Coding in R For Healthcare Informatics and Health Data Science

Kimberly Deas, MS, PhD Candidate
BlackTIDES Informatics Lead

This presentation provides an introduction to coding in R for healthcare informatics and health data science. We'll cover the basics of R, including installation and setup, as well as data manipulation and analysis techniques. We'll end with live coding of a healthcare data set.

Part 1: Introduction to R

Part 1: Introduction to R

- What is R?
- Why R is Useful in the Healthcare Data Analytics Field?
- Why R is useful in the healthcare data analytics field.
- R Resources

Introduction to R: What is R?

- R is a versatile programming language and software environment, highly regarded for statistical computing, data analysis, and data visualization.
- As an open-source tool, it's supported by a robust community and offers a vast repository of packages for statistical and data manipulation tasks.
- It excels in handling large datasets and complex analyses, while its straightforward syntax makes it accessible to users at all levels.
- R's integration with other languages and various data formats enhances its adaptability in a wide range of data scenarios.

Introduction To R: Why R is used in Statistical and Data Analysis.

- R's popularity in statistical and data analysis is driven by its comprehensive statistical techniques, covering everything from linear and nonlinear modeling to time-series analysis.
- Its strength lies in a vast array of user-contributed packages that simplify complex tasks and adapt to evolving methodologies.
- R's advanced plotting systems like ggplot2 enable the creation of detailed and insightful visualizations.

Introduction To R: Why R is Useful in the Healthcare Data Analytics Field?

- R plays a crucial role in healthcare data analytics with its proficiency in handling large, complex datasets common in medical research and public health.
- Its specialized packages for biostatistics and epidemiology, such as 'survival' and 'lme4', are essential tools in medical research for analyzing trends and patterns in health data.
- R excels in predictive modeling and machine learning, key for developing personalized medicine and forecasting disease outbreaks.
- Additionally, its ability to visualize health data simplifies complex information, aiding in effective communication for healthcare decision-making.

Introduction To R : Resources

- IMO, THE definitive book on R: ***R for Data Science: Import, Tidy, Transform, Visualize, and Model Data 2nd Edition*** by Hadley Wickham (Author), Mine Çetinkaya-Rundel (Author), Garrett Grolemund (Author)
- R Course: https://www.udemy.com/course/r-programming/. Kirill is also a great Tableau instructor.
- Others?

