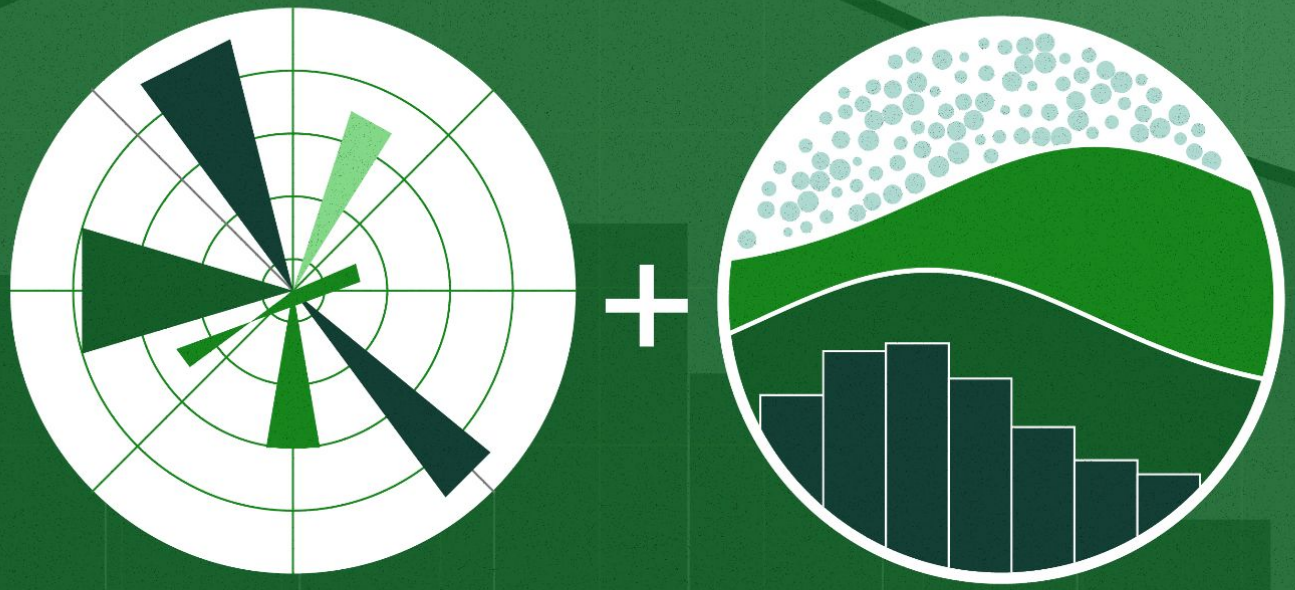
Curso Básico de

Visualizaciones de Datos con Matplotlib y Seaborn

Carlos Alarcón



La importancia de la visualización de datos

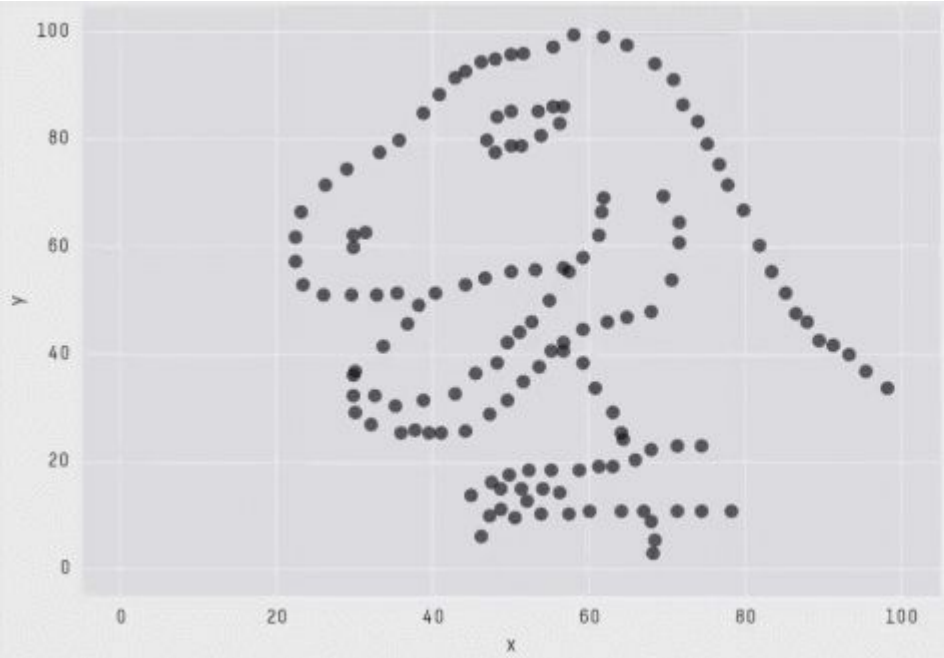
“

Visualization gives you answers to questions you didn’t know you had.

