Statistical Consulting Center
Department of Statistics – Franklin College of Arts and Sciences
UNIVERSITY OF GEORGIA

SCC Seminar on Data Analysis

Exploratory Data Analysis in R
Advice and Code for Getting Started

Dan Hall, PhD
Director, UGA Statistical Consulting Center

You have some data and some questions you hope to answer. How do you get started? This talk addresses that question with a focus on data analysis in R. I will cover such basic issues as (a) How and why should you look at the raw data file? (b) How do you read the data into R? (c) What are the best tools for efficiently summarizing data sets? (d) How do you quickly and compactly produce summary statistics and grouped summaries? and (e) What plots should you use to visualize the data and how do you create them? Most of these questions fall under the heading of Exploratory Data Analysis (EDA). Most people who work with data know that EDA is important and that it should be done before launching into model-fitting and complex methods of statistical analysis. But students are rarely trained in EDA in any systematic way. There are few rules to EDA, but there are steps to follow, techniques to adopt, and goals to focus on. By the end of this talk, listeners without extensive experience with EDA will have a better understanding of how to approach the task using R software. And even experienced data analysts will probably learn some new tools and techniques.

Monday, November 27, 2023, 4:00pm-5:00pm

ROOM 209, Sanford Hall

About the SCC Seminar Series: This series of non-technical talks, sponsored by the UGA Statistical Consulting Center, aims to present topics of interest to statisticians and users of statistics related to the practical use of statistics for data analysis.

Please share this announcement with colleagues who may be interested. To get announcements for future talks in this series, join the SCC listserv here.