multivariada03

Panosso AR & Oliveira JA

2023-07-12

**DESAFIO 1**. Promover o acesso universal à alimentação adequada e saudável, com prioridade para as famílias e pessoas em situação de insegurança alimentar e nutricional

Temas: Transferência de Renda; Alimentação escolar; Distribuição de alimentos.

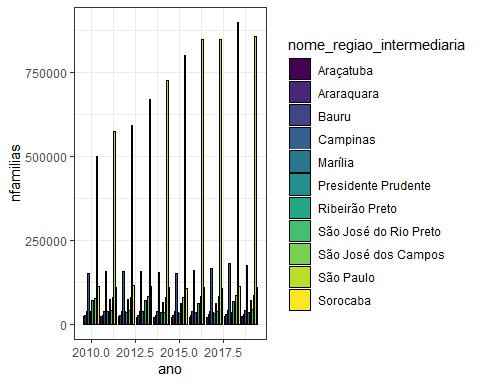
1.1 Programa Bolsa Família  
1.1.1 Pessoas Beneficiárias **(OK)**  
1.1.2 Famílias Beneficiárias **(OK)**  
1.1.3 Valor Pago **(OK)**

### Base de dados - “im”

# Carregando a base de dados  
im <- read\_rds("data/im.rds") %>%   
 select(nome,id\_municipio:valor\_pago\_pbf)  
  
im <- left\_join(im,  
 df\_nomes %>%   
 select(nome, nome\_regiao\_intermediaria),  
 by = "nome") %>%   
 relocate(nome, nome\_regiao\_intermediaria)  
  
# Média foi utilizada para agrupamento dos meses dentro do ano  
data\_set <- im %>% filter(ano >= 2010) %>%   
 group\_by(nome, ano) %>%   
 summarise(  
 familias\_beneficiarias\_pbf = mean(familias\_beneficiarias\_pbf,na.rm=TRUE),  
 pessoas\_beneficiarias\_pbf = mean(pessoas\_beneficiarias\_pbf,na.rm=TRUE),  
 valor\_pago\_pbf = mean(valor\_pago\_pbf,na.rm=TRUE)  
 )

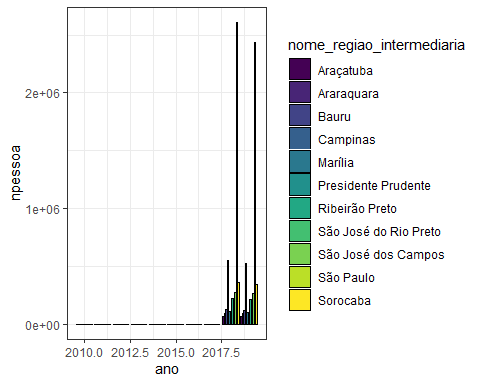
### Visualização de dados

# soma para cada ano das cidades  
tab\_im <- im %>% filter(ano >= 2010) %>%   
 group\_by(nome\_regiao\_intermediaria, nome, ano) %>%   
 summarise(  
 familias\_beneficiarias\_pbf = mean(familias\_beneficiarias\_pbf,na.rm=TRUE),  
 pessoas\_beneficiarias\_pbf = mean(pessoas\_beneficiarias\_pbf,na.rm=TRUE),  
 valor\_pago\_pbf = mean(valor\_pago\_pbf,na.rm=TRUE)  
 ) %>%   
 group\_by(ano, nome\_regiao\_intermediaria) %>% filter(ano >=2010, ano < 2020) %>%   
 summarise(  
 nfamilias = sum(familias\_beneficiarias\_pbf, na.rm=TRUE),  
 npessoa = sum(pessoas\_beneficiarias\_pbf, na.rm=TRUE),  
 valorpago = sum(valor\_pago\_pbf, na.rm=TRUE),  
 )  
  
tab\_im %>%   
 ggplot(aes(x=ano, y=nfamilias, fill=nome\_regiao\_intermediaria)) +  
 geom\_col(col="black", position = "dodge") +  
 scale\_fill\_viridis\_d()

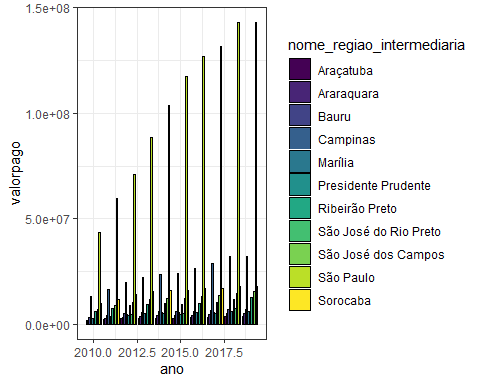


# geom\_point()+  
 # geom\_line()

tab\_im %>%   
 ggplot(aes(x=ano, y=npessoa, fill=nome\_regiao\_intermediaria)) +  
 geom\_col(col="black", position = "dodge") +  
 scale\_fill\_viridis\_d()



tab\_im %>%   
 ggplot(aes(x=ano, y=valorpago, fill=nome\_regiao\_intermediaria)) +  
 geom\_col(col="black", position = "dodge") +  
 scale\_fill\_viridis\_d()

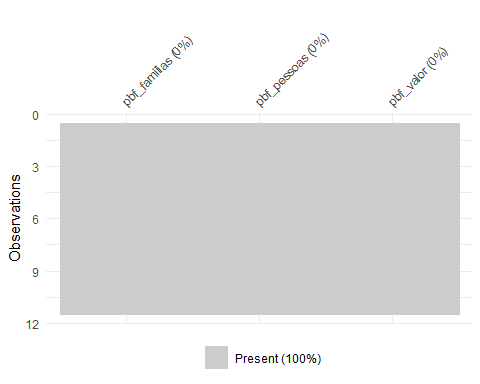


#### Criando a base resumo

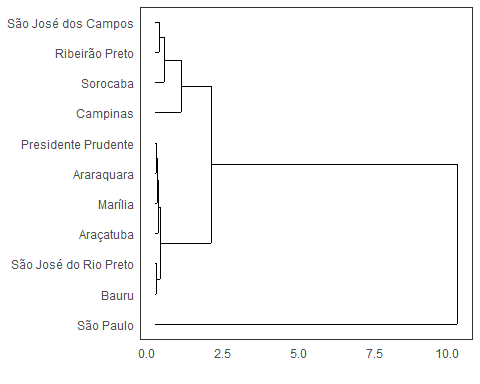
data\_set\_muni <- tab\_im %>%   
 filter(ano>= 2015, ano < 2020) %>%   
 group\_by(nome\_regiao\_intermediaria) %>%   
 summarise(  
 pbf\_familias = sum(nfamilias, na.rm=TRUE),  
 pbf\_pessoas = sum(npessoa, na.rm=TRUE),  
 pbf\_valor = sum(valorpago, na.rm=TRUE)  
 )

#### Agrupamento e Correlação

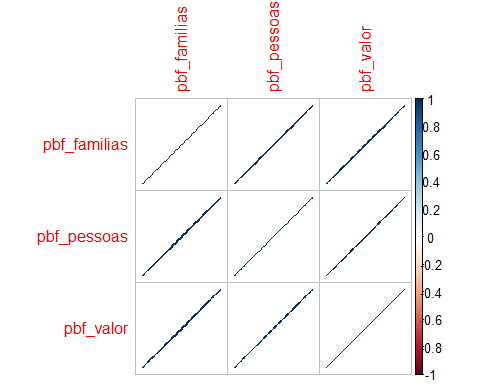
nome <- data\_set\_muni$nome\_regiao\_intermediaria  
da <- data\_set\_muni  
da\_pad <- decostand(da[-1] ,   
 method = "standardize",  
 na.rm=TRUE)  
df <- da\_pad  
visdat::vis\_miss(da\_pad)



row.names(df) <- nome  
da\_pad\_euc<-vegdist(df,"euclidean",na.rm=TRUE)   
da\_pad\_euc\_ward<-hclust(da\_pad\_euc, method="ward.D")  
grupo<-cutree(da\_pad\_euc\_ward,3)  
n\_grupo <- length(unique(grupo))  
ggdendrogram(da\_pad\_euc\_ward, rotate = TRUE, size = 2)

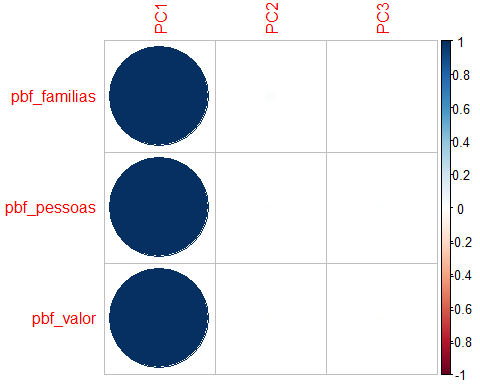


cor\_matrix <- cor(df,use = "na.or.complete")  
corrplot(cor\_matrix, method="ellipse")

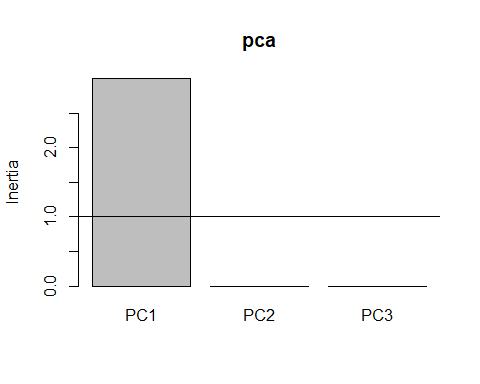


* Variáveis altamente correlacionadas, apenas uma delas representa bem esse banco de dados

pca <- prcomp(df,  
 scale=T)  
# Autovalores  
eig<-pca$sdev^2  
print(round(eig,3))  
#> [1] 3 0 0  
ve<-eig/sum(eig)  
print(round(ve,4))  
#> [1] 0.9999 0.0001 0.0000  
print(round(cumsum(ve),4)\*100)  
#> [1] 99.99 100.00 100.00  
mcor<-cor(df,pca$x)  
corrplot(mcor)

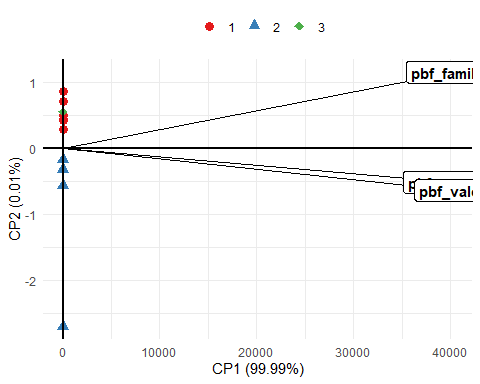


screeplot(pca)  
abline(h=1)



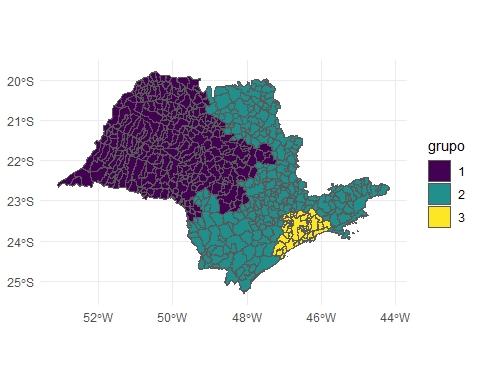
pc1V<-cor(df,pca$x)[,1]/sd(cor(df,pca$x)[,1])  
pc2V<-cor(df,pca$x)[,2]/sd(cor(df,pca$x)[,2])  
pc3V<-cor(df,pca$x)[,3]/sd(cor(df,pca$x)[,3])  
pc1c<-pca$x[,1]/sd(pca$x[,1])  
pc2c<-pca$x[,2]/sd(pca$x[,2])  
pc3c<-pca$x[,3]/sd(pca$x[,3])  
nv<-ncol(df)

bip<-data.frame(pc1c,pc2c,pc3c,nome,grupo=as.factor(grupo))  
texto <- data.frame(x = pc1V, y = pc2V,z = pc3V,label = names(df)  
)  
  
bip %>%   
 ggplot(aes(x=pc1c, y=pc2c))+  
 geom\_point() +  
 geom\_point(aes(shape = grupo, color = grupo), size = 3) + theme\_minimal()+  
 scale\_shape\_manual(values=16:(15+n\_grupo)) +  
 scale\_colour\_brewer(palette = "Set1") +  
 geom\_vline(aes(xintercept=0),  
 color="black", size=1) +  
 geom\_hline(aes(yintercept=0),  
 color="black", size=1) +  
 annotate(geom="segment",  
 x=rep(0,length(df)),  
 xend=texto$x,  
 y=rep(0,length(df)),  
 yend=texto$y,color="black",lwd=.5) +  
 geom\_label(data=texto,aes(x=x,y=y,label=label),  
 color="black",angle=0,fontface="bold",size=4,fill="white") +  
 labs(x=paste("CP1 (",round(100\*ve[1],2),"%)",sep=""),  
 y=paste("CP2 (",round(100\*ve[2],2),"%)",sep=""),  
 color="",shape="") +  
 theme(legend.position = "top")



ck<-sum(pca$sdev^2>=1)  
tabelapca<-vector()  
for( l in 1:ck) tabelapca<-cbind(tabelapca,mcor[,l])  
colnames(tabelapca)<-paste(rep(c("PC"),ck),1:ck,sep="")  
pcat<-round(tabelapca,3)  
tabelapca<-tabelapca[order(abs(tabelapca[,1])),]  
print(tabelapca)  
#> pbf\_familias pbf\_valor pbf\_pessoas   
#> 0.9999170 0.9999566 0.9999627

df\_grupo <- data.frame(nome\_regiao\_intermediaria = nome, grupo)  
  
df\_grupo <- left\_join(df\_nomes %>%   
 select(id\_municipio, nome, nome\_regiao\_intermediaria),  
 df\_grupo,  
 by ="nome\_regiao\_intermediaria")  
  
d\_sf\_municipio <- st\_read("shp/35MUE250GC\_SIR.shp", quiet = TRUE)  
  
d\_sf\_municipio <- d\_sf\_municipio %>%  
 rename(id\_municipio = CD\_GEOCMU) %>%  
 inner\_join(df\_grupo %>%  
 relocate(id\_municipio), "id\_municipio") %>%   
 mutate(grupo = as.factor(grupo))  
  
ggplot(d\_sf\_municipio) +  
 geom\_sf(aes(fill = grupo))+  
 theme\_minimal() +  
 scale\_fill\_viridis\_d()



for(i in 1:n\_grupo){  
 print("===============")  
 print(paste0("Grupo ",i,""))  
 print("===============")  
 print(paste(nome[grupo == i],collapse = "; "))  
 cat("\n")  
}  
#> [1] "==============="  
#> [1] "Grupo 1"  
#> [1] "==============="  
#> [1] "Araraquara; Araçatuba; Bauru; Marília; Presidente Prudente; São José do Rio Preto"  
#>   
#> [1] "==============="  
#> [1] "Grupo 2"  
#> [1] "==============="  
#> [1] "Campinas; Ribeirão Preto; Sorocaba; São José dos Campos"  
#>   
#> [1] "==============="  
#> [1] "Grupo 3"  
#> [1] "==============="  
#> [1] "São Paulo"

**DESAFIO 2**. Combater a insegurança alimentar e nutricional e promover a inclusão produtiva rural em grupos populacionais específicos, com ênfase em povos e comunidades tradicionais e outros grupos sociais vulneráveis no meio rural.

