

# Learning R

## From an undergraduate perspective

University of Glasgow, March 2018  
Moa Schafer

# Introduction

- 3rd year Undergraduate Student
- Psych & Business/Management
- New to R prior to Psychology



# Talk Overview

- The Change to R and the Benefits
- Practice: Weekly Exercises
- Support: Slack
- Resources: Explore!

# Talk Overview

- **The Change to R and the Benefits**
- Practice: Weekly Exercises
- Support: Slack
- Resources: Explore!

# The Change to R

1st year: SPSS



2nd year: Labs using R



3rd year: Weekly R sessions



# The Benefits

- Available for free
- Transferable skills
- Interactive approach of analysing
- More comprehensive understanding of stats



# Talk Overview

- The Change to R and the Benefits
- **Practice: Weekly Exercises**
- Support: Slack
- Resources: Explore!

# Practice: Weekly Homework Exercises

- Statistics Lecture
- Homework Exercises in R
- Feedback on Homework
- Interactive Sessions





# Practice: Weekly Homework Exercises

## Submit your homework

<http://talklab.psy.gla.ac.uk/L3stats/ps02>

## Correlation matrices using stat software

Click [here](#) to download the data for this exercise.

This (real!) dataset contains activity on the social media site twitter.com for 50 users, whose activity was measured during a month in 2014. Each row of data is for a different Twitter user. The variables are:

### Variable Description

n\_su Number of status updates posted  
fol\_chg Number of followers the user gained over the month  
fav\_chg Number of tweets the user "liked" over the month  
fri\_chg Number of "friends" (people the user started following) over the month

Create a correlation matrix for all bivariate relationships (use Spearman correlations) and then answer the following questions.

How many unique bivariate relationships are represented in the matrix (not including variables correlated with themselves)?



## Entering data into R

- For now, we will just type the data into Excel and save as csv
  - ▶ CSV = comma-separated values
- Alternative: just type into a text file as shown below
- Then import into R using the read.csv() function
- See [Appendix](#) for how to type it directly into R

### How data.csv looks

```
"SubjID", "Cond", "Mood", "SelfEst", "PosOut"  
1, "E", 75, 52, 65  
2, "E", 62, 69, 65  
...  
19, "C", 58, 56, 51  
20, "C", 50, 74, 27
```

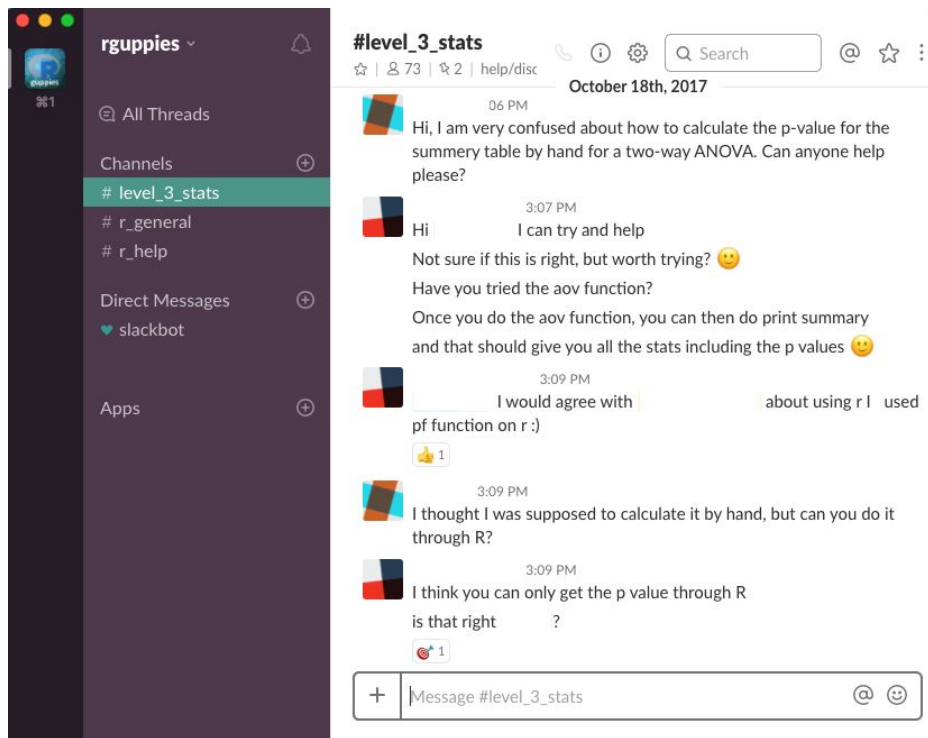
10/22



# Talk Overview

- The Change to R and the Benefits
- Practice: Weekly Exercises
- **Support: Slack**
- Resources: Explore!

