

# Práctica Rendimiento Automotriz

Luis Manuel Rodarte Solórzano

9/11/2021

## DESCRIPCION:

El conjunto de datos corresponde a una muestra que contiene 356 observaciones. Las variables de la base de datos son las siguientes:

Columna 1: No. (número de identificación)

Columna 2: Marca

Columna 3: MAKE MODEL

Columna 4: CLASS

Columna 5: ENGINE SIZE (L)

Columna 6: CYLINDERS

Columna 7: TRANSMISSION

Columna 8: FUEL TYPE

Columna 9: CONSUMPTION CITY (L/100 KM)

Columna 10: CONSUMPTION HIGHWAY (L/100 KM)

Columna 11: CONSUMPTION COMBINED (L/100 KM)

Columna 12: \$ PER YEAR

Columna 13: CO2 EMISSIONS (g/km)

Columna 14: CO2 RATING

---

Resuelva: Basado en el vídeo de la práctica rendimiento, realice lo que se pide a continuación:

1. Tabla de frecuencias de la variable “Class” para la marca “Toyota” y grafique en un diagrama de barras. Interprete.
2. Tabla de frecuencias de la variable “Transmission” para la marca “Mazda” y grafique en un diagrama de barras. Interprete.
3. Obtenga las estadísticas básicas (media y varianza) de la variable “CO2 EMISSIONS (g/km)” de la marca “Porsche”. Realice el histograma correspondiente. Interprete.
4. Obtenga las estadísticas básicas (media y varianza) de la variable “CONSUMPTION COMBINED (L/100 KM)” de la marca “Dodge”. Realice el histograma correspondiente. Interprete.
5. De toda la base de datos ¿Qué vehículo le parece el mejor? Justifique su respuesta.

## Paso 1. Se carga la base de datos

```
library(readxl)
Rendimiento <- read_excel("Rendimiento.xlsx")
head(Rendimiento)

## # A tibble: 6 x 14
##       No Marca      `MAKE MODEL` CLASS `ENGINE SIZE (L)` CYLINDERS TRANSMISSION
##   <dbl> <chr>      <chr>      <chr>          <dbl>      <dbl> <chr>
## 1     1 Acura      ILX          C              2.4         4 AM8
## 2     2 Acura      RLX          M              3.5         6 AS6
## 3     3 Acura      RLX HYBRID   M              3.5         6 AM7
## 4     4 Acura      TLX          C              2.4         4 AM8
## 5     5 Acura      TLX SH-AWD   C              3.5         6 AS9
## 6     6 ALFA ROMEO 4C          T              1.8         4 AM6
## # ... with 7 more variables: FUEL TYPE <chr>,
## # CONSUMPTION CITY (L/100 KM) <dbl>, CONSUMPTION HIGHWAY (L/100 KM) <dbl>,
## # CONSUMPTION COMBINED (L/100 KM) <dbl>, $ PER YEAR <dbl>,
## # CO2 EMISSIONS (g/km) <dbl>, CO2 RATING <dbl>
```

## Paso 2. Análisis y resultados.

### 1. Tabla de frecuencias de la variable “Class” para la marca “Toyota”.

```
# Tabla de frecuencias
TOYOTA <- Rendimiento[Rendimiento$Marca == "TOYOTA",]
table(TOYOTA$CLASS)
```

```
##
##  C  M WS
##  4 12 6
```

Compact (Compactos)	Mid Size (Medianos)	Station Wagon Small (Station Wagon pequeños)
4	12	6

```
# Tabla de proporciones
tabla_1_prop<-prop.table(table(TOYOTA$CLASS))
Class_Toyota_1_prop<-as.data.frame(tabla_1_prop)
Class_Toyota_1_prop
```

```
##   Var1      Freq
## 1    C 0.1818182
## 2    M 0.5454545
## 3   WS 0.2727273
```

