Building Data Visualizations Using Matplotlib

WORKING WITH THE MATPLOTLIB AND PYPLOT APIS

Janani Ravi
CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

Seamless integration with other packages in the PyData stack

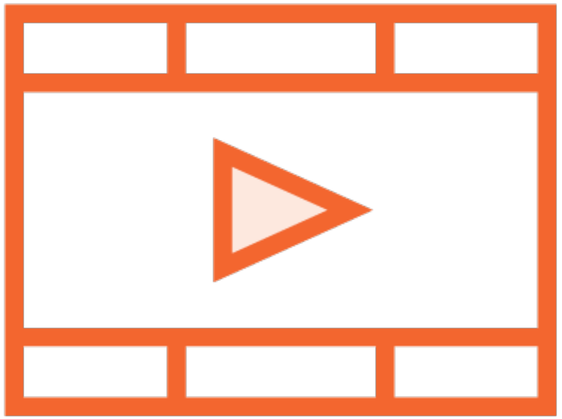
Anatomy of a Matplotlib figure

Basic plots, labels, titles, markers and watermarks

Figures, axes, subplots

Prerequisites and Course Outline

Prerequisite Courses



Basic understanding of Python programming (Python 3)

Working with Jupyter notebooks

Basic understanding of NumPy arrays

Software and Skills

Basic understanding of Python programming (Python 3)

Working with Jupyter notebooks

Basic understanding of NumPy arrays

Course Outline

Introducing Matplotlib and Pyplot

- Anatomy of a figure
- Plots, titles, labels, lines, markers, watermarks

Basic, intermediate and advanced plots

- Shapes and curves
- Text and annotations
- Different units on the same axis
- Scaled axis

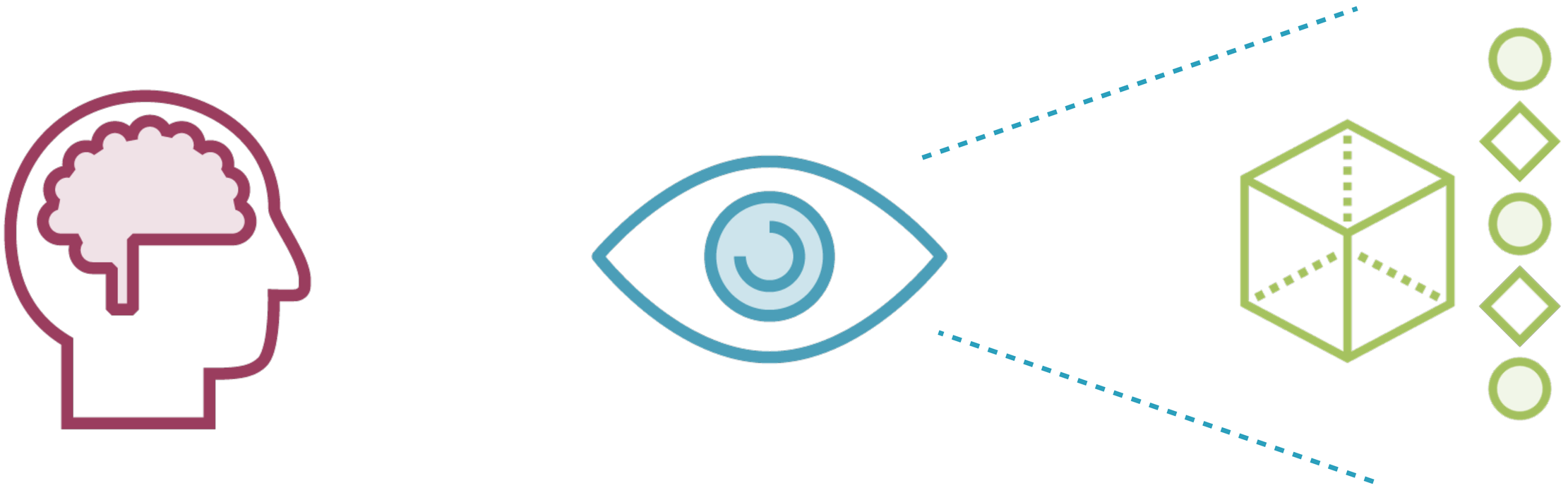
Visualizing data

- Boxplots and violin plots
- Histograms and pie charts
- Stem charts
- Autocorrelations
- Stackplots

Visualization

Visualizations are the easiest way to absorb information

Visualization



Visualizations are the easiest way to absorb information

Visualization in Data Science

Important step in data exploration

Helps developing an intuition for relationships in data

Precursor to higher level data analysis using ML techniques

Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms.

matplotlib.org

Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms.

matplotlib.org

Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms.

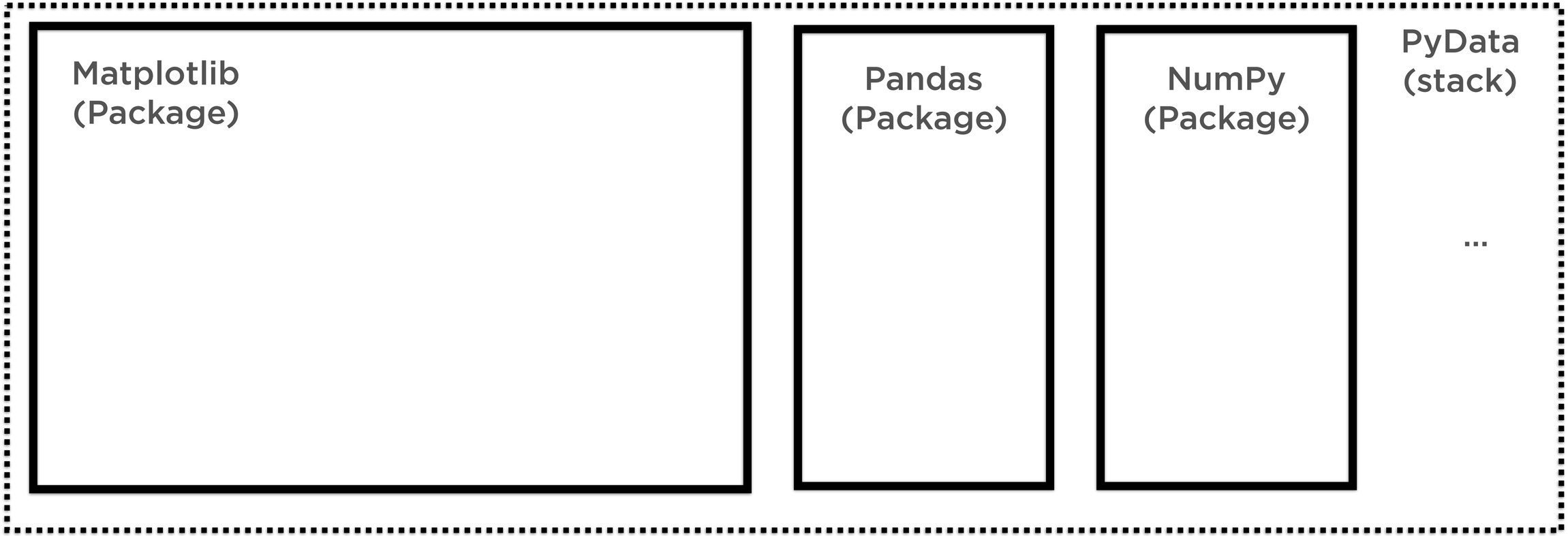
matplotlib.org

Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms.

