

# ECS3356 Applied Computational Finance

Sylvia Gottschalk

Lecture 17: Introduction to R



- 1 Outline
- 2 R Studio
- 3 Football leagues
- 4 Russian Twitter trolls
- 5 Basic commands in R

- 1 Outline
- 2 R Studio
- 3 Football leagues
- 4 Russian Twitter trolls
- 5 Basic commands in R

# Outline

- Posit and R Studio
- Exercises: Football Analytics and Russian Twitter trolls
- Basic R commands

- 1 Outline
- 2 R Studio**
- 3 Football leagues
- 4 Russian Twitter trolls
- 5 Basic commands in R

# Getting RStudio

- 1 If you want to install R on your laptop

Download the R software from:

<https://www.r-project.org/>

Download RStudio from:

<https://posit.co/>

- 2 If you are using a tablet or are on the go, use R Cloud:

<https://posit.cloud>

During the workshops we will use R Cloud, and Google Colab for Python.

# Posit Cloud

The screenshot displays the Posit Cloud web interface for a user named Sylvia Gottschalk. The browser address bar shows the URL `https://posit.cloud/content/yours?sort=name_asc`. The interface includes a navigation bar with 'Your Workspace' and tabs for 'Content', 'Usage', and 'About'. On the left, a sidebar shows 'Your Content' with sub-items 'Archive' and 'Trash'. The main area, titled 'Your Content (3)', lists three RStudio Projects: 'ECS3356', 'LSBS', and 'tests'. Each project entry includes an R logo, the name 'RStudio Project', a lock icon indicating it is 'Private', and the creation date. To the right of each project name are icons for adding, deleting, downloading, and sharing. At the top right of the content area is a 'New Project' button. A filter bar at the top of the content list allows for filtering by 'TYPE', 'ACCESS', and 'SORT' (currently set to 'A Z').

← → ↻ [https://posit.cloud/content/yours?sort=name\\_asc](https://posit.cloud/content/yours?sort=name_asc) ☆ 🔒 ⬇️ 👤 📁

☰ Your Workspace Sylvia Gottschalk Content Usage About SG Sylvia Gottschalk

👤 Your Content Your Content (3) New Project ▾

📁 Archive TYPE \* ACCESS \* SORT A Z 🔍

🗑️ Trash

**ECS3356** RStudio Project Private Created Jan 28, 2023 8:26 PM

**LSBS** RStudio Project Private Created Feb 26, 2023 1:34 PM

**tests** RStudio Project Private Created May 1, 2023 2:53 PM

# RStudio in Posit Cloud

≡ Your Workspace / ECS3356

RAM ⚙️ 🔍 SG Sylvia Gottschalk

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

R 4.2.3

ECS3356\_L17\_consumer.Rmd

Source Visual

```
1 ---
2 title: "ECS3356 L17 2024: Consumer behaviour"
3 author: "sylvia gottschalk"
4 date: "2024-01-25"
5 output:
6   html_document: default
7   pdf_document: default
8 ---
9
10 ```{r echo=F, results='hide', quietly=T}
11 #install.packages("forecast")
12 ```
13
14 ## Load the libraries
```

11:2 Chunk 1

R Markdown

Environment History Connections Tutorial

R - Global Environment

Environment is empty

Files Plots Packages Help Viewer Presentation

Folder Blank File Upload Delete Rename

Cloud project

Name	Size	Modified
..		
rhistory	0 B	Jan 28, 2023, 8:27 PM
Applied_computational_finance_get...	7.2 KB	May 9, 2023, 3:10 PM
Applied_computational_finance_get...	7.5 KB	May 2, 2023, 1:36 PM
Applied_computational_finance_get...	9.3 KB	Mar 30, 2023, 10:38 AM
consumer_behaviour.csv	223.1 KB	Jan 25, 2024, 2:03 PM
correlation1plot.jpeg	192.8 KB	Mar 30, 2023, 10:50 AM
correlation2plot.jpeg	181.7 KB	Mar 30, 2023, 10:50 AM
correlation3plot.jpeg	186.2 KB	Mar 30, 2023, 10:50 AM
correlation4plot.jpeg	192.3 KB	Mar 30, 2023, 10:50 AM
correlation5plot.jpeg	187.5 KB	Mar 30, 2023, 10:50 AM