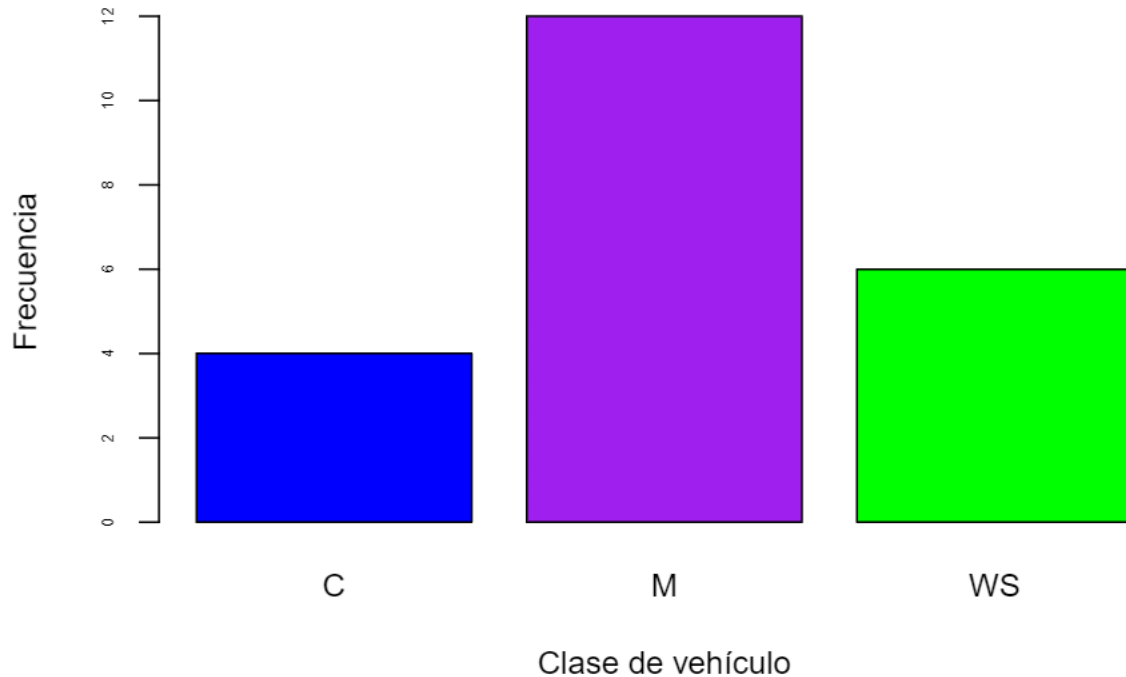
Compact (Compactos)	Mid Size (Medianos)	Station Wagon Small (Station Wagon pequeños)
0.18	0.55	0.27

### Diagrama de barras de la variable “Class” para la marca “Toyota”.

```
barplot(table(TOYOTA$CLASS),cex.axis=0.5,
        xlab="Clase de vehículo",
```

```
ylab="Frecuencia",
main="Diagrama de barras de la variable Class para la marca Toyota",
cex.main=1.1,
col =c("blue","purple","green"))
```

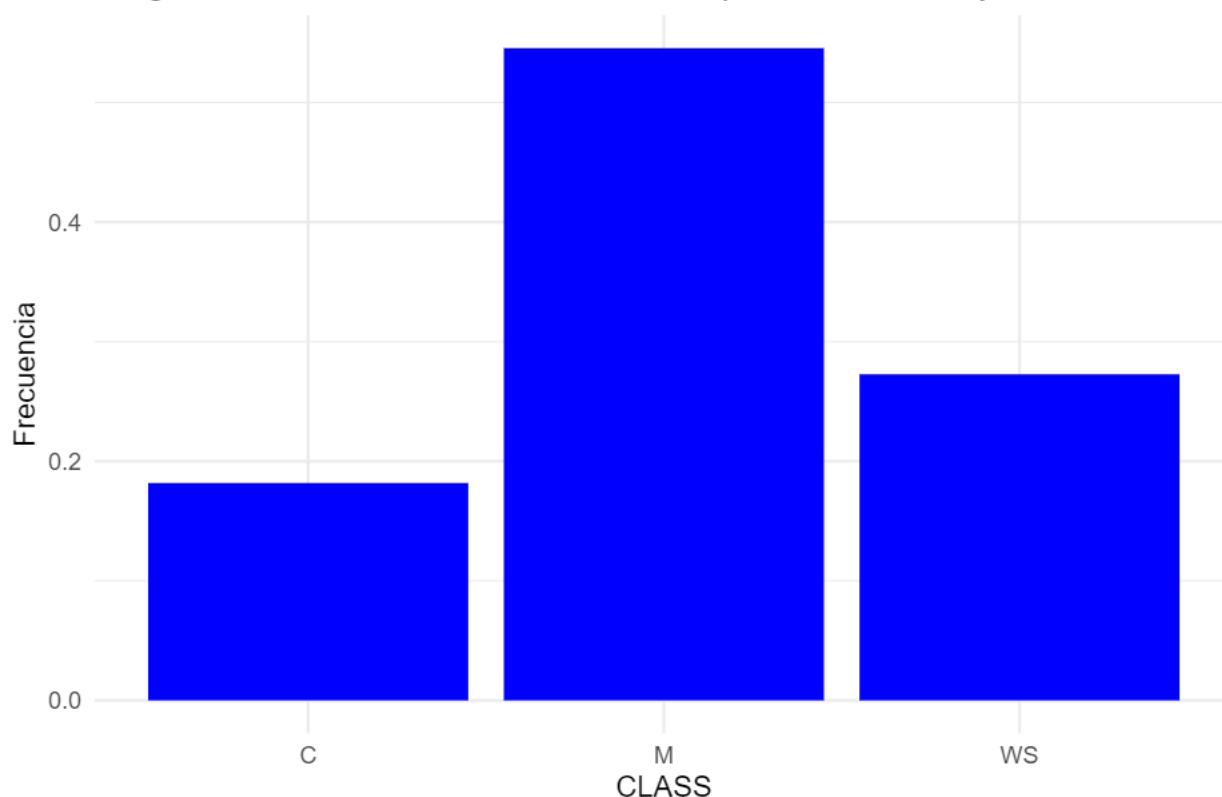
**Diagrama de barras de la variable Class para la marca Toyota**



```
library(ggplot2)

ggplot(TOYOTA) +
  aes(x = CLASS) +
  geom_bar(fill = "blue",aes(y = (..count..)/sum(..count..))) +
  theme_minimal()+
  labs(title = "Diagrama de barras de la variable Class para la marca Toyota")+
  xlab("CLASS")+
  ylab("Frecuencia")
```