Temas: Insegurança alimentar e nutricional; Inclusão produtiva rural; Acesso à terra e gestão territorial; Biodiversidade; Saúde indígena; extrativistas e ribeirinhos; Acesso a políticas públicas.

*MACRO DESAFIO*: Promoção de Sistemas Alimentares Saudáveis e Sustentáveis Desafios: o conjunto dos desafios 3, 4 e 5 contemplam este macro desafio.

**DESAFIO 3**. Promover a produção de alimentos saudáveis e sustentáveis, a estruturação da agricultura familiar e o fortalecimento de sistemas de produção de base agroecológica

Temas: Fortalecimento da agricultura familiar; Reforma agrária; Transição agroecológica; Mulheres; Juventude; Sementes; Mudanças climáticas.

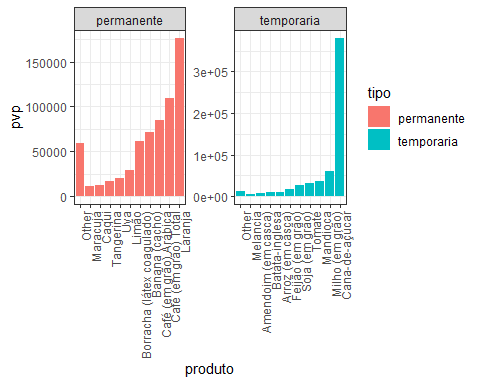
3.1 Produção de alimentos  
3.2 Área Plantada **(OK)**  
3.2.1 Área Plantada Agricultura Familiar  
3.2.2 Área Plantada Agricultura não familiar  
3.3. Rendimento Médio **(OK)**  
3.4 Proporção Área Colhida **(OK)**  
3.5 Proporção Valor da Produção **(OK)**

### Base de dados “lavouras”

lavoura\_permanente <- read\_rds("data/permanente.rds")   
lavoura\_temporaria <- read\_rds("data/temporaria.rds")  
lavouras <- rbind(lavoura\_permanente %>%   
 mutate(tipo = "permanente"),  
 lavoura\_temporaria %>%   
 mutate(tipo = "temporaria")  
)  
lavouras$na <- apply(  
 lavouras[5:12],  
 1,   
 function(x) sum(is.na(x))  
)

### Visualização de dados

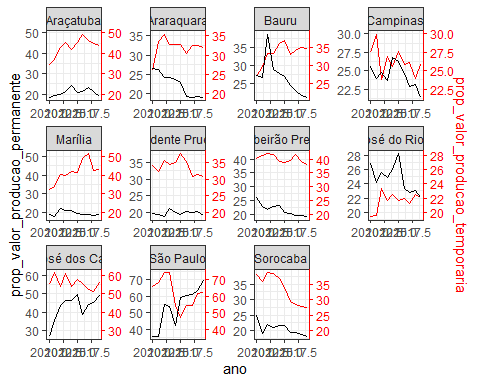
lavouras %>%   
 filter(na != 8, ano >= 2010) %>%   
 arrange(ano) %>%   
 relocate(tipo) %>%   
 group\_by(tipo, produto) %>%   
 summarise(pvp = sum(prop\_valor\_producao,na.rm = TRUE)) %>%   
 mutate(produto = produto %>% fct\_lump(n=10, w=pvp) %>% fct\_reorder(pvp)) %>%   
 ggplot(aes(y=pvp,x=produto,fill=tipo)) +  
 geom\_col()+  
 facet\_wrap(~tipo,scales = "free") +   
 theme\_bw() +  
 theme(axis.text.x = element\_text(angle = 90, hjust=1))



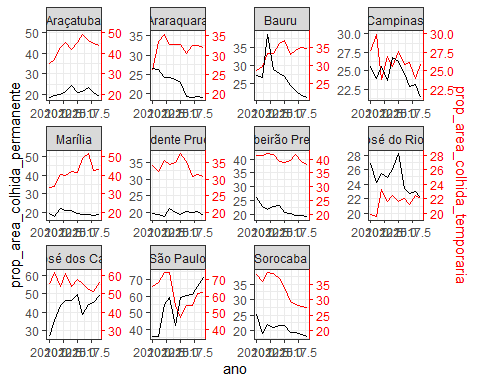
lavouras <- lavouras %>%   
 filter(na != 8, ano >= 2010,  
 produto != "Café (em grão) Arábica",  
 produto != "Café (em grão) Canephora") %>%   
 select(-na)  
  
lavouras <- lavouras %>%   
 group\_by(tipo, ano, nome) %>%   
 summarise(  
 n\_produtos = n(),  
 area\_plantada = sum(area\_plantada,na.rm=TRUE),  
 rendimento\_medio = mean(rendimento\_medio,na.rm=TRUE),  
 prop\_area\_colhida = mean(prop\_area\_colhida,na.rm=TRUE) ,  
 prop\_valor\_producao = mean(prop\_valor\_producao,na.rm=TRUE)  
 ) %>%   
 pivot\_wider(names\_from = tipo,  
 values\_from = n\_produtos:prop\_valor\_producao)  
  
data\_set <- left\_join(  
 data\_set,  
 lavouras,  
 by = c("ano","nome")  
)  
  
lavouras <- left\_join(lavouras, df\_nomes %>%   
 select(nome, nome\_regiao\_intermediaria),  
 by = "nome") %>%   
 relocate(ano, nome, nome\_regiao\_intermediaria)

coeff <-1  
tabe\_lavouras <- lavouras %>%   
 group\_by(ano, nome\_regiao\_intermediaria) %>%  
 summarise(  
 n\_produtos\_permanente =   
 mean(n\_produtos\_permanente,na.rm=TRUE),  
 n\_produtos\_temporaria =   
 mean(n\_produtos\_temporaria,na.rm=TRUE),  
 area\_plantada\_permanente =   
 mean(area\_plantada\_permanente,na.rm=TRUE),  
 area\_plantada\_temporaria =   
 mean(area\_plantada\_temporaria,na.rm=TRUE),  
 rendimento\_medio\_permanente =   
 mean(rendimento\_medio\_permanente,na.rm=TRUE),  
 rendimento\_medio\_temporaria =   
 mean(rendimento\_medio\_temporaria,na.rm=TRUE),  
 prop\_area\_colhida\_permanente =  
 mean(prop\_area\_colhida\_permanente,na.rm=TRUE),  
 prop\_area\_colhida\_temporaria =  
 mean(prop\_area\_colhida\_temporaria,na.rm=TRUE),  
 prop\_valor\_producao\_permanente =  
 mean(prop\_valor\_producao\_permanente,na.rm=TRUE),  
 prop\_valor\_producao\_temporaria =  
 mean(prop\_valor\_producao\_temporaria,na.rm=TRUE)  
 )

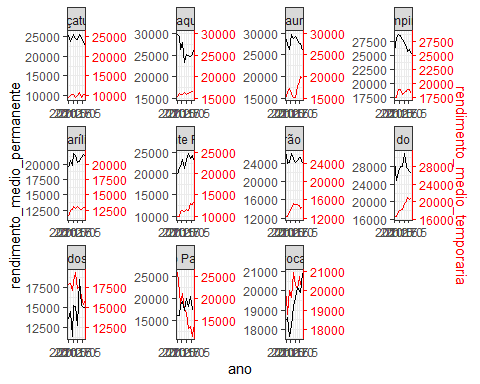
tabe\_lavouras %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = prop\_valor\_producao\_permanente), color="red") +  
 geom\_line(aes(y = prop\_valor\_producao\_temporaria/coeff)) +  
 scale\_y\_continuous(  
 name = "prop\_valor\_producao\_permanente",  
 sec.axis = sec\_axis(~.\*coeff, name="prop\_valor\_producao\_temporaria")  
 ) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red"))+  
 facet\_wrap(~nome\_regiao\_intermediaria, scales = "free", ncol=4)



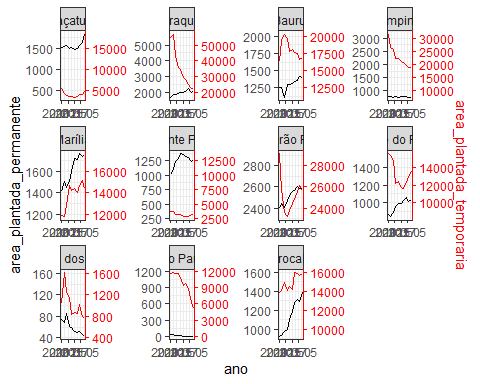
tabe\_lavouras %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = prop\_area\_colhida\_permanente), color="red") +  
 geom\_line(aes(y = prop\_area\_colhida\_temporaria/coeff)) +  
 scale\_y\_continuous(  
 name = "prop\_area\_colhida\_permanente",  
 sec.axis = sec\_axis(~.\*coeff, name="prop\_area\_colhida\_temporaria")  
 ) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red")) +  
 facet\_wrap(~nome\_regiao\_intermediaria, scales = "free", ncol=4)



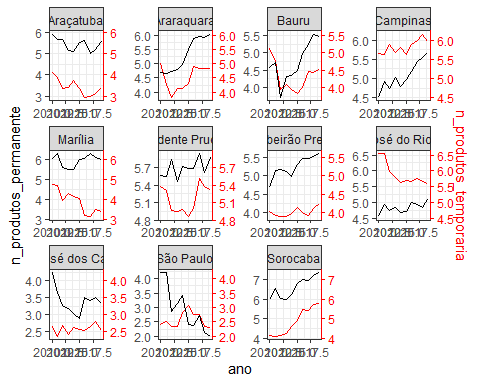
tabe\_lavouras %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = rendimento\_medio\_permanente), color="red") +  
 geom\_line(aes(y = rendimento\_medio\_temporaria/coeff)) +  
 scale\_y\_continuous(  
 name = "rendimento\_medio\_permanente",  
 sec.axis = sec\_axis(~.\*coeff, name="rendimento\_medio\_temporaria")  
 ) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red"))+  
 facet\_wrap(~nome\_regiao\_intermediaria, scales = "free")



tabe\_lavouras %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = area\_plantada\_permanente), color="red") +  
 geom\_line(aes(y = area\_plantada\_temporaria/10)) +  
 scale\_y\_continuous(  
 name = "area\_plantada\_permanente",  
 sec.axis = sec\_axis(~.\*10, name="area\_plantada\_temporaria")  
 ) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red")) +  
 facet\_wrap(~nome\_regiao\_intermediaria, scales = "free", ncol=4)



tabe\_lavouras %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = n\_produtos\_permanente), color="red") +  
 geom\_line(aes(y = n\_produtos\_temporaria/coeff)) +  
 scale\_y\_continuous(  
 name = "n\_produtos\_permanente",  
 sec.axis = sec\_axis(~.\*coeff, name="n\_produtos\_temporaria")  
 ) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red")) +  
 facet\_wrap(~nome\_regiao\_intermediaria, scales = "free", ncol=4)

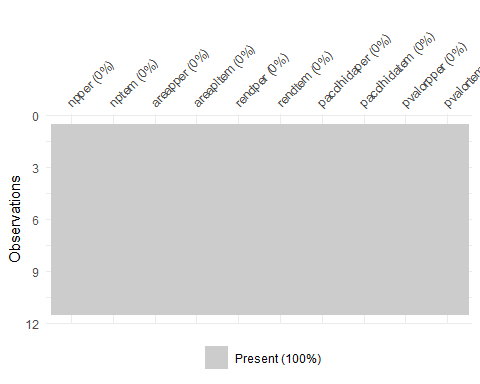


#### Criando a base resumo

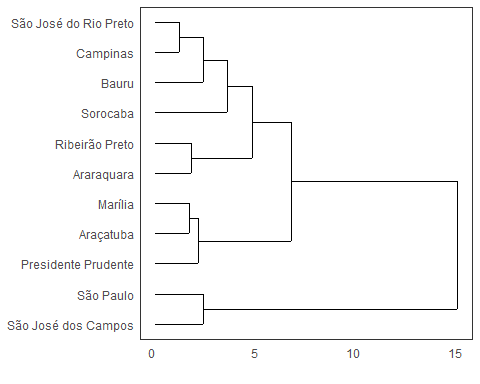
lavouras\_resu <- lavouras %>%   
 group\_by(nome\_regiao\_intermediaria) %>%   
 summarize(  
 npper = mean(n\_produtos\_permanente,na.rm=TRUE),   
 nptem = mean(n\_produtos\_temporaria,na.rm=TRUE),   
 areapper = mean(area\_plantada\_permanente,na.rm=TRUE),   
 areapltem = mean(area\_plantada\_temporaria,na.rm=TRUE),   
 rendper = mean(rendimento\_medio\_permanente,na.rm=TRUE),   
 rendtem = mean(rendimento\_medio\_temporaria,na.rm=TRUE),   
 pacolhidaper = mean(prop\_area\_colhida\_permanente,na.rm=TRUE),   
 pacolhidatem = mean(prop\_area\_colhida\_temporaria,na.rm=TRUE),   
 pvalorpper = mean(prop\_valor\_producao\_permanente,na.rm=TRUE),   
 pvalortem = mean(prop\_valor\_producao\_temporaria,na.rm=TRUE)   
 )

