

# Package ‘casen’

June 21, 2022

**Type** Package

**Title** Metodos De Estimacion Con Diseno Probabilistico y Estratificado en Encuesta CASEN (Estimation Methods with Probabilistic Stratified Sampling in CASEN Survey)

**Version** 0.2.0

**URL** <https://pacha.dev/casen/>

**BugReports** <https://github.com/pachadotdev/casen/issues/>

**Description** Funciones para realizar estadística descriptiva e inferencia con el diseño complejo de la Encuesta CASEN (Encuesta de Caracterización Socio-Económica). Incluye datasets que permiten armonizar los códigos de comunas que cambian entre años y permite convertir a los códigos oficiales de SUBDERE.  
(Functions to compute descriptive and inferential statistics with CASEN Survey [Socio-Economic Characterization Survey] complex design. Includes datasets to harmonize commune codes that change across years and allows to convert to official SUBDERE codes.)

**License** Apache License (>= 2)

**Depends** R (>= 3.5.0)

**Imports** rlang, stats, utils, dplyr, haven, labelled, glue, magrittr, tidy, survey, srvyr, purrr

**Suggests** knitr, rmarkdown, covr, testthat, broom, janitor

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.2

**VignetteBuilder** knitr

**NeedsCompilation** no

**Author** Mauricio Vargas [aut, cre] (<<https://orcid.org/0000-0003-1017-7574>>),  
Ricardo Aravena [ths],  
Klaus Lehmann [ctb],  
Pontificia Universidad Católica de Chile [cph],  
Ministerio de Desarrollo Social (MDS) [dte],  
Subsecretaría de Desarrollo Regional (SUBDERE) [dte]

**Maintainer** Mauricio Vargas <mavargas11@uc.cl>

**Repository** CRAN

**Date/Publication** 2022-06-21 07:30:10 UTC

## R topics documented:

armonizar_oficio . . . . .	2
armonizar_rama . . . . .	3
armonizar_region . . . . .	4
casen_2017_los_rios . . . . .	4
codigos_casen . . . . .	5
codigos_subdere . . . . .	6
configuracion_disenio . . . . .	6
descargar_casen_github . . . . .	7
descargar_casen_mds . . . . .	8
mediana_agrupada . . . . .	8
media_agrupada . . . . .	9
modelo_lineal_generalizado . . . . .	9
percentiles_agrupados . . . . .	10

<b>Index</b>	<b>11</b>
--------------	-----------

---

armonizar_oficio	<i>Armoniza los codigos de oficio</i>
------------------	---------------------------------------

---

### Description

Convierte las etiquetas de los oficios para usar de manera uniforme los nombres de CASEN 2017, debiendo fusionar algunas categorias que se intersectan (ejemplo: "Directivos de la Adm. Publica y Empresas" con "Miembros del Poder Ejecutivo, Legislativo y Directivos de la Adm. Publica"). El procedimiento consiste en buscar la columna oficio (antes del año 2011) u oficio1 (desde el año 2011) y aplicar expresiones regulares de acuerdo a la codificacion respectiva.

### Usage

```
armonizar_oficio(d)
```

### Arguments

d una encuesta CASEN en formato tibble o data.frame

### Value

un tibble con la columna de oficio transformada

## Examples

```
## Not run:  
casen1990 <- readRDS("casen1990.rds")  
armonizar_oficio(casen1990)  
  
## End(Not run)
```

---

armonizar_rama	<i>Armoniza los codigos de rama ocupacional</i>
----------------	---

---

## Description

Convierte las etiquetas de la rama ocupacional para usar de manera uniforme los nombres de CASEN 2017, debiendo fusionar algunas categorías que se intersectan (ejemplo: "Establecimientos Financieros" y "Servicios de Gobierno y Financieros"). El procedimiento consiste en buscar la columna rama (antes del año 2011) o rama1 (desde el año 2011) y aplicar expresiones regulares de acuerdo a la codificación respectiva.

## Usage

```
armonizar_rama(d)
```

## Arguments

d una encuesta CASEN en formato tibble o data.frame

## Value

un tibble con la columna de rama transformada

## Examples

```
## Not run:  
casen1990 <- readRDS("casen1990.rds")  
armonizar_rama(casen1990)  
  
## End(Not run)
```

armonizar\_region      *Armoniza los codigos de regiones*

---

### **Description**

Convierte las etiquetas de las regiones para usar de manera uniforme los nombres de CASEN 2017. El procedimiento consiste en buscar la columna r (antes del año 2000) o region (desde el año 2000) y aplicar expresiones regulares de acuerdo a la codificación respectiva.