Console Terminal Background Jobs

```
R 4.2.3 ~/cloud/project/
> library(rmgarch)
> library(xts)
> library(zoo)
> library(scales)
> library(yfrr)
Session restored from your saved work on 2024-Jan-25 20:04:51 UTC (1 hour ago)
```



# packages

Install a package: `install.packages("tidyverse")`

- Load a package: `library(tidyverse)`
- Load data from a package: `data(iris)`

- 1 Outline
- 2 R Studio
- 3 Football leagues**
- 4 Russian Twitter trolls
- 5 Basic commands in R

# Football infographics

<https://fbref.com/en/> is a website that collates statistics about worldwide football leagues:

- leagues
- matches: scores, away/home, ranking
- players: passes, expected passes, position, team

The exercise consists of:

- 1 uploading data about different leagues
- 2 create a plot showing passes vs expected passes and a plot showing assists vs expected assists.
- 3 creating an Rmarkdown file from the R code
- 4 creating a pdf and an html file from the RMarkdown code

## Exercise: Football Analytics

- 1 Download from MyLearning Week 17 the files `epl_player_summary.csv`, `laliga_player_summary.csv`, `ligue1_player_summary.csv`, `seriea_player_summary.csv`; and the code file `ECS3356_L17_Football.R`

- 2 Upload all these files into Posit

`ECS3356_L17_Football.R` creates the plots for EPL. Create the same plots for La Liga, Bundesliga, Serie A, and Ligue 1.

- 1 Outline
- 2 R Studio
- 3 Football leagues
- 4 Russian Twitter trolls**
- 5 Basic commands in R

## 3 million Russian disinformation campaign tweets

A Github directory contains data on nearly 3 million tweets sent from Twitter accounts connected to the Internet Research Agency, the Russian “troll factory” that interfered in the 2016 US election.

The tweets in this database were sent between February 2012 and May 2018, with the vast majority posted from 2015 through 2017.

Source: <https://github.com/fivethirtyeight/russian-troll-tweets>

# Twitter trolls exercise

- Download from MyLearning Week 17 the files:  
IRAhandle\_tweets\_1.csv and ECS3356\_L17\_Russian\_Twitter.r
- Upload all these files onto Posit
- Run the R code
- Create an Rmarkdown file from the R code
- Create a pdf and an html file from the RMarkdown code

# Twitter trolls code

The code produces tables answering the questions:

- ① What are the most frequent and second-most frequent languages?
- ② On average, how many followers did each tweet reach in each region?
- ③ How many tweets are retweets in each language?
- ④ How many tweets are not retweets in each language?
- ⑤ How often are “Trump” and “Clinton” mentioned in the tweets?



- 1 Outline
- 2 R Studio
- 3 Football leagues
- 4 Russian Twitter trolls
- 5 Basic commands in R

# Basics

Calculate:  $(2 + 5) / 7$

- Create a variable:  
`var_1 ← c(1, 4, 5, 7)`
- Create another variable:  
`var_2 ← c(1, 15, 25, -1)`
- Elements of a vector: `var_1[1:3]`
- Arithmetic with variables:  
`var_1 * var_2`  
`var_1 * 2`
- Create a data.frame
- `D ← data.frame(var_1, x c(1, 2, 3, 4))` `D ← data.frame(var_1, c(1, 2, 3))`

# Dataframe commands

Variables in dataset D: `names(D)`

- Index variable within a data.frame:  
`D$var1[2:4]`
- Create new variable within existing data.frame:  
`D$new_var ← c(100, 200, 300, 400)`
- Basic sampling functions:  
`sample(1:6, 10)`  
`rnorm(10, mean = 5, sd = 0.25)`

# Next week

- data visualisation in finance
- create a blog in Wix or Wordpress