

# Python для Data Science

**Заняття 4.**

Візуалізація даних

# План заняття

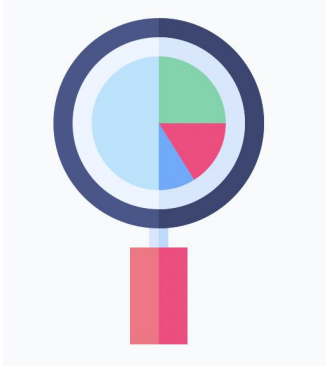
---



- Візуалізація даних за допомогою Matplotlib:
  - Базові типи візуалізації (line chart, scatter plot, bar plot, histogram)
  - Декілька графіків на одній візуалізації
- Візуалізація даних в Pandas

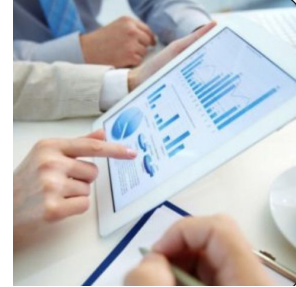
# Data visualization

---



Extremely helpful for EDA (exploratory data analysis)

Reporting



Presentation

# Matplotlib

---

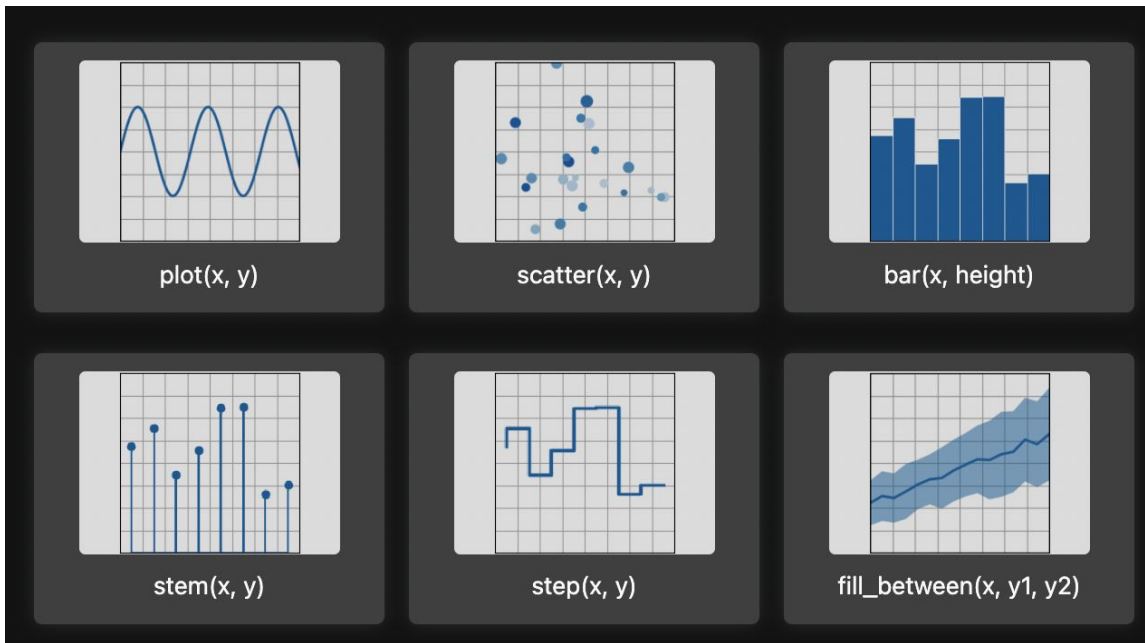
**Matplotlib** — це комплексна бібліотека для створення статичних, анімованих та інтерактивних візуалізацій на Python. Це чудова бібліотека візуалізації в Python для двовимірних графіків масивів. Matplotlib — це багатоплатформна бібліотека візуалізації даних, створена на основі масивів **NumPy** і призначена для роботи з ширшим стеком **SciPy**.

The Matplotlib logo is a circular icon containing several colored segments (orange, yellow, green, blue) radiating from a central point, resembling a stylized sun or a data visualization.

matplotlib

# Matplotlib Pyplot Interface

**matplotlib.pyplot** — це набір функцій, завдяки яким matplotlib працює як MATLAB. Кожна функція pyplot вносить певні зміни у фігуру: наприклад, створює фігуру, створює область побудови на фігурі, малює деякі лінії в області побудови, прикрашає графік мітками тощо.



