INSTRUCTOR: Jin Man Lee

    Office: Loop Campus (1 E Jackson DPC 6230 at the Department of Economics)
    Phone: (312)362-5970
    Email: jlee141@depaul.edu

COMMUNICATION

The best way to reach me is to send an email. Please use ECO520 as a prefix on the subject line to get my attention. If you don't receive my reply within 24 hours, please remind me again. Due to some email filters, your email might be lost. Before sending the email, read the discussions on D2L. You will find the most frequent questions and answers there. It is required to use the DePaul email address to communicate with the instructor for security.

ONLINE HYBRID

This is an Online Hybrid course. Online Hybrid blends asynchronous and synchronous online formats, with 50% of the learning activities designated as asynchronous online, and the other 50% synchronous online via Zoom. Attendance is required in the live Zoom sessions. Zoom classes are recorded, and recordings will be posted to D2L at my discretion. Please note: There is no option to attend live sessions entirely asynchronously - the course is designed for you to actively contribute in the live classes.

CLASS ETIQUETTE

- Students are expected to have their webcams switched on and microphones muted (unless talking). Students may use avatar images for the webcams or vitual backgroud for privacy.

- Students are expected to upload a Zoom profile photograph or acceptable avatar image.

- Students are expected to contribute in class. You must be in an environment where you are not distracted and can talk. You may be asked to share your computer screen and present to the class.

- Students connecting to Zoom must use their DePaul Zoom accounts. Students using non-DePaul Zoom accounts may be blocked from connecting to class. Your name in Zoom must match the preferred name on your university record and D2L.

- Class attendance will be automatically recorded by Zoom and credited to your grade.

ZOOM CLASS and OPEN OFFICE HOURS

- ZOOM CLASS: Every Wednesday 5:45-7:15 PM. This is a required class to attend. Attendance will be automatically recorded by ZOOM for the attendance credit. Make sure to use your full name as the name on D2L, and attend the whole time for attendance credit.

- OFFICE HOUR: ZOOM Office Hour Every Wednesday 4:00-5:00 PM. Any suggested topics are welcome via email prior to the meeting as a group. It can be a personal office hour and serve as first-come-first-serve. You can make an appointment via email if you are not available during the office hours.

- ZOOM Address for the class and office hour:
  https://depaul.zoom.us/j/95326652941

COURSE OBJECTIVES

This course focuses on developing advanced data analytics skills, enabling students to identify meaningful data patterns and transform these patterns into statistical models to inform profitable business decisions. The course covers a range of analytics methods, including descriptive, predictive, and prescriptive analytics. Key topics include advanced clustering analysis, discrete choice modeling, multivariate regression, random forest, and neural network analysis. These techniques will be applied to determine the most effective methods for driving data-driven business decisions. Throughout the course, students will extensively use SAS and R to develop analytical models using a variety of business data, including social media, housing, healthcare, banking, and publicly available datasets such as the American Community Survey (ACS) and the Current Population Survey (CPS).

PREREQUISITE

Applied Quantitative Analysis or passed the equivalent tests at the graduate school level

TEXTBOOK

- Applied Analytics through Case Studies Using SAS and R: Implementing Predictive Models and Machine Learning Techniques, Author: Deepti Gupta, ISBN: 9781484235249, Publisher: Apress (We will use many examples from this book)

- More required readings and article will be available on D2L every week

SOFTWARE USED in CLASS (required for Assignments)

- SAS Studio from SAS academy , SAS Studio from SAS University Edition, or SAS for Windows using apporto.com cloud service (https://depaul.apporto.com/)

- R with Rstudio for any platform. Current version: 4.1.0. This is a free public domain statistical program available to install. We will use R studio server from http: bigblue.depaul.edu:8787

GRADE

Weekly Assignments (40%), Attendance (5%), Midterm Exam (25%), Final Project Presentation (5%), and Final Project (25%)
Scale of grade: A: 93 or above, A-: 88-92.9, B+: 85-87.9, B: 80-84.9, B-: 77-79.9,
C+: 75-76.9, C: 70-74.9, C-: 68-69.9, D+: 65-67.9, D: 60-64.9, F: Below 60

EXAMS and PROJECT

- Midterm Exam (Online-Exam and Take-home SAS Project from 2/12 to 2/18 at 10:00 PM)

- Final Project Proposal via ZOOM presentation (PPT required) on 3/12 (each group has about 6 minutes of presentation and 4 minutes of discussion)

- Final Project Due on 3/22/2025 at 10:00 PM, upload all files to D2L

ASSIGNMENTS

Weekly assignments are related to review of computational work using many real world example data.

- All assignments are to be prepared individually. You risk an academic integrity violation if you submit the same work and answers with others. Group study is encouraged but not the submission of Assignments.

- All statistical code needs to be submitted as a txt file, For example, assignment.sas.txt or assignment.R.txt.

- All submitted codes should be tested without any error. If there are any errors, the assignment will get zero credit.

- All weekly assignments should be uploaded to D2L by Tuesday at 10:00 PM.

- Any late submission after the due dates will get a steep penalty without prior approval by the instructor.

- After the first submission, revise and resubmissions will be allowed until the following Tuesday at 10:00 PM.

DISCUSSIONS in D2L

We will have weekly discussions in D2L. You can ask any questions related to the material we covered each week, including homework. This is an excellent online-community space, so you are welcome to give answers or explanations to the questions. I will review the discussion board and leave comments if needed.

ACADEMIC HONESTY

Work done for this course must adhere to the University Academic Integrity Policy. Violations include but are not limited to the following categories: cheating; plagiarism; fabrication and academic misconduct.

- Cheating: any action that violates University norms or an instructor's guidelines for the preparation and submission of assignments. Such actions may include using or providing unauthorized assistance or materials on course assignments, or possessing unauthorized materials during an examination.

