STRINGFEST
ANALYTICS

Introduction to Statistics in R: One-day workshop

Lesson 1: Descriptive statistics

Lesson 2: Statistical data wrangling

Lesson 3: Foundations of inferential statistics

Lesson 4: Beyond t-tests

Lesson 5: Correlation

Lesson 6: Regression

Recommended preparation: Introduction to R, one-day workshop

Learning Objectives

- Student can classify, visualize and explore variables
- Student can prepare data for statistical analysis in R
- Student can compare the means of two groups
- Student can compare multiple groups, dependent samples and multiple categories
- Student can correlate two or more variables
- Student can conduct and interpret a univariate linear regression

Lesson 1: Descriptive statistics
Objective: Student can classify, visualize and explore variables
Description:
- What is tidy data?
- A tour of the tidy galaxies
- The tidy workflow

Time: 40 minutes
Assets needed: none

Lesson 2: Statistical data wrangling
Objective: Student can prepare data for statistical analysis in R
Description:
- Introduction to the tibble
- Importing text files
- Importing Excel workbooks

Time: 40 minutes
Assets needed: Baseball records

Lesson 3: Foundations of inferential statistics
Objective: Student can compare the means of two groups
Description:
- Pivoting and un-pivoting datasets
- Delimiting columns

Time: 60 minutes
Assets needed: Baseball records

Lesson 4: Beyond t-tests
Objective: Student can compare multiple groups, dependent samples and multiple categories
- Manipulating rows & columns
- Aggregating & summarizing data
- Piping functions

Time: 75 minutes
Assets needed: Baseball records

Lesson 5: Correlation
Objective: Student can correlate two or more variables
- Appending two or more tables
- Joining two tables: left, right, inner, outer

Time: 75 minutes
Assets needed: Baseball records

Lesson 6: Regression
Objective: Student can conduct and interpret a univariate linear regression
- Formatting, replacing and splitting strings
- Ordering and modifying factors
- Generating, calculating and resampling dates

Time: 60 minutes
Assets needed: Flight records