Honors Computational Social Science (ISS 4304), 3 credits

Instructor: Kevin Lanning

Term: Fall 2019

Office: WB 213; Office hours Monday & Tuesday 2-5

Class Meeting Days/Time: MW 1230-150

Class Location: WB 105
Phone: 561-799-8652

Email: lanning@fau.edu
instructor website: wise.fau.edu/~lanning
course website: will be available at <https://kevinlanning.github.io>

syllabus: **honors computational social science**

# course description and objectives

New technologies have created new opportunities for the understanding of social behavior, and, with it, new threats to privacy and equality. In ISS 4304, we will (a) review selected examples of computational research from the social sciences, and (b) gain training in methods such as natural language analysis and social network analysis in the service of (c) research projects in which students will collaboratively implement these techniques in case studies of social phenomena. Objectives include greater mastery of the aforementioned techniques, of the R programming language and reproducible methods, and consideration of the responsibilities of data analysis and data stewardship in the contemporary world.

The course is intended to build on the Introduction to Data Science course (COP 3073) and is intended to serve students in all concentrations of the Wilkes Honors College, particularly those in the behavioral and social sciences of economics, psychology, linguistics, geography, anthropology, and sociology, as well as students anticipating working in data-intensive careers. As with the Data Science course, we will primarily be working using the statistical and graphical language R. In addition to R, we'll use a range of other tools, including the Slack platform for communication and collaboration, markdown editors such as Typora for writing, and spreadsheets such as Excel or Google Sheets.

# note of honors distinction

The course receives honors credit by virtue of its small class size, by virtue of a dialectic approach in the classroom structure, and by the fact that students receive extensive exposure to supplementary materials and primary sources. This course differs substantially from a non-Honors course in that (a) the expectations for participation in class discussions will be greater than in a typical undergraduate course with a larger number of students, (b) class projects will be undertaken in heterogeneous groups in which students will be teaching and learning from their peers as well as the instructor, (c) assignments and expectations will be, to some extent, tailored to the backgrounds and interests of the individual student (d) the data sets we will collaboratively examine will be chosen to foster disciplinary breadth.

# course prerequisites/co-requisites

STA 2023 (or equivalent) and COP 3073 (or equivalent), or permission of instructor, are prerequisites.

#

# required texts and materials

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press. (<http://www.cs.cornell.edu/home/kleinber/networks-book/>.)

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York: Springer. (<http://www-bcf.usc.edu/~gareth/ISL/>)

Salgarnik (2017). *Bit by Bit: Social Research in the Digital Age.* (<http://www.bitbybitbook.com/en>).

All of these are available for free online or not-for-free in hardcopy. Additional sources are listed in [references](https://kevinlanning.github.io/DataSciSpring2018/references) below.

You will also need access to a laptop computer running either Windows or Mac OS. (If you don’t have access to a laptop for everyday use, please see the instructor as soon as possible; I will secure one on your behalf).

# requirements / course evaluationand grading

Grades will be based on a 100 point scale, with points earned by participation, homework and quizzes, a term project, and a final exam.

Participation (20 points). Attendance is a necessary but not sufficient part of class participation. Your participation grade will be based also on the extent to which you contribute to our class by asking constructive questions and helping us solve the numerous challenges which we collectively will face.

Homework/quizzes (30 points). You can anticipate brief quizzes on an intermittent basis throughout the term.

Two term projects (40 points). Learning is social. The term projects will be collaborative, data-based projects, which you will undertake with two to four of your peers and which you will submit as fully-contained R markdown documents, that is, as reproducible texts which include your argument, commented code, and the results of statistically appropriate analyses. These results will typically include tables and/or figures (data visualizations). The project will be empirical, typically from data that I provide you with or that we find together. Although these datasets may appear “large” (e.g., more than one million data points), you will be able to investigate them on your laptops in R.

In order for us to assess your individual contributions and to minimize social loafing, I ask that all meetings and communications among group members be undertaken on the [Slack](https://faudatascience.slack.com/) platform, and that, in addition to the paper, all group members sign a 1-page cover page describing the primary contribution and percent effort of each person. You'll present your projects in class; the first near the middle of the term, the second towards the end of the term.

A final exam (10 points). This will include questions about the technical and ethical challenges in working with social data.

Above and beyond (10 points). You'll have the opportunity to earn an additional 10 points by solving one or more data challenges that we will develop as the class goes forward.

|  |
| --- |
| course grading scale |
| A 93 or more | C 73 to 76 | *note that in borderline cases, students may receive the higher of two grades if there is evidence of sustained effort and/or improvement over the course of the term*  |
| A- 90 to 92 | C- 70 to 72 |
| B+ 87 to 89 | D+ 67 to 69 |
| B 83 to 86 | D 63 to 66 |
| B- 80 to 82 | D- 60 to 62 |
| C+ 77 to 79 | F 59 or less |

# classroom etiquette policy

In order to enhance and maintain a productive atmosphere for education, personal communication devices, such as cell phones, are to be disabled in class sessions. Laptops and tablets should be used only for appropriate purposes.

# attendance policy

As noted above, attendance is expected and will contribute to the portion of grades assessed as “class participation.”

