

The data.R script

Preprocess data, write TAF data tables

Colin Millar
Cecilia Kvaavik

WKTAF, Copenhagen
14-15 June 2022



Science for sustainable seas

The data.R script



After running `taf.bootstrap()`, all the underlying data files have been made ready inside the `bootstrap/data` folder

The purpose of the `data.R` script is to process the underlying data to create the input data that are used in the assessment

It also produces TAF tables showing the main input data

The data.R script

The underlying data files may:

- ▶ include years and ages that are not used in the stock assessment
great, let's document in a script how the data are filtered
- ▶ be at an earlier stage than the input data used in the stock assessment
great, let's document in a script how the data are calculated/smoothed/aggregated
- ▶ be in any format and are sometimes hard to read, both for humans and machines
no problem, let's also make the data available as standardized CSV files

Outline of data.R



```
## Preprocess data, write TAF data tables

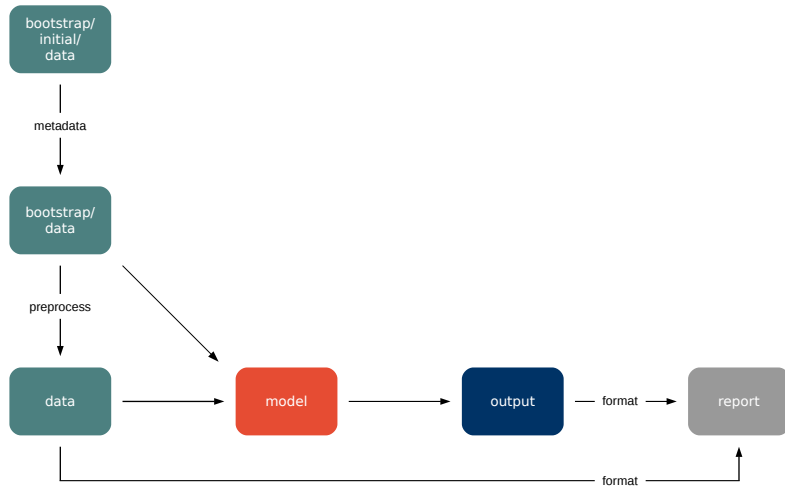
## Before: surveys_all.csv (bootstrap/data)
## After:  survey.csv (data)

library(icesTAF)

mkdir("data")

## 1  Read underlying data from bootstrap/data
...
## 2  Preprocess data
...
## 3  Write TAF tables to data directory
...
```

Data flow



Explore existing data.R scripts



<https://github.com/ices-taf>