#### Agrupamento e Correlação

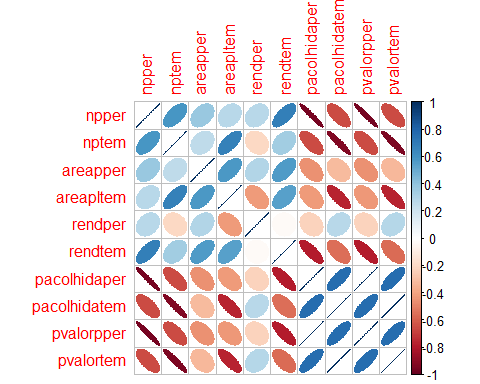
da <- lavouras\_resu  
da <- da %>%   
 mutate(across(where(is.numeric), function(x) {  
 med <- mean(x,na.rm=TRUE)  
 x[is.na(x)] <- 0  
 x  
 }))  
  
da\_pad <- decostand(da[-1] ,   
 method = "standardize",  
 na.rm=TRUE)  
df <- da\_pad %>%   
 drop\_na()  
nome <- da %>% drop\_na() %>% pull(nome\_regiao\_intermediaria)  
visdat::vis\_miss(da\_pad)



row.names(df) <- nome  
da\_pad\_euc<-vegdist(df,"euclidean",na.rm=TRUE)   
da\_pad\_euc\_ward<-hclust(da\_pad\_euc, method="ward.D")  
grupo<-cutree(da\_pad\_euc\_ward,3)  
n\_grupo <- length(unique(grupo))  
ggdendrogram(da\_pad\_euc\_ward, rotate = TRUE, size = 2)

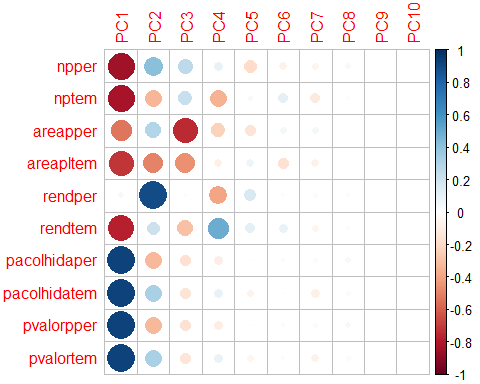


cor\_matrix <- cor(df,use = "na.or.complete")  
corrplot(cor\_matrix, method="ellipse")

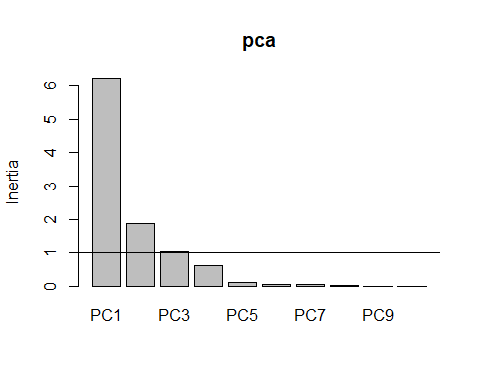


* Variáveis altamente correlacionadas, apenas uma delas representa bem esse banco de dados

pca <- prcomp(df,  
 scale=T)  
# Autovalores  
eig<-pca$sdev^2  
print(round(eig,3))  
#> [1] 6.218 1.889 1.047 0.629 0.110 0.055 0.045 0.008 0.000 0.000  
ve<-eig/sum(eig)  
print(round(ve,4))  
#> [1] 0.6218 0.1889 0.1047 0.0629 0.0110 0.0055 0.0045 0.0008 0.0000 0.0000  
print(round(cumsum(ve),4)\*100)  
#> [1] 62.18 81.07 91.54 97.82 98.92 99.47 99.92 100.00 100.00 100.00  
mcor<-cor(df,pca$x)  
corrplot(mcor)

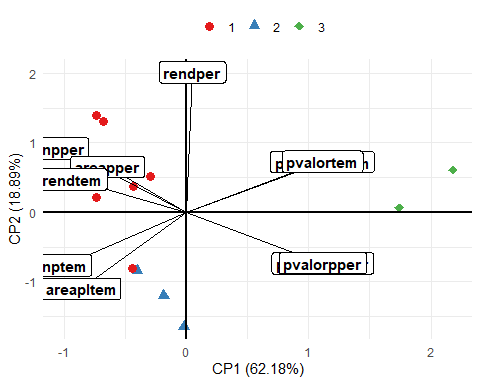


screeplot(pca)  
abline(h=1)



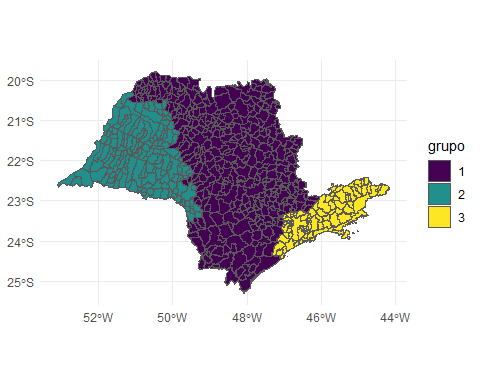
pc1V<-cor(df,pca$x)[,1]/sd(cor(df,pca$x)[,1])  
pc2V<-cor(df,pca$x)[,2]/sd(cor(df,pca$x)[,2])  
pc3V<-cor(df,pca$x)[,3]/sd(cor(df,pca$x)[,3])  
pc1c<-pca$x[,1]/sd(pca$x[,1])  
pc2c<-pca$x[,2]/sd(pca$x[,2])  
pc3c<-pca$x[,3]/sd(pca$x[,3])  
nv<-ncol(df)

bip<-data.frame(pc1c,pc2c,pc3c,nome,grupo=as.factor(grupo))  
texto <- data.frame(x = pc1V, y = pc2V,z = pc3V,label = names(df)  
)  
  
bip %>%   
 ggplot(aes(x=pc1c, y=pc2c))+  
 geom\_point() +  
 geom\_point(aes(shape = grupo, color = grupo), size = 3) + theme\_minimal()+  
 scale\_shape\_manual(values=16:(15+n\_grupo)) +  
 scale\_colour\_brewer(palette = "Set1") +  
 geom\_vline(aes(xintercept=0),  
 color="black", size=1) +  
 geom\_hline(aes(yintercept=0),  
 color="black", size=1) +  
 annotate(geom="segment",  
 x=rep(0,length(df)),  
 xend=texto$x,  
 y=rep(0,length(df)),  
 yend=texto$y,color="black",lwd=.5) +  
 geom\_label(data=texto,aes(x=x,y=y,label=label),  
 color="black",angle=0,fontface="bold",size=4,fill="white") +  
 labs(x=paste("CP1 (",round(100\*ve[1],2),"%)",sep=""),  
 y=paste("CP2 (",round(100\*ve[2],2),"%)",sep=""),  
 color="",shape="") +  
 theme(legend.position = "top")



ck<-sum(pca$sdev^2>=1)  
tabelapca<-vector()  
for( l in 1:ck) tabelapca<-cbind(tabelapca,mcor[,l])  
colnames(tabelapca)<-paste(rep(c("PC"),ck),1:ck,sep="")  
pcat<-round(tabelapca,3)  
tabelapca<-tabelapca[order(abs(tabelapca[,1])),]  
print(tabelapca)  
#> PC1 PC2 PC3  
#> rendper 0.04154814 0.8999120 0.01771652  
#> areapper -0.53671231 0.2930022 -0.74027429  
#> areapltem -0.71716071 -0.4908367 -0.45200322  
#> rendtem -0.77334546 0.2122231 -0.29430966  
#> nptem -0.82716478 -0.3398300 0.22764444  
#> npper -0.84007172 0.4101621 0.26619101  
#> pacolhidatem 0.92358102 0.3238861 -0.14002759  
#> pvalortem 0.92552884 0.3254123 -0.13644855  
#> pacolhidaper 0.92594199 -0.3271964 -0.15323853  
#> pvalorpper 0.92670066 -0.3257764 -0.15149391

df\_grupo <- data.frame(nome\_regiao\_intermediaria = nome, grupo)  
  
df\_grupo <- left\_join(df\_nomes %>%   
 select(id\_municipio, nome, nome\_regiao\_intermediaria),  
 df\_grupo,  
 by ="nome\_regiao\_intermediaria")  
  
d\_sf\_municipio <- st\_read("shp/35MUE250GC\_SIR.shp", quiet = TRUE)  
  
d\_sf\_municipio <- d\_sf\_municipio %>%  
 rename(id\_municipio = CD\_GEOCMU) %>%  
 inner\_join(df\_grupo %>%  
 relocate(id\_municipio), "id\_municipio") %>%   
 mutate(grupo = as.factor(grupo))  
  
ggplot(d\_sf\_municipio) +  
 geom\_sf(aes(fill = grupo))+  
 theme\_minimal() +  
 scale\_fill\_viridis\_d()



for(i in 1:n\_grupo){  
 print("===============")  
 print(paste0("Grupo ",i,""))  
 print("===============")  
 print(paste(nome[grupo == i],collapse = "; "))  
 cat("\n")  
}  
#> [1] "==============="  
#> [1] "Grupo 1"  
#> [1] "==============="  
#> [1] "Araraquara; Bauru; Campinas; Ribeirão Preto; Sorocaba; São José do Rio Preto"  
#>   
#> [1] "==============="  
#> [1] "Grupo 2"  
#> [1] "==============="  
#> [1] "Araçatuba; Marília; Presidente Prudente"  
#>   
#> [1] "==============="  
#> [1] "Grupo 3"  
#> [1] "==============="  
#> [1] "São José dos Campos; São Paulo"

**DESAFIO 4**. Promover o abastecimento e o acesso regular e permanente da população brasileira à alimentação

Temas: Compras públicas; Abastecimento; Legislação sanitária; Economia solidaria; Perdas e desperdícios de alimentos; Equipamentos públicos de SAN; Agricultura urbana.

**4.1 Programa da Alimentação Escolar**  
4.1.1 Número de Alunos Beneficiários **(OK)**  
4.1.2 Valor gasto com Compras da Agricultura Familiar 30%  
4.1.3 Valor Formalizado  
4.1.4 Valor Executado **(OK)**

### Base de dados - “pnae”

essa base precisa ser previamente construída.

pnae\_alunos <- read\_rds("data/pnae\_alunos\_atendidos.rds")  
lista\_etapas <- pnae\_alunos$etapa\_ensino %>% unique()

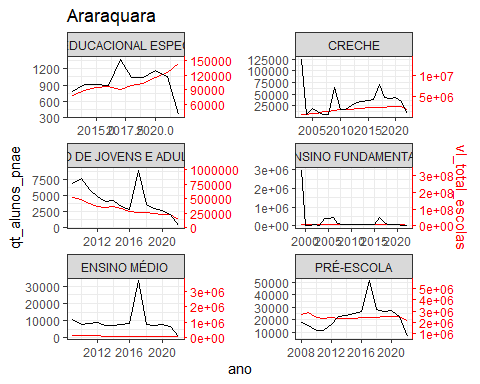
pnae\_recurso <- read\_rds("data/pnae\_recurso.rds")   
pnae\_recurso <- pnae\_recurso %>%   
 mutate(  
 modalidade\_ensino = ifelse(modalidade\_ensino == "EJA",  
 "EDUCAÇÃO DE JOVENS E ADULTOS (EJA)",  
 modalidade\_ensino)  
 )

pnae\_recurso <- pnae\_recurso %>%   
 group\_by(nome, ano, esfera\_governo, modalidade\_ensino) %>%   
 summarise(vl\_total\_escolas = mean(vl\_total\_escolas)) %>%   
 rename(etapa\_ensino = modalidade\_ensino) %>%   
 filter(etapa\_ensino %in% lista\_etapas)  
#visdat::vis\_miss(pnae\_recurso)

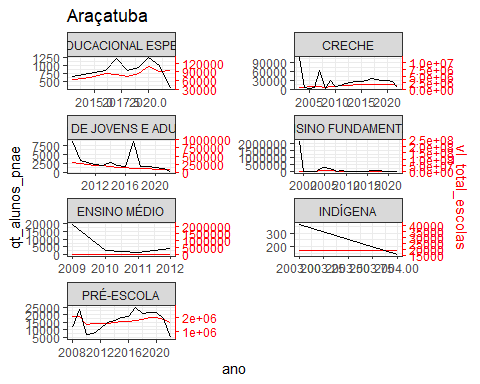
### Visualização de dados

pnae <- left\_join(pnae\_alunos %>% filter(esfera\_governo == "MUNICIPAL") %>%  
 select(ano:qt\_alunos\_pnae,nome),  
 pnae\_recurso,  
 by =c("nome","ano","esfera\_governo","etapa\_ensino"))  
  
pnae <- left\_join(pnae, df\_nomes %>%   
 select(nome, nome\_regiao\_intermediaria),  
 by = "nome")   
  
tab\_pane <- pnae %>%   
 group\_by(ano,nome\_regiao\_intermediaria,etapa\_ensino) %>%   
 summarise(  
 qt\_alunos\_pnae = sum(qt\_alunos\_pnae, na.rm=TRUE),  
 vl\_total\_escolas = sum(vl\_total\_escolas, na.rm=TRUE)  
 )

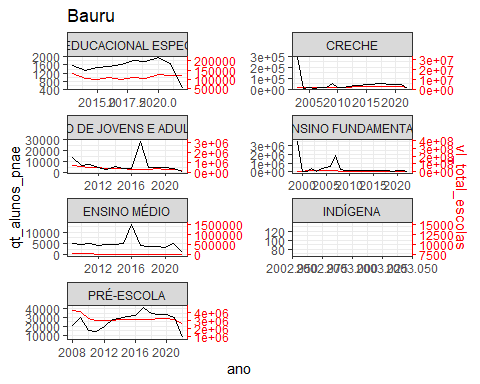
coeff <- 1/.009  
nri <- tab\_pane$nome\_regiao\_intermediaria %>% unique()  
map(nri, ~{tab\_pane %>% filter(nome\_regiao\_intermediaria == .x) %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = qt\_alunos\_pnae), color="red") +  
 geom\_line(aes(y = vl\_total\_escolas/coeff)) +  
 scale\_y\_continuous(  
 name = "qt\_alunos\_pnae",  
 sec.axis = sec\_axis(~.\*coeff, name="vl\_total\_escolas")  
 ) +  
 facet\_wrap(~ etapa\_ensino, scale="free", ncol=2) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red"))+  
 labs(title = .x)})  
#> [[1]]



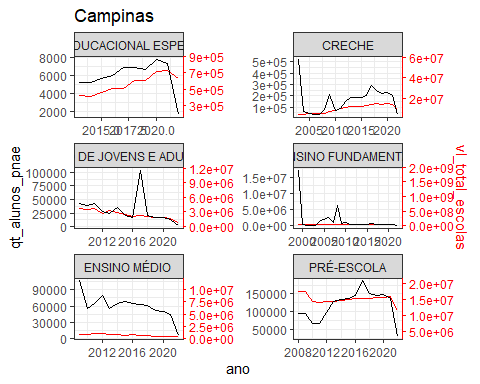
#>   
#> [[2]]