Diagrama de barras de la variable Class para la marca Toyota



Clase de vehículo	Frecuencia	Frecuencia relativa
Compact (Compactos)	4	0.18
Mid Size (Medianos)	12	0.55
Station Wagon Small (Station Wagon pequeño)	6	0.27

## Interpretación

Como se puede observar en las Tabla de frecuencias y proporciones, así como en los diagrama de barras anteriores, hay en la muestra, respecto a la marca Toyota, un mayor número de vehículos de tipo Mid Size (Mediano), siendo estos el 54.5% del total de vehículos marca Toyota en la muestra, siendo del doble de los de tipo Station Wagon Small (de los cuales hay 6 en la muestra, lo que representa el 27% de los vehículos marca Toyota en la muestra); en tanto que vehículos de tipo compacto de la marca Toyota solo hay 4 (18%), es decir, hay menos compactos que medianos y Station Wagon pequeños, apenas cerca de uno de cada cinco vehículos marca Toyota en la muestra, son de tipo compacto.

## 2. Tabla de frecuencias de la variable “Transmission” para la marca “Mazda”.

```
# Tabla de frecuencias
MAZDA <- Rendimiento[Rendimiento$Marca == "MAZDA",]
table(MAZDA$TRANSMISSION)
```

```
##
##  A9 AS6 AS7  M6
##   1   6  13   3
```

Transmisión	Frecuencia
A9 (Automatica, 9 cambios o vel.)	1
AS6 (Automática SelectShift de 6 vel.)	6
AS7 (Automática SelectShift de 7 vel.)	13
M6 (Manual o estándar, 6 vel.)	3

*# Tabla de proporciones*

```
tabla_2_prop<-prop.table(table(MAZDA$TRANSMISSION))
Transm_Mazda_2_prop<-as.data.frame(tabla_2_prop)
Transm_Mazda_2_prop
```

