

Seaborn

Creación de visualizaciones

CertiDevs

Índice de contenidos

1. Introducción	1
2. Datasets de seaborn	1
3. Cargar un dataset	1
4. Diagrama de dispersión (scatterplot)	1
5. Gráfico de barras (barplot)	1
6. Gráfico de línea (lineplot)	2
7. Histograma (histplot)	2
8. Gráfico de violín (violinplot)	2
9. Gráfico de caja (boxplot)	3
10. Gráfico de densidad (kdeplot)	3
11. Diagrama de dispersión con línea de regresión (regplot)	4
12. Gráfico de enjambre (swarmplot)	4
13. Gráfico de correlación (heatmap)	4

1. Introducción

Seaborn viene con una serie de **conjuntos de datos** integrados que se pueden utilizar para aprender y practicar cómo crear gráficos.

2. Datasets de seaborn

Seaborn incorpora distintos **datasets** para poder practicar y aprender a crear visualizaciones.

Para ver la lista de **datasets disponibles**, podemos ejecutar el siguiente código:

```
import seaborn as sns

print(sns.get_dataset_names())
```

3. Cargar un dataset

Para **cargar un dataset**, podemos usar la función `load_dataset()` de seaborn.

Por ejemplo, para cargar el dataset 'tips', podemos ejecutar el siguiente código:

```
import seaborn as sns

data = sns.load_dataset('tips')
```

4. Diagrama de dispersión (scatterplot)

Un **diagrama de dispersión** muestra la **relación entre dos variables continuas**.

Usaremos el conjunto de datos 'tips' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('tips')

sns.scatterplot(x='total_bill', y='tip', data=data)
plt.show()
```

5. Gráfico de barras (barplot)

Un **gráfico de barras** muestra la relación entre una **variable categórica** y una **variable numérica**.

Usaremos el conjunto de datos 'titanic' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('titanic')

sns.barplot(x='class', y='fare', data=data)
plt.show()
```

6. Gráfico de línea (lineplot)

Un **gráfico de línea** muestra la relación entre dos variables numéricas, donde una de ellas suele ser el tiempo.

Usaremos el conjunto de datos 'flights' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('flights')
data = data.pivot('month', 'year', 'passengers')

sns.lineplot(data=data)
plt.show()
```

7. Histograma (histplot)

Un **histograma** muestra la **distribución** de una variable numérica.

Usaremos el conjunto de datos 'penguins' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('penguins')

sns.histplot(x='bill_length_mm', data=data, bins=20)
plt.show()
```

8. Gráfico de violín (violinplot)

Un **gráfico de violín** muestra la distribución de una variable numérica a lo largo de diferentes categorías.

Usaremos el conjunto de datos 'iris' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('iris')

sns.violinplot(x='species', y='sepal_length', data=data)
plt.show()
```

9. Gráfico de caja (boxplot)

Un **gráfico de caja** muestra la distribución de una variable numérica a lo largo de diferentes categorías e indica la mediana, los cuartiles y los valores atípicos.

Usaremos el conjunto de datos 'diamonds' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('diamonds')

sns.boxplot(x='cut', y='price', data=data)
plt.show()
```

10. Gráfico de densidad (kdeplot)

Un **gráfico de densidad** estima y muestra la distribución de probabilidad de una variable numérica.

Usaremos el conjunto de datos 'penguins' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('penguins')

sns.kdeplot(x='bill_length_mm', data=data)
plt.show()
```

11. Diagrama de dispersión con línea de regresión (regplot)

Un **diagrama de dispersión** con línea de regresión muestra la relación entre dos variables numéricas y ajusta una línea de regresión lineal.

Usaremos el conjunto de datos 'tips' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('tips')

sns.regplot(x='total_bill', y='tip', data=data)
plt.show()
```

12. Gráfico de enjambre (swarmplot)

Un **gráfico de enjambre** muestra la distribución de una variable numérica a lo largo de diferentes categorías, distribuyendo los puntos de datos de manera que no se superpongan.

Usaremos el conjunto de datos 'iris' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('iris')

sns.swarmplot(x='species', y='sepal_length', data=data)
plt.show()
```

13. Gráfico de correlación (heatmap)

Un **gráfico de correlación** muestra la matriz de correlación entre variables numéricas utilizando colores para representar los valores de correlación.

Usaremos el conjunto de datos 'mpg' para este ejemplo:

```
import seaborn as sns
import matplotlib.pyplot as plt

data = sns.load_dataset('mpg')

# Calcular la matriz de correlación
```

```
corr_matrix = data.corr(numeric_only=True)

sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')
plt.show()
```