# Support: Slack



The screenshot shows a Slack interface with a sidebar on the left and a main chat area on the right. The sidebar includes a profile for 'rguppies', 'All Threads', a list of channels with '# level\_3\_stats' selected, 'Direct Messages' with 'slackbot', and 'Apps'. The main chat area is titled '#level\_3\_stats' and shows a conversation from October 18th, 2017. The messages are as follows:

- 06 PM: A user asks, "Hi, I am very confused about how to calculate the p-value for the summery table by hand for a two-way ANOVA. Can anyone help please?"
- 3:07 PM: A user replies, "Hi I can try and help Not sure if this is right, but worth trying? 😊 Have you tried the aov function? Once you do the aov function, you can then do print summary and that should give you all the stats including the p values 😊"
- 3:09 PM: A user replies, "I would agree with [redacted] about using r I used pf function on r :)" with a thumbs up icon and '1'.
- 3:09 PM: A user replies, "I thought I was supposed to calculate it by hand, but can you do it through R?"
- 3:09 PM: A user replies, "I think you can only get the p value through R is that right ?" with a thumbs up icon and '1'.

The bottom of the chat area shows a message input field with a plus sign on the left, the text "Message #level\_3\_stats", and a send button with a smiley face icon on the right.

- Variety of online channels
  - Guidance from staff
  - Peer support
- Plus: UofG PAL  
(Peer Assisted Learning) scheme



Cloud-based chat & communication tool

# Support: Slack

Channels

# level\_3\_stats

# r\_general

# r\_help

#level\_3\_stats

☆ | 👤 73 | 🔔 2 | help/discussion for PSYC4037

November 3rd, 2017

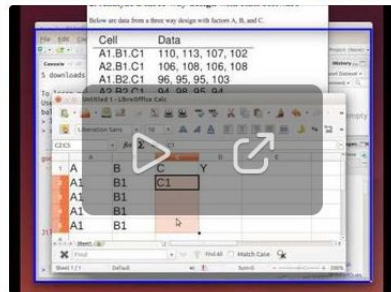


2:40 PM

@channel Hi everyone, I just posted a video walkthrough for homework 5, useful for anyone struggling with <https://youtu.be/AJDo9gpkEcg>

YouTube datahowler

Developing an analysis in R/RStudio: L3 stats homework 5 ▾



2 replies Last reply 4 months ago

#r\_help

☆ | 👤 725 | 🔔 3 | This is for help on R/RStudio. Open to all levels/classes

Friday, January 19th



10:41 AM

How should a categorical factor with four levels be coded? I understand dummy and deviation coding for two levels, but am lost for factors with more than two.



3:20 PM

Dummy coding would look like this (category d is the reference category):

```
cat4 cat_a cat_b cat_c
a 1 0 0
b 0 1 0
c 0 0 1
d 0 0 0
```

(edited)



1 reply 2 months ago



3:29 PM

if you create dummy codes and then center each of the variables (subtract the means) then you end up with deviation codes, which is important if you have interactions

# Talk Overview

- The Change to R and the Benefits
- Practice: Weekly Exercises
- Support: Slack
- **Resources: Explore!**

# Resources: Explore!


## Resources for learning R

- Dale's R/RStudio walkthrough for newbies (website, includes info on installing, and links to videos!)
- swirl
- codeschool
- datacamp