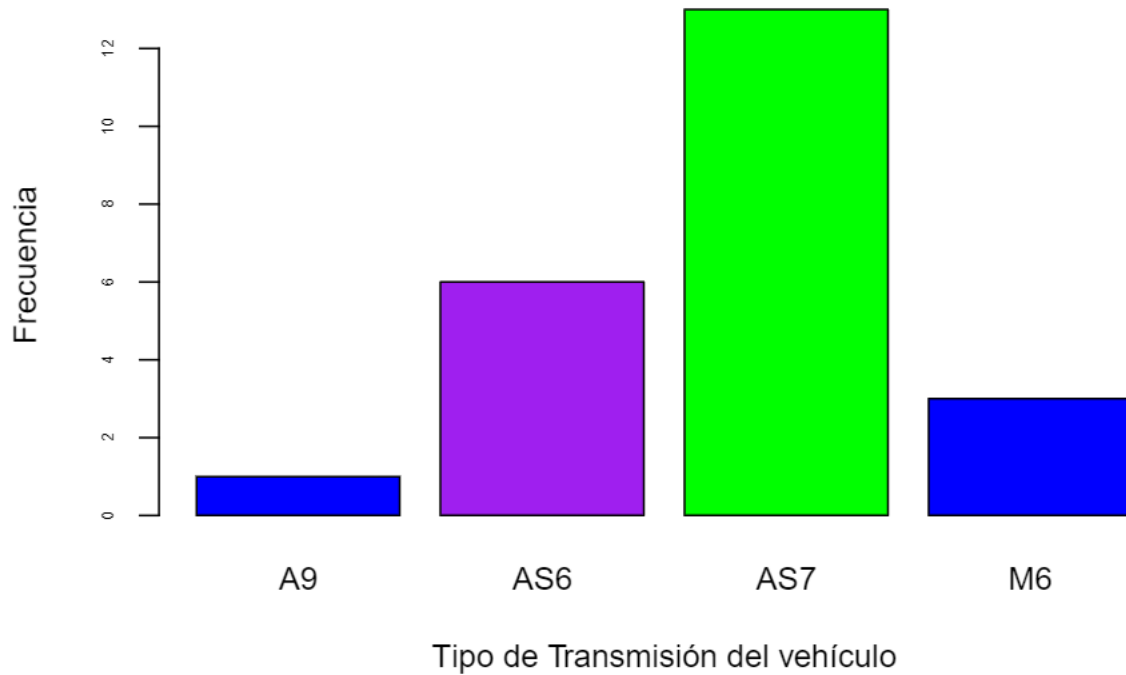
```
##   Var1      Freq
## 1   A9 0.04347826
## 2  AS6 0.26086957
## 3  AS7 0.56521739
## 4   M6 0.13043478
```

Transmisión	Frecuencia relativa
A9 (Automatica, 9 cambios o vel.)	0.04
AS6 (Automática SelectShift de 6 vel.)	0.26
AS7 (Automática SelectShift de 7 vel.)	0.57
M6 (Manual o estándar, 6 vel.)	0.13

Diagrama de barras de la variable “Transmission” para la marca “Mazda”.

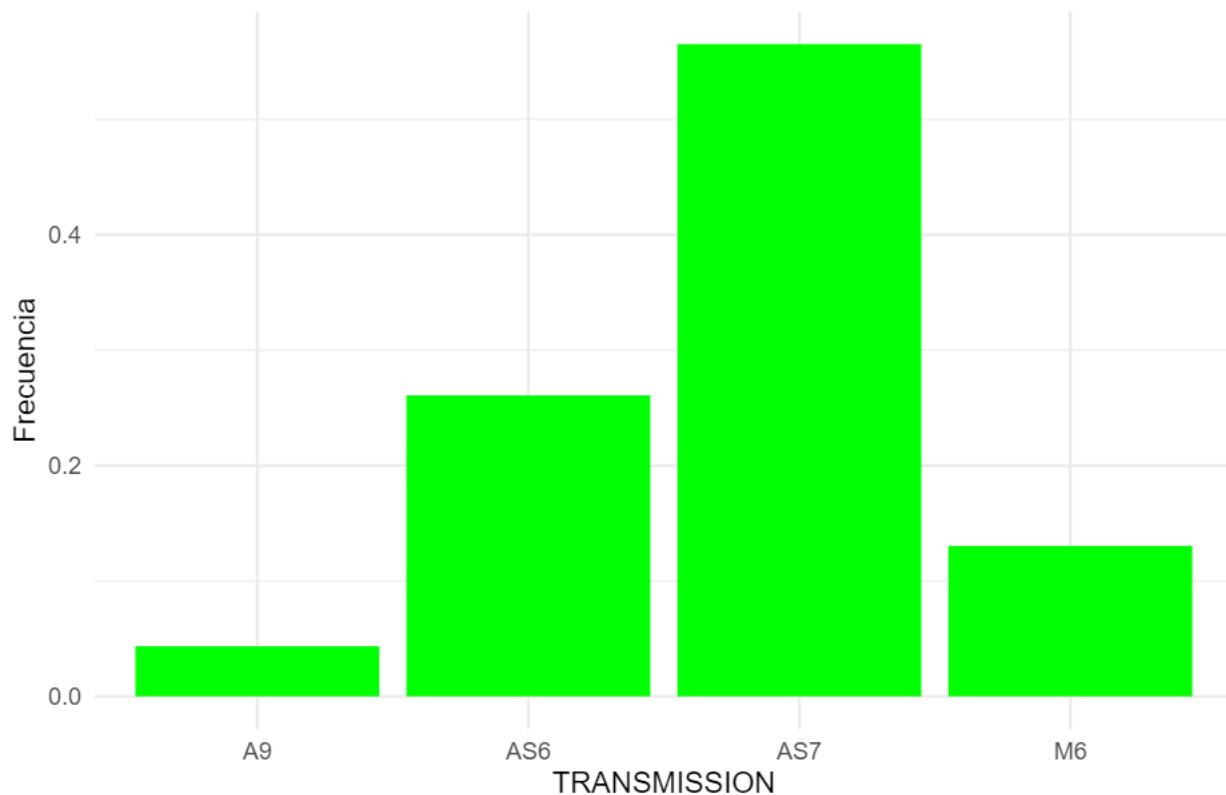
```
barplot(table(MAZDA$TRANSMISSION),cex.axis=0.5,
        xlab="Tipo de Transmisión del vehículo",
        ylab="Frecuencia",
        main="Diagrama de barras de la variable Transmisión para la marca Mazda",
        cex.main=1.1, col =c("blue","purple","green"))
```

