How to better structure code using own R packages to analyse advertising campaigns

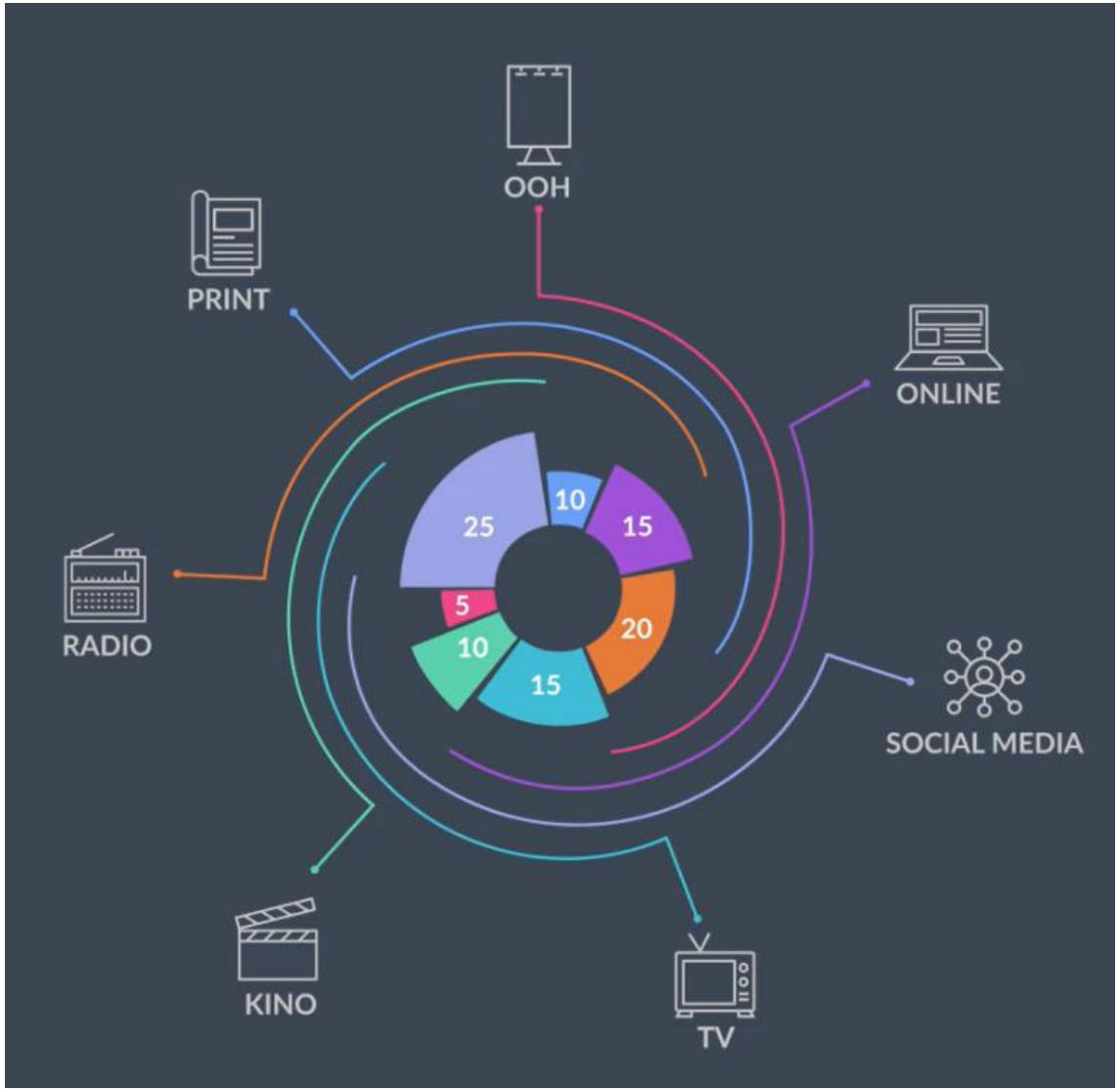
Núria Duran Adroher
Data scientist at Signifikant

Lucerne, November 25, 2020

- Signifikant is a startup located at Technopark Luzern (Root D4)

- We analyse campaigns launched by customers from different fields (consumer goods, telecommunication, online shopping, banking…)

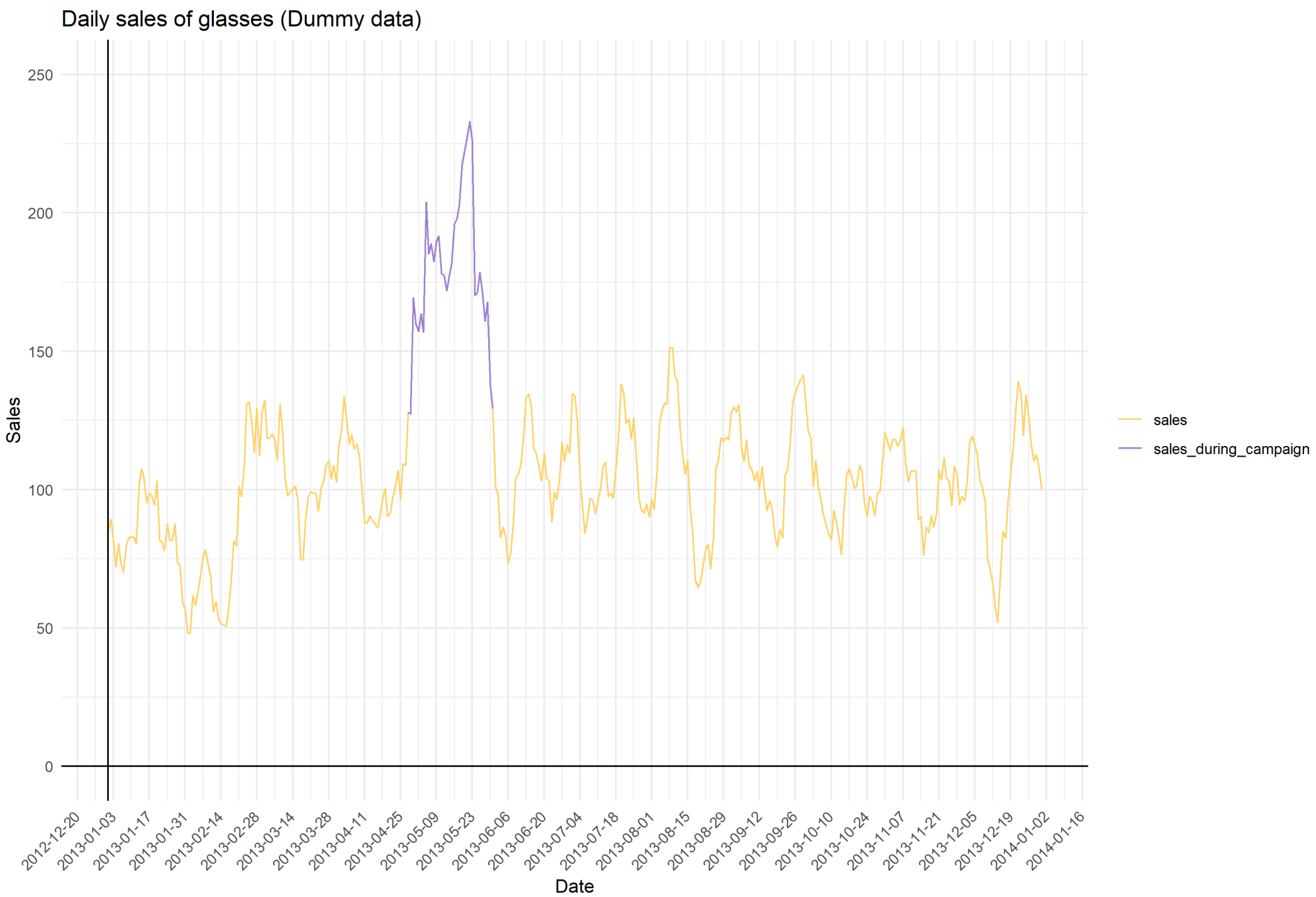
- Using data-driven algorithms, we answer questions such as:
  - How successful was the campaign?
  - How did the different media channels (online, tv, print, …) perform in the campaign?
  - …

