

UNIVERSIDAD METROPOLITANA CASTRO CARAZO
FACULTAD DE INFORMÁTICA

Técnico en análisis de datos

Seaborn

Eduardo Campos Guerrero
2 0680 0447

TRABAJO DE INVESTIGACION
CURSO DE Análisis de Datos

SAN JOSÉ, COSTA RICA
Agosto – 2023

¿Qué es Seaborn?

Seaborn es una librería para el lenguaje de programación de Python diseñada para la parte de creación de gráficos la cual fue construida sobre la famosa librería Matplotlib.

Seaborn tiene un diseño más moderno que Matplotlib con múltiples opciones de diseños y un API más intuitiva.

Si bien Matplotlib es una excelente herramienta para la generación de visualizaciones básicas, los utilizados con esta librería no atractivos estéticamente. Seaborn está pensada y diseñada para atacar ese punto negativo de Matplotlib ya que posee una gran cantidad de estilos y de paletas de colores que facilitan la personalización de las visualizaciones.

Seaborn también ofrece una característica la cual Matplotlib no posee, que sería la capacidad de poder trabajar estadísticas complejas con variables categóricas.

¿Como se instala Seaborn y que dependencias tiene?

Para poder utilizar la herramienta de Seaborn antes se deben de instalar varias dependencias como los son:

- Poseer la versión de Python 3.6 o superior
- Poseer la Librería de Numpy
- Poseer la librería de Pandas
- Poseer la Librería de Matplotlib

Una vez que se tenga esos requisitos ya se puede proceder con la instalación de la librería de Seaborn de la siguiente manera:

- En la ventana del terminal usamos el comando “pip install seaborn”
- Y otra forma sería utilizando el comando conda desde la terminal, “conda install seaborn”

Ejemplos de la utilización de la Librería de Seaborn

En este ejemplo vamos a ver cómo se puede generar un Histograma por grupos utilizando la librería de Seaborn en conjunto con numpy y pandas

Realizamos las importaciones de las librerías de numpy y pandas y generamos los datos de los grupos:

```
import numpy as np
import pandas as pd

# Simulación de datos
rng = np.random.RandomState(1)
x1 = rng.normal(0, 1, size = 500)
x2 = rng.normal(3, 1.5, size = 500)
x = np.concatenate((x1, x2), axis = 0)
grupo = np.repeat(np.array(["G1", "G2"]), [500, 500], axis = 0)
df = {'x': x, 'grupo': grupo}

# Data frame pandas
df = pd.DataFrame(data = df)
```

X	Grupo
1,62	G1
-0,61	G1
-0,53	G1
-1,07	G1
0,87	G1
-2,30	G1

1-6 de 1000 filas < 1 2 3 4 5 ... 167 >

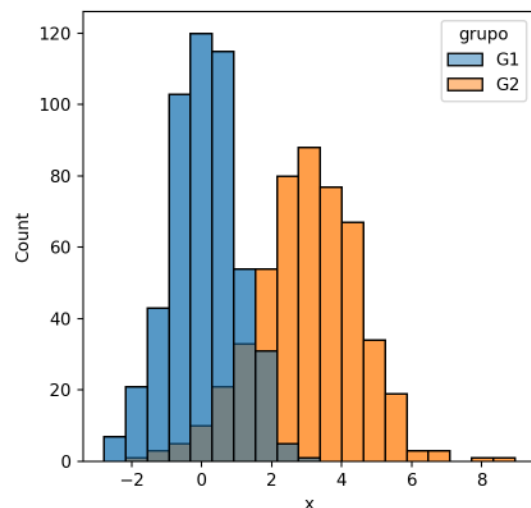
En este ejemplo se utilizará varias funciones de la librería de Seaborn como lo es "histplot", antes de generar el código de la grafica se debe de importar la librería de Seaborn con la línea de código "import seaborn as sns"

Para crear un histograma por grupo en seaborn tendrás que pasar tus datos a la función `histplot` y la variable categórica que represente cada grupo al argumento `hue`.

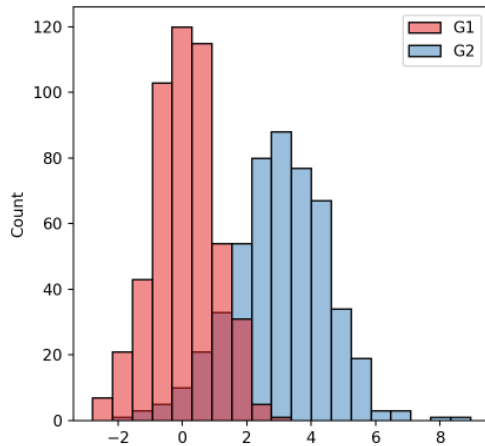
```
import seaborn as sns

# Histograma
sns.histplot(x = x, hue = grupo)

# Equivalente a:
sns.histplot(x = "x", hue = "grupo", data = df)
```



Siguiendo ese mismo ejemplo, el histograma se puede personalizar con diferentes tipos de argumentos, en este caso se le dará una personalización de color con el comando "palette"

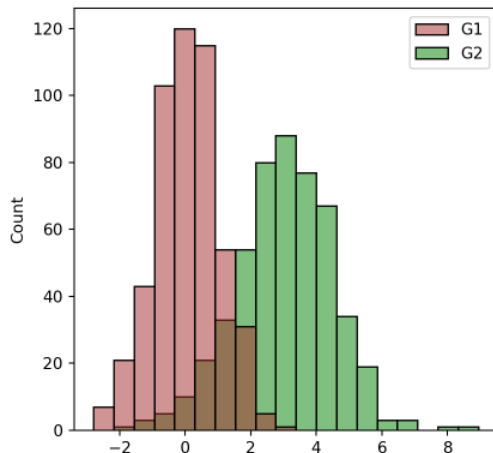


Paleta de colores

```
import seaborn as sns

# Histograma
sns.histplot(x = x, hue = grupo,
             palette = "Set1")
```

O bien, se puede indicar los tipos de colores que se desean para cada caso



Colores personalizados

```
import seaborn as sns

colores = {'G1': 'brown', 'G2': 'green'}

# Histograma
sns.histplot(x = x, hue = grupo,
             palette = colores)
```

En conclusión, Seaborn es una poderosa librería de visualización de datos en Python que proporciona gráficos personalizables para explorar y comprender de una manera diferente los datos. Seaborn además es fácil de usar y ofrece varias características avanzadas que la hacen una elección perfecta para poder representar gráficamente los datos.