”

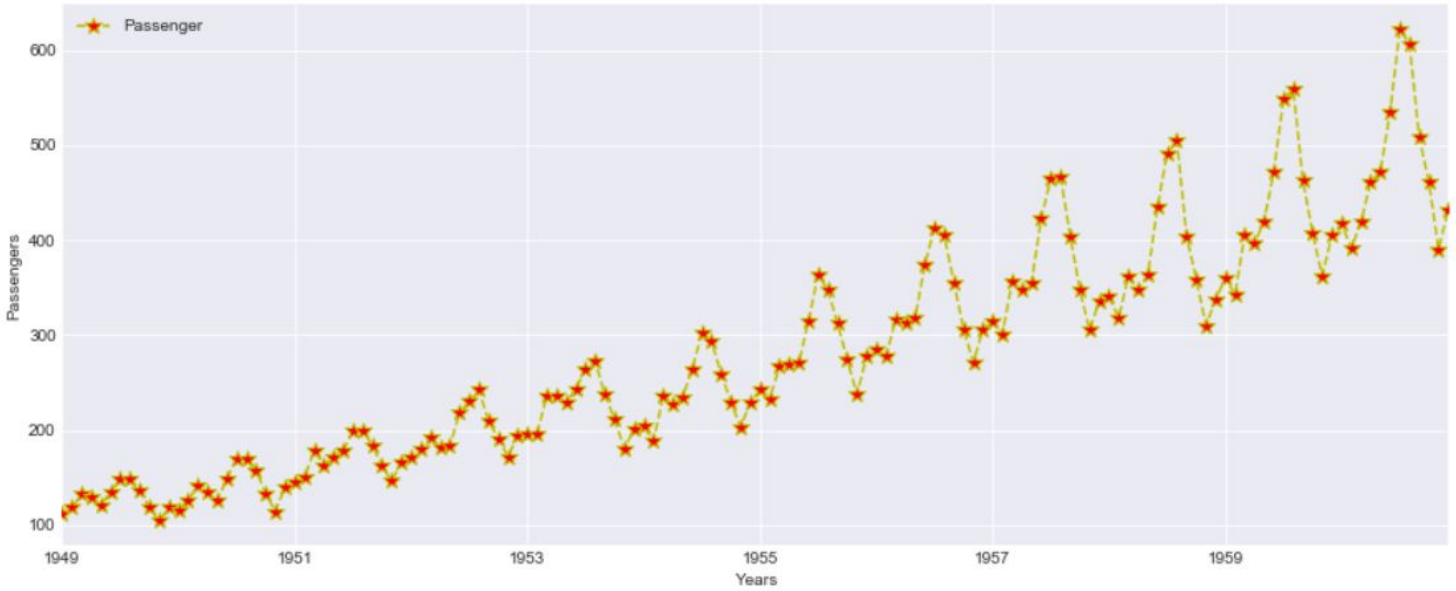
Ben Schneiderman

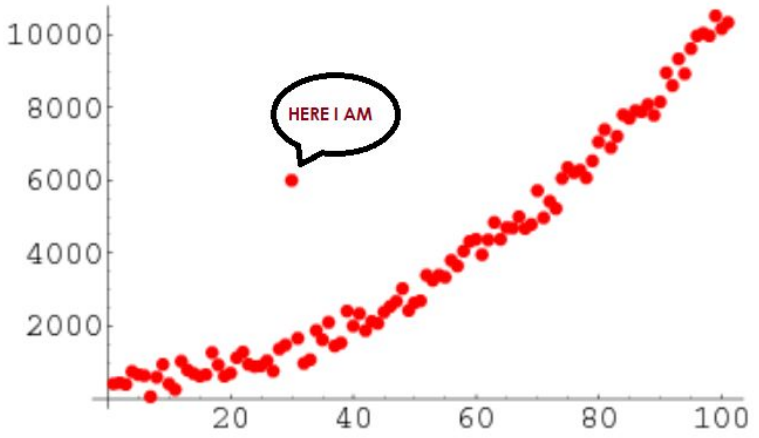
Mayor contexto

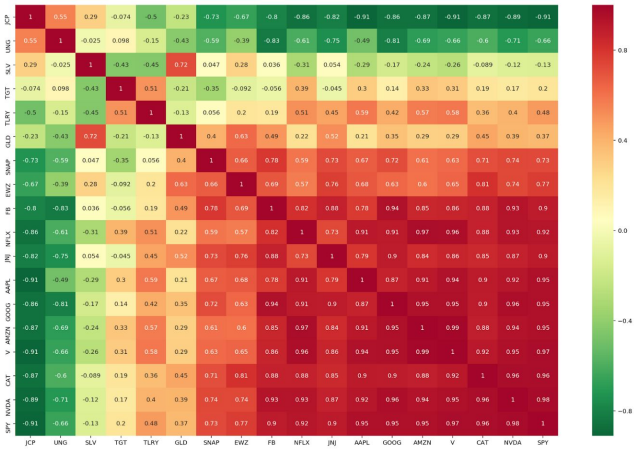


X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526

