Data Visualization in Python: Seaborn

A Reproducible Research Workshop

Simon Stone
Research Data Services
Dartmouth College

About the Reproducible Research Group

- Joint venture of **Research Computing @ ITC** and **Research Data Services @ Library**
- Consult with **experts** on
  - research data management,
  - data visualization,
  - biomedical research support,
  - spatial data and GIS,
  - high performance and research computing,
  - statistical analysis,
  - economics and social sciences data
- **Meet** the people on campus that support your reproducible research lifecycle
- **Engage** in community discussions to learn from other researchers on campus
- Attend a workshop to **learn** practical tools and tips

About Research Data Services

Research Data Management

Data Management Plans (DMPs) for sponsored projects

Finding and using 3rd party data

Collection and cleaning of data

Organization and documentation

Publishing and Repositories

Data Analysis/Visualization

Textual, numeric, spatial data

Reproducible research workflows

Scripting in R: tidyverse core package (i.e. ggplot, dplyr, tydr, tibble, etc.)

Scripting in Python: NumPy, SciPy, Pandas, Scikit-learn, Matplotlib, Seaborn, (OpenCV, PyTorch, TensorFlow, Tesseract, NLTK, etc.)

Computational Scholarship

Computational project planning

Collections as Data

Storytelling with data and visualizations

Text and data mining

Digital Humanities support

Computational Pedagogy

Work with us

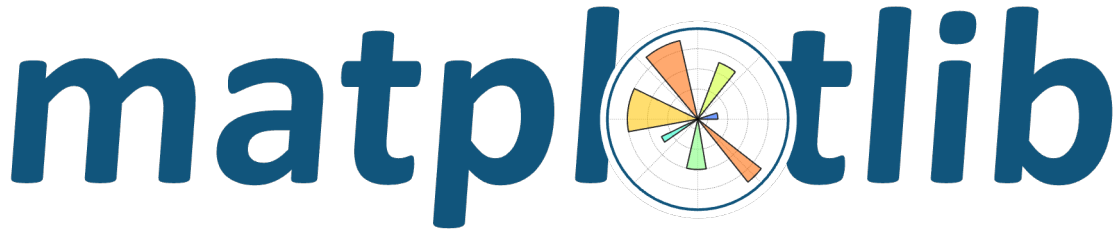
ResearchDataHelp@groups.dartmouth.edu

Jeremy Mikecz
Research Data Science Specialist
jeremy.m.mikecz@dartmouth.edu
dartgo.org/jeremyappts

Simon Stone
Research Data Science Specialist
simon.stone@dartmouth.edu
dartgo.org/meetwithsimon

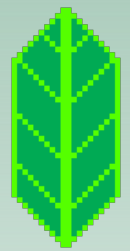
Lora Leligdon
Head of Research Data Services
lora.c.leligdon@dartmouth.edu
dartgo.org/lora

The Data Visualization Ecosystem in Python



Tabular data

Geospatial data



Folium

Dashboards

and many, many more...

Seaborn is...

📊 ...a data visualization library with a high-level interface specializing in statistical graphics.

 ...built on matplotlib.

🐼 ...the perfect match if you work with tabular data in pandas.

🎉 ...a popular choice to create publication-ready visualizations.

What you will learn in this workshop

- **How** seaborn is organized
- **How** to best organize your data to work with seaborn
- **How** to get started with seaborn's interface
- **How** to create visualizations showing multiple facets of your data

What we will work with in this workshop

- Platform: https://jhub.Dartmouth.edu
- Python
- Pandas (data storage)
- Seaborn
- Materials: www.dartgo.org/rr-seaborn

Let’s start plotting…

Next steps

- Explore seaborn's official tutorials and example gallery, covering a vast array of topics from basic to advanced!
- Learn more about styling and theming in seaborn!
- Go interactive and explore dashboarding with plotly express!

Thank you.