



# MSc in Analytics

## Course Syllabus

---

### **INTRODUCTION TO STATISTICAL CONCEPTS**

MSCA 31000 – 1:30 – 4:30 PM

Anil

Chaturvedi

[anilchaturvedi@uchicago.edu](mailto:anilchaturvedi@uchicago.edu)

Session 1: Section 2: March 5th

Session 2: Section 2: March 12th

Session 3: Section 2: March 19th

Session 4: Section 2: March 26th

Session 5: Section 2 9 am -12 pm March 28th

Check the classroom assignment at the 28<sup>th</sup> floor of NBC Tower Lobby on the first day of the class.

This course provides general exposure to basic statistical concepts that are necessary for students to understand the content presented in more advanced courses in the program. The course covers theoretical distributions and the way these distributions are used to assign probabilities to events in some depth. The course also introduces students to descriptive statistical methods to explore and summarize data, methodologies for sampling units for measurement or analysis, drawing inferences on the basis of knowledge gained from samples to populations, assessing relationships between variables, and making predictions based upon relationships between variables. This 50-credit course is either required or strongly recommended for all students entering the MSc in Analytics program. Letter grades will be assigned for this course and the course will appear on your official transcript. Students must obtain a passing grade of at least B-.

## **BOOK**

David Lane (Editor). *Introduction to Statistics*.(Version 2.0). Free online book at <http://onlinestatbook.com/>.

Developed by Rice University, University of Houston Clear Lake, and Tufts University.

You must **use the web version**, do not use the PDF version for the homework or readings.

**Internet Explorer still supports Java Applets.**

Use the mobile version of the book if you are using an iPad.

## **SOFTWARE**

R and related packages (<http://cran.r-project.org>)

## LEARNING OBJECTIVES

After completing this course, students should be able to:

- Design studies to collect valid and reliable data.
- Produce and interpret graphical summaries of data.
- Produce and interpret numerical summary statistics.
- Understand and use theoretical distributions to assign probabilities to events.
- Graphically and numerically describe the relations between quantitative variables.
- Infer properties of a population from a sample.
- Predict values of an outcome variable by using a set of indicator variables.

## EVALUATION:

Your course grade will be calculated as follows:

- 80% Weekly Assignments
- 20% Class Participation
- 5 points bonus - Final Project (optional)

## GRADING SCALE

A = 94%–100%

A- = 90%–93%

B+ = 87%–89%

B = 84%–86%

B- = 80%–83%

C+ = 77%–79%

C = 74%–76%

C- = 70%–73%

F = 0%–69%

## ATTENDANCE

This course will meet five times. All course goals, session learning objectives, and assessments are supported through in class lecture, activities, and discussions. Your attendance is thus required and paramount to your success in this course. You are allowed to miss at most one session, provided that you make arrangements with the instructor in advance.

You must sign in on the sheet.

## FINAL PROJECT (Due March 28th)

Final project is a three to five pages (excluding appendices) written report summarizing the results of an analysis of a data set presented in the case study “Angry Moods.” The case study is about controlling anger and can be found at the end of the book.

You will analyze the Angry Moods data set throughout the course in several steps. The final project will put together all of your findings from earlier sessions with the findings from additional analyses you will complete at the end of the course. Therefore, it is important that you start writing the final project as you obtain results in earlier sessions. The goal of the project is to develop a coherent and concise story or knowledge on how different people deal with anger. It is important that your story is substantiated on the basis of results of your data analyses. Put your detailed data analyses results in an appendix and submit together with the written report.

In addition to a brief executive summary of the entire project at the beginning, the written report must have the following sections.

- Background: Brief description of the study, the data set, and the general research question (10 points).
- Hypotheses: Specific questions that you would like to answer by analyzing the data set (20 points).
- Findings: Summary of your results (30 points).
- Discussion: Your interpretation of the findings, e.g., overall story of how different people deal with anger (30 points).
- Conclusion: The (possibly actionable) insight(s) you would like your readers to take away (5 points).
- References: List of sources you have consulted.

The final project will be graded out of 100 points. You may use R or any other statistical package you feel comfortable with to analyze the data. Yet, it is recommended that you use R since relevant syntax for the analysis required in each session will be provided. Make sure that your writing is clear and free of grammatical and spelling mistakes and that you use APA format for reporting statistical results and citations (5 points). An example project on a different topic is posted on the Canvas site for the course.

## WEEKLY ASSIGNMENTS

There are five set of weekly assignments. Assignments 1 to 3 have two parts. You will do the part a before the session after reading the respective chapters and submit it on the due date at the Canvas site for the course. The part b is due on the following session and will also be submitted to the Canvas site. For instance, you will submit assignment 1a on the first session but assignment 1b on the second session. Please name your assignment as name of the assignment+ your name. For instance, the name for assignment 1a might be: Assignment1a\_Barlas. Each assignment is about answering the questions listed in the course schedule for the respective session. Whenever necessary, R syntax will be provided. Most assignments have questions relevant for the final project. Your answers to these questions will be graded as part of the assignment, not the final project.