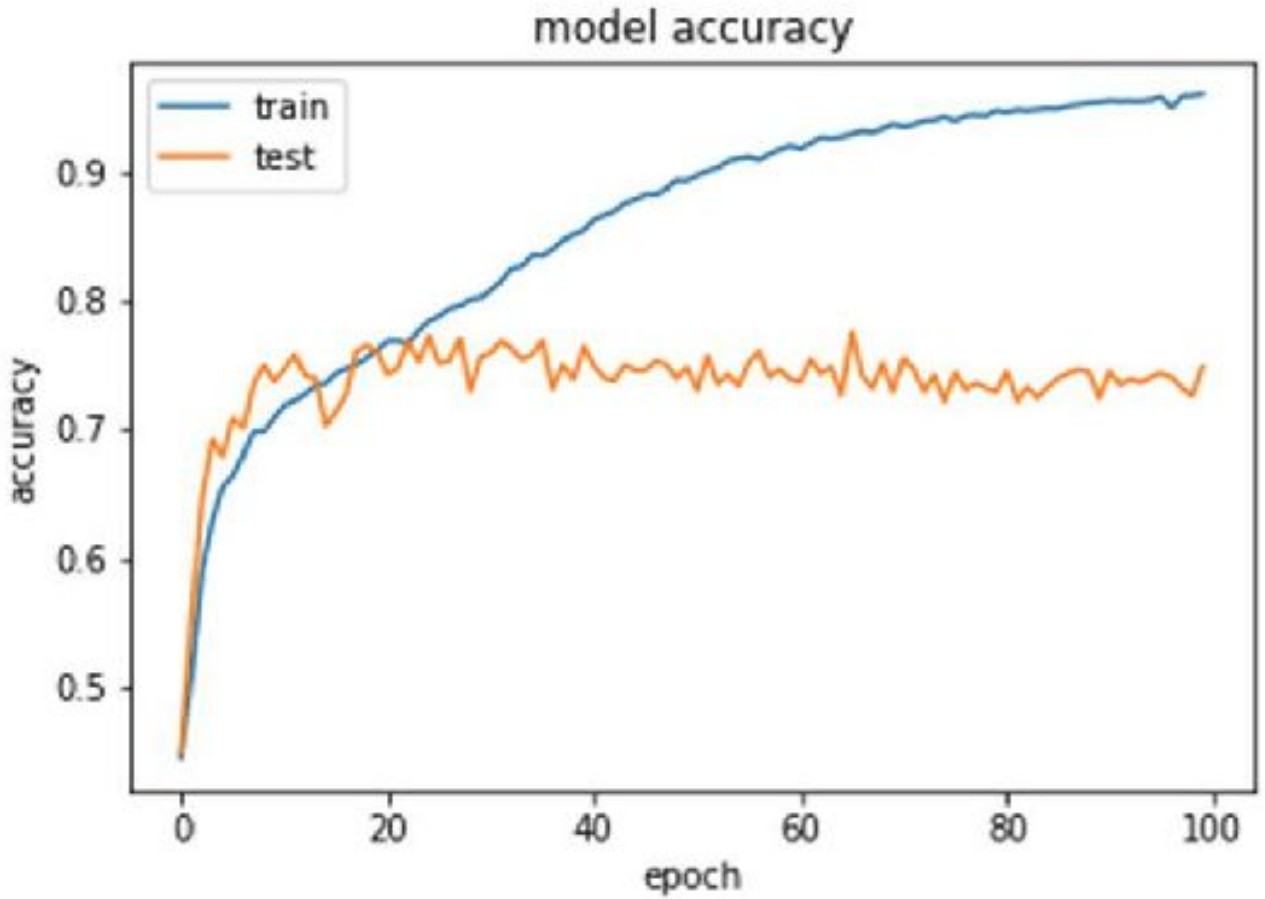
Hallazgos en nuestros datos







Mayor claridad en nuestro código



Matplotlib

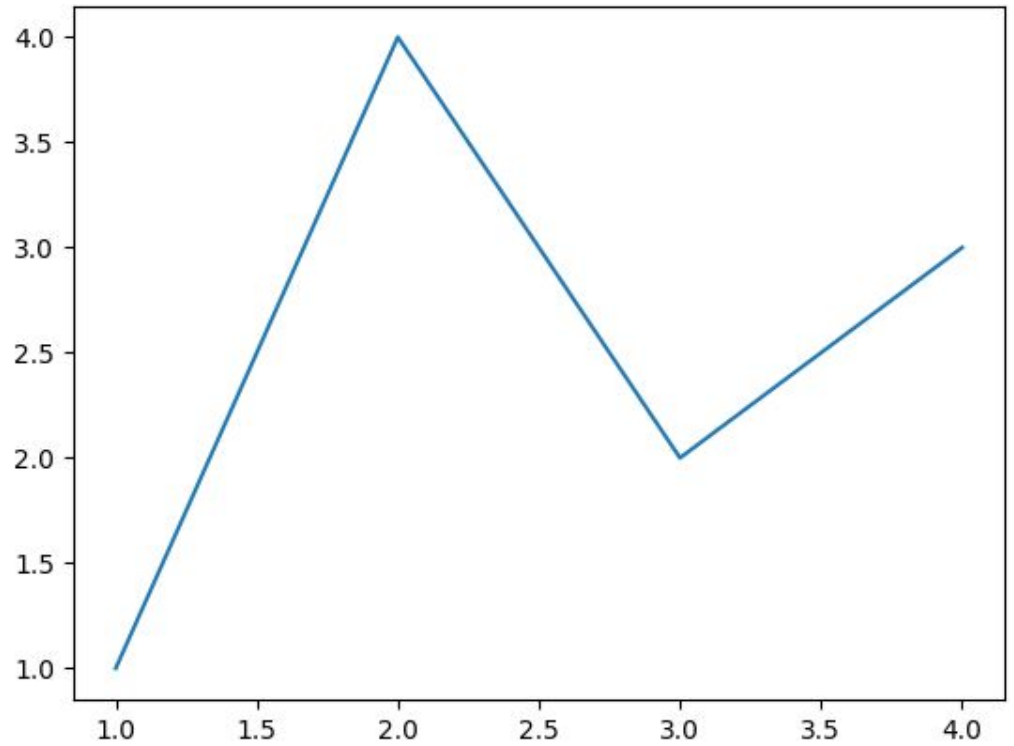
- Escrita por John D. Hunter.
- Creada en 2003.

Matplotlib

- Emula comandos de MATLAB.
