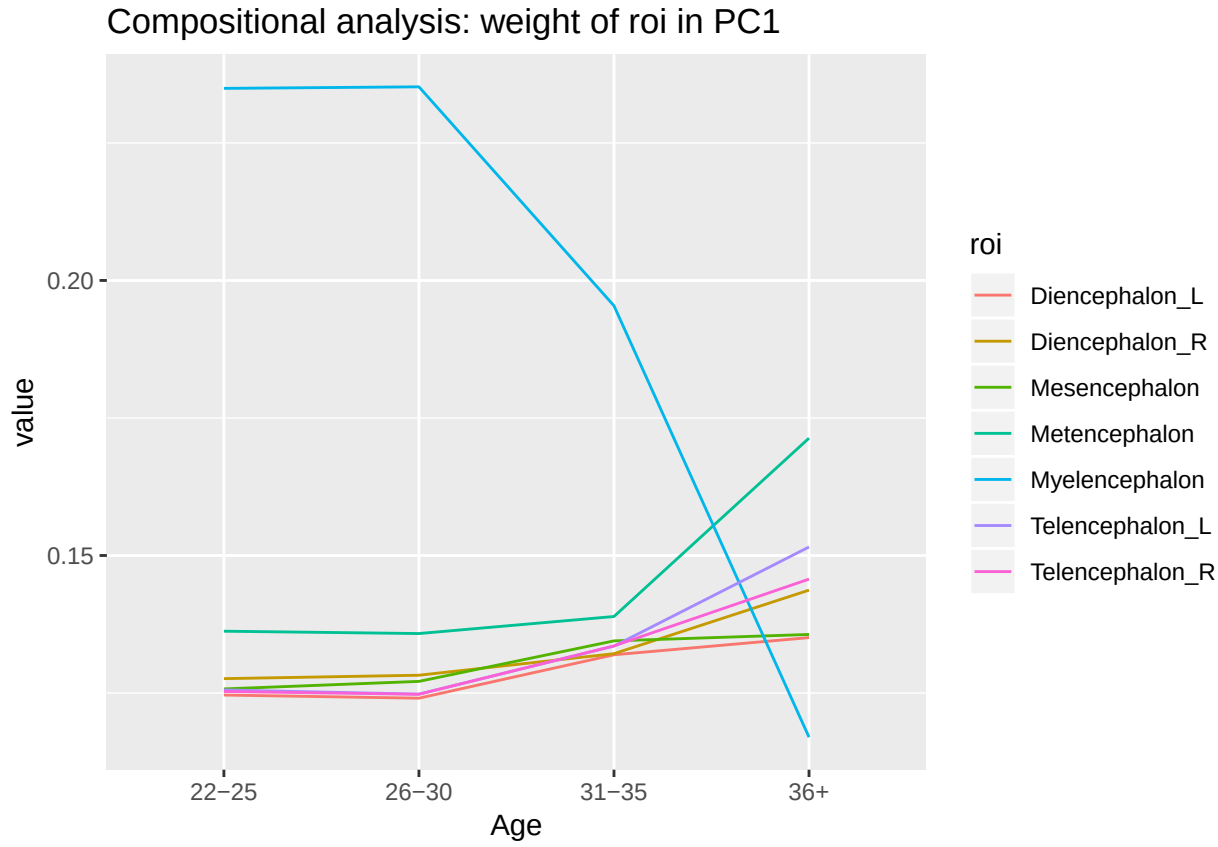
```r
library(tidyverse)
library(MRIcloudT1volumetrics)

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


#type = 1, level = 1
# dat = dat %>% mutate(volume_demean = volume - mean(volume))%>% select(roi,volume_demean)%>% spread(ro
dat = c()
dat = list()
for (i in 1:length(fileList)){
  fullPath = paste(roiDir, fileList[i], sep = "/")
  raw_dat = readSubject(fullPath) %>% subject2df()
  dat_type1 = raw_dat %>% filter(type ==1)
  # unique(dat_type1$level
  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)
  }

  # dat_each = raw_dat %>% filter(type == 1, level == 1) %>% select(roi,volume) %>% spread(roi,volume)
  # dat = rbind(dat,dat_each)
}


#extract rawid (integer)
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
```

