



Tecnológico de Monterrey

Campus Estado de México

Actividad Evaluable 5: Análisis de datos

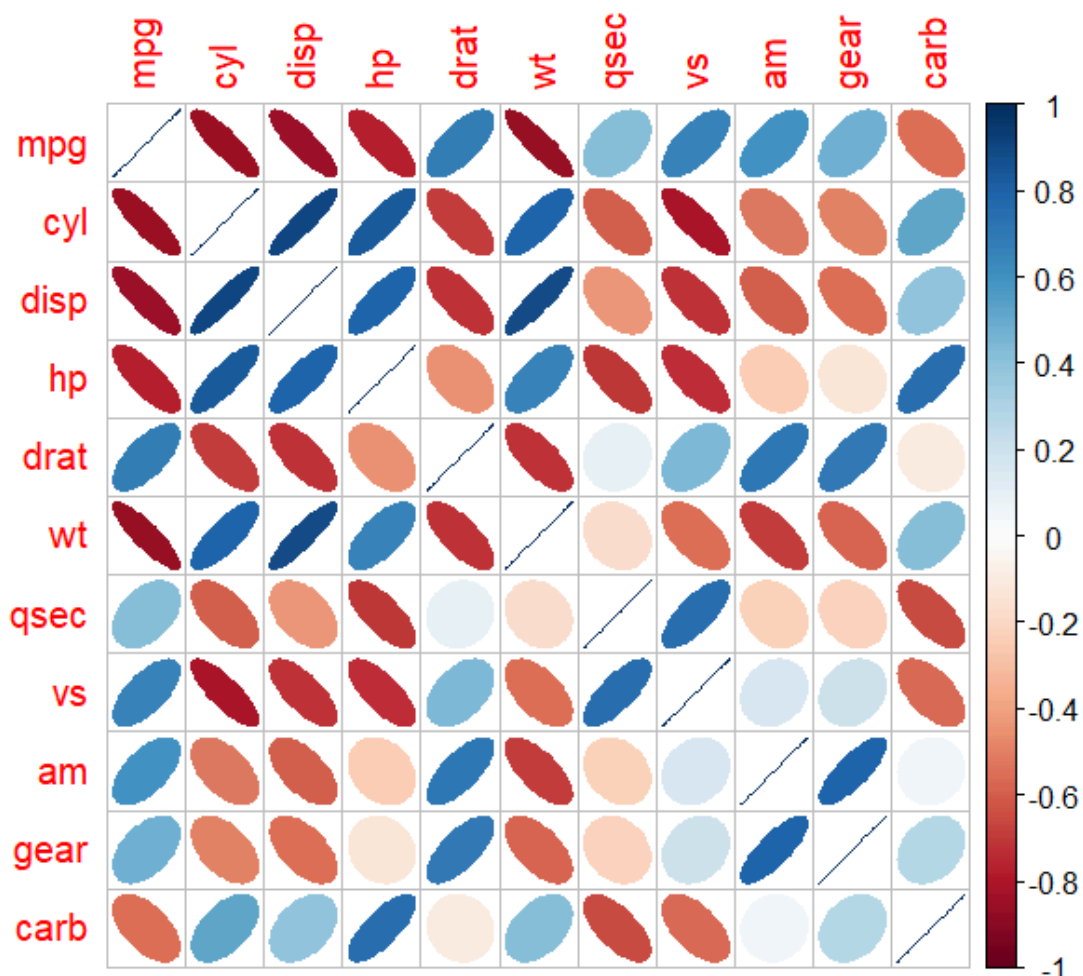
Profesores

Jorge Adolfo Ramírez Uresti

