

Wayne State University - Department of Economics
ECO 5100 - Introductory Statistics and Econometrics (Winter 2020)

Instructor: Vitor Kamada

Class: TT, 2:30 - 04:10 pm in Old Main, Room 1145

Office: 2139 Faculty Administration Building (FAB)

Send all e-mails to: econometrics.methods@gmail.com

Cel: 678 644 5511

Office hours: MW, 2:00 – 2:20 pm and 4:10 – 5:20 pm in my office, or by appointment.

1) Course Description

Econometrics is a set of methods that combines Statistics and Computer Science to solve economic problems. Data Science is just a modern name for this fusion among Statistics and Computer Science. In this sense, Econometrics is just Data Science applied to economic problems. Google's Chief Economist Hal Varian keeps saying the sexy job in the next ten years will be Statistician (Data Scientist). Modern society generates huge amount of data. The scarce factor is the skill to understand data and extract value from it.

This course will introduce the main tools from Data Science frequently used in economic consulting, big corporations, and academic sector. An important emphasis is put on practical application and on the use of Python Computer Language to analyze real-world datasets. In short, this course is based on three pillars: inferential thinking, computational thinking, and real-world relevance.

2) Learning Outcomes

This course teaches computer programming and statistical inference together with hands-on analysis of real-world datasets. The main goal of this course is to develop statistical reasoning and coding skills. Statistical reasoning involves understanding the logic behind the statistical procedures and being able to fully interpret the results. Coding skills involves converting symbolic Math and Statistical concepts in computer language, that is, the implementation of the statistical analysis.

After this course students will become proficient in using Python to access, process, visualize, understand, and extract value from data. Students are expected to produce statistical analysis and not only being consumers of statistical studies or results.

3) Required Free Textbooks

Adhikari, A., DeNero, J. (2020). Computational and Inferential Thinking: The Foundations of Data Science. Available for free at: <https://www.inferentialthinking.com/chapters/intro.html>

James, G., Witten, D., Hastie, T., Tibshirani, R. (2017). An Introduction to Statistical Learning with Applications in R. Springer. Available for free in the Author website: <http://faculty.marshall.usc.edu/gareth-james/ISL/>

Siegel, E. (2016). Predictive analytics: the power to predict who will click, buy, lie, or die. Wiley. Available for free via WSU library at: <https://elibrary.wayne.edu/record=b5588358~S47>

4) Required Free Software

4.1) Python

Python is an open-source programming language. It tends to be the dominant language in many branches of Data Science, such as: Machine Learning, Deep Learning, Natural Language Processing, Network Analysis, deployment of Big Data infrastructure; etc.

There are several ways to run Python Code. I will use the Google Colab, a free Jupyter notebook environment that runs entirely in the cloud.

If you don't have a Google Account, you will need to create one, before accessing Google Colab at: <https://colab.research.google.com/>

5) Course Schedule

Part I –Data Science based on Adhikari & DeNero (2020)

Date	Topics/Key Concepts	Chapters
Week 1 Jan 7	1) Data Science Python 3, Data Visualization, Plotting the Classics.	1
Week 1 Jan 9	2) Causality and Experiments Observational Study, Confounding, Randomization.	2
Week 2 Jan 14	3.1) Programming in Python 3.2) Data Types 3.3) Sequences	3, 4, and 5
Week 2 Jan 16	4) Tables Sorting and Selecting Rows, Population Trends.	6

Week 3 Jan 21	5) Data Visualization Scatter Plots, Line Graphs, Bar Chart, Histogram.	7
Week 3 Jan 23	6) Functions and Tables Syntax, Signature, Documentation, Body, Indentation, Arguments Methods, Joining Tables by Columns.	8
Week 4 Jan 28	7) Bike Sharing in the Bay Area Drawing Maps, OpenStreetMap, Latitude, Longitude.	8.5
Week 4 Jan 30	8) Conditional Statements, Iteration, and Simulation Randomness, “if”, “elif”, “for”, The Monty Hall Problem.	9
Week 5 Feb 4	9) Sampling and Empirical Distributions Deterministic and Probability Samples, Random Sampling, Population, Statistic.	10
Week 5 Feb 6	10) Testing Hypotheses Assessing Models, Predicting the Statistic, Multiple Categories.	11
Week 6 Feb 11	11.1) A/B Testing 11.2) Causality	12
Week 6 Feb 13	12.1) Bootstrap 12.2) Confidence Intervals	13
Week 7 Feb 18	13) Why the Mean Matters Standard Units, Normal Curve, Central Limit Theorem, Variability of the Sample Mean	14
Week 7 Feb 20	14.1) Correlation 14.2) The Regression Line	15
Week 8 Feb 25	15.1) Least Squares Regression 15.2) Visual Diagnostics	15
Week 8 Feb 27	16) K-Nearest Neighbors Classification, Training and Test data, Decision Boundary.	17
Week 9 Mar 3	17.1) Implementing the Classifier 17.2) Multiple Linear Regression	17
Week 9 Mar 5	Midterm	All
Week 10 Mar 10	Holiday - No Classes (Spring Break)	
Week 10 Mar 12	Holiday - No Classes (Spring Break)	

