



Universidad  
Tecnológica  
de Bolívar

# Visualización 2

Ejecutor técnico: Jorge Luis Villalba Acevedo

[www.utb.edu.co/talento-tech](http://www.utb.edu.co/talento-tech)

# Plotting and visualization 2

Tipo de Gráfico (kind)	Descripción	Método en Pandas
<b>line</b>	Gráfico de líneas. Ideal para mostrar tendencias a lo largo del tiempo o cambios en datos continuos.	<code>df.plot.line()</code>
<b>scatter</b>	Gráfico de dispersión. Útil para visualizar la relación entre dos variables numéricas.	<code>df.plot.scatter(x='col1', y='col2')</code>
<b>bar</b>	Gráfico de barras verticales. Perfecto para comparar valores entre diferentes categorías.	<code>df.plot.bar()</code>
<b>barh</b>	Gráfico de barras horizontales. Mejor para categorías con nombres largos.	<code>df.plot.barh()</code>

Tipo de Gráfico (kind)	Descripción	Método en Pandas
<b>hist</b>	Histograma. Utilizado para mostrar la distribución de una variable numérica.	<code>df.plot.hist()</code>
<b>area</b>	Gráfico de área. Muestra la contribución a lo largo del tiempo.	<code>df.plot.area()</code>
<b>box</b>	Gráfico de caja. Ideal para mostrar la distribución de datos y detectar outliers.	<code>df.plot.box()</code>
<b>pie</b>	Gráfico de torta. Representa la proporción de diferentes categorías en un todo.	<code>df.plot.pie(y='col1')</code>

Tipo de Gráfico (kind)	Descripción	Método en Pandas
<b>kde</b>	Gráfico de densidad. Muestra la estimación de la función de densidad de probabilidad de una variable numérica.	<code>df.plot.kde()</code>
<b>hexbin</b>	Gráfico hexagonal. Muestra la densidad de puntos en un espacio bidimensional.	<code>df.plot.hexbin(x='col1', y='col2')</code>

Argumento	Descripción
<b>x</b>	Nombre de la columna que se usará como eje x. Si no se especifica, se usa el índice del DataFrame.
<b>y</b>	Nombre de la columna o lista de columnas que se usarán como eje y.
<b>kind</b>	Tipo de gráfico a crear. Ejemplos: 'line', 'bar', 'barh', 'scatter', 'hist', 'box', 'pie', 'area', 'kde', 'hexbin'.
<b>title</b>	Título del gráfico.

Argumento	Descripción
<b>xlabel</b>	Etiqueta del eje x.
<b>ylabel</b>	Etiqueta del eje y.
<b>color</b>	Color de las líneas o barras. Puede ser un solo color o una lista de colores.
<b>legend</b>	Booleano que indica si se debe mostrar la leyenda.
<b>grid</b>	Booleano que indica si se debe mostrar la cuadrícula.

Argumento	Descripción
<b>alpha</b>	Nivel de transparencia de las líneas o barras (0.0 a 1.0).
<b>figsize</b>	Tamaño de la figura en pulgadas (ancho, alto).
<b>style</b>	Estilo de la línea (por ejemplo, '-', '--', '-.', ':').
<b>xticks</b>	Especifica las posiciones de las etiquetas del eje x.



Argumento	Descripción
<b>yticks</b>	Especifica las posiciones de las etiquetas del eje y.
<b>xlim</b>	Límite del eje x como una tupla (min, max).
<b>ylim</b>	Límite del eje y como una tupla (min, max).
<b>marker</b>	Tipo de marcador para los puntos en gráficos de dispersión (por ejemplo, 'o', '^').

Argumento	Descripción
<b>linewidth</b>	Ancho de la línea.
<b>markersize</b>	Tamaño del marcador en gráficos de dispersión.
<b>fontsize</b>	Tamaño de la fuente para las etiquetas y la leyenda.
<b>rot</b>	Grados de rotación de las etiquetas del eje x.