matplotlib.org

Matplotlib Ecosystem

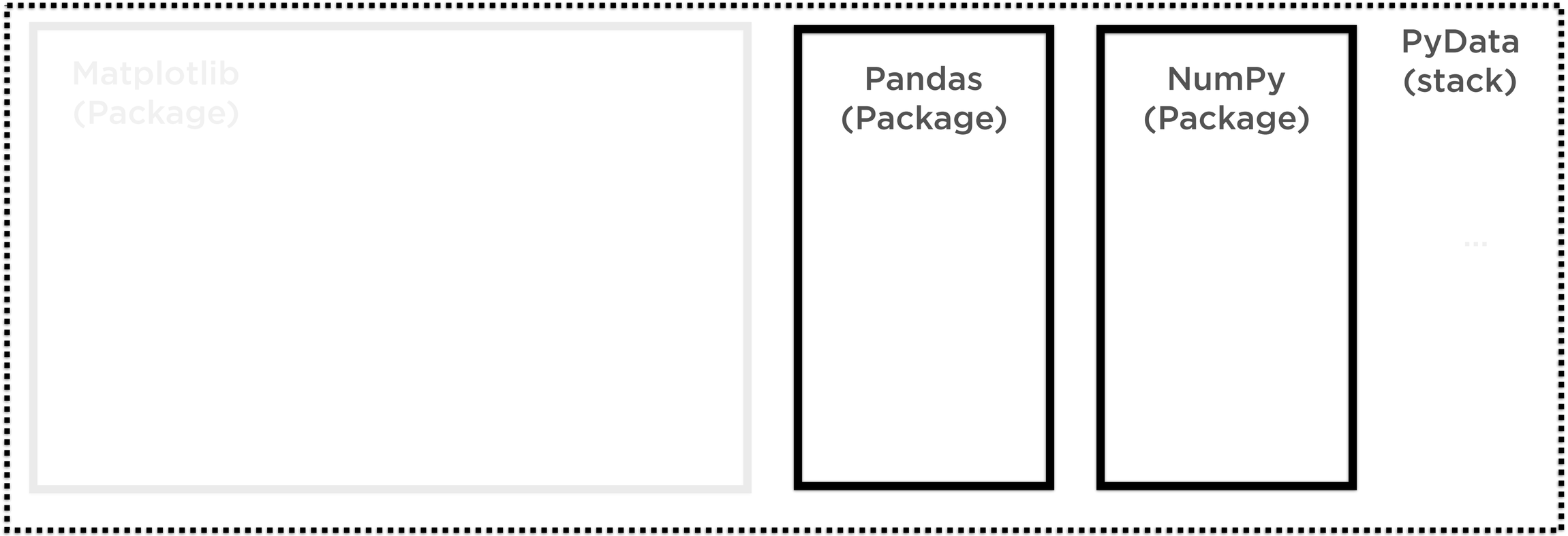


Matplotlib Ecosystem

Tightly integrates
with PyData stack

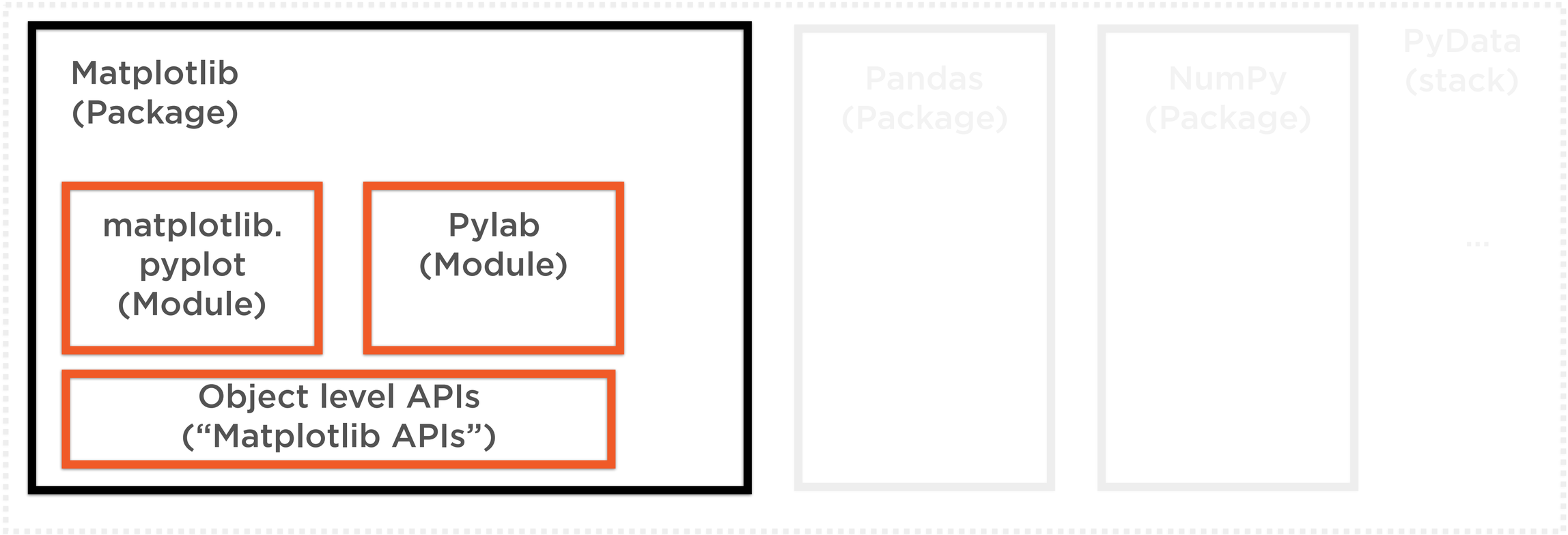
Matplotlib Ecosystem

Inter-operates with
Pandas, NumPy...



