Contents

[2.1 Video: 1](#_Toc57586161)

[Project, path, data formats basics 15:00 1](#_Toc57586162)

[Packages 23:21 2](#_Toc57586163)

[Reading delimited data 33:36 3](#_Toc57586164)

[2.2 Video: 3](#_Toc57586165)

[Excel, SAS, SPSS questions 7:25 3](#_Toc57586166)

[JSON, API questions 16:22 4](#_Toc57586167)

[Databases 30:13 4](#_Toc57586168)

[Homework Notes 4](#_Toc57586169)

# 2.1 Video:

## Project, path, data formats basics 15:00

What is a path to a file?

* The location of a file on your computer (or a URL).
* The method in which the file was created.
* The method for accessing the metadata about the file.
* The destination for your output to the console.

Feedback

* Right!
* The path to the file correspond to the location of the file. This is generally a location on your computer or at a URL.
* The path to the file correspond to the location of the file. This is generally a location on your computer or at a URL.
* The path to the file correspond to the location of the file. This is generally a location on your computer or at a URL.

Why do we prefer to use a “local path” to a file rather than a “full path” to a file?

Feedback

When using a full path (such as “C:/…/file.ext”), sharing code is more difficult. The person you share the code with would need to edit the code. When using a local path (corresponding to current R session’s working directory), code can often be run without changing anything.

What function can help you determine your working directory? Simply put the name of the function or the name of the function followed by (). For example, if the answer were 'hist' you can put hist or hist().

getwd, getwd()

Feedback

None

## Packages 23:21

Check all of the following that will read an R package into your session.

* library
* require
* install
* read.package
* use
* use.package

Feedback

None

Installing a package should be done

* every time you open R and want to use the package
* once on each computer, generally
* each time you want to use a function from the package

Feedback

* Installing a package is essentially the downloading (or unpacking of) the files associated with the package. This only needs to be done once, generally speaking.
* Right!
* Installing a package is essentially the downloading (or unpacking of) the files associated with the package. This only needs to be done once, generally speaking.

What does the following function call do?

stats::filter(…)

* Explicitly call the filter function from the stats package.
* Explicitly call the stats function from the filter package.

Feedback

* Right! If you don’t want to read in all of a package, you can use the :: operator. You still need to install that package prior to doing so.
* The first argument is the package and the second the function form that package. If you don’t want to read in all of a package, you can use the :: operator. You still need to install that package prior to doing so.

## Reading delimited data 33:36

What is a delimiter?

* A raw data value.
* A character that separates data values.
* A type of file.
* The function used to read raw data into R.

Feedback

* The delimiter is the character (or characters) that are intended to separate the raw data values.
* Right!
* The delimiter is the character (or characters) that are intended to separate the raw data values. Most of the time a delimited file will have an extension of .csv, .tsv, .dat, or .txt.
* The delimiter is the character (or characters) that are intended to separate the raw data values. There are many function in R to read in different types of delimited files (such as read.csv).

Why do many prefer to use packages from the tidyverse?

Feedback

The functions in the package generally work with the same syntax, work well together, and have many users creating use-case examples.

# 2.2 Video:

## Excel, SAS, SPSS questions 7:25

Check all of the following that are true about the readxl package and the read\_excel() function.

* Both.xls and .xlsx files can be read.
* Excel files can be read from a URL.
* A specific Excel sheet can be chosen to be read in
* Multiple Excel sheets can be read in using a single function call
* Not all columns must be read in from a particular Excel sheet

## JSON, API questions 16:22

Roughly speaking, what is an API and why are they important to know about?

Feedback

An API is a defined communication protocol between computers. If you can create a proper query of an API, you can return useful information.

Often APIs exist to allow users to return data of interest.

What is the usual process of contacting an API through R?

* Functions are used to build a URL needed to return appropriate information.
* You connect via the database connect functionality RStudio has.
* You scrape the data from the web using HTML tags.
* You use the readr package or the haven package.

Feedback

* Right! Usually you build a URL with the appropriate syntax to obtain information from the API.
* Usually you build a URL with the appropriate syntax to obtain information from the API.
* Usually you build a URL with the appropriate syntax to obtain information from the API.
* Usually you build a URL with the appropriate syntax to obtain information from the API.

## Databases 30:13

Roughly speaking, what is a database?

Feedback

A collection of data, usually many related rectangular datasets.

# Homework Notes