

▼ Actividad

-Natalia Taboada A01570464

-José M. López A00571541

-Bruno Rojas Torres A01283786

Entregar

Archivo PDF de la actividad y la liga de la actividad en su repostitorio.

Nota:

Todas las tareas entregadas fuera de la fecha limite se califican sobre 50 de los 100 puntos posibles.

```
# Si trabajamos en Google Colaboratory corremos las siguientes lineas de código
from google.colab import drive
drive.mount('/content/drive/')
```

🔗 Mounted at /content/drive/

```
# Nos cambiamos a la carpeta donde tengamos el repositorio
%cd 'drive/MyDrive/SemanaTec1/Repos/arte-analitica'
```

```
/content/drive/MyDrive/SemanaTec1/Repos/arte-analitica
```

▼ Highway MPG dataset

Este dataset contiene variables medidas por la agencia de protección ambiental de 38 modelos de vehículo diferentes de 1999 a 2008. Las variables que se registraron fueron:

- **manufacturer:** nombre del fabricante.
- **model:** nombre del modelo.
- **displ:** desplazamiento del motor, en litros.
- **year:** año de fabricación.
- **cyl:** número de cilindros.
- **trans:** tipo de transmisión.
- **drv:** tipo de tracción, f-delantera, r-trasera, 4- 4 llantas
- **cty:** rendimiento del motor en ciudad, en millas por galón.

- **hwy**: rendimiento del motor en carretera, en millas por galón.
- **fl**: tipo de combustible.
- **class**: tipo de vehículo.

```
# Carga las librerías y datos
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
sns.set()
```

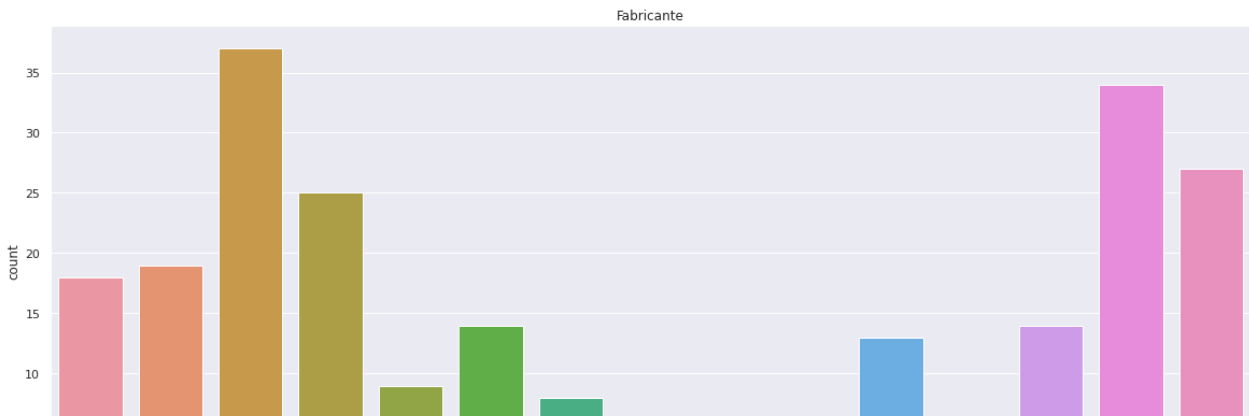
```
mpg_df = pd.read_csv('data/auto-mpg.csv')
```

```
mpg_df.head()
```

	manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
0	audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
1	audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
2	audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
3	audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
4	audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact

```
# ¿Cuántas observaciones hay de cada fabricante? Muéstralo en un gráfico.
fig = plt.figure(figsize=(20, 8))
mpg_df['manufacturer'].value_counts()
sns.countplot(data=mpg_df , x='manufacturer')
plt.title('Fabricante')
```

```
Text(0.5, 1.0, 'Fabricante')
```

