CoGrammar

Advanced Data Visualisation and Data Analytics

SKILLS FOR LIFE
SKILLS BOOTCAMPS

Department for Education

Data Science Lecture Housekeeping

- The use of disrespectful language is prohibited in the questions, this is a supportive, learning environment for all - please engage accordingly. **(FBV: Mutual Respect.)**
- No question is daft or silly - **ask them!**
- There are **Q&A sessions** midway and at the end of the session, should you wish to ask any follow-up questions. Moderators are going to be answering questions as the session progresses as well.
- If you have any questions outside of this lecture, or that are not answered during this lecture, please do submit these for upcoming Open Classes. You can submit these questions here: **Open Class Questions**

Data Science Lecture Housekeeping cont.

- For all **non-academic questions**, please submit a query: **www.hyperiondev.com/support**
- Report a **safeguarding** incident: **www.hyperiondev.com/safeguardreporting**
- We would love your **feedback** on lectures: **Feedback on Lectures**

Lecture Objectives

- Generate graphs in Python using **Matplotlib and Seaborn**.
- Gain an understanding of more **advanced graphing techniques**.

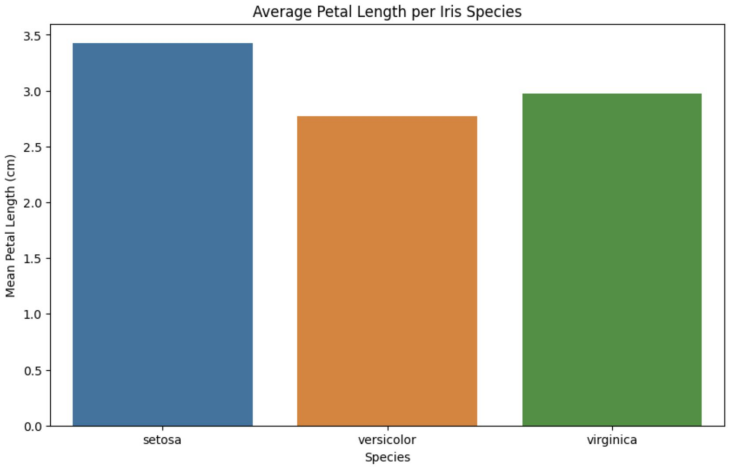
Key Goals of Data Visualization

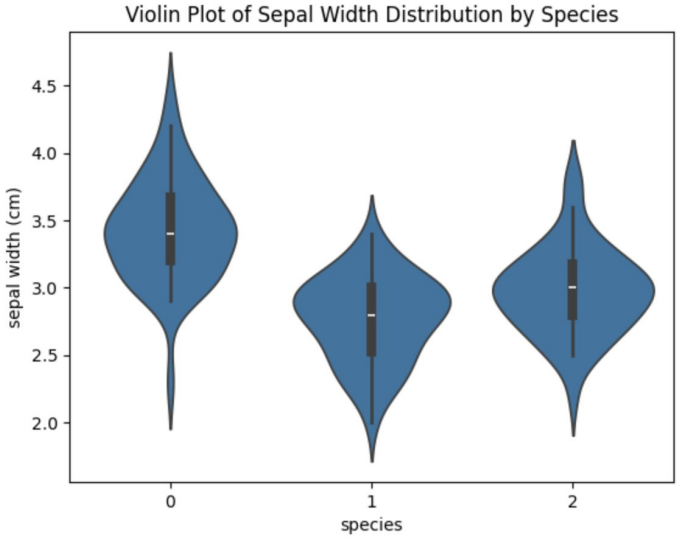
- ★ Expose **underlying trends, outliers, and relationships within data**.
- ★ Support **intuitive exploration** and **rapid insight generation.**
- ★ Create **visual narratives** that resonate with your audience.

Why Advanced Visualization?

- ★ Complex data demands **more sophisticated and flexible visual analysis** methods.
- ★ Answering **specific questions** sometimes requires **going beyond standard plots**.