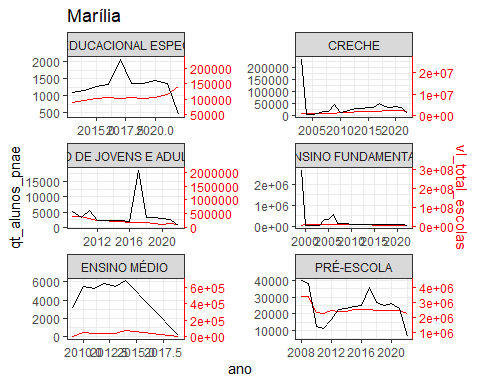
#>   
#> [[3]]



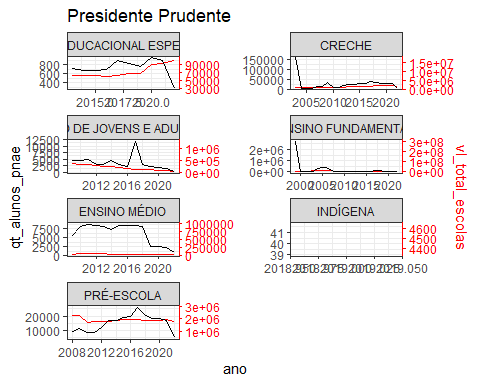
#>   
#> [[4]]



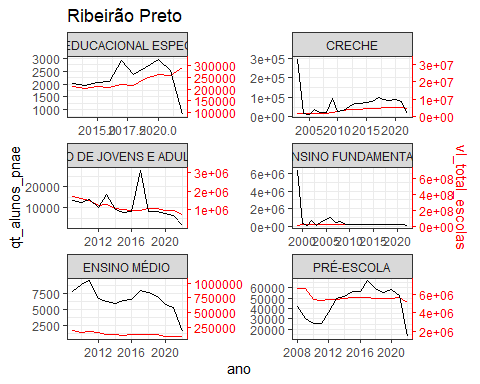
#>   
#> [[5]]



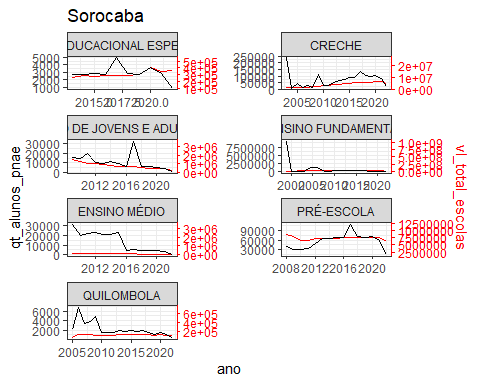
#>   
#> [[6]]



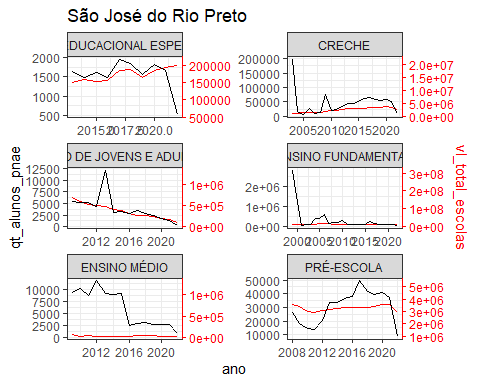
#>   
#> [[7]]



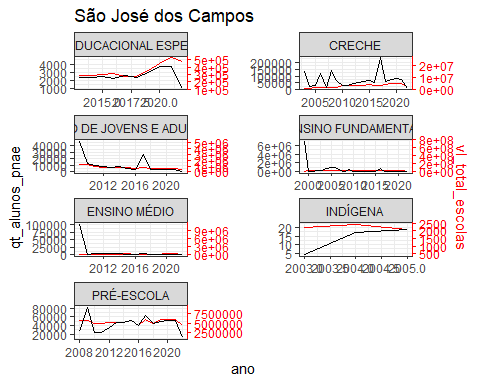
#>   
#> [[8]]



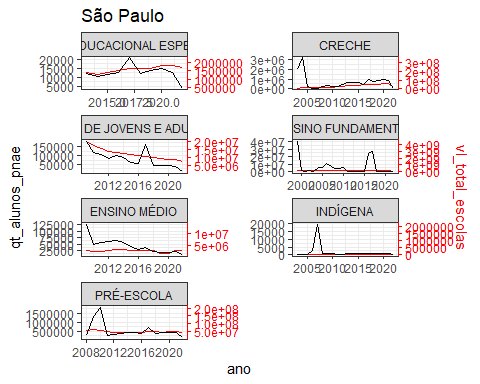
#>   
#> [[9]]



#>   
#> [[10]]



#>   
#> [[11]]



#### Criando a base resumo

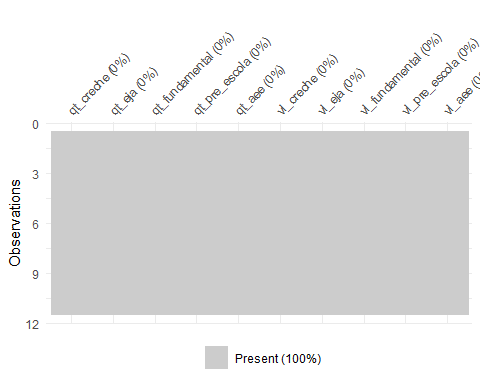
data\_set <- left\_join(data\_set,   
 pnae %>%   
 drop\_na() %>%   
 group\_by(ano, nome) %>%   
 summarise(  
 qt\_alunos\_pnae = sum(qt\_alunos\_pnae,na.rm=TRUE),  
 vl\_total\_escolas = sum(vl\_total\_escolas,na.rm=TRUE)  
 ), by=c("ano","nome"))

pnae\_resumo <- pnae %>%   
 # filter(ano >= 2010, ano <= 2020) %>%   
 # drop\_na() %>%   
 group\_by(ano, nome\_regiao\_intermediaria, etapa\_ensino) %>%   
 summarise(  
 qt\_alunos\_pnae = sum(qt\_alunos\_pnae,na.rm=TRUE),  
 vl\_total\_escolas = sum(vl\_total\_escolas,na.rm=TRUE)  
 ) %>%   
 pivot\_wider(  
 names\_from = etapa\_ensino,  
 values\_from = c(qt\_alunos\_pnae,vl\_total\_escolas)) %>%   
 group\_by(nome\_regiao\_intermediaria) %>%   
 mutate(  
 across(qt\_alunos\_pnae\_CRECHE:`vl\_total\_escolas\_ATENDIMENTO EDUCACIONAL ESPECIALIZADO (AEE)`, function(x) mean(x,na.rm=TRUE))) %>%   
 janitor::clean\_names() %>%   
 filter(ano == 2007) %>%   
 select(-ano)

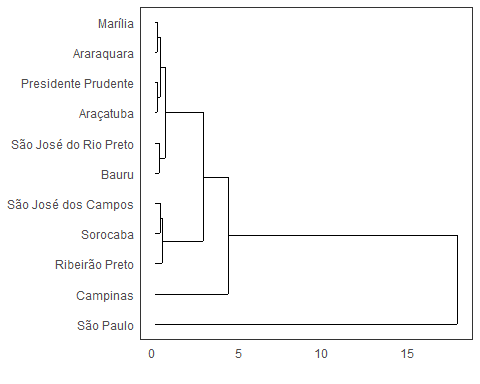
pnae\_resumo <- pnae\_resumo %>%   
 select(-qt\_alunos\_pnae\_ensino\_medio,  
 -qt\_alunos\_pnae\_indigena,  
 -qt\_alunos\_pnae\_quilombola,  
 -vl\_total\_escolas\_ensino\_medio,  
 -vl\_total\_escolas\_indigena,  
 -vl\_total\_escolas\_quilombola  
 )  
nomes\_antigos <- names(pnae\_resumo)  
nomes\_novos <- c(  
 "nome", "qt\_creche", "qt\_eja", "qt\_fundamental",   
 "qt\_pre\_escola", "qt\_aee", "vl\_creche", "vl\_eja",   
 "vl\_fundamental", "vl\_pre\_escola", "vl\_aee")  
pnae\_resumo <- pnae\_resumo %>%   
 rename\_at(vars(nomes\_antigos),~nomes\_novos)

#### Agrupamento e Correlação

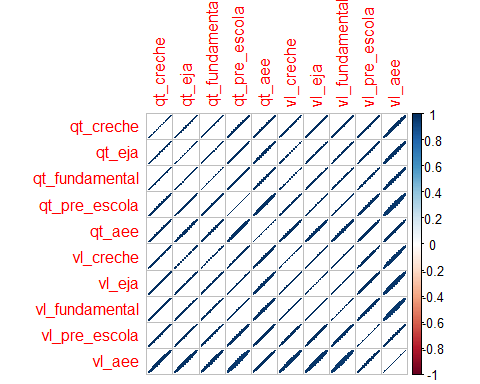
nome <- pnae\_resumo$nome  
da <- pnae\_resumo  
da\_pad <- decostand(da[-1] ,   
 method = "standardize",  
 na.rm=TRUE)  
df <- da\_pad  
visdat::vis\_miss(da\_pad)



row.names(df) <- nome  
da\_pad\_euc<-vegdist(df,"euclidean",na.rm=TRUE)   
da\_pad\_euc\_ward<-hclust(da\_pad\_euc, method="ward.D")  
grupo<-cutree(da\_pad\_euc\_ward,3)  
n\_grupo <- length(unique(grupo))  
ggdendrogram(da\_pad\_euc\_ward, rotate = TRUE, size = 2)

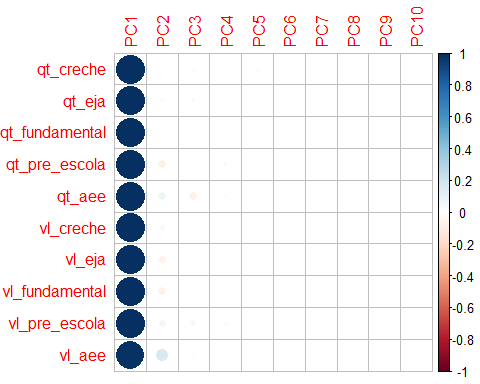


cor\_matrix <- cor(df,use = "na.or.complete")  
corrplot(cor\_matrix, method="ellipse")

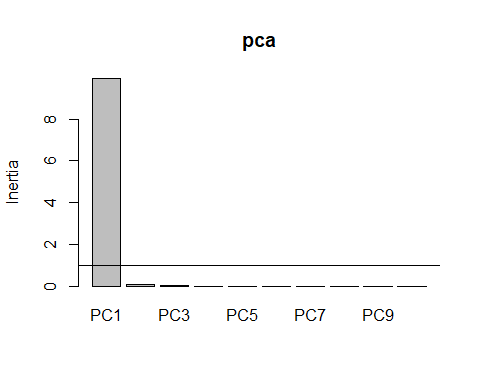


* Variáveis altamente correlacionadas, apenas uma delas representa bem esse banco de dados

pca <- prcomp(df,  
 scale=T)  
# Autovalores  
eig<-pca$sdev^2  
print(round(eig,3))  
#> [1] 9.929 0.056 0.010 0.002 0.001 0.001 0.000 0.000 0.000 0.000  
ve<-eig/sum(eig)  
print(round(ve,4))  
#> [1] 0.9929 0.0056 0.0010 0.0002 0.0001 0.0001 0.0000 0.0000 0.0000 0.0000  
print(round(cumsum(ve),4)\*100)  
#> [1] 99.29 99.86 99.96 99.98 99.99 100.00 100.00 100.00 100.00 100.00  
mcor<-cor(df,pca$x)  
corrplot(mcor)

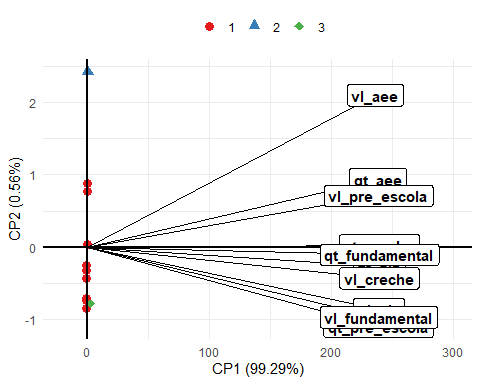


screeplot(pca)  
abline(h=1)