- Usa NumPy.
- Escrita en Python.

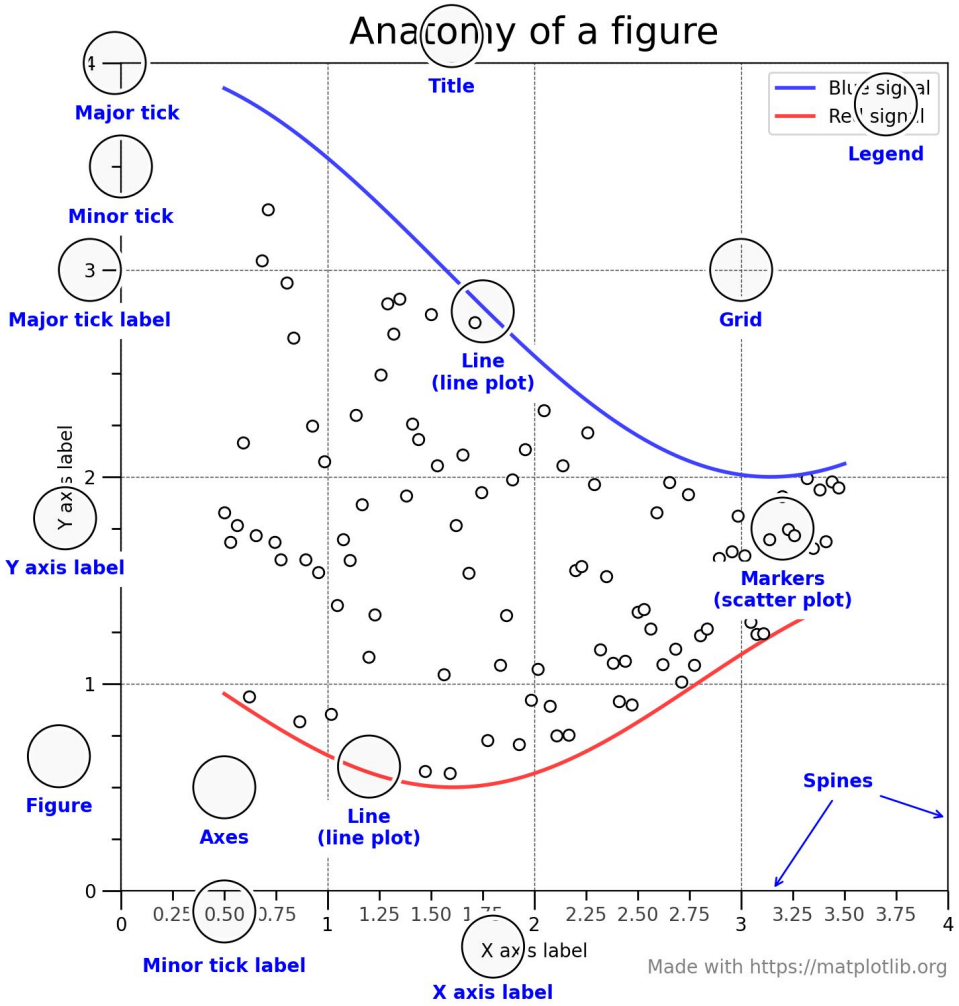
Simple

```
fig, ax = plt.subplots()  # Create a figure containing a single axes.
ax.plot([1, 2, 3, 4], [1, 4, 2, 3]);  # Plot some data on the axes.
```