Argumento	Descripción
<b>secondary_y</b>	Booleano o lista de columnas para usar un segundo eje y.
<b>subplots</b>	Booleano que indica si se deben crear subgráficos (por cada columna) en lugar de un solo gráfico.
<b>sharex</b>	Booleano que indica si los ejes x deben ser compartidos entre subgráficos.
<b>sharey</b>	Booleano que indica si los ejes y deben ser compartidos entre subgráficos.

Tipo de Gráfico (kind)	Descripción
<b>line</b>	Análisis de series temporales, seguimiento de métricas de rendimiento, y visualización de tendencias en datos financieros.
<b>scatter</b>	Estudio de correlaciones, análisis de regresión y exploración de relaciones entre variables en investigación científica.
<b>bar</b>	Informes de ventas, comparaciones de resultados de encuestas y análisis de desempeño entre diferentes grupos o categorías.

Tipo de Gráfico (kind)	Descripción
<b>barh</b>	Análisis de datos con nombres de categoría extensos, como nombres de productos o regiones geográficas.
<b>hist</b>	Análisis de la distribución de datos, identificación de la forma de la distribución y análisis de frecuencias.
<b>area</b>	Visualización de la evolución de componentes de un total a lo largo del tiempo, como la participación de mercado.

Tipo de Gráfico (kind)	Descripción
<b>box</b>	Análisis de la variabilidad de datos, comparación de distribuciones entre diferentes grupos y detección de valores atípicos.
<b>pie</b>	Comparación de partes de un total en presentaciones, visualización de la composición de un conjunto de datos.
<b>kde</b>	Visualización de la distribución de datos continuos, suavizando el histograma para facilitar la interpretación.
<b>hexbin</b>	Análisis de grandes conjuntos de datos, exploración de patrones en relaciones complejas y visualización de densidades espaciales.

# Resumen de Usos Prácticos

- **Gráficos de líneas y dispersión:** Comunes en análisis de series temporales y estudios de correlación.
- **Gráficos de barras y tortas:** Frecuentemente utilizados en informes y presentaciones para resumir datos categóricos.
- **Histogramas y gráficos de caja:** Herramientas esenciales en la estadística descriptiva para analizar la distribución de datos.

# Resumen de Usos Prácticos

- **Gráficos de área:** Útiles para visualizar contribuciones a lo largo del tiempo.
- **Gráficos de densidad y hexagonales:** Ideales para explorar la distribución y densidad de grandes conjuntos de datos.



# Activar librerías

```
import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt
```

# Código para Generar la Base de Datos

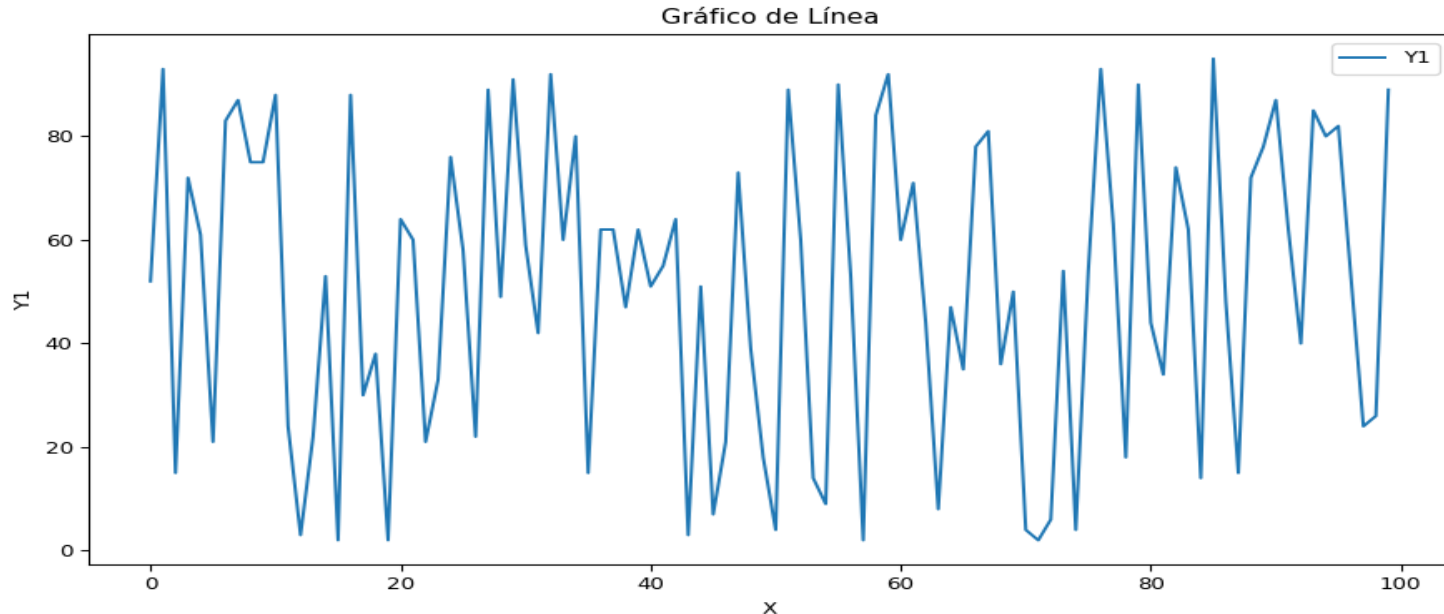
```
# Semilla
np.random.seed(42)

# Generar una base de datos aleatoria
data_size = 100
data = {
    'X': np.arange(data_size),
    'Y1': np.random.randint(1, 100, size=data_size),
    'Y2': np.random.normal(50, 15, size=data_size),
    'Category': np.random.choice(['A', 'B', 'C', 'D'],
size=data_size)
}

df = pd.DataFrame(data)
```