Herramientas computacionales: el arte de la analítica

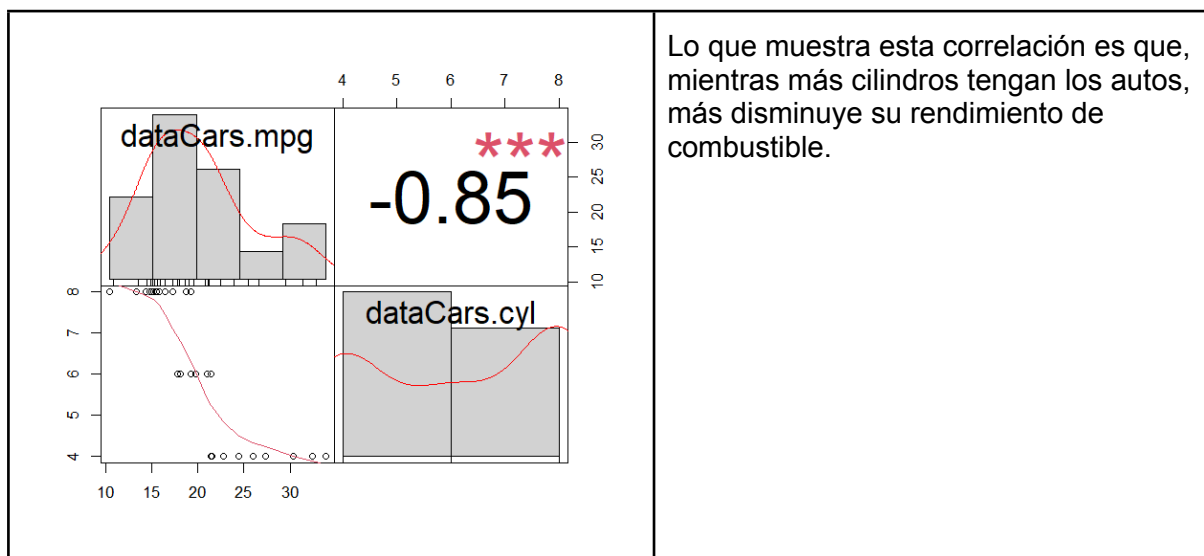
Arturo Barrios Mendoza		A01168331
Lucio Arturo Reyes Castillo		A01378985
Leyberth Jaaziel Castillo Guerra		A01749505

