KU CENTER FOR RESEARCH METHODS & DATA ANALYSIS

College of Liberal & Sciences

2019 SUMMER STATISTICAL INSTITUTE

May 20 - June 7, 2019

Build Your Toolkit at Stats Camp

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

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.

SESSIONS

- **Statistical Software: R**
  May 20-24, 2019 | Wescoe Hall 4066
  1445 Jayhawk Blvd. Lawrence, KS 66045

- **Python Data Science**
  May 28-31, 2019 | Wescoe Hall 4066
  1445 Jayhawk Blvd. Lawrence, KS 66045

- **Structural Equation Models**
  June 3-7, 2019 | Watson Library 455
  1425 Jayhawk Blvd. Lawrence, KS 66045

crmda.ku.edu ■ Watson Library, Room 470 ■ 785.864.3353

KU SUMMER STATISTICAL INSTITUTE:
WEEK 1: R   MAY 20-24, 2019

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 28-31, 2019

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 3-7, 2019

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 in SEM: Multi-group Models
- The SEM Example Collection: Overview
- The Examples fitted with lavaan

Day 3—Creating Models and Fitting Them (Part 2)
Instructor: CRMDA Staff

- Introduction to M*plus*
- Ordinal Logistic Regression
- SEM with Ordinal Data
- Problems associated with Chi-Square Difference Tests
- Presenting Results: The CRMDA semTable Function

Day 4—SEM Advanced Topics
Instructor: CRMDA Staff

- Missing Data: Full Information Maximum Likelihood (FIML) versus Multiple Imputation
- Mediation: Path Models

Day 5—New Modeling Opportunities
Instructor: CRMDA Staff

- Multi-Level Modeling and SEM
- Formative Constructs: Rethinking CFA