Datasets in "sdam" package

Antonio Rivero Ostoic

September 2022

Preliminaries

Install and load one version of "sdam" package.

```
install.packages("sdam") # from CRAN
devtools::install_github("sdam-au/sdam") # development version
devtools::install_github("mplex/cedhar", subdir="pkg/sdam") # a legacy version R 3.6.x
```

```
# load and check version
library(sdam)
packageVersion("sdam")
```

```
[1] '1.0.0'
```

Built-in datasets

Package "sdam" comes with a suite of datasets and external data to execute different functions available in the package and to perform analysis.

For a list of built-in datasets in "sdam" use the "utils" function data() or utils::data() with the 'package' argument.

The CRAN distribution has four built-in datasets, while the development and legacy distributions add three more built-in datasets.

```
# pop-up a new window
data(package="sdam")
```

```
# Data sets in package 'sdam':
#
# retn      Roman Empire transport network and Mediterranean sea
# rp        Roman province names and acronyms as in EDH
# rpcp      Roman provinces chronological periods
# rpd       Roman provinces dates from EDH
# rpmcd     Caption maps and affiliation dates of Roman provinces
```

```
# Additional built-in datasets in 'sdam':
#
# EDH       Epigraphic Database Heidelberg Dataset
# rpmp      Maps of ancient Roman provinces and Italian regions
```

A description of each dataset is available in the manual that from the R console is accessible as e.g. the EDH dataset in a non-CRAN distribution.

```
# Epigraphic Database Heidelberg Dataset help
?EDH
```

Ancient Mediterranean built-in datasets

The `EDH` dataset in `"sdam"` has information about Latin epigraphy retrieved from the Epigraphic Database Heidelberg API repository from the Roman world during the antiquity period.

A list of Roman provinces and regions in this dataset is available in dataset `"rp"`, and use again function `data()` to load this built-in dataset to look at its internal structure with `utils::str()` function.

- Dataset `"rp"` is a named list with Roman provinces and regions with acronyms according to the Epigraphic Database Heidelberg.

```
# load dataset
data("rp")

# obtain object structure
str(rp)
```

```
List of 66
 $ Ach: chr "Achaia"
 $ Aeg: chr "Aegyptus"
 $ Aem: chr "Aemilia (Regio VIII)"
 $ Afr: chr "Africa Proconsularis"
 $ AlC: chr "Alpes Cottiae"
 $ AlG: chr "Alpes Graiae"
 $ AlM: chr "Alpes Maritimae"
 $ AlP: chr "Alpes Poeninae"
 $ ApC: chr "Apulia et Calabria (Regio II)"
 $ Aqu: chr "Aquitania"
 $ Ara: chr "Arabia"
 $ Arm: chr "Armenia"
 $ Asi: chr "Asia"
 $ Ass: chr "Assyria"
 $ Bae: chr "Baetica"
 $ Bar: chr "Barbaricum"
 $ Bel: chr "Belgica"
 $ BiP: chr "Bithynia et Pontus"
 $ BrL: chr "Bruttium et Lucania (Regio III)"
 $ Bri: chr "Britannia"
 $ Cap: chr "Cappadocia"
 $ Cil: chr "Cilicia"
 $ Cor: chr "Corsica"
 $ Cre: chr "Creta"
 $ Cyp: chr "Cyprus"
 $ Cyr: chr "Cyrene"
 $ Dac: chr "Dacia"
 $ Dal: chr "Dalmatia"
 $ Epi: chr "Epirus"
 $ Etr: chr "Etruria (Regio VII)"
 $ Gal: chr "Galatia"
 $ GeI: chr "Germania inferior"
```

```
 $ GeS: chr "Germania superior"
 $ HiC: chr "Hispania citerior"
 $ Inc: chr "Provincia incerta"
 $ Iud: chr "Iudaea"
 $ LaC: chr "Latium et Campania (Regio I)"
 $ Lig: chr "Liguria (Regio IX)"
 $ Lug: chr "Lugdunensis"
 $ Lus: chr "Lusitania"
 $ LyP: chr "Lycia et Pamphylia"
 $ MaC: chr "Mauretania Caesariensis"
 $ MaT: chr "Mauretania Tingitana"
 $ Mak: chr "Macedonia"
 $ Mes: chr "Mesopotamia"
 $ MoI: chr "Moesia inferior"
 $ MoS: chr "Moesia superior"
 $ Nar: chr "Narbonensis"
 $ Nor: chr "Noricum"
 $ Num: chr "Numidia"
 $ PaI: chr "Pannonia inferior"
 $ PaS: chr "Pannonia superior"
 $ Pic: chr "Picenum (Regio V)"
 $ Rae: chr "Raetia"
 $ ReB: chr "Regnum Bospori"
 $ Rom: chr "Roma"
 $ Sam: chr "Samnium (Regio IV)"
 $ Sar: chr "Sardinia"
 $ Sic: chr "Sicilia, Melita"
 $ Syr: chr "Syria"
 $ Thr: chr "Thracia"
 $ Tra: chr "Transpadana (Regio XI)"
 $ Tri: chr "Tripolitania"
 $ Umb: chr "Umbria (Regio VI)"
 $ Val: chr "Valeria"
 $ VeH: chr "Venetia et Histria (Regio X)"
```

