Geofaceting Argentina

Juan Pablo Ruiz Nicolini

Universidad Torcuato Di Tella, Buenos Aires, Argentina
`juan.ruiznicolini@mail.utdt.edu`

Abstract. Application of the `geofacet` package to plot data from Argentina in grids that simulates geography.

Keywords: *ggplot2* · data visualization · georeferencing

Introduction

Wickham (2009) designed the application of Wilkinson's classic "The Grammar of Graphics" for R language (`ggplot2`). Its advantages are many and known. One of them is the ability to display faceted grids based on some variables. The `geofacet` package (Hafen, 2018a) proposes an extension of ggplot2 using a geographic id variable. Specifically, it adds the possibility of fitting `ggplot2` native grids as if they were maps. This introductory note summarizes some advantages of this approach and illustrates the technique with data from an Argentine province (Tucumán) as if it was georeferenced.

Why *geofacet?*

The `geofacet` proposal is better understood when compared to some other geographical data visualization techniques, such as tile-grams, choropleth or cartograms. In brief, Hafen (2018b) points out that `geofacet` is useful to surpass some limits of other techniques. Faceting is, in general, a powerful visualization technique in which we can (i) plot multiple variables or values per geographic entity; (ii) have more effective visual encoding schemes than just color, and (iii) each entity is given the same amount of screen real estate[1].

---

[1] Although this may not be desirable for all cases, this feature may be useful when districts that are geographically very small, but that are relevant in terms of the variables that are illustrated, may be underrepresented.

2   J.P.Ruiz Nicolini

Example

This example can illustrate several of the distinctive features of `geofacet`[2]. First, the grid bears a strong resemblance to the geographical map of the province. Secondly, we can graph more than one variable in each entity (department): `year` of observations and `vote share` for both parties. In this sense, we can appreciate a color key (to distinguish between political parties) and a `geom_line` that clearly shows the movements of the variable of interest. Finally, we address the size of the entities by focusing on the case of *San Miguel de Tucuman*, the capital district. If we had plotted it proportional to its geographical size, this entity would have low visual relevance despite accounting for $30 - 50\%$ of voters during this period.

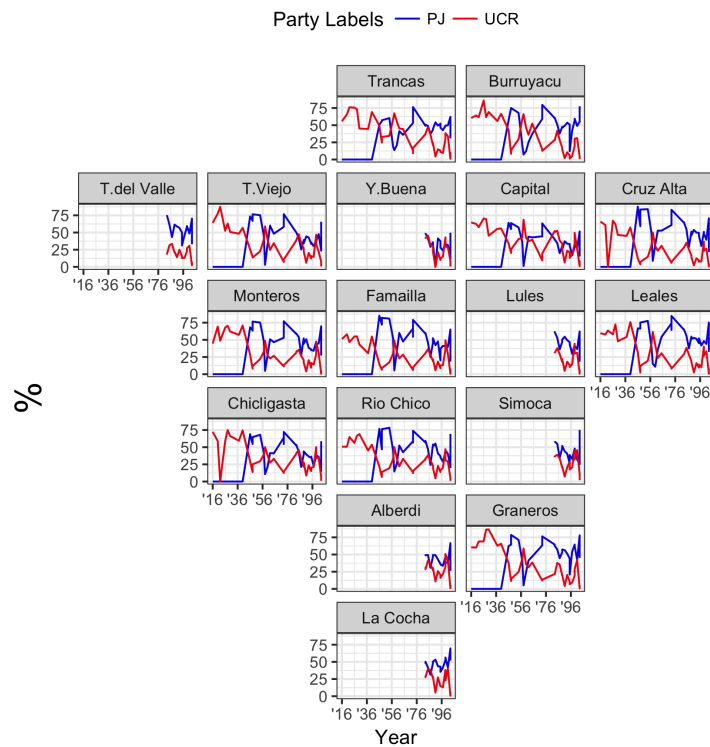**Main parties vote share in Tucumán elections (1916-2003)**

Fig. 1. Partido Justicialista and Unión Cívica Radical vote share in local political entities.

---

[2] The code and data needed to replicate this example are available at https://github.com/TuQmano/geofacet_ARG/tree/master/.LatinR

Bibliography

Ryan Hafen. geofacet. Available at https://github.com/hafen/geofacet (2018/04/2), 2018a.

Ryan Hafen. geofacet. Available at http://ryanhafen.com/blog/geofacet (2018/04/5), 2018b.

Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009. ISBN 978-0-387-98140-6. URL http://ggplot2.org.