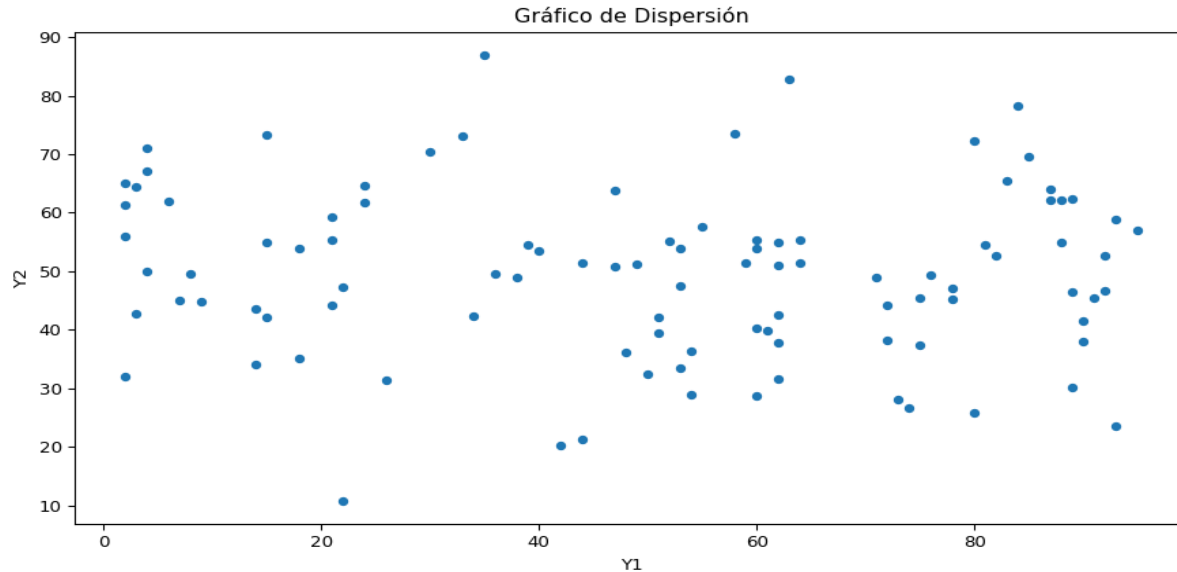
# Código para Generar los Gráficos

1. ¿Cómo se muestra la relación entre  $X$  y  $Y1$  en un gráfico de línea, y para qué es útil observarlo?



