CORRELATION BETWEEN OBESITY AND RESTAURANT DENSITY IN ENGLAND

Fabio Vaianella

February 2019

PREDICTING OBESITY RISKS IS USEFUL FOR THE DEPARTMENT OF HEALTH AND SOCIAL CARE OF UNITED KINGDOM

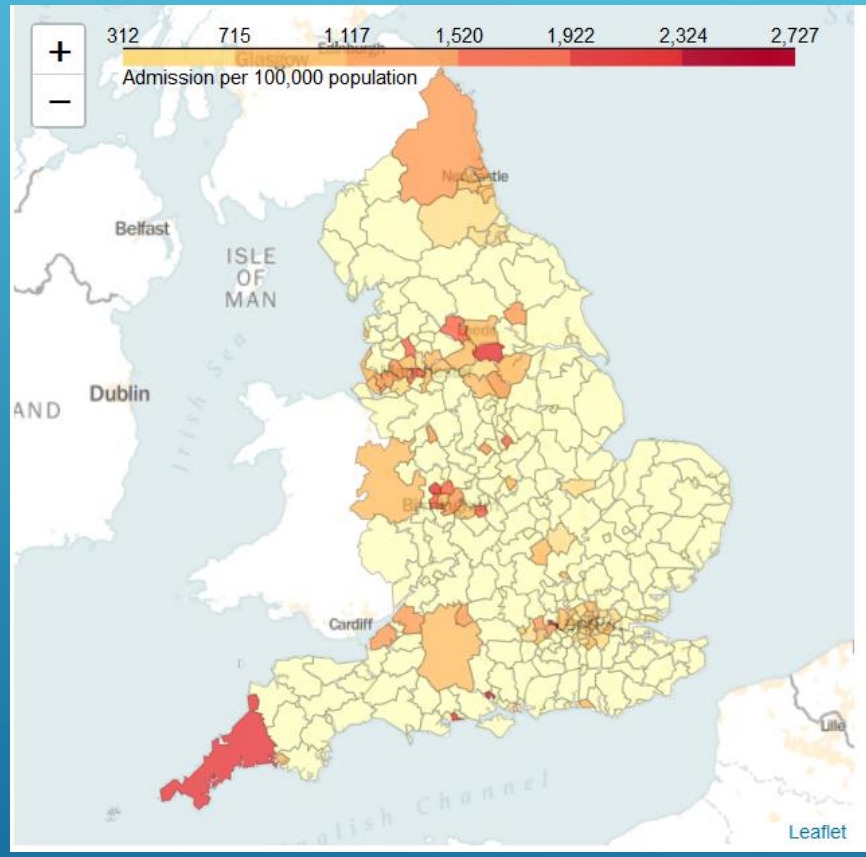
Is there any link between obesity and the density of restaurants in the surrounding of where a person lives in England?"