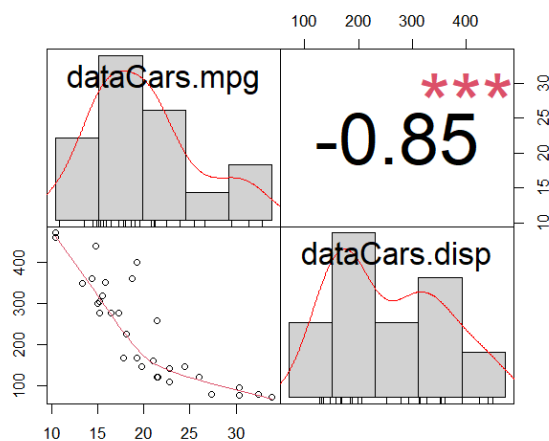
Fecha de Entrega:
8 de Mayo de 2024



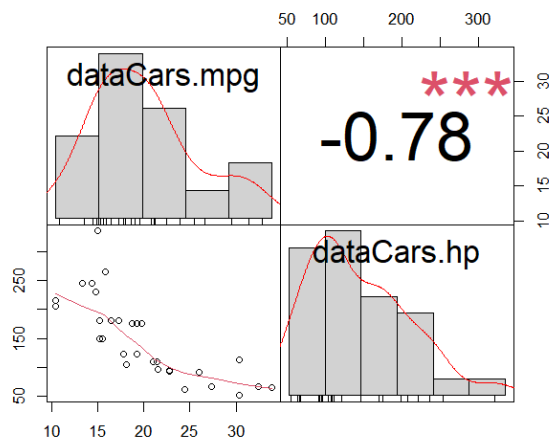
Posteriormente se detectaron los atributos con una mayor correlación, es decir, con una correlación mayor a 0.75 o menor a -0.75:

```
> cor(dataCars$mpg,dataCars$cyl)
[1] -0.852162
> cor(dataCars$mpg,dataCars$disp)
[1] -0.8475514
> cor(dataCars$mpg,dataCars$hp)
[1] -0.7761684
> cor(dataCars$mpg,dataCars$wt)
[1] -0.8676594
> cor(dataCars$cyl,dataCars$disp)
[1] 0.9020329
> cor(dataCars$cyl,dataCars$hp)
[1] 0.8324475
> cor(dataCars$cyl,dataCars$wt)
[1] 0.7824958
> cor(dataCars$cyl,dataCars$vs)
[1] -0.8108118
> cor(dataCars$disp,dataCars$hp)
[1] 0.7909486
> cor(dataCars$disp,dataCars$wt)
[1] 0.8879799
> cor(dataCars$am,dataCars$gear)
[1] 0.7940588
```

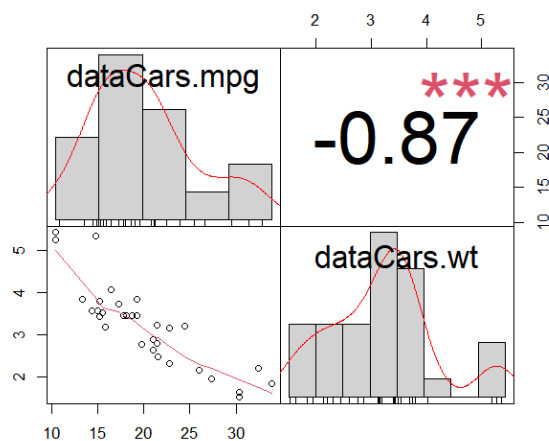




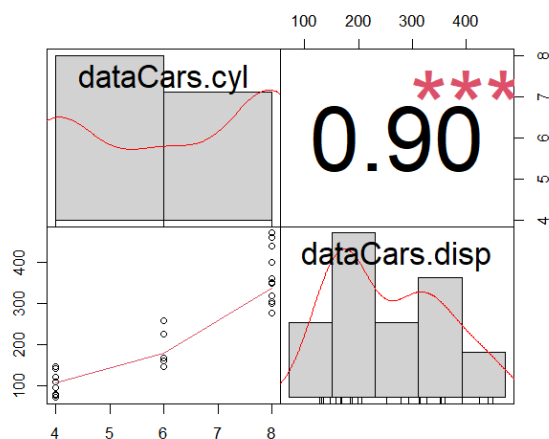
Esta correlación es similar a la anterior, indica que mientras más grande es el motor, más baja el rendimiento del auto.



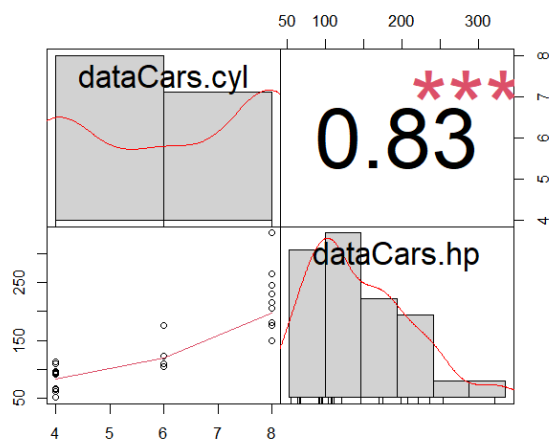
la correlación negativa entre el consumo de combustible y la potencia del motor sugiere una relación inversa, lo que indica que mientras más potente sea el motor, más bajo será el rendimiento del combustible



