

Indent and Comment Are Critical for Readability and Reproducibility

Huanyuan Zhang, Xiongjie Deng

Environmental Change Institute, School of Geography and the Environment, University of Oxford

huanyuan.zhang@ouce.ox.ac.uk xiongjie.deng@ouce.ox.ac.uk



Do

- Use snake case (e.g. my_variable)
- Use "<-" to assign
- Add a space after commas
- Use TRUE / FALSE
- Put "{" on the same line and "}" on its own line
- Use double quotation (") mark for strings
- Add spaces around infix operators

Don't

- Use dotted variable (e.g. my.variable)
- Use "=" to assign
- Explicitly type integers (e.g. 6L)
- Use T / F
- Terminate lines with semicolon
- Use single quotation (') mark for strings
- Use tab to indent

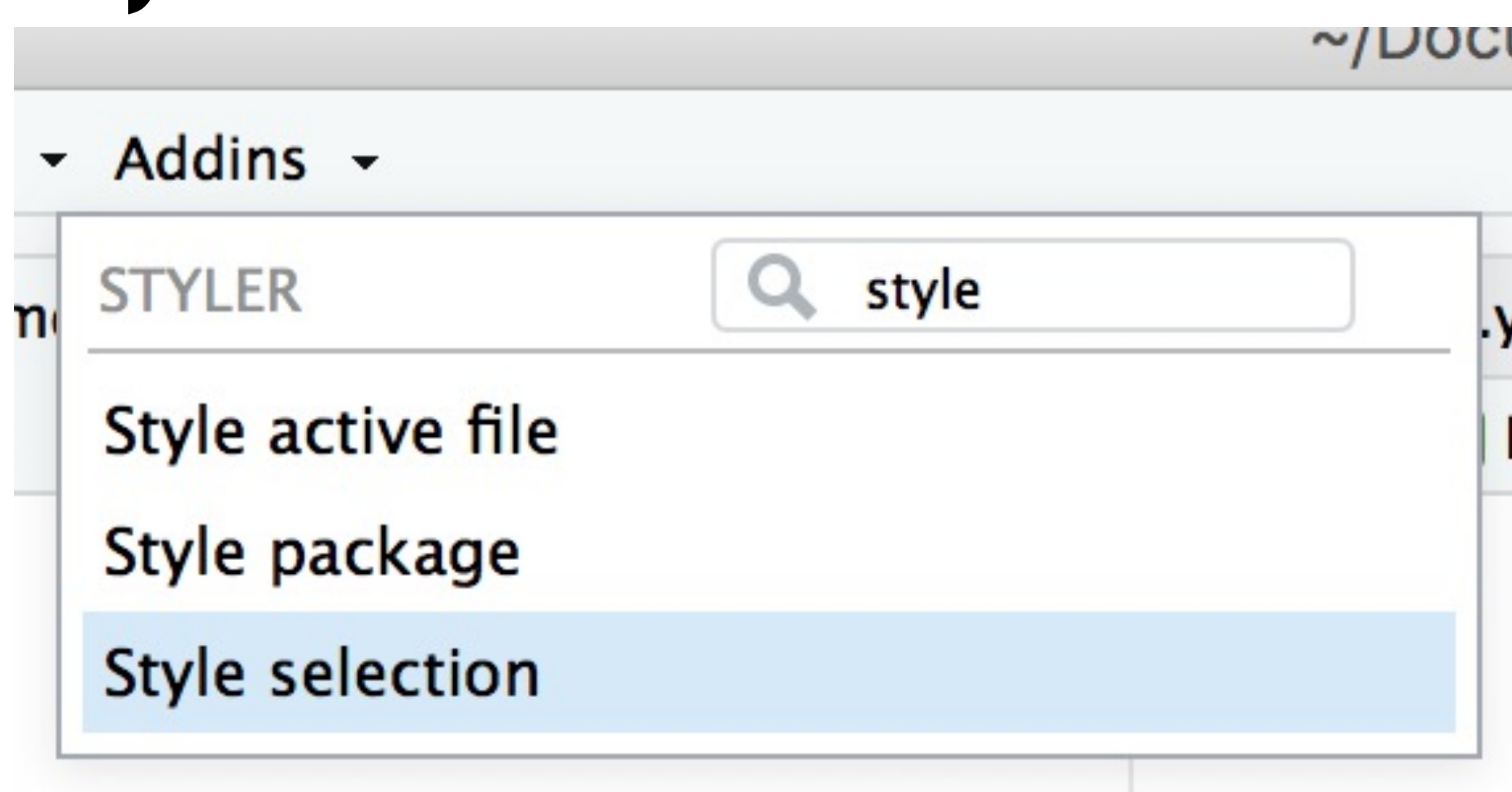
Good Example

```
soft.func <- function(m_v, leaky = FALSE) {  
  if (leaky) {  
    warnings("using leaky RELU!")  
    return(ifelse(m_v > 0, m_v * 3))  
  }  
  return(log(1 + exp(m_v)))  
}
```

Bad Example

```
softFunc=function(m.v,leaky=F)  
{  
  if(leaky)  
  {  
    warnings('using leaky RELU!')  
    return(ifelse(m.v>0L,m.v*3)); }  
  return(log(1L+exp(m.v))); }
```

'styler' for auto indent



Code structure

Method 1: Rmarkdown + Rstudio

Method 2: ARTofR (no Rstudio)

Tips:

- sub-titles ✓ navigation panel ✓ divider ✓
- in-line comments ✓ Paragraph comments ✓
- inline comments are critical for reproduction
- Good to copy and paste paragraph from email/paper to explain codes
- Read me embedded in the R script (Author; Contacts; Time; Instructions; Problems)

Reference

There is no right in styling, the suggestions above are based on surveys and online guides:

Yen, C. Y., Chang, M. H. W., & Chan, C. H. (2022). A Computational Analysis of the Dynamics of R Style Based on 108 Million Lines of Code from All CRAN Packages in the Past 21 Years. *R JOURNAL*, 14(1), 6-21.

Wickham, The tidyverse style guide <https://github.com/tidyverse/style>