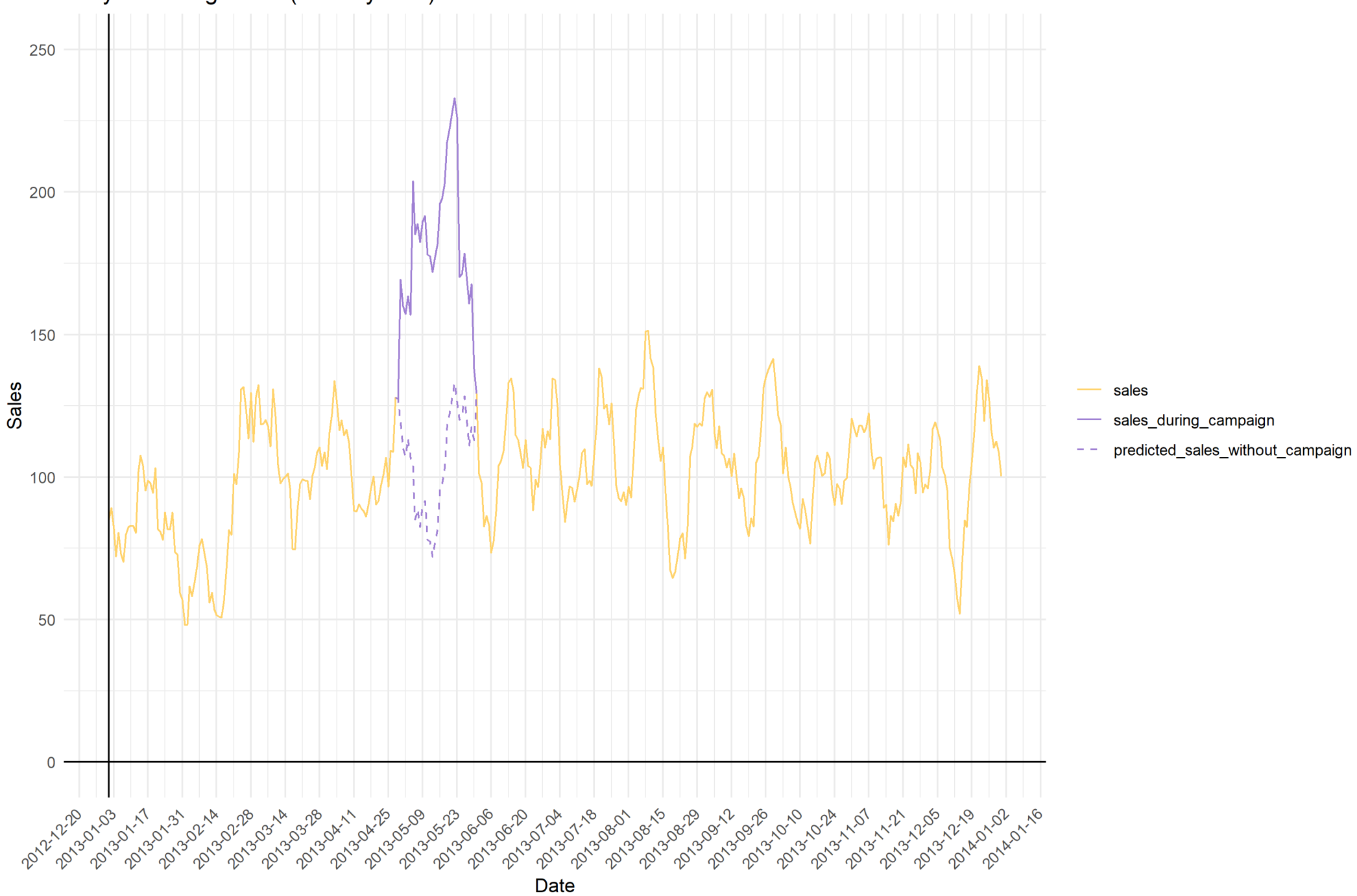


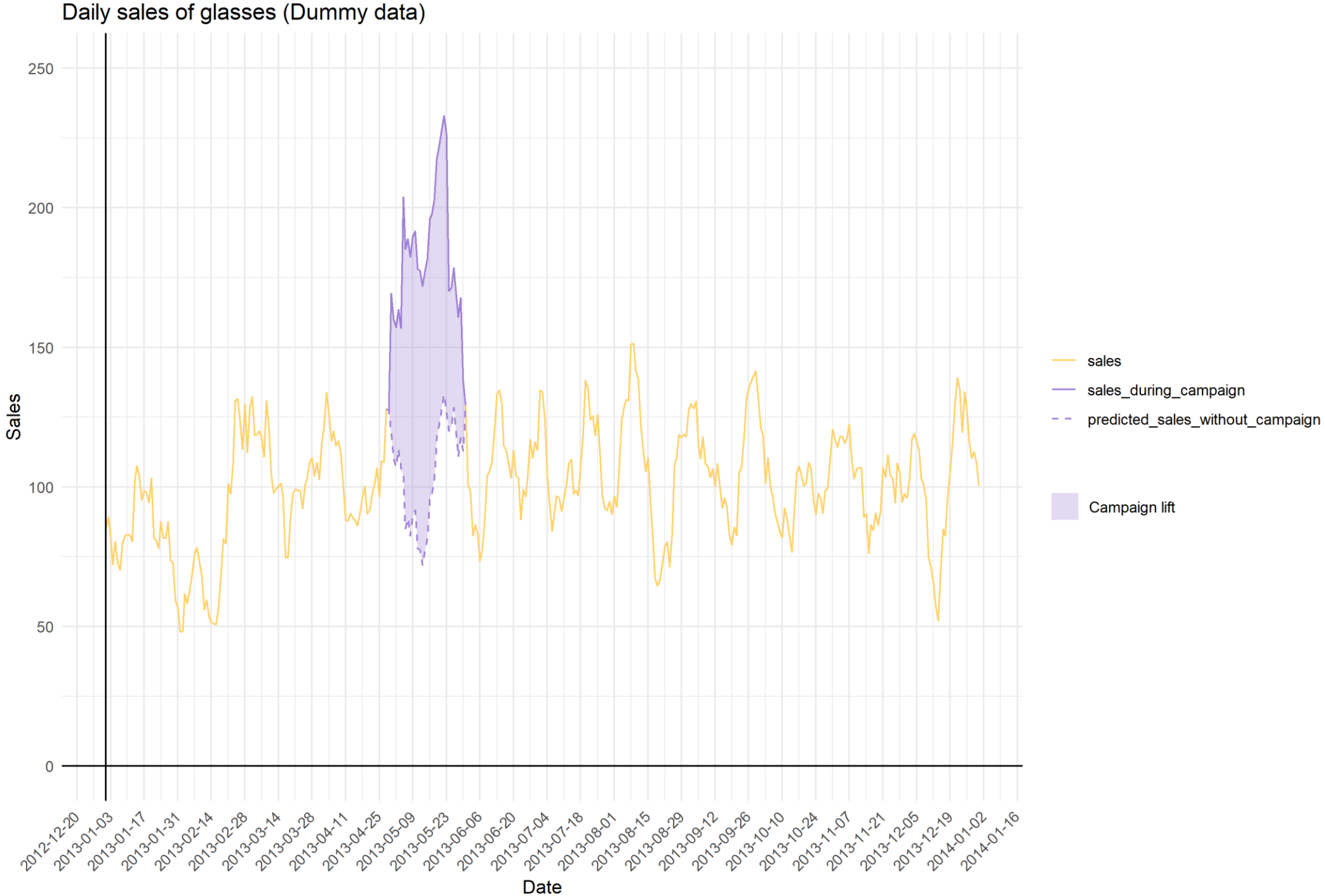
https://www.signifikant.com/

Example of a print campaign

https://www.strategies.fr/creations/campagnes/212344/keloptic-vente-de-lunettes-en-ligne-keloptic-com-voyez-plus-clair-pour-moins-cher-mai-2013.html

