STOR 320: Introduction to Data Science

Contents

This course is an application-driven introduction to data science. Statistical and computational tools are valued throughout the modern workplace from Silicon Valley startups, to marine biology labs, to Wall Street firms. These tools require technical skills such as programming and statistics. They also require professional skills such as communication, teamwork, problem solving, and critical thinking.

- Instructor: Mario Giacomazzo
- Instructional Assistant: Adam Waterbury
- Graduate Research Consultant: Varun Goel

See the **course syllabus** for more information.

Course Material

Date	Lecture	Slides	Notes & HW
August 23	Course Overview, Install R	Lecture 1	R Downloads
August 25	Visualization and ggplot2	Lecture 2	Homework 1
August 28	ggplot Wrap-up and Workflow Basics	Lecture 3	Reading for 8-28
August 30	Data Transformations with Dplyr	Lecture 4	Homework for 8-30
Sept. 1	Data Transformations with Dplyr	Class Activity	Homework for Sept 1
Sept. 6	Data Transformations with Dplyr	Still Lecture 4	Homework for Sept 6
Sept. 8	Workflow - Scripts, RMarkdown	Lecture 5	Data Analysis 1
Sept. 11	Exploratory Data Analysis	Lecture 6	

Date	Lecture	Slides	Notes & HW
Sept. 13	Exploratory Data Analysis	Lecture 7	Homework
Sept. 15	Tibbles and R Projects	Class Activity	Activity as .rmd
Sept. 18	Importing Data	Lecture 9	heights.csv
Sept. 20	Tidy Data	Lecture 10	MOMA_art.csv MOMA_artists.csv
Sept. 22	Tidy Data –> Joins	Lecture 11	Joins Homework
Sept. 25	Tidy Data - Wrap-up	Lecture 10 Redux	HW Due Weds. Script last class
Sept. 27	Owning Strings	Notes	Slides_____ HW for Friday
Sept. 29	Owning Strings	Slides as Rmd	Data Analysis 2
Oct. 2	Factors	Slides Lecture 12	
Oct. 4	Programming and Vectors	Notes L13	Slides
Oct. 6	Programming Lab	Lab	
Oct. 9	Lists and Loops	Slides	HW
Oct. 11	Lists and Loops	Slides	Loops HW
Oct. 13	Intro to Models	Slides	Reading
Oct. 16	Intro to Models	Slides	HW Due 10-23
Oct. 23	Intro to Shiny	Shiny Notes Shiny Rmd	Apps for Class
Oct. 25	Data Ethics	Lecture	data privacy _____NYT Uber Oneill Big Data
Oct. 27	Visualizing Models	Lecture	Final Proj Gps Final Project
Oct. 30	Visualizing Models	Lecture	
Nov. 01	Interaction	Commands	Homework

Date	Lecture	Slides	Notes & HW
Nov. 03	Interaction (ctd)	Lecture	
Nov. 06	Modeling Real Data	Activity	
Nov. 10	Predictive Modeling	Lecture	Project Exp An Data Analysis 3
Nov. 13	Get an A on the Project	Lecture	Lecture as ppt
Nov. 15	Classification	Slides	Everything
Nov. 20	Donner Destiny	Slides	Classification Folder Classification Activity
Dec. 4	Project Presentations and Donners	Surveys	Donner_Modified
Dec. 4	Project Presentations and Donners	Surveys for Exam Day	

- most of the course material is in the lecture notes (linked to above) and reading.

Reading

- R for Data Science (r4ds)
- R Programming for Data Science (RPDS)

Additional resources

Miscellaneous

This course was made possible by a grant from the Data@Carolina initiative and a ton of **input from lots of very smart people**.

This page was last updated on 2018-08-14 13:10:32 Eastern Time.