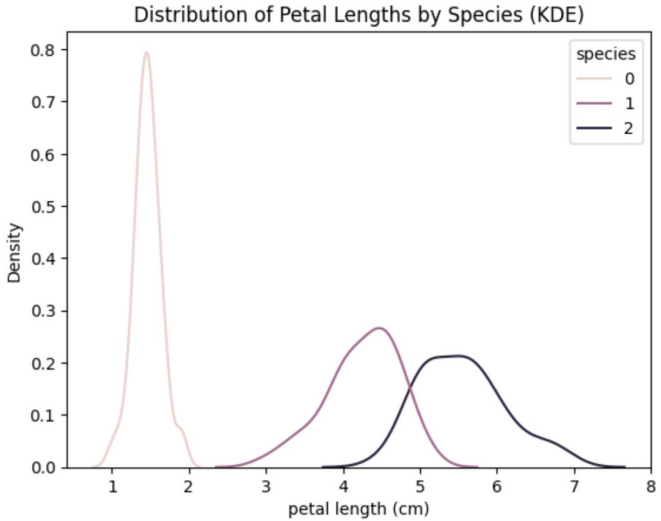
Beyond Standard Charts - Distributions





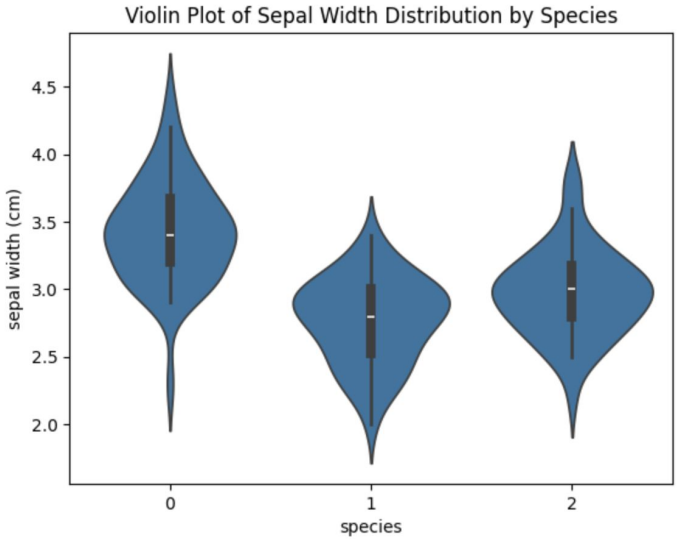
Kernel Density Plots (KDEs)

★ Smooth curves representing the **density of data points, great for comparing distributions of several groups**.



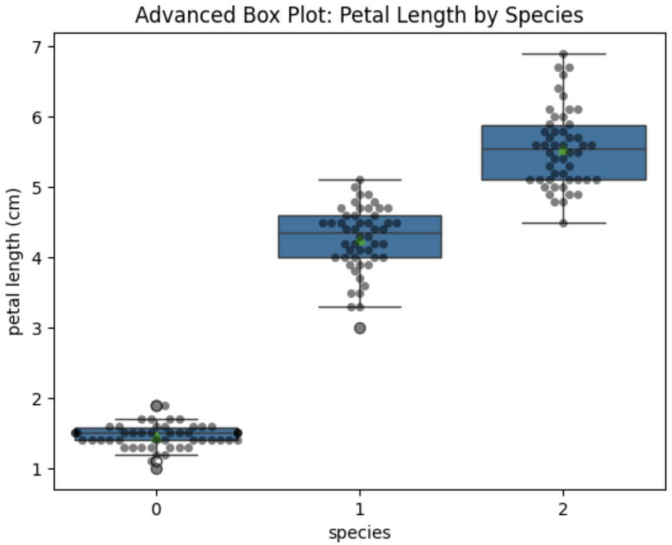
Violin Plots

★ Combine aspects of KDEs and boxplots, **ideal for showing density alongside summary statistics**.

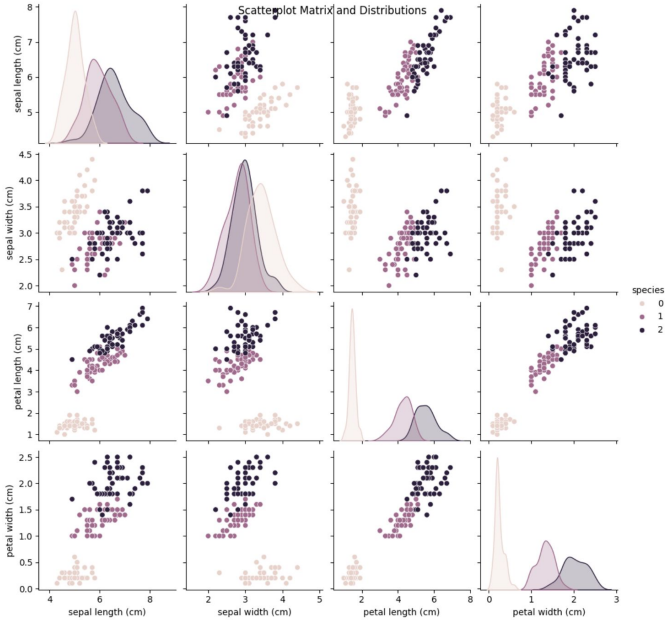


Advanced Box Plots

★ Overlay raw data points or swarm plots on top of standard box plots to **convey more intricate distribution details**.