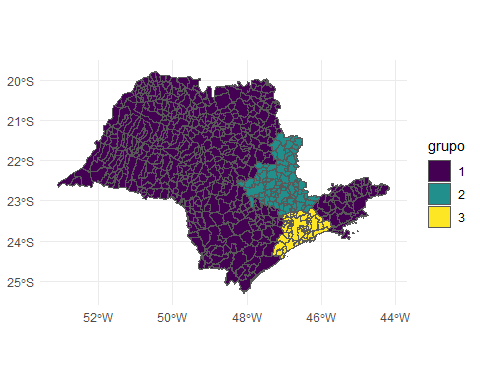
pc1V<-cor(df,pca$x)[,1]/sd(cor(df,pca$x)[,1])  
pc2V<-cor(df,pca$x)[,2]/sd(cor(df,pca$x)[,2])  
pc3V<-cor(df,pca$x)[,3]/sd(cor(df,pca$x)[,3])  
pc1c<-pca$x[,1]/sd(pca$x[,1])  
pc2c<-pca$x[,2]/sd(pca$x[,2])  
pc3c<-pca$x[,3]/sd(pca$x[,3])  
nv<-ncol(df)

bip<-data.frame(pc1c,pc2c,pc3c,nome,grupo=as.factor(grupo))  
texto <- data.frame(x = pc1V, y = pc2V,z = pc3V,label = names(df)  
)  
  
bip %>%   
 ggplot(aes(x=pc1c, y=pc2c))+  
 geom\_point() +  
 geom\_point(aes(shape = grupo, color = grupo), size = 3) + theme\_minimal()+  
 scale\_shape\_manual(values=16:(15+n\_grupo)) +  
 scale\_colour\_brewer(palette = "Set1") +  
 geom\_vline(aes(xintercept=0),  
 color="black", size=1) +  
 geom\_hline(aes(yintercept=0),  
 color="black", size=1) +  
 annotate(geom="segment",  
 x=rep(0,length(df)),  
 xend=texto$x,  
 y=rep(0,length(df)),  
 yend=texto$y,color="black",lwd=.5) +  
 geom\_label(data=texto,aes(x=x,y=y,label=label),  
 color="black",angle=0,fontface="bold",size=4,fill="white") +  
 labs(x=paste("CP1 (",round(100\*ve[1],2),"%)",sep=""),  
 y=paste("CP2 (",round(100\*ve[2],2),"%)",sep=""),  
 color="",shape="") +  
 theme(legend.position = "top")+  
 coord\_cartesian(xlim=c(-20,300))



ck<-sum(pca$sdev^2>=1)  
tabelapca<-vector()  
for( l in 1:ck) tabelapca<-cbind(tabelapca,mcor[,l])  
colnames(tabelapca)<-paste(rep(c("PC"),ck),1:ck,sep="")  
pcat<-round(tabelapca,3)  
tabelapca<-tabelapca[order(abs(tabelapca[,1])),]  
print(tabelapca)  
#> vl\_aee qt\_aee qt\_pre\_escola vl\_pre\_escola vl\_fundamental   
#> 0.9856991 0.9943210 0.9957766 0.9967815 0.9968977   
#> vl\_eja vl\_creche qt\_creche qt\_eja qt\_fundamental   
#> 0.9975330 0.9991835 0.9992132 0.9992737 0.9998276

df\_grupo <- data.frame(nome\_regiao\_intermediaria = nome, grupo)  
  
df\_grupo <- left\_join(df\_nomes %>%   
 select(id\_municipio, nome, nome\_regiao\_intermediaria),  
 df\_grupo,  
 by ="nome\_regiao\_intermediaria")  
  
d\_sf\_municipio <- st\_read("shp/35MUE250GC\_SIR.shp", quiet = TRUE)  
  
d\_sf\_municipio <- d\_sf\_municipio %>%  
 rename(id\_municipio = CD\_GEOCMU) %>%  
 inner\_join(df\_grupo %>%  
 relocate(id\_municipio), "id\_municipio") %>%   
 mutate(grupo = as.factor(grupo))  
  
ggplot(d\_sf\_municipio) +  
 geom\_sf(aes(fill = grupo))+  
 theme\_minimal() +  
 scale\_fill\_viridis\_d()



for(i in 1:n\_grupo){  
 print("===============")  
 print(paste0("Grupo ",i,""))  
 print("===============")  
 print(paste(nome[grupo == i],collapse = "; "))  
 cat("\n")  
}  
#> [1] "==============="  
#> [1] "Grupo 1"  
#> [1] "==============="  
#> [1] "Araraquara; Araçatuba; Bauru; Marília; Presidente Prudente; Ribeirão Preto; Sorocaba; São José do Rio Preto; São José dos Campos"  
#>   
#> [1] "==============="  
#> [1] "Grupo 2"  
#> [1] "==============="  
#> [1] "Campinas"  
#>   
#> [1] "==============="  
#> [1] "Grupo 3"  
#> [1] "==============="  
#> [1] "São Paulo"

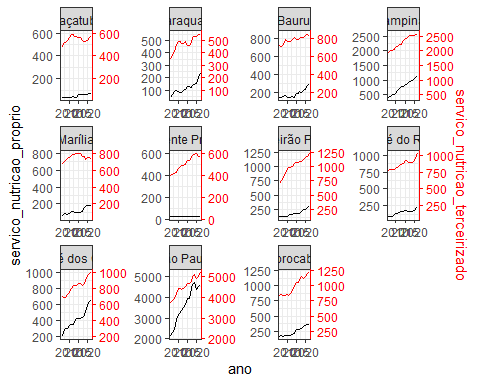
**4.2 Banco de Leite**   
4.2.1 Próprio **(OK)**  
4.2.2 Terceirizado **(OK)**  
**4.3 Serviço Lactário**  
4.3.1 Próprio **(OK)**  
4.3.2 Terceirizado **(OK)**  
**4.4 Serviço Nutrição**  
4.4.1 Próprio **(OK)**  
4.4.2 Terceirizado **(OK)**

### Base de dados - “sisvan”

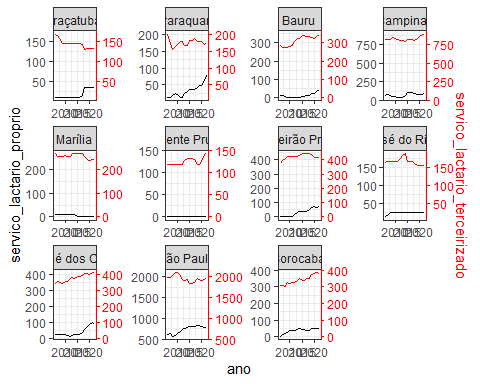
sisvan\_estab <- read\_rds("data/sisvan\_estab.rds")  
sisvan <- sisvan\_estab %>%   
 group\_by(municipio, ano) %>%   
 summarise(  
 servico\_nutricao\_proprio =   
 sum(indicador\_servico\_nutricao\_proprio),  
 servico\_nutricao\_terceirizado =   
 sum(indicador\_servico\_nutricao\_terceirizado),  
 servico\_lactario\_proprio =   
 sum(indicador\_servico\_lactario\_proprio), servico\_lactario\_terceirizado =   
 sum(indicador\_servico\_lactario\_terceirizado), servico\_banco\_leite\_proprio =   
 sum(indicador\_servico\_banco\_leite\_proprio), servico\_banco\_leite\_terceirizado=  
 sum(indicador\_servico\_banco\_leite\_terceirizado),  
 # total = trunc(n()/12)  
 ) %>% mutate(nome = municipio)  
  
sisvan\_estab <- left\_join(sisvan\_estab %>%   
 mutate(nome = municipio),   
 df\_nomes %>%  
 select(nome, nome\_regiao\_intermediaria),  
 by = "nome")

### Visualização de dados

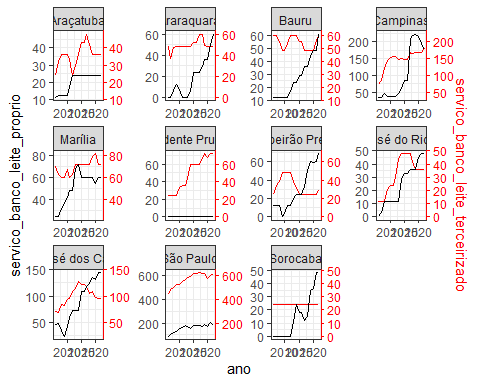
sisvan <- left\_join(sisvan,  
 df\_nomes %>%   
 select(nome, nome\_regiao\_intermediaria),  
 by = "nome")  
  
glimpse(sisvan)  
#> Rows: 10,965  
#> Columns: 10  
#> Groups: municipio [645]  
#> $ municipio <chr> "Adamantina", "Adamantina", "Adamanti~  
#> $ ano <dbl> 2006, 2007, 2008, 2009, 2010, 2011, 2~  
#> $ servico\_nutricao\_proprio <int> 24, 24, 24, 24, 24, 24, 24, 24, 24, 4~  
#> $ servico\_nutricao\_terceirizado <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~  
#> $ servico\_lactario\_proprio <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~  
#> $ servico\_lactario\_terceirizado <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~  
#> $ servico\_banco\_leite\_proprio <int> 0, 0, 0, 0, 9, 12, 12, 12, 23, 24, 24~  
#> $ servico\_banco\_leite\_terceirizado <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~  
#> $ nome <chr> "Adamantina", "Adamantina", "Adamanti~  
#> $ nome\_regiao\_intermediaria <chr> "Presidente Prudente", "Presidente Pr~  
coeff <- 1  
sisvan %>%   
 group\_by(ano,nome\_regiao\_intermediaria) %>%   
 summarise(  
 servico\_nutricao\_proprio = sum(servico\_nutricao\_proprio,na.rm=TRUE),  
 servico\_nutricao\_terceirizado = sum(servico\_nutricao\_terceirizado,na.rm=TRUE)   
 ) %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = servico\_nutricao\_proprio), color="red") +  
 geom\_line(aes(y = servico\_nutricao\_terceirizado/coeff)) +  
 scale\_y\_continuous(  
 name = "servico\_nutricao\_proprio",  
 sec.axis = sec\_axis(~.\*coeff, name="servico\_nutricao\_terceirizado")  
 ) +  
 facet\_wrap(~ nome\_regiao\_intermediaria, scale="free", ncol=4) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red"))



coeff <- 1  
sisvan %>%   
 group\_by(ano,nome\_regiao\_intermediaria) %>%   
 summarise(  
 servico\_lactario\_proprio = sum(servico\_lactario\_proprio,na.rm=TRUE),  
 servico\_lactario\_terceirizado = sum(servico\_lactario\_terceirizado,na.rm=TRUE)   
 ) %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = servico\_lactario\_proprio), color="red") +  
 geom\_line(aes(y = servico\_lactario\_terceirizado/coeff)) +  
 scale\_y\_continuous(  
 name = "servico\_lactario\_proprio",  
 sec.axis = sec\_axis(~.\*coeff, name="servico\_lactario\_terceirizado")  
 ) +  
 facet\_wrap(~ nome\_regiao\_intermediaria, scale="free", ncol=4) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red"))



coeff <- 1  
sisvan %>%   
 group\_by(ano,nome\_regiao\_intermediaria) %>%   
 summarise(  
 servico\_banco\_leite\_proprio = sum(servico\_banco\_leite\_proprio,na.rm=TRUE),  
 servico\_banco\_leite\_terceirizado = sum(servico\_banco\_leite\_terceirizado,na.rm=TRUE)   
 ) %>%   
 ggplot(aes(x = ano)) +  
 geom\_line(aes(y = servico\_banco\_leite\_proprio), color="red") +  
 geom\_line(aes(y = servico\_banco\_leite\_terceirizado/coeff)) +  
 scale\_y\_continuous(  
 name = "servico\_banco\_leite\_proprio",  
 sec.axis = sec\_axis(~.\*coeff, name="servico\_banco\_leite\_terceirizado")  
 ) +  
 facet\_wrap(~ nome\_regiao\_intermediaria, scale="free", ncol=4) +  
 theme(axis.line.y.right = element\_line(color = "red"),   
 axis.ticks.y.right = element\_line(color = "red"),  
 axis.text.y.right = element\_text(color = "red"),  
 axis.title.y.right = element\_text(color = "red"))



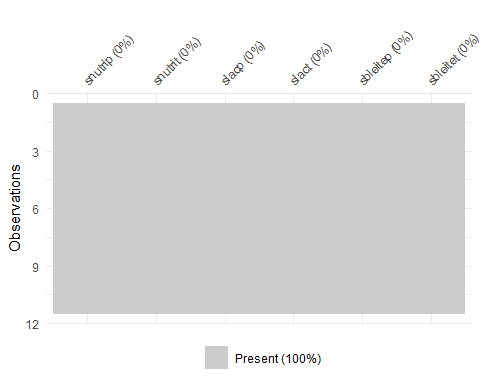
data\_set <- left\_join(data\_set, sisvan, by = c("ano","nome"))  
# data\_set

#### Criando a base resumo

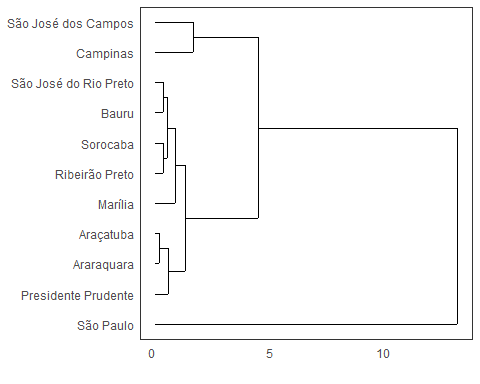
sisvan\_resumo <- sisvan %>%   
 group\_by(nome\_regiao\_intermediaria, nome) %>%   
 summarise(  
 snutrip = mean(servico\_nutricao\_proprio,na.rm=TRUE),  
 snutrit = mean(servico\_nutricao\_terceirizado,na.rm=TRUE),   
 slacp = mean(servico\_lactario\_proprio,na.rm=TRUE),  
 slact = mean(servico\_lactario\_terceirizado,na.rm=TRUE),   
 sbleitep = mean(servico\_banco\_leite\_proprio,na.rm=TRUE),  
 sbleitet = mean(servico\_banco\_leite\_terceirizado,na.rm=TRUE),  
 ) %>%   
 group\_by(nome\_regiao\_intermediaria) %>%  
 summarise(  
 snutrip = sum(snutrip,na.rm=TRUE),  
 snutrit = sum(snutrit,na.rm=TRUE),   
 slacp = sum(slacp,na.rm=TRUE),  
 slact = sum(slact,na.rm=TRUE),   
 sbleitep = sum(sbleitep,na.rm=TRUE),  
 sbleitet = sum(sbleitet,na.rm=TRUE),  
 ) %>%   
 rename(nome = nome\_regiao\_intermediaria)  
   
# data\_set\_muni <- left\_join(data\_set\_muni, sisvan\_resumo, by = "nome")

#### Agrupamento e Correlação

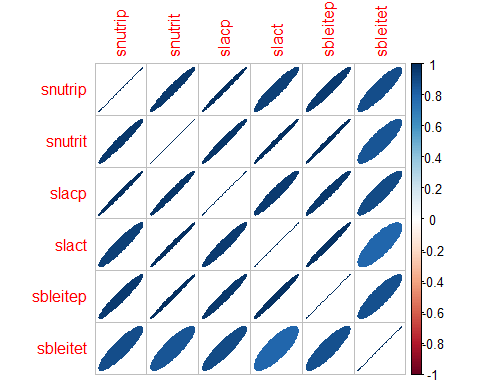
nome <- sisvan\_resumo$nome  
da <- sisvan\_resumo  
da\_pad <- decostand(da[-1] ,   
 method = "standardize",  
 na.rm=TRUE)  
df <- da\_pad  
visdat::vis\_miss(da\_pad)



row.names(df) <- nome  
da\_pad\_euc<-vegdist(df,"euclidean",na.rm=TRUE)   
da\_pad\_euc\_ward<-hclust(da\_pad\_euc, method="ward.D")  
grupo<-cutree(da\_pad\_euc\_ward,3)  
n\_grupo <- length(unique(grupo))  
ggdendrogram(da\_pad\_euc\_ward, rotate = TRUE, size = 2)

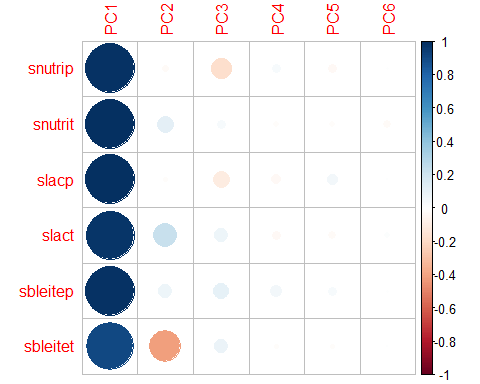


cor\_matrix <- cor(df,use = "na.or.complete")  
corrplot(cor\_matrix, method="ellipse")

