CODEBOOK

Tingting

7 5 2019

https://github.com/rubenarslan/codebook
(https://github.com/rubenarslan/codebook)

Generate markdown codebooks from the attributes of the variables in data frame

```
knitr::opts_chunk$set(
  warning = TRUE, # show warnings during codebook generation
  message = TRUE, # show messages during codebook generation
  error = TRUE, # do not interrupt codebook generation in case of errors,
                # usually makes debugging easier, and sometimes half a codebook
                # is better than none
  echo = FALSE  # don't show the R code
)
ggplot2::theme_set(ggplot2::theme_bw())
```

```
## Registered S3 methods overwritten by 'ggplot2':
##   method          from
##   [.quosures      rlang
##   c.quosures      rlang
##   print.quosures rlang
```

```
pander::panderOptions("table.split.table", Inf)

load("C:/Users/tingtinf/Downloads/t1.RData")
codebook_data = t1
```

##Sometimes, the metadata is not set up in such a way that codebook can leverage it fully. These functions help fix this.

```
## Registered S3 methods overwritten by 'codebook':
##   method              from
##   as_factor.character forcats
##   as_factor.numeric   forcats
```

If you are not using formr, the codebook package needs to guess which items

form a scale. The following line finds item aggregates with names like this:

scale = scale_1 + scale_2R + scale_3R

identifying these aggregates allows the codebook function to

automatically compute reliabilities.

However, it will not reverse items automatically.

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

#generating a codebook is as simple as calling codebook from a chunk in an rmarkdown document.

```
## No missing values.
```

Metadata

Description

Dataset name: codebook_data

The dataset has N=205 rows and 9 columns. 205 rows have no missing values on any column.

- Metadata for search engines
  - **Date published**: 2019-05-07
    - **keywords**: *time*, *status*, *sex*, *age*, *year*, *thickness*, *ulcer*, *exit.date* and *site*

Variables

time

| Distribution | Summary statistics |

time

0 missing values.

status
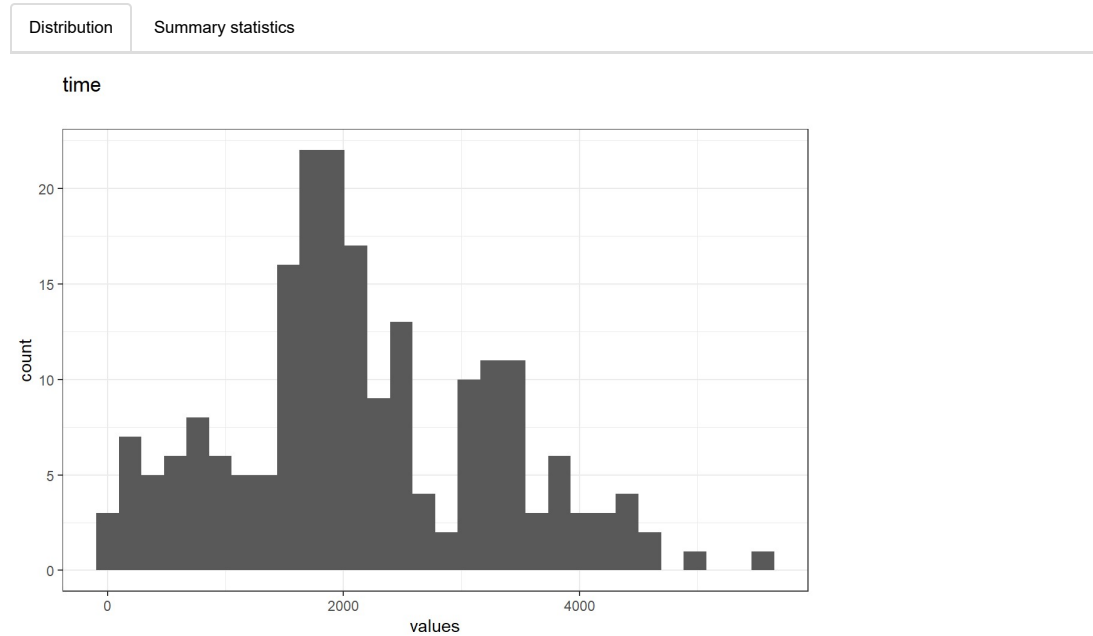
| Distribution | Summary statistics |

status

0 missing values.