Rápida

Personalizable



Seaborn

Conocimientos previos

- Tipos de visualizaciones de datos.
- Python.
- Pandas y Numpy.

Pyplot básico

Subplot

Método orientado a objetos

¿Para qué sirven?

Pyplot

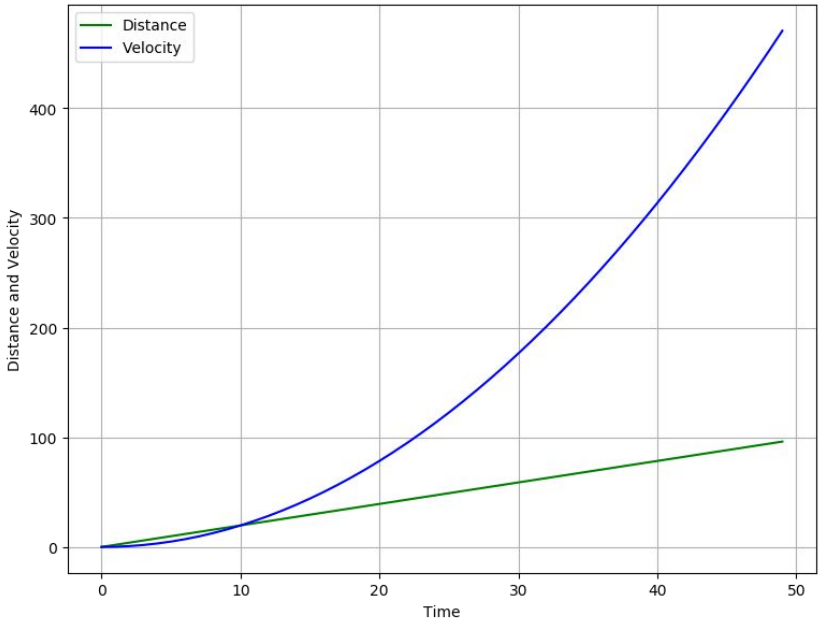
- Rápido
- Fácil
- Una sola figura

Object Oriented

- Mayor personalización.
- Más amigable a múltiples diagramas.
- Más código.

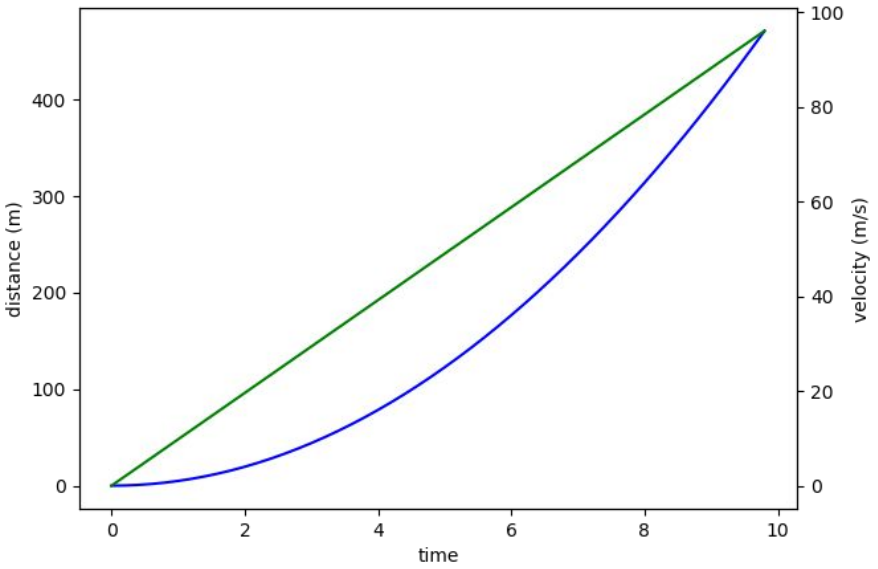
Pyplot

```
plt.figure(figsize=(9,7),
dpi=100)
plt.plot(time, velocity,'g-')
plt.plot(time, distance,'b-')
plt.ylabel("Distance and
Velocity")
plt.xlabel("Time")
plt.grid(True)
```

