

Day 5: Wrapping Up

DIMA Units of Regional Bureaux: EHAGL, SA and WCA

18 February 2022

Today will be about:

- Recap of Last 4 Days
- Small Things to Know
 - R Coding Style
 - Strings and Country Names
- Not Yet Asked Questions
- Potential Areas to Cover in the Future
- Feedback and Building Community

Recap of Last 4 Days

- Introduction
- Basics and Data Import
- Data Manipulation
- R Markdown and Visualisation with ggplot2

Small Things to Know

R Coding Style: Notation and Naming (Reference: [R Coding Style](#))

File names

File names should be meaningful and end in **.R**.

```
# Good
fit-models.R
utility_functions.R

# Bad
foo.r
stuff.r
```

If files need to be run in sequence, prefix them with numbers:

```
0-download.R
1-parse.R
2-explore.R
```

Small Things to Know

R Coding Style: Notation and Naming (Reference: [R Coding Style](#))

Object names

Variable and function names should be lowercase. Use an underscore (_) to separate words within a name. Generally, **variable** names should be **nouns** and **function** names should be **verbs**.

```
# Good
day_one
day_1
dayOne

# Bad
first_day_of_the_month
dayone
djm1
```

Where possible, avoid using names of existing functions and variables. Doing so will cause confusion for the readers of your code.

```
# Bad
T <- FALSE
c <- 10
mean <- function(x) sum(x)
```

Small Things to Know

R Coding Style: Syntax (Reference: [R Coding Style](#))

Spacing

Place spaces around all infix operators (=, +, -, <-, etc.).
The same rule applies when using = in function calls.
Always put a space after a comma, and never before (just like in regular English).

```
# Good
average <- mean(feet / 12 + inches, na.rm = TRUE)

# Bad
average<-mean(feet/12+inches,na.rm=TRUE)
```

There's a small exception to this rule: :, :: and ::: don't need spaces around them.

```
# Good
x <- 1:10
base::get

# Bad
x <- 1 : 10
base :: get
```

Small Things to Know

String Problem of Country Names

- Use country codes where possible
- Example: Different country names between Dataport and ASR

```
## # A tibble: 10 x 2
##   dataport_coo_name      asr_coo_name
##   <chr>                <chr>
## 1 Central African Republic Central African Rep.
## 2 Congo, Republic of the Congo
## 3 Democratic Republic of the Congo Dem. Rep. of the Congo
## 4 Gambia, Republic of The Gambia
## 5 United Kingdom of Great Britain and North. Ireland United Kingdom of Great B~
## 6 Côte d'Ivoire Cote d'Ivoire
## 7 Iran, Islamic Republic of Iran (Islamic Rep. of)
## 8 Macau China, Macao SAR
## 9 Syrian Arab Republic Syrian Arab Rep.
## 10 United Republic of Tanzania United Rep. of Tanzania
```

- Encoding issue: Côte d'Ivoire

Small Things to Know

String Problem of Country Names

- Example: Different country codes between UNHCR code and ISO3

```
## # A tibble: 94 x 3
##   iso3  unhcr unsd_name
##   <chr> <chr> <chr>
## 1 DZA   ALG   Algeria
## 2 AGO   ANG   Angola
## 3 ATG   ANT   Antigua and Barbuda
## 4 AUS   AUL   Australia
## 5 AUT   AUS   Austria
## 6 BHR   BAH   Bahrain
## 7 BRB   BAR   Barbados
## 8 BLZ   BZE   Belize
## 9 BMU   BER   Bermuda
## 10 BTN  BHU   Bhutan
## # ... with 84 more rows
```

- Use `countrycode` package

```
install.packages(countrycode)
library(countrycode)
```


Not Yet Asked Questions

Potential Areas to Cover in the Future

Feedback and Building Community

