

# Package ‘regioncode’

October 14, 2022

**Title** Convert Region Names and Division Codes of China Over Years

**Version** 0.1.1

**Author** Yue Hu [aut, cre],  
Yufei Sun [aut],  
Wenquan Wu [aut],  
Yuyang Shi [ctb],  
Yujia Xu [ctb],  
Yuxin Pan [ctb],  
Haiting Tian [ctb],  
Weihang Shao [ctb],  
Yuanqian Chen [ctb],  
Xueyan Liu [ctb]

**Maintainer** Yue Hu <yuehu@tsinghua.edu.cn>

**Description** A fast tool to conquer the difficulties to convert various region names and administration division codes of Chinese regions. The current version enables seamlessly converting Chinese regions' formal names, common-used names, and codes between each other at the city level from 1986 to 2019.

**License** MIT + file LICENSE

**Encoding** UTF-8

**BugReports** <https://github.com/sammo3182/regioncode/issues>

**Depends** R(>= 3.6.0)

**Imports** dplyr, knitr, pinyin, textshaping

**Suggests** testthat (>= 3.0.0), rmarkdown

**LazyData** true

**RoxygenNote** 7.1.1

**Config/testthat/edition** 3

**VignetteBuilder** knitr

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2021-08-02 13:40:02 UTC

## R topics documented:

corruption . . . . .	2
regioncode . . . . .	3
<b>Index</b>	<b>5</b>

---

corruption	<i>China's Corruption Investigations Dataset</i>
------------	--

---

### Description

A dataset containing information on almost 20,000 officials who were investigated during Xi Jinping's anti-corruption campaign.

### Usage

corruption

### Format

A data frame with 6 variables:

**province** 2-digit province number

**prefecture** Prefecture name in Chinese

**county** County name in Chinese

**province\_id** 6-digit province number

**prefecture\_id** 6-digit province number

**county\_id** 6-digit province number

### Source

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/9QZRAD> #  
link where you got the data if there is

---

regioncode	<i>regioncode</i>
------------	-------------------

---

## Description

regioncode is developed to conquer the difficulties to convert various region names and administration division codes of Chinese regions. In the current version, regioncode enables seamlessly converting Chinese regions' formal names, common-used names, and geocodes between each other at the prefectural level from 1986 to 2019.

## Usage

```
regioncode(
  data_input,
  year_from = 1999,
  year_to = 2015,
  convert_to = "code",
  incomplete_name = "none",
  zhixiashi = FALSE,
  to_dialect = "none",
  to_pinyin = FALSE,
  province = FALSE
)
```

## Arguments

data_input	A character vector for names or a six-digit integer vector for division codes to convert.
year_from	A integer to define the year of the input. The default value is 1999.
year_to	A integer to define the year to convert. The default value is 2015.
convert_to	A character indicating the converting methods. At the prefectural level, valid methods include converting between codes in different years, from codes to region names, from region names to division codes, from region names or division codes to sociopolitical area names, and between names in different years. The current version automatically detect the type of the input. Users only need to choose the output to be codes (code), names (name) or area (area). The default option is code. When province is TRUE, one can also choose abbre, abbreToCode, abbreToName, and abbreToArea to convert between names/codes and abbreviations of provinces.
incomplete_name	A character to specify if a short name of region is used. See the Details for more information. The default is "none". Other options are "from", "to", and "both".
zhixiashi	A logic string to indicate whether treating division codes and names of municipality directly under the central government (Only makes a difference for prefectural-level conversion). The default value is FALSE.

<code>to_dialect</code>	A character indicating the language transformation. At the prefectural level, valid transformation include <code>dia_group</code> , <code>dia_sub_group</code> . At the province level, valid transformation is <code>dia_super</code> . The default value is "none". When <code>province</code> is TRUE, one can also choose <code>dia_super</code> to get the language zone of provinces.
<code>to_pinyin</code>	A logic string to indicate whether the output is in pinyin spelling instead of Chinese characters. The default is FALSE.
<code>province</code>	A logic string to indicate the level of converting. The default is FALSE.

### Details

In many national and regional data in China studies, the source applies incomplete names instead of the official, full name of a given region. A typical case is that "Xinjiang" is used much more often than "Xinjiang Weiwuer Zizhiqu" (the Xinjiang Uygur Autonomous Region) for the name of the province. In other cases the "Shi" (City) is often omitted to refer to a prefectural city. `regioncode` accounts this issue by offering the argument `incomplete_name`. The argument has four options: "none", "from", "to", and "both".

- "none": no short name will be used for either input or output;
- "from": input data is short names instead of the full, official ones;
- "to": output results will be short names;
- "both": both input and output are using short names.

The argument makes a difference only when code or name are chose in `convert_to`. Users can use this argument together with `name` to convert between names and incomplete names.

### Value

The function returns a character or numeric vector depending on what method is specified.

### Examples

```
# library(regioncode)

# regioncode(data_input = corruption$prefecture_id,
#            year_from = 2016,
#            year_to = 2017)
```

# Index

\* **datasets**

corruption, [2](#)

corruption, [2](#)

regioncode, [3](#)