Recitation 1: Introduction to Research Methods for Politics

Dept. of Politics, NYU

POL-850

Fall 2020

Agenda

- Logistics
- Rstudio: an introduction
- Basic functions: length(), mean(), max(),...
- Rmarkdown: how generate a beautiful and professional document

Logistics

- Our Friday section: 11:00 am - 12:15 pm;
- Remote (Online) for the whole semester
- Attendence?
- Homeworks and exams
- Find all material:
  https://github.com/Jiawei-Fu/UA850-FALL2020

Office Hours

Office hours:

- TBD
- Ideally, tell me in advance via email, so I can have an idea of how many you are and give you a precise time to connect.

Expectations

My expectations of you:

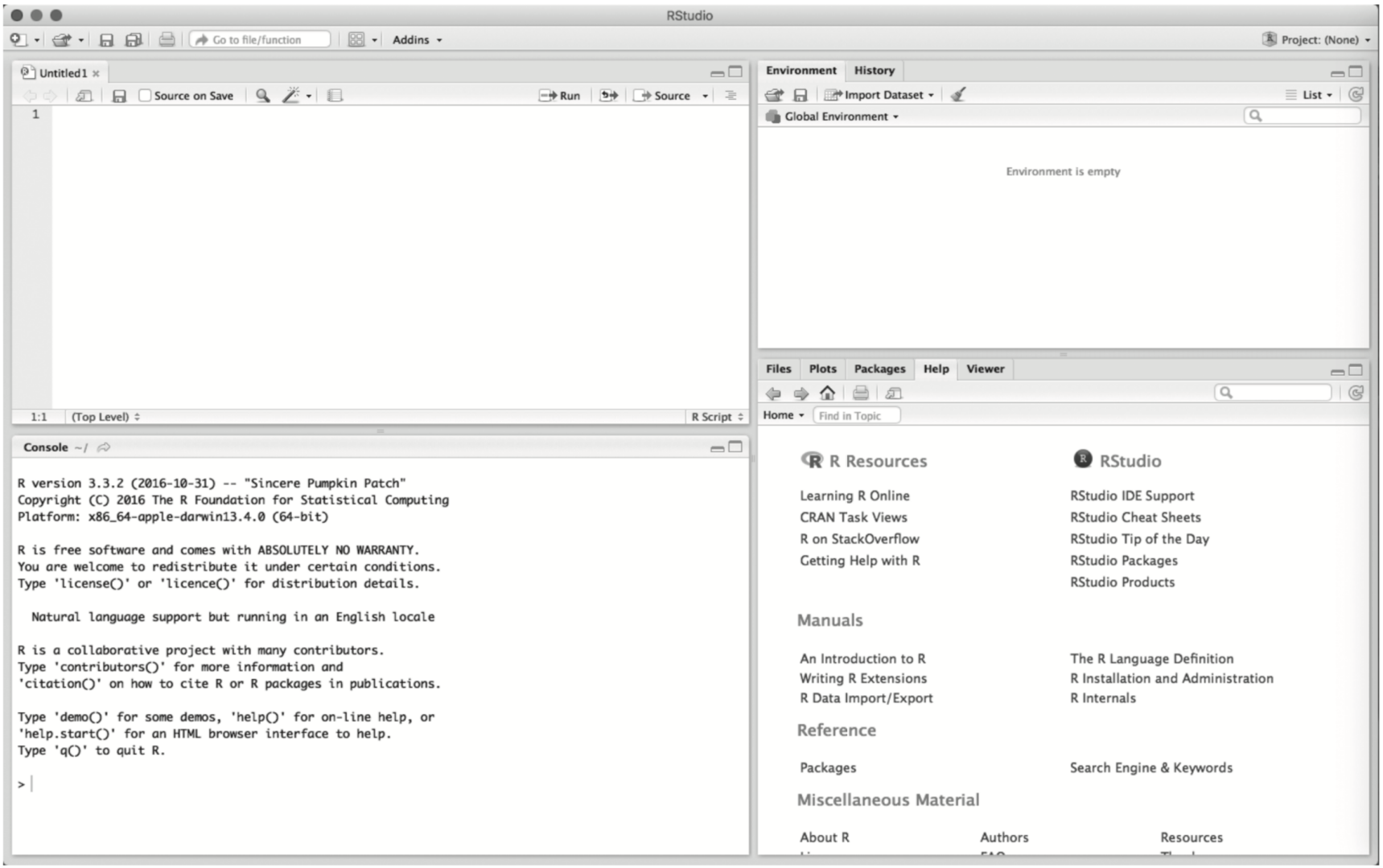
1. Attend both class and recitation
2. Turn in homework on-time
3. Ask me questions early and often
4. Check out the textbook for additional insights