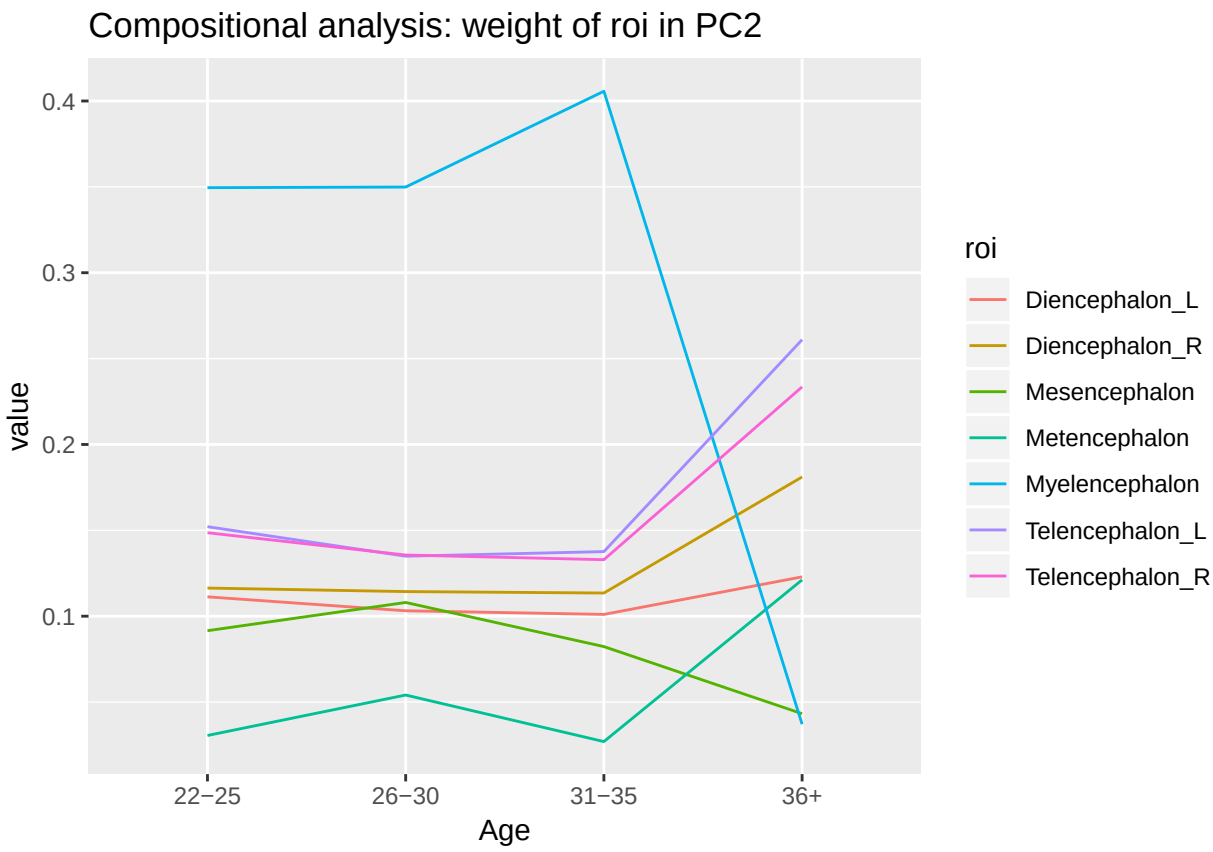
```r
#PCA after getting rid of CSF
library(compositions)
library(factoextra)


dat_age_noCSF = list()
#compositional analysis
for(i in Age){
  dat_age_noCSF[[i]] = dat_age[[i]] %>%
    select(-CSF) %>% acomp()
}

res.pca = list()
res.rotations_noCSF = list()
for(i in Age){
  res.pca[[i]] = prcomp(dat_age_noCSF[[i]],scale. = T)
  dat_visualization = prop.table(abs(res.pca[[i]]$rotation),margin = 2) %>% melt()
  colnames(dat_visualization) = c('roi','PC','value')
```

```
  #further work on PC1
  res.rotations_noCSF[[i]] = dat_visualization %>% filter(PC == 'PC1')

}


res.rotations_noCSF$compile = c()

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



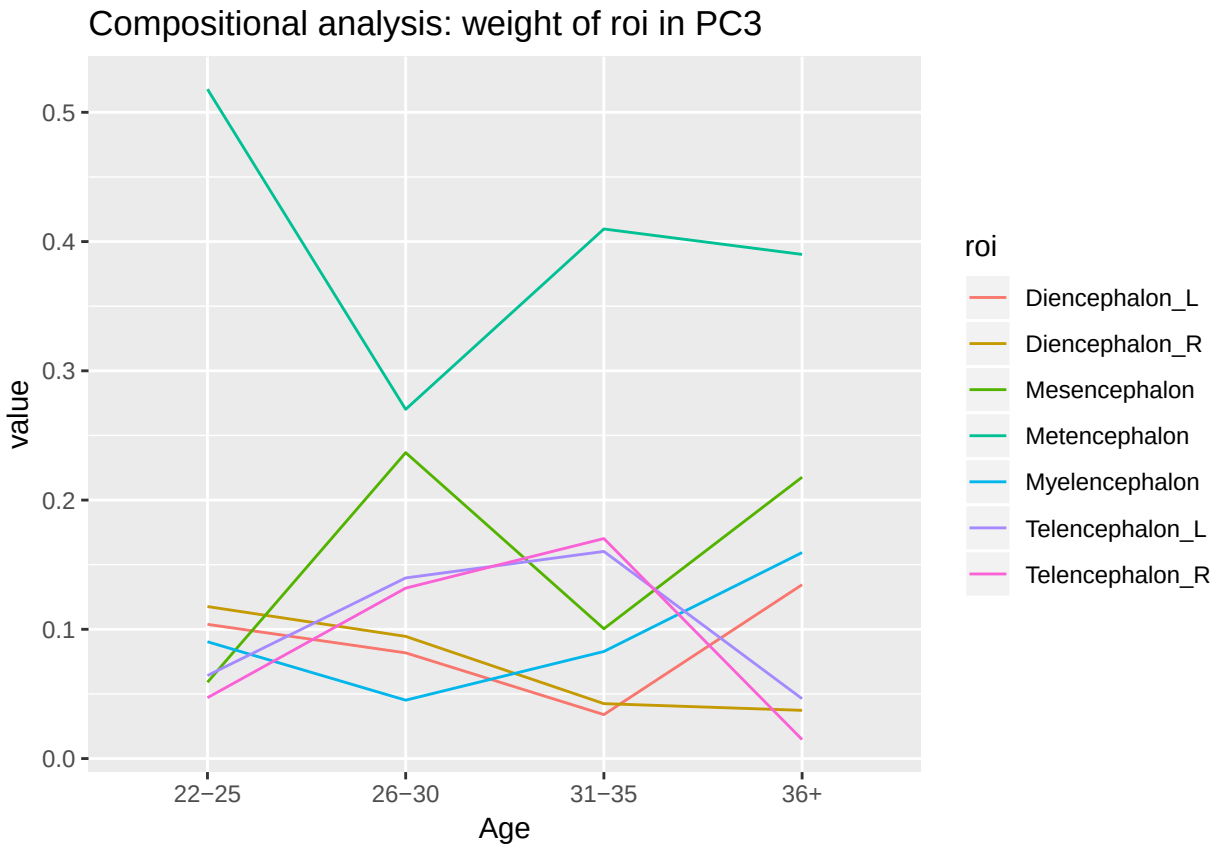

