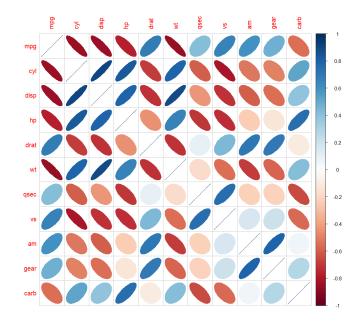
Actividad Evaluable 5: Análisis de datos

```
Actividad 5.R* X
                                                                                                              Source on Save | Q / V - | 1 #Actividad 5
                                                                                 Run D O O Source - E
     install.packages("factoextra")
    library(factoextra)
library(ggplot2)
     #Cargar dataset en R
data("mtcars")
  6
  8
      View(mtcars)
 10 kM <- kmeans(mtcars,3)</pre>
 11 kM
 12
 13 #Graficación de clusters
 14 fviz_cluster(kM,mtcars)
 15
 16  #Escalando datos
17  mtcarsEscalado <- scale(as.matrix(mtcars[, 1:4]))</pre>
 18 kM1 <- kmeans(mtcarsEscalado,4)
 19 fviz_cluster(kM1,mtcarsEscalado)
 20
 21 #Gráfica matriz de correlación
 22
    library(corrplot)
 23 corrCar <- cor(mtcars)
24 corrplot(corrCar,metho</pre>
     corrplot(corrCar,method = "ellipse")
 25
 26
```

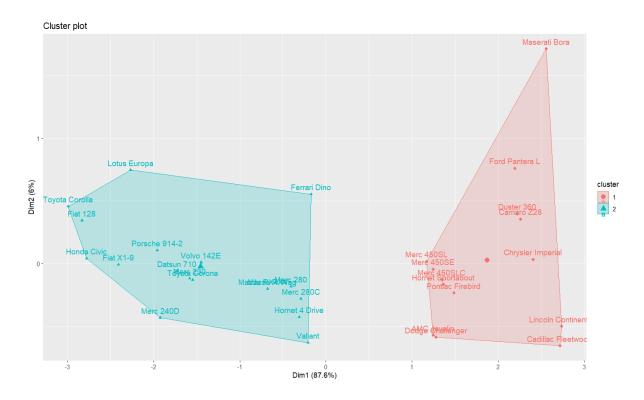
Correlación

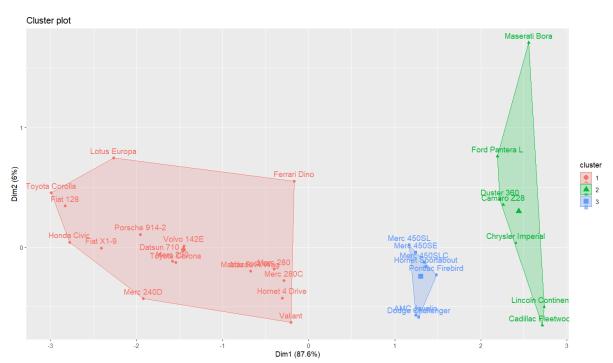


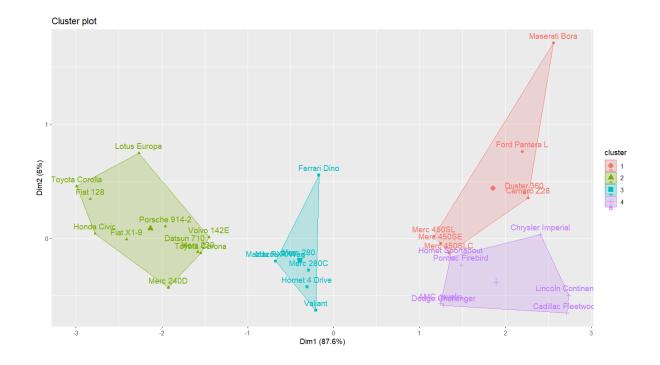
- ¿Cuáles variables están correlacionadas? cyl y mpg, disp y mpg, hp y mpg, wt y mpg, disp y cyl, vs y cyl, wt y disp.
- ¿Qué implicaciones tiene?

Una de las implicaciones es que entre más cilindros tiene un carro, rinde menos millas por galón. También entre más caballos de fuerza tiene un carro, rinde menos millas por galón.

K-means







• ¿Cuántos grupos son mejores para representar la información: 2, 3 o 4?

En 4 grupos es mejor, pues parece ser más específico el tema por el que se dividen (o al menos más enriquecedor), por ejemplo en el caso del cluster 1 (el verde) creemos que se agruparon según el mpg alto, hp bajo y el wt también bajo, pues los coches que incluye tienen valores similares en estos rubros, cluster 2 (azul) vimos una relación entre los valores de mpg bajo y hp alto, en el tercero (rojo) el hp alto, cyl similar y mgp bajo parecen tener relación entre ellos y en el último (morado) tienen un wt más alto, un hp medio y un mpg entre media y baja.

Conclusión

Después de haber estudiado las correlaciones entre las variables de los autos utilizando nuestras gráficas, pudimos concluir que los carros con mayor potencia que poseen más cilindros y caballos de fuerza, gastan más gasolina y por lo tanto rinden menos millas por galón.