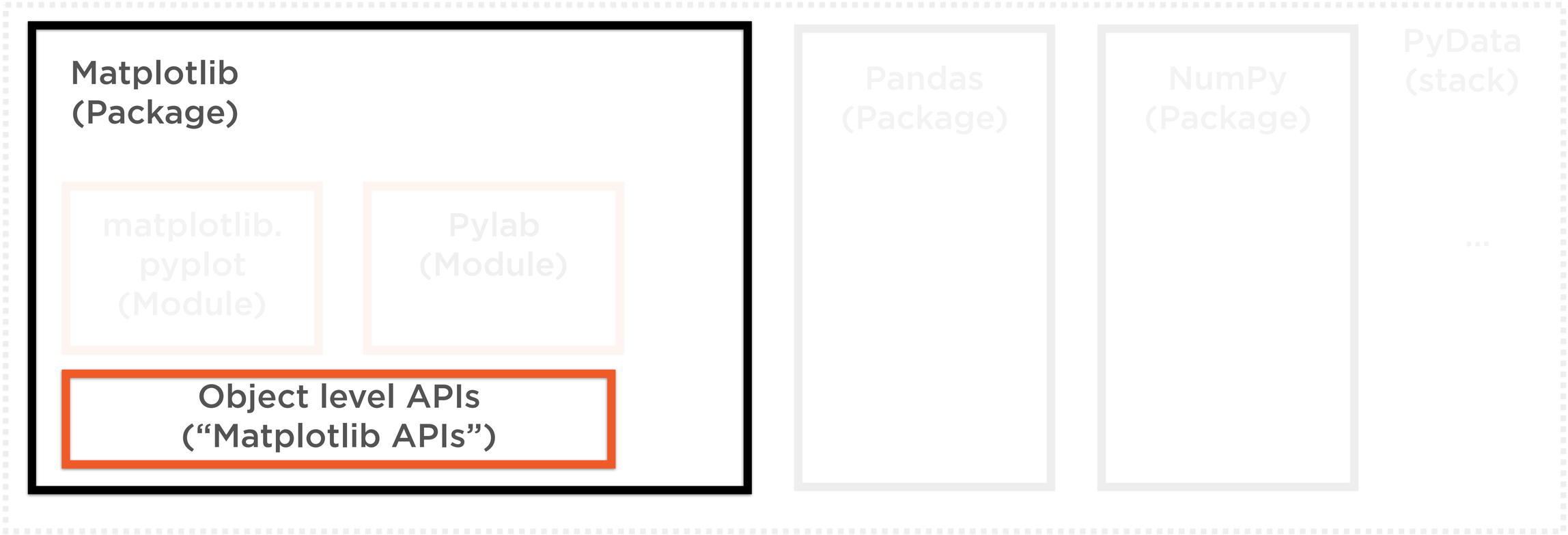
Matplotlib Ecosystem

Matplotlib is a complex package that includes multiple modules



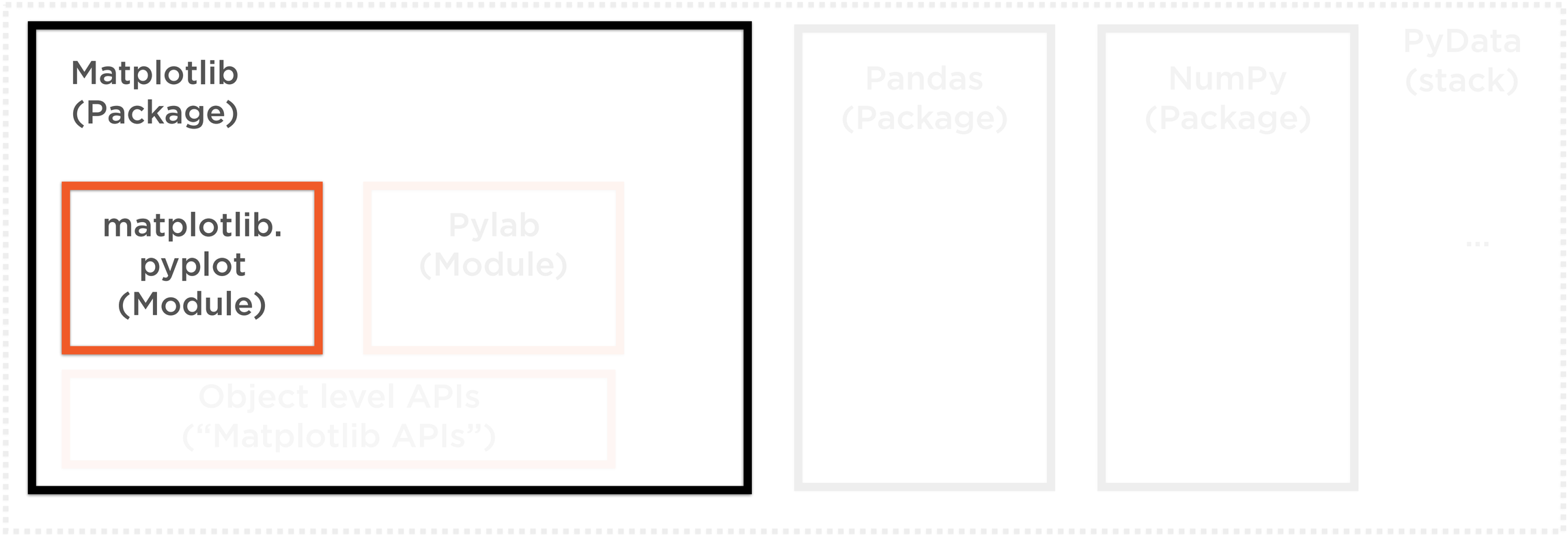
Matplotlib Ecosystem

Includes granular low-level APIs to control each object in a plot



Matplotlib Ecosystem

Also includes a higher level API that controls the “state-machine”



Matplotlib Ecosystem

Pylab is a convenience module imports portions of Matplotlib and NumPy to give users a Matlab-like access to functions

History

Matlab

Excels at making nice looking plots easy

But:

- Not free
- Difficult to scale
- Limited as a programming language

History

Emulates MATLAB's plotting capabilities

Matplotlib

Python plotting library

Easy to create plots

Embeddable GUI for application development

Can be used across platforms

Three Top-level Components

Pylab

Collection of command style functions that make matplotlib work like MATLAB

Frontend

Set of classes that abstract actual plotting mechanics and details using a high-level user API

Backend

Renderers, that transform the frontend representation to hardcopy or a display device

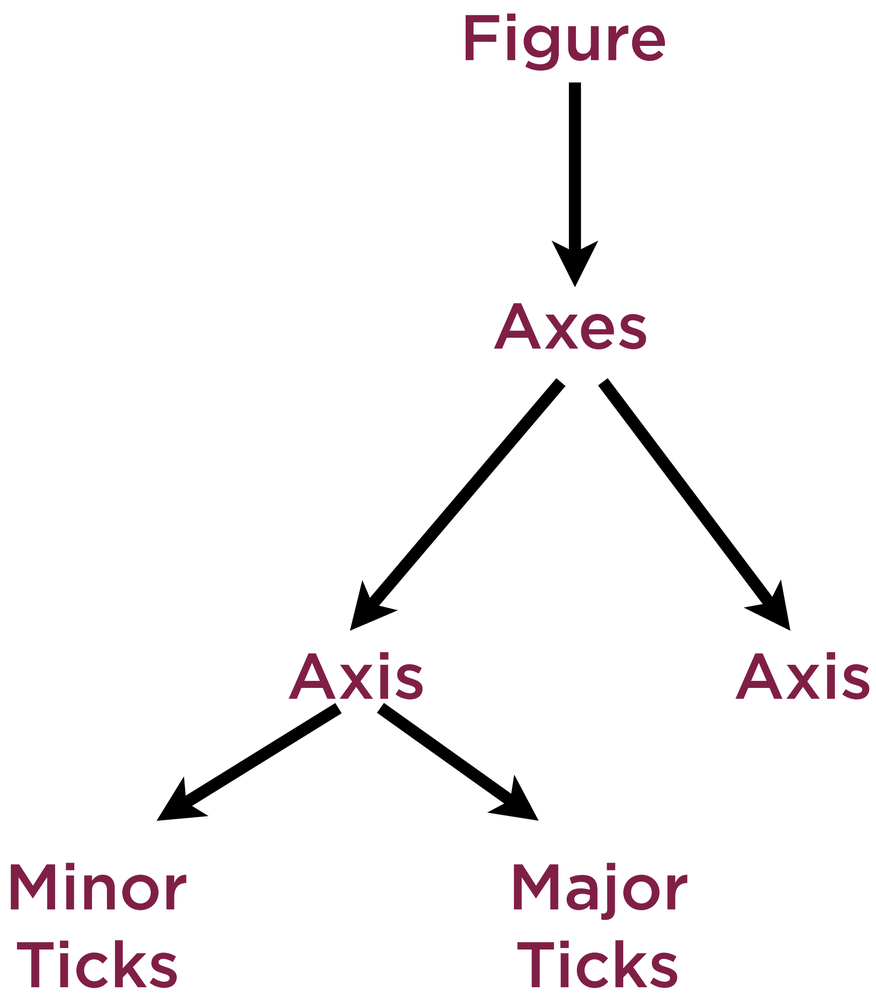
Hierarchy

Everything is an "Artist"