## Diagrama de barras de la variable Transmisión para la marca Mazda



```
ggplot(MAZDA) +  
  aes(x = TRANSMISSION) +  
  geom_bar(fill = "green", aes(y = (..count..)/sum(..count..))) +  
  theme_minimal()+  
  labs(title = "Diagrama de barras de la variable Transmisión para la marca Mazda")+  
  xlab("TRANSMISSION")+  
  ylab("Frecuencia")
```

Diagrama de barras de la variable Transmisión para la marca Mazda



Transmisión	Frecuencia	Frecuencia relativa
A9 (Automatica, 9 cambios o vel.)	1	0.04
AS6 (Automática SelectShift de 6 vel.)	6	0.26
AS7 (Automática SelectShift de 7 vel.)	13	0.57
M6 (Manual o estándar 6 vel.)	3	0.13

## Interpretación

Como se puede observar en las tablas de frecuencia y proporciones, así como en los diagramas de barras anteriores, respecto a la variable Transmisión en la marca Mazda, hay en la muestra un mayor número de vehículos con transmisión de tipo AS7, con un poco más de la mitad de los casos (57%), es decir, aproximadamente uno de cada dos vehículos de la marca Mazda tienen este tipo de transmisión; el 26% (aproximadamente uno de cada cuatro) tiene una transmisión de tipo AS6 (automática SelectShif de 6 velocidades); y solo el 13% (3 vehículos) poseen transmisión manual de 6 velocidades. Por último, solo un vehículo de la marca Mazda, en la muestra, posee una transmisión de tipo A9.

## 3. Estadísticas básicas (media y varianza) de la variable “CO2 EMISSIONS (g/km)” de la marca “Porsche”.

```
PORSCHE <- Rendimiento[Rendimiento$Marca == "PORSCHE",]
mean(PORSCHE$`CO2 EMISSIONS (g/km)`, na.rm = TRUE)
```

```
## [1] 268.5161
```

```
var(PORSCHE$`CO2 EMISSIONS (g/km)`, na.rm = TRUE)
```

```
## [1] 1435.525
```

```
sd(PORSCHE$`CO2 EMISSIONS (g/km)`, na.rm = TRUE)
```

```
## [1] 37.88832
```

```
summary(PORSCHE$`CO2 EMISSIONS (g/km)`, na.rm = TRUE)
```