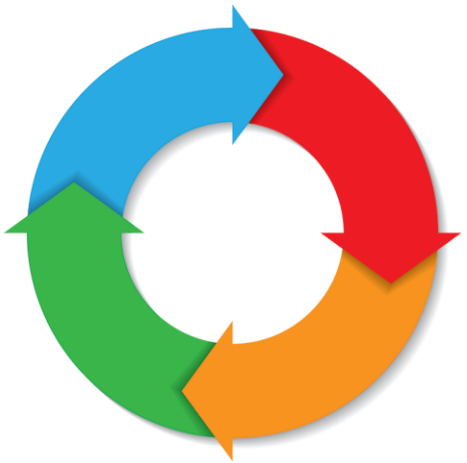

Daily sales of glasses (Dummy data)
Sales
250
200
150
100
50
0
2012-12-20
2013-01-03
2013-01-17
2013-01-31
2013-02-14
2013-02-28
2013-03-14
2013-03-28
2013-04-11
2013-04-25
2013-05-09
2013-05-23
2013-06-06
2013-06-20
2013-07-04
2013-07-18
2013-08-01
2013-08-15
2013-08-29
2013-09-12
2013-09-26
2013-10-10
2013-10-24
2013-11-07
2013-11-21
2013-12-05
2013-12-19
2014-01-02
2014-01-16
Date
sales
sales_during_campaign


Daily sales of glasses (Dummy data)
Sales
250
200
150
100
50
0
2012-12-20
2013-01-03
2013-01-17
2013-01-31
2013-02-14
2013-02-28
2013-03-14
2013-03-28
2013-04-11
2013-04-25
2013-05-09
2013-05-23
2013-06-06
2013-06-20
2013-07-04
2013-07-18
2013-08-01
2013-08-15
2013-08-29
2013-09-12
2013-09-26
2013-10-10
2013-10-24
2013-11-07
2013-11-21
2013-12-05
2013-12-19
2014-01-02
2014-01-16
Date
sales
sales_during_campaign
predicted_sales_without_campaign


Daily sales of glasses (Dummy data)
Sales
250
200
150
100
50
0
2012-12-20
2013-01-03
2013-01-17
2013-01-31
2013-02-14
2013-02-28
2013-03-14
2013-03-28
2013-04-11
2013-04-25
2013-05-09
2013-05-23
2013-06-06
2013-06-20
2013-07-04
2013-07-18
2013-08-01
2013-08-15
2013-08-29
2013-09-12
2013-09-26
2013-10-10
2013-10-24
2013-11-07
2013-11-21
2013-12-05
2013-12-19
2014-01-02
2014-01-16
Date
sales
sales_during_campaign
predicted_sales_without_campaign
Campaign lift

Data analysis steps

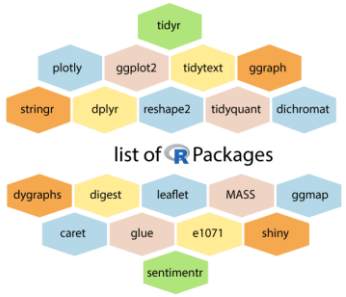
- Load and merge the different datasets
  - Sales
  - Media channels (online, TV, print, …)
  - Weather
  - Competition
  - …
- Predict sales in the hypothetical situation that the campaign had not taken place
- Calculate the performance of the media channels
- Transform the results to present them in a dashboard



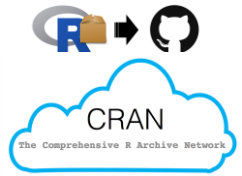
Developing own R packages to structure data analysis

https://r-pkgs.org/index.html

WHY



- Using existing R packages is very useful: it's nice to use functions from other people instead of having to code them by yourself



- Sharing code via R packages
  - R users can install it and learn how to use it

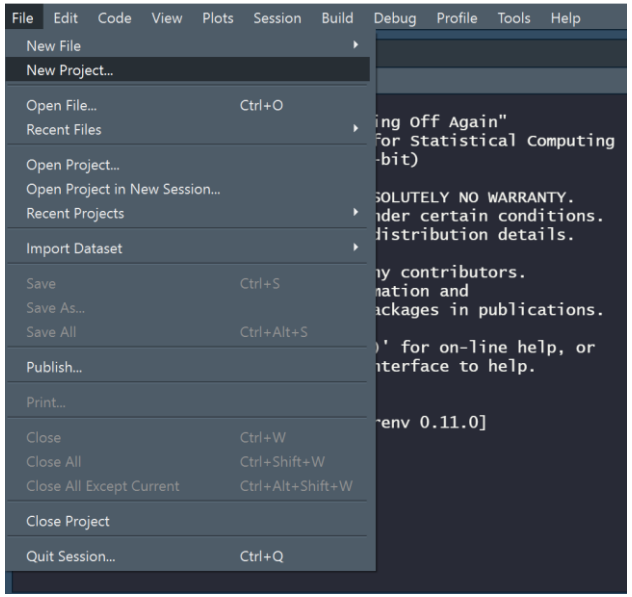
- Even if one does not want to share code, organising code in packages makes life easier
  - Reproducibility[1]
  - Automation
  - It dramatically improves usability and visibility for your work

[1] Ben Marwick, Carl Boettiger & Lincoln Mullen (2018) Packaging Data Analytical Work Reproducibly Using R (and Friends), The American Statistician, 72:1, 80-88, DOI: 10.1080/00031305.2017.1375986

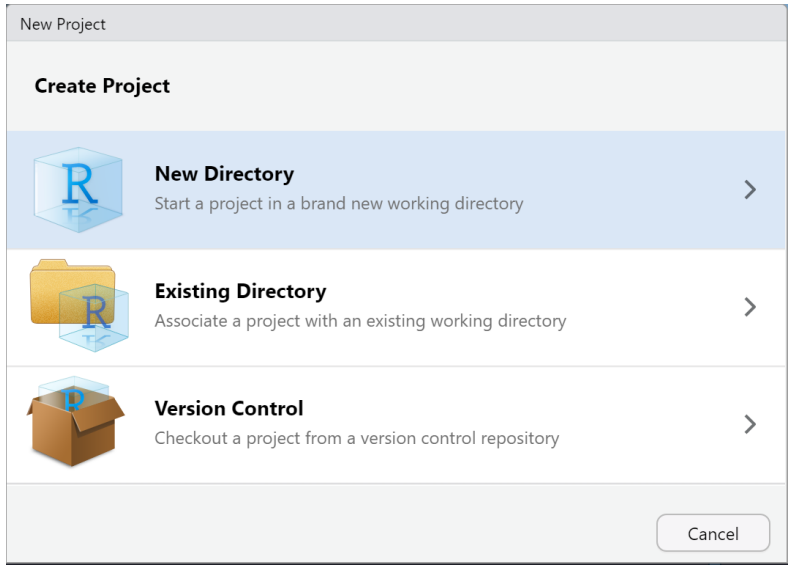
Developing own R packages to structure data analysis

