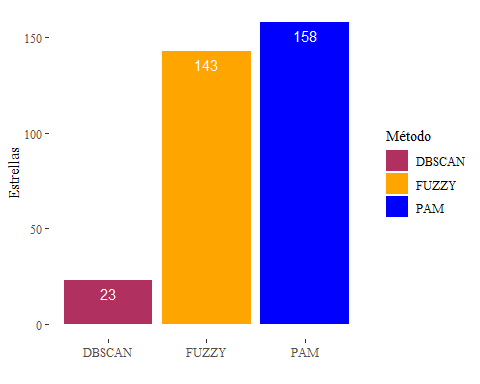
Seleccion Final de Candidatas

Fernando Menendez

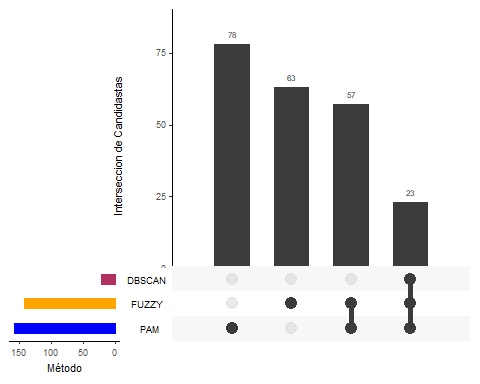
October 28, 2018

pam<-read.csv("hip\_pamk5\_clust.csv") %>% filter(Symbad\_Hyades==F)  
dbs<-read.csv("DBSCAN - Candidatas.csv")  
fzy<-read.csv("FUZZY - Candidatas.csv")

cand.list<- dbs %>% rbind(fzy) %>% rbind(pam[,c("HIP","source")])  
cand.list$source<-as.factor(cand.list$source)  
ggplot(data=cand.list,aes(x=source)) +  
 geom\_bar(aes(fill=source)) +  
 geom\_text(stat='count',aes(label=..count..),vjust=1.6,color="white") +  
 scale\_fill\_manual(name="Método",values=c("maroon","orange","blue")) +  
 xlab("") +  
 ylab("Estrellas") +  
 theme\_tufte()



listInput<-list(DBSCAN=dbs$HIP,FUZZY=fzy$HIP,PAM=pam$HIP)  
upset(fromList(listInput), order.by = "freq",point.size = 4, line.size = 2,   
 mainbar.y.label = "Interseccion de Candidastas", sets.x.label = "Método",  
 sets.bar.color=c("blue","orange","maroon"))



**OBS:** Como podemos ver en el grafico anterior, las estrellas detectadas en DBSCAN se encuentran tanto en PAM como en clustering difuso. Los 3 metodos comparten un total de **80** estrellas, entendemos a partir de esta validacion que estas tienen mayor chance de pertenecer efectivamante al grupo de Hyades.

cand.list.hip<- pam %>% select(HIP)  
write.csv(cand.list.hip, "Candidatas Hyades - Hipparcos.csv",row.names = F)  
cand.list.tyc<-read.csv("tyc\_pam\_clustk2.csv") %>% select(recno)  
write.csv(cand.list.tyc, "Candidatas Hyades - Tycho.csv", row.names = F)