

Institute of Systems Science

National University of Singapore

**GRADUATE CERTIFICATE
BUSINESS ANALYTICS PRACTICE**

Supplementary Workshop Guide

Subject: *NICF- Statistics Bootcamp (SF)*

Interactive R Workshops

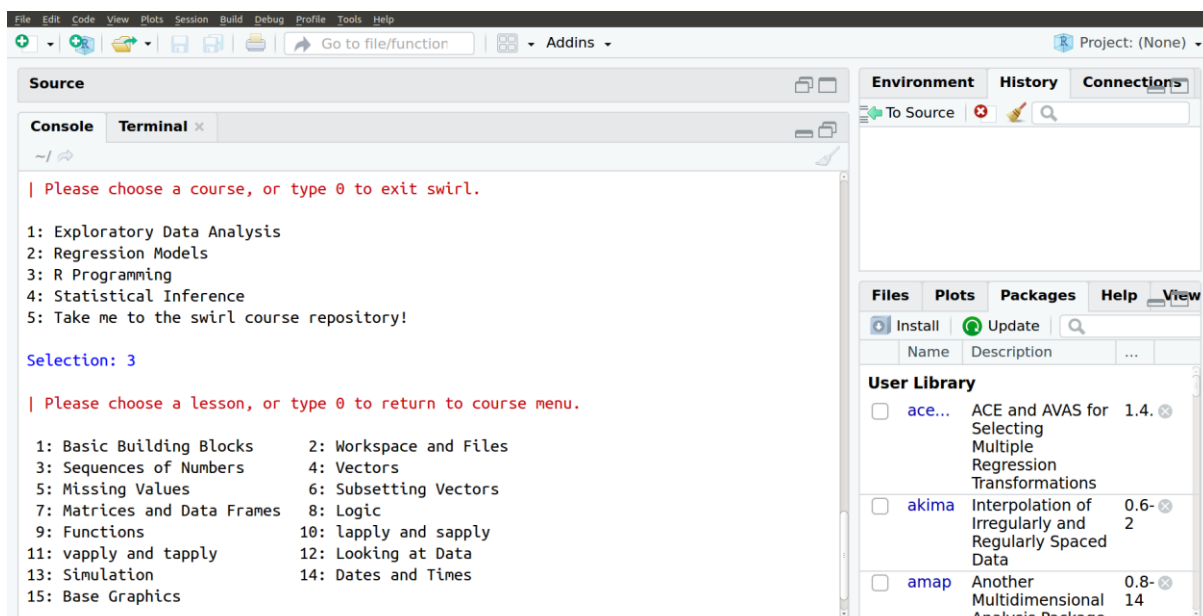
Home Learn Teach Contribute Blog FAQ Help



Learn R, in R.

swirl teaches you R programming and data science
interactively, at your own pace, and right in the R
console!

<https://swirlstats.com/>



 **Studio** Integrated Development Environment (IDE)

Step 1 - Install swirl package into R/RStudio:

```
install.packages("swirl")
```

Step 2 - Activate swirl package in R/RStudio:

```
library(swirl)
```

Step 3 - Install interactive workshops:

```
install_course("R Programming")
install_course("Exploratory Data Analysis")
install_course("Regression Models")
install_course("Statistical Inference")
```

Step 4 - Start interactive workshops:

```
swirl()
```

# "R Programming"	[Day 1]
# "Exploratory Data Analysis"	[Day 1 & 2]
# "Statistical Inference"	[Day 2]
# "Regression Models"	[Day 3 & 4 Optional]

Video Guides <https://github.com/telescopeuser/S-SB-Workshop>

NUS National University of Singapore | **ISS** INSTITUTE OF SYSTEMS SCIENCE

Home Executive Education ▾ Graduate Programmes ▾ Stackable Programmes ▾ Centres of Excellence ▾ Community ▾ About Us ▾ A- A A+ Q

NICF- Statistics Bootcamp

Overview

This course aims to train you to effectively use R in your statistical analysis for accurate and impactful data-driven decision making.

The Statistics Bootcamp is more than just an introduction to statistics, it teaches you how to use R, a command driven language for quick visual analytics. In addition, you will be presented with multiple business cases that will be solved in the course of lecturing and hands-on workshops.

The lecturers will give extensive illustrations of statistical principles and techniques. Each statistical technique taught will be demonstrated using R so that you can practise R programming in a series of workshops. In addition, you will get to dabble with data visualisation including creating graphs.

Ultimately, you will gain insights into how to use statistics to answer your business questions.

