1. What is the purpose and significance of geographic data analysis?
2. What variables will be the most important to analyze in the analysis of Lisbon's parish data?
3. When the geographic data obtained is too detailed in terms of geographic units, how can it be integrated to a higher level of geographic units? For example, how to integrate the street level data from BGRI2011\_1106\_all data to the parish level?
4. Do we need to verify the accuracy and authenticity of the data when we get some data from the web? For example, AirBnB's data does not use the exact coordinates of the user for information security reasons, how to deal with this?
5. According to my understanding of the article Clustering & Regionalization, this article is analyzed and studied on existing geographic data as well as some variables at the same time. How does this lead to a comparison and exploration between existing data and historical data?
6. The article concludes by mentioning that the improved geographical coherence comes at a pretty hefty cost in terms of feature goodness of fit. Should we abandon this step? Or is this step simply used as a reference point to demonstrate the superiority of a purely mathematical model?
7. Is there any way to collect some data accurate to the parish level? Because I tried to search for some data, I also used the website Pordata, but most of the data is only accurate to the city level.
8. Is our current goal to fully explore the data from the parish of Lisbon and then find some interesting information before using it as our final thesis?
9. Are there any databases that are only available to academics and not to the public? Perhaps some useful data could be found.