Artists are arranged in a hierarchy

Artist is an abstract base class

Figure is a container class

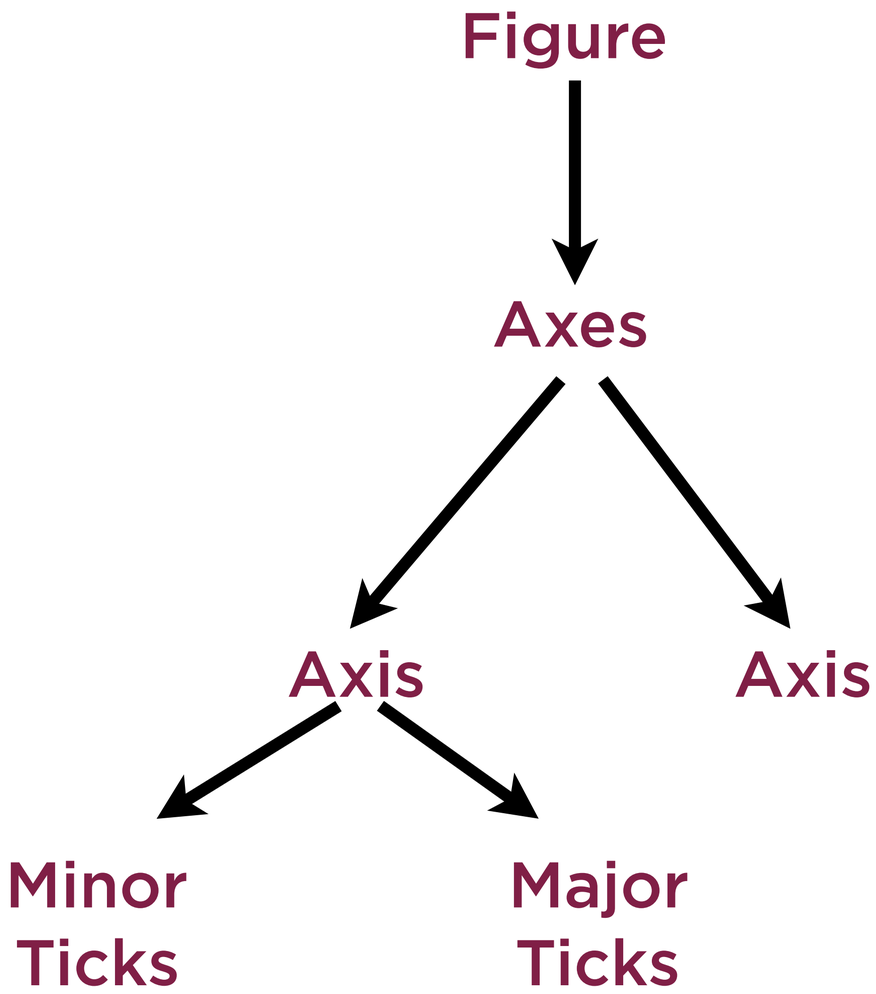


Hierarchy

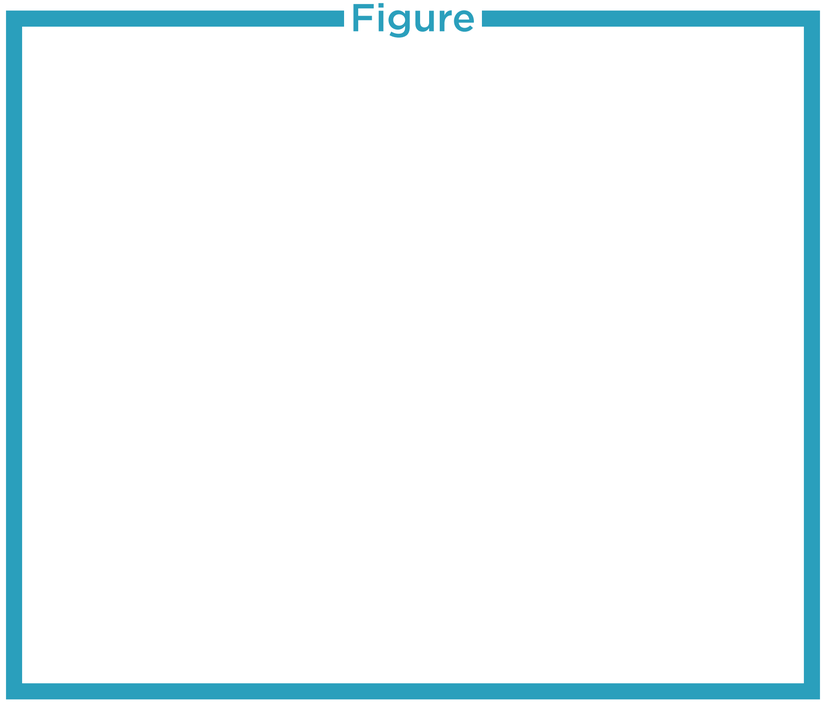
Figure is a top-level container

PyPlot APIs operate at higher levels

Matplotlib APIs at lower levels

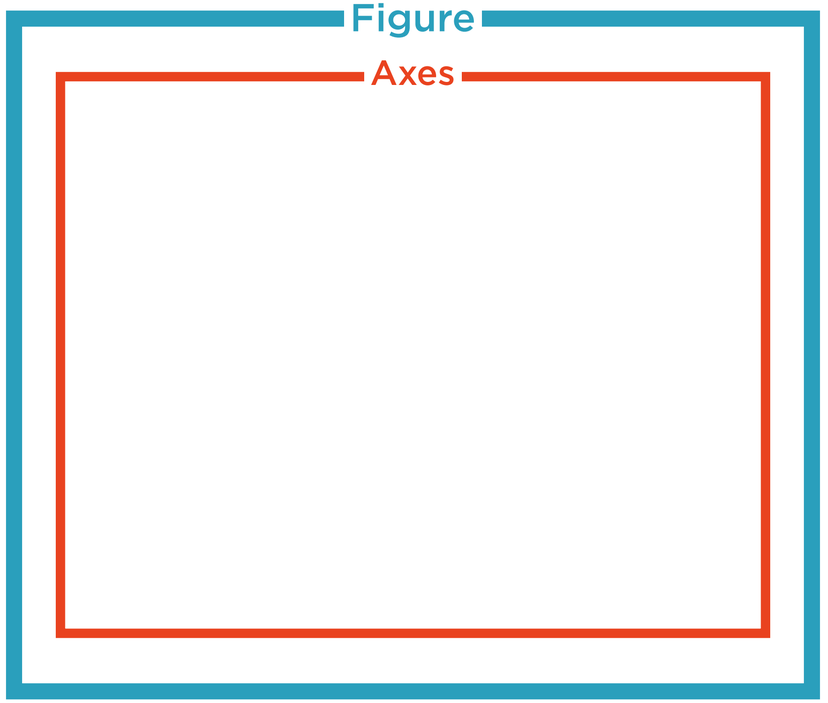


Matplotlib Object Hierarchy



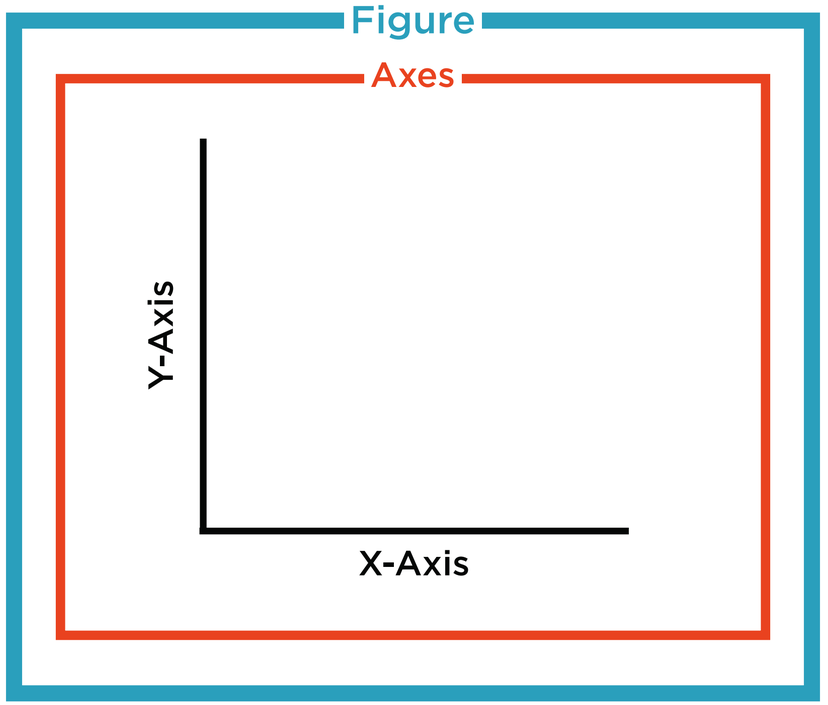
Overall window or page, within which all operations are performed

Matplotlib Object Hierarchy



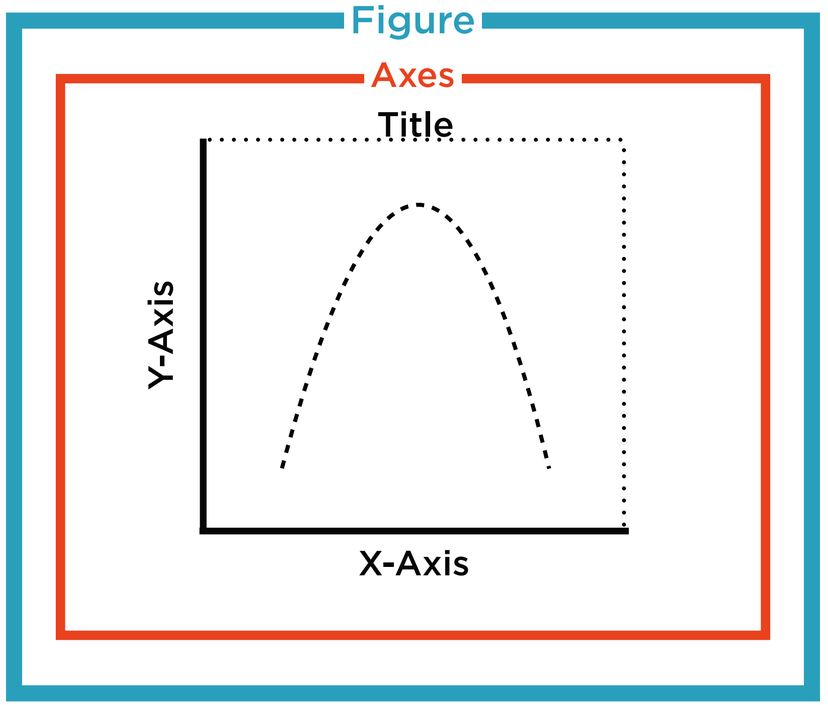
Area within a figure where actual graphs are plotted