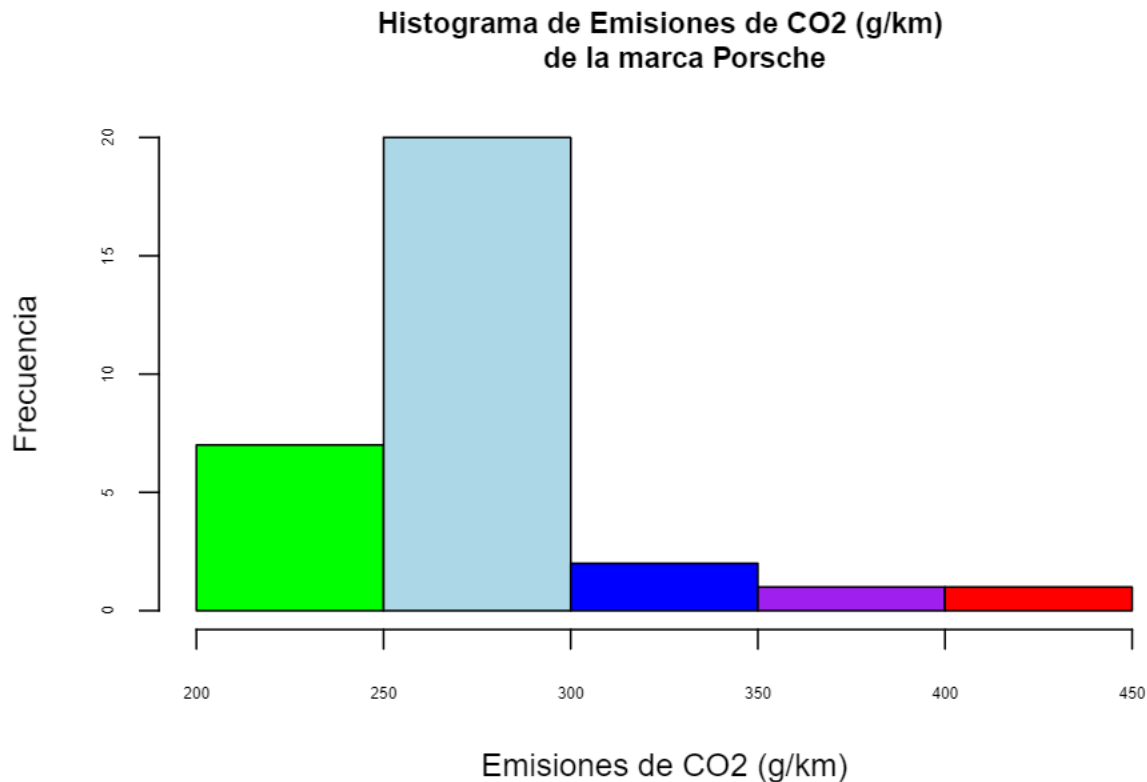
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      236.0   251.0   258.0   268.5   263.5   404.0
```

Estadístico	Valor
Media	268.5
Varianza	1435.5
Desviación estándar	37.9
Mínimo	236.0
Primer cuartil	251.0
Mediana	258.0
Tercer cuartil	263.5
Máximo	404.0

Histograma de la variable “CO2 EMISSIONS (g/km)” de la marca “Porsche”.

```
hist(PORSCHE$`CO2 EMISSIONS (g/km)`, breaks = 5,
     main="Histograma de Emisiones de CO2 (g/km)
de la marca Porsche",cex.main=0.9,
     xlab="Emisiones de CO2 (g/km)",
     ylab="Frecuencia",cex.axis=0.5,
     col=c("green","light blue","blue","purple","red"))
```






---

Estadístico	Media	Varianza	Desviación estándar	Mínimo	Q1	Mediana	Q3	Máximo
Valor	268.5	1435.5	37.9	236.0	251.0	258.0	263.5	404.0

---

## Interpretación

El promedio de emisiones de CO2 para la marca Porsche en la muestra es de 268.5 g/km, con una desviación estándar de 37.9 g/km. En esta marca de vehículo en la muestra, el valor mínimo de emisiones de CO2 fue de 236 g/km, mientras que el máximo fue de 404 g/km. Por otro lado, el 25% de los vehículos marca Porsche de la muestra, reportó un valor menor o igual a 251 g/km de emisiones de CO2, mientras que el 50% de estos vehículos (uno de cada dos) reportaron un valor menor o igual que 258 g/km de emisiones de CO2, y el 75% un valor menor o igual que 263.5 g/km, es decir, un 25% de los vehículos marca Porsche en la muestra, reportaron un valor de emisiones de CO2 entre los 263.5 y los 404 g/km.

## 4. Estadísticas básicas (media y varianza) de la variable “CONSUMPTION COMBINED (L/100 KM)” de la marca “Dodge”.

```
DODGE <- Rendimiento[Rendimiento$Marca == "DODGE",]
mean(DODGE$`CONSUMPTION COMBINED (L/100 KM)`, na.rm = TRUE)
```

