MBA 7098 – Statistics and Data Analysis, Fall 2014

Instructor: Ling-Chieh Kung
Department of Information Management
National Taiwan University

Statistics and data analysis are probably playing the most important roles in business analytics nowadays. With the ability to conduct scientific statistical studies and systematically analyze data, managers will be able to understand more about their customers, suppliers, competitors, and the business environment. The insights may then facilitate better decision making and help a company to attain competitive advantages. In this fundamental course in the Global MBA (GMBA) program, we will focus on the techniques for conducting basic statistical studies and data analysis. The hope is that students will be capable of doing scientific data analyses in their future GMBA courses and after graduations. Time will be spent on tools, applications, as well as theories. R, one of the most popular programming languages for data analysis, will be taught and used throughout this course. Students will learn how to write R programs to play with data. For at least part of this course, I plan to adopt the "flipped classroom" principle, which may be new to some students. Please pay attention to the syllabus to get an idea about the design of this course.

This is a required course offered in the GMBA program in National Taiwan University. The GMBA office does not allow non-GMBA students to take or audit this course.

Basic information

Instructor
- Ling-Chieh Kung (孔令傑). E-mail: lckung(AT)ntu.edu.tw.
- Office: Room 413, Management Building II. Tel: 02-3366-1176.
- Office hour: 5:30-6:00pm, Monday and 5:30-6:30pm, Wednesday.
- http://www.im.ntu.edu.tw/~lckung/

Teaching Assistants
- Ian Zhong (鍾冠宇). E-mail: r03725040(AT)ntu.edu.tw.
- Ho Ho (何禾). E-mail: r03725041(AT)ntu.edu.tw.

Lectures
- 2:20-5:20pm, Monday in E-Sun Hall, Management Building I.

References
- Ken Black, 2011, *Business Statistics: For Contemporary Decision Making* (7th edition).
- Steven Murray, *Learn R in a Day* (Amazon Kindle e-books only).
- Steven Levitt and Stephen Dubner, 2009, *Freakonomics* (revised and expanded edition).[1]
- Viktor Mayer-Schönberger and Kenneth Cukier, 2014, *Big Data*.[2]

On-line Resources
- For checking grades: CEIBA.
- For materials: http://www.im.ntu.edu.tw/~lckung/courses/SDA-Fa14/.
- For discussions: the bulletin board "NTUIM-lckung" on PTT.

Grading

Breakdown
- Class participation: 10%. Lecture problems: 20%.
- Three case reports: 15% (5% for each). One case presentation: 10%.
- Midterm Exam: 15%. Final project: 30%.

Conversion Rule
- The final letter grades will be given according to the following conversion rule:

Letter	Range	Letter	Range	Letter	Range	Letter	Range	Letter	Range
F	[0, 60)	C−	[60, 63)	C	[63, 67)	C+	[67, 70)	B−	[70, 73)
B	[73, 77)	B+	[77, 80)	A−	[80, 85)	A	[85, 90)	A+	[90, 100]

[1] Translated into Chinese with the book title "蘋果橘子經濟學".
[2] Translated into Chinese with the book title "大數據".

Course outline

There are four modules in this course: foundations, inferential statistics, advanced techniques, and applications. We will spend five, four, and five lectures in the first three modules to introduce basic theories and methods. Each module is then concluded with a case study. The last module includes a brief introduction to data mining and a guest speaker's talk about applying statistics and data analysis. Students' presentations for their final projects then conclude this course.

Policies

"Flipped Classroom"

- Before most Monday lectures, the instructor will upload videos containing some materials to be discussed on that Monday. The total length of those videos for one lecture will be around 60 to 90 minutes. Students must find their own time to watch the videos before the lecture.
- During lectures, we answer students' questions regarding materials in the videos, give examples, programming with R, and provide exercises for students to do on-site analyses, programming, and discussions.
- At the end of each regular lecture, the instructor will assign a ***lecture problem*** for students to solve. To earn points for lecture problems, students must work on the problems ***on-site*** and submit their answers ***by 5:20pm***.

Teams

- Students must form teams to do lecture problems, case studies, and the final project. One's teams for these three tasks need not to be identical.
  1. Lecture problems and case studies: Students will be randomly grouped into teams with around three students. A team will work together for lecture problems and the case study in a module. For different modules, one may have different teammates.
  2. Final project: Students form teams by themselves. We expect to have six or eight teams in total, where the number depends on the number of enrolled students.

Homework and Case Studies

- To give students more chances to do practices, five homework assignments will be given. However, thanks to in-class exercises and lecture problems, students ***do not*** need to submit homework! Solutions will be provided on the due dates.
- Three case studies will also be given in this semester. Students will be given either real stories or real data sets and related background information. They may then try to understand the story or analyze the data to discover insightful hidden messages. For a case study, each team submits a written report. Some teams will be selected for giving oral presentations. The instructor will make sure that each student participates in exactly one oral presentation.

Project

- Students will form teams to do a final project by applying the techniques learned in this course to a self-selected problem. Each team will make an oral presentation in one of the last two lectures and submit a report. The written report is due on the date the team makes the oral presentation. All team members must be in class for the team to present.

Class Participation

- We encourage class participation and include it in evaluating each student. During lecture time, students are more than welcome to ask or answer questions and provide comments. One gets good participation grades if her/his participation enhances the learning experiences of the class or she/he simply impresses the instructor with her/his passion and diligence.

Office Hour

- You are welcome to the instructor's office hour to ask him any question. You may ask him to clarify some concepts, give suggestions on case studies, or discuss the final project. Discussions not related to this course are also welcome. If you do not want to come in the designated time, feel free to send me an e-mail to schedule a meeting.

Midterm exam

- The midterm exam will be in-class and open whatever you have (including all kinds of electronic devices). However, no information is allowed to be transferred among students. Cheating will result in severe penalty. There is no final exam.

Tentative schedule

Week	Date	Lecture	Reading[3]	What is due?
1	9/15	Overview	Ch. 1	N/A
2	9/22	Descriptive statistics	Chs. 2 and 3	N/A
3	9/29	Probability distributions	Chs. 4, 5, and 6	(Homework 1)
4	10/6	***TA session:*** Advanced R programming	*Learn R in a Day*	N/A
5	10/13	***Case presentation 1***: Successful stories of data analysis	*Freakonomics*	Case report 1
6	10/20	Sampling and sampling distributions	Ch. 7	N/A
7	10/27	Estimations	Ch. 8	N/A
8	11/3	Hypothesis testing: Basic ideas	Chs. 9 and 16	(Homework 2)
9	11/10	***Midterm Exam***	N/A	N/A
10	11/17	Hypothesis testing: single population	Ch. 10	Case report 2
11	11/24	Hypothesis testing: multiple populations	Ch. 11	(Homework 3)
12	12/1	Regression: Simple regression	Ch. 12	N/A
13	12/8	Regression: Multiple regression	Chs. 13 and 14	(Homework 4)
14	12/15	***Case presentation 2***: Bike sharing and bank marketing	N/A	Case report 3
15	12/22	Correlation mining	*Big Data*	(Homework 5)
16	12/29	Guest talk and closing summary	N/A	N/A
17	1/5	***Project presentation 1***	N/A	Project report
18	1/12	***Project presentation 2***	N/A	Project report

[3] Chapter numbers are based on *Business Statistics: For Contemporary Decision Making* by Black. However, one may find related chapters in almost all introductory business statistics textbooks.