sex
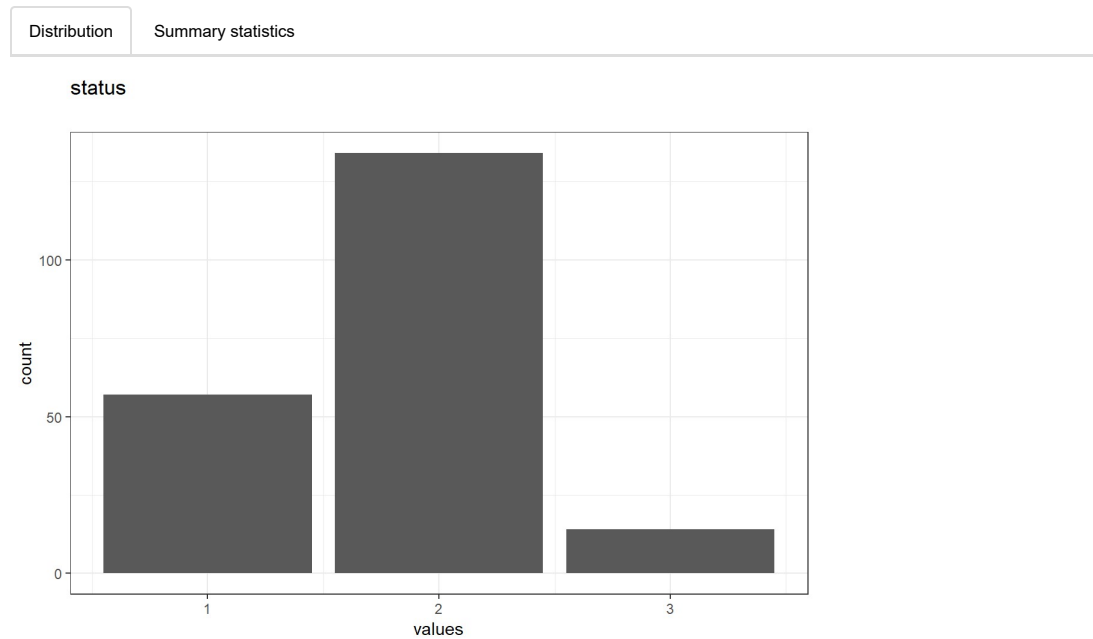
| Distribution | Summary statistics |

sex

0 missing values.

age
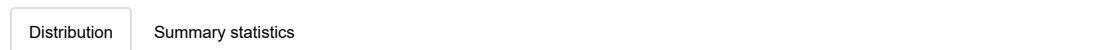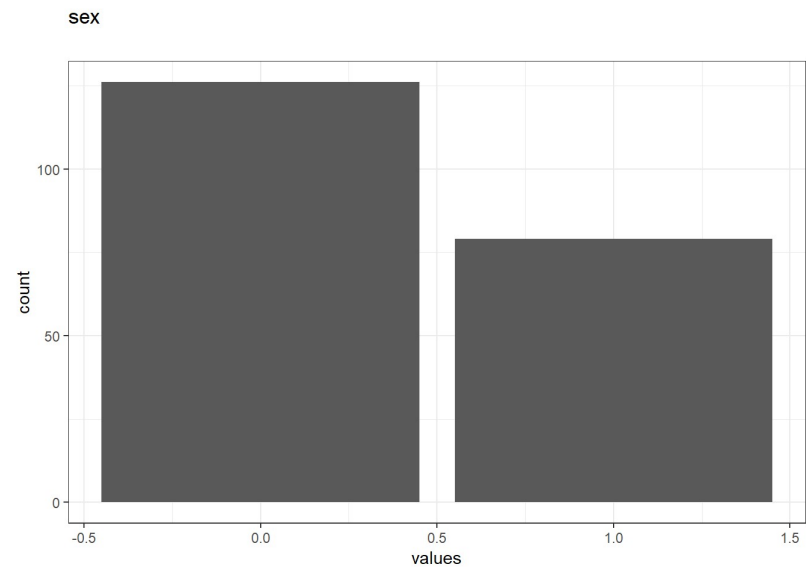
Distribution	Summary statistics

age

0 missing values.