```
# PC2 analysis

for ( i in Age){
  dat_visualization = prop.table(abs(res.pca[[i]]$rotation), margin = 2) %>% melt()
  colnames(dat_visualization) = c('roi','PC','value')
  res.rotations_noCSF[[i]] = dat_visualization %>% filter(PC == 'PC2')
}

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

```
for (i in Age){
  res.rotations_noCSF$compile = rbind(res.rotations_noCSF$compile,
                                      select(res.rotations_noCSF[[i]],c(roi,value))
                                      %>% mutate(Age = i))
}

ggplot(res.rotations_noCSF$compile, aes(x = Age,y = value,colour = roi,group = roi)) +
  geom_line() +
  labs(title = 'Compositional analysis: weight of roi in PC2')
```

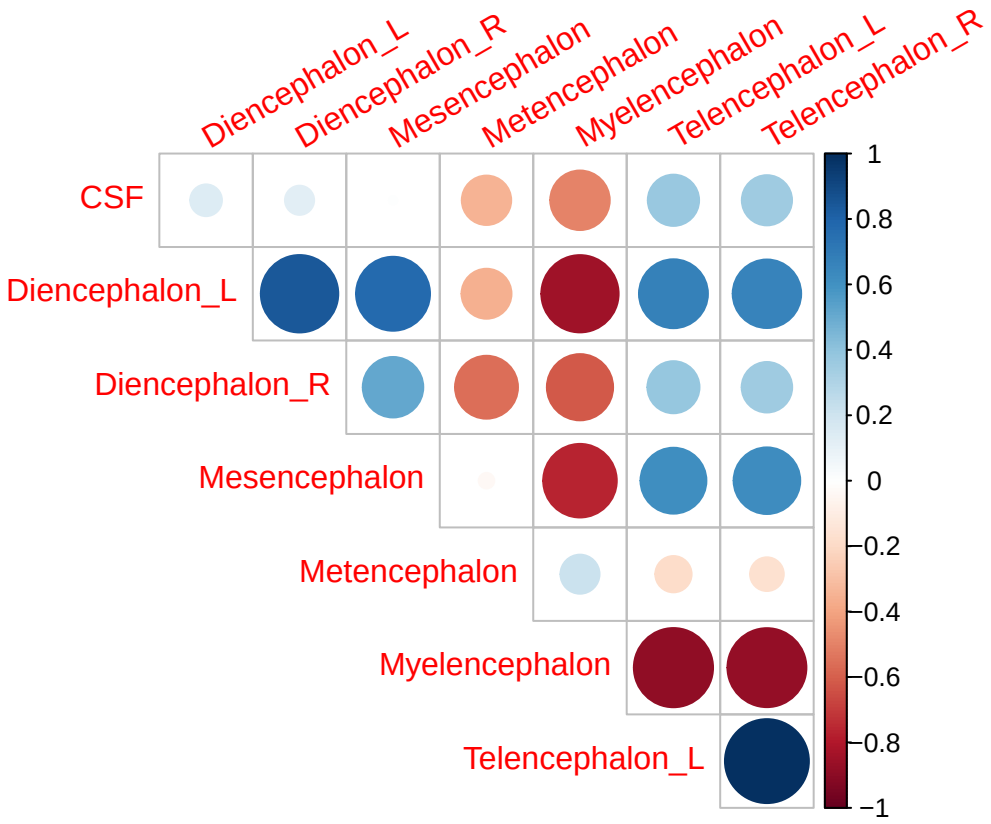


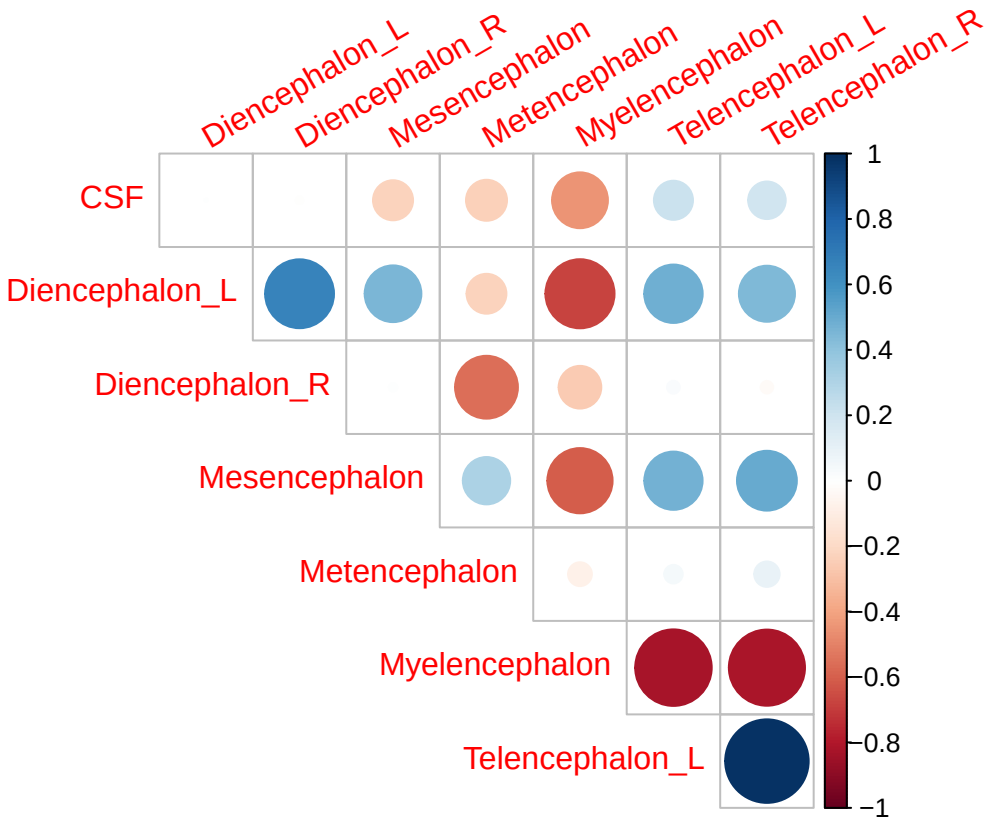

```
#PC3

for ( i in Age){
  dat_visualization = prop.table(abs(res.pca[[i]]$rotation), margin = 2) %>% melt()
  colnames(dat_visualization) = c('roi','PC','value')
  res.rotations_noCSF[[i]] = dat_visualization %>% filter(PC == 'PC3')
}

res.rotations_noCSF$compile = c()
for (i in Age){
  res.rotations_noCSF$compile = rbind(res.rotations_noCSF$compile,
                                      select(res.rotations_noCSF[[i]],c(roi,value))
                                      %>% mutate(Age = i))
}
ggplot(res.rotations_noCSF$compile, aes(x = Age,y = value,colour = roi,group = roi)) +
  geom_line() +
```

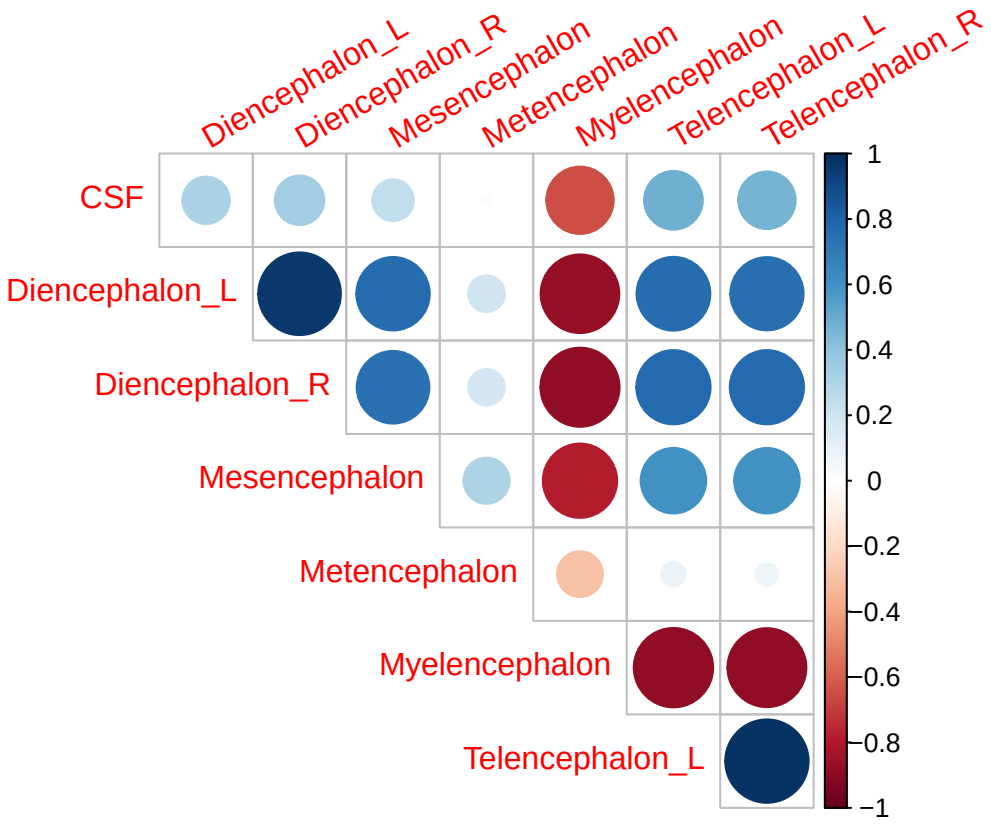
```
  labs(title = 'Compositional analysis: weight of roi in PC3')
```