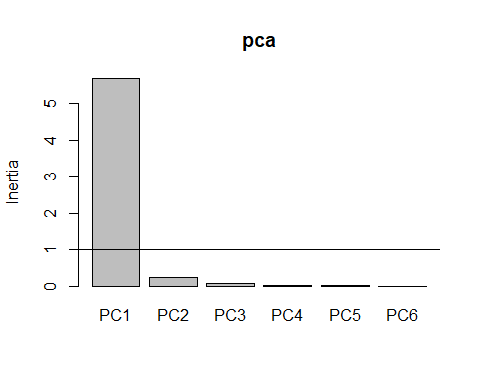


* Variáveis altamente correlacionadas, apenas uma delas representa bem esse banco de dados

pca <- prcomp(df,  
 scale=T)  
# Autovalores  
eig<-pca$sdev^2  
print(round(eig,3))  
#> [1] 5.680 0.240 0.066 0.008 0.006 0.001  
ve<-eig/sum(eig)  
print(round(ve,4))  
#> [1] 0.9466 0.0401 0.0110 0.0013 0.0009 0.0001  
print(round(cumsum(ve),4)\*100)  
#> [1] 94.66 98.67 99.77 99.89 99.99 100.00  
mcor<-cor(df,pca$x)  
corrplot(mcor)

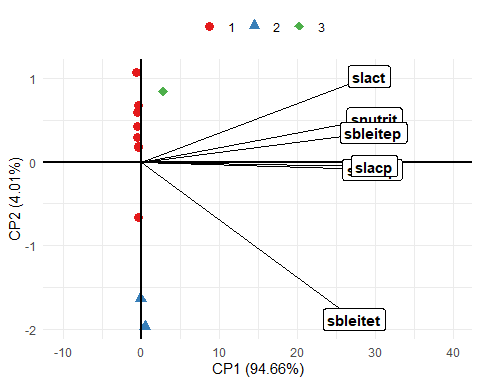


screeplot(pca)  
abline(h=1)



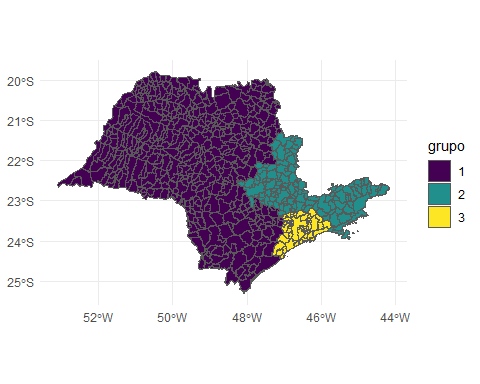
pc1V<-cor(df,pca$x)[,1]/sd(cor(df,pca$x)[,1])  
pc2V<-cor(df,pca$x)[,2]/sd(cor(df,pca$x)[,2])  
pc3V<-cor(df,pca$x)[,3]/sd(cor(df,pca$x)[,3])  
pc1c<-pca$x[,1]/sd(pca$x[,1])  
pc2c<-pca$x[,2]/sd(pca$x[,2])  
pc3c<-pca$x[,3]/sd(pca$x[,3])  
nv<-ncol(df)

bip<-data.frame(pc1c,pc2c,pc3c,nome,grupo=as.factor(grupo))  
texto <- data.frame(x = pc1V, y = pc2V,z = pc3V,label = names(df)  
)  
  
bip %>%   
 ggplot(aes(x=pc1c, y=pc2c))+  
 geom\_point() +  
 geom\_point(aes(shape = grupo, color = grupo), size = 3) + theme\_minimal()+  
 scale\_shape\_manual(values=16:(15+n\_grupo)) +  
 scale\_colour\_brewer(palette = "Set1") +  
 geom\_vline(aes(xintercept=0),  
 color="black", size=1) +  
 geom\_hline(aes(yintercept=0),  
 color="black", size=1) +  
 annotate(geom="segment",  
 x=rep(0,length(df)),  
 xend=texto$x,  
 y=rep(0,length(df)),  
 yend=texto$y,color="black",lwd=.5) +  
 geom\_label(data=texto,aes(x=x,y=y,label=label),  
 color="black",angle=0,fontface="bold",size=4,fill="white") +  
 labs(x=paste("CP1 (",round(100\*ve[1],2),"%)",sep=""),  
 y=paste("CP2 (",round(100\*ve[2],2),"%)",sep=""),  
 color="",shape="") +  
 theme(legend.position = "top") +  
 coord\_cartesian(xlim=c(-10,40))



ck<-sum(pca$sdev^2>=1)  
tabelapca<-vector()  
for( l in 1:ck) tabelapca<-cbind(tabelapca,mcor[,l])  
colnames(tabelapca)<-paste(rep(c("PC"),ck),1:ck,sep="")  
pcat<-round(tabelapca,3)  
tabelapca<-tabelapca[order(abs(tabelapca[,1])),]  
print(tabelapca)  
#> sbleitet slact snutrip sbleitep slacp snutrit   
#> 0.9068429 0.9706657 0.9833296 0.9894282 0.9919848 0.9925701

df\_grupo <- data.frame(nome\_regiao\_intermediaria = nome, grupo)  
  
df\_grupo <- left\_join(df\_nomes %>%   
 select(id\_municipio, nome, nome\_regiao\_intermediaria),  
 df\_grupo,  
 by ="nome\_regiao\_intermediaria")  
  
d\_sf\_municipio <- st\_read("shp/35MUE250GC\_SIR.shp", quiet = TRUE)  
  
d\_sf\_municipio <- d\_sf\_municipio %>%  
 rename(id\_municipio = CD\_GEOCMU) %>%  
 inner\_join(df\_grupo %>%  
 relocate(id\_municipio), "id\_municipio") %>%   
 mutate(grupo = as.factor(grupo))  
  
ggplot(d\_sf\_municipio) +  
 geom\_sf(aes(fill = grupo))+  
 theme\_minimal() +  
 scale\_fill\_viridis\_d()



for(i in 1:n\_grupo){  
 print("===============")  
 print(paste0("Grupo ",i,""))  
 print("===============")  
 print(paste(nome[grupo == i],collapse = "; "))  
 cat("\n")  
}  
#> [1] "==============="  
#> [1] "Grupo 1"  
#> [1] "==============="  
#> [1] "Araraquara; Araçatuba; Bauru; Marília; Presidente Prudente; Ribeirão Preto; Sorocaba; São José do Rio Preto"  
#>   
#> [1] "==============="  
#> [1] "Grupo 2"  
#> [1] "==============="  
#> [1] "Campinas; São José dos Campos"  
#>   
#> [1] "==============="  
#> [1] "Grupo 3"  
#> [1] "==============="  
#> [1] "São Paulo"

**DESAFIO 5**. Promover e proteger a alimentação adequada e saudável da população brasileira, com estratégias de educação alimentar e nutricional e medidas regulatórias

Temas: Promoção da alimentação saudável; Promoção da alimentação saudável no ambiente escolar; Ações regulatórias; Controle dos riscos relacionados ao consumo de alimentos e a exposição ao uso de agrotóxicos.

**DESAFIO 6**. Controlar e prevenir os agravos decorrentes da má alimentação Diretriz correspondente: 5

Temas: Desnutrição; excesso de peso e obesidade; doenças e agravos relacionados à má nutrição.

6.1 P/I abaixo do percentil 0,1: criança com peso muito baixo para a idade.  
6.2 P/I maior ou igual ao percentil 0,1 e menor que o percentil 3: criança com peso baixo para a idade.  
6.3 P/I maior ou igual ao percentil 3 e menor que o percentil 10: criança em risco nutricional.  
6.4 P/I maior ou igual ao percentil 10 e menor que o percentil 97: criança com peso adequado para a idade (eutrófica).  
6.5 P/I maior ou igual ao percentil 97: criança com risco de sobrepeso.  
6.6 P/A abaixo do percentil 3: criança com baixo peso para sua altura.  
6.7 P/A maior ou igual ao percentil 3 e menor que o percentil 10: criança com risco de baixo peso para sua altura.  
6.8 P/A maior ou igual ao percentil 10 e menor que o percentil 97: criança com peso adequado para sua altura.  
6.9 P/A maior ou igual ao percentil 97: criança com risco de sobrepeso para sua estatura.  
6.10 A/I abaixo do percentil 3: criança com altura baixa para a idade.  
6.11 A/I maior ou igual ao percentil 3 e menor que o percentil 10: criança com risco para altura baixa para a idade.  
6.12 A/I maior ou igual ao percentil 10 e menor que o percentil 97: criança com altura adequada para sua idade.  
6.13 A/I maior ou igual ao percentil 97: criança com altura elevada para sua idade  
6.14 Percentil de IMC por idade abaixo de 5: adolescente com baixo peso.  
6.15 Percentil de IMC por idade maior ou igual a 5 e menor que 85: adolescente com peso adequado (eutrófico).  
6.16 Percentil de IMC por idade maior ou igual a 85: adolescente com sobrepeso.  
6.17 Valores de IMC abaixo de 18,5: adulto com baixo peso.  
6.18 Valores de IMC maior ou igual a 18,5 e menor que 25,0: adulto com peso adequado (eutrófico).  
6.19 Valores de IMC maior ou igual a 25,0 e menor que 30,0: adulto com sobrepeso.  
6.20 Valores de IMC maior ou igual a 30,0: adulto com obesidade.  
6.21 índice de Massa Corporal (IMC) para avaliação do estado nutricional de idosos.  
6.22 Valores de IMC menor ou igual a 22,0: idoso com baixo peso.  
6.23 Valores de IMC maior que 22,0 e menor que 27,0: idoso com peso adequado (eutrófico).  
6.24 Valores de IMC maior ou igual a 27,0: idoso com sobrepeso.  
6.25 Valores de IMC por idade gestacional abaixo da curva inferior: gestante com baixo peso.  
6.26 Valores de IMC por idade gestacional entre a curva inferior e a segunda curva: gestante com peso adequado (eutrófica).  
6.27 Valores de IMC por idade gestacional entre a segunda curva e a curva superior: gestante com sobrepeso.  
6.28 Valores de IMC por idade gestacional acima da curva superior: gestante com obesidade

estado\_nutricional <- read\_rds("data/df\_final.rds") %>%   
 group\_by(ano,fase\_da\_vida,idade,indice,municipio) %>%   
 mutate(valor\_p = valor/sum(valor)\*100) %>%   
 ungroup()

estado\_nutricional <- left\_join(estado\_nutricional,   
 df\_nomes %>% select(nome, id\_municipio\_6,nome\_regiao\_intermediaria) %>%   
 rename(id\_municipio = id\_municipio\_6),  
 by = c("id\_municipio")) %>%   
 select(-regiao,-uf,-codigo\_uf,-id\_municipio,-municipio,-indice\_cri,  
 -indice\_ado,-valor) %>% relocate(ano,nome,nome\_regiao\_intermediaria) %>%   
 pivot\_wider(names\_from = c(fase\_da\_vida, idade, indice, classe),  
 values\_from = c(valor\_p)) %>%   
 janitor::clean\_names()

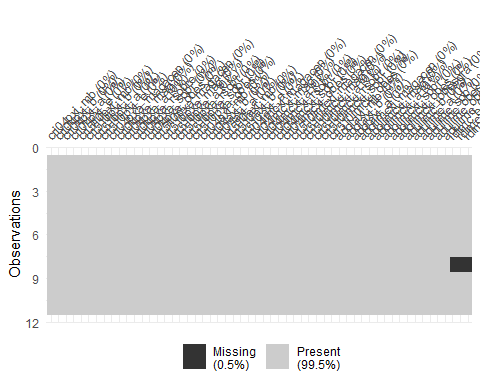
estado\_nutri\_resumo <- estado\_nutricional %>%   
 group\_by(nome) %>%   
 mutate(across(crianca\_0\_4\_peso\_x\_idade\_muito\_baixo:idoso\_imc\_sobrepeso,mean)) %>%   
 filter(ano == 2017) %>%   
 select(-ano)

# estado\_nutri\_resumo %>% names  
nomes\_antigos <- names(estado\_nutri\_resumo)  
novos\_nomes <- c(  
 "nome",   
 "nome\_regiao\_intermediaria" ,  
 "cri04pxi\_mb" ,   
 "cri04pxi\_b" ,   
 "cri04pxi\_a" ,   
 "cri04pxi\_e" ,   
 "cri510pxi\_mb" ,   
 "cri510pxi\_b" ,   
 "cri510pxi\_a" ,   
 "cri510pxi\_e" ,   
 "cri04pxa\_magacen",   
 "cri04pxa\_mag" ,   
 "cri04pxa\_a",   
 "cri04pxa\_rsobre",   
 "cri04pxa\_sobr",   
 "cri04pxa\_obes" ,   
 "cri510pxa\_magacen",  
 "cri510pxa\_mag" ,   
 "cri510pxa\_a" ,   
 "cri510pxa\_rsobr" ,   
 "cri510pxa\_sobr" ,   
 "cri510pxa\_obes",   
 "cri04axi\_mb" ,   
 "cri04axi\_b" ,   
 "cri04axi\_a" ,   
 "cri510axi\_mb",   
 "cri510axi\_b" ,   
 "cri510axi\_a",   
 "cri04imcxi\_magacen",   
 "cri04imcxi\_mag" ,   
 "cri04imcxi\_a" ,   
 "cri04imcxi\_rsobr" ,   
 "cri04imcxi\_sobr",   
 "cri04imcxi\_obes",   
 "cri510imcxi\_magacen",   
 "cri510imcxi\_mag" ,   
 "cri510imcxi\_a" ,   
 "cri510imcxi\_rsobr",   
 "cri510imcxi\_sobr",   
 "cri510imcxi\_obes",   
 "adolaxi\_mb",   
 "adolaxi\_b",   
 "adolaxi\_a" ,   
 "adolimcxi\_magacen",   
 "adolimcxi\_mag",   
 "adolimcxi\_a",   
 "adolimcxi\_sobr",   
 "adolimcxi\_obes" ,   
 "adolimcxi\_obesgra",   
 "adulimc\_b" ,   
 "adulimc\_a",   
 "adulimc\_sob" ,   
 "adulimc\_obesg\_i",   
 "adulimc\_obesg\_ii" ,   
 "adulimc\_obesg\_iii",   
 "idimc\_b",   
 "idimc\_a" ,   
 "idimc\_sob"  
)  
estado\_nutri\_resumo <- estado\_nutri\_resumo %>%   
 rename\_at(vars(nomes\_antigos), ~novos\_nomes)