Introduction to R Studio

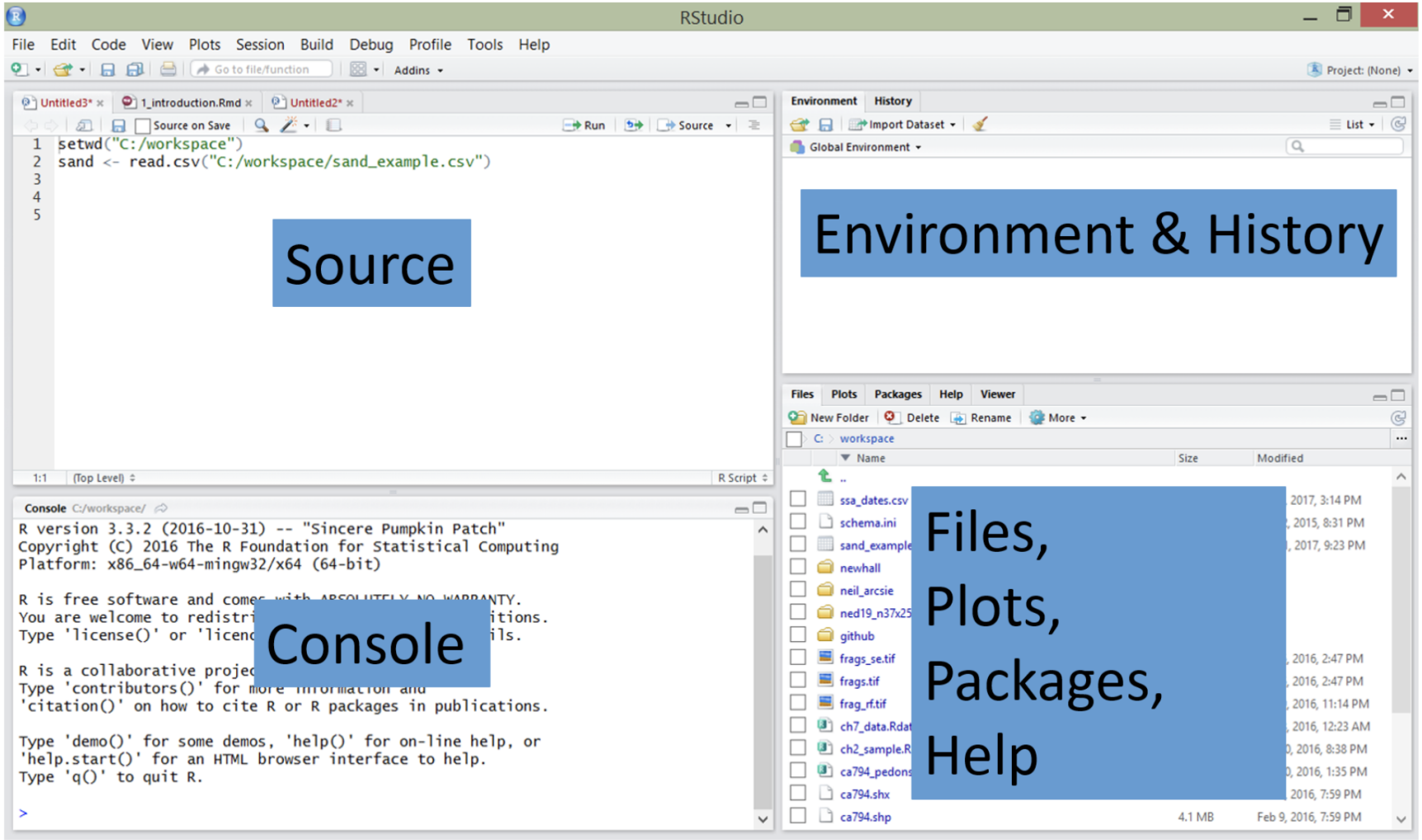
What is R Studio?