Part II - Statistical Learning based on James et al. (2017)

Week 11 Mar 17	18) Multiple Linear Regression Non-linear Relationships, Qualitative Predictors.	3.1, 3.2, 3.3
Week 11 Mar 19	19) Logistic Regression Odds, Maximum Likelihood, Making Predictions.	4.3

Week 12 Mar 24	20) Decision Trees Classification and Regression Tree, Recursive Binary Splitting, Tree Pruning, Gini Index, Entropy.	8.1
Week 12 Mar 26	21) Bagging, Random Forests, Boosting Bootstrap, Out-of-Bag, Variable Importance.	8.2
Week 13 Mar 31	22) Principal Components Analysis Total Variance, Proportion of Variance Explained.	10.2
Week 13 Apr 2	23) K-Means Clustering, and Hierarchical Clustering Dendrogram, Linkage, Dissimilarity Measure.	10.3

Part III – Student Presentations based on Siegel (2016)

Week 14 Apr 7	Chapter 2: With Power Comes Responsibility: Hewlett-Packard, Target, the Cops, and the NSA Deduce Your Secrets
	Chapter 3: The Data Effect: A Glut at the End of the Rainbow
Week 14 Apr 9	Chapter 4: The Machine That Learns: A Look inside Chase's Prediction of Mortgage Risk
	Chapter 5: The Ensemble Effect: Netflix, Crowdsourcing, and Supercharging Prediction
Week 15 Apr 14	Chapter 6: Watson and the Jeopardy! Challenge
	Chapter 7: Persuasion by the Numbers: How Telenor, U.S. Bank, and the Obama Campaign Engineered Influence

Part IV - Final Exam

Week 15 Apr 16	Draft: Empirical Report No lecture, but I will be available in my office for feedback.
Week 16 Apr 21	Study Day
Week 16 Apr 23	Final Students can decide between Empirical Report or Final Exam

6) Grading

6.1) Your final grade will be assessed as follows:

Assignment*	Composition**	Weight	Date
Surveys*	Individual	1%	Thursday, Feb 6 (at 2:30 pm)
Quizzes	Individual	20%	Check on Canvas
Labs	Group	39%	Check on Canvas

Midterm	Individual	15%	Thursday, Mar 5 (at 2:30 pm)
Presentation	Group	10%	Apr 7, Apr 9, and Apr 14
Final	Individual	15%	Thursday, Apr 23 (at 2:30 pm)
Total		100%	

*** Guidelines and detailed instructions for Assignments are on Canvas**

**** For Group Assignment, you must learn how to work in team, communicate properly, and negotiate with others. If you work alone, I will not accept the assignment and your grade will be automatically 0.**

Grading Scale

94+ = A	74+ = C
90+ = A-	70+ = C-
87+ = B+	67+ = D+
84+ = B	64+ = D
80+ = B-	61+ = D-
77+ = C+	Below 61 = F

6.2) Makeup Policy for any Assignment

If you miss any Assignment, I will provide a makeup activity in the case of an excused and unavoidable absence. Then it is YOUR RESPONSIBILITY to provide satisfactory written documentation of an excused and unavoidable absence as soon as possible. For example, if you are ill – the accompanying doctor's note must say that you cannot (or could not) do the Exam or Lab. If the doctor's note does not state this clearly, your score will be zero.

7. Course Expectations

7.1) Clarifying Expectations

To succeed in this course, you'll need to invest a good amount of time and energy doing exercises outside the class time. If at any time you feel you're investing the required time and energy but aren't learning the material or improving your skills, contact me and I'll do my best to help you and to suggest additional resources and options. If you have questions or concerns that you believe can be handled via e-mail, feel free to contact me that way. If I cannot adequately respond to your question via e-mail, I'll ask you to come to my regular office hours or make an appointment.

7.2) Academic Integrity

Wayne State University aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Student Code of Conduct, please see <https://doso.wayne.edu/conduct/codeofconduct.pdf>. Students who commit or assist in committing dishonest acts are subject to sanctions described in the Student Code of Conduct.

7.3) Special Accommodations

If you have a documented disability that requires accommodations, you will need to register with Student Disability Services (SDS) for coordination of your academic accommodations. The Student Disability Services (SDS) office is located at 1600 David Adamany Undergraduate Library in the Student Academic Success Services department. SDS telephone number is 313-577-1851 or 313-577-3365 (TDD only). Once you have your accommodations in place, I will be glad to meet with you privately during my office hours to discuss your special needs. Student Disability Services' mission is to assist the university in creating an accessible community where students with disabilities have an equal opportunity to fully participate in their educational experience at Wayne State University.