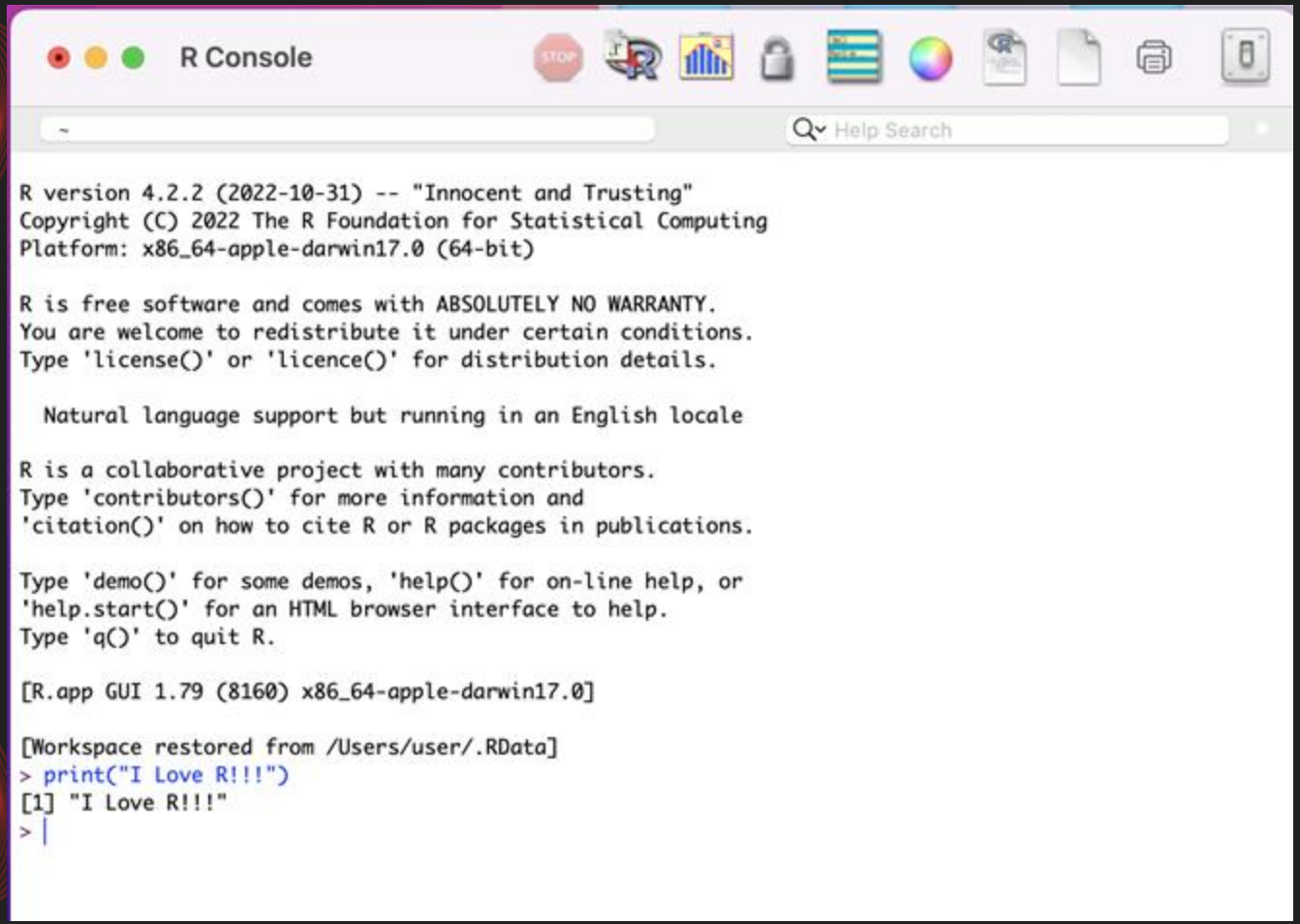
Part 2: Basics of R

Basics of R: Getting Started with R: Installation and Setup

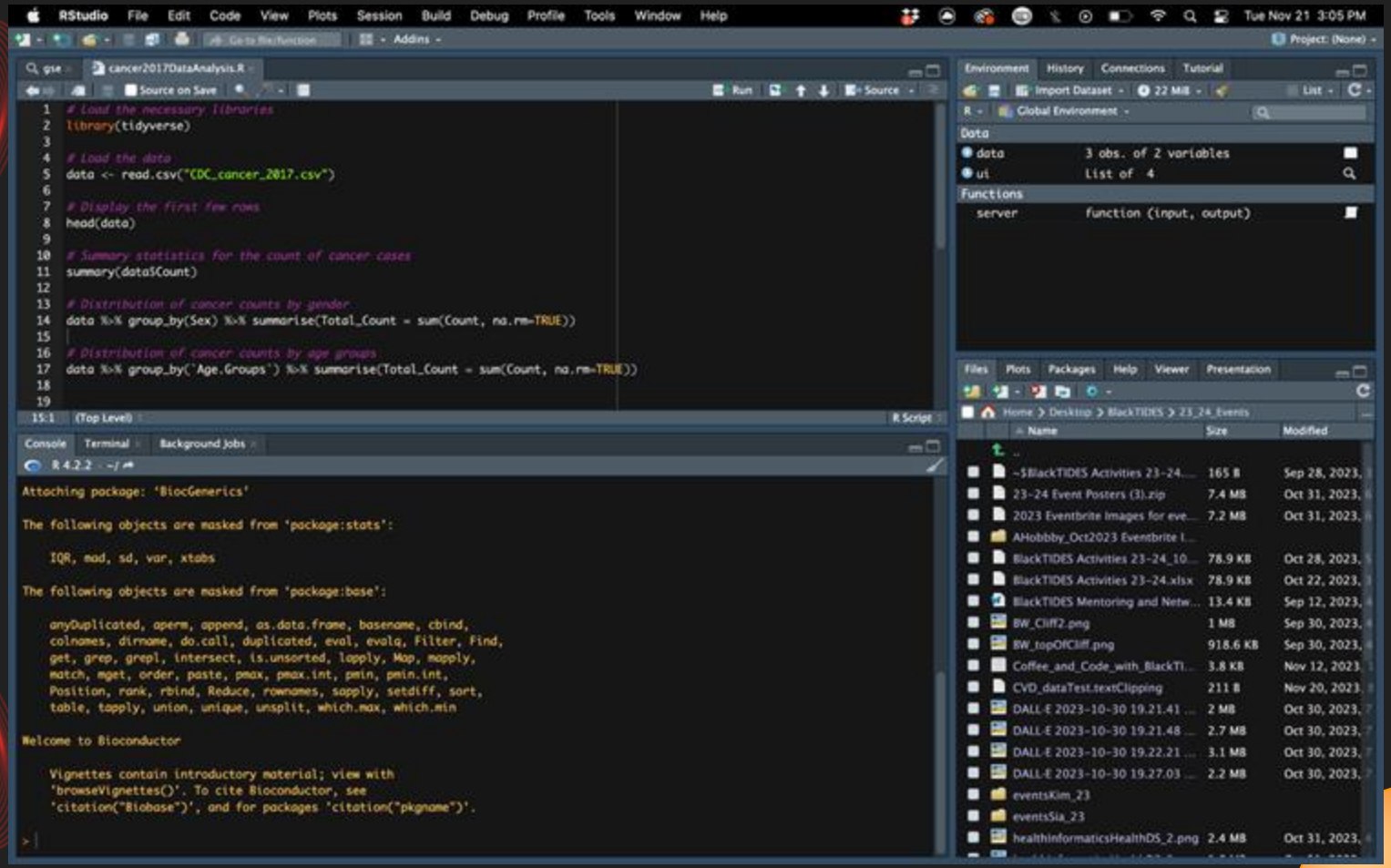
- Installing R and R studio: https://rstudio-education.github.io/hopr/starting.html

- Live demo of R and R studio

Basics of R: R Console



Basics of R: R Studio



Basics of R: Basic R Syntax and Commands

```r
# Variables and data types
age <- 30           # Numeric type
patient_id <- "P001"  # Character type
smoker <- TRUE       # Logical type

# Basic operations
new_age <- age + 1
print(new_age)

# Data structures: vector, matrix, list, data frame
vector <- c(1, 2, 3)
matrix <- matrix(1:9, nrow=3)
list <- list(age, patient_id, smoker)
data_frame <- data.frame(ID=c(1,2), Name=c("Alice", "Bob"))
```

Live demo R Studio - Basics

Part 3: Overview of the Dataset

Basics of R: Overview of the Dataset

- Loading the Dataset in R
- Exploratory Data Analysis
- Go to console

Part 4: Data Manipulation and Analysis

Data Manipulation and Analysis: Data Cleaning

- Handling missing values
- Data Transformation
- Go to console

Data Manipulation and Analysis: Basic Data Analysis Techniques

- Descriptive statistics
- Simple Visualizations
- Bar chart
- Scatter plot
- Go to console

Data Manipulation and Analysis: Using packages for Data Manipulation and Analysis

- dplyr and ggplot2
- Data Manipulation with dplyr
- Advanced data viz with ggplot2
- Go to console

Thank you. Please feel free to ask any questions.