Time Series for Data Science

Olga Scrivner

"Time series are analyzed to understand the past and to predict the future, enabling managers or policy makers to make properly informed decisions."
(Cowpertwait and Metcalfe, p.1, 2009)

Course Description

The goal of this course is to develop practical skills needed to do applied research in areas operating with time series data. We will cover time series regression and exploratory data analysis, ARMA/ARIMA models, model identification/estimation/linear operators, Fourier analysis, spectral estimation, and state space models. The analyses will be performed using the freely available package astsa, xts, zoo. Both R and RStudio are required for this class. R Basics is required. Lectures and reading are obligatory.

Required Reading

Cowpertwait, P. and Metcalfe, A. (2009). Introductory Time Series with R. Springer. [CM]

Shumway, R.H., and Stoffer, D.S. (2006) Time series analysis and its applications. Springer. [SS]

Shumway, R.H., and Stoffer, D.S. (2006) Time series analysis and its applications. Springer. [SS]

Prerequisites

Prerequisites: R Basics

Course Communication Channels

Piazza, Zoom, Canvas, DataCamp

Course Structure

Weekly lectures, readings, quiz and hands-on R practice with DataCamp platform. The instructor will provide students with a free academic access to DataCamp. Students are required to obtain digital copies of required books (free to all IU students via IU library) and do weekly reading.

Assessments

Each week will have a quiz based on lectures and reading. There will be four graded assignments from DataCamp.

Final Project

Proposal (3-4 pages, double spaced). First select a fairly challenging data set for a final project. You are required to select your own data. The proposal should include the following: (1) description of the dataset; (2) specific hypotheses; (3) data exploratory analysis including preliminary plots and summary tables; (4) plan for future analysis and modeling for your final report.

The written final report (10 pages including figures, double spaced, APA style references) in a professional journal APA style that includes an introduction to the problem (hypotheses of interest), a brief description of the methods used to collect the data and the measurements obtained, a result section summarizing the results of the analyses deemed appropriate for the final project, and a discussion section that relates the analyses to the hypotheses stated in the introduction. Also include references in APA style.

Grading Policy

No late assignments are accepted unless an official documentation is provided (e.g., medical documentation). Business trips, job interview, vacations are not considered as valid excuses.

- **15%** Midterm exam
- **5%** Proposal
- **30%** Final Project
- **40 %** 4 Assignments
- **10%** Quizzes

Assignments

All assignments are based on DataCAmp. You will be required to sign up for a free account on DataCamp using your IU email. For each completed assignment you will be required to provide a statement of completion from DataCamp. Each assignment will consist of several sections of DataCamp courses which will be distributed across several weeks.

Schedule

The schedule is tentative and subject to change.

Week	Theory	Practice
Week 1	Introduction to Course Characteristics of Time Series Data [CC 1.1-1.4; SS 1.1] Matrix Algebra 1	Exploratory time series data analysis
Week 2	Decomposition [CM 1.5-1.6; SS 1.2] Matrix Algebra 2	Introduction to eXtensible Time Series
Week 3	Correlation [CM 2.1-2.2; SS 1.3-1.6] Probability	Basic Manipulations Merging and modifying time series
week 4	Time Series Regression [SS ch.2, CC ch.2] Linear Regression	Apply and aggregate by time Extra features of xts
week 5	Forecasting Models [CM ch.3] Impulse Response	Apply and aggregate by time Predicting the future
week 6	Stochastic Models [CM Ch.4] Time Domain Stochastic Processes	Correlation analysis Autoregression
week 7	Regression Models [CM Ch.5; CC Ch.3] Frequency Domain Stochastic Processes	A simple moving average
week 8	Stationary Models [CM Ch.6; CC Ch.3] State Space Model	
week 9	ARIMA [SS Ch.3; CM Ch.7] State Space Model	Time Series Data and Models Fitting ARIMA models
week 10	ARMA [CM Ch.6] State Space Model	ARIMA Models Seasonal ARIMA
week 11	Spectral Analysis [SS Ch.4]	Exploring and visualizing time series
week 12	Univariate Time Domain modeling	Benchmark methods Forecast accuracy
week 13	Multivariate Time Domain modeling	Exponential smoothing
week 14	Multivariate Models [CM Ch.11]	Forecasting with ARIMA models Advanced methods
week 15	Final project	