edhw() interface with "rp" dataset

- Function `edhw()` is a wrapper to extract and transform the records in the `EDH` dataset that invokes `"rp"` dataset to retrieve the records from a specific Roman province or region in `EDH`.

```
# Armenian records in 'EDH'
edhw(province="Arm")[1]
```

```
Warning in edhw(province = "Arm"): "x" is for dataset "EDH".

Warning in edhw(province = "Arm"): "province" with no "vars" returns lists.

[[1]]
[[1]]$ID
[1] "015521"

[[1]]$commentary
[1] " Mehrere, teils aneinanderpassende Fragmente erhalten. Die Inschrift lief über mehrere Ta
```

```
[[1]]$country
[1] "Armenia"

[[1]]$depth
[1] "21 cm"

[[1]]$diplomatic_text
[1] "IMP CAESAR DIV[ ] NERVAE F[ ]ERVA TRAIANVS / OPTIMVS A[ ]G G[ ]RM DACI[ ]THICVS PONT MAX

[[1]]$edh_geography_uri
[1] "https://edh-www.adw.uni-heidelberg.de/edh/geographie/3407"

[[1]]$findspot_ancient
[1] "Artaxata, bei"

[[1]]$findspot_modern
[1] "Pokr Vedi"

[[1]]$height
[1] "80 cm"

[[1]]$id
[1] "HD015521"

[[1]]$language
[1] "Latin"

[[1]]$last_update
[1] "2015-10-22"

[[1]]$letter_size
[1] "20-16 cm"

[[1]]$literature
[1] "AE 1968, 0510.; B. Arakelean, RevPaz 126, 1968, 135-136; Foto u. Zeichnung. - AE 1968 "

[[1]]$material
[1] "limestone: rocks - clastic sediments"

[[1]]$military
[1] "data available"

[[1]]$not_before
[1] "0116"

[[1]]$people
[[1]]$people[[1]]
[[1]]$people[[1]]$person_id
[1] "1"

[[1]]$people[[1]]$name
[1] "div[i] Nervae f[il. N]erva Traianus"
```

```
[[1]]$people[[1]]$gender
[1] "male"

[[1]]$people[[1]]$cognomen
[1] "Nerva+ Traianus"



[[1]]$province_label
[1] "Armenia"

[[1]]$responsible_individual
[1] "Gräf"

[[1]]$transcription
[1] "Imp(erator) Caesar div[i] Nervae f[il(ius) N]erva Traianus / optimus A[u]g(ustus) G[e]rm(

[[1]]$trismegistos_uri
[1] "https://www.trismegistos.org/text/217430"

[[1]]$type_of_inscription
[1] "building/dedicatory inscription"

[[1]]$type_of_monument
[1] "tabula"

[[1]]$uri
[1] "https://edh-www.adw.uni-heidelberg.de/edh/inschrift/HD015521"

[[1]]$width
[1] "(205) cm"

[[1]]$work_status
[1] "provisional"

[[1]]$year_of_find
[1] "1967"
```

The `Warning` messages from `edhw()` are first because there is not an explicit input in `x`, it is assumed that the input data is from the `EDH` dataset. The second warning message just tells the type object to return is always a list for argument `province` alone.

`EDH` in data frames

All records in the `EDH` dataset have a list format and it is possible to transform this information into a dataframe format with the wrapper function `edhw()`. For instance, displaying the first record from `Arm` as a data frame in argument `'as'` is made by the record `'id'` number.

```
# record HD015521
edhw(id="15521", as="df")
```

However, it is easier to visualise in the screen only the variables related to people.

```
# record HD015521 with explicit variables
edhw(id="15521", vars="people", as="df")
```

```
        id        cognomen gender                                  name person_id
1 HD015521 Nerva+ Traianus   male div[i] Nervae f[il. N]erva Traianus         1
```

```
# record HD015521 with more explicit variables
edhw(id="15521", vars=c("people", "province_label"), as="df")
```

```
        id        cognomen gender                                  name person_id
1 HD015521 Nerva+ Traianus   male div[i] Nervae f[il. N]erva Traianus         1
  province_label
1        Armenia
```

Obtaining all `people` variables

Start by looking at the `people` variables in the `EDH` dataset for the Roman province of **Armenia**.

Armenia

Fig. 1: Roman province of Armenia (ca 117 AD).