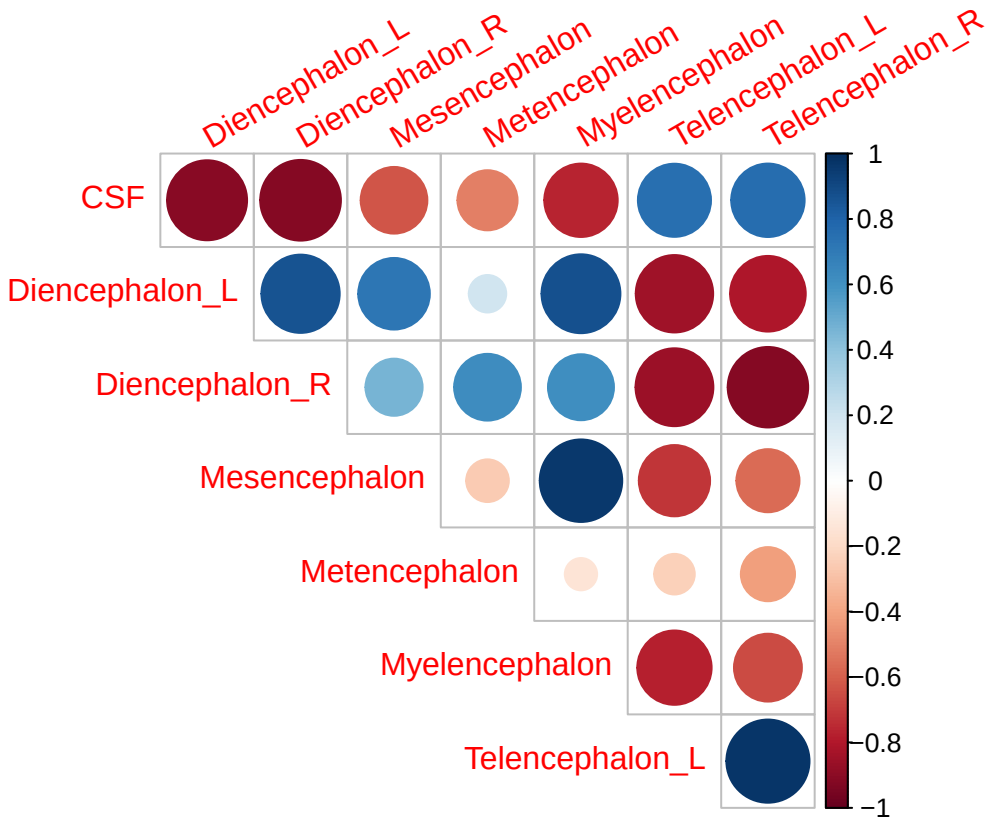



```
# correlation accross age in all genders

library(compositions)
library(corrplot)

res.allgender = list()
dat_all_gender = list()

for ( i in Age){
  dat_all_gender[[i]] = filter(dat_cor$level1, Age == i) %>% select(-c(rawid,Age,Gender))
  res.allgender[[i]] = acomp(dat_all_gender[[i]])

  A =   corrplot(cor(res.allgender[[i]]), type="upper",
              # order="hclust",

           # title = 'All genders',
           tl.srt=30,
           diag=FALSE)


}
```

