SMU
SINGAPORE MANAGEMENT UNIVERSITY
MASTERS

Master of Science in Quantitative Finance

QF627 Extras - Financial Analytics

Instructor Name: Sungjong ROH, PhD
Title: Assistant Professor
Phone: 6828 0377
Email: sroh@smu.edu.sg
Office: LKCSB #4077
URL: http://TalkToRoh.com

COURSE DESCRIPTION

The course is tailored for **MQF students** to equip them with the skills to understand and apply econometric methods for causal inference, as well as machine learning, to solve problems in finance and real estate. It offers step-by-step guidance on implementing these methods using Python.

Throughout this course, students will gain a comprehensive understanding of both causal and predictive analytics. This includes mastering econometric methods of causal inference as well as machine learning techniques. The course's practical approach is designed to address a range of real-world business challenges in finance and real estate. Over the duration of ten weeks, it aims to transform theoretical knowledge of causal and predictive analytics into actionable skills, equipping students to solve complex business problems effectively.

The teaching methodology of this course is a blend of interactive lectures and hands-on programming exercises. The lectures focus on the principles and applications of causal and predictive analytics, while the hands-on component involves students actively engaging with the material. This engagement is facilitated through the completion of segments of script templates provided in the course.

This combination of conceptual understanding and technical, experiential learning is key to the course's effectiveness. The emphasis on hands-on learning encourages students to learn by doing, a critical aspect of understanding and applying the course content in real-world situations. This approach is instrumental in preparing students to work efficiently with data science teams within their organizations.

To further enhance the learning process, the course incorporates weekly exercise questions in Python, starting from Week 3. These exercises are supplemented with a review session during the weekly consultation, ensuring a thorough understanding and application of the course material.

LEARNING OBJECTIVES

By the end of this course, students will be able to:

(a) Navigate through a comprehensive "workflow" of causal and predictive analytics using the Python programming language, demonstrating proficiency in handling analytical processes.
(b) Comprehend the fundamental concepts of econometric methods for causal inference and the principles of machine learning, establishing a strong theoretical foundation.
(c) Conduct causal and predictive analytics with Python.
(d) Generate reproducible reports that detail the processes and findings of causal and predictive analytics, specifically tailored to address challenges in the finance and real estate sectors.
(e) Apply the results of their analyses to derive insights and support business decision-making, effectively bridging the gap between technical analysis and practical application in a business context.

INSTRUCTIONAL METHODS AND EXPECTATIONS

The class sessions will consist of an instructional component and a variety of other learning activities, including short lectures, hands-on data extraction, wrangling, management, and analysis. To be specific, each session will include three components: (1) learning key concepts in Python programming; (2) doing essential Python programming; (3) testing the understanding and application of (1) and (2) in the practice domain. Students should bring their laptops to the online classroom and be prepared to perform hands-on practice with programs and complete given tasks while in class. We will use Jupyter Notebook for the programming environment in this course. Students should bring their own computers to the class (not company laptops or computers), and be sure they have administrative access.

ACADEMIC INTEGRITY

All acts of academic dishonesty (including, but not limited to, plagiarism, cheating, fabrication, facilitation of acts of academic dishonesty by others, unauthorized possession of exam questions, or tampering with the academic work of other students) are serious offences.

All work (whether oral or written) submitted for purposes of assessment must be the student's own work. Penalties for violation of the policy range from zero marks for the component assessment to expulsion, depending on the nature of the offence.

When in doubt, students should consult the course instructor. Details on the SMU Code of Academic Integrity may be accessed at http://www.smuscd.org/resources.html.

CONSULTATIONS

Students may secure consultations with the instructor. Students are expected to send an email to the instructor so that a meeting time can be arranged in case there are multiple students who wish to meet.

CLASS TIMINGS

TUESDAYS (07:00PM—10:15pm)

Please see LESSON PLANS below for the detailed dates of each week's lesson.

STUDY MATERIALS

Lecture notes written on Jupyter Notebook are accessible via eLearn. Students are not required to purchase any reading materials. All materials for the course are either open-source or available through the library webpage.

The following open-source online books are required reading for the course.

- Facure, M. (2023). *Causal inference for the brave and true.*
- McKinney, W. (2022). *Python for data analysis*: *Data wrangling with pandas, NumPy, & Jupyter.*
- VanderPlas, J. (2022). Python data science handbook: Essential tools for working with data.

LESSON PLANS

Below is a tentative course plan. The course plans to maximize students' actionable and tangible learning, not to concentrate on one-way instruction. To maximize students' learning progress, after observations of their learning, the course plan will take into account their progress and may be revised during the semester.

Week 1: Introduction to the Course

- Introduction of Course Members (Myself, Our TA, and You)
- Course Logistics
- Analytics for Data-Driven Decision Making and Problem Solving

Week 2: The Two Worlds of Statistical Modeling, Confounders, Colliders, and RCTs

- Installation of Python Programming Environment
- Modeling for Explanation and Prediction Purposes
- Confounders and Colliders
- Randomized Controlled Trials (RCTs)

Week 3: Foundations in Regression Analysis, The Grammar of Graphics, and Identification

- The Grammar of Graphics
- Regression and Decomposition of Variance
- Exogeneity (and Endogeneity), Measurement Errors, and Identification

Week 4: Matching with Propensity Scoring, Inverse Probability Weights

- Matching with Propensity Scoring
- Inverse Probability Weights

Problem Set 1 is Due

Problem-Solving Leadership Session 1

Week 5: Difference-in-Differences and Regression Discontinuity

- Causal Inference with Difference-in-Differences (DiD)
- Regression Discontinuity Design (RDD)

Problem Set 2 is Due

Problem-Solving Leadership Session 2

Week 6: Regression Discontinuity and Instrumental Variables

- RDD Revisited
- Instrumental Variables

Problem Set 3 is Due
Problem-Solving Leadership Session 3

Recess Recess Week

- Postgraduate Program Term Break

Week 7: Synthetic Control

- Causal Inference with Synthetic Counterfactual

Problem Set 4 is Due

Problem-Solving Leadership Session 4

Week 8: Modeling for Prediction Purposes: The y-hat Question

- Counterfactual Predictions: Modeling for Prediction vs. Explanation Revisited
- Guest Lecture: The World of Quantitative Hedge Funds and AI/ML

Problem Set 5 is Due

Problem-Solving Leadership Session 5

Week 9: Machine Learning Part 1: Supervised Learning

- Predicting Singapore Housing Price with Supervised Learning

Problem Set 6 is Due

Problem-Solving Leadership Session 6

Week 10: Machine Learning Part 2: Unsupervised Learning

- Association Rule Mining
- Clustering Singapore Housing Market (feat. Dimensionality Reduction)

Problem Set 7 is Due

Problem-Solving Leadership Session 7