

Capstone project - Presentation

Correlation between the Maryland Counties rental value and its surrounding venues

1. Introduction

The final idea of the Project is to know the social impact in Maryland, This is a first stage of the complete work that personally I will finish. In this part it will be shown the correlations between the rental value of the Maryland State's Cities and the venues registered in the Foursquare Data Base.

The principal points to analyze are:

- Are the rental values related with the surrounded venues?
- It there specific reasons for the value difference?
- In case that the relation exist, could we in a next step determine the social status correlation?

2. Data

The principal dataset was downloaded in csv files from <https://public.opendatasoft.com>

The GEOJson by the Maryland Government public REST API.

Pitifully, the data wasn't clean and has information from several years of the rental value and for different type, This issue made the analysis non accurate.

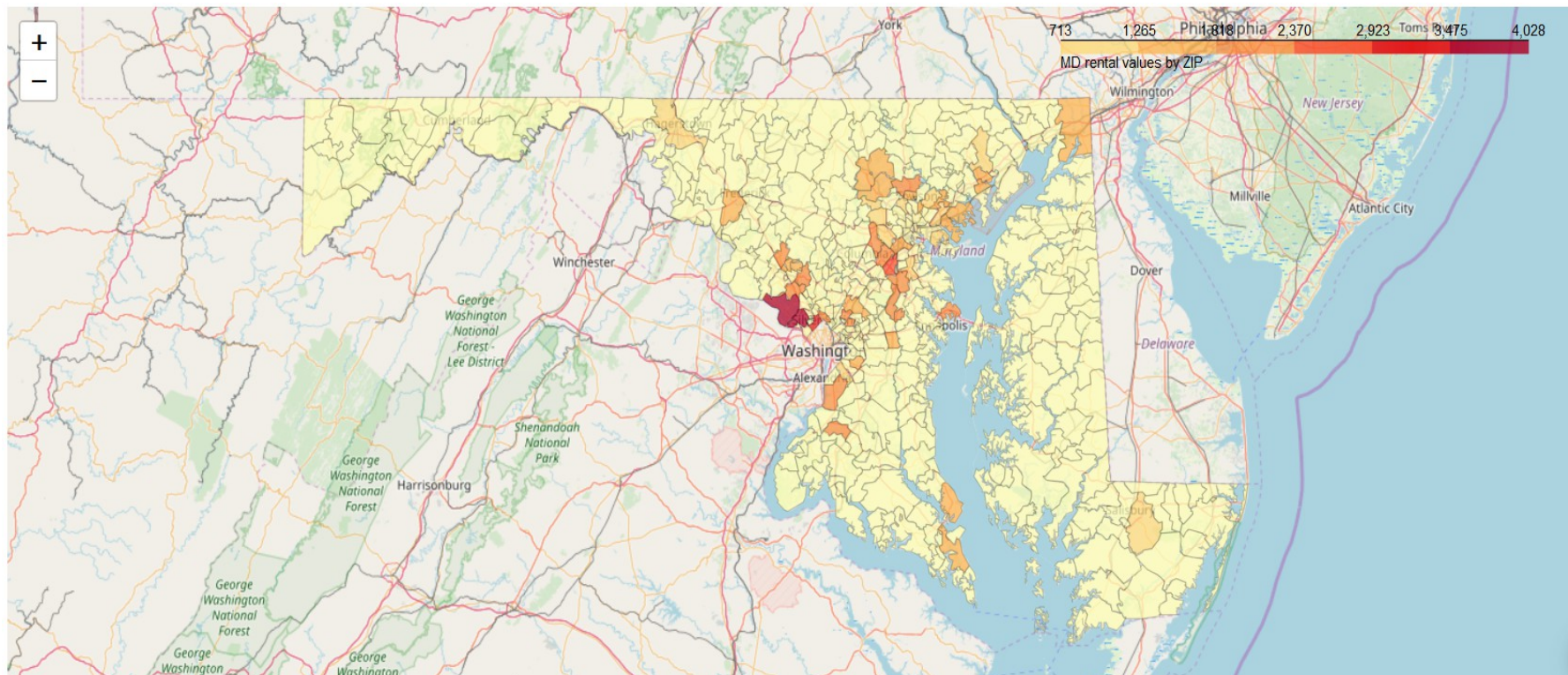
The csv files have the following structure:

rental_values.csv	Us-zip-code-latitude-and-longitude.csv
Zip	Zip
City	City
State	State
Metro	Latitude
County	Longitude
House type	Timezone
Date	Daylight
RentalValue	savings
Price	time
	flag
	geopoint

3. Estudy

Using Watson Studio and Jupyter Notebooks the data provides us the following information.

- Heating Map
Due the bad quality data doesn't show a proper definition of all the cities.



3. Estudy

- Linear Regression:

The coefficients and Scores are really small to get a good evaluation of the correlation, also the error score is too high.

R2 score: 0.15694132999421595

MSE: 337643.99620514584

4. Conclusions

- The data is incomplete (small sample size, missing deciding factors)
- The machine learning techniques are chosen or applied poorly.
- A few insights are relevant to consider after reviewing other data set related:
 - Population information provide by <https://public.opendatasoft.com>
 - News
 - Venues from Foursquare

INSIGHTS:

Cities with more minority social groups (Unemployment, poverty, immigration, etc.) are more likely to have a low rental value

THANK YOU

GRACIAS