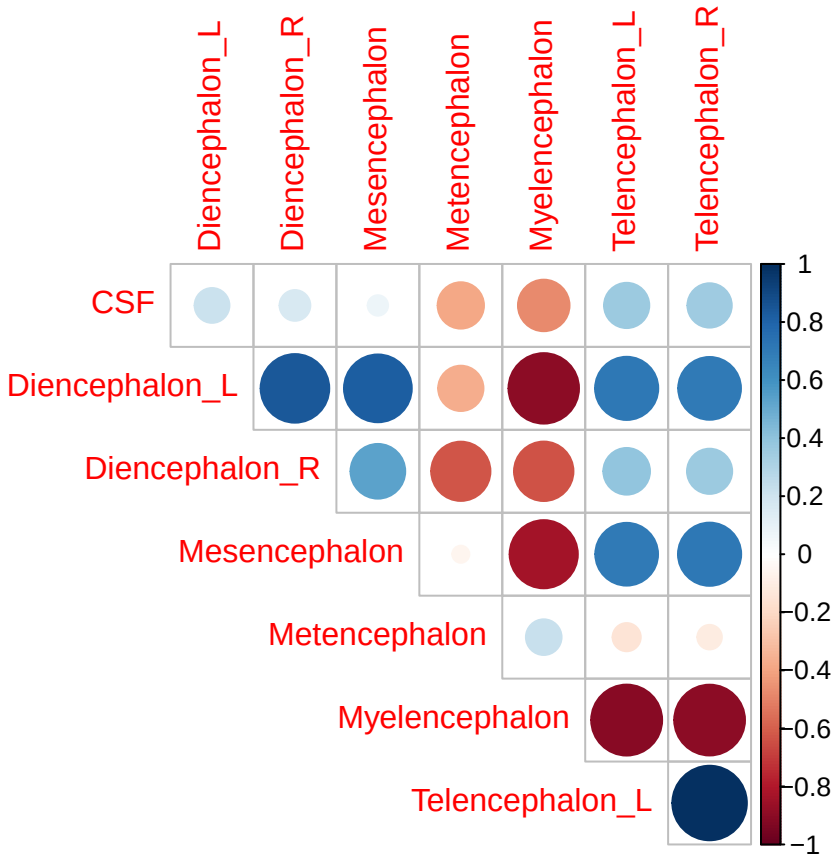

Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

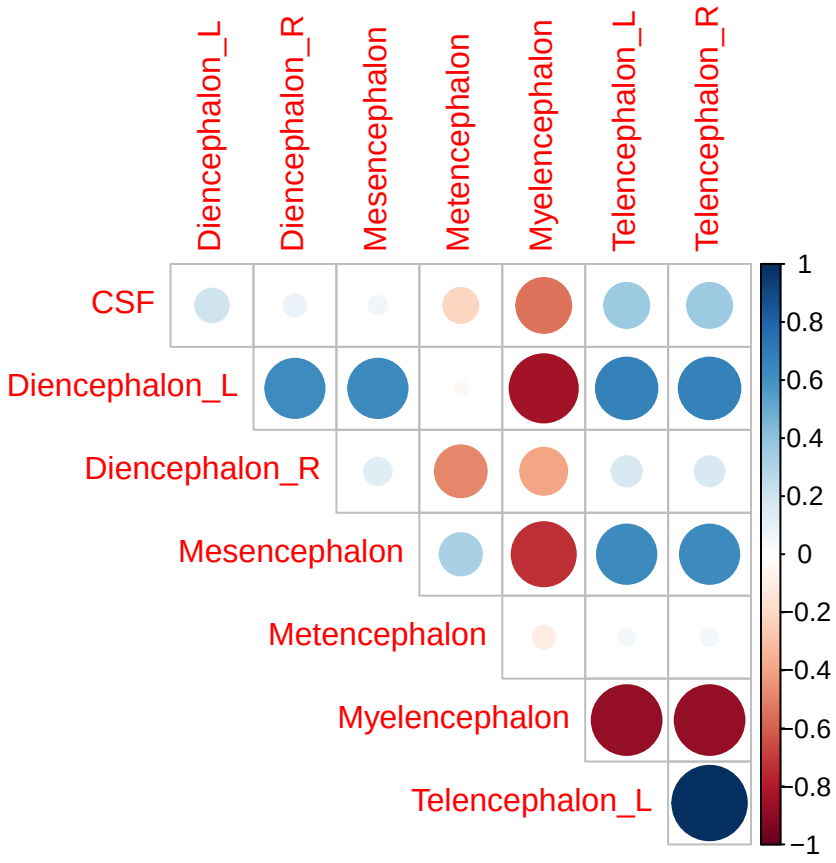

Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

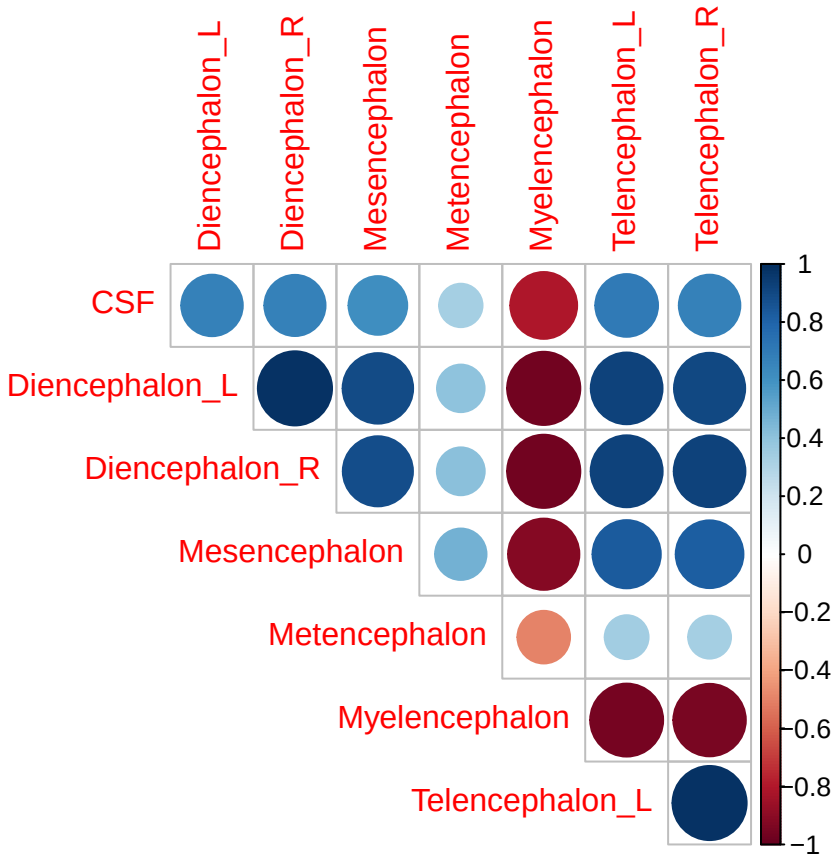

Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

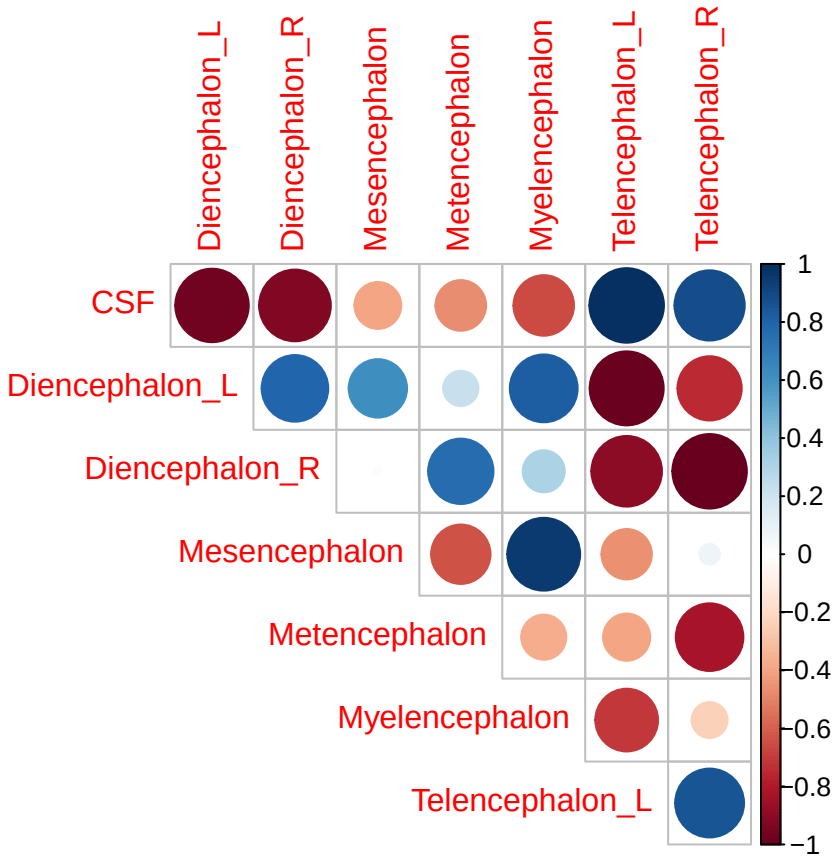



```
# correlation analysis across age
# gender = Male

library(compositions)
library(corrplot)

dat_male = list()
res.male = list()
for ( i in Age){
  dat_male[[i]] = filter(dat_cor$level1, Gender == 'M', Age == i) %>% select(-c(rawid,Age,Gender))
  res.male[[i]] = acomp(dat_male[[i]])

  A=  corrplot(cor(res.male[[i]]), type="upper",
              # order="hclust",
         # tl.srt=30,
         # title = 'Male',
         diag=FALSE)
}
```