### Lab 1 Preparatory Material

#### Table of Contents

- Video walkthrough
  - Part 1: Introduction to R and RStudio
  - Part 2: Library (RMarkdown)
- What is R?
  - The R language
  - The RStudio Integrated Development Environment (IDE)
- Installing R
  - Installing Base R
  - Installing RStudio
  - Additional resources you might want to try
- Identifying R packages to install
  - Conference Schedule for Maximum Reproducibility
  - Reproducible data science with RStudio and RMarkdown
- Writing R commands
  - R scripts are like R as a calculator
  - Showing results in a variable
  - Whitespace
  - The workspace
  - Simple and vectorized calculations
  - Getting feedback
- Add-on packages
  - Getting help
  - Showing directory, session environment, and reading/writing data files
  - Calling functions
  - R function syntax
- Some basic data types
  - Character vectors
  - Lists
  - Matrices
  - Data frames
- Learning R




Learn R, in R.

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


[Follow @swirls](#)

Get questions? Join our [discussion group!](#)

Try it! Sponsored By: **O'REILLY** Created By: **Code School**



R is a tool for statistics and data modeling. The R programming language is elegant, versatile, and has a highly expressive syntax designed around working with data. R is more than that, though – it also includes extremely powerful graphics capabilities. If you want to easily manipulate your data and present it in compelling ways, R is the tool for you.



FREE COURSE

## Introduction to R

[Start Course For Free](#)

4 hours | 10 videos | 62 Exercises | 743,059 Participants | 4,300 XP

[DOWNLOAD THE APP](#)

This course is part of these tracks:

- Data Analyst with R
- Data Scientist with R
- R Programmer
- R Programming

# Resources: UofG Psychology Resources

## Lab 1 Preparatory Material

### Table of Contents

- [Video walkthroughs](#)
  - [Part 1: Introducing R and RStudio](#)
  - [Part 2: Using RMarkdown](#)
- [What is R?](#)
- [Interacting with R](#)
  - [The Base R Console](#)
  - [The RStudio Integrated Development Environment \(IDE\)](#)
- [Installing R](#)
  - [Installing Base R](#)
  - [Installing RStudio](#)
  - [Additional tweaks you might want to try](#)
- [Developing Reproducible Scripts](#)
  - [Configure RStudio for Maximum Reproducibility](#)
  - [Reproducible reports with RStudio and RMarkdown](#)
- [Typing in commands](#)
  - [Warming up: Use R as a calculator](#)
  - [Storing results in a variable](#)
  - [Whitespace](#)
  - [The workspace](#)
  - [Vectors and vectorized calculations](#)
  - [Calling functions](#)
- [Add-on packages](#)
- [Getting help](#)
- [Working directory, session environment, and reading/writing data files](#)
- [Calling functions](#)
  - [Function syntax](#)
- [Some basic data types](#)
- [Container types](#)

### Some basic data types

Below are a list of different data types in R.

type	description	example
double	floating point value	.333337
integer	integer	-1, 0, 1
numeric	any real number (int,double)	1, .5, -.222

## Video walkthroughs

Here are video walkthroughs introducing R and RStudio. You may find it best to watch these videos first, and then consult the text in this document as needed to fill in any gaps in your knowledge.

### Part 1: Introducing R and RStudio

- The R script generated in this part is here: [scottish\\_babynames.R](#)

Developing an analysis using R and RStudio:  
Scottish babynames

Dale Barr

*Psychology*  
*University of Glasgow*

▶ 0:00 / 39:41 🔊 🔍 ⬇

- Fantastic resource to get started
- Learning the basics
- Video + text

# Resources: The Internet



Introduction to R

Complete the script to produce the output shown

## OUTPUT

```
[1] apple apple apple banana banana
Levels: apple banana
```

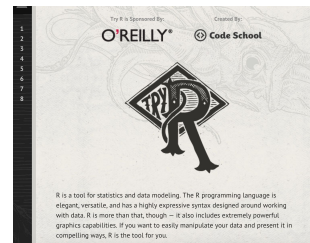
## SCRIPT

```
x <- c("a", "a", "a", "b", "b")
y <-  (x)
levels(y) <- c("apple", "banana")
y
```

Try Code (3)

Check

Lots of interactive online resources available





# Conclusion

Practice  **Studio**<sup>TM</sup> Weekly Exercises

Support  **slack** + PAL (Peer Assisted Learning)

Resources  **moodle** +  **DataCamp** + more!

Thank you for listening!