

# regioncode: Convert Region Names and Division Codes of China Over Years

Yue Hu<sup>1</sup>, Yufei Sun<sup>1</sup>, and Wenquan Wu<sup>1</sup>

<sup>1</sup> Department of Political Science, Tsinghua University

DOI: [10.21105/joss.0XXXX](https://doi.org/10.21105/joss.0XXXX)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Editor: [Editor Name](#) ↗

Submitted: 01 January XXXX

Published: 01 January XXXX

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Summary

The Chinese government gives unique geocodes for each county, city (prefecture), and provincial-level administrative unit. The so-called “administrative division codes” were consistently adjusted to matched national and regional plans of development. Geocode adjustments disturb researchers when they merge data with different versions of geocodes or region names. Especially, when researchers render statistical data on a Chinese map, different geocodes between map data and statistical data may cause mess-up data output or visualization.

The package is developed to conquer such difficulties to match regional data across years more conveniently and correctly. Inspired by Vincent Arel-Bundock’s well-known `countrycode` (Arel-Bundock et al., 2018), we created `regioncode` to achieve similar functions specifically for China studies. `regioncode` enables seamlessly converting formal names, common-used names, and division codes of Chinese prefecture regions between each other and across thirty-four years from 1986 to 2019.

In the current version, we provide three basic functions `2code`, `2name`, and `2sname` to convert formal names, common-used names, and division codes between each other. We also provide some useful features: `incompleteName`, which completes incomplete parameters; `2area`, which converts region codes and names of the region into the municipal area that they belong to; `topinyin`, which convert the names or areas into the form of pinyin; and `language_trans`, which offers a function to convert the name of prefecture from any year to language(local dialect) zone.

## Examples

```
library(regioncode)

# Division codes across years

## original geocodes. It's 2019 version
corruption$prefecture_id

## after conversion. It's 1999 version
regioncode(data_input = corruption$prefecture_id,
            year_from = 2019,
            year_to = 1999)

# Division codes to region name
```

```
regioncode(data_input = corruption$prefecture_id,  
            year_from = 2019,  
            year_to = 1999,  
            method = "2name")  
  
# Advanced Usages  
  
## Incomplete input  
  
### Full, official names  
corruption$prefecture  
  
### Incomplete names  
corruption$prefecture_sname  
  
### Converting  
regioncode(data_input = corruption$prefecture_sname,  
            year_from = 2019,  
            year_to = 1999,  
            method = "2code",  
            incompleteName = "from")  
  
## Language zone translation  
  
regioncode(data_input = corruption$prefecture_name,  
            year_from = 2019,  
            year_to = 1999,  
            language_zone = TRUE,  
            language_trans = 'dia_group')
```

## Acknowledgements

We acknowledge contributions from Yuyang Shi, Yujia Xu, and Yuxin Pan, Haiting Tian, Weihang Shao, and Yuanqian Chen.

## References

Arel-Bundock, V., Enevoldsen, N., & Yetman, C. (2018). Countrycode: An r package to convert country names and country codes. *Journal of Open Source Software*, 3(28), 848. <https://doi.org/10.21105/joss.00848>