

Geographic Data Science

The need to group data

Dani Arribas-Bel

The need to group data

Everything should be made as simple as possible, but not simpler

Albert Einstein

The need to group data

- The world is complex and multidimensional
- Univariate analysis focuses on only one dimension
- Sometimes, world issues are best understood as multivariate. E.g.
 - Percentage of foreign-born Vs. *What is a neighborhood?*
 - Years of schooling Vs. *Human development*
 - Monthly income Vs. *Deprivation*

Grouping as simplifying

- Define a given number of categories based on **many characteristics** (multi-dimensional)
- Find the **category** where each observation *fits best*
- **Reduce complexity**, keep all the **relevant information**
- Produce **easier-to-understand outputs**

Types of grouping

- Non-spatial clustering
- Regionalisation



A course on Geographic Data Science by Dani Arribas-Bel is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.