Transformation of the entire province from the `EDH` dataset requires extracting first a list with the province content. Function `edhw()` is to obtain available inscriptions per province from `EDH` and all data attributes from `people` variable. The default outputs are a list and a dataframe for the first and the second instance of the function.

```
# people in Armenia
edhw(province="Arm") |>
  edhw(vars="people")
```

```
        id         age: years        cognomen gender
1 HD015521               <NA> Nerva+ Traianus   male
2 HD015524 data not available        Cre(---)   male
3 HD015524               <NA>           [---]   male
                                name     nomen person_id praenomen             status
1 div[i] Nervae f[il. N]erva Traianus      <NA>         1      <NA>               <NA>
2                       C. Val. Cre. Valerius*         1        C. military personnel
3                              [---]     [---]         2       [-] military personnel
```

People attribute variables in inscriptions for `Armenia` are `age: years`, `cognomen`, `gender`, `name`, `nomen`, `person_id`, `praenomen`, and, `status`, but any inscription with `tribus` or `origo` as in the case of other provinces.

For `Armenia`, two inscriptions have people variables and all people scripted are `male`, where record `HD015524` spans two rows because there are two persons where one have `nomen`, `cognomen`, and `name` ineligible.

Datasets for cartographical maps

The plotting of the Roman province in the previous section requires other datasets. Apart from "rp". In "sdam", there are other three datasets invoked for plotting cartographical maps related to the Roman Empire and the Mediterranean basin, which are "rpmp", "rpmcd", and "retn".

Function plot.map() calls dataset "rpmp" for the shapes and colours in the plotting of the cartographical maps of different regions of the Roman Empire. For the caption and province dates with this function shapes and colours are in dataset "rpmcd".

- Dataset "retn" bears the shapes of places and routes of an ancient transportation system in the Mediterranean region and political divisions of the Roman Empire. It also has it contours and parts of the European continent.

```
# land contour around Mediterranean
plot.map(type="plain")
```



```
# display settlements and shipping routes
plot.map(type="plain", settl=TRUE, shipr=TRUE)
```

Vignette Cartographical maps and networks has more about transportation networks in the ancient Mediterranean.

Datasets with dates

There are built-in datasets in "sdam" related to dates as well that are either displayed in a cartographical map or used for other computations.

- Dataset "rpd" that has dates for provinces from the EDH dataset. It serves for performing a restricted imputation on data subsets in EDH or in another dataset.

```
# dates from EDH
data("rpd")

# three provinces in object structure
str(rpd[1:3])
```

```
List of 3
 $ Ach: 'HD001917' num [1:4] 105 -190 571 761
 $ Aeg: 'HD000741' num [1:4] 144 -71 500 571
 $ Aem: 'HD000010' num [1:4] 121 -201 771 972
```

From this set of three Roman provinces in the EDH, the longest timespan is for Aem, and on average Ach has the oldest incriptions, while Aeg has incriptions with the newest dates.

- Dataset "rpcp" with chronological periods for regions with early and later Roman influence per province.

```
# periods for Roman provinces
data("rpcp")

# object structure
str(rpcp)
```

```
List of 2
 $ Early:'data.frame':  45 obs. of  3 variables:
  ..$ Province: chr [1:45] "Italia (Final Consolidation)" "Sicilia" "Sardinia & Corsica" "Hisp
  ..$ EarInf  : num [1:45] -509 -241 -238 -206 -206 -206 -202 -202 -188 -188 ...
  ..$ OffPrv  : num [1:45] -272 -241 -238 -197 -197 -197 -146 -81 43 -133 ...
 $ Late :'data.frame':  45 obs. of  3 variables:
  ..$ Province: Factor w/ 45 levels "Achaea","Aegyptus",..: 30 43 42 27 28 26 3 23 32 9 ...
  ..$ LateInf : num [1:45] 476 436 436 409 409 ...
  ..$ Fall    : num [1:45] 476 436 436 409 409 409 409 418 1400 1500 ...
```

The early and later Roman influence in the 45 ancient provinces and regions are timespans with a *terminus ante quem* and a *terminus post quem*.

Vignette Dates and missing dating data has the visualisation of these and other dates.

External data

Apart from the built-in datasets, it is attached as external data the semi-colon separated file StraussShipwrecks.csv with the Shipwrecks dataset for performing analyses: Reference and documentation in

Strauss, J. (2013). *Shipwrecks Database*. Version 1.0. Accessed (07-12-2021) from oxrep.classics.ox.ac.uk/databa

Built from Parker, A.J. *Ancient Shipwrecks of the Mediterranean and the Roman Provinces* (Oxford: BAR International Series 580, 1992)

Details about the access to the database are in:

- Shipwrecks network in the Mediterranean Basin (23-June-2022)
- Vignettes Dates and missing dating data and Cartographical maps and networks also use the Shipwrecks dataset.