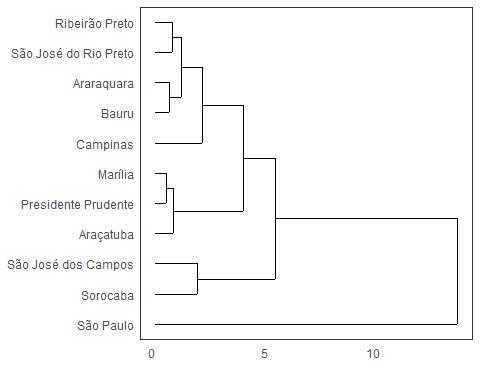
estado\_nutri\_resumo %>%  
 group\_by(nome\_regiao\_intermediaria) %>%   
 mutate(across(where(is.numeric), function(x) {  
 med <- mean(x,na.rm=TRUE)  
 x[is.na(x)] <- med  
 x  
 }))  
#> # A tibble: 645 x 58  
#> # Groups: nome\_regiao\_intermediaria [11]  
#> nome nome\_regiao\_intermed~1 cri04pxi\_mb cri04pxi\_b cri04pxi\_a cri04pxi\_e  
#> <chr> <chr> <dbl> <dbl> <dbl> <dbl>  
#> 1 Adamanti~ Presidente Prudente 0.993 2.26 88.8 7.98  
#> 2 Adolfo São José do Rio Preto 0.353 1.20 78.0 20.5   
#> 3 Aguaí Campinas 1.04 2.92 87.1 8.93  
#> 4 Águas da~ Campinas 2.14 5.38 70.5 22.0   
#> 5 Águas de~ Campinas 0.754 2.46 89.0 7.80  
#> 6 Águas de~ Sorocaba 0.637 1.31 92.3 5.70  
#> 7 Águas de~ Campinas 2.14 5.38 70.5 22.0   
#> 8 Agudos Bauru 1.35 3.22 84.9 10.5   
#> 9 Alambari Sorocaba 1.14 1.01 89.5 8.35  
#> 10 Alfredo ~ Presidente Prudente 0.948 1.70 86.5 10.8   
#> # i 635 more rows  
#> # i abbreviated name: 1: nome\_regiao\_intermediaria  
#> # i 52 more variables: cri510pxi\_mb <dbl>, cri510pxi\_b <dbl>,  
#> # cri510pxi\_a <dbl>, cri510pxi\_e <dbl>, cri04pxa\_magacen <dbl>,  
#> # cri04pxa\_mag <dbl>, cri04pxa\_a <dbl>, cri04pxa\_rsobre <dbl>,  
#> # cri04pxa\_sobr <dbl>, cri04pxa\_obes <dbl>, cri510pxa\_magacen <dbl>,  
#> # cri510pxa\_mag <dbl>, cri510pxa\_a <dbl>, cri510pxa\_rsobr <dbl>, ...

#### Agrupamento e Correlação

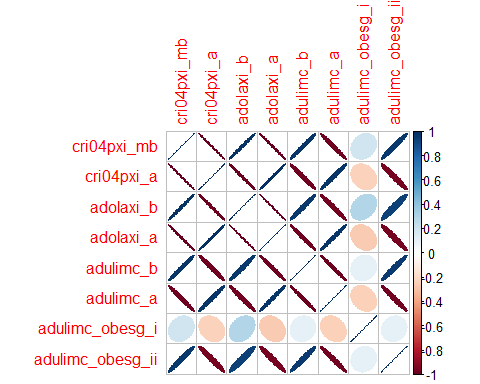
da <- estado\_nutri\_resumo %>%  
 group\_by(nome\_regiao\_intermediaria) %>%   
 mutate(across(where(is.numeric), function(x) {  
 mean(x,na.rm=TRUE)  
 })) %>% select(-nome) %>%   
 distinct() %>% ungroup()  
  
da\_pad <- decostand(da[-1] ,   
 method = "standardize",  
 na.rm=TRUE)  
visdat::vis\_miss(da\_pad)



df <- da\_pad %>%   
 select(cri04pxi\_mb,cri04pxi\_a,adolaxi\_b,adolaxi\_a,adulimc\_b,adulimc\_a,adulimc\_obesg\_i,adulimc\_obesg\_ii)  
row.names(df) <- da$nome\_regiao\_intermediaria  
da\_pad\_euc<-vegdist(df,"euclidean",na.rm=TRUE)   
da\_pad\_euc\_ward<-hclust(da\_pad\_euc, method="ward.D")  
grupo<-cutree(da\_pad\_euc\_ward,3)  
n\_grupo <- length(unique(grupo))  
ggdendrogram(da\_pad\_euc\_ward, rotate = TRUE, size = 2)

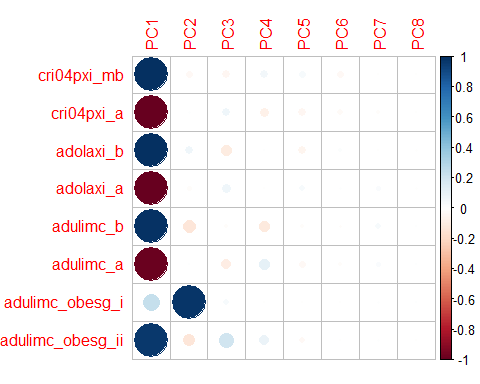


cor\_matrix <- cor(df,use = "na.or.complete")  
corrplot(cor\_matrix, method="ellipse")

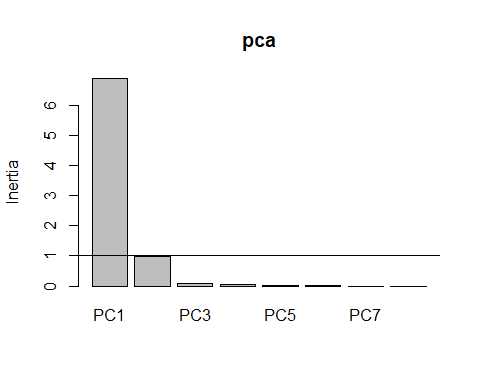


* Variáveis altamente correlacionadas, apenas uma delas representa bem esse banco de dados

pca <- prcomp(df,  
 scale=T)  
# Autovalores  
eig<-pca$sdev^2  
print(round(eig,3))  
#> [1] 6.891 0.984 0.069 0.039 0.010 0.003 0.002 0.000  
ve<-eig/sum(eig)  
print(round(ve,4))  
#> [1] 0.8614 0.1230 0.0087 0.0049 0.0013 0.0004 0.0003 0.0000  
print(round(cumsum(ve),4)\*100)  
#> [1] 86.14 98.45 99.32 99.80 99.93 99.97 100.00 100.00  
mcor<-cor(df,pca$x)  
corrplot(mcor)

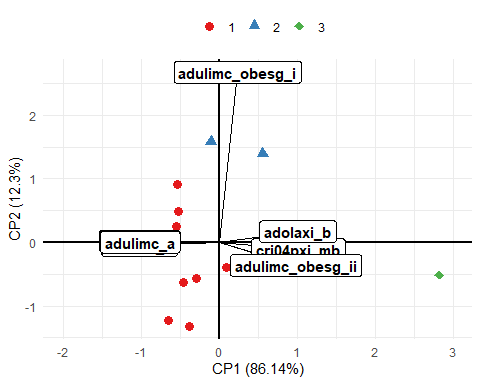


screeplot(pca)  
abline(h=1)



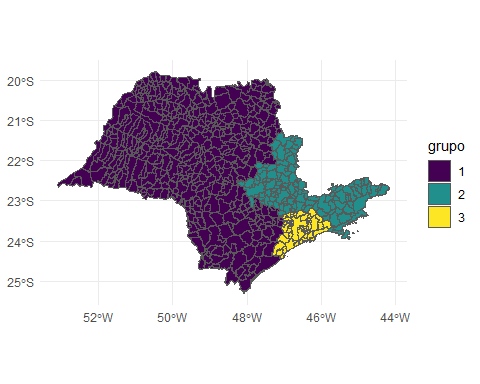
pc1V<-cor(df,pca$x)[,1]/sd(cor(df,pca$x)[,1])  
pc2V<-cor(df,pca$x)[,2]/sd(cor(df,pca$x)[,2])  
pc3V<-cor(df,pca$x)[,3]/sd(cor(df,pca$x)[,3])  
pc1c<-pca$x[,1]/sd(pca$x[,1])  
pc2c<-pca$x[,2]/sd(pca$x[,2])  
pc3c<-pca$x[,3]/sd(pca$x[,3])  
nv<-ncol(df)

bip<-data.frame(pc1c,pc2c,pc3c,nome,grupo=as.factor(grupo))  
texto <- data.frame(x = pc1V, y = pc2V,z = pc3V,label = names(df)  
)  
  
bip %>%   
 ggplot(aes(x=pc1c, y=pc2c))+  
 geom\_point() +  
 geom\_point(aes(shape = grupo, color = grupo), size = 3) + theme\_minimal()+  
 scale\_shape\_manual(values=16:(15+n\_grupo)) +  
 scale\_colour\_brewer(palette = "Set1") +  
 geom\_vline(aes(xintercept=0),  
 color="black", size=1) +  
 geom\_hline(aes(yintercept=0),  
 color="black", size=1) +  
 annotate(geom="segment",  
 x=rep(0,length(df)),  
 xend=texto$x,  
 y=rep(0,length(df)),  
 yend=texto$y,color="black",lwd=.5) +  
 geom\_label(data=texto,aes(x=x,y=y,label=label),  
 color="black",angle=0,fontface="bold",size=4,fill="white") +  
 labs(x=paste("CP1 (",round(100\*ve[1],2),"%)",sep=""),  
 y=paste("CP2 (",round(100\*ve[2],2),"%)",sep=""),  
 color="",shape="") +  
 theme(legend.position = "top") +  
 coord\_cartesian(xlim=c(-2,3))



ck<-sum(pca$sdev^2>=1)  
tabelapca<-vector()  
for( l in 1:ck) tabelapca<-cbind(tabelapca,mcor[,l])  
colnames(tabelapca)<-paste(rep(c("PC"),ck),1:ck,sep="")  
pcat<-round(tabelapca,3)  
tabelapca<-tabelapca[order(abs(tabelapca[,1])),]  
print(tabelapca)  
#> adulimc\_obesg\_i adulimc\_obesg\_ii adulimc\_b adulimc\_a   
#> 0.2363062 0.9674039 0.9838531 -0.9890383   
#> adolaxi\_b cri04pxi\_a cri04pxi\_mb adolaxi\_a   
#> 0.9910596 -0.9935463 0.9950909 -0.9970059

df\_grupo <- data.frame(nome\_regiao\_intermediaria = nome, grupo)  
  
df\_grupo <- left\_join(df\_nomes %>%   
 select(id\_municipio, nome, nome\_regiao\_intermediaria),  
 df\_grupo,  
 by ="nome\_regiao\_intermediaria")  
  
d\_sf\_municipio <- st\_read("shp/35MUE250GC\_SIR.shp", quiet = TRUE)  
  
d\_sf\_municipio <- d\_sf\_municipio %>%  
 rename(id\_municipio = CD\_GEOCMU) %>%  
 inner\_join(df\_grupo %>%  
 relocate(id\_municipio), "id\_municipio") %>%   
 mutate(grupo = as.factor(grupo))  
  
ggplot(d\_sf\_municipio) +  
 geom\_sf(aes(fill = grupo))+  
 theme\_minimal() +  
 scale\_fill\_viridis\_d()



for(i in 1:n\_grupo){  
 print("===============")  
 print(paste0("Grupo ",i,""))  
 print("===============")  
 print(paste(nome[grupo == i],collapse = "; "))  
 cat("\n")  
}  
#> [1] "==============="  
#> [1] "Grupo 1"  
#> [1] "==============="  
#> [1] "Araraquara; Araçatuba; Bauru; Marília; Presidente Prudente; Ribeirão Preto; Sorocaba; São José do Rio Preto"  
#>   
#> [1] "==============="  
#> [1] "Grupo 2"  
#> [1] "==============="  
#> [1] "Campinas; São José dos Campos"  
#>   
#> [1] "==============="  
#> [1] "Grupo 3"  
#> [1] "==============="  
#> [1] "São Paulo"

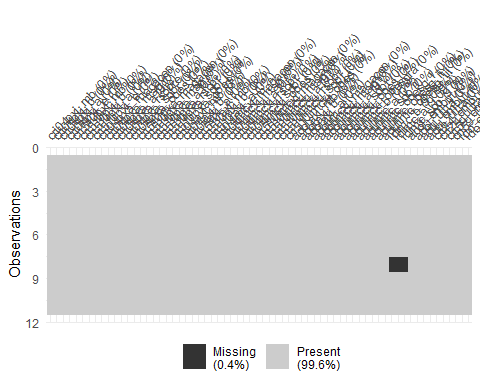
1. Consumo  
   7.1 Consumo de ultra processados **(OK)** 7.1.2 Consumo embutidos **(OK)** 7.1.3 Consumo processados **(OK)**

consumo <- read\_rds("data/consumo/consumo.rds")  
consumo <- left\_join(consumo, df\_nomes %>% select(nome, id\_municipio\_6, nome\_regiao\_intermediaria) %>%   
 rename(codigo\_ibge = id\_municipio\_6) %>%  
 mutate(codigo\_ibge = as.numeric(codigo\_ibge)),  
 by = c("codigo\_ibge"))  
#consumo$tipo\_relatorio %>% unique()  
  
tipos\_relat <- c("CONS\_ULTRA","CONS\_EMBUT")  
  
consumo <- consumo %>% select(-faixa\_etaria, -codigo\_ibge, -municipio,  
 -uf, -total, -monitorados) %>%   
 filter(tipo\_relatorio %in% tipos\_relat,  
 fase\_da\_vida != "entre-6-meses-23-meses")  
  
consumo <- consumo %>%   
 pivot\_wider(names\_from = c(fase\_da\_vida,tipo\_relatorio),  
 values\_from = percent) %>%   
 janitor::clean\_names()