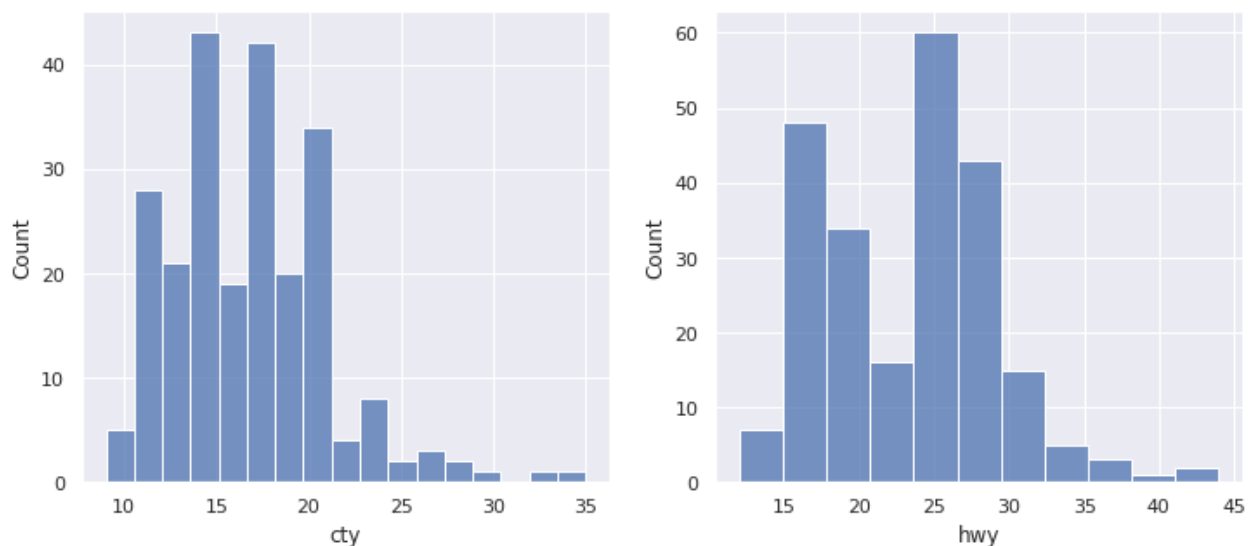


```
# Haz un histograma de las variables numéricas cty, hwy y haz una comparación en un gráfico 1
fig, axs = plt.subplots(1,2, figsize=(12, 5))
```

```
datos = mpg_df
sns.histplot(datos, x='cty', ax=axs[0])
```

```
datos = mpg_df
sns.histplot(datos, x='hwy', ax=axs[1])
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f8a985b4f90>
```

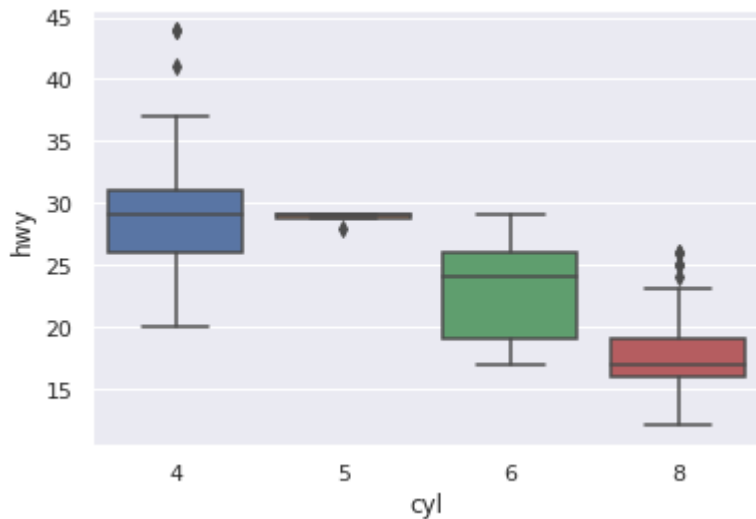


```
# ¿Cómo se comparar el rendimiento en carretera (hwy) contra el número de cilindros?
# Haz un gráfico que represente esta relación.
```

```
# cyl es una variable numérica, pero como tiene pocos valores podemos tomarla como si fuera c
# ¿De qué otra forma se puede representar la relación de la pregunta anterior?
```

