

Resources

DATA 322: Machine Learning for Data Science

Resources

This page provides links to essential course resources including the textbook, Python software downloads, and reference materials.

Lectures

Lectures - Access all course lectures organized by week and day.

Textbook

Required Textbook: *An Introduction to Statistical Learning with Applications in Python*

Available free online or for purchase in print. This textbook provides a broad and less technical treatment of key topics in statistical learning, with practical applications in Python.

Python Software Downloads

DATA 322 uses Python for machine learning and data analysis. You can use Python in several ways:

Python (Base Installation)

Download Python: [Python.org](https://www.python.org)

Python is free and open-source. The base Python installation includes the core language and standard library.

- **Windows:** Download the installer from [python.org](https://www.python.org)
- **Mac:** Python may be pre-installed, or download from [python.org](https://www.python.org)
- **Linux:** Install via your distribution's package manager or from [python.org](https://www.python.org)

Python Data Science Libraries

Commonly used libraries for this course include: - `numpy` - Numerical computing - `pandas` - Data manipulation and analysis - `scikit-learn` - Machine learning algorithms - `matplotlib` - Data visualization - `seaborn` - Statistical data visualization

Install via pip: `pip install numpy pandas scikit-learn matplotlib seaborn`

Python in Google Colab

Access Python in Google Colab: [Google Colab](https://colab.research.google.com/)

Google Colab provides free access to Python in a cloud-based notebook environment with many data science libraries pre-installed. This is useful if you don't want to install Python on your computer or need to work from different devices.

Python in Jupyter Notebook

Install Jupyter: [Jupyter Installation Guide](https://jupyter.org/install)

Jupyter Notebook is a popular environment for interactive Python development and data analysis.

1. Install Python first
2. Install Jupyter: `pip install jupyter`
3. Launch Jupyter: `jupyter notebook`

Note: Python and commonly used data science libraries are installed on all campus computers, so you can use the computer labs if you prefer not to install software on your personal computer.

Course Reference Materials

Python Reference

Python Quick Reference - A quick reference guide for common Python commands and functions used in this course for machine learning and data analysis.

Methods Map

Methods Map - A reference table of machine learning methods covered in the course, organized by topic.

Additional Resources

Campus Resources

- **Computer Labs:** Python and commonly used data science libraries are installed on all campus computers
- **ROSE (Reusable Office Supply Exchange):** [Free used computers](#) may be available if you need a computer

Getting Help

- **Office Hours:** See the [Syllabus](#) for instructor office hours
- **Canvas:** Check Canvas for announcements and course communications
- **Email:** Contact the instructor at rho3@humboldt.edu