DEPAUL UNIVERSITY CDM
IT 223 DATA ANALYSIS (Fall/Autumn 2015)
SYLLABUS, CLASS SCHEDULE & POLICIES

COURSE: IT 223- Data Analysis **START AND END DATES:** September 9– November 24, 2015

LOCATION:

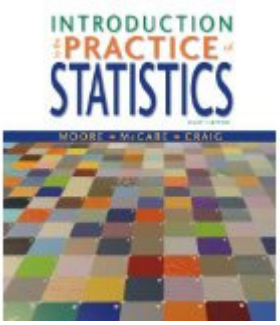
SECTION 402 (M/W 11:50 AM-1:20 AM)	In-Class session	Lewis Center Room 1510
SECTION 401 (M/W 1:30 PM-3 PM)	In-Class session	Lewis Center Room 1510

INSTRUCTOR:
RAY Partha Sarathy Email: rpartha1@cdm.depaul.edu
Office hours: 9:50 AM- 11:20 AM [Mondays & Wednesdays]
Office location: CDM 705 *(come to the CDM building seventh floor lobby and call the phone number provided above)*

- No appointment is necessary to meet instructor during office hours- feel free to visit.
- Voice mail is NOT available at this phone number at this point in time- do not leave me voice mails.
- Emails will be responded to within two business days- plan accordingly (most times I will respond sooner but be aware that it could take up to two business days to get a response).
- Please use the contact information provided above at all times. Information provided elsewhere may or may not have been updated.

COURSE MATERIALS:

Textbook (Required):

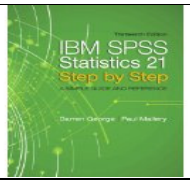


Moore, D.S., McCabe, G.P., & Craig, B. (2014). Introduction to the Practice of Statistics (8th Ed.). W.H. Freeman (McMillan) Publication.

Mode	ISBN-10	ISBN-13
Print/Paper Book (your bookstore, amazon.com, ebay.com etc.).	1464158932	978-1464158933
Other Options	eBook (Kindle)	
Rent book from chegg.com *(check other textbook rental websites as well to get the best deal)*	http://www.chegg.com	

- Access to the textbook website is required—I will assign homework assignment problems from the textbook. Data sets for some questions may be found on the textbook website.

Supplementary Readings/References-Optional [you are <u>not required</u> to buy these books, if you choose to buy them, you could buy older editions from e-bay or amazon---no need to buy the latest editions]

	George, D. & Mallery, P. (2013). IBM SPSS Statistics Step-by-Step	ISBN-13: 978-0205985517

SPSS Software: SPSS software is available to you on all CDM computer lab terminals. You do NOT need to buy your own copy of SPSS. However if you wish to do so for your personal use, here is the web link/further information: http://www-03.ibm.com/software/products/en/spss-stats-gradpack.

Other Materials-Required:

- A simple non-programmable calculator (should NOT have statistical functions, the ability to be programmed, or the ability to remember any information such as formulae in its memory; should HAVE a square root/square function in addition to routine math/arithmetic functions)
- Ruler and Pencil (graphs/diagrams will need to be drawn with ruler and pencil often times during the course, in addition to drawing them using software)

COURSE DESCRIPTION:
The aim of the course is to illustrate data analysis methods through the use of Statistics concepts and Probability Theory. Some key course topics include: descriptive statistics, data visualization, an introduction to statistical inference, and linear regression models. Students will use the statistical package SPSS to conduct statistical analyses of data sets. *SPSS will be used as a means to an end (to simplify certain statistical calculations), and not as an end in itself (the primary goal in this course is to teach data analysis, not SPSS).*

PREREQUISITE COURSES: None. However, students are expected to have a sound knowledge of basic mathematical notations and calculations and college algebra concepts. Knowledge of MS Excel will be useful. If you need to brush up your Math and Algebra, there are so many websites available these days in addition to many basic books. Here are a few useful website addresses:

http://www.wtamu.edu/academic/anns/mps/math/mathlab/col_algebra/index.htm.
http://www.mygretutor.com/tutorials/arithGRE.aspx
https://www.khanacademy.org/math

COURSE OBJECTIVES:

- to assist the students in developing an understanding of the basic concepts of data analysis through the use of statistics and probability,
- to help the students become informed and critical readers of quantitative arguments,
- to provide sufficient skills to apply simple statistical techniques
- to help the students gain an appreciation for the role of statistics in empirical research and scientific study,
- to help the students to gain flexible problem-solving skills applicable to unfamiliar statistical settings.

COURSE SCHEDULE, TOPICS, AND ACTIVITIES:

The following course schedule is a tentative schedule subject to changes as needed throughout the term. When changes are made, the updated document will be posted in D2L. It is the student's responsibility to keep track of changes to this schedule by viewing this schedule in D2L every week.

Module	Main Topics
Module#1	Introductions/Course Overview Introduction to Data Analysis, Statistics and Statistical terminology, Descriptive and Inferential Statistics, Types of data and their treatment, Exploratory data analysis, Use of graphs such as histograms, bar charts, box plots, stem-and-leaf plots to analyze and display data, Statistics for central tendency and spread, Types of variables, Categorical variable, Two way tables and their analysis, Simpson's Paradox, Introduction to distributions and examining distributions, Time Plots, Chebyshev's rule and its application, How to present the findings of data exploration
Module#2	Density curves, Random variable, Normal distributions and their properties, Normal distribution applications, Testing the normality assumption, Shapiro-Wilk's test, reading SPSS output for tests of normality
Module#3	Data relationships, Scatter Plots, Correlation, Pearson vs Spearman Correlation, Coefficient of determination, Correlation vs causation, Lurking variable and confounding variable, Introduction to regression analysis, Simple Linear Regression/Least Squares Regression, model fitting, Residuals, Treatment of outliers and extrapolations in regression
Module#4	Experimental designs, Data types, Surveys, Confounding, Factors and treatments, Comparative experiments, Randomization, Matched pairs, Blocked designs, Principles and types of sampling, Bias and variability
Module#5	Probability-Probability models and rules, Random variable and its applications, Statistical estimation, Law of large numbers, Venn diagrams, Tree diagrams, Bayes theorem
Module#6	Sampling distribution for averages/means, Central limit theorem, Sampling distributions for counts and proportions, Binomial distribution and Binomial probability, Normal approximation to the Binomial distribution and its applications.
Module#7	Introduction to statistical inference, Confidence interval and confidence levels, Sample size calculations, P value, Hypothesis testing
Module#8	An introduction to more advanced concepts in data analysis and statistics Course review for the final exam

FOR FURTHER DETAILS REFER TO THE DETAILED SYLLABUS POSTED IN D2L OR CONTACT THE INSTRUCTOR.