# policy on accommodations

In compliance with the Americans with Disabilities Act Amendments Act (ADAAA), students who require reasonable accommodations due to a disability to properly execute coursework must register with Student Accessibility Services (SAS) and follow all SAS procedures. SAS has offices across three of FAU’s campuses – Boca Raton, Davie and Jupiter – however disability services are available for students on all campuses. For more information, please visit the SAS website at [www.fau.edu/sas/](http://www.fau.edu/sas/).

# policy on late work, makeup tests, and incompletes

late work will not be accepted. Make-up examinations will be provided in the case of medical excuse or dire need. Incomplete grades will not be granted unless judged absolutely necessary.

# code of academic integrity policy statement

Students at Florida Atlantic University are expected to maintain the highest ethical standards. Academic dishonesty is considered a serious breach of these ethical standards, because it interferes with the university mission to provide a high quality education in which no student enjoys an unfair advantage over any other. Academic dishonesty is also destructive of the university community, which is grounded in a system of mutual trust and places high value on personal integrity and individual responsibility. Harsh penalties are associated with academic dishonesty. For more information, see University Regulation 4.001.

# policy recognizing that we are human (counseling and psychological services)

Life as a university student can be challenging physically, mentally and emotionally. Students who find stress negatively affecting their ability to achieve academic or personal goals may wish to consider utilizing FAU’s Counseling and Psychological Services (CAPS) Center. CAPS provides FAU students a range of services – individual counseling, support meetings, and psychiatric services, to name a few – offered to help improve and maintain emotional well-being. For more information, go to <http://www.fau.edu/counseling/>

# course outline / schedule

**With the exception of scheduled exam and paper due dates, all content is subject to change.** Please monitor the Slack channel for the latest updates (<https://honorscollege.slack.com/>).

|  |  |  |  |
| --- | --- | --- | --- |
| Week | Meeting type\* | Topic | Assignment |
| 1 | Introduction | Course introduction, refresheR | Discussion of Lazer et al (2009), Wickham (2017) |
| 2 | Discussion | Ethics and privacy | Tufecki (2014), Wang & Kosinski (2018) |
| 3 | Discussion | Networks: Basic concepts including distance, centrality, weak ties, power law | Easley & Kleinberg (2010) Chapters 1-5.Borgatti (2009) |
| 4 | Discussion | Empirical studies of networks; contagion | Christakis, N. A., & Fowler, J. H. (2010). Nishi et al., (2015). Eagle et al., (2009)Apicella et al. (2012) |
| 5 | Laboratory | Network analysis and visualization using GEPHI | Bastian et al., (2009). |
| 6 | Laboratory | Network analysis in R with tidygraph | See http://bit.ly/tidygraphTalk  |
| 7 | Laboratory | Community structure; bibliometric networks | Palla et al., (2005)Lanning (2017) |
| 8 | Discussion | Collaboration and crowdsourcing | Salgarnik, Chapter 5 |
| 9 |  | Presentations | --- |
| 10 | Discussion | Text analysis: From word clouds to the LIWC | Pennebaker et al. (2015)Lanning (2018) |
| 11 | Discussion | n-grams in social media | Schwartz et al (2013) |
| 12 | Laboratory | Preparing text for analysis in R | R packages: Quanteda, tidytext, wordcloud |
| 13 | Laboratory | Text analysis: Inferring meaning using computational methods; sentiment analysis | James et al (2013) Chapters 1 and 2 |
| 14  | Presentations |  | --- |
| 15 | Review | Ethics revisited; review | Salgarnik, Chapters 6 and 7 |

\* Nearly every class will include lecture, discussion, and laboratory (hands-on experience programming). The “meeting type” describes the anticipated primary focus for the week.

references

Apicella, C. L., Marlowe, F. W., Fowler, J. H., & Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. Nature, 481(7382), 497.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. Icwsm, 8, 361-362.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. science, 323(5916), 892-895.

Christakis, N. A., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks. PloS one, 5(9), e12948.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... & Jebara, T. (2009). Life in the network: the coming age of computational social science. Science (New York, NY), 323(5915), 721.

Eagle, Nathan, Alex Sandy Pentland, and David Lazer. "Inferring friendship network structure by using mobile phone data." Proceedings of the national academy of sciences 106.36 (2009): 15274-15278.

Easley, D., & Kleinberg, J. (2010). Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press. (<http://www.cs.cornell.edu/home/kleinber/networks-book/>).

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning. New York: Springer. (<http://www-bcf.usc.edu/~gareth/ISL/>).

Lanning, K. (2017). What is the Relationship Between “Personality” and “Social” Psychologies? Network, Community, and Whole Text Analyses of The Structure of Contemporary Scholarship. Collabra: Psychology, 3(1).

Lanning, K., Pauletti, R. E., King, L. A., & McAdams, D. P. (2018). Personality development through natural language. Nature Human Behaviour, 1.

Nishi, A., Shirado, H., Rand, D. G., & Christakis, N. A. (2015). Inequality and visibility of wealth in experimental social networks. Nature, 526(7573), 426.

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. nature, 435(7043), 814.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015.

Salgarnik (2017). Bit by Bit: Social Research in the Digital Age. (<http://www.bitbybitbook.com/en>).

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. PloS one, 8(9), e73791.

Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. ICWSM, 14, 505-514.

Wickham, H., & Grolemund, G. (2016). R for data science: import, tidy, transform, visualize, and model data. " O'Reilly Media, Inc."