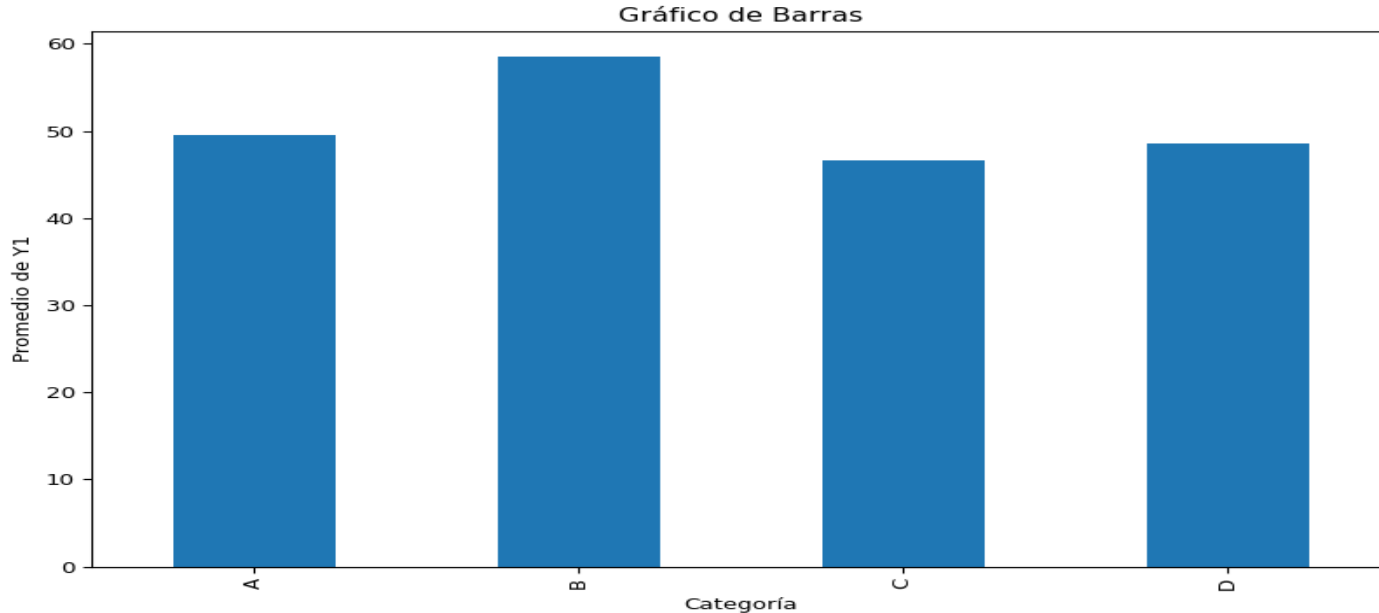
# Código para Generar los Gráficos

2. ¿Qué relación se visualiza en un gráfico de dispersión entre  $Y1$  y  $Y2$ , y por qué es ideal para ver correlaciones?