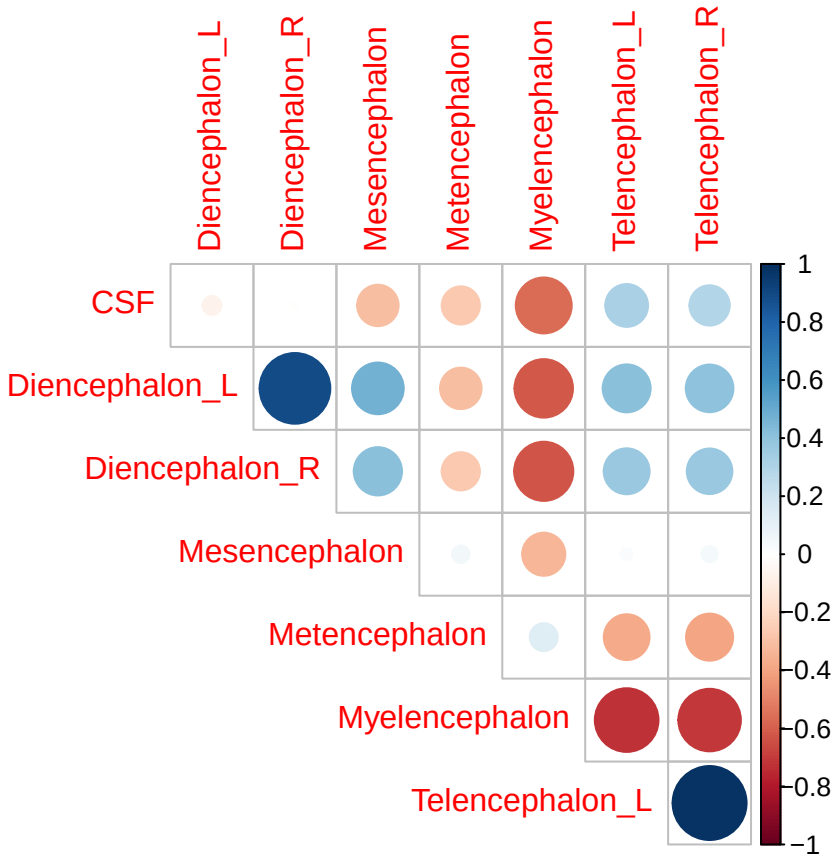

Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

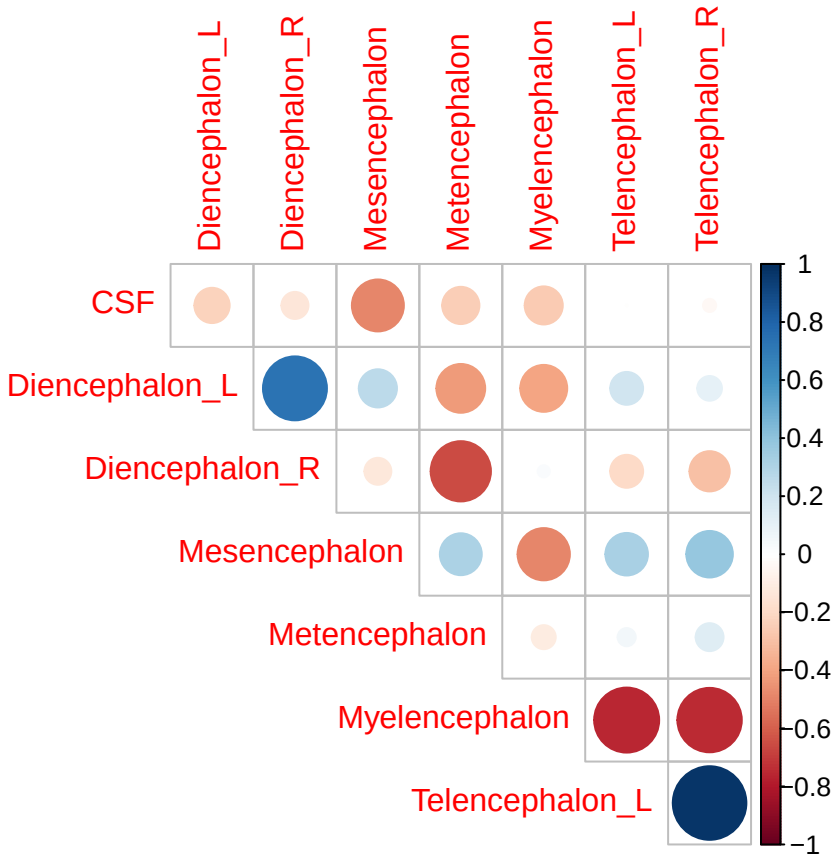

Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

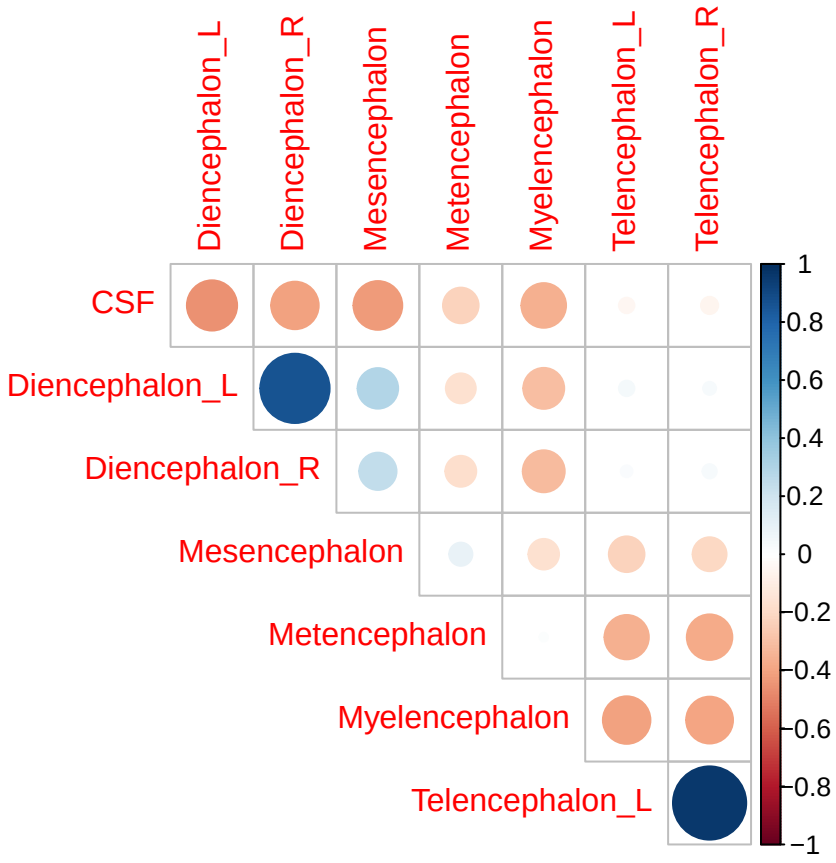

Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

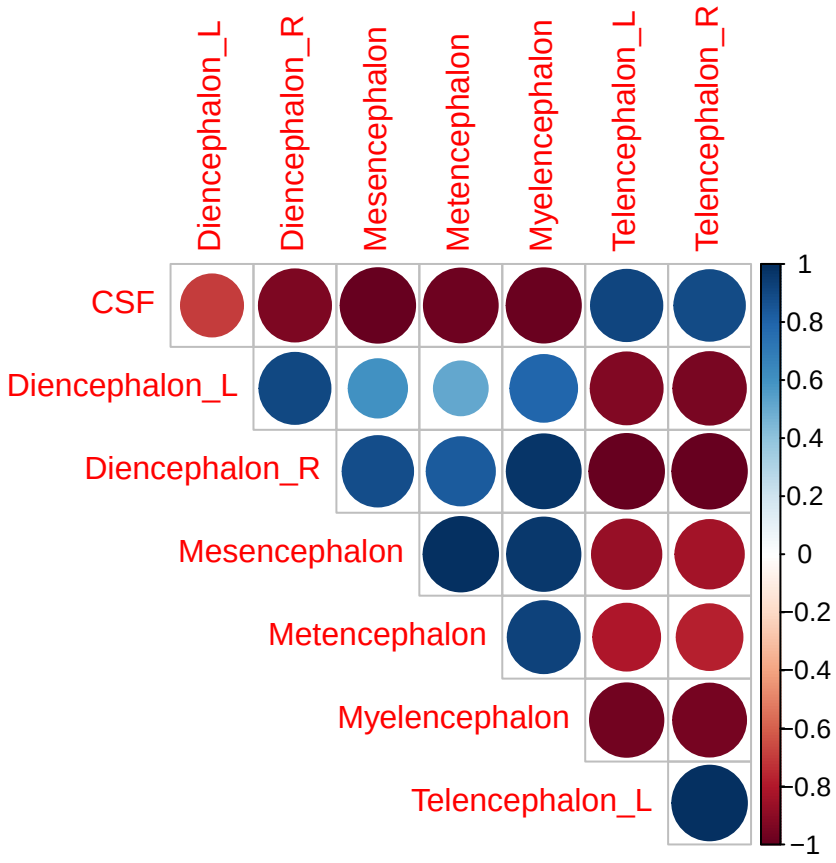



```
# correlation analysis across age
# gender = Female

dat_female = list()
res.female = list()
for ( i in Age){
  dat_female[[i]] = filter(dat_cor$level1, Gender == 'F', Age == i) %>% select(-c(rawid,Age,Gender))
  res.female[[i]] = acomp(dat_female[[i]])

  A =   corrplot(cor(res.female[[i]]), type="upper",
             # order="hclust",
         # tl.srt=30,
         # title = 'Female',
         diag=FALSE)
}
```


Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1


Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1


Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1


Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
Telencephalon_R
CSF
Diencephalon_L
Diencephalon_R
Mesencephalon
Metencephalon
Myelencephalon
Telencephalon_L
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1