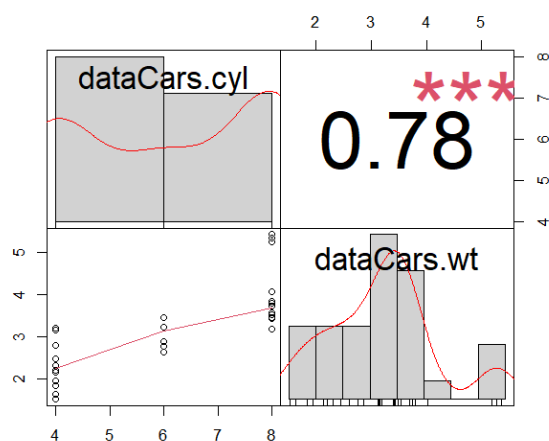
Con esta correlación también se observa que el peso tiene un impacto negativo en el rendimiento de combustible del auto, pues es necesario mover una mayor masa.



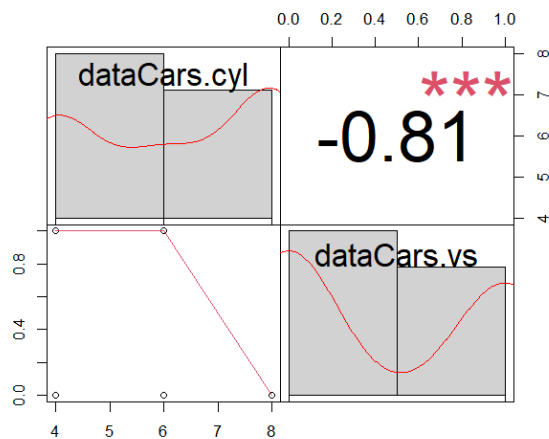
Esta relación es la más fuerte encontrada en este análisis; indica que, mientras más cilindros tiene un auto, su motor tiende a ser de un mayor tamaño, en cuestión de volumen.



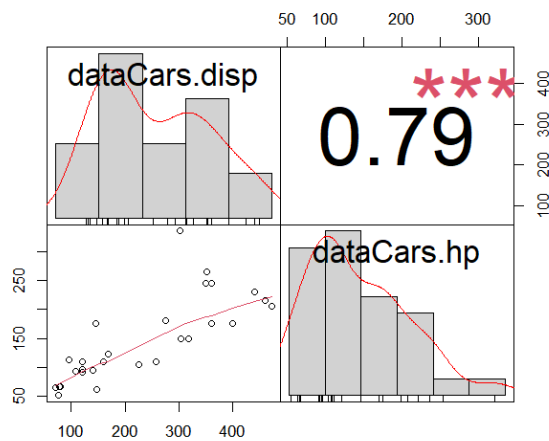
La correlación alta indica que cuantos más cilindros tenga el motor es más probable que el vehículo tenga una potencia más alta.



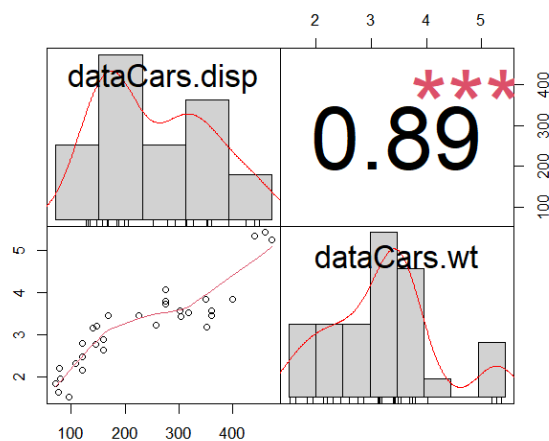
Esta correlación indica que los autos más pesados tienden a tener un mayor número de cilindros.