- Open-source program that facilitates the use of R
- For your homework, you can save your R codes in R Studio
- You can also generate PDF files with R codes and results

What is R Studio?



What is R Studio?



Accessing R Studio: Jupyter Hub

Getting into Jupyter Hub

This semester, we give all of you the possibility to use R Studio online instead of setting it up on your own machine.

- We use a platform called Jupyter Hub, where a page dedicated to POL-850 is available for you
- To access it: login at `https://polua-850-fall.rcnyu.org/hub/login`

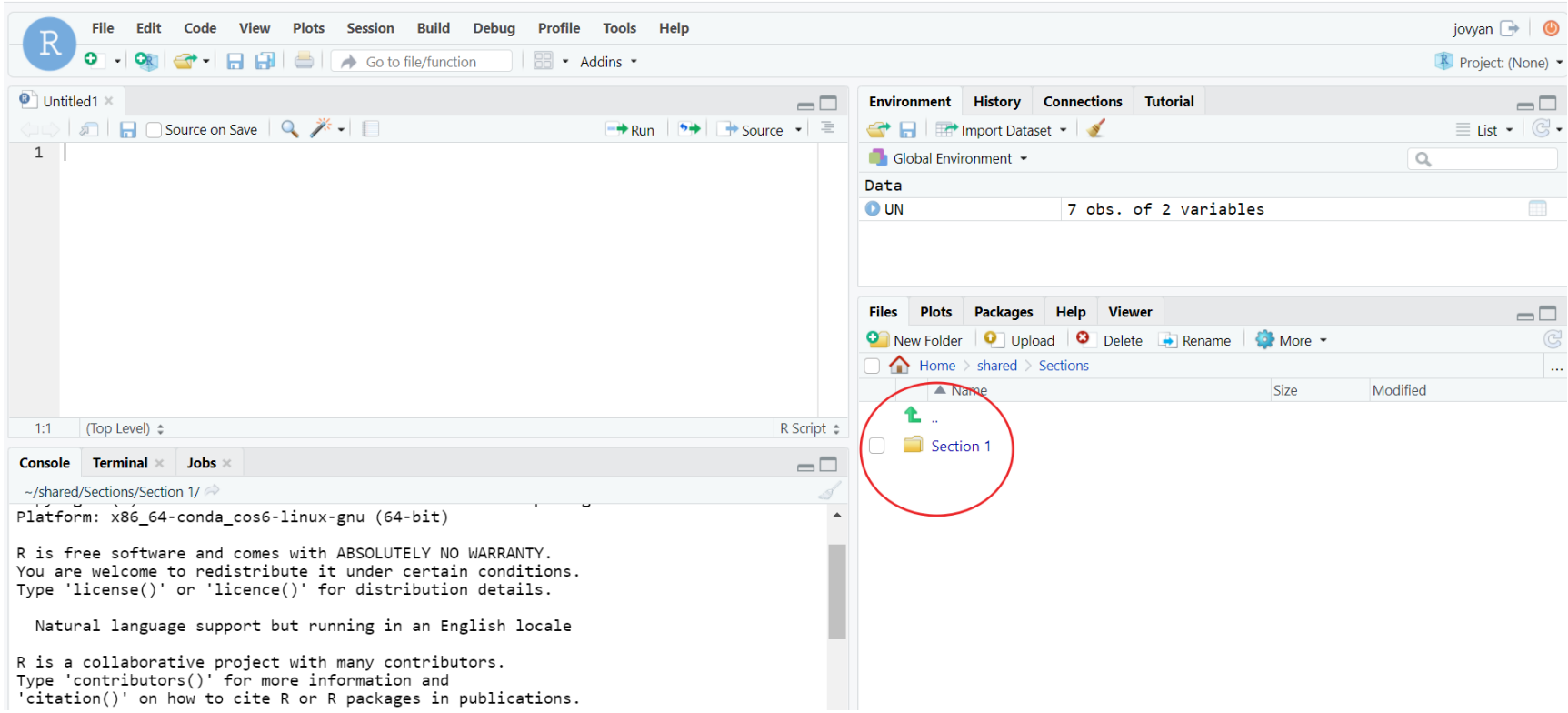
Getting into Jupyter Hub

Once you are there:

- Click on the orange button "Sign in with New York University"
- Click on the orange button "Spawn" and wait for the bar to charge
- The RStudio server opens up directly, and you can work from it!