Matplotlib Object Hierarchy



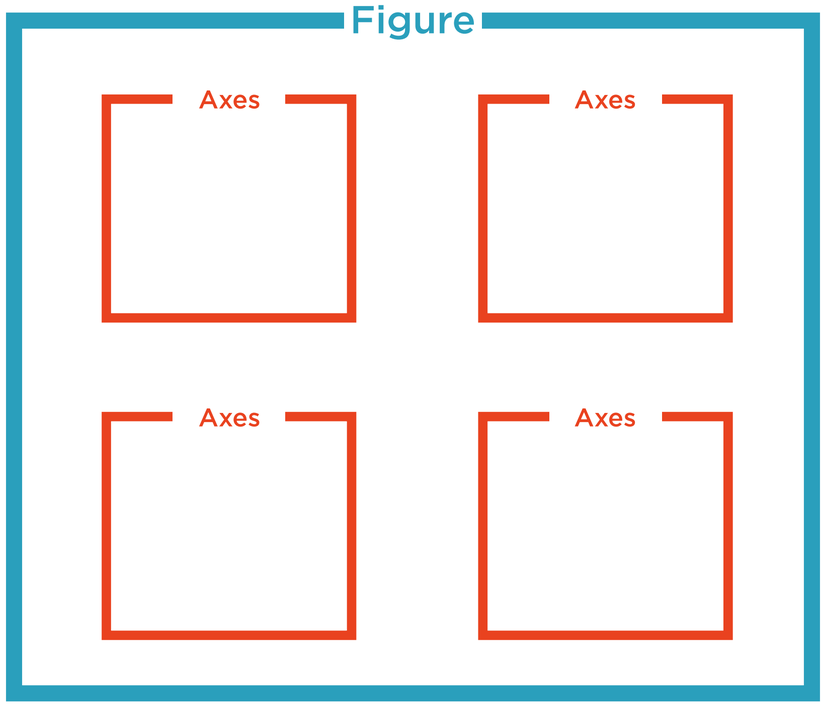
Axes has an X-Axis and a Y-Axis

Matplotlib Object Hierarchy



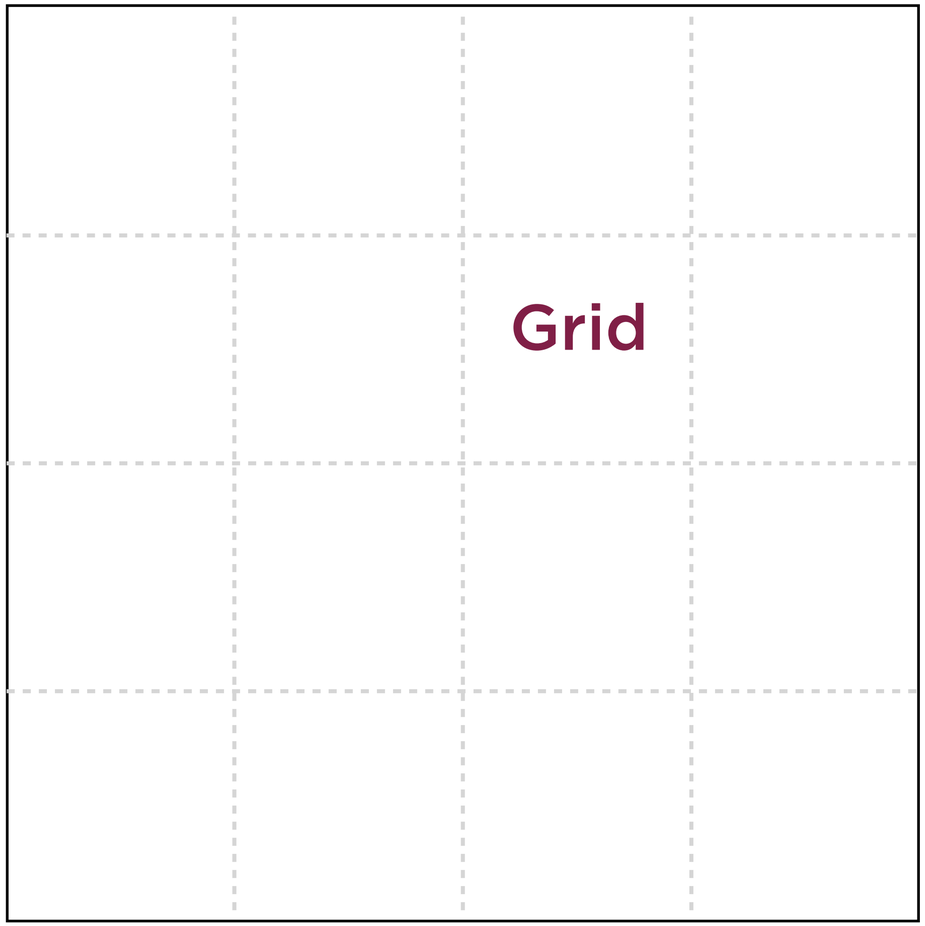
Contains ticks, tick locations, labels

Matplotlib Object Hierarchy

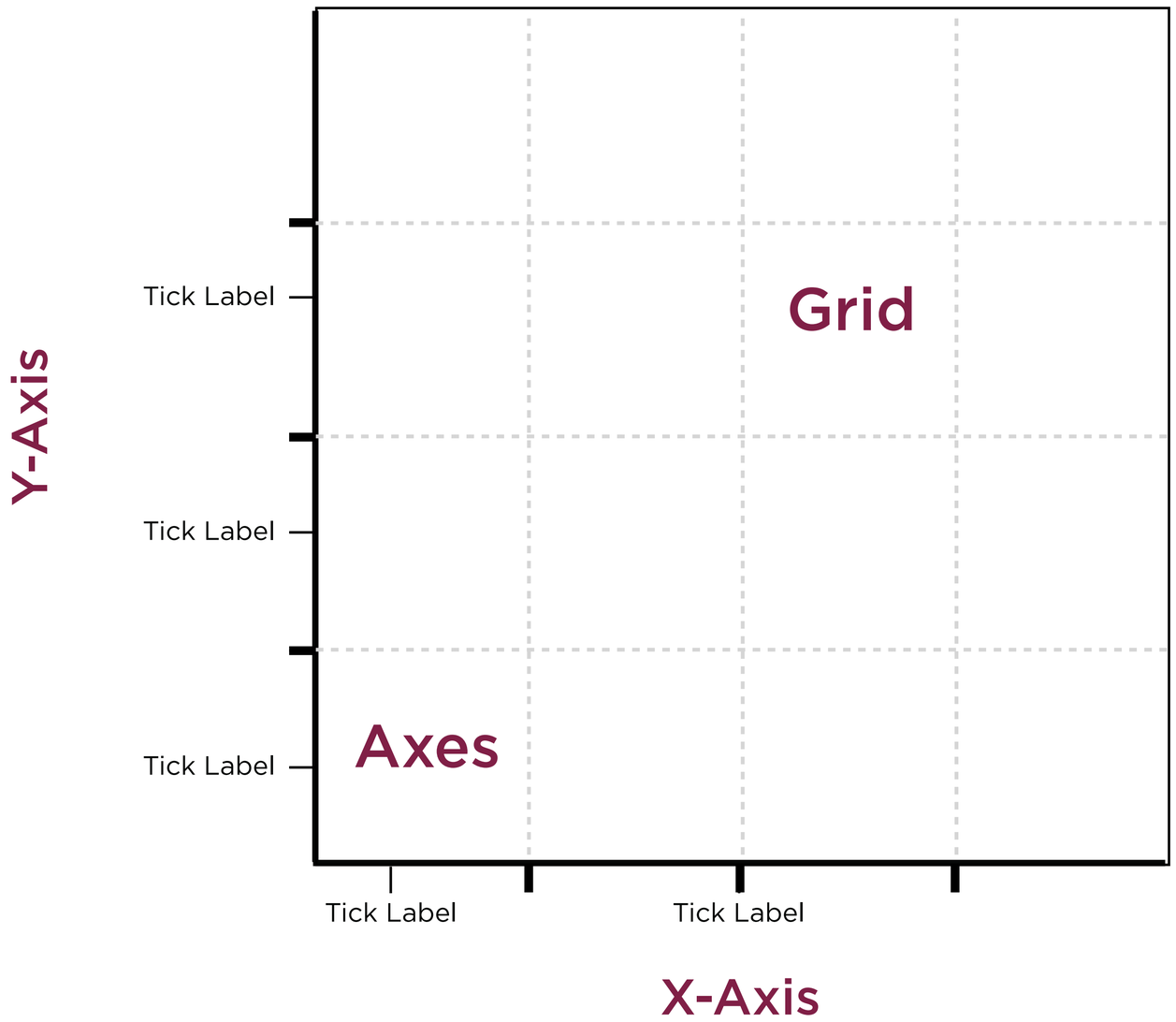


A figure can have more than one axes

