STA 199: Intro to Data Science

Intro to data science and statistical thinking. Learn to explore, visualize,and analyze data to understand natural phenomena, investigate patterns, model outcomes,and make predictions, and do so in a reproducible and shareable manner. Gain experience in data wrangling and munging, exploratory data analysis, predictive modeling, data visualization, and effectively communicating results. Work on problems and case studies inspired by and based on real-world questions and data. The course will focus on the R statistical computing language.

Course info

Lectures

Soc Sci 139 Tue and Thu 10:05a - 11:20a

Labs

Lab 01 Old Chem 003 Fri 10:05a - 11:20p

Lab 02 Soc Psych 127 Fri 11:45a - 1p

Lab 03 Old Chem 003 Fri 1:25p - 2:40p

Teaching team and office hours

Instructor	Prof. Maria Tackett (http://stat.duke.edu/~mt324/)	(mailto:maria.tackett@duke.edu) (https://github.com/matackett)	Thu 1p - 2:30p	Old Chem 118B
TAs	Salvador Arellano (https://www.linkedin.com/in/salvador-chavero-arellano-405969168/)	(mailto:salvador.chavero.arellano@duke.edu) (https://github.com/salvadorchavero)	Wed 5p - 7p	Old Chem 203B
	Max Bartlett (http://maxbartlett.com/)	(mailto:maxwell.bartlett@duke.edu) (https://github.com/MaxBartlett)	Tue 6:30p - 8:30p	Old Chem 203B
	Meredith Brown (https://www.linkedin.com/in/meredith-brown-807964172/)	(mailto:meredith.brown@duke.edu) (https://github.com/meredithb3)	Mon 12:30p - 2:30p	Old Chem 203B
	Steven Herrera (https://www.linkedin.com/in/rosvidstevenherrera/)	(mailto:rosvid.herrera.tenorio@duke.edu) (https://github.com/stevenherrera24)	Thu 5:30p - 7:30p	Old Chem 203B
	Malavi Ravindran (https://www.linkedin.com/in/malavi-ravindran-332035175)	(mailto:malavi.ravindran@duke.edu) (https://github.com/MalaviRavindran)	Mon 3p - 5p	Old Chem 203B
	Becky Tang (https://beckytang.rbind.io/)	(mailto:becky.tang@duke.edu) (https://github.com/beckytang)	Wed 3p - 5p	Old Chem 203B

Texts

All books are freely available online. Hardcopies are also available for purchase.

R for Data Science (http://r4ds.had.co.nz/)	Grolemund, Wickham	O'Reilly, 1st edition, 2016
OpenIntro Statistics (https://www.openintro.org/stat/textbook.php)	Diez, Barr, Çetinkaya-Rundel	CreateSpace, 4th Edition, 2019
Introductory Statistics with Randomization and Simulation (https://www.openintro.org/stat/textbook.php?stat_book=isrs)	Diez, Barr, Çetinkaya-Rundel	CreateSpace, 1st Edition, 2014

Materials

You should bring a fully-charged laptop, tablet with keyboard, or comparable device to every lecture and lab session.

This course has achieved Duke's Green Classroom Certification. The certification indicates that the faculty member teaching this course has taken significant steps to green the delivery of this course. Your faculty member has completed a checklist indicating their common practices in areas of this course that have an environmental impact, such as paper and energy consumption. Some common practices implemented by faculty to reduce the environmental impact of their course include allowing electronic submission of assignments, providing online readings and turning off lights and electronics in the classroom when they are not in use. The eco-friendly aspects of course delivery may vary by faculty, by course and throughout the semester. Learn more at https://sustainability.duke.edu/action/certification/classroom (https://sustainability.duke.edu/action/certification/classroom).

DUKE STATSCI (HTTP://STAT.DUKE.EDU/) RSTUDIO (HTTPS://RSTUDIO.CLOUD/)

GITHUB (HTTPS://GITHUB.COM/STA199-FA19) (./)