

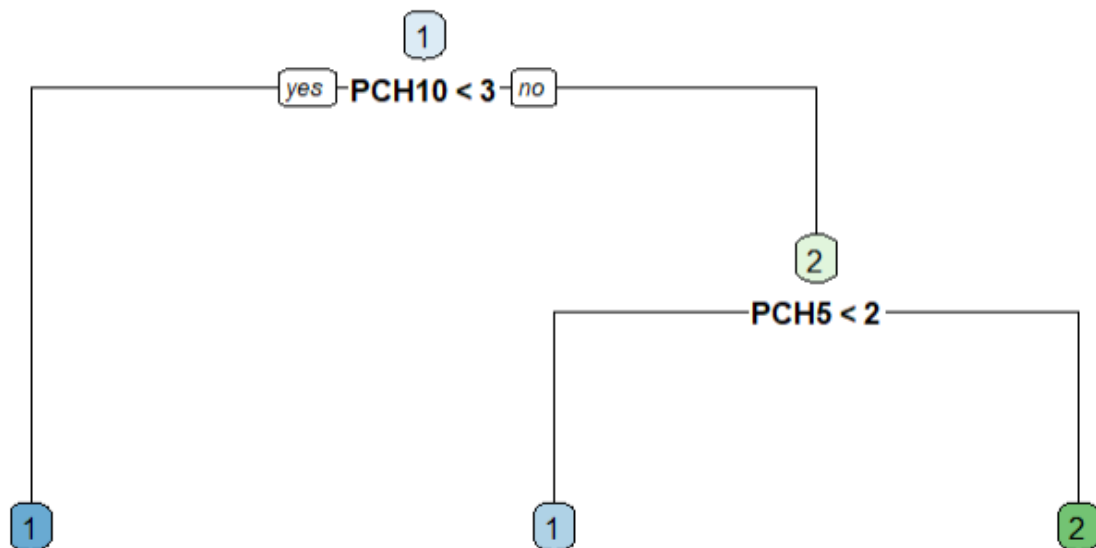
Arbol de decision

Para el hogar se hizo un análisis completo de las variables más relevantes

```
arbol <- rpart(AREA ~  
  PCH1+  
  PCH2+  
  PCH3+  
  PCH4+  
  PCH5+  
  PCH6+  
  PCH7+  
  PCH8+  
  PCH9_A+  
  PCH9_B+  
  PCH9_C+  
  PCH9_D+  
  PCH9_E+  
  PCH9_F+  
  PCH9_G+  
  PCH9_H+  
  PCH9_I+  
  PCH9_J+  
  PCH9_K+  
  PCH9_L+  
  PCH9_M+  
  PCH10+  
  PCH11+  
  PCH12+  
  PCH13+  
  PCH14,  
  data = hogar, method = "class")  
  
rpart.plot(arbol, type=2, extra=0, under = TRUE, fallen.leaves = TRUE, box.palette = "BuGn", main ="Predicción de area", cex = 0.5)
```

Esto nos llevó al resultado de que sólo dos variables son altamente influyentes en conocer si el hogar es urbano o rural. Estas variables son **PCH10** qué es ¿Cómo elimina la mayor parte de la basura en el hogar? y **PCH5** la cuál es ¿Qué tipo de servicio sanitario tiene este hogar?

Predicción de area



Cómo podemos ver dice que el primer factor de decisión que decide si es urbano es que la persona elimine su mayor parte de la basura de dos formas, siendo **Servicio Municipal y Servicio Privado**. Después podemos ver que el segundo factor de mayor importancia para decidir si es rural definitivamente o tiene una pequeña probabilidad de ser urbana aunque su forma de eliminar basura sea *quemarla, enterrarla, tirarla a un río quebrada o mar, tirar en cualquier lado, reciclaje u otro*, es que si su servicio sanitario es **un inodoro conectado a redes de drenajes**, sigue teniendo probabilidad de ser urbano, pero si esto tampoco se cumple y su servicio sanitario se encuentra entre *inodoro conectado a fosa séptica, excusado lavable, letrina o pozo ciego, o no tiene*, es casi seguro que es rural.

Lo interesante de esto es que si la persona es propietaria o no, o si cuenta con tele, cable, internet, carro o muchas variables más, no tienen importancia en descubrir si es urbano o rural cómo lo tienen estas dos variables, lo cuál es lógico e interesante de reafirmar, ya que estas dos variables demuestran la gran desigualdad que hay entre el área urbana y rural, ya que el manejo de basura y el servicio sanitario son dos servicios que deberían ser básicos para todos, esto por lo importante que es para la salud de los integrantes y la higiene de un hogar.

Link del repositorio: <https://github.com/Marito-R-T/data-mining-r>