Week 1 Setting up a data science working environment

- anaconda, Rstudio, basics of python and R

Week 2 Probability and Statistics: a first look at the normal distribution

- normal distribution, sampling from a normal distribution, t-test, confidence intervals, p-values

Week 3 Programming with Data in R and Python

- data in R and python; dataframes - indexing, selecting,summarizing in R and python numpy; indexing; files and file I/O; functions in R and Python

Week 4 Discrete Probability and Bayes Theorem

- discrete probability, Bayes Theorem, NaiveBayes filtering

Week 5 Programming with Data in R and Python

- Pandas (Python) grouping, summarizing, selecting data; basics of plotting data
- R tidyverse grouping summarizing selecting data; basics of plotting
- means, variances, medians, categorical data
- more on functions in R and Python

Week 6 Linear Algebra

- Geometry of n-dimensional space, vectors, addition and scalar multiplication of vectors, the dot product, orthogonality, matrices; matrix multiplication, rank; computational examples in R and Python.

Week 7-8 Multivariate calculus

- discussion of functions of several variables; graphs and contour plots; review of the derivative in one variable and partial derivatives; directional derivatives; gradient descent. Illustrated in R and Python
    –

Week 9 A deeper dive into Visualization

- more on ggplot and its capabilities
- python plotting packages such as seaborn and bokeh

Week 10 Statistical models

- What is a statistical model?
- Likelihood and model parameters
- Maximum likelihood estimation and gradient descent
- Illustrated with OLS and Logistic Regression

Week 11-12 More on Statistics

- Another look at the normal distribution; the multivariate normal
- covariance, correlation
- estimation of parameters for linear regression and logistic regression
- significance and confidence intervals
- null and alternative hypotheses
- p-values

Week 13 Version Control

- Git as a tool (command line and through R studio)
- commits, branches
- remotes and github
- using github to host a web page for a project
- collaboration using Git – pull requests; contributing to open source projects

Week 14 Advanced topics in programming

- Data structures; object oriented concepts
- Essential notions from data structures:
    - stacks, lists, hashing
- Python classes
    - data attributes and methods

Week 14 Databases

- What is a relational database? tables, keys, indices, joins
- Basic SQL for getting data from a database