# Matplotlib Object-Oriented Interface

---

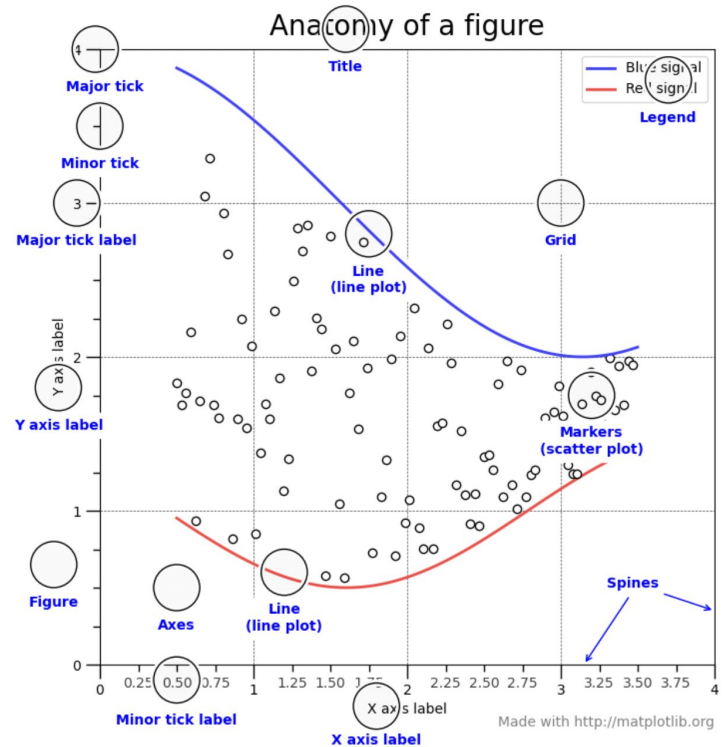
Під час використання **OO interface** корисно знати, як matplotlib структурує свої графіки. Остаточним графіком, який ми бачимо як результат, є об'єкт «Фігура» (“Figure”). Об'єкт Figure є контейнером верхнього рівня для всіх інших елементів, які складають графічне зображення. Ці «інші» елементи називаються художниками (artists). Об'єкт «Фігура» можна розглядати як полотно, на якому діють різні художники, створюючи кінцеве графічне зображення. Ця фігура може містити будь-яку кількість різних художників.

Отже, зовнішній контейнер у графіку/діаграмі Matplotlib називається **figure** (фігурою). Кожна фігура може містити **одну або кілька Axes, які є реальними графіками (plots)**. Кожна з осей матиме додаткові підкомпоненти, такі як **axis (x & y), title, legend, axis labels, major & minor ticks** тощо.

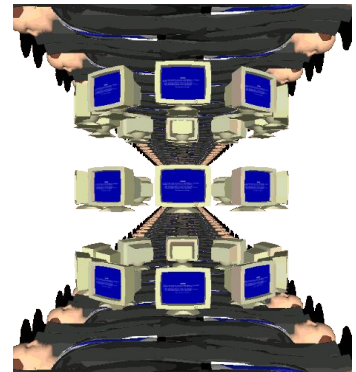
# Matplotlib: anatomy of a figure

Варто звернути увагу на:

- Усі пункти, позначені синім кольором, є **Artists**. Artists — це в основному всі елементи, які відображаються на фігурі. Може включати текст, патчі (наприклад, стрілки та форми) тощо. Таким чином, усі наступні Figure, Axes та Axis об'єкти також є Artists. Кожен графік, який ми бачимо на малюнку, є об'єктом **Axes**. Об'єкт Axes містить актуальні дані, які ми збираємося відобразити. Він також міститиме X- та Y-axis labels, title. Кожен об'єкт Axes міститиме два або більше об'єктів Axis.
- Об'єкти **Axis** встановлюють обмеження даних. Також містить ticks та ticks labels. ticks – це позначки, які ми бачимо на axis.



# Live coding / Практика





# Для самостійного вивчення:

---



- [Seaborn](#) – бібліотека візуалізації. Це оболонка matplotlib із більш зручним інтерфейсом у деяких випадках і приємнішим стандартним стилем (але ми розглянемо це під час наступної лекції)
- [Plotly](#) – графічна бібліотека, яка створює інтерактивні графіки високої якості

**Питання**

???

???