### **Usage**

```
armonizar_region(d)
```

### **Arguments**

d                      una encuesta CASEN en formato tibble o data.frame

### **Value**

un tibble con la columna de region transformada

### **Examples**

```
## Not run:  
casen1990 <- readRDS("casen1990.rds")  
armonizar_region(casen1990)  
  
## End(Not run)
```

---

casen\_2017\_los\_rios      *Sub-conjunto CASEN 2017*

---

### **Description**

Datos de demostración limitados a las comunas de la Región de Los Ríos con representatividad comunal (Valdivia, Los Ríos, Paillaco, Panguipulli, La Unión y Río Bueno).

### **Usage**

```
casen_2017_los_rios
```

**Format**

Un data frame con 8055 observaciones en las siguientes 7 variables.

expc factor de expansion comunal  
 varstrat estratos de varianza  
 varunit conglomerados de varianza  
 ytotcorh ingreso total del hogar corregido  
 sexo sexo  
 comuna comuna  
 pobreza situacion de pobreza por ingresos

**Source**

Elaboracion propia con base en datos de Ministerio de Desarrollo Social (MDS).

**Examples**

casen\_2017\_los\_rios

---

codigos_casen	<i>Codigos de comunas encuesta CASEN</i>
---------------	--

---

**Description**

Proporciona todos los codigos de comunas usados en la encuesta CASEN y registra sus cambios en el tiempo.

**Usage**

codigos\_casen

**Format**

Un tibble con 683 filas y las siguientes cuatro columnas.

valido\_desde primer anio en que se usa el codigo  
 valido\_hasta ultimo anio en que se usa el codigo  
 codigo\_casen codigo de comuna, valido solo en la encuesta  
 codigo\_subdere\_2017 codigo oficial de comuna, tiene validez legal

**Source**

Elaboracion propia con base en datos de Ministerio de Desarrollo Social (MDS).

**Examples**

codigos\_casen

---

codigos_subdere	<i>Codigos Unicos Territoriales (CUT)</i>
-----------------	---

---

**Description**

Corresponde a los codigos oficiales de comunas.

**Usage**

```
codigos_subdere
```

**Format**

Un tibble con 346 filas y las siguientes seis columnas.

```
codigo_comuna  codigo oficial de comuna
nombre_comuna  nombre oficial de comuna
codigo_provincia codigo oficial de provincia
nombre_provincia nombre oficial de provincia
codigo_region  codigo oficial de region
nombre_region  nombre oficial de region
```

**Source**

Elaboracion propia con base en datos de Subsecretaria de Desarrollo Regional (SUBDERE).

**Examples**

```
codigos_subdere
```

---

configuracion_disenio	<i>Diseño complejo para estadística descriptiva e inferencia</i>
-----------------------	--

---

**Description**

Usa los factores de expansion, conglomerados y estratos para definir un objeto de disenio complejo.

**Usage**

```
configuracion_disenio(
  datos,
  variable = "ytotcorh",
  agrupacion = "region",
  peso = "expr",
  conglomerado = "varunit",
  estrato = "varstrat"
)
```

**Arguments**

datos	un data.frame o tibble con la encuesta CASEN (o un subconjunto acotado a una region, etc)
variable	una columna de tipo numerico, por ejemplo ytotcorh que es la opcion por defecto
agrupacion	una columna de tipo texto/factor, por ejemplo region que es la opcion por defecto
peso	una columna de tipo numerico, por defecto es expr que corresponde al factor de expansion regional de acuerdo al manual CASEN 2017
conglomerado	una columna de tipo numerico, por defecto es varunit de acuerdo al manual CASEN 2017
estrato	una columna de tipo numerico, por defecto es varunit de acuerdo al manual CASEN 2017

**Value**

Una lista con el diseño y los grupos.

**Examples**

```
cd <- configuracion_disenio(casen_2017_los_rios, "ytotcorh", c("comuna", "sexo"), "expc")
cd$disenio
cd$grupos
```

---

descargar\_casen\_github

*Descarga la encuesta CASEN de GitHub*

---

**Description**

Obtiene los archivos de la encuesta en formato R.

**Usage**

```
descargar_casen_github(anios = NULL, carpeta = NULL)
```

**Arguments**

anios	si no se indica un anio, descarga todos los anios disponibles
carpeta	se debe especificar una carpeta de descarga

**Value**

Los archivos rds de la encuesta CASEN descargados desde GitHub.