- Plagiarism: the representation of others' work as your own. You are to prepare your own assignments. Violations may result in the failure of the assignment, failure of the course, and/or additional disciplinary actions.

- Misconduct: This includes but is not limited to attempts to bribe an instructor for academic advantage; persistent hostile treatment of, or any act or threat of violence against, an instructor, advisor or other students. Violations may result in additional disciplinary actions by other university officials and possible civil or criminal prosecution.

You may review the Academic Integrity Policy in the Student Handbook or by visiting Academic Integrity at DePaul University (http://academicintegrity.depaul.edu)

STUDENT WITH DISABILITY
Students with Disability may register The Productive Learning Strategies (PLuS) Program. You may request your exam schedule arrangement by requesting through the PLuS program. For more information on the PLuS program, you may visit http://studentaffairs.depaul.edu/plus/ or call: 312-362-8000.

SUMMARY OF WEEKLY SCHEDULE

1. Monday 10:00 AM : Weekly Material will be posted in D2L

2. Wendesday 5:45-7:15 PM: ZOOM Online Class (Require to attend)

3. Tuesday 10:00 PM : Weekly Assignment Due

4. Wednesday 4:00-5:00 PM: ZOOM Open Office Hours

TENTATIVE SCHEDULE OF TOPICS

(The instructor may change the order or contents by needs, any special material needs for the class will be available on D2L)

- WEEK 1. Introduction to Business Analytics and SAS

    - TOPIC 1: Introduction to Business Analytics (BA)
        * Business Analytics, Data Analytics, and Big Data
    - TOPIC 2: Introduction to SAS Programming
        * SAS Data procedure syntax
        * SAS Descriptive Analytics
        * Assignment 1 (SAS Coding Assignment)

- WEEK 2. Descriptive Analytics and Data Cleaning

    - TOPIC 3: Statistics and Data Exploration for BA
      We will explore the topics on more issues on descriptive statistics and data issues using SAS. Following topics will be included:
        * Statistics and Inference: Descriptive Statistics, Inference, and Hypothesis Tests on Business Decision
    - TOPIC 4: Validation and Sanitation of Data
      We will explore the topics on more issues on descriptive statistics and data issues using SAS. Following topics will be included:
        * Data Exploration and Sanitation: Understand the challenge of real data and enhance the ability to programming skills
        * Missing Value and Outlier Treatment
        * Assignment 2 (SAS: Descriptive Analytics and Cleaning Data)

- WEEK 3. Classification and Clustering Analysis

    - TOPIC 5: Unsupervised Learning: Classification and Clustering Analysis
      Clustering Analysis in Big Data: Cluster effect depending upon the geographical, political, and socio-economic environment. Since the process of identifying, defining, and describing is not standard, the student will need to work on their data to understand how to cluster their data for further data analysis.

* K-Means
    * Hierarchical Clustering
    * Data Visualization and Statistics by Clustering
    * Assignment 3 (Clustering Analysis)

- WEEK 4-5. Predictive Analytics: Regression Analysis

  - TOPIC 6: Regression Analysis I
    Regression and forecasting techniques can yield new insight for managers by uncovering patterns and relationships that they had not previously noticed or considered. The student will learn how to make appropriate regression models reach a business decision by allowing various alternative models.
    * Correlation Coefficient and Causation
    * Simple and Multiple regression models
    * Inference of Regression
    * Assignment 4 (Regression Analysis I)
  - TOPIC 6: Regression Analysis II
    * Dummy and Categorical Variables
    * Nonlinear and Functional Forms
    * Regression Diagnostic Tests
    * Prediction and Power Tests
    * Assignment 5 (Regression Analysis II)

- WEEK 6. Midterm Exam Due (Tuesday) Feb 18 10:00 PM

- WEEK 7. Introduction to R

  - TOPIC 7: Predictive Analytics in R
    * Introduction to R for Predictive Analytics
    * Bring your own data (BYOD) proposal and data collection
    * Assignment 6 (Regression Analysis using R)

- WEEK 8. Binary Classification Model

  - TOPIC 8: Linear Probability and Logistic Model for Binary Classification
    models might be used for supply chain, investment decision, response to marketing communication, payment of credit card, shoppers' brand choice model, consumer purchasing decision model, drug choice model, and readmission of patients in a hospital.
    * Linear Probability Model
    * Binomial Logistic Regression
    * Assignment 7 (Discrete Choice Model)

- WEEK 9. Machine Learning (ML) and Deep Learning Models (DL)

  - TOPIC 9: Random Forest Model for ML

* Random Forest Model (RFM) is a supervised classification and prediction algorithm. Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.
    * Assignment 8 (RFM)
  - TOPIC 10: Artificial Neural Network for DL
    * Artificial Neural Network Model (ANN) provides the solutions to many problems in classification, clustering, regression, Image recognition, speech recognition, and natural language processing. Deep neural network analysis considered as components of machine-learning applications.
    * Assignment 9 (RFM, ANN)

- WEEK 10. Case Study and Project Management

  - TOPIC 11 Project Management and Final Project Presentation
    * Case Study: Data Analytics and Strategy - Chicago Housing Price Models, Loan Performance Model
    * Descriptive, Predictive, Prescriptive Analysis using clustering, regression, and machine learning models.
  - FINAL PROJECT Proposal Presentation: March 12 (Wednesday)
    Final project can be a group project with maximum of three students in each group. Each group can find their own topic and data of interest. The proposal presentation should have the PowerPoint to discuss the topics of interest, introduce the data using the descriptive analytics, and explain how to finish the project.

- FINAL PROJECT DUE : March 22 (Saturday) 10:00 PM