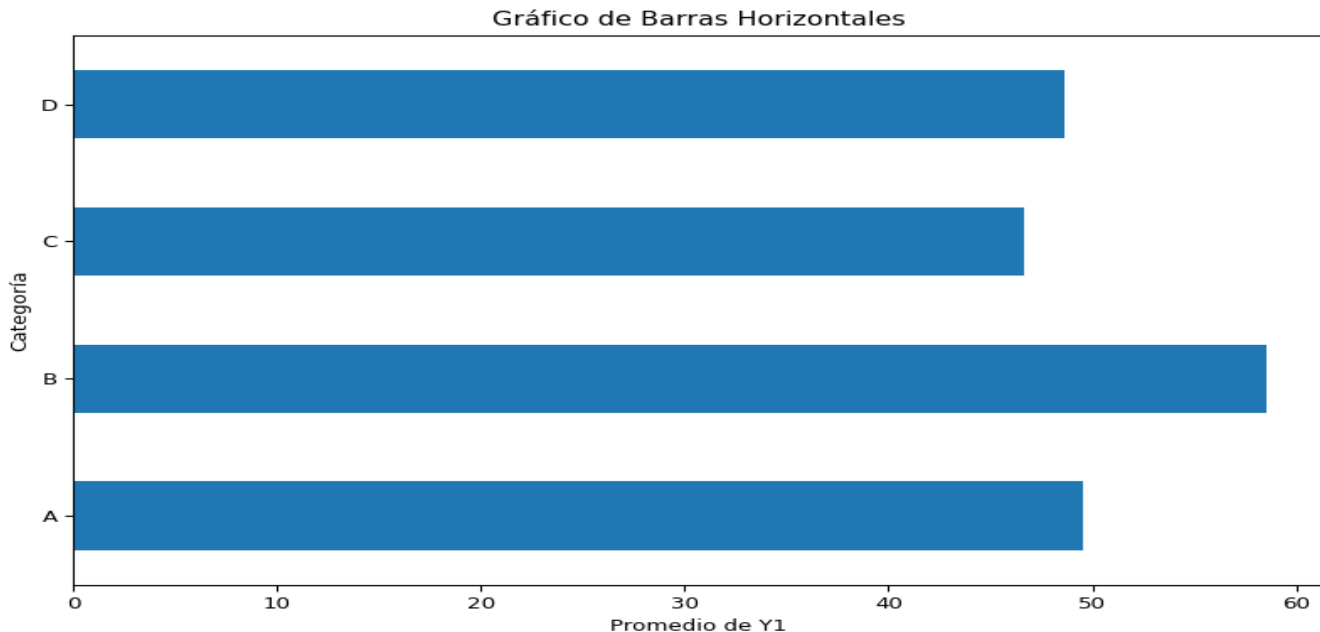
# Código para Generar los Gráficos

3. ¿Cómo se compara el promedio de Y1 entre diferentes categorías en un gráfico de barras?



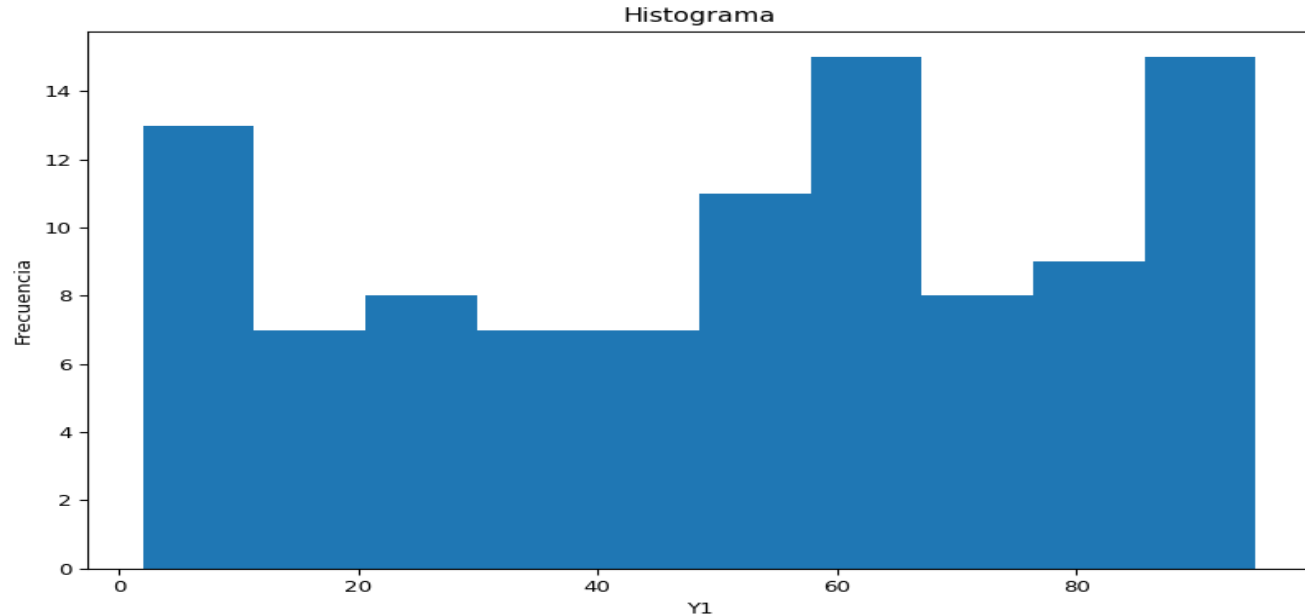
# Código para Generar los Gráficos

4. ¿En qué se diferencia un gráfico de barras horizontales de un gráfico de barras verticales?