**Examples**

```
# descargar CASEN 1990 en carpeta temporal
descargar_casen_github(1990, tempdir())
```

---

descargar_casen_mds	<i>Descarga la encuesta CASEN del sitio web del Ministerio de Desarrollo Social</i>
---------------------	---

---

**Description**

Obtiene los archivos de la encuesta en formato RAR. No descomprime ni modifica los datasets originales.

**Usage**

```
descargar_casen_mds(anios = NULL, carpeta = NULL)
```

**Arguments**

anios	si no se indica un año, descarga todos los años disponibles
carpeta	se debe especificar una carpeta de descarga

**Value**

Los archivos comprimidos de la encuesta CASEN descargados desde el sitio web del Ministerio de Desarrollo Social.

**Examples**

```
# descargar CASEN 1990 en carpeta temporal
descargar_casen_mds(1990, tempdir())
```

---

mediana_agrupada	<i>Mediana agrupada usando diseño complejo</i>
------------------	--

---

**Description**

Usa los factores de expansion, conglomerados y estratos para calcular correctamente las medianas comunales o regionales.

**Usage**

```
mediana_agrupada(disenio)
```

**Arguments**

disenio	la salida de 'configuracion_disenio()' que provee además los grupos y las variables en forma de lista
---------	---



**Value**

Una tabla con las medianas agrupadas y su intervalo de confianza.

**Examples**

```
cd <- configuracion_diseño(casen_2017_los_rios, "ytotcorh", c("comuna", "sexo"), "expc")
mediana_agrupada(cd)
```

---

media_agrupada	<i>Media agrupada usando diseño complejo</i>
----------------	--

---

**Description**

Usa los factores de expansión, conglomerados y estratos para calcular correctamente las medias comunales o regionales.

**Usage**

```
media_agrupada(diseño)
```

**Arguments**

diseño            la salida de 'configuracion\_diseño()' que provee además los grupos y las variables en forma de lista

**Value**

Una tabla con las medias agrupadas y su intervalo de confianza.

**Examples**

```
cd <- configuracion_diseño(casen_2017_los_rios, "ytotcorh", c("comuna", "sexo"), "expc")
media_agrupada(cd)
```

---

modelo_lineal_generalizado	<i>Modelo lineal generalizado con diseño complejo</i>
----------------------------	---

---

**Description**

Usa los factores de expansión, conglomerados y estratos para ajustar un modelo lineal generalizado con las variables definidas por el usuario.

**Usage**

```
modelo_lineal_generalizado(diseño, modelo = "ytotcorh ~ sexo")
```

**Arguments**

disenio	la salida de 'configuracion_disenio()' que provee ademas los grupos y las variables en forma de lista
modelo	una expresion simbolica escrita como cadena de texto, consulta [stats::formula()] para los detalles de modelos con componentes logaritmicas, cuadraticas, etc.

**Value**

Una lista cuyas clases son svyglm, glm y lm.

**Examples**

```
cd <- configuracion_disenio(casen_2017_los_rios, "ytotcorh", c("comuna", "sexo"), "expc")
modelo_lineal_generalizado(cd, "ytotcorh ~ comuna + sexo")
```

---

percentiles\_agrupados *Percentiles agrupados usando diseño complejo*

---

**Description**

Usa los factores de expansion, conglomerados y estratos para calcular correctamente los percentiles comunales o regionales.

**Usage**

```
percentiles_agrupados(disenio, percentiles = 0.7)
```

**Arguments**

disenio	la salida de 'configuracion_disenio()' que provee ademas los grupos y las variables en forma de lista
percentiles	percentiles a calcular, si no se especifica calcula el percentil 70

**Value**

Una tabla con los percentiles y su error estandar.

**Examples**

```
cd <- configuracion_disenio(casen_2017_los_rios, "ytotcorh", c("comuna", "sexo"), "expc")
percentiles_agrupados(cd, 0.7)
```

# Index

## \* datasets

- casen\_2017\_los\_rios, [4](#)
- codigos\_casen, [5](#)
- codigos\_subdere, [6](#)

- armonizar\_oficio, [2](#)
- armonizar\_rama, [3](#)
- armonizar\_region, [4](#)

- casen\_2017\_los\_rios, [4](#)
- codigos\_casen, [5](#)
- codigos\_subdere, [6](#)
- configuracion\_diseno, [6](#)

- descargar\_casen\_github, [7](#)
- descargar\_casen\_mds, [8](#)

- media\_agrupada, [9](#)
- mediana\_agrupada, [8](#)
- modelo\_lineal\_generalizado, [9](#)

- percentiles\_agrupados, [10](#)