Applied Statistics Workshops

Fall 2025

The **Center for Statistical Computing** (CSC) invites all graduate students, staff, and faculty to participate in our Applied Statistics Workshops. Sessions are offered both via Zoom and in **Lab C** on the upper level (UL) of Healey Library. These workshops focus on recently developed statistical methods and make use of tools such as SPSS, SAS, Stata, R, G*Power and AMOS. Participants will receive handouts, program files, and example data sets. Workshop descriptions, schedules, and registration links are provided below.

COVID-19 Data Analysis Using R: This workshop will involve downloading COVID-19 data for different states and Massachusetts from the Center for Systems Science and Engineering of Johns Hopkins University and the Department of Public Health (DPH) of Massachusetts. We will employ time series and spatial regression models to analyze the COVID-19 data, utilizing R packages such as ***forecast***, ***tseries***, ***spdep***, ***maptools***, and ***ggplot2***. Additionally, this workshop will demonstrate how to use R to generate reports for COVID data.

Sample Size Estimation and Power Calculations (SAS and G*Power): This workshop covers sample size determinations and power estimation for various statistical comparisons and tests using the PROC POWER procedure in SAS.

Introduction to HLM (Mixed Models) (SPSS): This workshop provides an overview of the fundamental principles of multilevel/hierarchical linear models. Topics include the necessity for appropriate methods to model dependencies (e.g., clustering of students within schools), formulating and interpreting two-level multilevel models and their relevant parameters, and using SPSS to estimate model parameters.

Missing Data Analysis (SAS & Stata): This workshop covers the mechanisms of missing data, analysis of non-random selection bias, and methods of single and multiple imputation (MI) using SAS and Stata. Missing data is a common issue in various datasets. Most statistical software packages automatically eliminate entire cases with missing data from analysis, potentially leading to low sample sizes and biased results.

Structural Equation Modeling I (AMOS & R): This workshop introduces techniques for structural equation modeling (SEM). SEM is employed to test complex relationships between observed (measured) and unobserved (latent) variables. Topics covered include fundamentals underlying SEM, SEM notation, path diagrams, data preparation, mediation analysis, path analysis, parameter estimation, and assessment of model fit. AMOS and R are used to demonstrate examples.

Structural Equation Modeling II (AMOS & R): The second SEM workshop delves into advanced topics including measurement error, latent variables analysis, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), development of structural equation models with estimation, and model testing. Additionally, this workshop introduces latent growth models for longitudinal data. An R program and AMOS are utilized to demonstrate model structures, parameter estimation, and model modification.

Applied Statistics Workshops

Fall 2025

Introduction to Statistical Learning (R): This is an introductory workshop in statistical learning focusing on the important elements of modern data analysis such as regression and classification methods. Topics covered include linear and logistic regression, linear discriminant analysis, cross-validation, principal components, and clustering. Data analysis examples in this workshop are demonstrated using R.

Event History Analysis (Survival Models) in SPSS: This workshop using SPSS presents statistical methods of survival analysis, specifically focusing on studies where the outcome is a time-to-event variable. It covers the estimation of survival time using the life table and Kaplan-Meier Methods, as well as modeling survival risk. It also assesses the relationship of risk factors and survival times using the Cox regression model. SPSS 28.0 will be used for data analysis.

Event-Study Regression using R: Event-study is a causal inference research design method for analyzing the impact of a specific event on a particular outcome or variable of interest over a defined time period. The event can be considered as the treatment in a Difference-in-Difference (DiD) analysis, and the dynamics of the impact can be assessed by comparing the changes in outcomes over the time between the treated and control groups. This workshop will make use of a variety of R packages, specifically, ***fixest***, ***plm***, and ***did*** for event-study regression. Topics covered include data preparation, DiD analysis, dynamic DiD model, and the graphic display of the dynamic event effects.

Logistic Regression and Random Forest in R, a Practical Introduction: This workshop is designed to introduce the fundamental principles of classification techniques in data science, with a particular emphasis on Logistic Regression and Random Forest using R programming language. Participants will develop a conceptual understanding of both models, gain hands on experience implementing them in R, and examine their respective strengths and limitations through real-world examples. The session aims to equip students with the analytical skills needed to critically evaluate and apply classification methods in research and data analysis across a range of academic disciplines.

Registration Procedures:
Seats and handouts are limited. Please register in advance.

1. Click the 'In-person Register' or 'On-Zoom Register' under Registration.
2. Fill out all the information requested and submit your registration.
3. Join the workshops via Zoom link in the confirmation email or attend in-person sessions for in-person workshops.

All in-person workshops will be held in ***Lab C*** on the upper level (UL) of Healey Library.

Please contact Mr. Inal Mashukov at inal.mashukov001@umb.edu for any questions regarding the workshops.

Web: https://www.umb.edu/academics/graduate/info_for_graduate_students/center_for_statistical_computing

Applied Statistics Workshops

Fall 2025

Location: **Healey Library, Lab C.** (From the main elevators in Healey Library, take the Upper level (UL). Turn right out of elevator, and you'll find Lab C on the right in the hallway)

Fall 2025 Applied Statistical Workshop Schedule:

Topic	*Date*	*Day*	*Time*	*Registration*
COVID-19 Data Analysis Using R	**Sep. 22**	**Monday**	**10:00-12:00 P.M.**	On-Zoom Register
Intro to HLM (Mixed Models) (SPSS)	**Sep. 24**	**Wednesday**	**11:00-1:00 P.M.**	In-Person Register --- On-Zoom Register
Structural Equation Modeling I (AMOS & R)	**Sep 26**	**Friday**	**11:00-1:00 P.M.**	On-Zoom Register
Structural Equation Modeling II (AMOS & R)	**Oct. 3**	**Friday**	**11:00-1:00 P.M.**	On-Zoom Register
Intro to Statistical Learning (R)	**Oct. 10**	**Friday**	**11:00-1:00 P.M.**	On-Zoom Register
Missing Data Analysis (SAS & Stata)	**Oct. 13**	**Monday**	**11:00-1:00 P.M.**	In-Person Register
Sample Size Estimation (SAS and GPower)	**Oct. 29**	**Wednesday**	**11:00-1:00 P.M.**	In-Person Register --- On-Zoom Register
Event History Analysis (Survival Models) in SPSS	**Oct. 30**	**Thursday**	**11:00-1:00 P.M.**	In-Person Register
Event-Study Regression using R	**Nov. 3**	**Monday**	**11:00-1:00 P.M.**	On-Zoom Register
Logistic Regression and Random Forest in R	**Nov. 4**	**Tuesday**	**11:00-1:00 P.M.**	In-Person Register --- On-Zoom Register