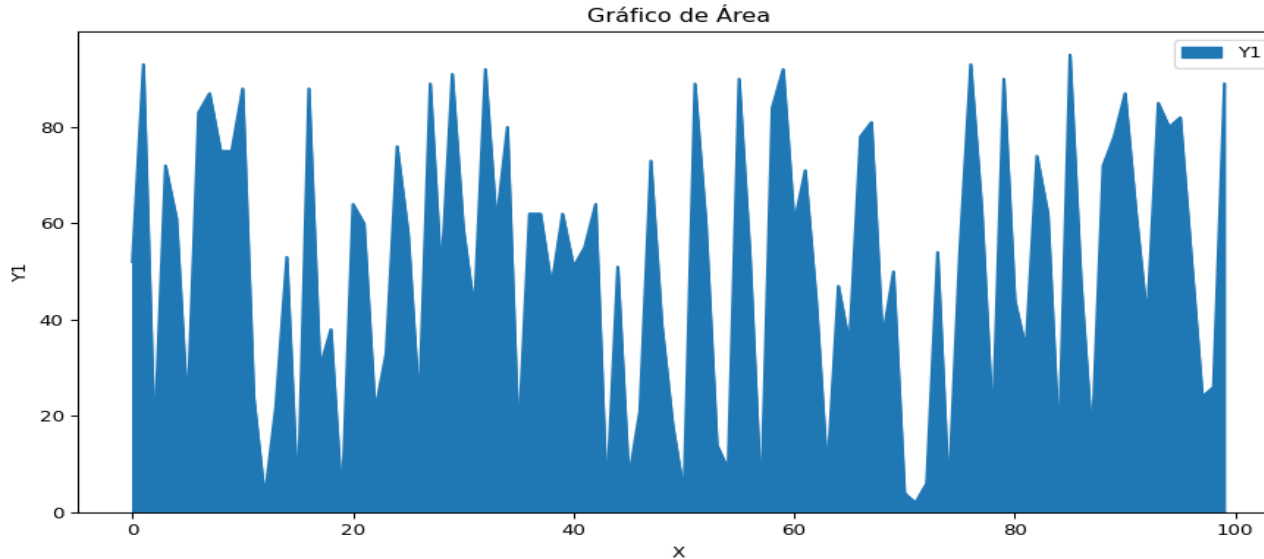
# Código para Generar los Gráficos

5. ¿Qué información proporciona un histograma sobre la distribución de los valores de Y1?



# Código para Generar los Gráficos

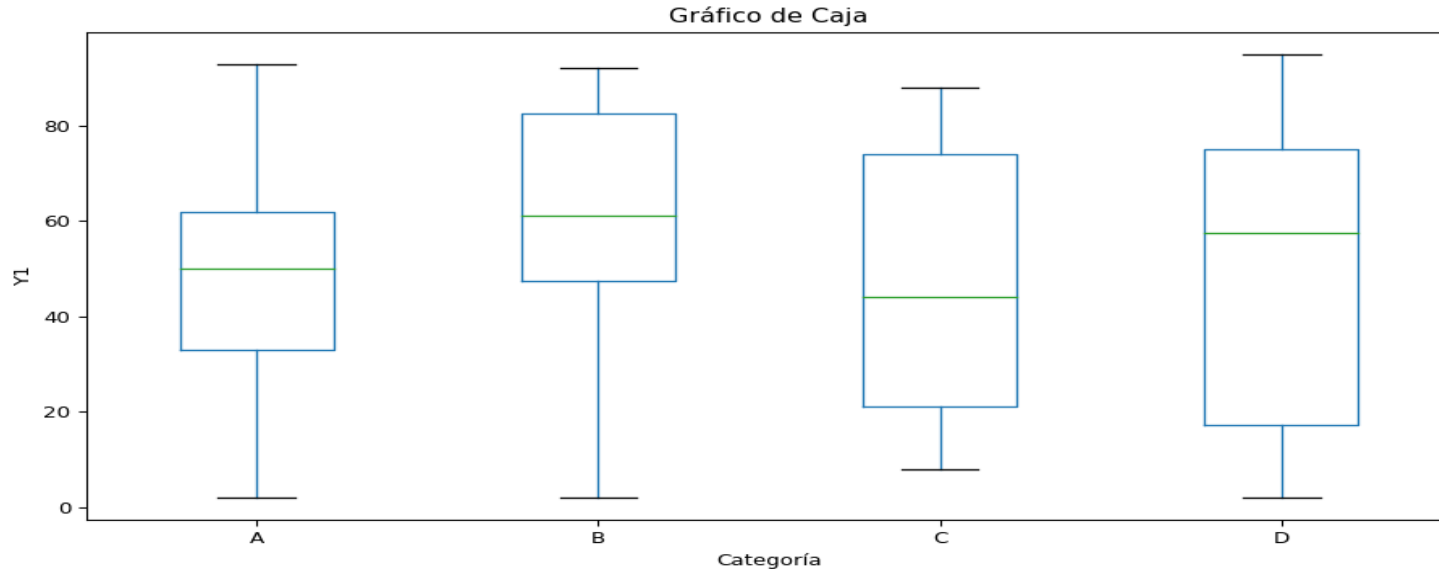
6. ¿Cómo representa un gráfico de área el área bajo la curva de Y1 a lo largo de X?





