P3: Clearing OpenStreetMap Data

by Lieby Cardoso

- Region selected for the project:

I selected Belo Horizonte, MG, Brazil because it is the city where I live and I believe it is interesting to know some local points, also the domain of previous information about the region can assist in the analysis of the information registered. Region: http://www.openstreetmap.org/relation/368782 (http://www.openstreetmap.org/relation/368782)

Content

1. Methods and tools
2. Problems found on the map
3. Data overview
4. Other Ideas About the Data
5. Conclusion
6. References

1. Methods and tools

Descriptive of the methods and tools used in this work
https://github.com/liebycardoso/OpenStreetMap/blob/master/Methods.pdf
(https://github.com/liebycardoso/OpenStreetMap/blob/master/Methods.pdf)

2. Problems found on the map

Location identification of street type

Unlike the exercises performed, where the map was from the city of Chicago and the street type could be identified by theLast name, on the map of Belo Horizonte, the name of the street consists of the type, abbreviated or not, followed by the name. To meet this particularity of the Brazilian nomenclature, the variable street_type_re has been reformulated to work with the first word of the sentence. The change consisted in including the ^ character, which always corresponds to The start of the string and deletion of the $ character, which corresponds to the end of the string.

Code snippet:

```
In [ ]:   street_type_re = re.compile(r'^\b(?u)\w\S+\.?', re.IGNORECASE)
```