year
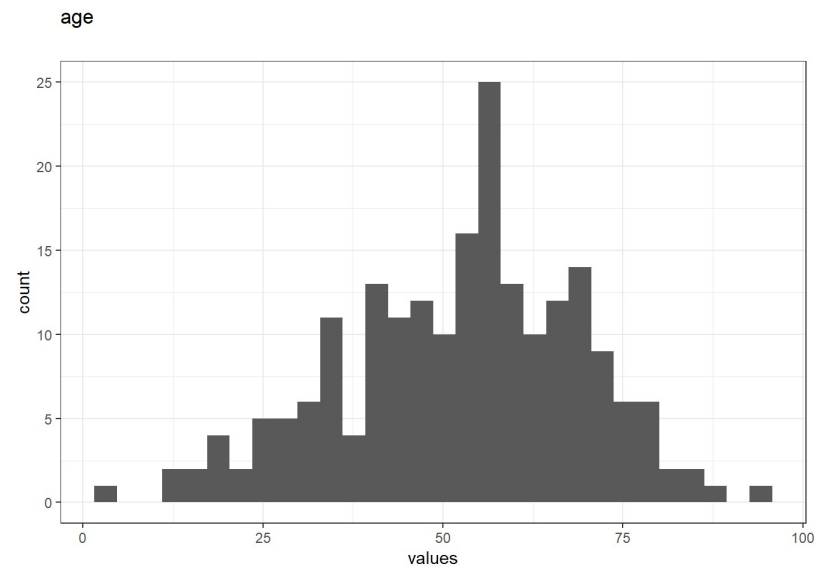
Distribution	Summary statistics

year
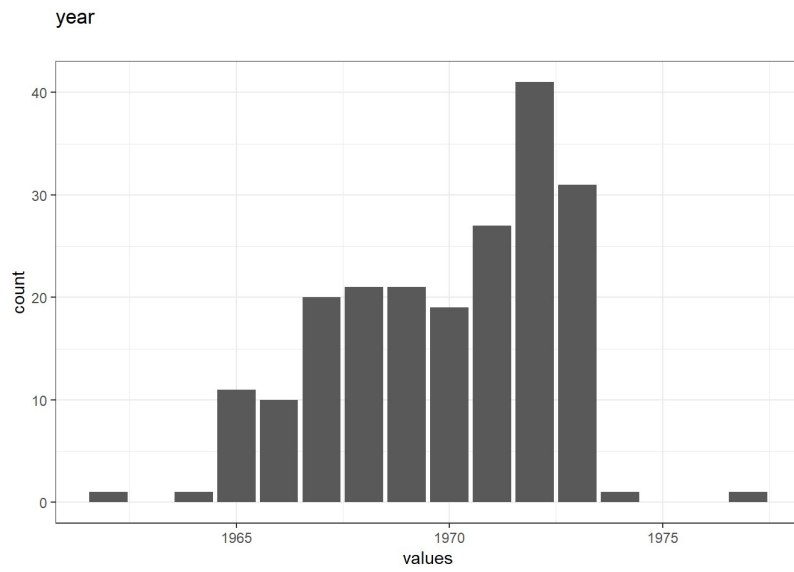
0 missing values.

thickness

| Distribution | Summary statistics |

thickness
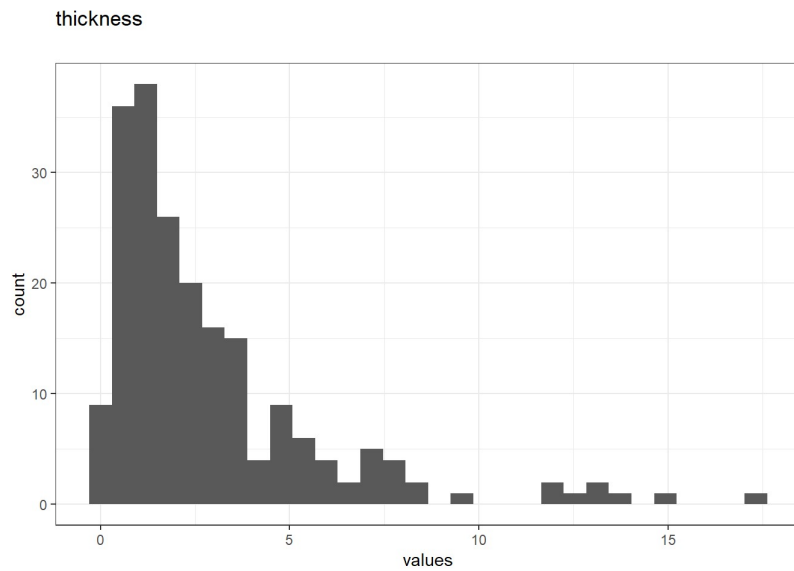
0 missing values.