What You Should See

All the files you need will be made available, find them in the relevant folder in the bottom right panel!



Basic R codes

Arithmetic Operations

```
5 + 3
```

```
## [1] 8
```

```
5 - 3
```

```
## [1] 2
```

```
5 / 3
```

```
## [1] 1.666667
```

```
5 ^ 3
```

```
## [1] 125
```

```
5 * (10 - 3)
```

```
## [1] 35
```

```
sqrt(4)
```

```
## [1] 2
```

Objects

```
> result <- 5 + 3
> result
[1] 8
> print(result)
[1] 8
> result <- 5 - 3
> result
[1] 2
>
> alex <- "instructor"
> alex
[1] "instructor"
> alex <- "one of the instructors"
> alex
[1] "one of the instructors"
```

Data Files

- R will load data from and save data to your **working directory**
- To check you current working directory, `getwd()`
- To reset your working directory, `setwd()`

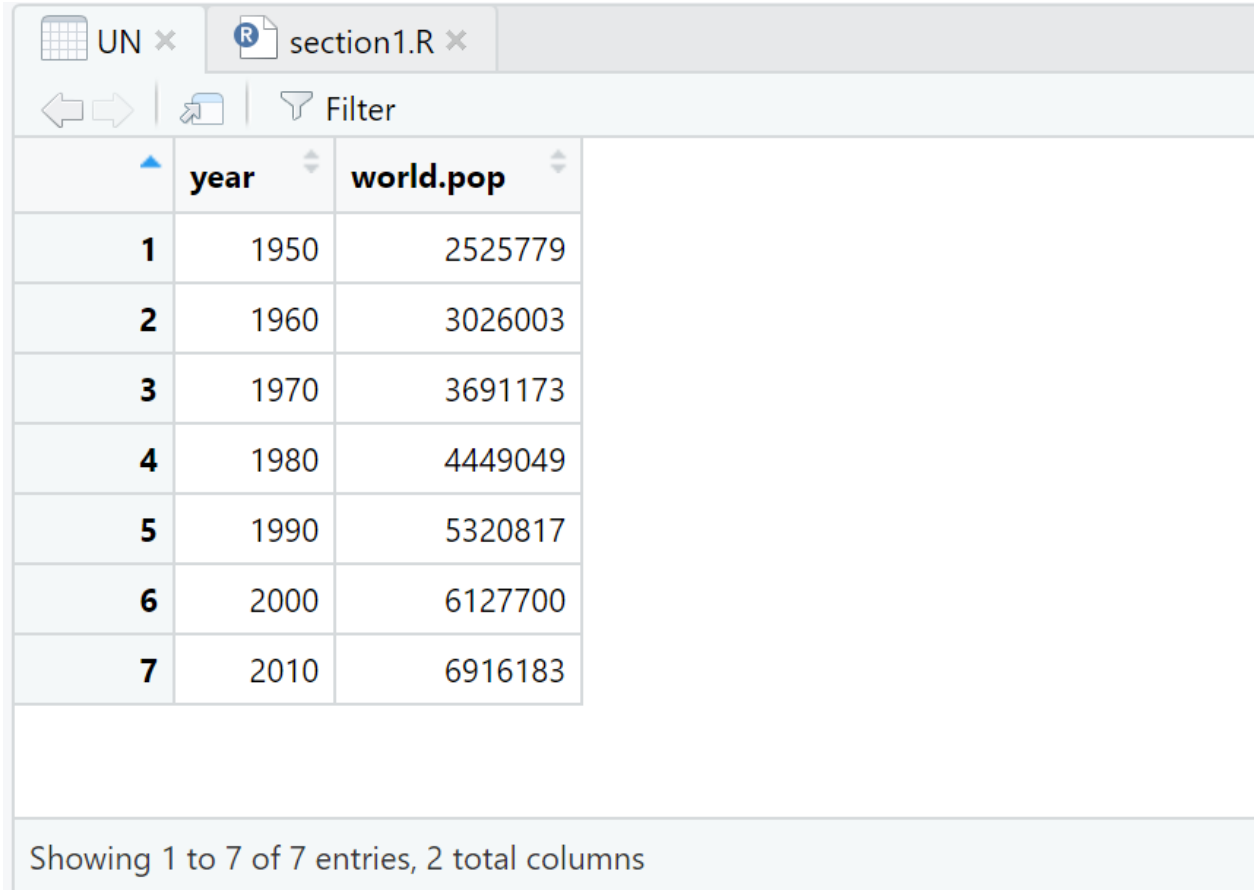
Data Description

Some basic commands to make sense of your data:

```
# Load data #
UN <- read.csv("UNpop.csv")
class(UN)
] "data.frame"
names(UN) # look at variable names
] "year"       "world.pop"
nrow(UN) # n. of observations
] 7
ncol(UN) # n. of variables
] 2
dim(UN) # dimensionality of the data set
] 7 2
```

A Data Set: Rows and Columns

To look at the whole table in Excel-style format, use the `View()` command.



	year	world.pop
1	1950	2525779
2	1960	3026003
3	1970	3691173
4	1980	4449049
5	1990	5320817
6	2000	6127700
7	2010	6916183

Showing 1 to 7 of 7 entries, 2 total columns

Data at a Glance: `summary()`

The `summary()` command provides summary statistics of all the variables in your data set:

```
> summary(UN) # summary statistics
      year          world.pop
 Min.   :1950   Min.   :2525779
 1st Qu.:1965   1st Qu.:3358588
 Median :1980   Median :4449049
 Mean   :1980   Mean   :4579529
 3rd Qu.:1995   3rd Qu.:5724258
 Max.   :2010   Max.   :6916183
```

Working on Variables

The key to call one or more variables in your data is to use the $ symbol after the name of the data set:

```
# Call a specific variable: use the $ sign!
UN$year
] 1950 1960 1970 1980 1990 2000 2010
UN$world.pop
] 2525779 3026003 3691173 4449049 5320817 6127700 6916183

head(UN$year)
] 1950 1960 1970 1980 1990 2000
```

Working on Variables

We can then transform and combine variables, and add these transformations to our original data set using the assignment arrow:

```
## Create a new variable
UN$unit.pop <- (UN$world.pop) * 1000

head(UN$unit.pop) # a new variable now in our data!
] 2525779000 3026003000 3691173000 4449049000 5320817000
```

Working on Variables

Oftentimes, we will need to compute ratios and percentage ratios. We just divide the relevant variables and multiply the result by 100:

```
## Percentage ratios

UN$popratio.mean <- UN$world.pop/mean(UN$world.pop)*100
head(UN$popratio.mean)
]  55.15368  66.07673  80.60158  97.15080 116.18699 133.80633
```

Functions

Built-in functions

```
length(UN$world.pop) # n. elements of a variable (column)
] 7
min(UN$world.pop)  # lowest value of a variable
] 2525779
max(UN$world.pop)  # highest value of a variable
] 6916183
range(UN$world.pop)  # range of a variable (max - min)
] 2525779 6916183
mean(UN$world.pop)   # mean of a variable
] 4579529
```

Saving Your Work

- Once you are done and satisfied with your progress, you can save the new version of your script
- Using the drop-down menu on the top is the easiest way
- Clicking sequence: *File* > *Save* As...
- And then just choose the location and the name you wanna give to your script

Introduction to R Markdown

Introduction to R Markdown

Why do you need R markdown?

- When you submit your homework, you need to submit both your R codes *and* results
- R markdown automatically does that for you
- Rather than copying and pasting R codes and results in Words, just use R markdown!

Introduction to R Markdown

If you use R Markdown, you can generate a PDF file like ...