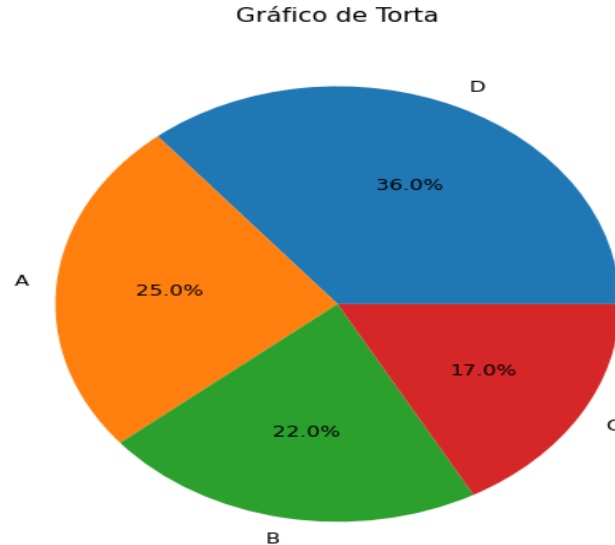
# Código para Generar los Gráficos

7. ¿Qué información muestra un gráfico de caja sobre la distribución de  $Y1$  por cada categoría, incluyendo mediana y cuartiles?



