2018 SUMMER STATISTICAL INSTITUTE

May 21 - June 8, 2018

Learn statistical terminology and computing tools
Obtain strategies for reproducible research
Enhance research techniques and develop new skills

Statistical Software: R
Python Data Science
Structural Equation Models

May 21-25, 2018 | Wescoe Hall 4066
1445 Jayhawk Blvd. Lawrence, KS 66045

May 29-June 1, 2018 | Watson Library 455
1425 Jayhawk Blvd. Lawrence, KS 66045

June 4-8, 2018 | Watson Library 455
1425 Jayhawk Blvd. Lawrence, KS 66045

REGISTRATION OPEN!
Go to crmda.ku.edu/statscamp

Suitable for research in academic, corporate & government settings

Can’t attend in person?
Attend remotely using zoom

All KU affiliates have a Zoom account.
Go to kansas.zoom.us to access your account.
If you are not affiliated with KU, Zoom is free to set-up.

crmda.ku.edu  Watson Library, Room 470  785.864.3353
KU SUMMER STATISTICAL INSTITUTE:
WEEK 1: R      MAY 21-25, 2018

Day 1—Installation & Getting to Know R
Instructor: Dr. Paul Johnson, CRMDA Director and Professor, Political Science
- Installation
- R documentation, interacting with the R help system, packages
- User interfaces: Comparing Windows R, Emacs, Notepad++, Macintosh
- Line aRt: Illustrating functions, create publication quality line art and graphic displays

Day 2—Establishing a Replicable Research Process
Instructor: Dr. Paul Johnson, CRMDA Director and Professor, Political Science
- Data impact and Recode data: Wrestling with numerical, text, and factor variables
- Graphical exploration and presentation scatterplots, barplots, boxplots, etc.
- Export tables for presentations in documents: cross tabulations regression, and other tables

Day 3—Statistical Analysis the R Way
Instructor: Dr. Paul Johnson, CRMDA Director and Professor, Political Science
- Regression & ANOVA
- Structural Equation Modeling
- Moderation and Mediation

Day 4—R Toolkit for Interacting with Data
Instructor: Dr. Paul Johnson, CRMDA Director and Professor, Political Science
- Matrix Algebra with R
- Iteration concepts in R: for, lapply
- Subsetting data, processing subsets and merging results
- Create R functions to customize analysis

Day 5—Monte Carlo Programming and Power Analysis
Instructors: Dr. Paul Johnson, CRMDA Director and Professor, Political Science
- Monte Carlo simulation in R
- Power analysis: definition and implications
- Use Monte Carlo simulation to estimate power
KU SUMMER STATISTICAL INSTITUTE:
WEEK 2: PYTHON DATA SCIENCE      MAY 29-JUNE 1, 2018

Day 1—Getting Started in Python
Instructor: Dr. Jonathan P. Lamb, CRMDA Faculty Fellow, Associate Professor of English
- What is Python?
- Installing Python and Anaconda
- Navigating around Jupyter Notebooks
- Discuss basic features and functions of Python

Day 2—Working with Text Data
Instructor: Dr. Jonathan P. Lamb, CRMDA Faculty Fellow, Associate Professor of English
- Work with texts and the question of data
- Introduce National Language Toolkit (NLTK)
- Discuss key functions and basic text analysis using NLTK
- Advanced Text Analysis, Corpus Analysis and Sentiment Analysis
- Topic modeling and measuring document similarity

Day 3—Data Science
Instructor: Dr. Jonathan P. Lamb, CRMDA Faculty Fellow, Associate Professor of English
- Introduce Pandas
- Discuss various functions of Pandas
- Perform statistical analysis

Day 4—Working with Internet Data
Instructor: TBD
- Scrape data from the Internet
- Understand web data formats
- Turn webpages into datasets
- Analyze Internet datasets
KU SUMMER STATISTICAL INSTITUTE:
WEEK 3: SEM       JUNE 4-8, 2018

Day 1—Introduction to Structural Equation Modeling
Instructor: Dr. Ed Merkle, University of Missouri, Department of Psychological Sciences
• From Regression to SEM
• Confirmatory Factor Analysis
• Estimation and Evaluation
• Advanced Topics: Multiple groups, Measurement Invariance

Day 2—Creating Models and Fitting Them (Part 1)
Instructor: CRMDA Staff
• Brief R Overview, including the bootstrap
• Estimating SEM with the lavaan package
• Moderators and Mediators in SEM
• The SEM Example Collection: Overview
• The Examples fitted with lavaan
• Acceptable Tables: The CRMDA semTable package

Day 3—Creating Models and Fitting Them (Part 2)
Instructor: CRMDA Staff
• Introduction to Mplus
• Ordinal Logistic Regression
• SEM with Ordinal Data
• Problems associated with Chi-Square Difference Tests

Day 4—Dealing with Missing/Incomplete Data
Instructor: CRMDA Staff
• The Problem of Missing Data: Listwise deletion is especially harmful in SEM
• FIML as a remedy for incomplete data
• No FIML for Ordinal Models? What Now?
• Multiple Imputation

Day 5—The New Wave: Bayesian Models for Structural Equations
Instructor: CRMDA Staff
• Bayesian Perspective
• Specifying SEM for Bayesian Estimation
• Examples from competing software (BUGS/JAGS and Stan)