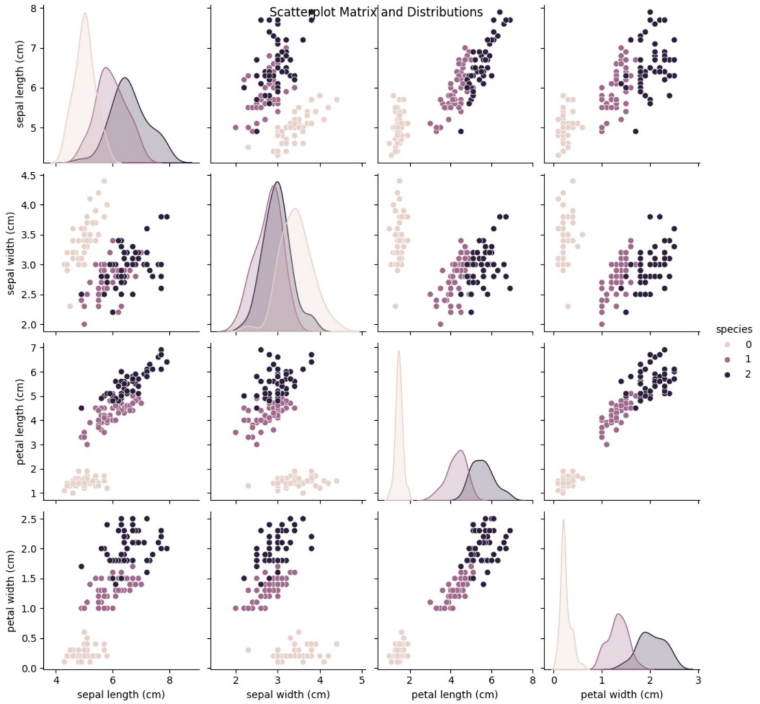


Beyond Standard Charts - Multidimensionality



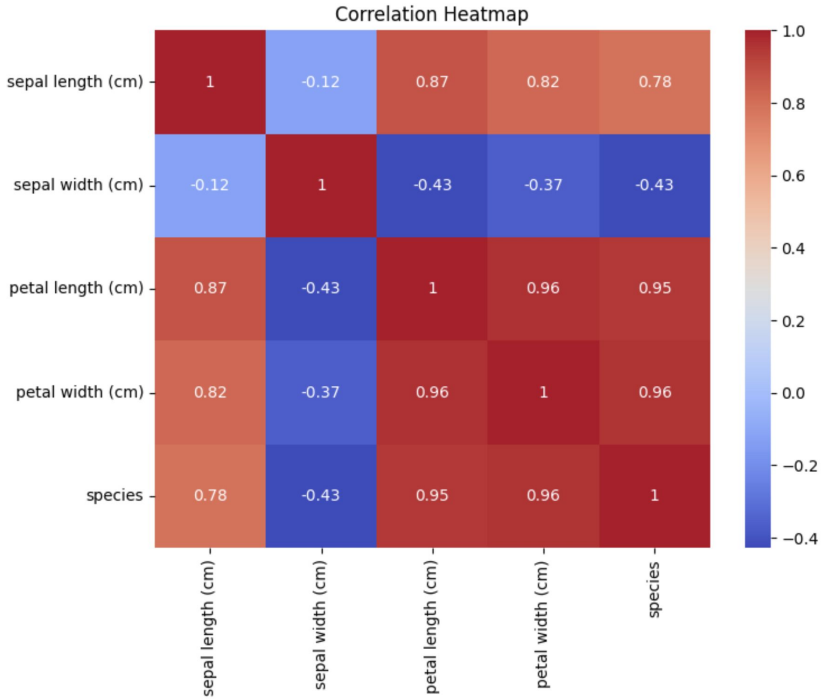
Scatter plot Matrices

★ Compact way to **depict pairwise relationships between several variables simultaneously**.



