2019 Ag-REEU Data Science Class Schedule

Schedule

(15-minute breaks in morning and afternoon, 1.5 hrs for lunch)

Oregon State University

Date	Morning 1 9:00-10:30	Morning 2 10:45a-12:00	Afternoon 1 1:30-3:00	Afternoon 2 3:15-4:45
Mon, June 17	1: Overview, Network Computing	2: Linux Commands	3: Shell Scripting	4: Shell Scripting
Tue, June 18	5: Python and Jupyter Intro: Basic Syntax, Comments, and Printing	6: Vars, Operators, F-Strings, and Lists	7: Command Line Arguments, Flow Control with Loops	8: Practicing Flow Control with Loops
Wed, June 19	9: Conditional Statements and Logical Operators	10: Practicing Conditional Statements and Logical Operators	11: File I/O: Reading and Writing Files	12: Practicing File I/O
Thu, June 20	13: Data Structures (Dictionaries)	14: Nested Data Structures	15: Processing Files in a Directory	16: Working with Strings and Pattern Matching
Fri, June 21	17: Other useful data structures (tuples, sets)	18: Defining Functions	19: Shell review, Jupyter tips & tricks	20: Computing outside the sandbox
Mon, June 24	21: Working w/ Python packages 22a: NumPy I	22b: NumPy II	23: Pandas, Data Frames	24: Matplotlib, PyPlot
Tue, June 25	25: Seaborn	26: Data management: principles and practices	Intro to team projects	Team Project Time
Wed, June 26	27: Plotly, Geo-plotting	28: SciPy, SciKit	Team Project Time	Team Project Time
Thu, June 27	29: Linear Regression	30: K-Means Clustering	Team Project Time	Team Project Time
Fri, June 28	31: PCA	32: A brief introduction to R	Team Project Review	Celebration! What's next?

Team Projects - Visualization and Analysis

- **Gene Expression:** Explore, visualize, and analyze maize leaf transcriptome data provided by the EMBL-EBI Expression Atlas project.
- **Sequence Motif Distribution:** Compute and plot the distribution of occurrences of various motifs in a given set of genomic sequences.
- **vis-NIR Soil Characterization:** Examine, graph, and characterize *in situ* soil properties gathered using visible and near-infrared spectroscopy.