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

Yue Hu[1], **Yufei Sun**[1], **and Wenquan Wu**[1]

1 Department of Political Science, Tsinghua University

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 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 some useful features, like `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 offers a parameter "topinyin" to convert the names or areas into the form of pinyin, `language_trans`, which offers a function to convert 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 = "code2name")

# 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 = "name2code",
                    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

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