Code for QSS Chapter 1: Introduction

Kosuke Imai

First Printing

Section 1.1: Overview of the Book

Section 1.2: How to Use this Book

```
install.packages("swirl") # install the package
library(swirl) # load the package
install_course_github("kosukeimai", "qss-swirl") # install the course
library(swirl)
swirl()
```

Section 1.3: Introduction to R

Section 1.3.1: Arithmetic Operations

```
5 + 3
```

```
## [1] 8
```

Introduction to R Markdown

R Markdown file template will be provided

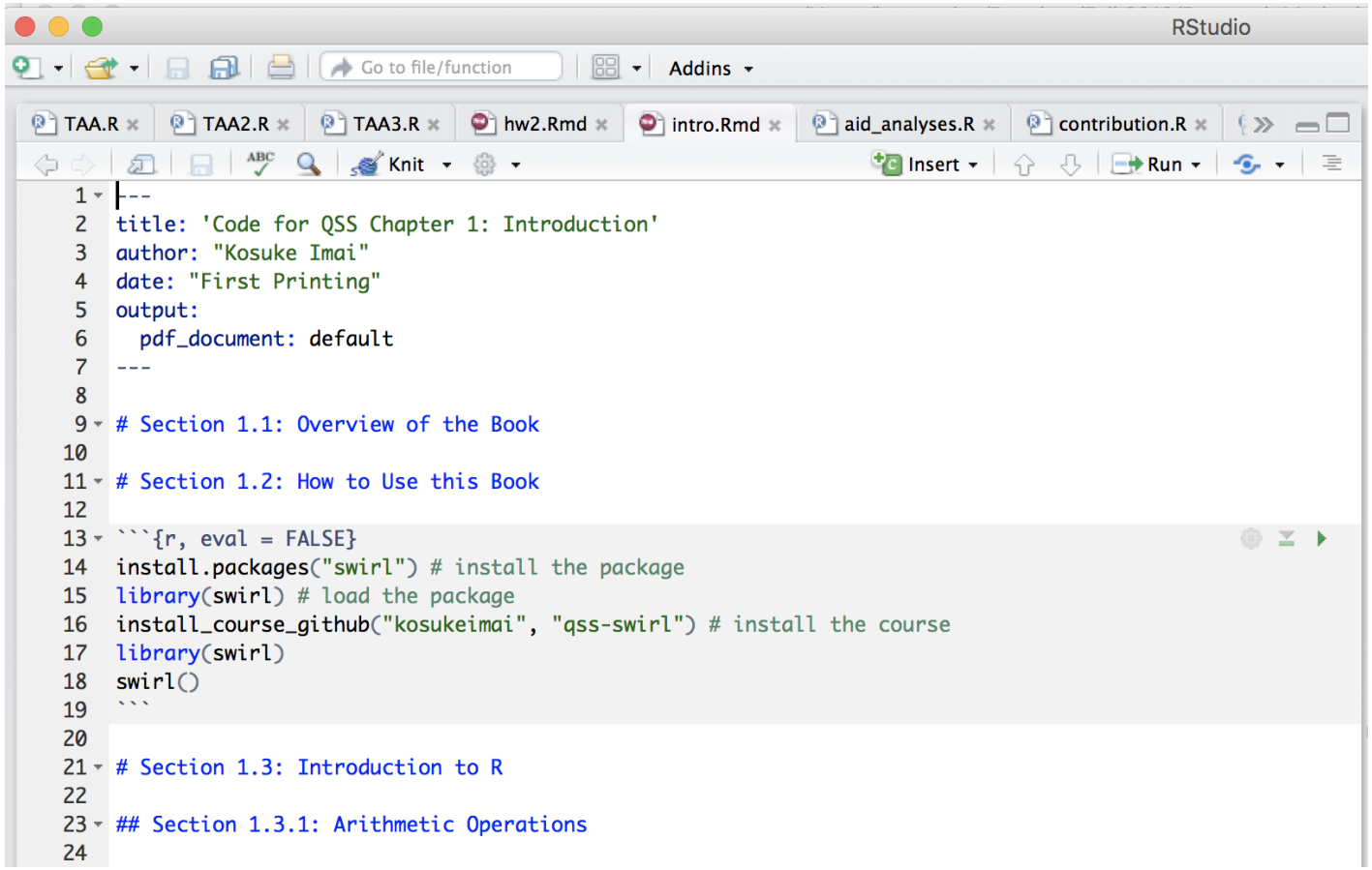
Elements of R Markdown file

1. Document information
2. Chunks of R code

Introduction to R Markdown

Document information (title/ author/ date/ output)

