Workflow: code style

Luis Francisco Gomez Lopez

2023-08-16

Contents

Names

- Must start with a letter
  - It is recommend to use only **ASCII** lower letters
  - You can use for example Ã, ã, Á, á, À, à, Ç, ç, Ü or ü but it is not a good practice
- Should use only lowercase letters, numbers, and _
- As a general rule of thumb it's better to prefer long and descriptive names that are easy to understand
- If you have a bunch of names for related things, do your best to be consistent

```
# Strive for:
short_flights <- flights |> filter(air_time < 60)

# Avoid:
ShfLÍ <- flights |> filter(air_time < 60)
```

Spaces

- Put spaces on either side of mathematical operators $+, -, ==, <, \ldots$
  - The exception is ^

```
# Strive for
z <- (a + b)^2 / d

# Avoid
z<-( a + b ) ^ 2/d
```

- Don't put spaces inside or outside parentheses for regular function calls

```
# Strive for
mean(x, na.rm = TRUE)

# Avoid
mean (x ,na.rm=TRUE)
```

Spaces

- You can add extra spaces if it improves alignment but it is not mandatory

```
flights |>
  mutate(
    speed      = (distance * 1.609344) / (air_time / 60),
    dep_hour   = dep_time %/% 100, # integer division
    dep_minute = dep_time %% 100 # modulus
  )
```

Pipes

- |> should always have a space before it and should typically be the last thing on a line
  - Easier to add new steps
  - Rearrange existing steps
  - Modify elements within a step
  - Bird's eye view by skimming the verbs on the left-hand side

```
# Strive for
flights |>
  filter(!is.na(arr_delay), !is.na(tailnum)) |>
  count(dest)

# Avoid
flights|>filter(!is.na(arr_delay), !is.na(tailnum))|>count(dest)
```

Pipes

- If the arguments to a function don't all fit on one line, put each argument on its own line and indent (Alexander Wickham, 2022, Chapter 4)

```
# Strive for
flights |>
  group_by(tailnum) |>
  summarize(
    delay = mean(arr_delay, na.rm = TRUE),
    n = n()
  )

# Avoid
flights|>
  group_by(tailnum) |>
  summarize(
             delay = mean(arr_delay, na.rm = TRUE),
             n = n()
           )

# Avoid
flights|>
  group_by(tailnum) |>
  summarize(
  delay = mean(arr_delay, na.rm = TRUE),
  n = n()
  )
```

ggplot2

- The same basic rules that apply to |> also apply to ggplot2 so treat + the same way as |>

```
flights |>
  group_by(dest) |>
  summarize(
    distance = mean(distance),
    speed = mean(distance / air_time, na.rm = TRUE)
  ) |>
  ggplot(aes(x = distance, y = speed)) +
  geom_smooth(
    method = "loess",
    span = 0.5,
    se = FALSE,
    color = "white",
    linewidth = 4
  ) +
  geom_point()
```

Sectioning comments

- Break up your file into manageable pieces by using sectioning comments

```
# Libraries ----


# Data sets ----


# Import data ----
        .
        .
        .
```

References I

Alexander Wickham, H. (2022). *The tidyverse style guide*. https://style.tidyverse.org/