HOW

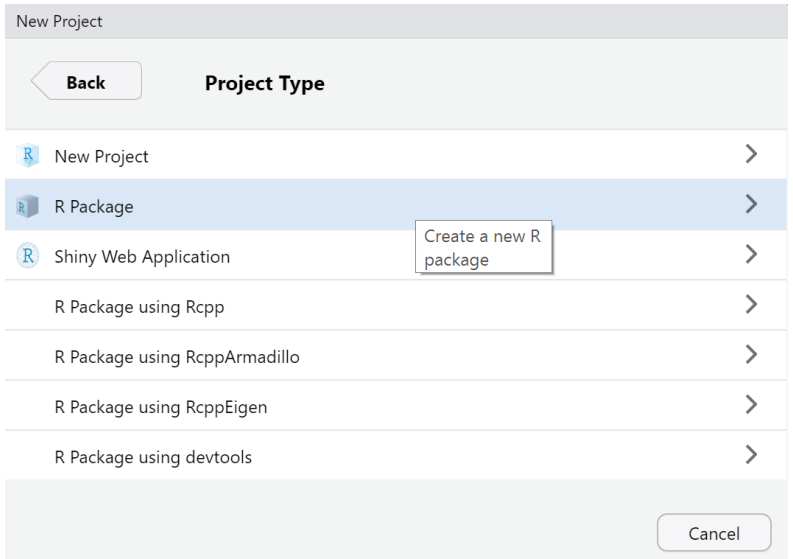
1



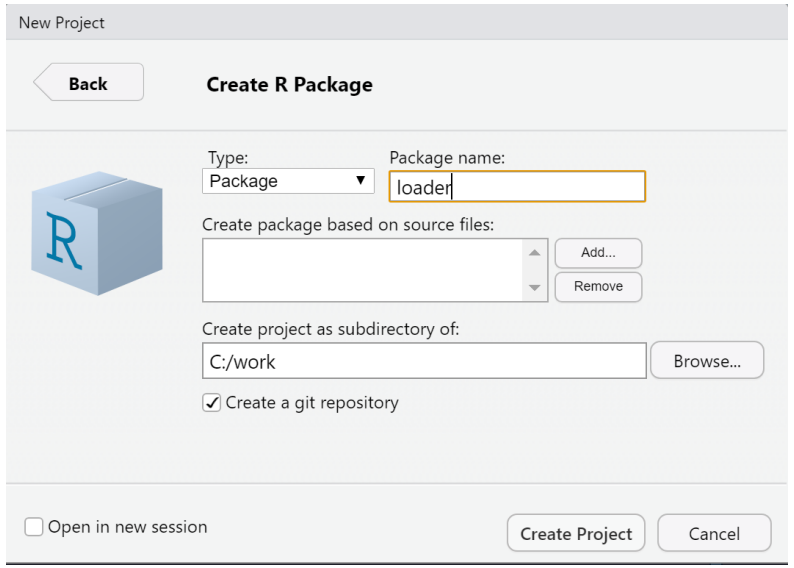
2



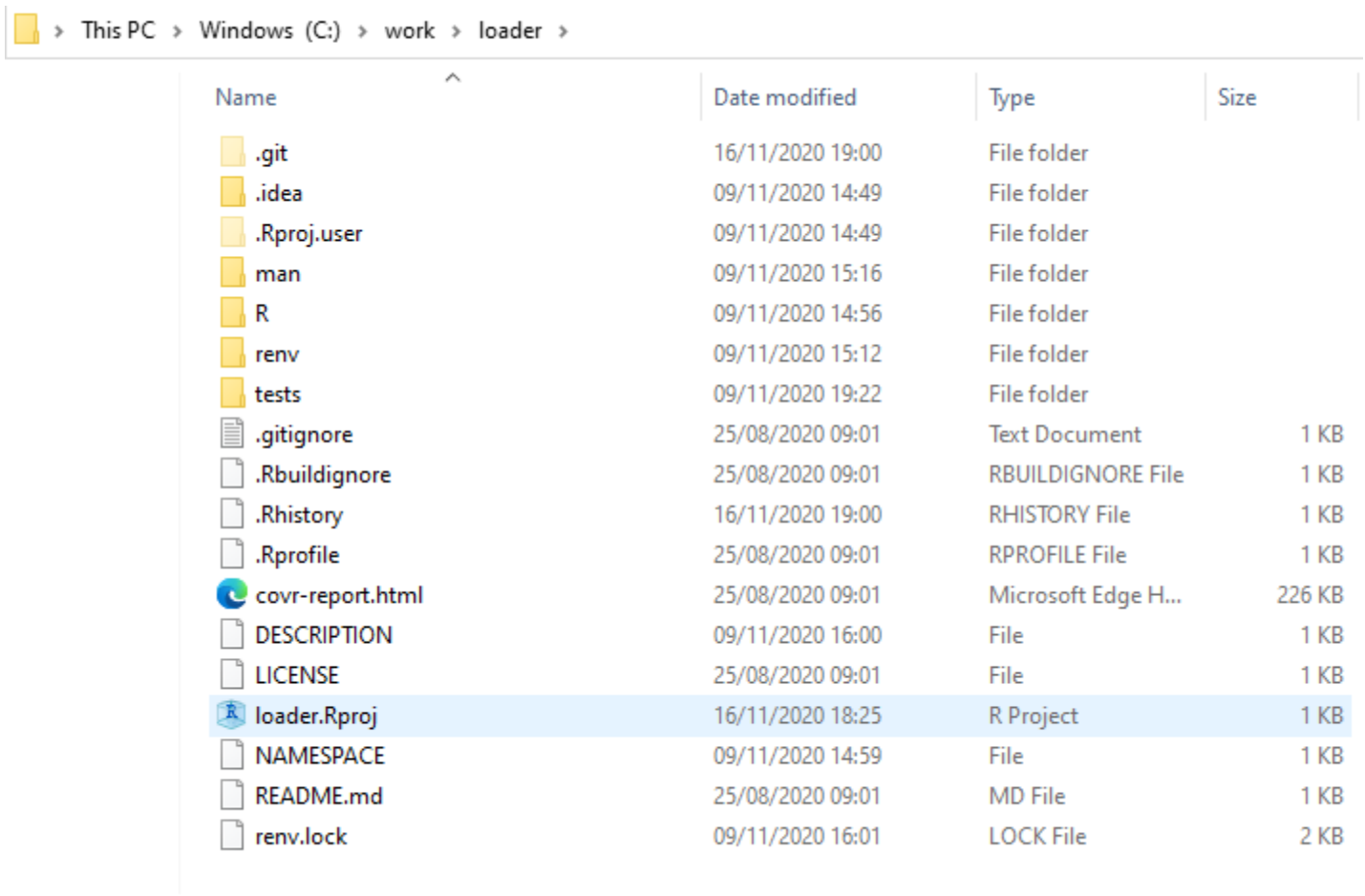
3



4

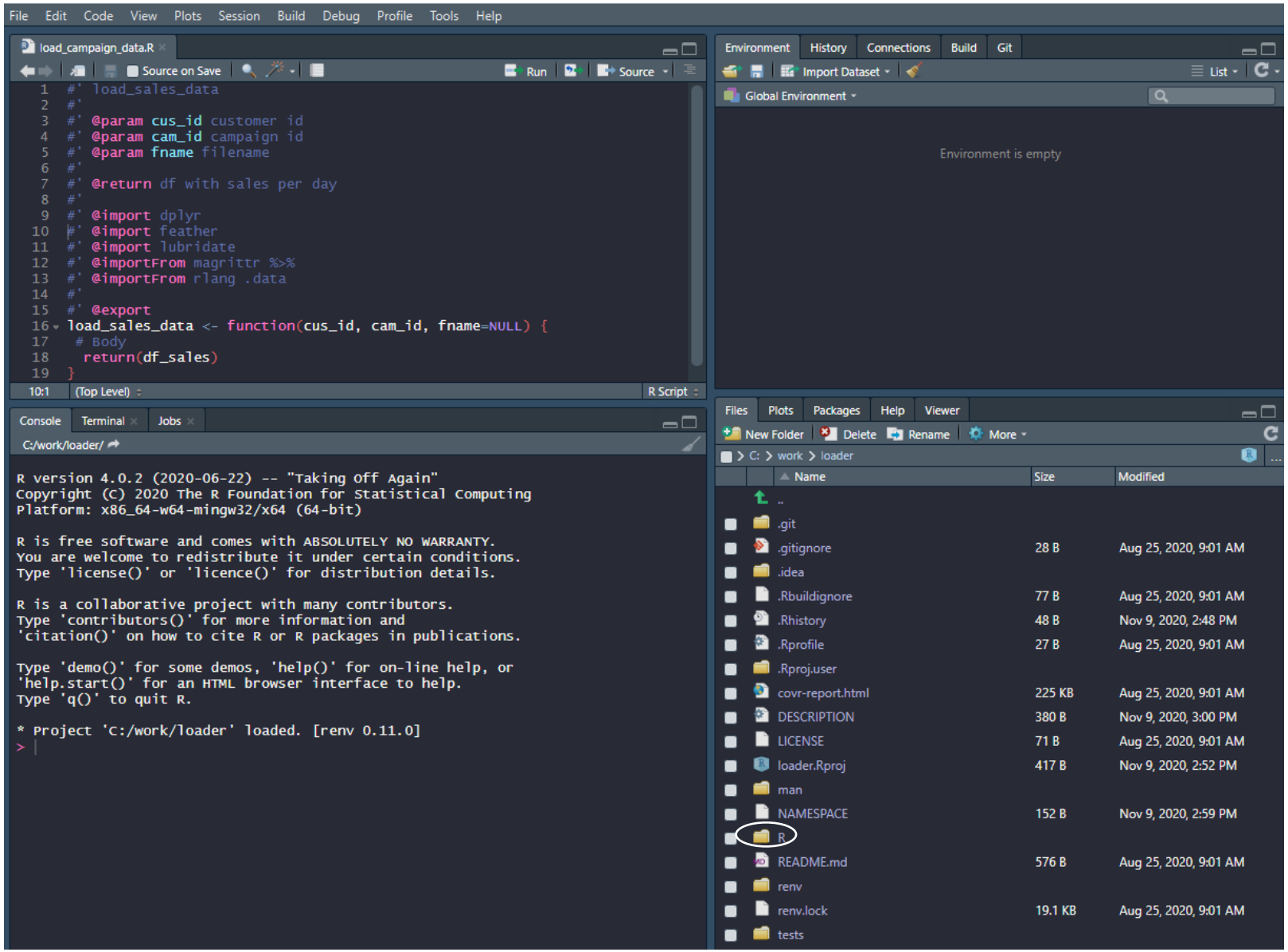


HOW

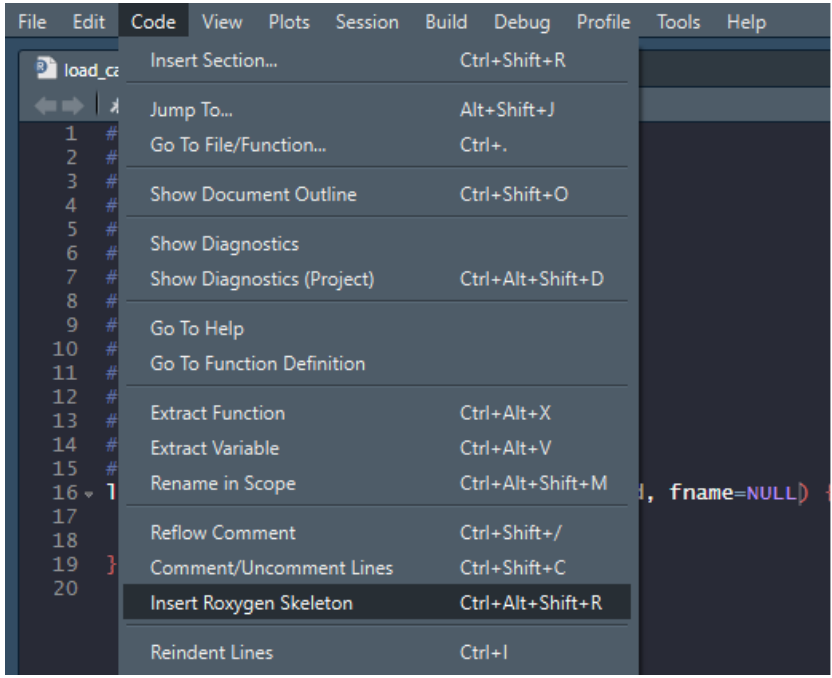

This PC › Windows (C:) › work › loader ›
Name
Date modified
Type
Size
.git 16/11/2020 19:00 File folder
.idea 09/11/2020 14:49 File folder
.Rproj.user 09/11/2020 14:49 File folder
man 09/11/2020 15:16 File folder
R 09/11/2020 14:56 File folder
renv 09/11/2020 15:12 File folder
tests 09/11/2020 19:22 File folder
.gitignore 25/08/2020 09:01 Text Document 1 KB
.Rbuildignore 25/08/2020 09:01 RBUILDIGNORE File 1 KB
.Rhistory 16/11/2020 19:00 RHISTORY File 1 KB