- R markdown file is saved as .Rmd, whereas R code file is saved as .R



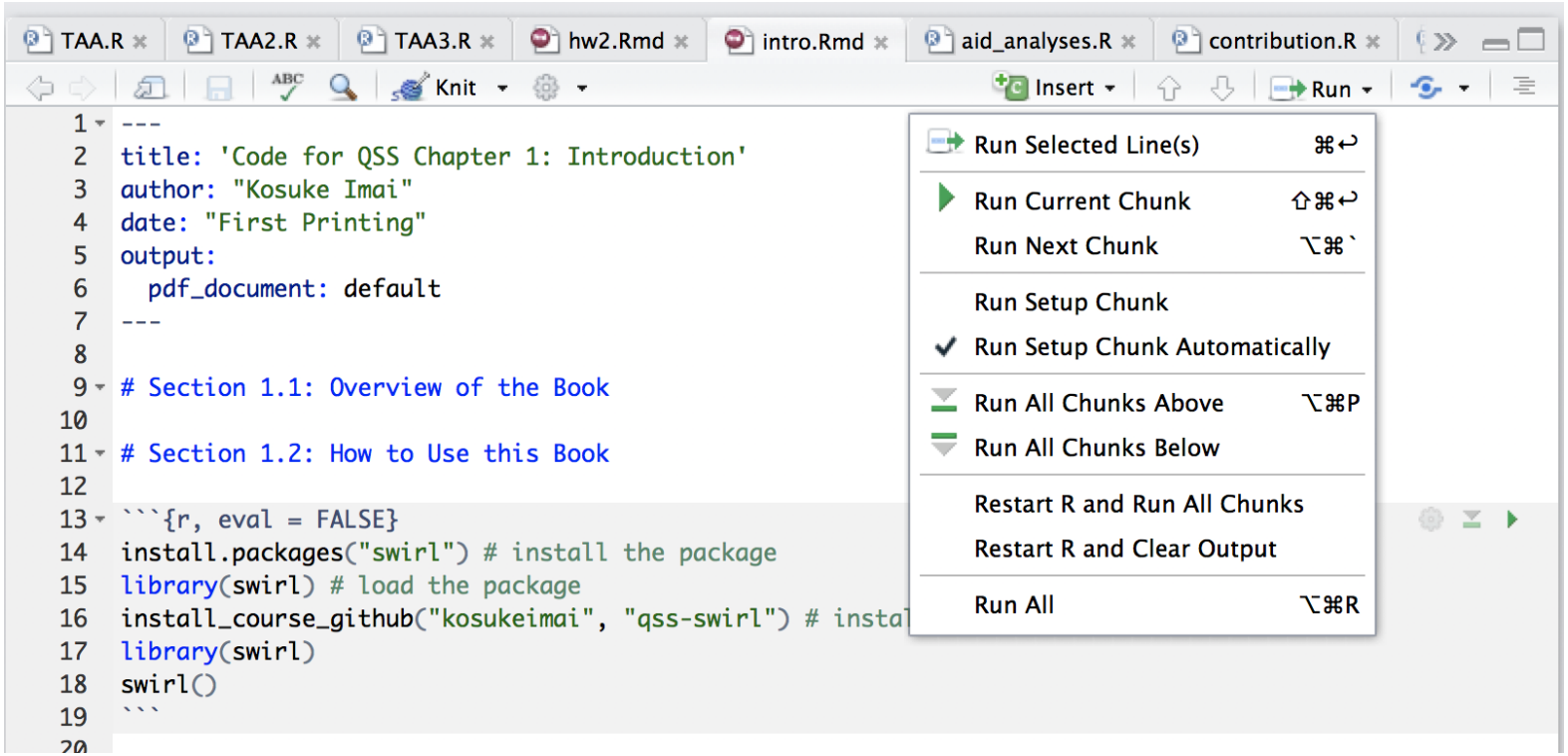
Introduction to R Markdown

Chunks of R code

````
23 ## Section 1.3.1: Arithmetic Operations
24
25 ```{r}
26 5 + 3
27 5 - 3
28 5 / 3
29 5 ^ 3
30 5 * (10 - 3)
31 sqrt(4)
32 ```
````

- The chunk should start with ‘‘‘{r } and end with ‘‘‘
- Make sure that each chunk has different names (*e.g.* r first, r second ...)
- To add texts (not R code), use #

Introduction to R Markdown



- To check whether your R codes run well, click Run button on the upper-right side of source window
- To generate pdf file from Rmd file, click Knit button