```
## [1] 11.79375
```

```
var(DODGE$`CONSUMPTION COMBINED (L/100 KM)`, na.rm = TRUE)
```

```
## [1] 5.312088
```

```
sd(DODGE$`CONSUMPTION COMBINED (L/100 KM)`, na.rm = TRUE)
```

```
## [1] 2.304797
```

```
summary(DODGE$`CONSUMPTION COMBINED (L/100 KM)`, na.rm = TRUE)
```

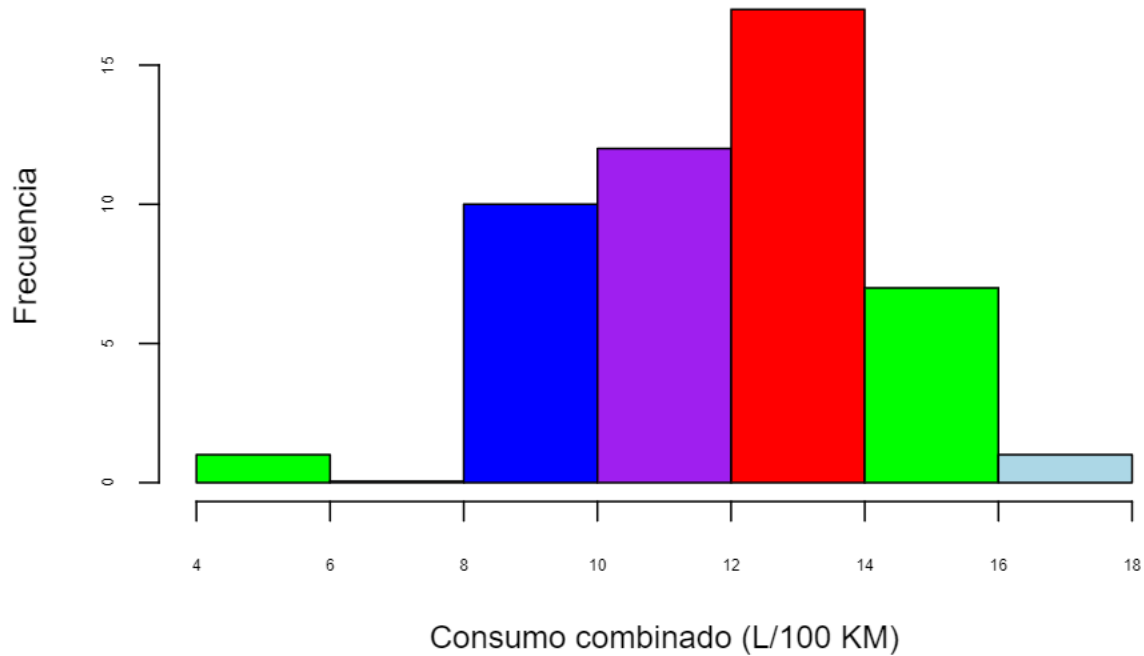
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      5.50   10.25   12.30   11.79   13.18   16.30
```

Estadístico	Valor
Media	11.79
Varianza	5.31
Desviación estándar	2.30
Mínimo	5.50
Primer cuartil	10.25
Mediana	12.30
Tercer cuartil	13.18
Máximo	16.30

Histograma de la variable “CONSUMPTION COMBINED (L/100 KM)” de la marca “Dodge”.

```
# Histograma
hist(DODGE$`CONSUMPTION COMBINED (L/100 KM)`,
     breaks = 5,
     main="Histograma de Consumo Combinado (L/100 KM) de la marca Dodge",
     cex.main=1,
     xlab="Consumo combinado (L/100 KM)",
     ylab="Frecuencia",cex.axis=0.5,
     col=c("green","light blue","blue","purple","red"))
```

## Histograma de Consumo Combinado (L/100 KM) de la marca Dodge



Estadístico	Media	Varianza	Desviación estándar	Mínimo	Q1	Mediana	Q3	Máximo
Valor	11.79	5.31	2.30	5.50	10.25	12.30	13.18	16.30

### Interpretación

Como se puede observar en el análisis descriptivo básico y en el histograma, el promedio del consumo combinado en los vehículos marca Dodge de la muestra es de 11.79 litros por cada 100 kilómetros, con una desviación estándar de 2.30 l/100 km. El consumo combinado mínimo es de 5.50 l/100 km, y el máximo de 16.30 l/100 km. Por otro lado, el 25% de los vehículos marca Dodge en la muestra tiene un consumo combinado menor o igual a 10.25 l/100 km, el 50% un consumo combinado menor o igual a 12.30 l/100 km, y el 75% un consumo combinado menor o igual a 13.18 l/100 km, es decir, un 25% de los vehículos marca Dodge tiene un consumo combinado entre 13.18 y 16.30 litros por cada 100 kilómetros.

Algo que llama la atención es que no hay en la muestra ningún vehículo de la marca Dodge que tenga un consumo combinado de entre 6.0 l/100 km y 8.0 l/100 km, lo cual también se observa en el histograma. Específicamente el segundo mínimo de consumo mínimo combinado es de 8.1 l/100 km y corresponde a un modelo Dart.