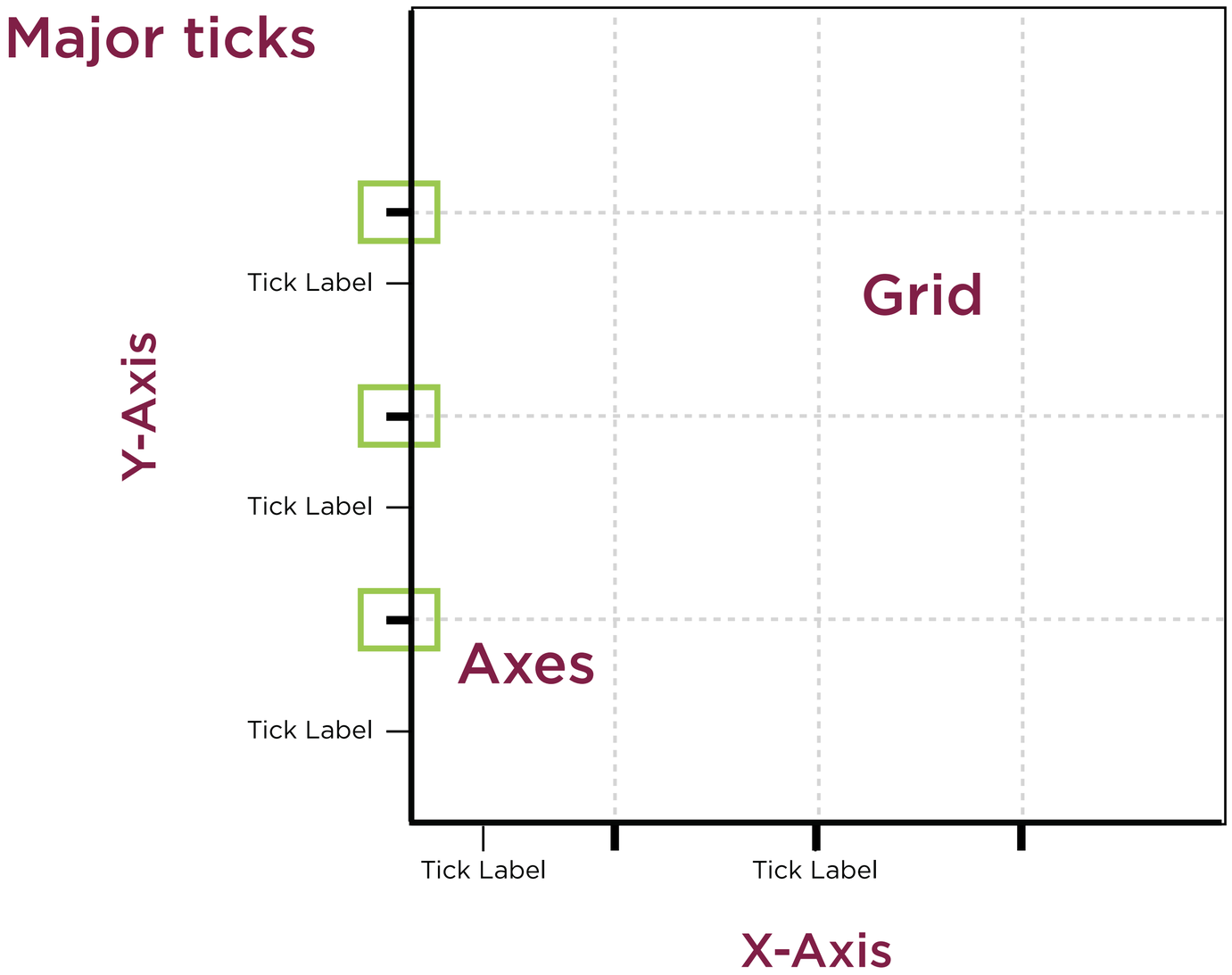
Anatomy of a Figure



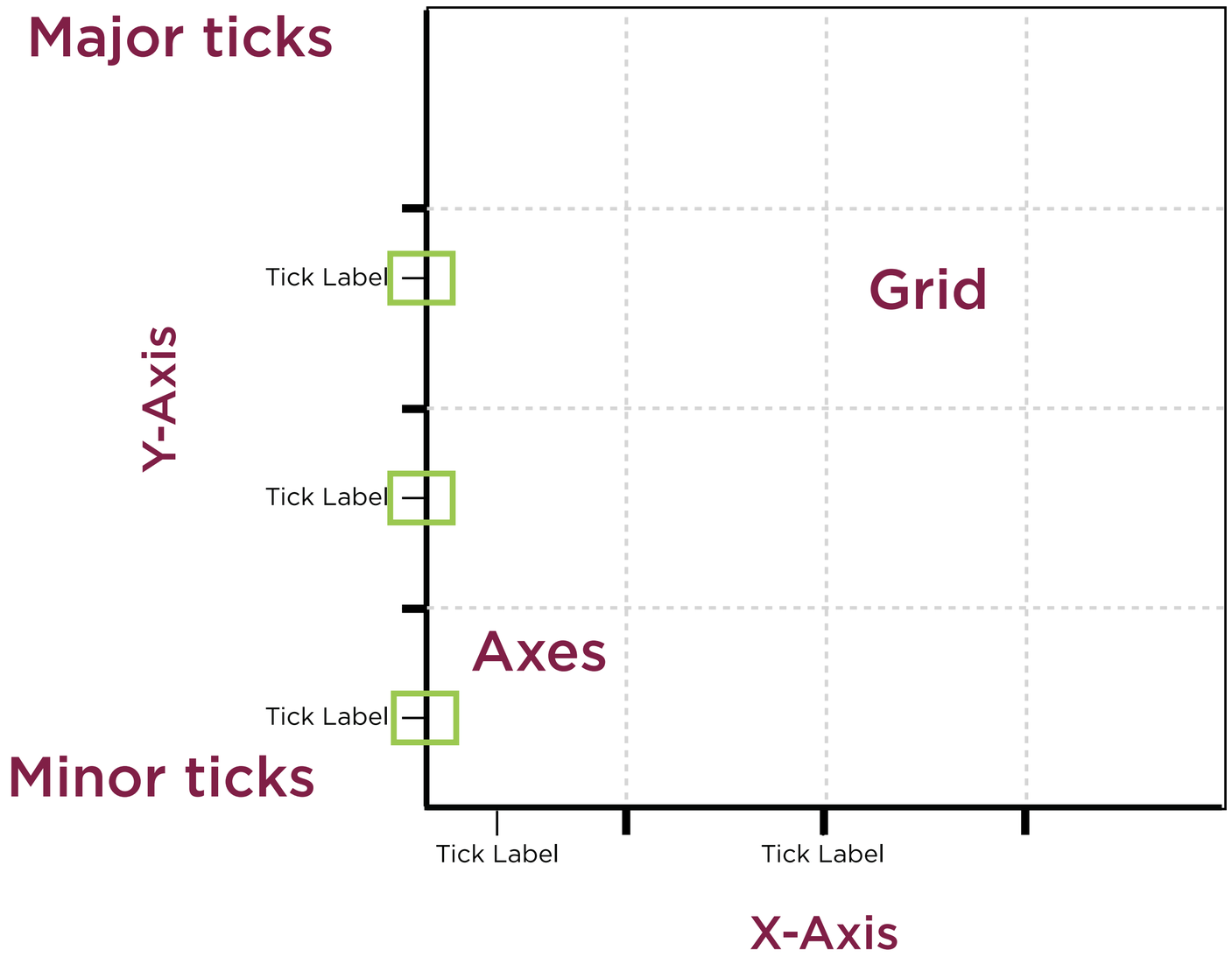
Anatomy of a Figure



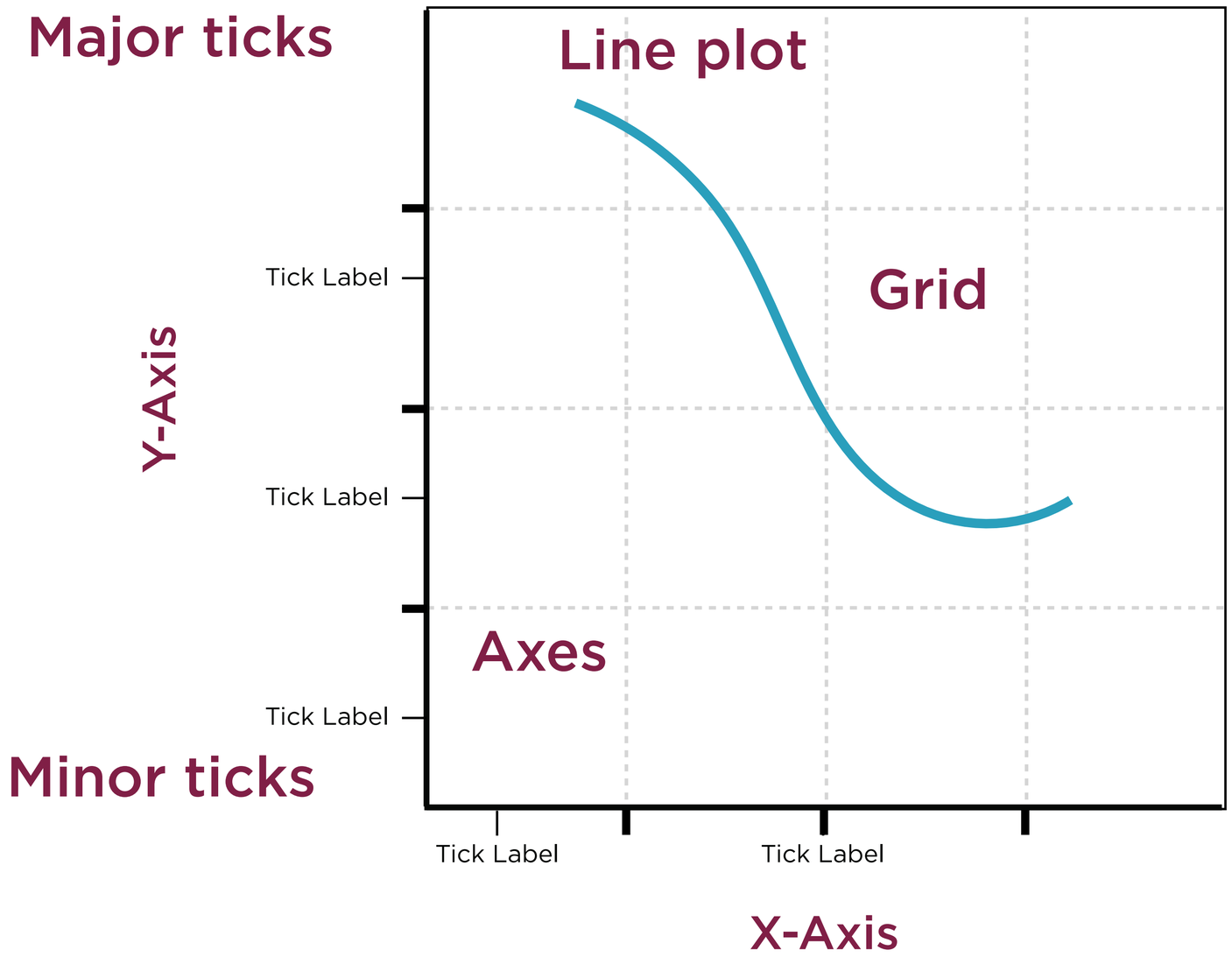
Anatomy of a Figure



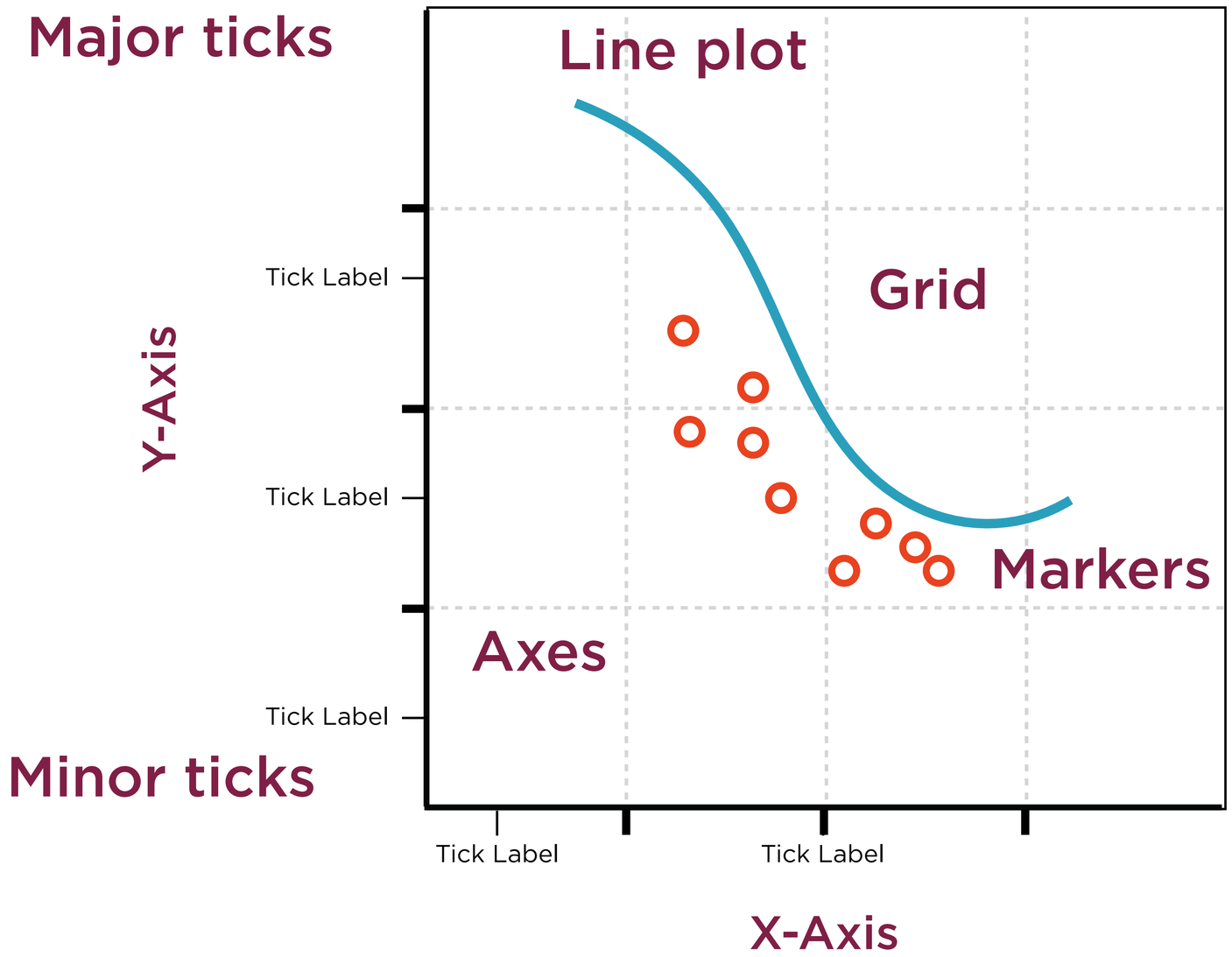
Anatomy of a Figure



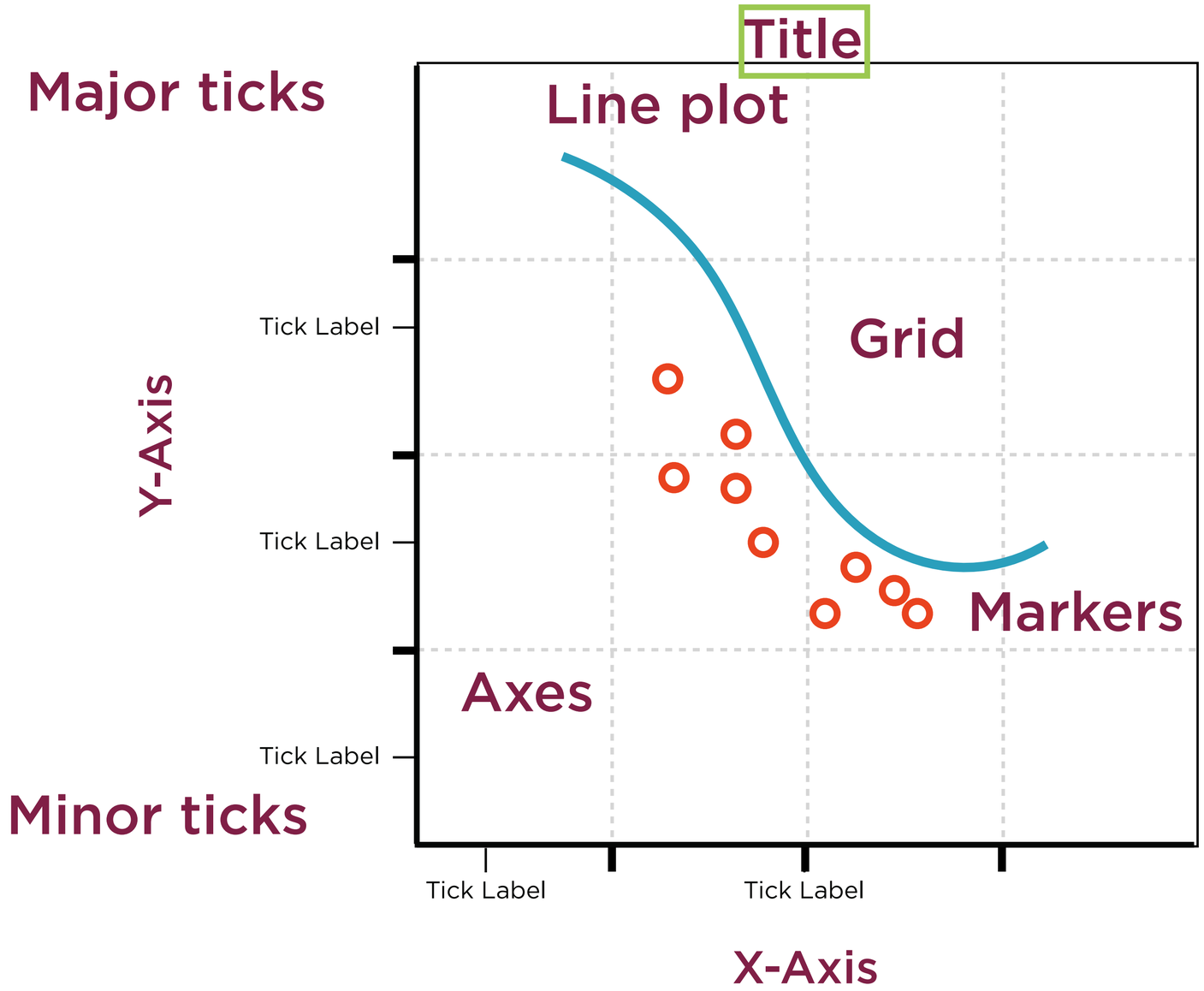
