# Package ‘socialrisk’

February 15, 2023

**Type** Package

**Title** Identifying Patient Social Risk from Administrative Health Care Data

**Version** 0.5.1

**Description** Social risks are increasingly becoming a critical component of health care research. One of the most common ways to identify social needs is by using ICD-10-CM ``Z-codes.'' This package identifies social risks using varying taxonomies of ICD-10-CM Z-codes from administrative health care data. The conceptual taxonomies come from:
Centers for Medicare and Medicaid Services (2021) <`https://www.cms.gov/files/document/zcodes-infographic.pdf`>,
Reidhead (2018) <`https://web.mhanet.com/`>,
A Arons, S DeSilvey, C Fichtenberg, L Gottlieb (2018) <`https://sirenetwork.ucsf.edu/tools-resources/resources/compendium-medical-terminology-codes-social-risk-factors`>.

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.2

**Imports** dplyr, magrittr, stringr, rlang, tidyselect, tidyr,

**URL** `https://github.com/WYATTBENSKEN/multimorbidity`

**BugReports** `https://github.com/WYATTBENSKEN/multimorbidity/issues`

**Suggests** rmarkdown, knitr

**VignetteBuilder** knitr

**Depends** R (>= 3.5.0)

**NeedsCompilation** no

**Author** Wyatt Bensken [aut, cre] (<`https://orcid.org/0000-0002-2597-9732`>)

**Maintainer** Wyatt Bensken <`wpb27@case.edu`>

**Repository** CRAN

**Date/Publication** 2023-02-15 21:40:02 UTC

# R topics documented:

---

## `clean_data` *Prepare our administrative data for analysis*

### Description

`clean_data` returns a dataset which has been transformed and cleaned for subsequent functions in this package.

### Usage

```
clean_data(dat = NULL, style = "long", id = NULL, prefix_dx = "dx")
```

### Arguments

| | |
|---|---|
| `dat` | dataset |
| `style` | long, the default, is one diagnosis column per row whereas wide is multiple diagnosis columns |
| `id` | unique patient identifier variable name |
| `prefix_dx` | the variable prefix for the diagnosis columns (defaults to "dx"), in quotes |

### Details

This function takes our raw administrative data, in a number of different forms, and prepares it in a way which allows the other functions in this package to easily work with it. It is recommended to run this package on all data regardless of setup.

### Value

dataframe with multiple rows per patient, which has re-structured their administrative data

### Examples

```
clean_data(dat = i10_wide, id = patient_id, style = "wide", prefix_dx = "dx")
```

---

## `i10_wide` *Example administrative data.*

---

### Description

A dataset with fake patient data for 5 patients with ICD-10 diagnosis codes.

### Usage

```
data(i10_wide)
```

### Format

A data frame with 29 rows and 11 variables:

| | |
|---|---|
| **patient_id** | patient_id |
| **sex** | patient's sex (male or female) |
| **date_of_serv** | the date of service for the fake claim |
| **dx1** | first diagnosis |
| **dx2** | second diagnosis |
| **dx3** | third diagnosis |
| **dx4** | fourth diagnosis |
| **dx5** | fifth diagnosis |
| **visit_type** | inpatient (ip) or outpatient(ot) |
| **hcpcs** | HCPCS code |
| **icd_version** | Which version of ICD the row is. 9 = ICD-9, 0 = ICD-10 |

### Source

This was created by the package author.

---

## `socialrisk` *Social Risk*

---

### Description

`socialrisk` returns a summary dataset containing indicators of social risk, which vary based on the taxonomy command, for each patient.

### Usage

```
socialrisk(dat = NULL, id = NULL, dx = "dx", taxonomy = "cms")
```

## Arguments

| | |
|---|---|
| `dat` | dataset which has been properly prepared in long format |
| `id` | variable of the unique patient identifier |
| `dx` | the column with the diagnoses (defaults to 'dx') |
| `taxonomy` | the taxonomy one wishes to use for social risk, with options of "cms" (default), "mha", and "siren" |

## Details

This function uses data which has been properly prepared to identify and flag social risks.

## Value

dataframe with one row per patient, a column for their patient id, a column with whether they have any social risk, a column with the number of social risk domains, and columns with indicator variables for each social risk

## Examples

```
data <- clean_data(dat = i10_wide, id = patient_id, style = "wide", prefix_dx = "dx")
socialrisk(dat = data, id = patient_id, dx = dx, taxonomy = "cms")
```

# Index