### 5. De toda la base de datos ¿Qué vehículo le parece el mejor? Justifique su respuesta.

Para responder esta pregunta y considerando a las variables Consumo combinado, Emisiones de CO2 y Costo anual; como las variables principales o más importantes, se realizaron los siguientes análisis:

```
min(Rendimiento$`CONSUMPTION COMBINED (L/100 KM)`, na.rm = TRUE)
```

```
## [1] 5.5
```

```

min(Rendimiento$`CO2 EMISSIONS (g/km)`, na.rm = TRUE)

## [1] 130

min(Rendimiento$`$ PER YEAR`, na.rm = TRUE)

## [1] 1199

Rendimiento_li <- na.omit(Rendimiento)

library(tidyverse)
filtro <- Rendimiento_li %>%
  select(Marca, `MAKE MODEL`, CLASS,
         `ENGINE SIZE (L)`, CYLINDERS,
         `CONSUMPTION COMBINED (L/100 KM)`,
         `CO2 EMISSIONS (g/km)`,
         `$ PER YEAR`)
head(filtro)

## # A tibble: 6 x 8
##   Marca      `MAKE MODEL` CLASS `ENGINE SIZE (L)` CYLINDERS `CONSUMPTION COMBI-
##   <chr>      <chr>      <chr>      <dbl>      <dbl>      <dbl>
## 1 Acura      ILX            C          2.4        4          8.1
## 2 Acura      RLX            M          3.5        6          10
## 3 Acura      RLX HYBRID     M          3.5        6          7.7
## 4 Acura      TLX            C          2.4        4          8.3
## 5 Acura      TLX SH-AWD     C          3.5        6          9.6
## 6 ALFA ROMEO 4C          T          1.8        4          8.4
## # ... with 2 more variables: CO2 EMISSIONS (g/km) <dbl>, $ PER YEAR <dbl>

Res <- filtro[filtro$`CONSUMPTION COMBINED (L/100 KM)`<=6
               & filtro$`CO2 EMISSIONS (g/km)`<=150
               & filtro$`$ PER YEAR`<=1300,]
head(Res)

## # A tibble: 4 x 8
##   Marca      `MAKE MODEL` CLASS `ENGINE SIZE (L)` CYLINDERS `CONSUMPTION C-
##   <chr>      <chr>      <chr>      <dbl>      <dbl>      <dbl>
## 1 DODGE      FUSION HYBRID M          2          4          5.5
## 2 LINCOLN    MKZ HYBRID     M          2          4          5.8
## 3 TOYOTA     CAMRY HYBRID LE M          2.5        4          5.7
## 4 TOYOTA     CAMRY HYBRID XLE/SE M 2.5        4          5.9
## # ... with 2 more variables: CO2 EMISSIONS (g/km) <dbl>, $ PER YEAR <dbl>

```

## Conclusión

Ordenando los datos de menor a mayor en función de la variable Consumo Combinado, se observa que los de menor consumo combinado y menor costo anual son los vehículos híbridos, siendo **el vehículo modelo Fusion Hybrid de la marca Dodge** el de menor consumo combinado, con 5.5 litros por cada 100 kilómetros, menor costo anual (1,199 dólares) y también el de menores emisiones de CO2 (130 g/km).

Por otro lado, y quizás considerando también variables como tamaño o potencia del motor e incluso velocidad de depreciación (variable no medida en el estudio), quizás a mi en particular me parecería mejor **el vehículo modelo Camry Hybrid LE de la marca Toyota**, ya que tiene un consumo combinado de 5.7 litros por cada 100 km, un motor 2.5 l. de 4 cilindros, un costo anual de 1,243 dólares y solo 134 g/km de emisiones de CO2.

Finalmente, es necesario mencionar que para dar una mejor respuesta respecto a qué vehículo parece ser

el mejor, sería conveniente tomar una muestra más grande y separar también por híbridos y no híbridos, posteriormente realizar comparaciones por grupos de acuerdo a clase de vehículo, transmisión, potencia del motor, cilindros, híbrido/no híbrido, consumo combinado, costo anual y emisiones de CO<sub>2</sub>; ya que como se observó en la base de datos y en este último análisis un motor pequeño no significa necesariamente menor consumo combinado de combustible, pero quizás si menor potencia, tal es el caso del modelo Fit de la marca Honda, cuyo motor es 1.5 l., de cuatro cilindros, pero, al menos en el caso de los dos vehículos de este tipo que hay en la muestra, estos tienen un consumo combinado mayor de 6 l/100 km, específicamente de 6.4 y 6.8 l/100 km, y un costo anual de 1,395 y 1,482 dólares.