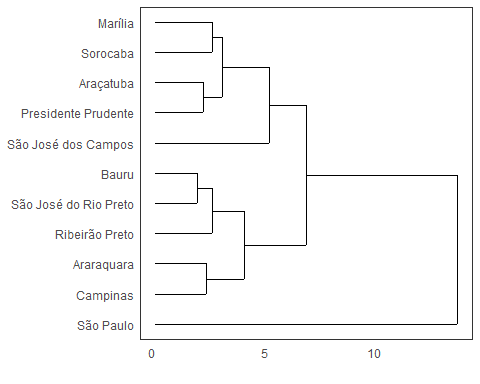
consumo\_estado\_nutricional <- left\_join(consumo, estado\_nutricional,  
 by=c("ano","nome","nome\_regiao\_intermediaria"))  
  
consumo\_resumo <- consumo %>%   
 group\_by(nome\_regiao\_intermediaria, nome) %>%   
 summarise(  
 adol\_embut = mean(adolecentes\_cons\_embut,na.rm=TRUE),   
 adol\_ultra = mean(adolecentes\_cons\_ultra,na.rm=TRUE),   
 adul\_embut = mean(adultos\_cons\_embut,na.rm=TRUE),   
 adul\_ultra = mean(adultos\_cons\_ultra,na.rm=TRUE),   
 cri24\_embut = mean(criancas\_2\_4\_cons\_embut,na.rm=TRUE),  
 cri24\_ultra = mean(criancas\_2\_4\_cons\_ultra,na.rm=TRUE),   
 cri59\_embut = mean(criancas\_5\_9\_cons\_embut,na.rm=TRUE),  
 cri59\_ultra = mean(criancas\_5\_9\_cons\_ultra,na.rm=TRUE),  
 ido\_embut = mean(idosos\_cons\_embut,na.rm=TRUE),   
 ido\_ultra = mean(idosos\_cons\_ultra,na.rm=TRUE)   
 )  
  
consumo\_estado\_nutricional\_resumo <- left\_join(estado\_nutri\_resumo, consumo\_resumo,  
 by = c("nome","nome\_regiao\_intermediaria"))

#### Agrupamento e Correlação

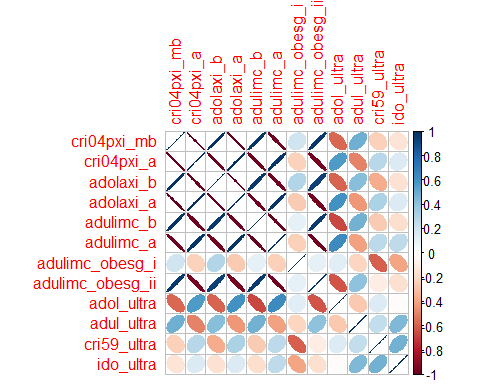
da <- consumo\_estado\_nutricional\_resumo %>%  
 group\_by(nome\_regiao\_intermediaria) %>%   
 mutate(across(where(is.numeric), function(x) {  
 mean(x,na.rm=TRUE)  
 })) %>% select(-nome) %>%   
 distinct() %>% ungroup()  
  
da\_pad <- decostand(da[-1] ,   
 method = "standardize",  
 na.rm=TRUE)  
visdat::vis\_miss(da\_pad)



df <- da\_pad %>%   
 select(cri04pxi\_mb,cri04pxi\_a,adolaxi\_b,adolaxi\_a,adulimc\_b,adulimc\_a,adulimc\_obesg\_i,adulimc\_obesg\_ii,adol\_ultra,adul\_ultra,cri59\_ultra,ido\_ultra)  
row.names(df) <- da$nome\_regiao\_intermediaria  
da\_pad\_euc<-vegdist(df,"euclidean",na.rm=TRUE)   
da\_pad\_euc\_ward<-hclust(da\_pad\_euc, method="ward.D")  
grupo<-cutree(da\_pad\_euc\_ward,3)  
n\_grupo <- length(unique(grupo))  
ggdendrogram(da\_pad\_euc\_ward, rotate = TRUE, size = 2)

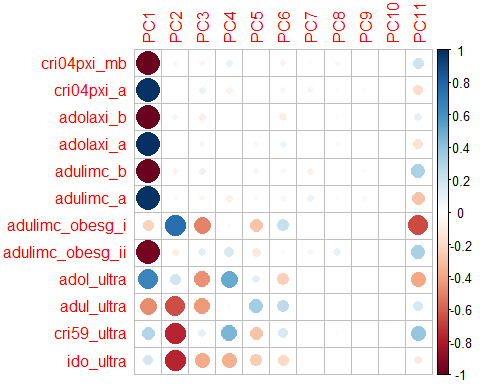


cor\_matrix <- cor(df,use = "na.or.complete")  
corrplot(cor\_matrix, method="ellipse")

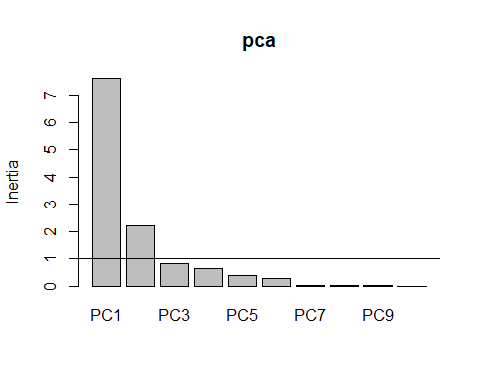


* Variáveis altamente correlacionadas, apenas uma delas representa bem esse banco de dados

pca <- prcomp(df,  
 scale=T)  
# Autovalores  
eig<-pca$sdev^2  
print(round(eig,3))  
#> [1] 7.611 2.224 0.837 0.634 0.380 0.272 0.022 0.016 0.004 0.001 0.000  
ve<-eig/sum(eig)  
print(round(ve,4))  
#> [1] 0.6342 0.1853 0.0698 0.0528 0.0317 0.0227 0.0018 0.0013 0.0003 0.0000  
#> [11] 0.0000  
print(round(cumsum(ve),4)\*100)  
#> [1] 63.42 81.95 88.93 94.21 97.38 99.65 99.83 99.96 100.00 100.00  
#> [11] 100.00  
mcor<-cor(df,pca$x)  
corrplot(mcor)

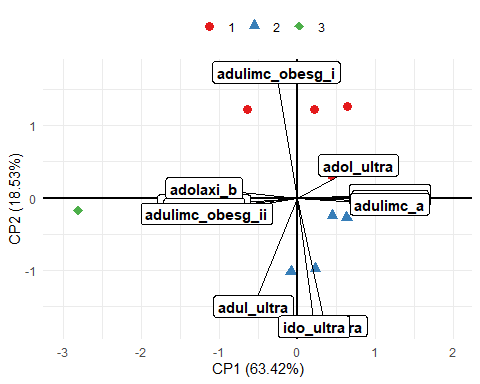


screeplot(pca)  
abline(h=1)



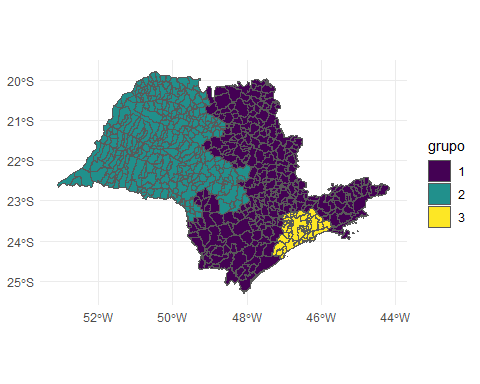
pc1V<-cor(df,pca$x)[,1]/sd(cor(df,pca$x)[,1])  
pc2V<-cor(df,pca$x)[,2]/sd(cor(df,pca$x)[,2])  
pc3V<-cor(df,pca$x)[,3]/sd(cor(df,pca$x)[,3])  
pc1c<-pca$x[,1]/sd(pca$x[,1])  
pc2c<-pca$x[,2]/sd(pca$x[,2])  
pc3c<-pca$x[,3]/sd(pca$x[,3])  
nv<-ncol(df)

bip<-data.frame(pc1c,pc2c,pc3c,nome,grupo=as.factor(grupo))  
texto <- data.frame(x = pc1V, y = pc2V,z = pc3V,label = names(df)  
)  
  
bip %>%   
 ggplot(aes(x=pc1c, y=pc2c))+  
 geom\_point() +  
 geom\_point(aes(shape = grupo, color = grupo), size = 3) + theme\_minimal()+  
 scale\_shape\_manual(values=16:(15+n\_grupo)) +  
 scale\_colour\_brewer(palette = "Set1") +  
 geom\_vline(aes(xintercept=0),  
 color="black", size=1) +  
 geom\_hline(aes(yintercept=0),  
 color="black", size=1) +  
 annotate(geom="segment",  
 x=rep(0,length(df)),  
 xend=texto$x,  
 y=rep(0,length(df)),  
 yend=texto$y,color="black",lwd=.5) +  
 geom\_label(data=texto,aes(x=x,y=y,label=label),  
 color="black",angle=0,fontface="bold",size=4,fill="white") +  
 labs(x=paste("CP1 (",round(100\*ve[1],2),"%)",sep=""),  
 y=paste("CP2 (",round(100\*ve[2],2),"%)",sep=""),  
 color="",shape="") +  
 theme(legend.position = "top") +  
 coord\_cartesian(xlim=c(-3,2))



ck<-sum(pca$sdev^2>=1)  
tabelapca<-vector()  
for( l in 1:ck) tabelapca<-cbind(tabelapca,mcor[,l])  
colnames(tabelapca)<-paste(rep(c("PC"),ck),1:ck,sep="")  
pcat<-round(tabelapca,3)  
tabelapca<-tabelapca[order(abs(tabelapca[,1])),]  
print(tabelapca)  
#> PC1 PC2  
#> ido\_ultra 0.1814009 -0.76929233  
#> adulimc\_obesg\_i -0.2207344 0.75220535  
#> cri59\_ultra 0.2997798 -0.76579065  
#> adul\_ultra -0.4623185 -0.64876991  
#> adol\_ultra 0.6539460 0.19543090  
#> adulimc\_obesg\_ii -0.9595405 -0.09447909  
#> adulimc\_b -0.9882449 -0.06433571  
#> adulimc\_a 0.9883348 -0.03685391  
#> adolaxi\_b -0.9885752 0.05748265  
#> cri04pxi\_a 0.9891557 0.01912702  
#> cri04pxi\_mb -0.9897903 -0.04546575  
#> adolaxi\_a 0.9940881 -0.01759425

df\_grupo <- data.frame(nome\_regiao\_intermediaria = nome, grupo)  
  
df\_grupo <- left\_join(df\_nomes %>%   
 select(id\_municipio, nome, nome\_regiao\_intermediaria),  
 df\_grupo,  
 by ="nome\_regiao\_intermediaria")  
  
d\_sf\_municipio <- st\_read("shp/35MUE250GC\_SIR.shp", quiet = TRUE)  
  
d\_sf\_municipio <- d\_sf\_municipio %>%  
 rename(id\_municipio = CD\_GEOCMU) %>%  
 inner\_join(df\_grupo %>%  
 relocate(id\_municipio), "id\_municipio") %>%   
 mutate(grupo = as.factor(grupo))  
  
ggplot(d\_sf\_municipio) +  
 geom\_sf(aes(fill = grupo))+  
 theme\_minimal() +  
 scale\_fill\_viridis\_d()



for(i in 1:n\_grupo){  
 print("===============")  
 print(paste0("Grupo ",i,""))  
 print("===============")  
 print(paste(nome[grupo == i],collapse = "; "))  
 cat("\n")  
}  
#> [1] "==============="  
#> [1] "Grupo 1"  
#> [1] "==============="  
#> [1] "Araraquara; Campinas; Ribeirão Preto; Sorocaba; São José dos Campos"  
#>   
#> [1] "==============="  
#> [1] "Grupo 2"  
#> [1] "==============="  
#> [1] "Araçatuba; Bauru; Marília; Presidente Prudente; São José do Rio Preto"  
#>   
#> [1] "==============="  
#> [1] "Grupo 3"  
#> [1] "==============="  
#> [1] "São Paulo"

**DESAFIO 7**. Ampliar a disponibilidade hídrica e o acesso à água para a população, em especial a população pobre no meio rural

Temas: Água para consumo humano; Água para a produção de alimentos; Recursos hídricos; Saneamento básico rural

**DESAFIO 8**. Consolidar a implementação do sistema nacional de segurança alimentar e nutricional (SISAN), aperfeiçoando a gestão federativa, a intersetorialidade e a participação social Diretrizes correspondentes: 3 e 8 (envolve ainda as diretrizes da LOSAN)

Temas: Intersetorialidade nas ações federativas; Participação social; Gestão e financiamento do Sistema; Formação, pesquisa e extensão em SAN e DHAA; Exigibilidade e monitoramento do DHAA

**DESAFIO 9**. Apoio às iniciativas de promoção da soberania, segurança alimentar e nutricional, do direito humano à alimentação adequada e de sistemas alimentares democráticos, saudáveis e sustentáveis em âmbito internacional, por meio do diálogo e da cooperação internacional

Temas: Governança global; Cooperação internacional; Participação da sociedade civil

**DESAFIO 10**. Socioeconômica

10.1Índice de Vulnerabilidade Social Dimensão Capital Humano  
10.2 Índice de Vulnerabilidade Social Dimensão Renda e Trabalho  
10.3 Índice de Desenvolvimento Humano Municipal  
10.4 ndice de Desenvolvimento Humano Municipal Dimensão Longevidade  
10.5 Índice de Desenvolvimento Humano Municipal Dimensão Educação  
10.6 Índice de Desenvolvimento Humano Municipal Dimensão Renda  
10.7 Índice de Desenvolvimento Humano Municipal Dimensão Sub Escolaridade  
10.8 Índice de Desenvolvimento Humano Municipal Dimensão Frequência Escolar  
10.9 Índice Gini  
10.10 Renda Per capita  
10.11 Proporção de empregados com carteira assinada com 18 anos ou mais  
10.12 Proporção de empregados sem carteira assinada com 18 anos ou mais  
10.13 Proporção de trabalhadores do setor público com 18 anos ou mais  
10.14 Proporção de trabalhadores por conta própria com 18 anos ou mais  
10.15 Proporção de empregadores com 18 anos ou mais  
10.6 Proporção do grau formalização dos ocupados com 18 anos ou mais  
10.7 Proporção de atividade das pessoas de 10 a 14 anos de idade  
10.8 População Economicamente Ativa (PEA) com 10 anos ou mais  
10.9 População Economicamente Ativa (PEA) com 10 a 14 anos  
10.10 População Economicamente Ativa (PEA) com 15 a 17 anos  
10.11 População Economicamente Ativa (PEA) com 18 anos ou mais