| **Quantitative Research Methods**  *Updated March 1 7, 2020* | |
| --- | --- |
| Course: GSR - 71100-01  Hunter College, Spring 2020  Wednesdays - 5:35PM-7:25PM via Zoom | Instructor: Nga Than  Office Hours: Thursdays at 5:30PM  Email: [nthan@gradcenter.cuny.edu](mailto:nthan@gradcenter.cuny.edu) |

**COURSE DESCRIPTION**

This course will introduce you to basic concepts and methods used in survey research, research design, measurement, questionnaire construction, sampling, and data analysis. Students will learn classical machine learning methods such as k-means clustering, principal components analysis, random forests, support vector machines, and neural networks. Students will also design and collect their original survey data through Amazon Mechanical Turk or Facebook Ads Manager and use this data to write a research paper. The course will also introduce students to the R-programming language, LaTex and Git/Github.

**LEARNING OUTCOMES**

* Familiarity with survey data, cleaning survey data, creating a survey, and creating visualizations
* Familiarity with creating surveys through Qualtrics and Pol.is
* Familiarity with Amazon Mechanical Turk and Facebook Ads Manager
* Familiarity with packages Tidyverse, ggplot2, caret in R, understanding how Git and Github work
* Research and writing about a topic of interest.

**REQUIRED BOOKS:**

*Wickham, H., & Grolemund, G. (2016). R for data science: import, tidy, transform, visualize, and model data. " O'Reilly Media, Inc.".* [*https://r4ds.had.co.nz/*](https://r4ds.had.co.nz/)

**Extra Texts, Data and Assignments*:*** Blackboard

**Exercises:** DataCamp

**COURSE REQUIREMENTS**

**Grades:**

| **Item** | **Date** | **Weight (%)** |
| --- | --- | --- |
| Homework Assignments | Varies | 30.0 |
| Mid-term | Oct 20 | 30.0 |
| Data & Research Proposal | Nov 30 | 10.0 |
| Final Research Paper | May 15 | 30.0 |

The grading system in this class is as follows:

| A+ | 97-100 |
| --- | --- |
| A | 93-96 |
| A- | 90-92 |
| B+ | 87-89 |
| B | 83-86 |
| B- | 80-82 |
| C+ | 76-79 |
| C | 74-75 |
| C- | 70-73 |
| D | 60-69 |
| F | <=59 |

***Attendance/class participation***

Regular attendance is one of the most important parameters to successful completion of the course requirements. Even though I will not take attendance, I expect that you read required readings before class, and participate in classroom discussions.

***Research Hypotheses/Theories***

One page research proposal which outlines a research question and hypotheses. If the data already exists, the statement should provide the data source. If the data does not exist, the essay should describe how the data could be collected (through conducting a survey, online experiments, or digital trace data).

**Survey Design**

Use either Google Form or Qualtrics or [Pol.is](http://pol.is/) to examine the above stated research question. What kinds of questions do you want to ask? Collect 10-15 responses for the pilot phase, and write a 3-page report on the findings from these responses. What challenges do you face while collecting survey responses? If the survey were to be launched, what else would you include, or change?

***Mturk Exercise/ Facebook Ads Manager***

Use either Amazon Mechanical Turk or Facebook Ads Manager to collect real-world data for either Survey No. 1 or Survey No. 2. The number of responses should be at least between 150 to 250 responses.

**3 Assignments**

Complete 3 assignments to demonstrate that you understand concepts, and how build models, and how to evaluate the model results

**Final Research project -> Project Reflection**

From the Mturk and Facebook Ads Manager datasets, you can write your final research project. The tasks in these exercises include cleaning data, getting it in shape, and choosing and fine-tuning the right models for your final research project.

***Privacy***

Students who participate in this class with their camera on or use a profile image are agreeing to have their video or image recorded solely for the purpose of creating a record for students enrolled in the class to refer to, including those enrolled students who are unable to attend live. If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate orally are agreeing to have their voices recorded. If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live.

***Hunter College Policy on Academic Integrity***

Hunter College regards acts of academic dishonesty (e.g., plagiarism, cheating on examinations, obtaining unfair advantage, and falsification of records and official documents) as serious offenses against the values of intellectual honesty. The College is committed to enforcing the CUNY Policy on Academic Integrity and will pursue cases of academic dishonesty according to the Hunter College Academic.