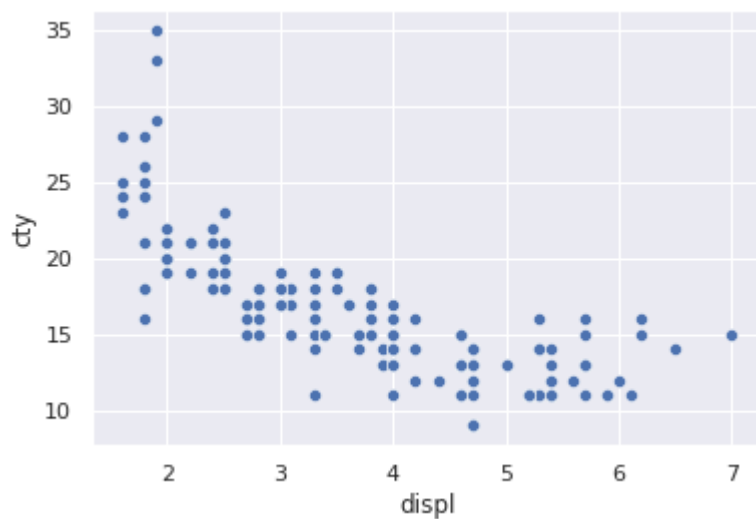
```
sns.boxplot(data=mpg_df, v= 'hwy', x='cyl')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f8a98278690>
```



#Cuál es la relación entre el desplazamiento del motor y el rendimiento dentro de la ciudad?
 sns.scatterplot(data=datos, x='displ', y='cty')

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f8a98181dd0>
```



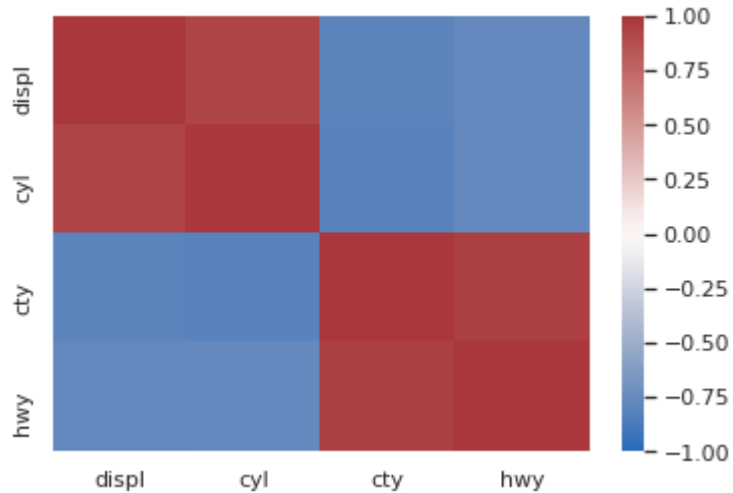
¿Cómo se correlacionan las variables numéricas? Muestra esta correlación en un gráfico. (No # variable numérica)

```
# Sugerencia: usa la paleta de colores 'vlag' o una paleta DIVERGENTE
matriz_corr= mpg_df.drop('year',axis=1).corr()
matriz_corr
```

	displ	cyl	cty	hwy
displ	1.000000	0.930227	-0.798524	-0.766020

```
sns.heatmap(data=matriz_corr, vmin=-1, vmax=1, cmap='vlag')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f8a98049190>
```



```
# ¿Cuales variables tienen una fuerte relación positiva entre sí y cuáles tienen una fuerte r
+= hwy, displ, cyl
-= hwy, cty, hwy
```

```
# ¿Los fabricantes mejoran el rendimiento en ciudad conforme pasa el tiempo? Haz una tabla re
# rendimiento promedio y grafica los resultados.
```

```
fig = plt.figure(figsize=(15, 8))
sns.scatterplot(data=datos, x='manufacturer', y='cty', hue='year', palette='viridis')
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f8a8f4a06d0>

