vegtable: An R object for vegetation-plot data sets

Miguel Alvarez

29 October 2020

Introduction

Why R?

- Collaborative data assessment
  - (R-script + R-image).
- Documenting and ensuring repeatability.
- Teaching purposes.

`taxlist-vegtable` complex

- `taxlist`: taxonomic lists
- `vegtable`: bio-diversity records

Introduction

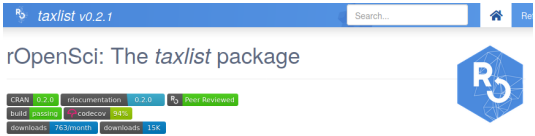
https://docs.ropensci.org/taxlist/

https://github.com/kamapu/vegtable

- Both packages developed on **GitHub**
- Both packages accessible at **CRAN**

```r
## Installing from CRAN
install.packages("vegtable", dependencies=TRUE)

## Installing from GitHub
library(remotes)
install_github("ropensci/taxlist",
        build_vignettes=TRUE)
install_github("kamapu/vegtable")
```



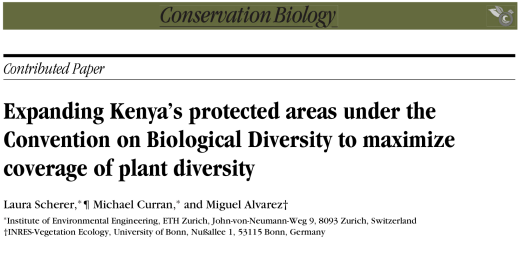
Introduction

vegtable in retrospective

- **2015:** Experiments on the basis of `vegdata`
- **2017:** First version at ***CRAN***
- **2018:** Publication of `taxlist` by Alvarez & Luebert in ***Biodiversity Data Journal***
- **2020:** `taxlist` accepted in ***rOpenSci***

SWEA-Dataveg

https://kamapu.github.io/sweadataveg
https://www.givd.info/ID/AF-00-006



Applications of vegtable

Suggested Applications

- Cleaning data from (messy) spreadsheet applications previous database storage
- Importing data from relational databases
- Merging data from different databases
- Documenting data assessment (R-scripts or R-markdown documents)

Recommended Functions

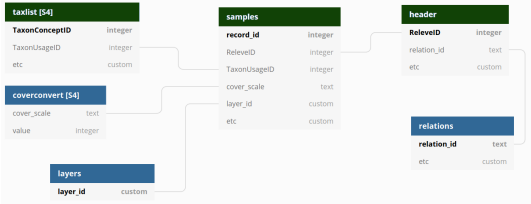
- `backup_object()`: Use of R-images as full backups with timestamp
- `clean()`: Get rid of orphan entries
- `count_taxa()`: Counting number of taxa at plot or cluster levels
- `cross2db()`: Formating cross tables into column oriented tables
- `crosstable()`: Writting cross tables
- `df2vegtable()`: Coercing data frames into `vegtable` objects
- `load_last()`: Load the latest backup written by `backup_object()`
- `match_names()`: Matching strings with a taxonomic list
- `subset()`: Produce sub-sets of `vegtable` objects
- `summary()`: Briefing of `vegtable` and `taxlist` objects
- `trait_stats()`: Statistics for numeric functional traits
- `trait_proportion()`: Proportion of categorical traits
- `tv2vegtable()`: Import data sets from **Turboveg 2**
- `used_concepts()`: List of taxonomic concepts applied in the records
- `used_synonyms()`: List of synonyms applied in the records
- `vegtable2kml()`: Display plot locations in **Google Earth**
- `write_juice()`: Export tables to **Juice**

Dissecting vegtable

vegtable objects

S4 objects (definition, prototype, validation, methods) organized in slots.

- **description** (metadata)
- **species** (a `taxlist` object)
- **header** (plot information)
- **relations** (tables with relationships to *header*)
- **samples** (record table)
- **layers** (tables with relationships to *samples*)
- **coverconvert** (coverage/abundance conversions)



Dissecting vegtable

Slot species (`obj@species`)

Taxonomic information of *recorded organisms* stored in a `taxlist` object.

- Relationships of usage names and taxon concepts
- Parent-child relationships and taxonomic ranks
- Functional traits
- Connected through **TaxonUsageID**

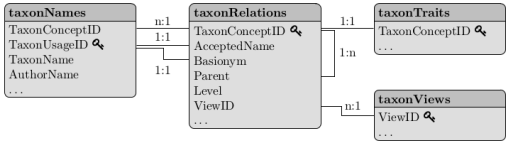
`taxlist` supports different degrees of information completeness/resolution

```
Kenya_veg@species
```

```
## object size: 554 Kb
## validation of 'taxlist' object: TRUE
##
## number of taxon usage names: 3164
## number of taxon concepts: 2392
## trait entries: 102
## number of trait variables: 1
## taxon views: 3
##
## concepts with parents: 2237
## concepts with children: 957
##
## hierarchical levels: form < variety < subspecies < speci
## number of concepts in level form: 0
## number of concepts in level variety: 65
## number of concepts in level subspecies: 52
## number of concepts in level species: 1422
## number of concepts in level complex: 0
## number of concepts in level genus: 699
## number of concepts in level family: 154
```

Dissecting vegtable

Slot species (`obj@species`)