***Hunter College Policy on Sexual Misconduct***

In compliance with the CUNY Policy on Sexual Misconduct, Hunter College reaffirms the prohibition of any sexual misconduct, which includes sexual violence, sexual harassment, and gender-based harassment retaliation against students, employees, or visitors, as well as certain intimate relationships. Students who have experienced any form of sexual violence on or off campus (including CUNY-sponsored trips and events) are entitled to the rights outlined in the Bill of Rights for Hunter College. a. Sexual Violence: Students are strongly encouraged to immediately report the incident by calling 911, contacting NYPD Special Victims Division Hotline (646-610-7272) or their local police precinct, or contacting the College's Public Safety Office (212-772-4444). b. All Other Forms of Sexual Misconduct: Students are also encouraged to contact the College's Title IX Campus Coordinator, Dean John Rose (jtrose@hunter.cuny.edu or 212-650-3262) or Colleen Barry (colleen.barry@hunter.cuny.edu or 212-772-4534) and seek complimentary services through the Counseling and Wellness Services Office, Hunter East 1123. CUNY Policy on Sexual Misconduct Link: <https://www.cuny.edu/wp-content/uploads/si>

***American Disability Act Policy***

In compliance with the American Disability Act of 1990 (ADA) and with Section 504 of the Rehabilitation Act of 1973, Hunter College is committed to ensuring educational parity and accommodations for all students with documented disabilities and/or medical conditions. It is recommended that all students with documented disabilities (Emotional, Medical, Physical, and/or Learning) consult the Office of AccessABILITY, located in Room E1214B, to secure necessary academic accommodations. For further information and assistance, please call: (212) 772-4857 or (212) 650-3230.

**TENTATIVE COURSE SCHEDULE**

**Feb 03:** **Introduction – Why study quantitative methods?**

Overview of DataCamp, Github, R, R-studio, Pol.is, Amazon Mechanical Turk

Create accounts for Github, Pol.is

HW: Create the introduction README.md file for Github, Installing necessary software (R, Rstudio)

**Feb 10: Data Ethics/ Research Ethics; R - Importing data, Transforming Data, and Cleaning Data**

Guest Speaker: [Kasia Chmielinski](https://cyber.harvard.edu/people/kasia-chmielinski) ( [The Data Nutrition Project](https://datanutrition.org/))

Salganik, M. J. (2019). *Bit by bit: Social research in the digital age*. Princeton University Press. [Chapter 6: Ethics](https://www.bitbybitbook.com/en/1st-ed/ethics/)

Breiman, Leo. 2001. [“Statistical Modeling: The Two Cultures.”](https://projecteuclid.org/journals/statistical-science/volume-16/issue-3/Statistical-Modeling--The-Two-Cultures-with-comments-and-a/10.1214/ss/1009213726.full) Statistical Science 16 (3): 199-231. 2

Lab: Introduction to R, Importing data, transforming data, and cleaning data in R

HW: DataCamp ([Introduction to R)](https://learn.datacamp.com/courses/free-introduction-to-r)

**Feb 17: Unsupervised Learning - K-means clustering**

Molina, M., & Garip, F. (2019). [Machine learning for sociology](https://osf.io/a6r9g/download). *Annual Review of Sociology*.

Garip, F. (2012). [Discovering diverse mechanisms of migration: The Mexico–US Stream 1970–2000](https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1728-4457.2012.00510.x?casa_token=cu4m036gw9YAAAAA:0aaBcmvlquF14uK219Fcb5Mc6z5Ubz9TKdnWA9Qv2yuz1br4m5cdP3DcySOFKxM4K9rCAC1FHbNcXA). *Population and Development Review*, *38*(3), 393-433.

* Pre-publication version: <https://projects.iq.harvard.edu/files/wcfia/files/clustering_manuscript_full.pdf>

HW: DataCamp ([Cluster Analysis in R](https://learn.datacamp.com/courses/cluster-analysis-in-r)), Assignment 1 (BB)

**Feb 24: Unsupervised Learning - Principal Component Analysis**

Reynolds, J., & Xian, H. (2014). Perceptions of meritocracy in the land of opportunity. *Research in Social Stratification and Mobility*, *36*, 121-137.

HW: DataCamp ([Dimensionality Reduction Analysis](https://learn.datacamp.com/courses/dimensionality-reduction-in-r)), Assignment 2

**March 03: Supervised Learning - Random Forests**

Rodriguez, M. Y., DePanfilis, D., & Lanier, P. (2019). Bridging the gap: Social work insights for ethical algorithmic decision-making in human services. *IBM Journal of Research and Development*, *63*(4/5), 8-1.

HW: Data Camp ([Supervised Learning with caret](https://learn.datacamp.com/courses/machine-learning-with-caret-in-r))

**March – 10: Supervised Learning - Neural Networks**

Davidson, Thomas. "[Black-box models and sociological explanations: Predicting high school grade point average using neural networks](https://journals.sagepub.com/doi/pdf/10.1177/2378023118817702)." *Socius* 5 (2019): 2378023118817702.

HW: Data Camp ([Hyperparameter tuning in R](https://learn.datacamp.com/courses/hyperparameter-tuning-in-r), [Neural Nets with Tensorflow](https://learn.datacamp.com/courses/introduction-to-tensorflow-in-r)), Assignment 3

**March 17: Survey Design (Qualtrics, Google Form and Pol.is)**

Guest Speaker:[Darshana Narayanan](https://www.linkedin.com/in/darshananarayanan/) (Pol.is)

Salganik, M. J., & Levy, K. E. (2015). [Wiki surveys: Open and quantifiable social data collection](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0123483). *PloS one*, *10*(5), e0123483.