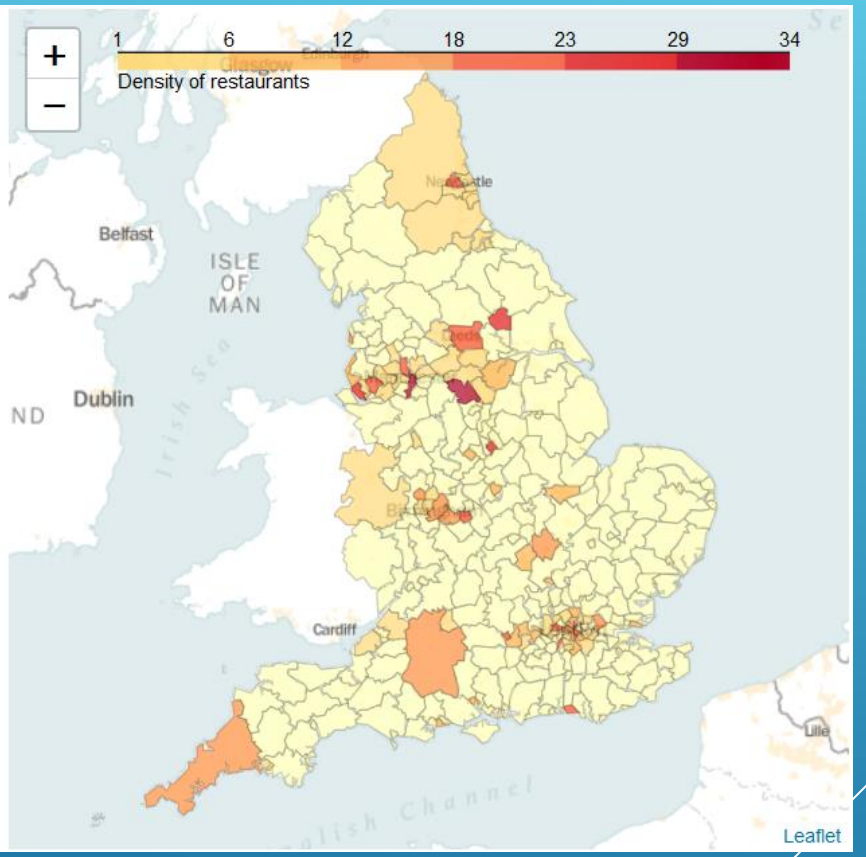
DATA ACQUISITION AND CLEANING

- Obesity admission on 152 English district for male and female on Kaggle. 12 columns before cleaning, 4 after cleaning
- Geographical location with Geopy using district name (+2 columns)
- Restaurant densities from Foursquare using district name (+1 column)

	District	All	Male	Female	Latitude	Longitude	Restaurants
0	County Durham	647.0	502.0	787.0	53.872616	-1.705963	2
1	Darlington	623.0	541.0	698.0	51.536151	-0.134964	4
2	Gateshead	716.0	427.0	1003.0	54.958554	-1.605700	2
3	Middlesbrough	657.0	496.0	812.0	54.576042	-1.234405	4
4	Newcastle upon Tyne	737.0	451.0	1038.0	54.973847	-1.613157	18