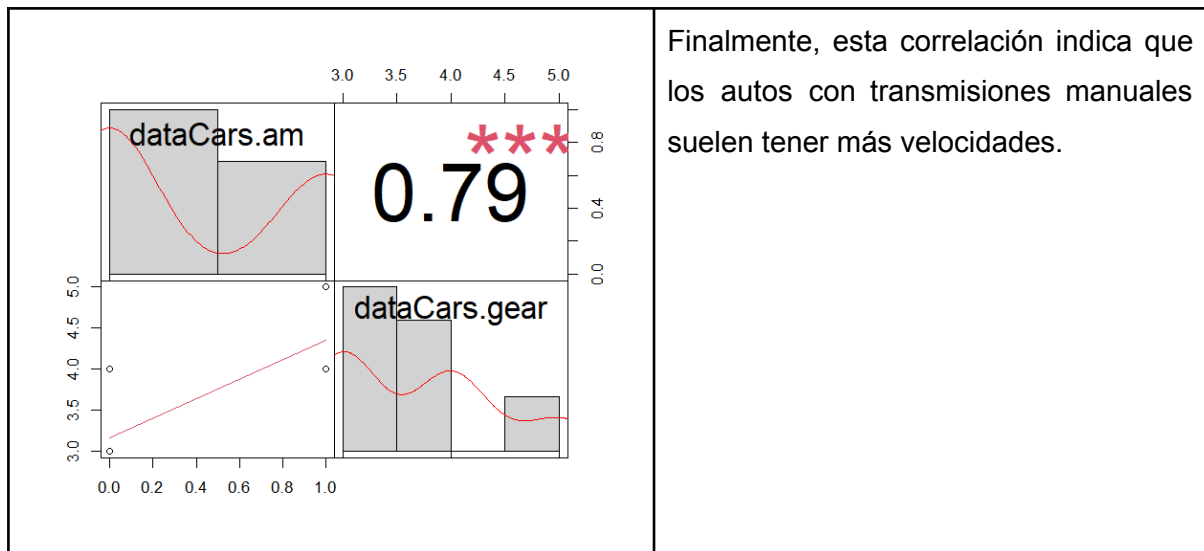
Indica que existe una relación inversa, si los vehículos tienen un mayor número de cilindros, tienden a tener motores en V.



Esta correlación indica que mientras más grande es el motor, en cuestión de desplazamiento, mayor es la potencia en caballos de potencia.



Esta correlación indica que mientras más grande es el motor del auto, es decir mientras mayor es su desplazamiento, mayor suele ser el peso del mismo.



Una vez realizada la correlación entre todas las variables obtenemos la siguiente tabla:

Columna1	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
mpg	1	-0.85216196	-0.84755138	-0.77616837	0.68117191	-0.86765938	0.41868403	0.66403892	0.59983243	0.48028476	-0.55092507
cyl	-0.85216196	1	0.90203287	0.83244745	-0.69993811	0.78249579	-0.59124207	-0.8108118	-0.52260705	-0.4926866	0.52698829
disp	-0.84755138	0.90203287	1	0.79094859	-0.71021393	0.88797992	-0.43369788	-0.71041589	-0.59122704	-0.5555692	0.39497686
hp	-0.77616837	0.83244745	0.79094859	1	-0.44875912	0.65874789	-0.70822339	-0.72309674	-0.24320426	-0.12570426	0.74981247
drat	0.68117191	-0.69993811	-0.71021393	-0.44875912	1	-0.71244065	0.09120476	0.44027846	0.71271113	0.69961013	-0.0907898
wt	-0.86765938	0.78249579	0.88797992	0.65874789	-0.71244065	1	-0.17471588	-0.55491568	-0.69249526	-0.583287	0.42760594
qsec	0.41868403	-0.59124207	-0.43369788	-0.70822339	0.09120476	-0.17471588	1	0.74453544	-0.22986086	-0.21268223	-0.65624923
vs	0.66403892	-0.8108118	-0.71041589	-0.72309674	0.44027846	-0.55491568	0.74453544	1	0.16834512	0.20602335	-0.56960714
am	0.59983243	-0.52260705	-0.59122704	-0.24320426	0.71271113	-0.69249526	-0.22986086	0.16834512	1	0.79405876	0.05753435
gear	0.48028476	-0.4926866	-0.5555692	-0.12570426	0.69961013	-0.583287	-0.21268223	0.20602335	0.79405876	1	0.27407284
carb	-0.55092507	0.52698829	0.39497686	0.74981247	-0.0907898	0.42760594	-0.65624923	-0.56960714	0.05753435	0.27407284	1