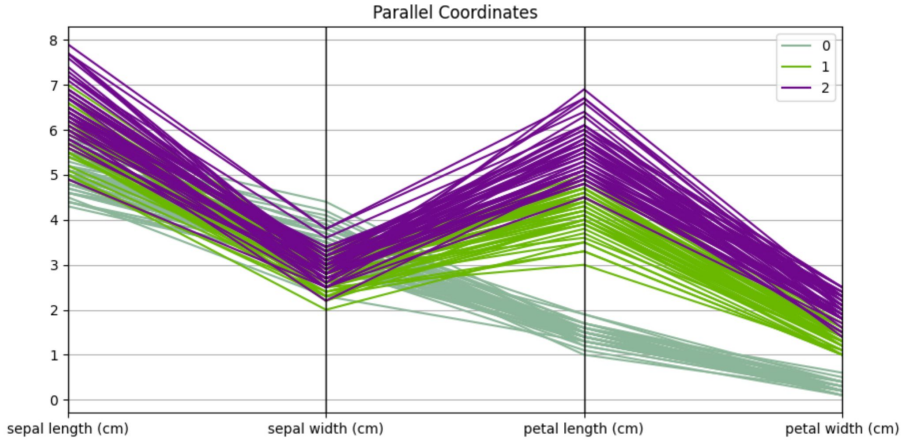
Heatmaps

- ★ Color-coded matrices excellent for **revealing structure, highlighting correlations, and identifying clusters**.
- ★ **Customization** in heatmaps is key for optimal interpretation.

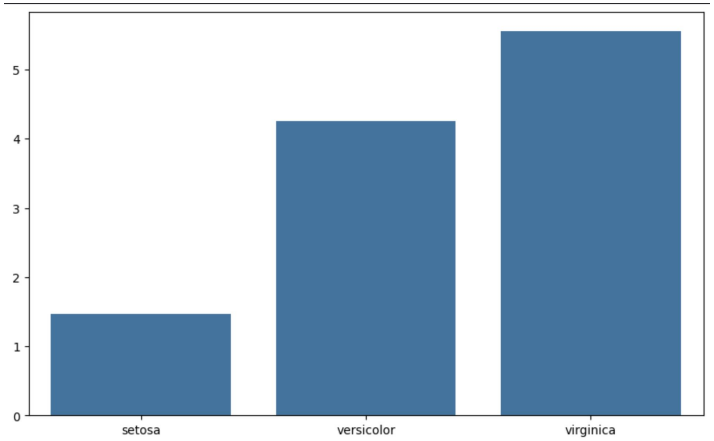


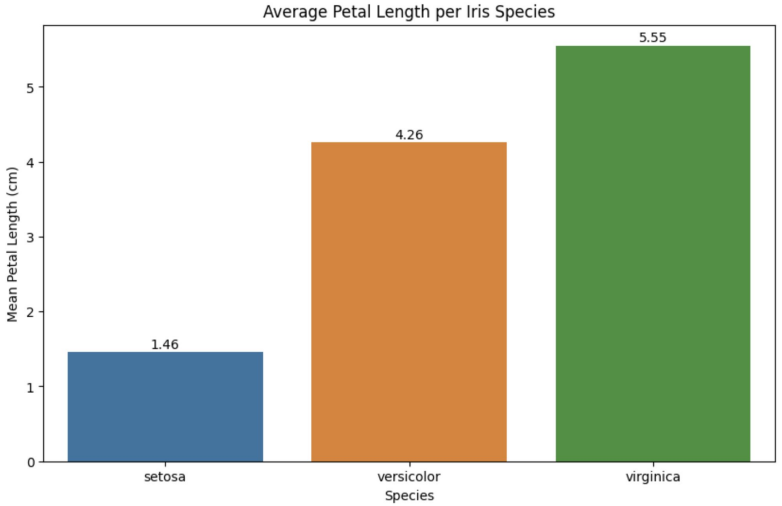
Parallel Coordinates

★ Show **many dimensions on the same plot**; each feature has a vertical axis, data points become lines crossing them. **Ideal when you have many interrelated variables and need to spot outliers or group characteristics**.



Principles of Effective Visualization





1. Visual Perception

- **Pre-attentive attributes** (color, position, size, shape) are quickly processed by our brains. **Use them deliberately for emphasis.**

- **Our brains group visual elements automatically**; leverage this fact by using **spatial arrangement, similarity, and clear layout to reinforce the key takeaways of your visualization**.

2. Chart Choice vs. Your Question

- ★ Comparing **individual values**? Consider **bar charts, or dot plots**.
- ★ Analyzing **trends over time**? **Line charts** are well-suited.
- ★ Interested in **proportions of a whole**? **Pie charts** (sparingly), treemaps, etc.
- ★ Focus on **relationships between variables**? **Scatter plots** (and their extensions) and **heatmaps** excel.

3. Less is More

- **Clear labels, thoughtful color choice, and minimal clutter improve the impact of your visual message**. Unnecessary embellishment creates a distraction.

CoGrammar

Q & A SECTION

Please use this time to ask any questions relating to the topic, should you have any.

CoGrammar

Thank you for joining!