Alvarez & Luebert (2018)

```
summary(Kenya_veg@species, "Cyperus involucratus",
        secundum="secundum")
```

```
## ------------------------------
## concept ID: 51757
## view ID: 1 - African Plant Database (2012)
## level: species
## parent: 54853 Cyperus L.
##
## # accepted name:
## 51757 Cyperus involucratus Rottb.
##
## # synonyms (1):
## 53973 Cyperus flabelliformis Rottb.
## ------------------------------
```

```
summary(Kenya_veg@species, 54853, secundum="secundum")
```

```
## ------------------------------
## concept ID: 54853
## view ID: 2 - Taxonomic Name Resolution Service (2018)
## level: genus
## parent: 55959 Cyperaceae NA
##
## # accepted name:
## 54855 Cyperus L.
## ------------------------------
```

Dissecting vegtable

Slot header (`obj@header`)
Main table (data frame) including information on plot observations (relevés).

- Identifiers
- Time and location of records
- Environmental information (e.g. soil sample analyses)
- Remarks
- *Statistics*

Variable **ReleveID** is mandatory in header.

```
head(Kenya_veg@header)
```

```
##     ReleveID COUNTRY REFERENCE TABLE_NR NR_IN_TAB
## 358      358      KE      2974        1         1
## 359      359      KE      2974        1         2
## 360      360      KE      2974        1         3
## 361      361      KE      2974        1         4
## 362      362      KE      2974        1         5
## 363      363      KE      2974        1         6
##     COVERSCALE DATE SURF_AREA ALTITUDE EXPOSITION
## 358         01 <NA>        NA       NA        <NA>
## 359         01 <NA>        NA       NA        <NA>
## 360         01 <NA>        NA       NA        <NA>
## 361         01 <NA>        NA       NA        <NA>
## 362         01 <NA>        NA       NA        <NA>
## 363         01 <NA>        NA       NA        <NA>
##     INCLINATIO COV_TOTAL TREE_HIGH     REMARKS
## 358         NA        NA        NA Mount Nyiro
## 359         NA        NA        NA Mount Nyiro
## 360         NA        NA        NA Mount Nyiro
## 361         NA        NA        NA Mount Nyiro
## 362         NA        NA        NA Mount Nyiro
## 363         NA        NA        NA Mount Nyiro
##     LONGITUDE LATITUDE PH_H2O
## 358   36.8167   2.1833     NA
## 359   36.8167   2.1833     NA
```

Dissecting vegtable

Slot header (`obj@samples`)

Data frame incluging the records of taxa (inserted as taxon usage names) in plot observations (relevés).

- Occurrence, frequency, abundance, sociability, etc.
- Relationships to layers.
- Relationships to collected specimens.
- Relationships to individuals.

Variables ** releveID** and **TaxonUsageID** are mandatory.

```
head(Kenya_veg@samples)
```

```
##      ReleveID TaxonUsageID COVER_CODE LAYER
## 5658      358           18          +     0
## 5659      358          220          r     0
## 5660      358          233          +     0
## 5661      358          287          +     0
## 5662      358          407          r     0
## 5663      358        54983          r     0
##      SOCIABILIT INDIVID br_bl b_bbds
## 5658       <NA>    <NA>     +   <NA>
## 5659       <NA>    <NA>     r   <NA>
## 5660       <NA>    <NA>     +   <NA>
## 5661       <NA>    <NA>     +   <NA>
## 5662       <NA>    <NA>     r   <NA>
## 5663       <NA>    <NA>     r   <NA>
```

Dissecting vegtable

Slot coverconvert (`obj@coverconvert`)

An S4 object used for conversion of abundance scale to cover percentage.

- Cover convesion tables.
- Information on class limits (bottom and top)
- Conversion by three alternative rules (bottom, mean, top)
- Only records that make sense to convert

```
names(Kenya_veg@coverconvert)
```

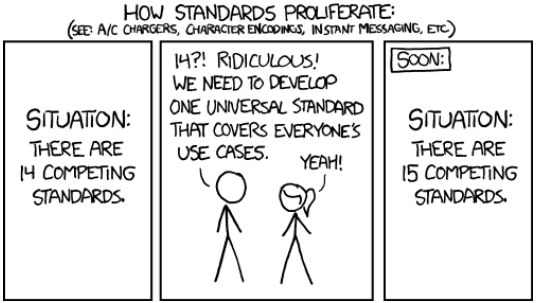
```
## [1] "br_bl"  "b_bbds" "ordin."
```

```
summary(Kenya_veg@coverconvert$br_bl)
```

```
## ## Number of cover scales: 1
##
## * scale 'br_bl':
##   Levels    Range
## 1      r    0 - 1
## 2      +    0 - 1
## 3      1    1 - 5
## 4      2   5 - 25
## 5      3  25 - 50
## 6      4  50 - 75
## 7      5 75 - 100
```

For the Future

- Multiple taxon views implemented in `taxlist`
- Import/export functions for **Veg-X**
- Better interaction with **Juice** (ideally in Linux)
- Better interaction with **Turboveg**
- Implementation in **sPlot**
- Contributions to **GIVD**
- Calculation of diversity indices (Shannon's H, Constancy values, Mean cover, etc.)
- Summaries for communities



xkcd.com

THANKS!

For discussion:

https://stackoverflow.com/users/5846398/miguel-alvarez

https://github.com/kamapu/vegtable/issues

https://www.facebook.com/groups/ecologyinr

...and also visit me at https://kamapu.github.io/