This course is part of the Analytics & Intelligent Systems Series and MTech EBAC Stackable foundation certificate

INSTITUTE OF SYSTEMS SCIENCE

Overview

Key Takeaways

Workshop 1.6 [Fundamental] R Programming

=====	
1: Basic Building Blocks	[Day 1A Data Processing]
2: Workspace and Files	[Day 1A Data Processing]
3: Sequences of Numbers	[Day 1A Data Processing]
4: Vectors	[Day 1A Data Processing]
5: Missing Values	[Day 1A Data Processing]
6: Subsetting Vectors	[Day 1A Data Processing]
7: Matrices and Data Frames	[Day 1A Data Processing]
8: Logic	[Day 1A Exercise]
9: Functions	[Optional]
10: lapply and sapply	[Day 1B Exercise]
11: vapply and tapply	[Optional]
12: Looking at Data	[Day 1B Data Processing]
13: Simulation	[Day 2B Sample & Norm]
14: Dates and Times	[Optional]
15: Base Graphics	[Day 1B Charting]
=====	

Workshop 2.6 [Fundamental] Exploratory Data Analysis

=====		
1: Principles of Analytic Graphs		[Optional]
2: Exploratory Graphs	[Day 1B Exercise]	
3: Graphics Devices in R [PDF PNG SVG]		[Optional]
4: Plotting Systems	[Day 2A Charting]	
5: Base Plotting System	[Day 2A Exercise]	
6: Lattice Plotting System		[Optional]
7: Working with Colors		[Optional]
8: GGPlot2 Part1 [qplot]	[Day 2A Charting]	
9: GGPlot2 Part2 [ggplot]	[Day 2A Charting]	
10: GGPlot2 Extras [qplot]		[Optional]
11: Hierarchical Clustering [Machine Learning]		[Optional]
12: K Means Clustering [Machine Learning]		[Optional]
13: Dimension Reduction [Machine Learning]		[Optional]
14: Clustering Example [Machine Learning]		[Optional]
15: CaseStudy [pm2.5 air pollution]		[Optional]
=====		

[Advanced] Statistical Inference

=====	
1: Introduction	[Day 2A Exercise]
2: Probability1 [dice, playing cards]	[Optional]
3: Probability2 [PMF PDF CDF]	[Optional]
4: ConditionalProbability [medic test]	[Optional]
5: Expectations [E(X), CLT]	[Day 2B Sample & Norm]
6: Variance [Var = $E(X^2) - (E(X))^2$]	[Optional]
7: CommonDistros [Bernoulli Normal Poisson]	[Optional]
8: Asymptotics [central limit theorem, CI]	[Optional]
9: T Confidence Intervals [nitty-gritty]	[Optional]
10: Hypothesis Testing [t-test, z score]	[Day 2B Exercise]
11: P Values [nitty-gritty]	[Day 2B Exercise]
12: Power [false negative, Type II error]	[Optional]
13: Multiple Testing [confusion matrix]	[Optional]
14: Resampling [bootstrap]	[Optional]
=====	

[Intermediate] Regression Models

=====		
1: Introduction	[regression to mean]	[Optional]
2: Residuals	[$\text{var}(\text{data}) = \text{var}(\text{estmt}) + \text{var}(\text{resdls})$]	[Optional]
3: Least Squares Estimation		[Optional]
4: Residual Variation	[$R^2 = 1 - s_{\text{Res}}/s_{\text{Tot}} = \text{cor}()^2$]	[Optional]
5: Introduction to Multivariable Regression		[Optional]
6: MultiVar Examples		[Optional]
7: MultiVar Examples2		[Optional]
8: MultiVar Examples3	[MultipleLinearRegression]	[Optional]
9: Residuals Diagnostics and Variation		[Optional]
10: Variance Inflation Factors	[VIF]	[Optional]
11: Overfitting and Underfitting	[ANOVA, F-test]	[Optional]
12: Binary Outcomes		[Optional]
13: Count Outcomes		[Optional]
=====		

More Interactive Workshops

Follow below for: Step 3 - Install interactive workshops:
 # Download ZIP [Download ZIP](#) file; save it to R working directory
 # <https://github.com/telescopeuser/S-SB-Workshop>

telescopeuser / S-SB-Workshop

3 commits 1 branch 0 releases 1 contributor GPL-3.0

Branch: master New pull request

Create new file Upload files Find file Clone or download

Gu Zhan initial

- Exploratory_Data_Analysis
- Getting_and_Cleaning_Data
- Mathematical_Biostatistics_Boot_Camp
- Open_Intro/Overview_of_Statistics
- R_Programming
- Regression_Models
- Statistical_Inference
- .gitignore
- LICENSE
- README.md

Clone with HTTPS Use SSH

Use Git or checkout with SVN using the web URL.

<https://github.com/telescopeuser/S-SB-Workshop>

Open in Desktop Download ZIP

File	Author	Time
R_Programming	initial	an hour ago
Regression_Models	initial	an hour ago
Statistical_Inference	initial	an hour ago
.gitignore	initial	an hour ago
LICENSE	initial	an hour ago
README.md	initial	an hour ago

```
install_course_zip("S-SB-Workshop-master.zip", multi=TRUE)
```

Step 4 - Start interactive workshops:

```
swirl()
```

```
[Intermediate] Open Intro
=====
1: Overview of Statistics [Optional]
=====
```

```
[Intermediate] Mathematical Biostatistics Boot Camp
=====
1: One Sample t-test [nitty-gritty] [Optional]
2: Two Sample t-test [nitty-gritty] [Optional]
3: Errors Power and Sample Size [Optional]
=====
```

```
[Advanced] Getting and Cleaning Data
=====
1: Manipulating Data with dplyr [Optional]
2: Grouping and Chaining with dplyr [Optional]
3: Tidying Data with tidyr [Optional]
4: Dates and Times with lubridate [Optional]
=====
```