ulcer

| Distribution | Summary statistics |

ulcer
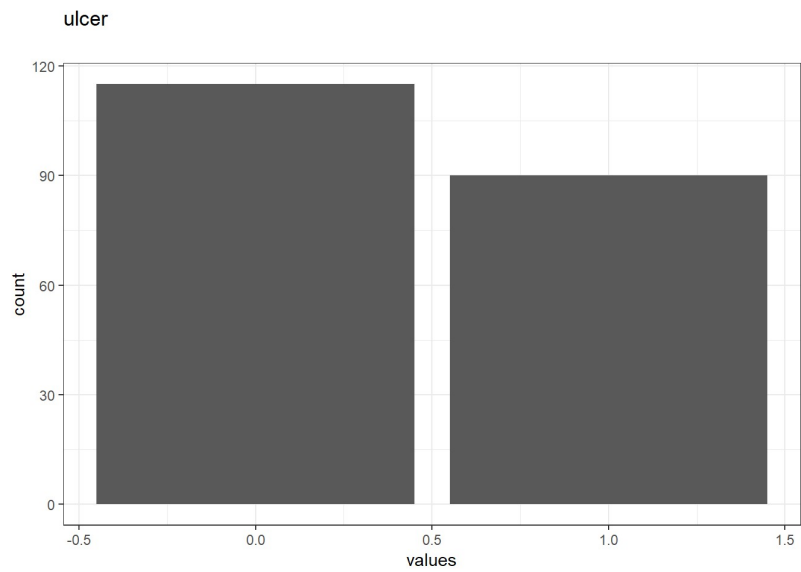
0 missing values.

exit.date

Distribution	Summary statistics

```
## 205  unique, categorical values, so not shown.
```

0 missing values.

site

Distribution	Summary statistics

site
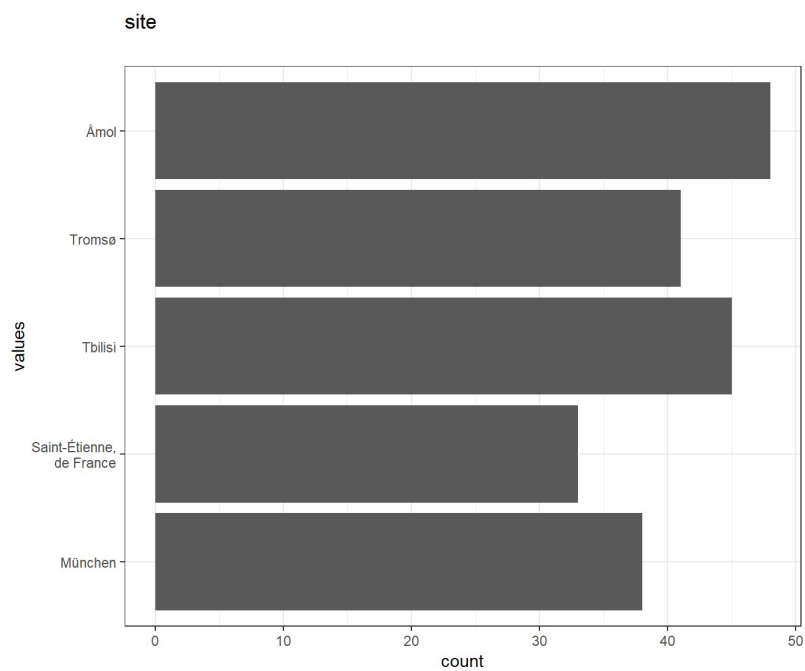
0 missing values.