Tomamos como una fuerte correlación con valores mayores a 0.75 y como débil correlación a valores menores a 0.45, obteniendo los siguientes resultados:

Columna1	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
mpg	1	-0.85216196	-0.84755138	-0.77616837	0.68117191	-0.86765938	0.41868403	0.66403892	0.59983243	0.48028476	-0.55092507
cyl	-0.85216196	1	0.90203287	0.83244745	-0.69993811	0.78249579	-0.59124207	-0.8108118	-0.52260705	-0.4926866	0.52698829
disp	-0.84755138	0.90203287	1	0.79094859	-0.71021393	0.88797992	-0.43369788	-0.71041589	-0.59122704	-0.5555692	0.39497686
hp	-0.77616837	0.83244745	0.79094859	1	-0.44875912	0.65874789	-0.70822339	-0.72309674	-0.24320426	-0.12570426	0.74981247
drat	0.68117191	-0.69993811	-0.71021393	-0.44875912	1	-0.71244065	0.09120476	0.44027846	0.71271113	0.69961013	-0.0907898
wt	-0.86765938	0.78249579	0.88797992	0.65874789	-0.71244065	1	-0.17471588	-0.55491568	-0.69249526	-0.583287	0.42760594
qsec	0.41868403	-0.59124207	-0.43369788	-0.70822339	0.09120476	-0.17471588	1	0.74453544	-0.22986086	-0.21268223	-0.65624923
vs	0.66403892	-0.8108118	-0.71041589	-0.72309674	0.44027846	-0.55491568	0.74453544	1	0.16834512	0.20602335	-0.56960714
am	0.59983243	-0.52260705	-0.59122704	-0.24320426	0.71271113	-0.69249526	-0.22986086	0.16834512	1	0.79405876	0.05753435
gear	0.48028476	-0.4926866	-0.5555692	-0.12570426	0.69961013	-0.583287	-0.21268223	0.20602335	0.79405876	1	0.27407284
carb	-0.55092507	0.52698829	0.39497686	0.74981247	-0.0907898	0.42760594	-0.65624923	-0.56960714	0.05753435	0.27407284	1

