

Workflow: scripts and projects

Luis Francisco Gomez Lopez

2023-09-01

Contents

- Scripts
- Projects
- References

Scripts

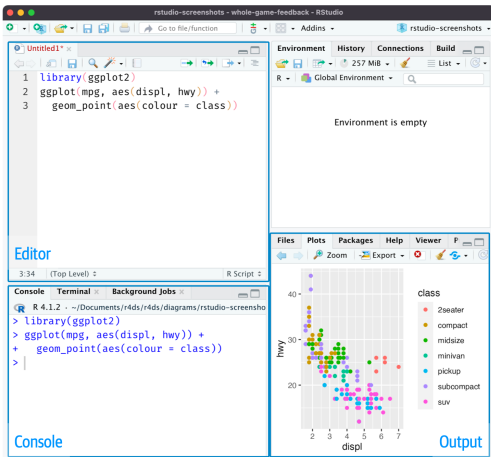


Figure 1: Editor script pane

Scripts

- Create script Cmd/Ctrl + Shift + N
- Run line with Cmd/Ctrl + Enter
- Run complete script Cmd/Ctrl + Shift + S

```
not_cancelled <- flights |>
  filter(!is.na(dep_delay), !is.na(arr_delay))

not_cancelled |>
  group_by(year, month, day) |>
  summarize(mean = mean(dep_delay))
```

Scripts

- Diagnostics
 - Problems
 - Potential problems

```
x ← 10  
3 == NA
```

Scripts

- Saving and naming
 - File names should be machine readable
 - Avoid spaces, symbols and special characters
 - Don't rely on case sensitivity
 - File names should be human readable
 - Use file names to describe what's in the file
 - File names should play well with default ordering
 - Start file names with numbers so that alphabetical sorting puts them in the order they get used

```
000_images
000_data_sets
000_articles
001_import
002_data_wrangling
...
008_communication
```

Projects

- Working directory
 - Where does your analysis live?
 - Fix the working directory using R projects
 - File > New Project

```
getwd()
```

```
[1] "C:/Users/Usuario/Documents/r_for_data_science_2_edition_umng/007_workflow_scripts_and_projects"
```

- Specifying the relative path because the working directory is fixed using R projects
 - You can avoid absolute paths and share your work with others

```
ggplot(diamonds, aes(x = carat, y = price)) +  
  geom_hex()  
ggsave("../000_images/007_diamonds.png")  
  
write_csv(diamonds, "../000_data_sets/007_diamonds.csv")
```

References I