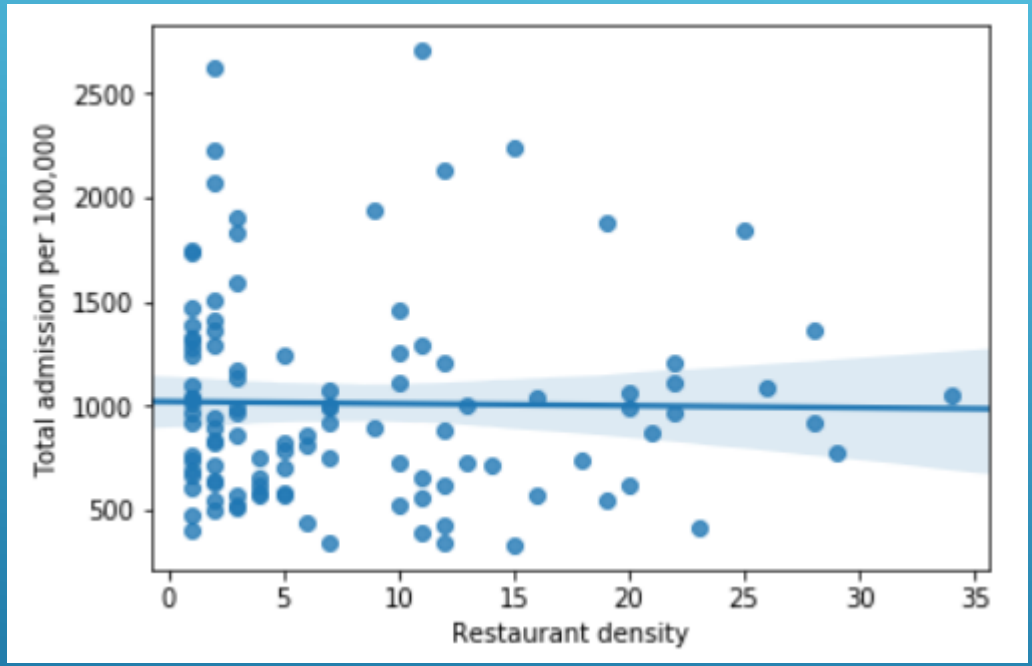
VISUALIZATION- CHOROPLETH MAPS

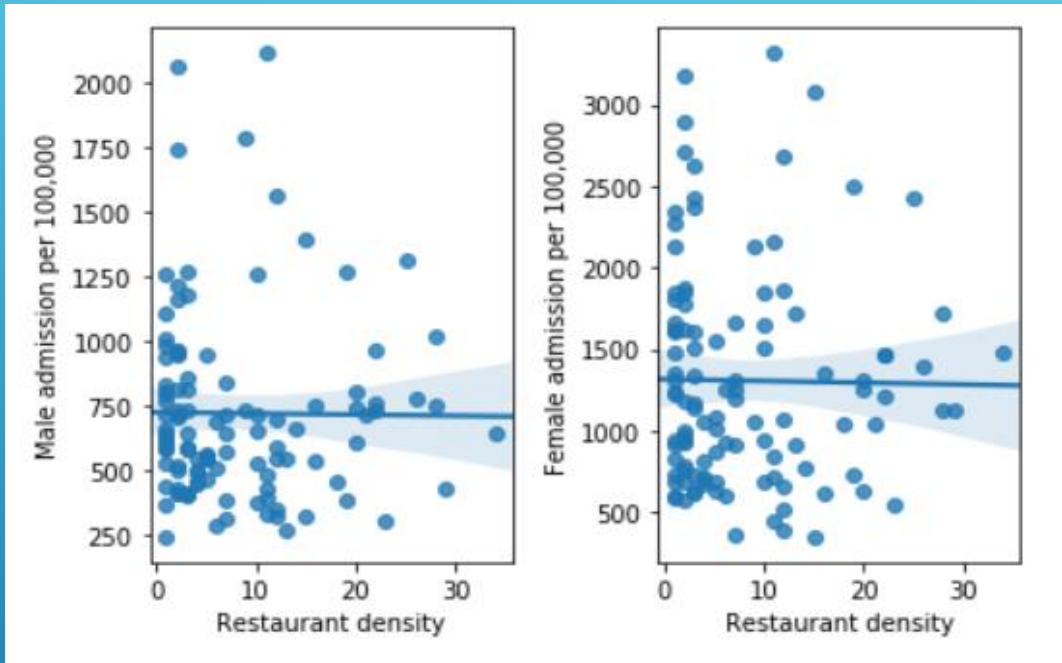




Similar in appearance, but maxima of admission do not correspond at all with maxima of restaurants

REGRESSION PLOTS





Data are spread out around the linear regression prevision, probably not correlation between restaurant densities and obesity. No trending can be observed

ACCURACY METRICS OF THE LINEAR MODEL

Accuracy on the training set (80% of data)

```
Mean absolute error: 381.55
Residual sum of squares (MSE): 262759.26
R2-score: 0.00
```

Accuracy of the prediction on the test set (20% of data)

```
Mean absolute error: 332.31
Residual sum of squares (MSE): 160457.10
R2-score: -0.03
```

ACCURACY METRICS OF THE LINEAR MODEL

Accuracy on the training set (80% of data)

```
Mean absolute error: 381.55
Residual sum of squares (MSE): 262759.26
R2-score: 0.00
```

Accuracy of the prediction on the test set (20% of data)

```
Mean absolute error: 332.31
Residual sum of squares (MSE): 160457.10
R2-score: -0.03
```

Linear regression very bad model
→ Probability that there is no link between obesity and density of restaurants very high

DISCUSSION AND CONCLUSION

- no link between obesity and the number of restaurants in a district of England
- Take into account only restaurants that are more fat providers (such as fast-food, steakhouse,...) could improve the model
- Take into account that the number of gyms or parks in a district can reduce the obesity rate
- Limit on the number of venues we can explore (100) can be a problem too