Missingness report

Codebook table

Search:

name	data_type	missing	complete	n	empty	n_unique	median	min	max	mean
	A		/			/				
time	integer	0	205	205						2152.8
status	integer	0	205	205						1.79
sex	integer	0	205	205						0.39
age	integer	0	205	205						52.46
year	integer	0	205	205						1969.91
thickness	numeric	0	205	205						2.92
ulcer	integer	0	205	205						0.44
exit.date	Date	0	205	205		205	1999-07-02	1999-01-01	2000-01-01	
site	character	0	205	205	0	5		4	24	

Showing 1 to 9 of 9 entries

Previous   1   Next

- JSON-LD metadata
  The following JSON-LD can be found by search engines, if you share this codebook publicly on the web.

```
{
  "name": "codebook_data",
  "datePublished": "2019-05-07",
  "description": "The dataset has N=205 rows and 9 columns.\n205 rows have no missing values on any column.\n\n## Table o
f variables\nThis table contains variable names, labels, their central tendencies and other attributes.\n\n|name        |data_
type |missing |complete |n    |empty |n_unique |median   |min        |max         |mean    |sd       |p0      |p25  |p50  |p75
 |p100   |hist      |\n|:--------|:--------|:------|:-------|:---|:-----|:-------|:--------|:--------|:--------|:--
----|:------|:---|:---|:---|:---|:----|:------|:-------|\n|time        |integer   |0        |205      |205 |NA    |NA       |N
A      |NA         |2152.8  |1122.06 |10   |1525 |2005 |3042 |5565 |▁▃█▅▂▂▁▁_____|\n|status     |integer
 |0     |205      |205 |NA    |NA       |NA      |NA         |NA      |1.79    |0.55 |1    |1    |2    |2    |
3       |____█__    |\n|sex        |integer   |0        |205      |205 |NA    |NA      |NA       |NA         |N
A      |0.39    |0.49 |0    |0    |0    |1    |1    |█_____▅|\n|age        |integer   |0        |205      |205 |NA    |N
A      |NA         |NA       |52.46   |16.67 |4    |42   |54   |65   |95   |___▃▅▇██▅▂___|\n|year       |
integer   |0        |205      |205 |NA    |NA       |NA      |NA         |NA      |1969.91 |2.58    |1962 |1968 |1970 |
1972 |1977 |___▂▅▇█▅___|\n|thickness |numeric   |0        |205      |205 |NA    |NA       |NA       |N
A      |2.92    |2.96      |0.1  |0.97 |1.94 |3.56 |17.42 |█▅▂▁▁_____|\n|ulcer     |integer   |0        |205      |205
 |NA    |NA       |NA         |NA       |0.44    |0.5  |0    |0    |0    |1    |1    |█_____▅|\n|ex
it.date |Date      |0        |205      |205 |NA    |205      |1999-07-02 |1999-01-01 |2000-01-01 |NA      |NA       |NA   |N
A    |NA   |NA   |NA   |\n|site       |character |0        |205      |205 |NA    |NA       |NA       |NA         |4       |
24   |NA      |NA         |NA   |NA   |NA   |NA   |NA   |\n\n### Note\nThis dataset was automatically describe
d using the [codebook R package](https://rubenarslan.github.io/codebook/) (version 0.8.0).",
  "keywords": ["time", "status", "sex", "age", "year", "thickness", "ulcer", "exit.date", "site"],
  "@context": "http://schema.org/",
  "@type": "Dataset",
  "variableMeasured": [
    {
      "name": "time",
      "@type": "propertyValue"
    },
    {
      "name": "status",
      "@type": "propertyValue"
    },
    {
      "name": "sex",
      "@type": "propertyValue"
    },
    {
      "name": "age",
      "@type": "propertyValue"
    },
    {
      "name": "year",
      "@type": "propertyValue"
    },
    {
      "name": "thickness",
      "@type": "propertyValue"
    },
    {
      "name": "ulcer",
      "@type": "propertyValue"
    },
    {
      "name": "exit.date",
      "@type": "propertyValue"
    },
    {
      "name": "site",
      "@type": "propertyValue"
    }
  ]
}`
```