.Rprofile 25/08/2020 09:01 RPROFILE File 1 KB
covr-report.html 25/08/2020 09:01 Microsoft Edge H... 226 KB
DESCRIPTION 09/11/2020 16:00 File 1 KB
LICENSE 25/08/2020 09:01 File 1 KB
loader.Rproj 16/11/2020 18:25 R Project 1 KB
NAMESPACE 09/11/2020 14:59 File 1 KB
README.md 25/08/2020 09:01 MD File 1 KB
renv.lock 09/11/2020 16:01 LOCK File 2 KB

Developing own R packages to structure data analysis

HOW

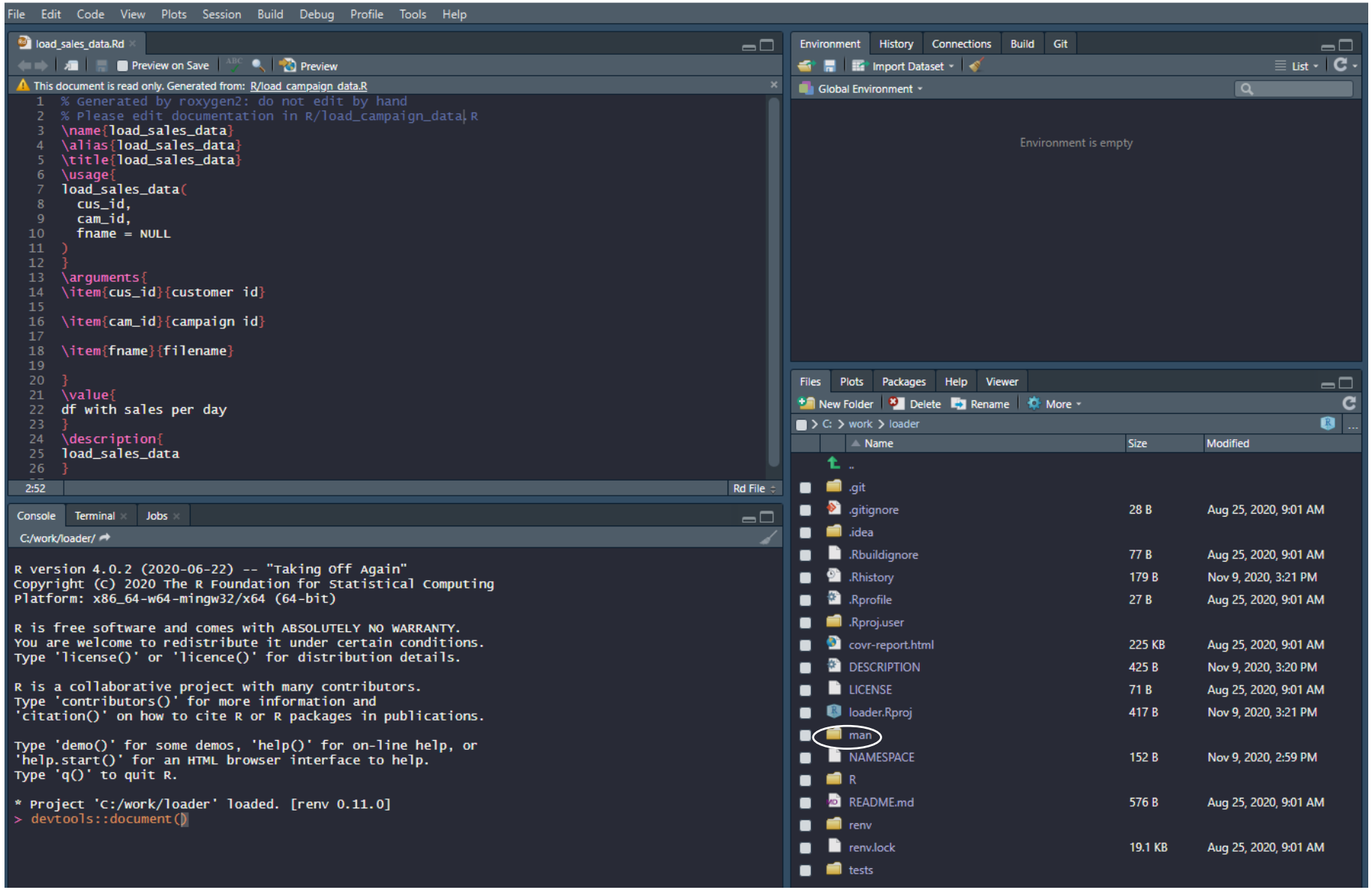


- Use ***devtools***, ***usethis***, and ***roxygen2*** packages in RStudio
- Follow a template: a package is a R project (e.g., *'loader.Rproj')* with predefined files and folders
- Functions must be stored under the *R* folder
- Functions must be documented using *Roxygen Skeleton*



