School of Economics
Academic Year 2022-23
Term 2

DSA301: : **Time Series Data Analysis**
Instructor Name : Benjamin Ee
Email : benjaminee@smu.edu.sg
Tel : +65 9277 9294

COURSE DESCRIPTION

This course seeks to provide students with a statistical toolkit for analysis, modelling and forecasting of time series data. We begin with a discussion of the distinguishing properties of time series, before progressing through an ascending sequence of more general linear parametric models including AR, MA, ARIMA, ECMs and VAR models (these acronyms will be second nature by the end of the course!).

Time series classification techniques will also be covered including support vector machines (supervised) and clustering techniques (unsupervised learning). We will pay special regard to the distinguishing properties of time series data such as temporal ordering, contiguous datapoints, etc and implications to the forgoing algorithms. Prior discussions on time series forecasting will also be extended with a discussion on the application of neural networks.

As part of the course, we will discuss choice of techniques or models given data availability, limitations and research or commercial intent. This is because in practice, there could be multiple options available to the data analyst.

Students will be provided with basic "example" implementations of time series algorithms using two major industry relevant programming languages (R & Python); students will then have the opportunity to enhance these code chunks using what they have learned, and also to apply the algorithms to current data sourced from major public and private sector databases. Doing this will also help students build a greater intuition for the practical 'quirks' found in working with time series in practice such as "data wrangling" and cleaning, dealing with missing data, data exploration and visualization prior to modelling. Given the prevalence of both R and Python, it is also hoped that the code from this course will be useful in future research projects or commercial settings. With this goal in mind, the code provided focuses on employing widely available and open source software packages such as R's vars and seasonal as well as Python's pandas and numpy. We also use industry standard Keras and Tensorflow packages for both languages. This maximizes the capacity of student code written to work in the broadest possible context. The main language of the course, and the language to be tested, is R.

PRE-REQUISITE/CO-REQUISITE/MUTUALLY EXCLUSIVE COURSE(S)

Please use the class search function at OASIS > Study > BOSS > BOSS Bidding (Plan & Bid page > Add to Cart and Perform Course Search) or the course catalogue in OASIS to check the most updated attributes of this course.

LEARNING OBJECTIVES

By the end of this course, students will be able to
- Understand and describe the differences between time series and cross-sectional data
- Further categorize time series trends into cyclical and non-cyclical components
- Build theoretical models of time series data based on MA, AR and ARMA components
- Appreciate and describe the differences between cointegration and correlation
- Understand the concept of stationarity.
- Transform non stationary data to stationary data (e.g. through differencing)
- Build univariate ARIMA and ECM models, with climate change modelling as a case study for ECMs
- Build multi-endogenous VAR models
- Perform time series classification using support vector machines, and clustering techniques

• Perform time series regression, and appreciate CAPM as a case study with both statistical and economic perspectives
• Perform time series modelling / forecasting using both parametric linear models (e.g. ARIMA-X and dynamic regression approaches, with GARCH as a case study) or nonlinear approaches such as neural networks
• Understand the general significance of domain specific relationships in time series modelling using the forgoing application in volatility modelling as example
• Be proficient (by the end of the course) in utilizing the R programming language for time series data analysis

ASSESSMENT METHODS

Basis of Assessment

Class participation and discussion	20%
Homework assignments	20%
Group Project	20%
Final exam	40%

Final Examination (40%): 2 hour paper. Students will need to bring a laptop computer for R programming.

Assignments (20%): There will be 3 take home assignments spanning the topics covered. The problem sets includes project like questions designed to give students the opportunity to review and enhance their understanding of the class material. Two take home assignments (5% each) will be done in groups of up to 3 students with class presentations, while the last will be individual (10%). All assignments will be distributed through eLearn.

Group project (20%): Students may chose from amongst a preselected set of project topics spanning finance, climate, healthcare, technology and other areas of economics. They will work on these projects in groups of up to 3 students. Project milestones will be in the following order: (i) a mid-term written proposal around week 4 (with ungraded feedback provided by the instructor shortly after) and (ii) a written report due shortly before the final examination. Overall, the group projects are a platform for students to appreciate the significance of domain specific insights in modelling, and also to build a portfolio for internships or job interviews.

Class Participation (20%): Students are encouraged to ask questions and provide suggestions throughout the course. An internal class discussion forum will also be setup for students to discuss data and coding, with helpful participation being recognized.

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 offense.

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

COPYRIGHT NOTICE

ACCESSIBILITY
SMU strives to make learning experiences accessible for all. If students anticipate or experience physical or academic barriers due to disability, please let the instructor know immediately. Students are also welcome to contact the university's disability services team if they have questions or concerns about academic provisions: included@smu.edu.sg.

Please be aware that the accessible tables in the seminar room should remain available for students who require them.

DIGITAL READINESS FOR TEACHING AND LEARNING (DRTL)
As part of emergency preparedness, instructors may conduct lessons online via the Zoom platform during the term, to prepare students for online learning. During an actual emergency, students will be notified to access the Zoom platform for their online lessons. The class schedule will mirror the current face-to-face class timetable unless otherwise stated.

CLASS TIMINGS
Class sessions are of 3-hour duration per week. Each session will involve a lecture and a discussion of assignments and readings.

RECOMMENDED TEXT AND READINGS

Required Text
R.J. Hyndman and G. Athanasopolos, Forecasting: Principles and Practice, 3rd edition, OTEXTS 2022; available online at https://otexts.com/fpp3/

WEEKLY LESSON PLAN

Week	**Topic**	**Readings & R packages**
Week 1	Introduction to Time Series Data and Exploratory Analysis	Chapter 2.1 to 2.7
Week 2	Autocorrelations & Data Transformations	Chapters 2.8 Chapter 3.1
Week 3	Benchmark Models & Forecast Evaluation	Chapter 5
Week 4 (Project proposal due)	Stationarity, Differencing and Time Series Decom-position	Chapter 9.1 Chapter 3
Week 5 Assignment 1 due	ARIMA Models (including seasonal frameworks)	Chapter 9.2 to 9.5, 9.9
Week 6	ARIMA Estimation	Chapter 9.6 to 9.8
Week 7	Vector ARIMA Models and Estimation	Chapter 12.3
Recess	Recess Week	
Week 8 Assignment 2 due	Time Series Regression Models	Chapter 7 Chapter 12.2
Week 9	Case Study: CAPM and Beta Estimation	Class slides
Week 10	Dynamic Regression Case Study: GARCH and Value at Risk	Chapter 10 Class slides
Week 11 Assignment 3 due	Error Correction Models Case Study: Climate Change Modelling	Class slides
Week 12	Machine Learning for Time Series Classification and Forecasting	Chapter 12.4 Class Slides
Revision week	Revision Week Zoom based review session on Saturday of 1st week. May be moved depending on final examination date	
Finals	Final Examination	