Syllabus

Last updated: 16 August 2020

Course coordinator. Assoc. Prof. Nelson Uhan 🖂 uhan@usna.edu

Course objectives. By the end of this course, students will be able to:

- (i) Wrangle (i.e. clean and manipulate) large, messy data sets into forms suitable for modeling and analysis (in particular, optimization, simulation, and statistical models).
- (ii) Create sophisticated visualizations of large data sets that provide useful insights for decision-making as well as further modeling and analysis.
- **Textbook.** There is no required textbook for this course. You will be provided with lesson notes and links to supplementary readings throughout the semester.

Schedule. Here is a *very* tentative schedule.

Week Topics

Warm up

1 A very brief introduction to Pandas Method chaining

Data visualization with Altair

- 2 Altair basics: data types, encoding channels, graphical marks
- 3 Basic data transformations: binning and counting, aggregation Advanced data transformations: calculate, filter, window
- 4 Basic top-level chart configuration Scales: axes, colors Advanced top-level chart configuration
- 5 Multi-view composition: layers, concatenation, facets
- 6 Review Exam 1
- 7 Interactive visualizations: tooltips Interactive visualizations: dynamic queries
- 8 Cartographic visualization: geoshape marks, point maps, symbol maps, chloropleth maps, lookup transforms

Data wrangling with Pandas

Brief Python review: functions and lambda functions; classes, methods, and attributes; for loops; f-strings

9 The Series and DataFrame objects
Basic arithmetic operations on Series and DataFrame objects, broadcasting
Applying functions to Series and DataFrame objects

1

10	Selecting and dropping data Adding new columns Dealing with missing data
11	Grouped operations: split-apply-combine Tidy data: long vs. wide data, reshaping, pivoting
12	Review Exam 2
13	Relational data: merge, join, and concatenate Dealing with dates and times Dealing with strings
Additional topics	
14	Scraping data Passing data between R and Python
15	Review or additional topics
16	Review or additional topics
17	Review or additional topics