

DATA 101 Topics

Topic	Suggested Number of Weeks
Intro to Data Science • What it is • Why it is important • Current examples	1/2
Data Tools Set-Up & Intro • R, RStudio, R Script • Tidyverse ◦ How to use libraries • RMarkdown • Git, GitHub	2
Data Ingestion • Read data in via ◦ Downloaded CSV ◦ Live Google Sheet	1/2
Data Storage • CSV • Database introduction ◦ MS Access	1 1/2

Exploratory Analysis • Data Types - numerical, categorical, discrete, continuous • Summary stats • Graphs - how to and when appropriate ◦ Bar graph vs histogram ◦ Boxplot ◦ Scatter plot ◦ Pie chart ◦ Crosstabs (contingency table) ◦ Line graph • Locate missing data	2 1/2
Data Cleaning & Management • R libraries for cleaning - tidyverse • Data types, factors, converting data types, functions, etc. • What to do with missing values • Finding and handling outliers ◦ Ethical concerns • Documenting (for example, if you deleted data) • Subsetting data	2 1/2
Statistical Analysis • Normal distribution (LLN, CLT, SD) • Empirical rule • Z scores • P values • Sampling methods • Confidence intervals • Hypothesis testing ◦ Assumptions for testing and CI • Simulation in R - computational hypothesis testing	3 1/2
Machine Learning • Linear Regression in one-variable	1

****Emphasize interpreting analysis results****

****Emphasize when/which analysis methods are appropriate for given data****