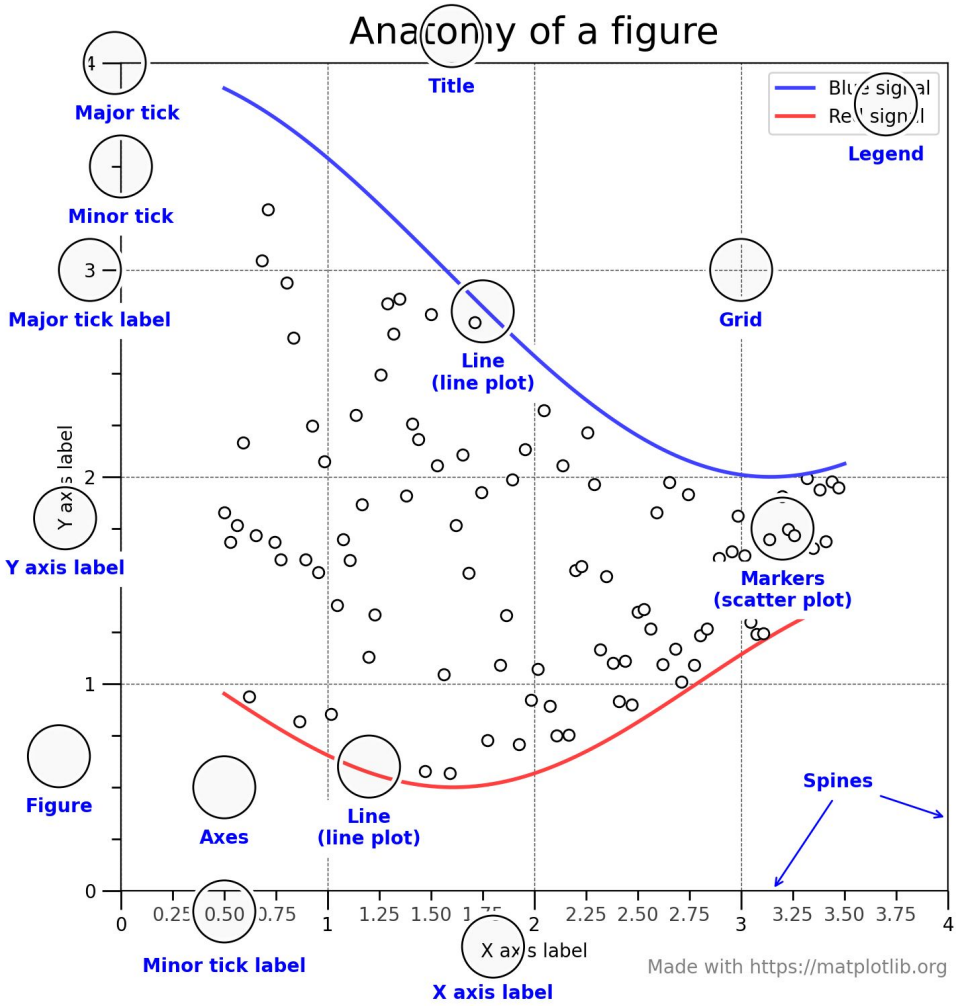


Object oriented

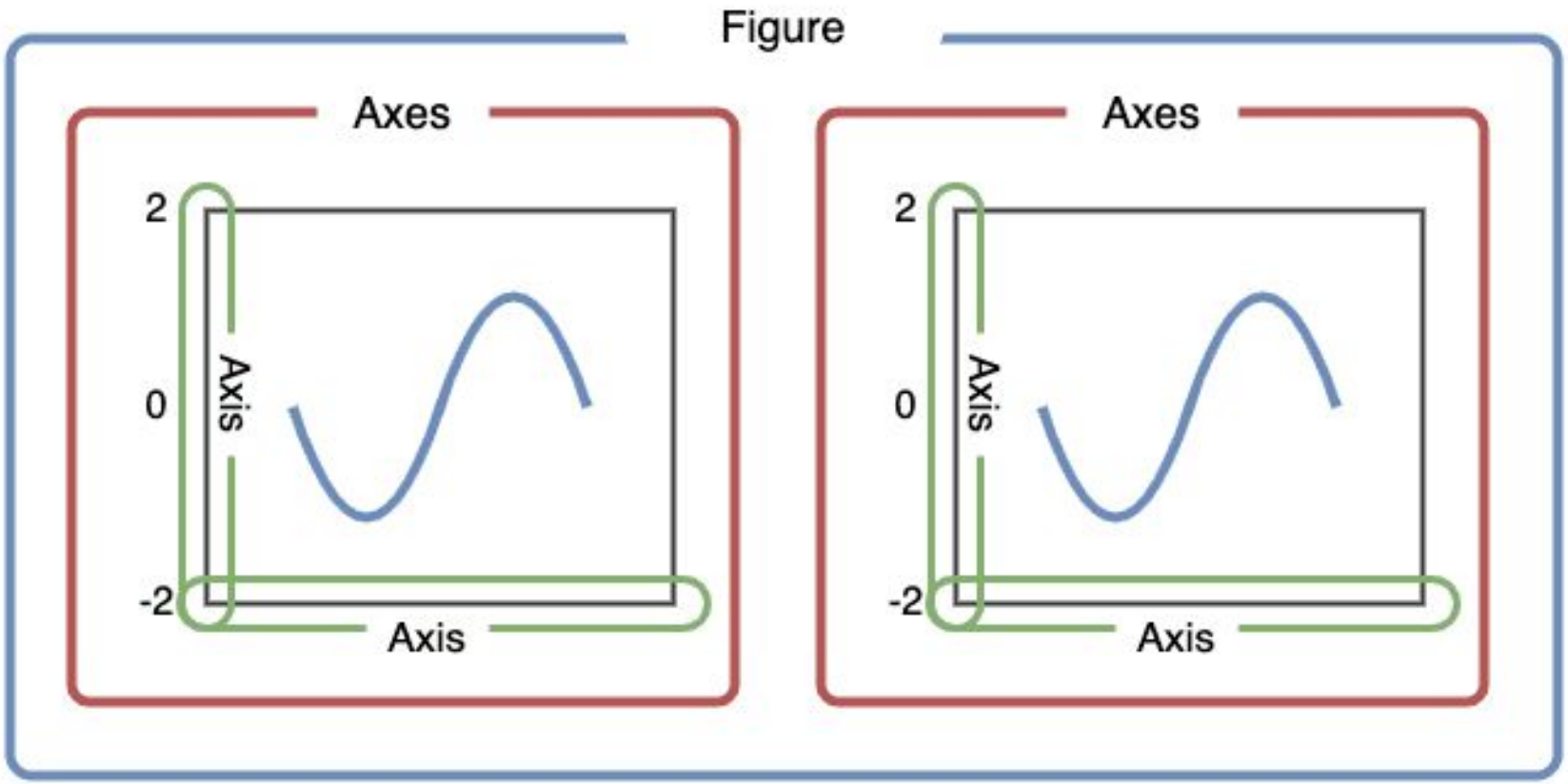
```
fig, ax1 = plt.subplots()
ax1.set_ylabel("distance (m)")
ax1.set_xlabel("time")
ax1.plot(time, distance, "blue")
ax2 = ax1.twinx()
ax2.set_ylabel("velocity (m/s)")
ax2.set_xlabel("time")
ax2.plot(time, velocity, "green")
fig.set_size_inches(7,5)
fig.set_dpi(100)
```



Personalizable



Estructura



Subplots

Leyendas, etiquetas, títulos, tamaño

Colores y estilos

Bar Plot

Otras gráficas

Seaborn

Seaborn

- Escrita por Michael Waskom.
- Construida sobre Matplotlib.
- Integrada para estructuras de Pandas.

