# Package 'codebookr'

October 12, 2022

**Title** Create Codebooks from Data Frames

**Version** 0.1.5

**Maintainer** Brad Cannell <brad.cannell@gmail.com>

**Description** Quickly and easily create codebooks (i.e. data dictionaries) directly from a data frame.

**License** MIT + file LICENSE

**Encoding** UTF-8

**RoxygenNote** 7.2.1

**URL** https://github.com/brad-cannell/codebookr,
https://brad-cannell.github.io/codebookr/

**BugReports** https://github.com/brad-cannell/codebookr/issues

**Depends** R (>= 2.10)

**LazyData** true

**Suggests** hms, testthat

**Imports** haven (>= 2.5.0), flextable, dplyr, officer, purrr, rlang, stringr, tibble, tidyr

**NeedsCompilation** no

**Author** Brad Cannell [aut, cre, cph]

**Repository** CRAN

**Date/Publication** 2022-08-10 00:20:02 UTC

## R topics documented:


cb_add_col_attributes . . . . . 2
codebook . . . . . 2
study . . . . . 4

**Index** **5**

## cb_add_col_attributes *Add Attributes to Columns*

### Description

Add arbitrary attributes to columns (e.g., description, source, column type). These attributes can later be accessed to fill in the column attributes table. Note: Column type is the type you say it is: categorical or numeric, Data type is what the computer says it is.

### Usage

```
cb_add_col_attributes(df, .x, ...)
```

### Arguments

| | |
|---|---|
| df | Data frame of interest |
| .x | Column of interest in df |
| ... | Arbitrary list of attributes (i.e., attribute = "value") |

### Value

Data frame

## codebook *Automate creation of a data codebook*

### Description

The codebook function assists with the creation of a codebook for a given data frame.

### Usage

```
codebook(
  df,
  title = NA,
  subtitle = NA,
  description = NA,
  keep_blank_attributes = FALSE,
  no_summary_stats = NULL
)
```

## Arguments

| | |
|---|---|
| df | The data frame the codebook will describe |
| title | An optional title that will appear at the top of the Word codebook document |
| subtitle | An optional subtitle that will appear at the top of the Word codebook document |
| description | An optional text description of the dataset that will appear on the first page of the Word codebook document |
| keep_blank_attributes | TRUE or FALSE. By default, the column attributes table will omit the Column description, Source information, Column type, value labels, and skip pattern rows from the column attributes table in the codebook document if those attributes haven't been set. In other words, it won't show blank rows for those attributes. Passing TRUE to the keep_blank_attributes argument will cause the opposite to happen. The column attributes table will include a Column description, Source information, Column type, and value labels row for every column in the data frame - even if they don't have those attributes set. |
| no_summary_stats | A character vector of column names. The summary statistics will not be added to column attributes table for any column passed to this argument. This can be useful when a column contains values that are sensitive or may be used to identify individual people (e.g., names, addresses, etc.) and the individual values for that column should not appear in the codebook. |

## Details

Codebook expects that df is a data frame that you have read into memory from a saved data file. Please provide the path to the saved data file. This function gets selected attributes about file saved at path and stores those attributes in a data frame, which is later turned into a flextable and added to the codebook document.

## Value

An rdocx object that can be printed to a Word document

## Examples

```
study_codebook <- codebook(
  df = study,
  title = "My Example Study",
  subtitle = "A Subtitle for My Example Study Codebook",
  description = "Brief (or long) description of the data."
)

# Create the Word codebook document
print(study_codebook, path = "example_codebook.docx")
```

`study` *Simulated study data.*

## Description

This is the code to create the study data - a simulated dataset that can be used to demonstrate how to use the codebook package.

## Usage

```
study
```

## Format

A data frame with 20 rows and 10 variables:

**id** Participant's study identification number

**address** Participant's home address

**sex** Biological sex of the participant assigned at birth, female/male

**date** Participant's date of enrollment

**time** Participant's time of enrollment

**date_time** Participant's date and time of enrollment

**days** Total number of days the participant was enrolled in the study

**height** Participant's height in inches at date of enrollment

**likert** An example Likert scale item, 1-5

**outcome** Participant experienced the outcome of interest, TRUE or FALSE

# Index

∗ **datasets**
study, 4

cb_add_col_attributes, 2
codebook, 2

study, 4