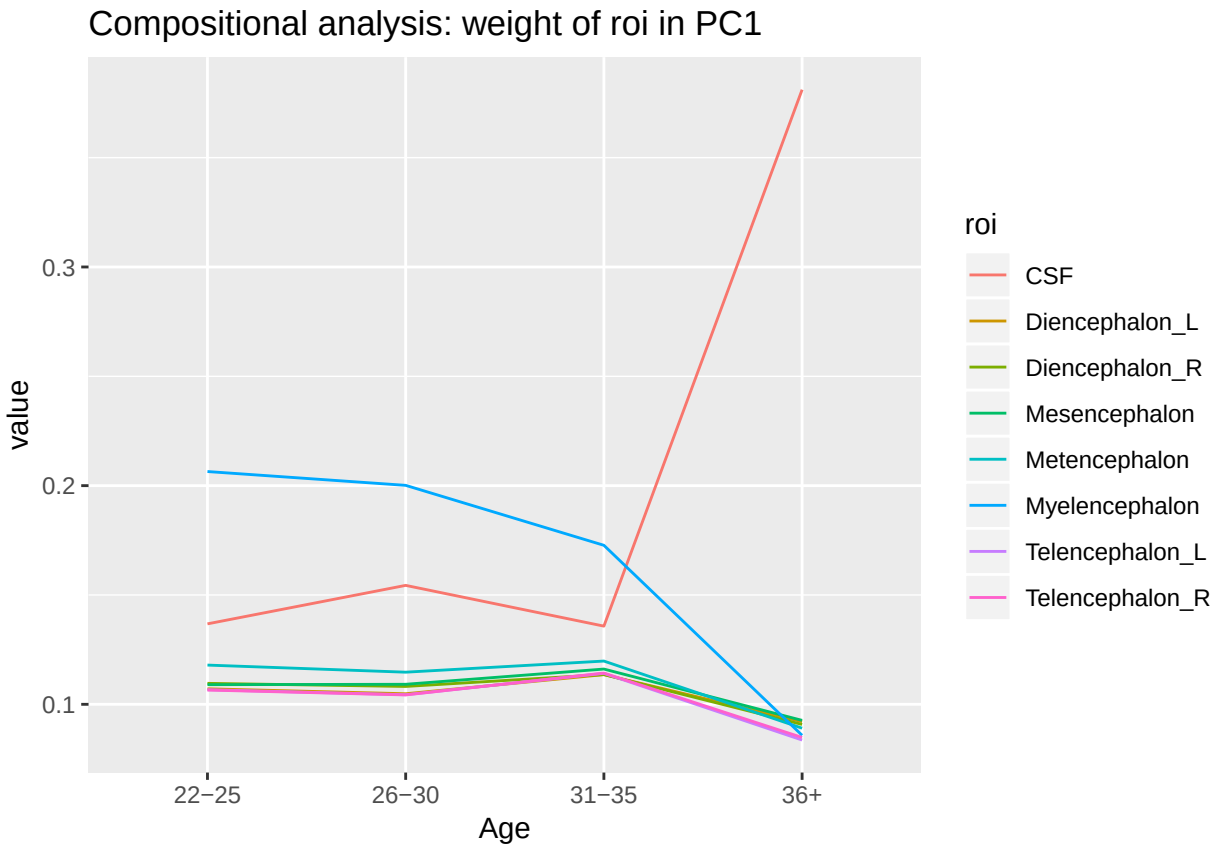

Scree plot
Percentage of explained variances
75
50
25
0
1
2
3
4
5
6
7
8
Dimensions


compositional PCA across age 31−35
1.00
0.75
0.50
0.25
0.00
value
PC1
PC2
PC3
PC4
PC5
PC6
PC7
PC8
Comp
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R


Scree plot
Percentage of explained variances
60
40
20
0
1
2
3
4
5
6
Dimensions




```
res.rotations$compile = c()

#PC1 visualization
for (i in names(dat_age)){
  res.rotations$compile = rbind(res.rotations$compile,
                                select(res.rotations[[i]] %>% filter(PC == 'PC1'),c(roi,value))
                                %>% mutate(Age = i))
}
ggplot(res.rotations$compile, aes(x = Age,y = value,colour = roi,group = roi)) +
  geom_line() +
  labs(title = 'Compositional analysis: weight of roi in PC1')
```


Compositional analysis: weight of roi in PC1
value
0.3
0.2
0.1
22−25
26−30
31−35
36+
Age
roi
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R