Data Visualisation with Python

Level: Intermediate

Duration: 6 hours

Python has a number of packages for the effective creation of graphics to communicate your data insights. This course will examine two popular libraries for creating static 2D plots: Matplotlib and Seaborn. During the training session, we'll cover plotting basics and customisation of figures with Matplotlib, before moving onto complex statistical visualisations with Seaborn.

Course Outline

- **Matplotlib**: Introduction to the most widely used visualisation library for Python, covering basic plotting and formatting.
- **Plot building**: Using Matplotlib's object-oriented interface to build more complex plots made up of multiple plot panels.
- **Customisation**: Creating visually-appealing figures by customising fonts, axes and colours, and defining custom style sheets.
- **Seaborn**: Introduction to Seaborn, a very useful graphics package built on top of Matplotlib to aid in easy creation of beautiful statistical graphs.
- **Statistical visualisation**: Exploring your data using Seaborn's statistical functions, including regression models, kernel density estimation, bivariate distributions and pairwise plots.

Learning Outcomes

Session 1:

By the end of session 1, participants will...

- be familiar with the graphics landscape in Python:
  - Matplotlib related packages
  - Alternatives to Matplotlib
- understand Matplotlib's object oriented plotting interface.
- have explored different types of plots:
  - Line
  - Scatter
  - Histograms
  - Bar charts
  - Sub plots
- be comfortable constructing complex plot layouts with `GridSpec`.
- be able to customise Matplotlib figures:
  - Legends
  - Axes
  - Fonts
  - Colourmaps
  - Using predefined and custom style sheets

Session 2:

By the end of session 2, participants will...

- be introduced to basic plotting with Seaborn.
- be able to customise Seaborn figures using Matplotlib.
- be able to apply advanced statistical visualisations with Seaborn:
  - Bivariate relationships
  - Optimisation functions
  - Univariate and bivariate distributions
  - Kernel density estimation
  - Box plots
- understand how to create and customise multi-panel plots:
  - Facet grids
  - Pair grids

This course does not include:

- Animations and interactive plots.
- Data manipulation and data cleaning.
- Advanced data modelling techniques, see our website for courses on machine and deep learning.

Prior Knowledge

This course assumes you have some basic familiarity with Python programming and data structures including Pandas data objects. Completion of the Introduction to Python course or similar experience would be sufficient.

Attendee Feedback

- "The presenter really did an excellent job on explaining the content in the course and I believe that the actual content is very useful to know. I will definitely be using it"
- "It was well structured, the materials were clear and professionally produced, good enough to be followed on their own"

Contact

hello@jumpingrivers.com