Developing own R packages to structure data analysis

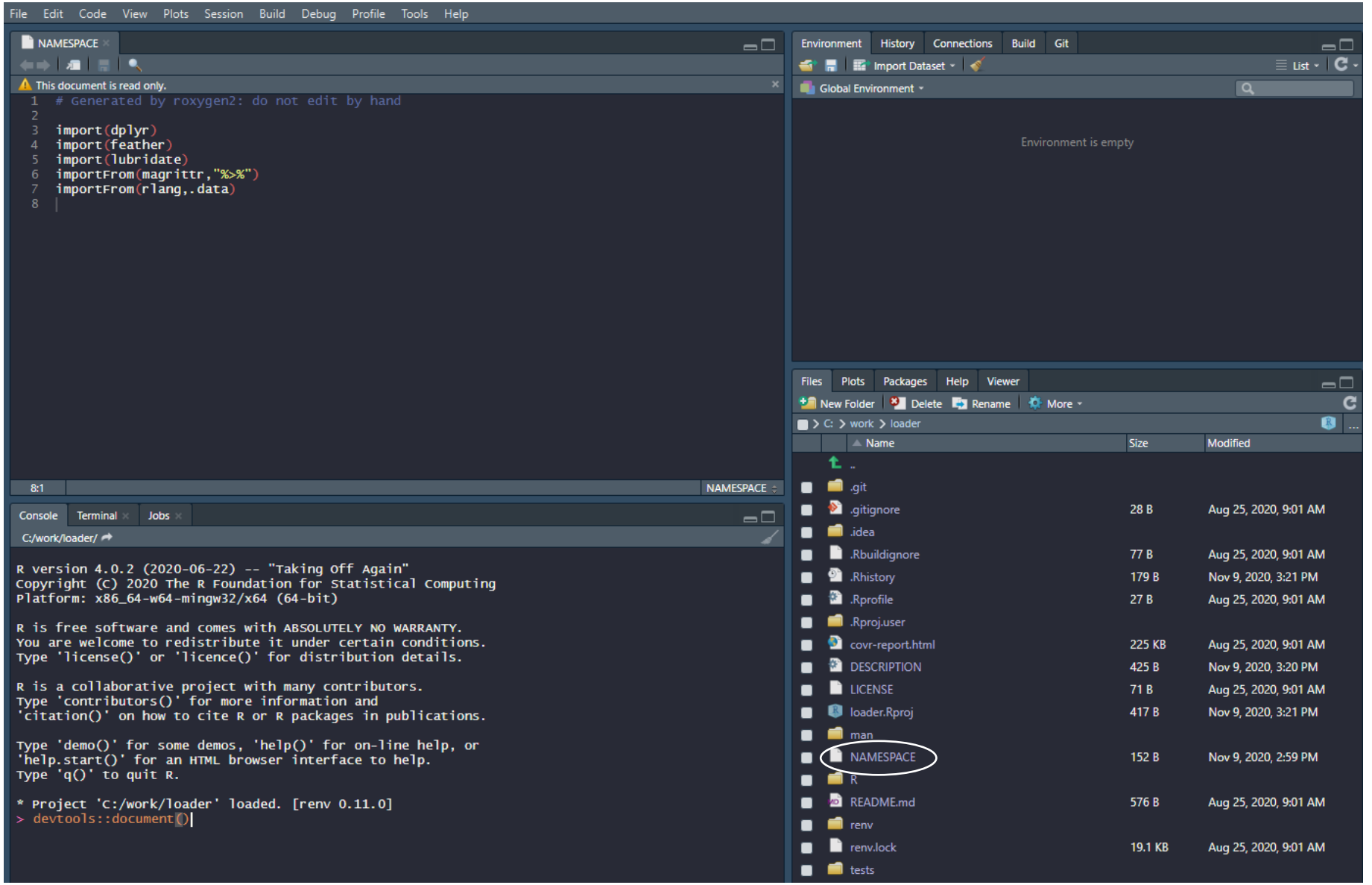
HOW



- The *man* folder contains one *.Rd* file for each function
- The *.Rd* file is automatically generated with the command *devtools::document()*

Developing own R packages to structure data analysis

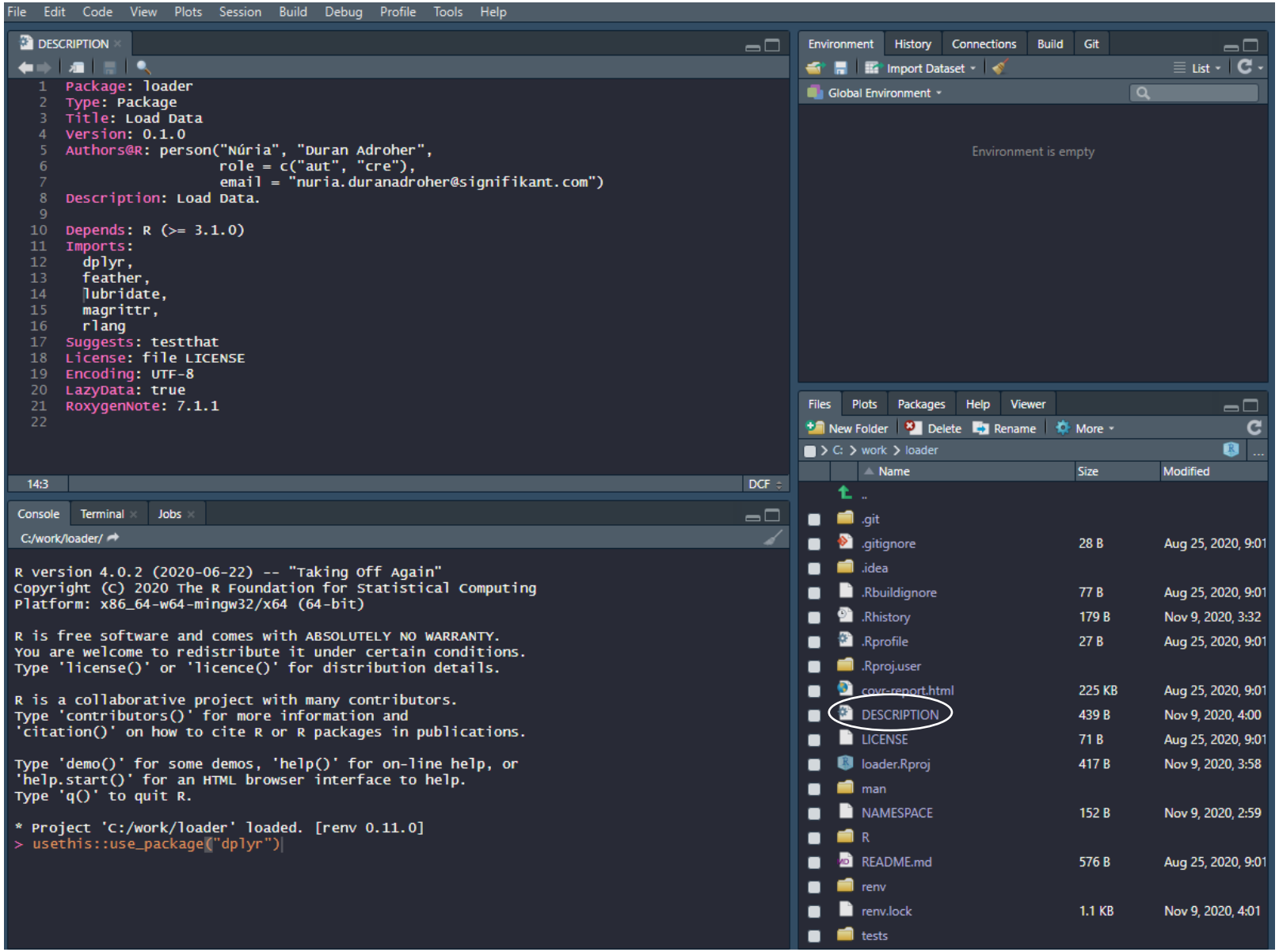
HOW



- The *NAMESPACE* file contains commands for importing or exporting functions, generated by roxygen2
- It is automatically updated using the command *devtools::document()*

