ISOM3530 Business Data Analytics

Course Outline

LECTURE:

Instructor: Dr. Baoqian PAN, Kris
Room: 5041 (LSK Business building); Email: **ismtpbq@ust.hk**
Office Hours: Wednesday 10:30am-11:30am (LSK5041).

Teaching Assistant: He Haolin, hheaz@connect.ust.hk
Kong Hao, hkongab@connect.ust.hk
Office Hours: Mon: 1:00pm-2:00pm; Room: 5017 (LSK Business building).
Lab section: Tutorial sections will be arranged to help learn programming language R and Python.

Class meets:
9am-10:20am, Mon and Wed (Rm1409, L25-26)

PC Laboratory:
10:30am-11:20am, Thu (LSKG005)

COURSE OBJECTIVE AND INTEDED LEARNING OUTCOMES

This is a course designed to rapidly get you to speed on doing data analytics for all business professionals. You will learn how Business analytics can be used to describe, predict, and inform business decisions in the areas of finance, marketing, operations, and human resources. By developing a strong understanding of data literacy and an analytic mindset, you'll be able to make more informed decisions and drive better business outcomes.

We'll focus on practical applications of data analytics, giving you hands-on experience with real-world data sets. Our goal is to help you develop the skills you need to derive insights from complex data and make appropriate business strategy recommendations.

After you complete this course, you will be able to:
• Understand and apply basic statistical analysis to practical problems and to present the result in a meaningful way.
• Interpret the results of business analytics and their implications to business administrations
• Make data driven decisions to optimize the business process and address issues in business administrations.

• Implement analytical models using R or Python.

PREREQUISITES

ISOM2500 Business Statistics or MATH2411 Applied Statistics

REFERENCE BOOK

Practical Multivariate Analysis, 5th edition, by Abdelmonem Afifi, Susanne May, Virginia A. Clark.
Introduction to Statistical Learning, by James, Witten, Hastie and Tibshirani.

SYLLABUS

This syllabus is subject to change without notice at the instructor's discretion.

Topic 1: Introduction

Topic 2: Data Exploration
- Summary statistic to spot data problem
- Visualization to spot data problem
- Different visualization techniques

Topic 3: Data Preparation/Preprocessing
- Data Transformations for individual predictors
- Data Transformations for multiple predictors
- Dealing with Missing Values
- Removing Predictors
- Adding Predictors
- Binning Predictors

Topic 4: Review - Multiple Linear Regression
- Multicolinearity
- F test and partial F test
- Regression Model Diagnosis

Topic 5: Fitting Curve to Data
- Quadric regression
- Smoothing splines

Topic 6: Time Series Analysis
- Stationary
- AR, MA, ARIMA model

Topic 7: Regression with Indicator and Interaction Variables
- Dichotomous and polytomous variables

- Modeling interaction effects with indicator variables

Topic 8: Variable Selection
- Subset Selection method (All/Best subset, Forward, Backward, Stepwise)
- Regulation or shrinkage method (Ridge regression, Lasso, Elastic Net)

Topic 9: Logistic Regression
- Building a linear model for binary response data
- Interpretation of the regression coefficient
- Statistical inference
- Assumption and diagnosis checking
- ROC curve
- Classification of new cases

Topic 10: Classification and Regression Tree (CART)

Topic 11: Tree Based Model
- Bagging and Random Forest
- Boosting (AdaBoost, XGBoost, LightGBM, CatBoost)

Topic 12: Remedies for Severe Class Imbalance

COURSE WEBSITE

http://canvas.ust.hk

GRADING

Your grade in the course is based on: Homework 32%; Midterm Exam 33%; Final Exam 35%.

Explanation of Assessment

1. **Homework (32%)**

   1) Throughout the semester, you will be given four assignments, which can be completed either individually or in teams of up to three members. Please ensure that you show your work clearly, as credit will not be given for incomplete or unclear submissions.
   2) The assignments will require you to use R or Python to perform data analyses. Please ensure that your solutions are typed and accompanied by computer outputs, such as graphs and computational results. All computer outputs must be original printouts, as photocopies will not be accepted.

3) To submit your assignments, please upload the code and softcopy of your solution to CANVAS before the deadline.
4) Your teaching assistant will collect the softcopies of your assignment solutions during the weekly tutorial sessions and return them to you in the next session.

2. Midterm Exam (33%) & Final Exam (35%)

1) Both exams will be closed book and closed notes, but don't worry - you will be allowed to prepare a summary sheet (hand-written, double-sided, A4-sized) to help you remember key concepts and formulas. We will also provide relevant statistical tables if necessary. Just be sure to include your name on your summary sheet and turn it in along with your exam.
2) To ensure fairness and accuracy, please bring your own summary sheet and calculator. Mobile phone, computer, iPad cannot be used as substitutes for calculators or writing instruments, and sharing is not allowed.
3) The final exam is not cumulative, we will not test the content of midterm exam directly.

3. Weekly Practice (0%)

1) At the end of each week, we will upload a set of Practice Questions and Answers to CANVAS in separate files. These questions will be in a similar format to the midterm and final exams, giving you valuable practice and helping you to prepare for the exams.
2) Please note that you do not need to submit your answers - these practice sessions are just for you to review the course content and test your understanding.

ACADEMIC INTEGRITY

Academic dishonesty includes, but is not limited to cheating, plagiarizing, fabricating of information facilitating acts of academic dishonesty by others, having unauthorized possession of examinations, submitting work of other groups, or tampering with the academic work of other groups. All exam answers must be your own, and you must not provide any assistance to other students during exams. Current university policy on academic dishonesty is "if a student is discovered cheating however minor the offence, the course grade will appear on the students' record with an X, to show that the grade resulted from cheating." This X grade stays on the record until graduation. If the student cheats again and "earns" another X grade, the student will be dismissed from the university. Please make sure you adhere to the HKUST Academic Honor Code at all time (see http://www.ust.hk/vpaao/integrity/).