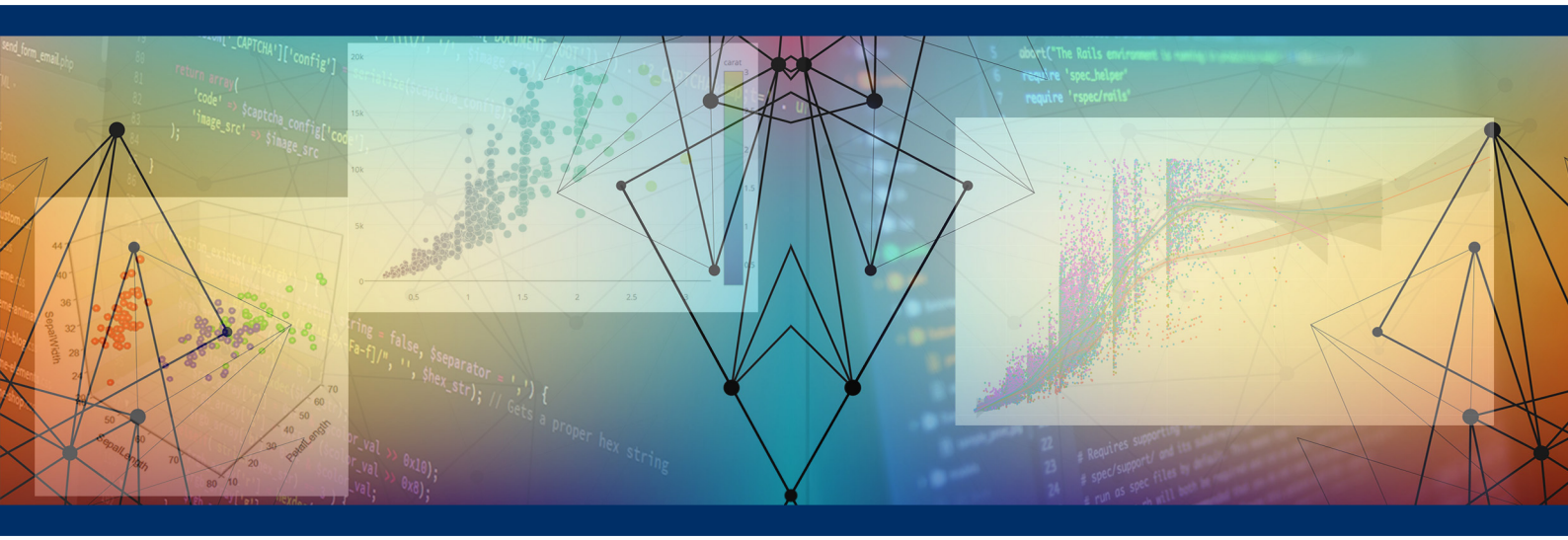
ENTERPRISES
University of Pretoria



An Introductory Data Science Workshop: R Fundamentals and Biostatistics

Hosted by the Forensic Anthropology Research Centre, Department of Anatomy, School of Medicine, Faculty of Health Sciences, University of Pretoria

20-24 August 2018

NW-1, Hatfield Campus, University of Pretoria
Mon-Thurs: 09:00 – 17:00
Fri: 09:00 – 13:00

FARC
Forensic Anthropology Research Centre

An Introductory Data Science Workshop: R Fundamentals and Biostatistics will provide you with the most important tools so you can do data science. Once you have the tools available to you, you can conduct a wide variety of data science challenges no matter the field or whether your data is small or big. Statistics lend credibility and validity to scientific research. While many statistical software programs are available, R is a free, versatile, open-source software program that can perform a broad and ever-growing range of statistical computing and data manipulation. Furthermore, R is renowned for its graphical capabilities.

This workshop provides an introduction to the R language and the environment, both of which are crucial for success using the program. Once the foundation is set, the course will transition to introducing statistical analyses that are commonly employed in biomedical research ranging from hypothesis testing and outlier detection to regression and classification statistics. The workshop will not focus on any one analysis or specific field; instead, the aim of the workshop is to teach the basic knowledge necessary to use R independently, thus helping participants initiate their own research projects and produce publications.

NEXT PAGE

ENTERPRISES
University of Pretoria

An Introductory Data Science Workshop: R Fundamentals and Biostatistics

Workshop content

* R and the RStudio environment
* Data Manipulation: Import, Cleaning, Transforming
* Data Visualizations with ggplot2
* R programming: The Tidyverse
* Descriptive Statistics, Outlier Detection, and Hypothesis testing
* Introduction to Supervised and Unsupervised Machine Learning Techniques
* Categorical Analyses, Linear and Logistic Regression, Discriminant Function Analysis, Multiple Correspondence Analysis, and Principal Component Analysis

Learning outcomes

The workshop is structured around hands-on practicals and interactive sessions that will ensure participants are comfortable with the R environment, running statistical analyses, and interpreting statistical outputs by the time the course is completed.

Who should enrol?

The workshop is ideal for anyone that is interested in becoming a better data scientist including postgraduate students and researchers.

Workshop requirements

No advanced statistics knowledge or prior programming or R experience is required.