Developing own R packages to structure data analysis

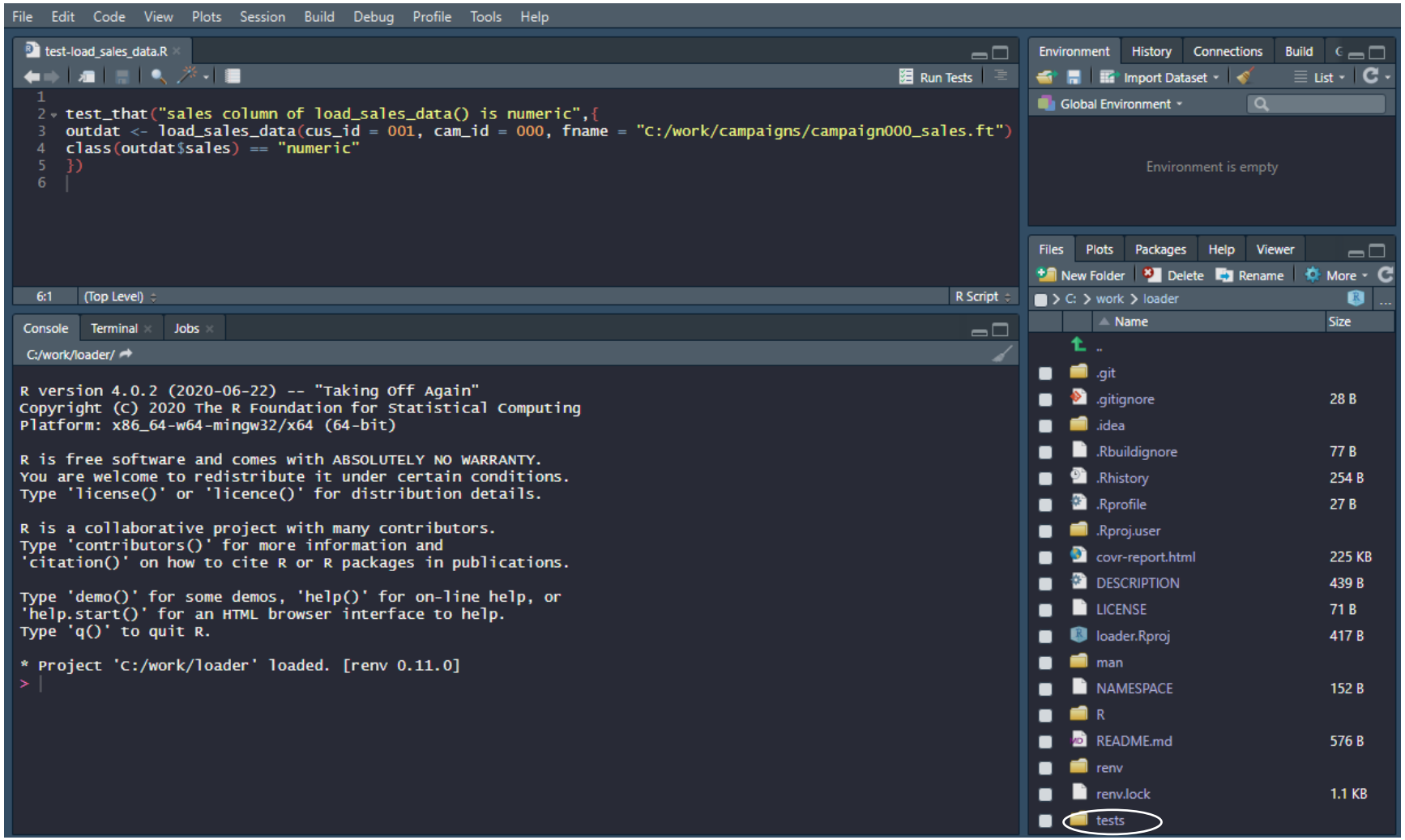
HOW



- The *DESCRIPTION* file contains metadata
- The *Imports* section is automatically updated using the command *usethis::use_package("pkg_name")*