Survey Design using Google form and Qualtrics

HW: Design a 20 question survey in either Pol.is, Google Form or Qualtrics to launch on MTurk and Facebook Ads

**March 24: Amazon Mechanical Turk**

Guest Speaker: [Siqi Tu](https://siqitu.com/), [Max Planck Institute for the Study of Religious and Ethnic Diversity](https://www.mmg.mpg.de/person/107747/2553)

Elliott-Negri, L., Tu, S., Zheng, W., & Lennon, M. C. (2021). [Hope, Emotional Charges, and Online Action: An Experimental Study of the DREAM Act](https://academic.oup.com/socpro/advance-article/doi/10.1093/socpro/spaa076/6086023?casa_token=xhh0hvnoqjsAAAAA:A6Qt3040BIOuJXfdHqtwy9fkfEZbQ0vt6oRt-qQbpH-Rv9JDLEimfZhSClZxOBenjB10mIAbgWar). *Social Problems*.

Schwemmer, C., Knight, C., Bello-Pardo, E. D., Oklobdzija, S., Schoonvelde, M., & Lockhart, J. W. (2020). [Diagnosing gender bias in image recognition systems](https://journals.sagepub.com/doi/pdf/10.1177/2378023120967171). Socius, 6, 2378023120967171.

Gray, M. L., & Suri, S. (2019). [*Ghost work: How to stop Silicon Valley from building a new global underclass*](https://ghostwork.info/). Eamon Dolan Books. (optional)

***In class exercise:*** Creating a Survey on MTurk (Not launch yet)

**March 31 - Spring Break**

**April 7: Facebook Ads Manager**

*Facebook Ads Manager:*

Schneider, D., & Harknett, K. (2019). [What’s to like? Facebook as a tool for survey data collection.](https://journals.sagepub.com/doi/pdf/10.1177/0049124119882477)  *Sociological Methods & Research*, 0049124119882477.

Griesbach, K., Reich, A., Elliott-Negri, L., & Milkman, R. (2019). Algorithmic control in platform food delivery work. *Socius*, *5*, 2378023119870041.

Milkman, R., Elliott-Negri, L., Griesbach, K., & Reich, A. (2020). Gender, Class, and the Gig Economy: The Case of Platform-Based Food Delivery. *Critical Sociology*, 0896920520949631.

***In class exercise:*** Creating a survey on Facebook Ads Manager;

Launch your MTurk Experiment & Facebook Ads Manager Experiment

**April 14: Importance of Piloting surveys - ggplot2 - Data Visualization Workshop**

**April - 21: Survey Research Clinic/ Support Vector Machine Workshop**

Guest Speaker: [Noam Segal](https://www.linkedin.com/in/noamsegal), UX Manager - Wealthfront

**April - 28: Survey Research Clinic**

Guest Speaker: Ameen Jauhar (tentative)

**May - 5: Presentations**

* RJ
* Cara and Alyx, Sustainability and Green Space in NYC Housing

**May - 12: Presentations**

* Madison, Kana, Anne, Katherine “News Consumption, Political Discussion, and Religion during COVID-19”
* Jacob, Amanda and Colleen, “Essential” Food Workers in NYC

**Deadlines:**

**May - 15: Facebook ads receipts (due)**

**May - 20: Final paper and Project reflection**

**Automatic Extension - May 27th:**

**Recommended DataCamp Exercises**

1. [Introduction to R](https://learn.datacamp.com/courses/free-introduction-to-r)
2. [Intermediate R](https://learn.datacamp.com/courses/intermediate-r)
3. [Introduction to Importing in R](https://learn.datacamp.com/courses/introduction-to-importing-data-in-r)
4. [Intermediate Importing in R](https://learn.datacamp.com/courses/intermediate-importing-data-in-r)
5. [Cleaning Data in R](https://learn.datacamp.com/courses/cleaning-data-in-r)
6. [Tidyverse R](https://learn.datacamp.com/courses/working-with-data-in-the-tidyverse)
7. [Exploratory Data Analysis](https://learn.datacamp.com/courses/case-study-exploratory-data-analysis-in-r)
8. [Correlation and Regression in R](https://learn.datacamp.com/courses/correlation-and-regression-in-r)
9. [Logistic Regression](https://learn.datacamp.com/courses/multiple-and-logistic-regression-in-r)
10. [Unsupervised Machine Learning](https://learn.datacamp.com/courses/unsupervised-learning-in-r)
11. [Supervised Machine Learning Classification](https://learn.datacamp.com/courses/supervised-learning-in-r-classification)
12. [Supervised Machine Learning: regression](https://learn.datacamp.com/courses/supervised-learning-in-r-regression)
13. [Intro to Visualization in R (ggplot2)](https://learn.datacamp.com/courses/introduction-to-data-visualization-with-ggplot2)
14. [Intermediate Data Visualization](https://learn.datacamp.com/courses/intermediate-data-visualization-with-ggplot2)