Note that you do not have to type your answers; yet, you are responsible for making yourself understood.

## LATE WORK

All assignments must be submitted to the Chalk site for the course on the due date before 11:59 pm. If you turn in an assignment late, 10% credit will be deducted from the total score for each day after the deadline. Assignments turned in more than one week late will not receive credit. In the case of unexpected events, you

must contact the instructor before the assignment due date in order to receive a grace period. Students can only receive up to two grace periods in the course.

## REQUESTING REASONABLE ACCOMMODATIONS

If you are interested in requesting disability accommodations, you may want to begin by reading through the information published on this website <https://disabilities.uchicago.edu/>. Also, please do communicate your requests as soon as possible to Gregory Moorehead, director of disability services, at 773.702.7776 or [gmoorehead@uchicago.edu](mailto:gmoorehead@uchicago.edu).

## ACADEMIC HONESTY & PLAGIARISM

It is contrary to justice, academic integrity, and to the spirit of intellectual inquiry to submit another's statements or ideas of work as one's own. To do so is plagiarism or cheating, offenses punishable under the University's disciplinary system. Because these offenses undercut the distinctive moral and intellectual character of the University, we take them very seriously.

Proper acknowledgment of another's ideas, whether by direct quotation or paraphrase, is expected. In particular, if any written or electronic source is consulted and material is used from that source, directly or indirectly, the source should be identified by author, title, and page number, or by website and date accessed. Any doubts about what constitutes "use" should be addressed to the instructor.

## COURSE SCHEDULE

**Important Note:** Changes may occur to the syllabus at the instructor's discretion. When changes are made, students will be notified via email and in-class announcement.

## SESSION 1

### Learning Objectives

After this session, the students should be able to understand, produce, and interpret:

- Distribution of data.
- Graphs.
- Measures of central tendency and dispersion.
- Correlation.

Reading: Chapter 1, 2, 3, and 4.

Assignment 1a (10 points, **due on March 5<sup>th</sup>, the first day of your class**): Chapter 1 to 4 – Odd numbered questions (questions from case studies are not included).

Assignment 1b (10 points, **due on March 12<sup>th</sup>**): Angry Moods Case - Questions 10 - 11, 17, 18, and 20 - 21.

## **SESSION 2**

### Learning Objectives

After this session, the students should be able to understand, evaluate, and/or implement:

- Concept of probability.
- Bayes theorem.
- Using theoretical distribution to assign probabilities to events.
- Sampling distributions.

Reading: Chapter 5, 7, and 9.

Assignment 2a (10 points, **due on March 12<sup>th</sup>**):

Chapter 5 – Odd numbered questions between 1 and 9.

Chapter 7 - Even numbered questions between 2 and 10.

Chapter 9 - Odd numbered questions between 1 and 21.

Assignment 2b (15 points, **due on March 19<sup>th</sup>**):

Chapter 5 – Odd numbered questions between 11 and 15.

Chapter 7 - Even numbered questions between 12 and 22.

Chapter 9 - Questions 23 – 25.

## **SESSION 3**

### Learning Objectives

After this session, the students should be able to understand, implement, and interpret results of:

- Statistical inference for single population.
- Statistical inference about two populations.
- Power.

Reading: Chapter 10, 11, 12, and 13. Assignment 3a (10 points, **due on March 19<sup>th</sup>**):

Chapter 10 - Questions 1, 5, 10, 15, and 20.

Chapter 11 - Questions 4, 8, 14, and 18.

Chapter 12 - Questions 8, 9, and 11.

Chapter 13 - Question 1 -3.

Assignment 3b (15 points, **due on March 26<sup>th</sup>**):

Chapter 10 - Questions 24 - 26.

Chapter 11 - Questions 20 - 24.

Chapter 12 - Questions 21 and 22.

Chapter 13 - Question 5.

## **SESSION 4**

### Learning Objectives

After this session, the students should be able to understand, implement, and interpret results of:

- Regression.
- Transformations.

Reading: Chapter 14 and 16.

Assignment 4 (15 points, **due on March 28<sup>th</sup>**):

Chapter 14 - Even numbered questions 2 – 18 and #19.

Chapter 16: Questions 1-5.

## **SESSION 5**

### Learning Objectives

After this session, the students should be able to understand, implement, and interpret results of:

- ANOVA between subject designs.
- ANOVA within subject designs.
- Effect size.

Reading: Chapter 15 and 19

Assignment 5 (15 points, **due on March 28**):

Chapter 15 - Odd numbered questions between 1 and 23.

Chapter 19 - Questions 3 - 6.

Angry Moods case – Question 19.