Developing own R packages to structure data analysis

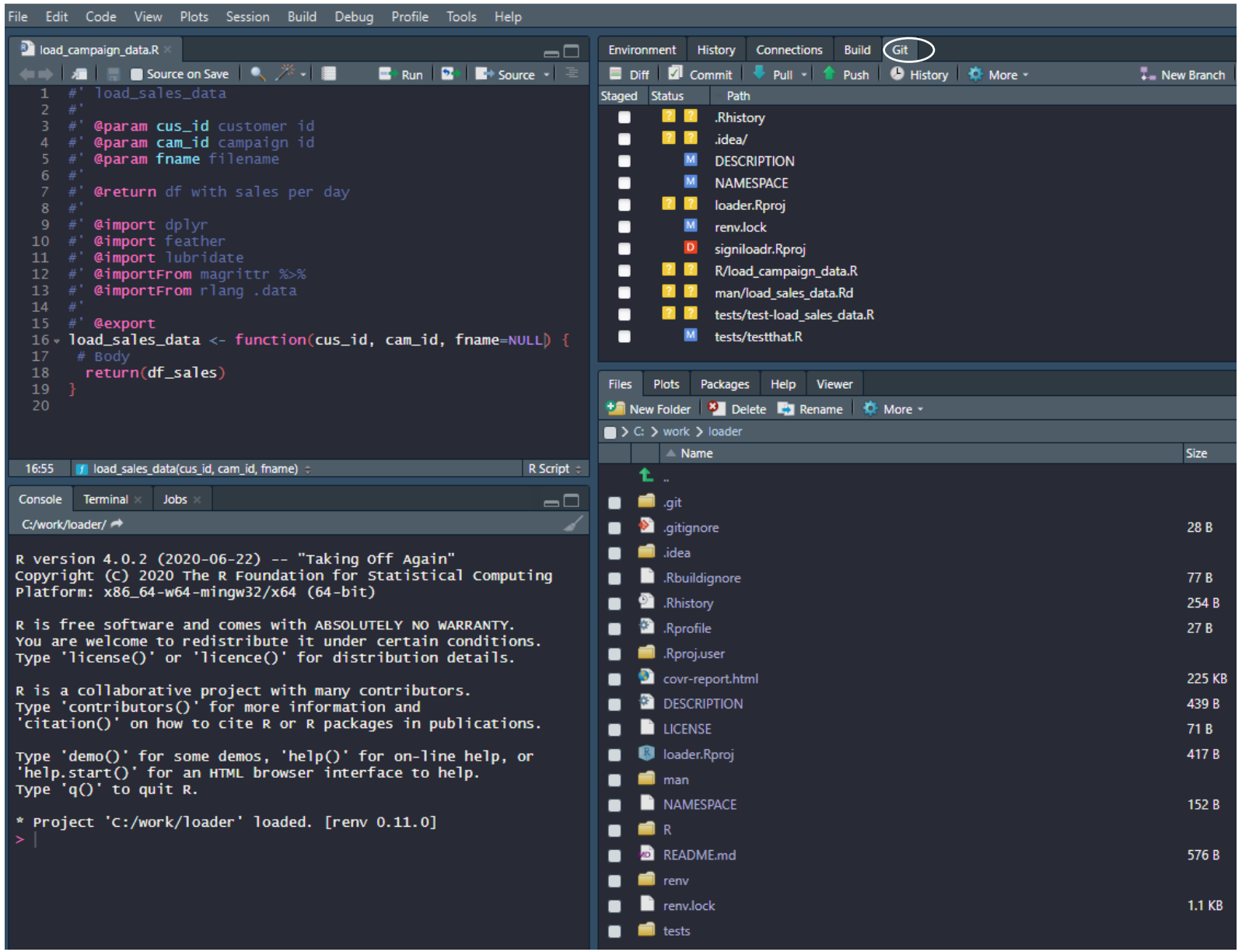
HOW



- Write unit tests under the *tests* folder

Developing own R packages to structure data analysis

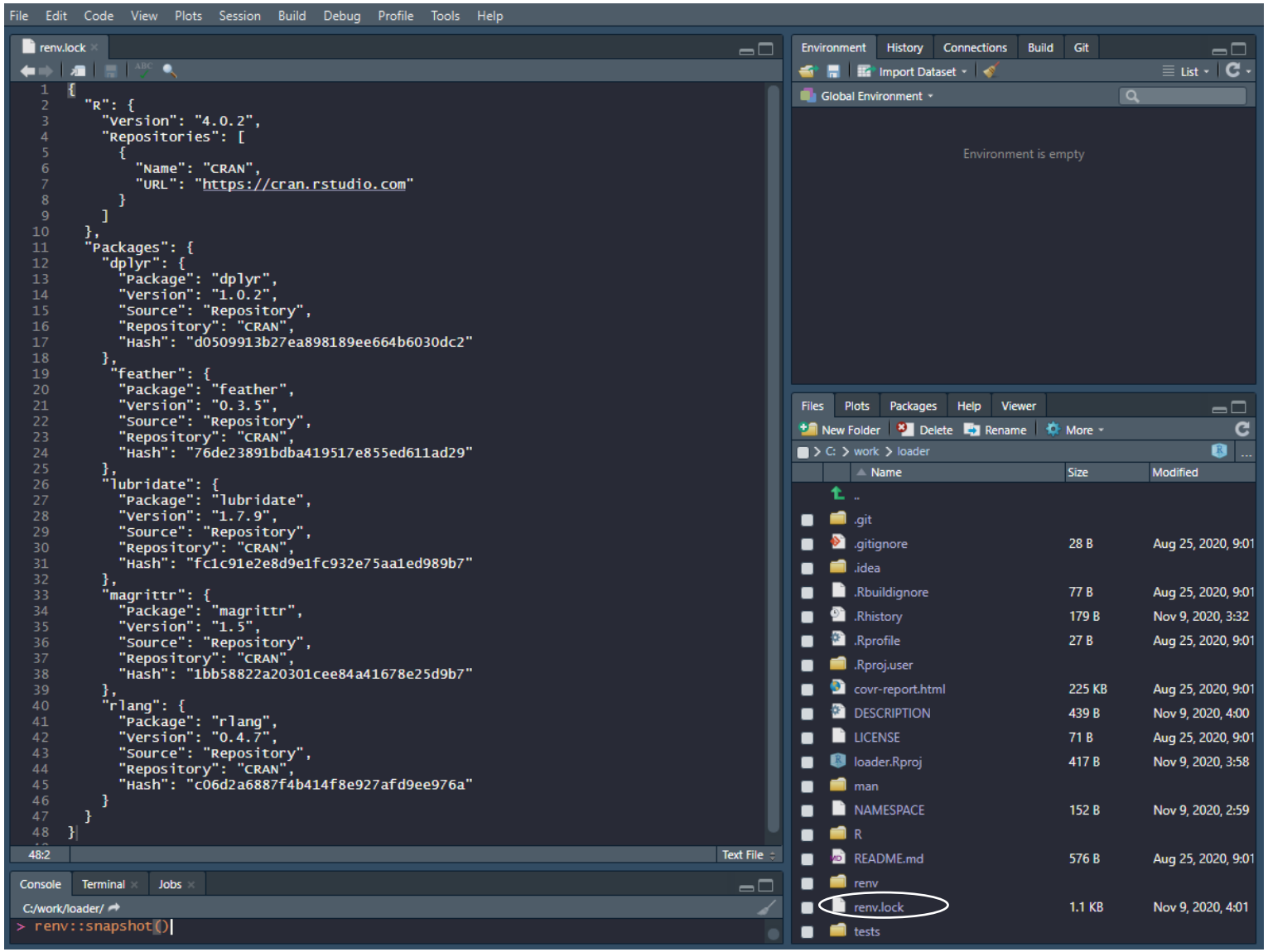
HOW



- Commit and push

Developing own R packages to structure data analysis

HOW (optional)

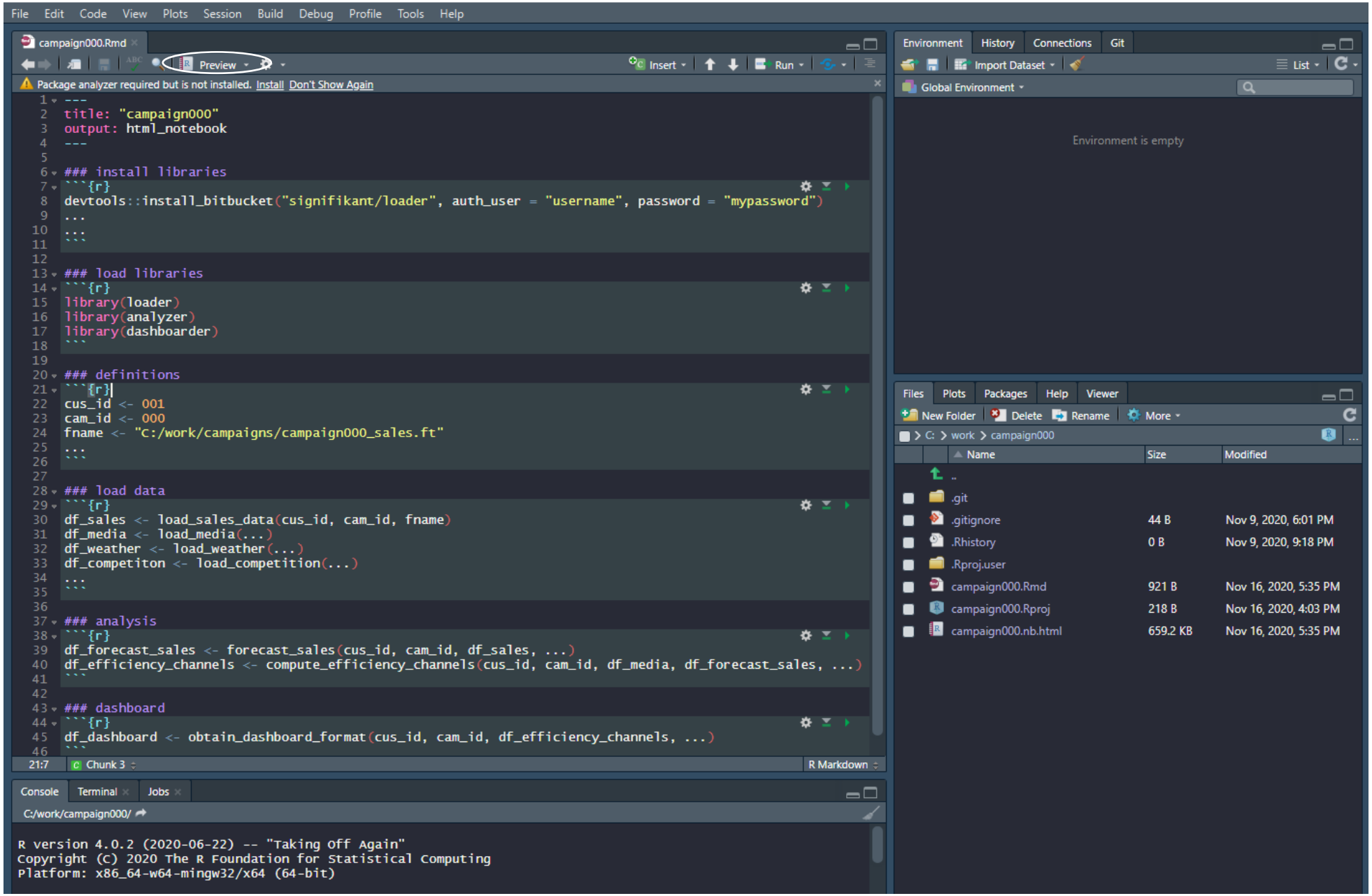


- Save and load the state of a project library using the package *renv*
  - The *renv.lock* file captures the state of the library at some point in time

Using own R packages for the analysis

- Upload them in **bitbucket** (similar repository to **github**)
- In a R script (.R) or R Markdown file (.Rmd), install them via the command *devtools::install_bitbucket(…)*

Using the packages for the analysis



- Create a .Rmd file containing the steps of the analysis process

C:/work/campaign000/campaign000.nb.html

campaign000.nb.html | Open in Browser | Find | Publish

campaign000

Code

install libraries

Hide

```
devtools::install_bitbucket("signifikant/loader", auth_user = "username", password = "mypassword")
...
...
```

load libraries

Hide

```
library(loader)
library(analyzer)
library(dashboarder)
```

definitions

Hide

```
cus_id <- 001
cam_id <- 000
fname <- "C:/work/campaigns/campaign000_sales.ft"
...
```

load data

Hide

```
df_sales <- load_sales_data(cus_id, cam_id, fname)
df_media <- load_media(...)
df_weather <- load_weather(...)
df_competiton <- load_competition(...)
...
```

analysis

Hide

```
df_forecast_sales <- forecast_sales(cus_id, cam_id, df_sales, ...)
df_efficiency_channels <- compute_efficiency_channels(cus_id, cam_id, df_media, df_forecast_sales, ...)
```

dashboard

Hide

```
df_dashboard <- obtain_dashboard_format(cus_id, cam_id, df_efficiency_channels, ...)
```

Summary

- R packages can be used for any kind of project; we've seen the example of defining functions which analyse ad campaign data

- To develop a R package one needs to create a project and adapt to the specified template of files and folders

- Reproducibility enhanced by using ***renv***
- Automation obtained by running a main R script / Markdown file which mainly calls functions defined in the own developed packages
- Usability is improved by documented functions using ***roxygen2***
- Visibility is improved by clean code in R script / Markdown file

Thanks for your attention!

nuria.duranadroher@signifikant.com