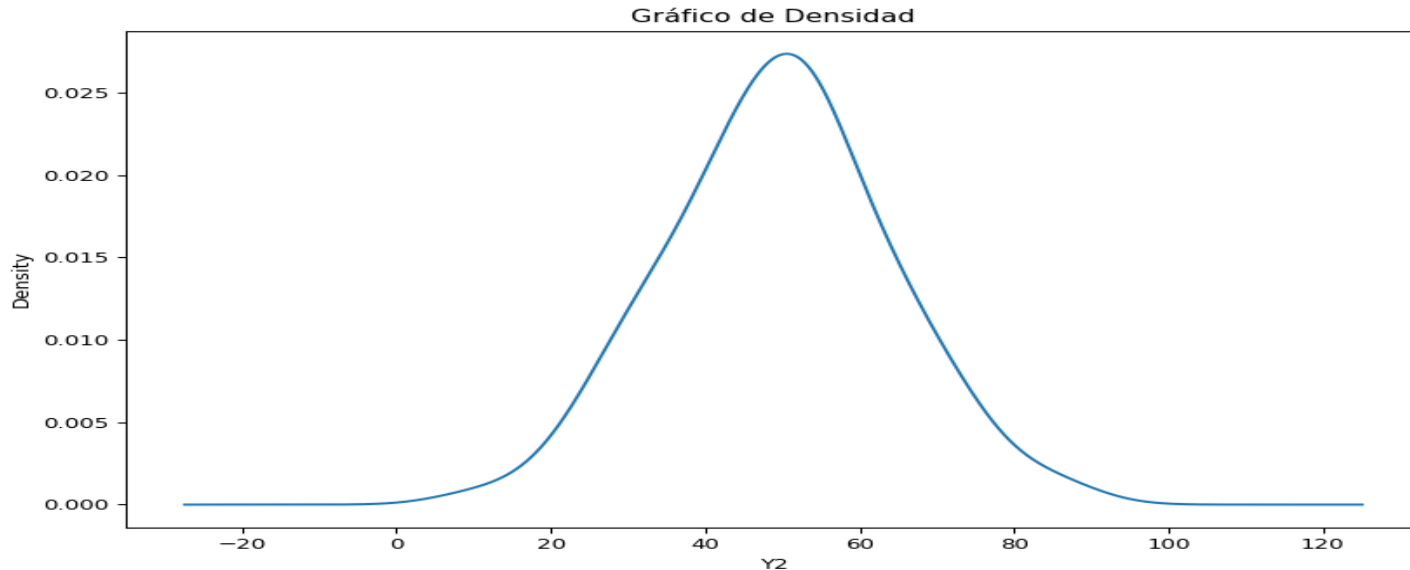
# Código para Generar los Gráficos

8. ¿Cómo visualiza un gráfico de torta la proporción de cada categoría en el total?



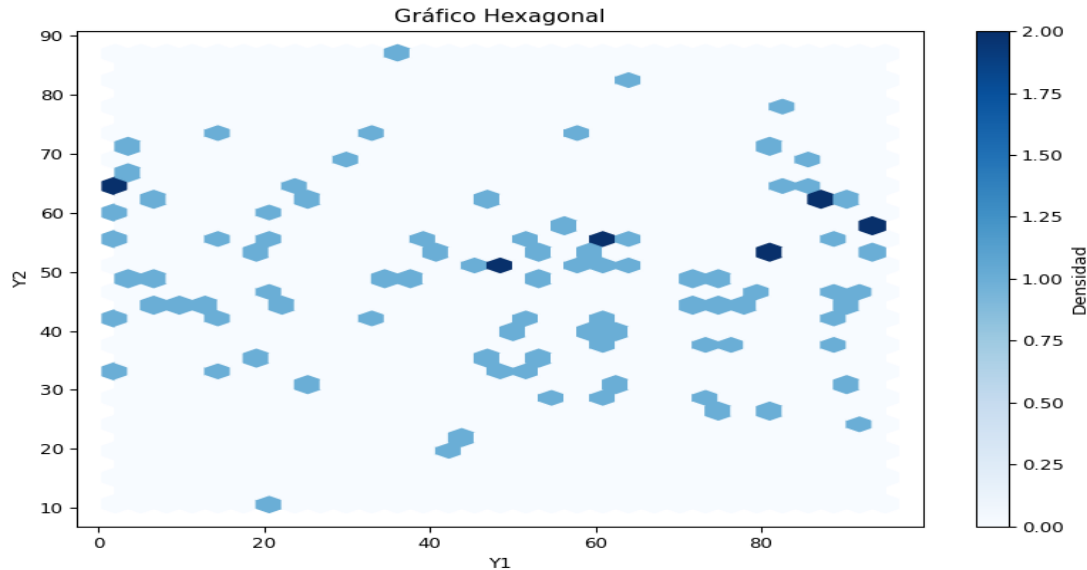
# Código para Generar los Gráficos

9. ¿Qué se estima en un gráfico de densidad respecto a la probabilidad de  $Y_2$ ?



# Código para Generar los Gráficos

10. ¿Qué información proporciona un gráfico hexagonal sobre la densidad de puntos en un espacio bidimensional, y por qué es útil para grandes conjuntos de datos?



# Referencias

1. McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython* (2nd ed.). O'Reilly Media.
2. Sweigart, A. (2019). *Automate the Boring Stuff with Python: Practical Programming for Total Beginners* (2nd ed.). Publicado de forma independiente.
3. González Duque, R. (2015). *Python para Todos*. Publicado de forma independiente.
4. Bosch, J. (2020). *Introducción a la Programación en Python: Algoritmos y Lógica para Principiantes*. Publicado de forma independiente.