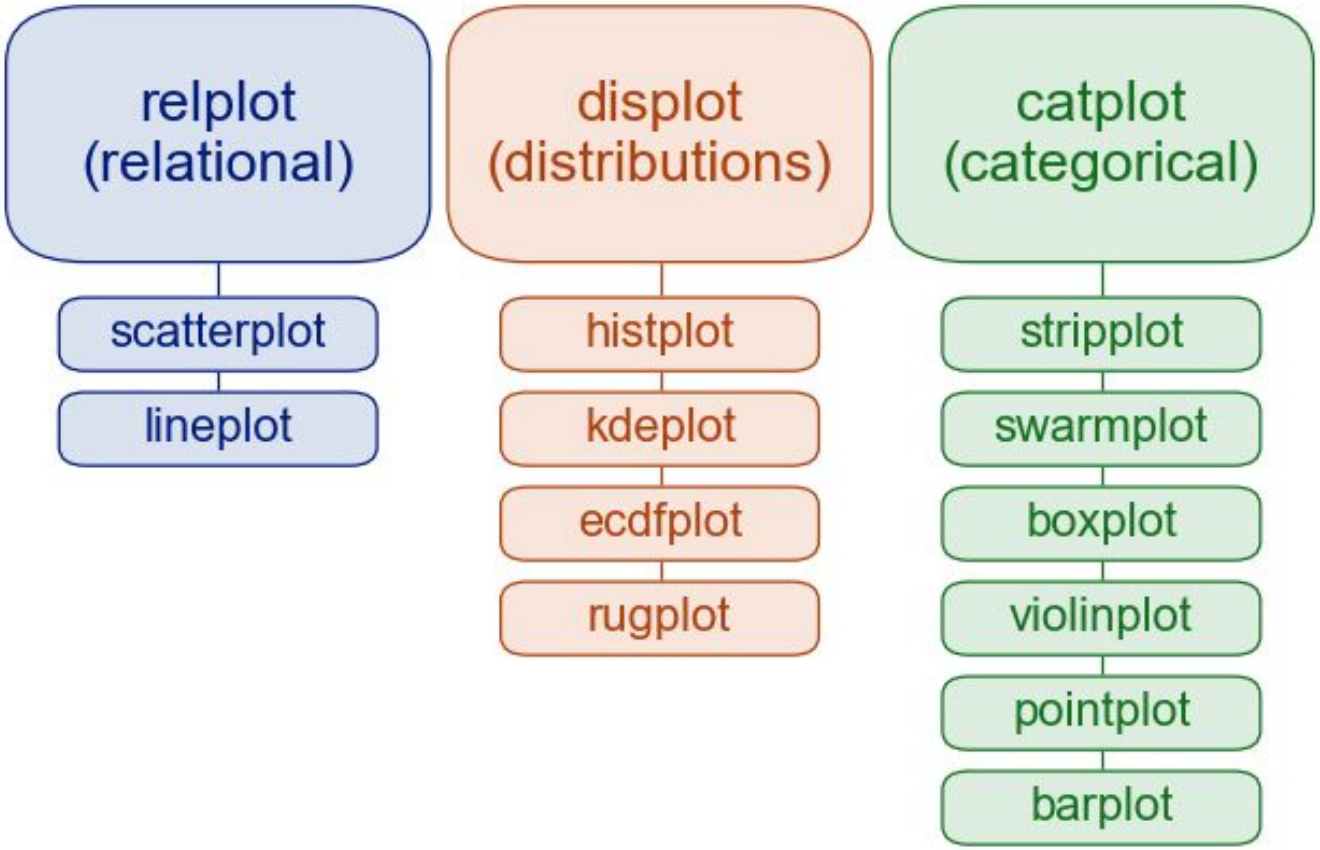
¿Por qué Seaborn?

- Velocidad
- Poco código
- Customizable

```
##Estructura básica

sns.”Tipo de gráfica”(
    data=”Dataset”,
    x=”Data en eje x”,
    y=”Data en eje y”,
    hue=”Variable de agrupamiento”)
```

Tipos de gráficas



Set

Parámetros más usados

Distribuciones

Categóricos

Relation

Jointplot / Pairplot

Heatmap

Heatmap

Posibilidades con Matplotlib y Seaborn

Matplotlib

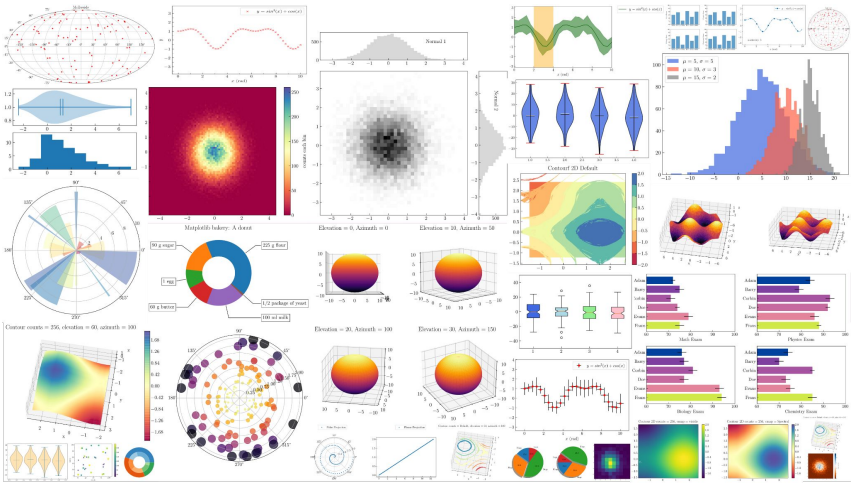
- Pyplot
- Orientado a objetos
- Subplots
- Tittle, legends, labels
- Colores y estilos
- Otros gráficos
- Tamaños

Seaborn

- Estructura
- Parámetros más usados
- Distributions
- Relational
- Categorical
- Jointplot - Pairplot
- Heatmap

Posibilidades infinitas

kaggle →



Carlos Andrés Alarcón

@Alarcon7a