Anatomy of a Figure



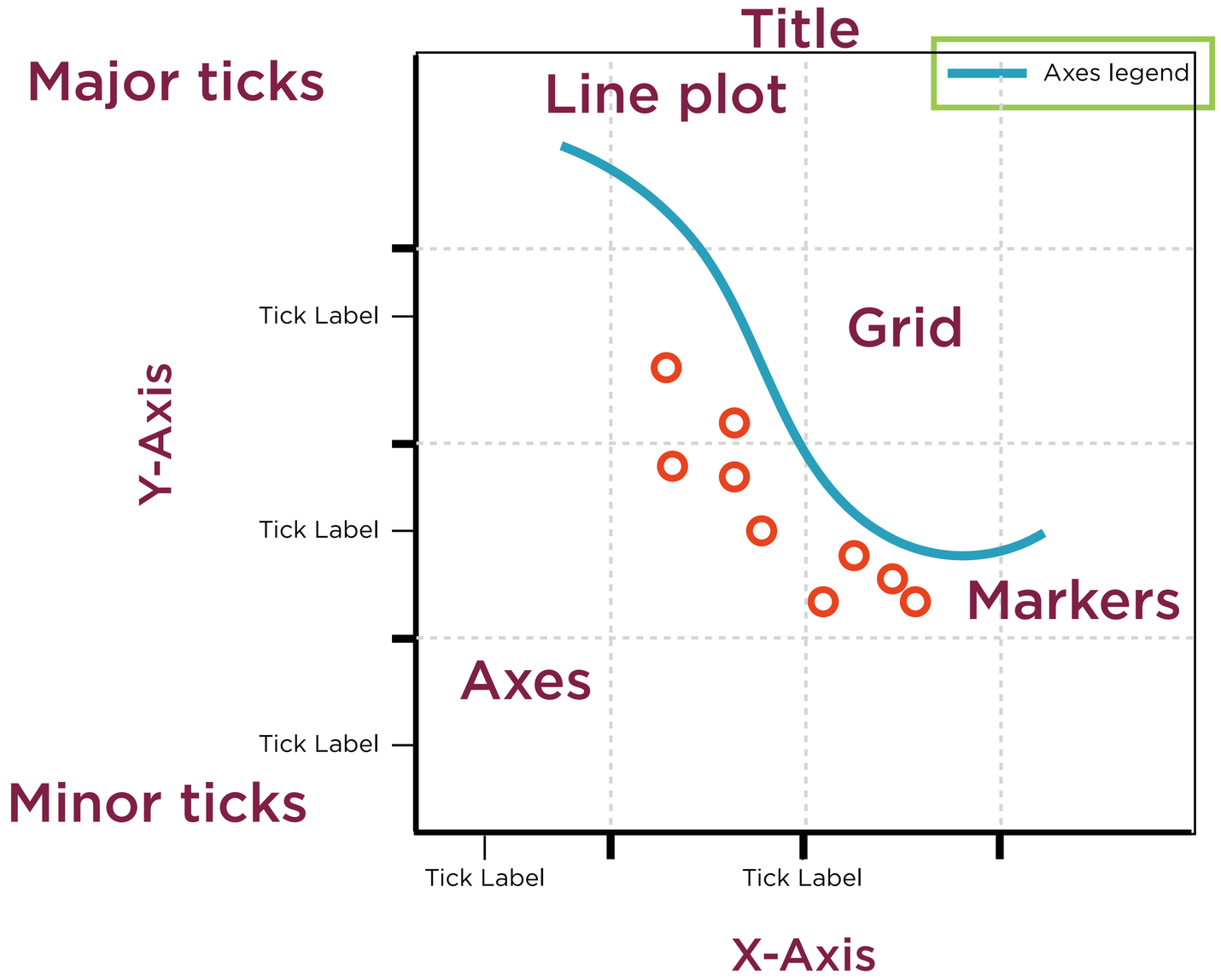
Anatomy of a Figure



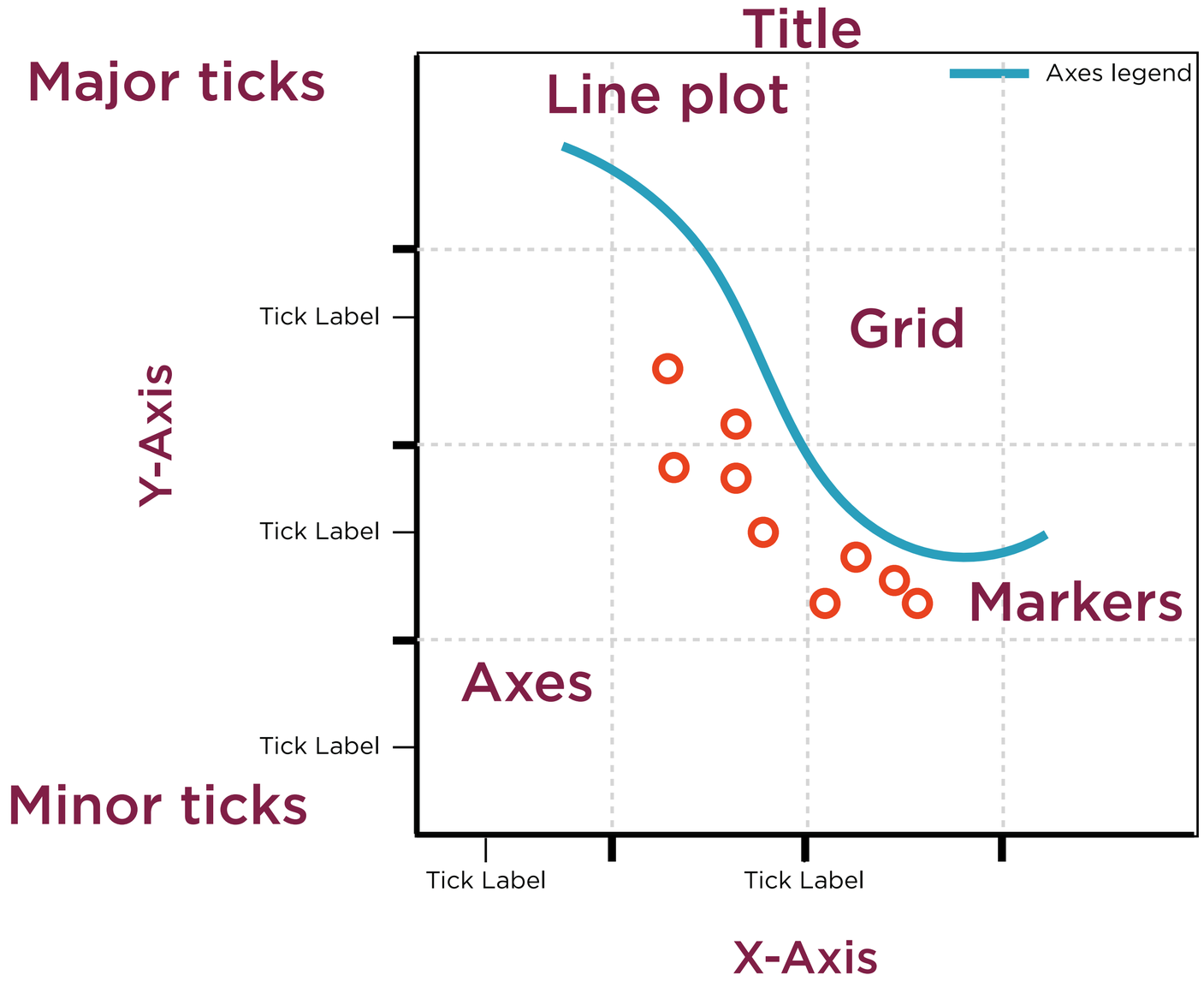
Anatomy of a Figure



Anatomy of a Figure



Anatomy of a Figure



Demo

Understanding the non-interactive mode in Matplotlib

Demo

Working with the interactive backend

Demo

Basic plots

Demo

Lines and markers

Demo

Figures

Demo

Watermarks

Demo

Bringing it all together with some real world data

Summary

Well integrated with other packages in the PyData stack

Anatomy of a Matplotlib figure

Basic plots, labels, titles, markers and watermarks

Figures, axes, subplots