In order to get as much out of the workshop as possible, you should bring a laptop that runs Windows (2000/XP/2003/ Vista/7/8/2012 Server/8.1/10) or Mac. The computer should have at least 32MB of RAM, a mouse, and enough disk space for recovered files, image files, etc. You should also bring your own laptop charger.

Workshop fees

R4 000.00 per delegate (VAT inlc.)
Workshop fees include all course material and refreshments during contact days.

Workshop fees must be paid in full 14 days prior to workshop start dates. Proof of payment can be submitted to enrolments@enterprises.up.ac.za.

Accreditation and certification

Enterprises University of Pretoria (Pty) Ltd is wholly owned by the University of Pretoria. As a public higher education institution, the University of Pretoria functions in accordance to the Higher Education Act 101 of 1997. Enterprises University of Pretoria offers short courses on behalf of the University and these short courses are not credit-bearing, and do not lead to formal qualifications on the National Qualifications Framework (NQF) – unless stated otherwise. Delegates who successfully complete a short course and comply with the related assessment criteria (where applicable) are awarded certificates of successful completion and/or attendance by the University of Pretoria.

Registration and enquiries

Workshop coordinator
Peggy Motswatswa
Tel: +27 (0)12 434 2640
Cell: +27 (0)83 282 6975
Email: peggy.motswatswa@enterprises.up.ac.za

Workshop enquiries
Gabriele C. Krüger
Forensic Anthropology Research Centre
Department of Anatomy
Tel: +27 (0)12 319 2539
Email: gabi.kruger@up.ac.za

NEXT PAGE

An Introductory Data Science Workshop: R Fundamentals and Biostatistics

Programme

Monday August 20
08:00 – 08:40 Registration
08:40 – 09:00 Introduction to the Course
09:00 – 10:00 Introduction: Introducing R, Markdown, and Importing Data
10:00 – 10:20 *Tea*
10:20 – 13:00 Introduction: Introducing R, Markdown, and Importing Data Data Munge
13:00 – 13:50 *Lunch*
13:50 – 15:20 Data Munge
15:20 – 15:30 *Break*
15:30 – 16:30 Data Munge

Tuesday August 21
09:00 – 10:30 Visualizations with ggplot2
10:30 – 10:45 *Tea*
10:45 – 13:00 Visualizations with ggplot2
13:00 – 13:50 *Lunch*
13:50 – 15:20 The Tidyverse
15:20 – 15:30 *Break*
15:30 – 17:00 In-Class Practical #1

Wednesday August 22
09:00 – 11:00 Exploring Assumptions & Outliers
11:00 – 11:15 *Tea*
11:15 – 12:00 Descriptive Statistics
12:00 – 13:00 Comparing Means
13:00 – 13:50 *Lunch*
13:50 – 14:20 Comparing Several Means
14:20 – 15:20 Categorical Data and Multiple Correspondence Analysis
15:20 – 15:30 *Break*
15:30 – 17:00 Correlations and In-class Practical #2

Thursday August 23
09:00 – 09:30 Machine Learning Introduction
09:30 – 10:30 Linear and Logistic Regression
10:30 – 10:45 *Tea*
10:45 – 13:00 Linear and Logistic Regression
13:00 – 13:50 *Lunch*
13:50 – 14:30 Discriminant Function Analysis (DFA)
14:30 – 14:45 *Break*
14:45 – 17:00 DFA

Friday August 24
09:00 – 10:30 Data Reduction Techniques: PCA
10:30 – 10:45 *Tea*
10:45 – 12:30 Cluster Analyses, Decision Trees, Shake-Rattle-Roll
12:30 – 13:00 Presentation of Certificates, Feedback, and Final Comments

NEXT PAGE

An Introductory Data Science Workshop: R Fundamentals and Biostatistics

Course presenters

Kyra E. Stull, PhD is an Assistant Professor in the Department of Anthropology at the University of Nevada, Reno. Dr. Stull earned her B.A. in Anthropology from the University of Tennessee, Knoxville (2006), a MS in Biological and Forensic Anthropology at Mercyhurst University (2008), and her PhD in Anatomy, with a concentration in Biological Anthropology, from the University of Pretoria (2014). Dr. Stull has research interests in forensic anthropology, human growth and development, modern human variation, and quantitative methods. Specifically, her research seeks to explore the predictive ability of age and sex indicators and to develop techniques to accurately identify unknown decedents. A large portion of her current research is dedicated to increasing the utility and accessibility of modern techniques and methods through intuitive graphical user interfaces that operate through R. She has provided professional workshops and has taught graduate and undergraduate courses on applied biostatistics using R since 2014.

kstull@unr.edu

Cortney N. Hulse, MA is currently a 3rd year PhD student in the Department of Anthropology at University of Nevada, Reno. Cortney earned her Associates of Arts (2011), B.A. in Anthropology (2013) and M.A. in Anthropology (2016) from Idaho State University. Cortney's research interests include trauma research, skeletal biomechanics, and human rights. Her current research utilizes advanced imaging techniques and computer modeling to create predictive models of bone failure.

chulse@nevada.unr.edu