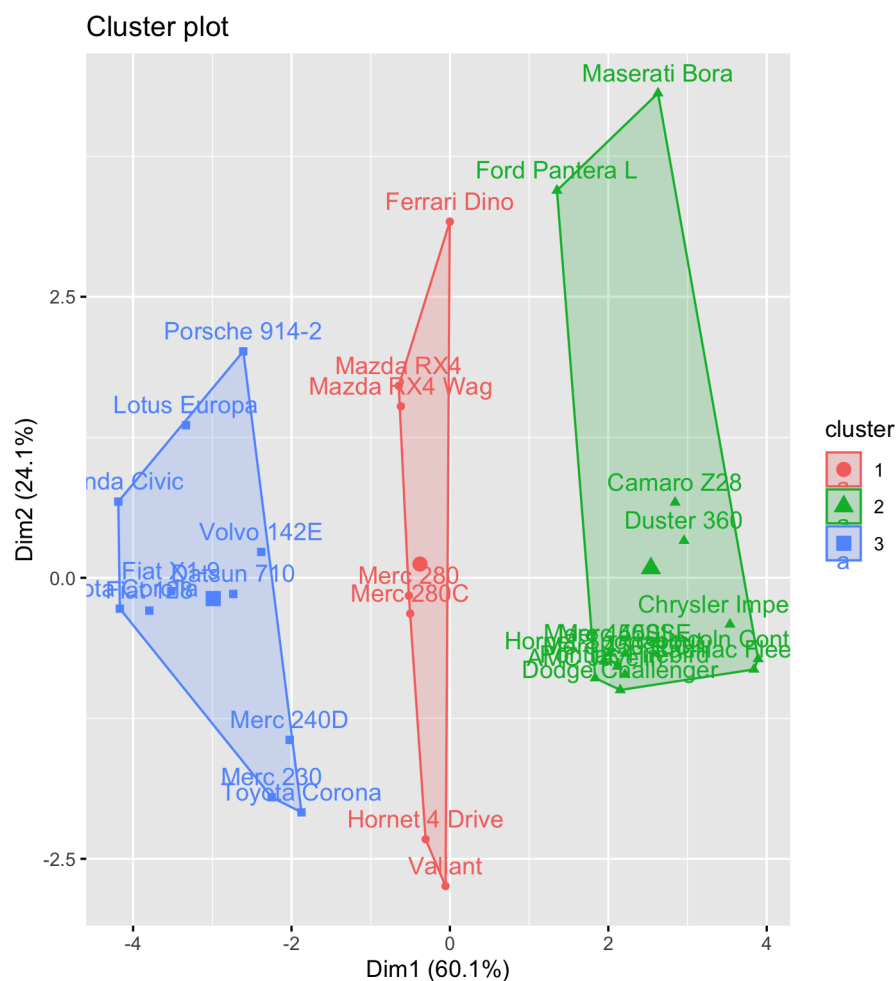
Cuando se presentan correlaciones con valores mayores a 0.75 nos sugiere una relación bastante fuerte entre las variables involucradas, dichas correlaciones significativas pueden proporcionar distintas inferencias o interpretaciones para entender mejor el significado de las relaciones fuertes y débiles, por ejemplo:

Relación entre cilindros y desplazamiento del motor (disp):

- Correlación: 0.902
- La fuerte correlación positiva entre el número de cilindros y el desplazamiento del motor nos dice que los vehículos con más cilindros tienden a tener motores de mayor capacidad (desplazamiento), en este caso es común que los vehículos con más cilindros, como los V8, tengan motores de mayor desplazamiento en comparación con vehículos con menos cilindros.

Relación entre peso del automóvil (wt) y desplazamiento del motor (disp):

- Correlación: 0.888
- La fuerte correlación positiva entre el peso del automóvil y el desplazamiento del motor se puede inferir que los vehículos más pesados tienden a tener motores de mayor capacidad, analizando esto, lógicamente se debe a que los vehículos más grandes y pesados requieren motores más potentes para proporcionar un buen rendimiento.



Referente al diagrama de agrupamiento o mejor conocido como Cluster Plot, nos ayuda a entender cómo es que son los conjuntos de datos, agrupándolos de manera de características similares. En este caso se dividió en tres clusters, los grandes implica que muchos con características similares, intermedias características y pocos con características similares. En el primer cluster son aquellos con una menor potencia y menor peso, en donde se podría colocar a los autos de uso diario, el segundo representa un menor peso en comparación con la potencia, es decir autos ligeros y con alta potencia, como autos deportivos, y finalmente el cluster alta potencia y alto peso, como camionetas y muscle cars.

Conclusiones

A través del análisis K-means podemos observar que los autos se pueden diferenciar en categorías distintas dependiendo de su peso y potencia. Por otra parte, con el análisis de correlación también es posible observar cómo influyen entre sí las características de los autos. La mayoría de estos atributos, como peso, potencia, desplazamiento y número de cilindros influyen principalmente en el rendimiento de combustible.

A través de las correlaciones entre peso y potencia, o peso y tamaño de combustible, es posible encontrar un sentido en la forma en la que se agrupan los autos en el diagrama de clúster.