Are Culturally Diverse Boroughs More Spatially Connected?

Ellen Zeng, Binxia Xu

Abstract

The aim of this research is to investigate whether there is a relationship between cultural diversity and physical connectivity in London boroughs. Three sources of datasets were used - Ethnic group indices from the 2011 census[1], the CRDC 2017 Ordnance Survey Greater London Geodata pack[2], and data on railway stations or bus stops in London from Google maps. We used various visualisation methods to conduct an exploratory analysis. We also run several regression models for deeper statistical analysis.

Introduction

There are plentiful studies on the relationship between ethnic diversity and socioeconomic indicators such as standard of living, affluence, and social cohesion. However, there is less research in understanding how level of diversity in an area can vary with physical connectivity. We measure borough cultural diversity by the extent of diversity in the aspects of ethnicity, religion, country of birth, languages spoken and passports of the borough's inhabitants. We quantify the physical connectivity of boroughs by the density of roads, bus stops and railway stations within the borough. We are also interested in the application of statistical methods to census data and new ways of combining census data with other data sources. In this project, besides using census data, we pull data from CRDC's Geodata packs and Google Maps API.

Method

In this section, we elaborate on the data preprocessing and experimental procedures.

Measuring Ethnic Diversity

From the ethnic group indices from the 2011 census, we use a standardised concentration index[3] to measure diversity in ethnicity, religion, country of birth, languages spoken and passports of the inhabitants of each borough. The concentration index is as follows:

$$HI = 1 - \sum_{i=1}^{n} s_i^2$$

where s_i is the share of a particular category, out of a total of n other categories. For example, in the case of ethnicity, s_i refers to a particular ethnic group, out of a total of n other ethnic groups. Higher scores of HI denote more heterogeneous populations.

Measuring Physical Connectivity

From the CRDC 2017 Ordnance Survey Greater London Geodata map, we extract the number of roads in each borough from the shapefiles. From this data, we calculate the density of roads in each borough. We also extract data on bus stops and railway stations from the Google Maps API. By locating the borough in which each bus stop and railway

[1] Census Information Scheme. 2011 Census Diversity. *London Data Store*, 2014. Web. https://data.london.gov.uk/dataset/2011-census-diversity

[2] Consumer Data Research Centre. 2017 Ordnance Survey Open Map Greater London Area Geodata Pack: London. Web. https://data.cdrc.ac.uk/dataset/cdrc-2017-ordnance-survey-open-map-greater-london-area-geodata-pack-london-e12000007

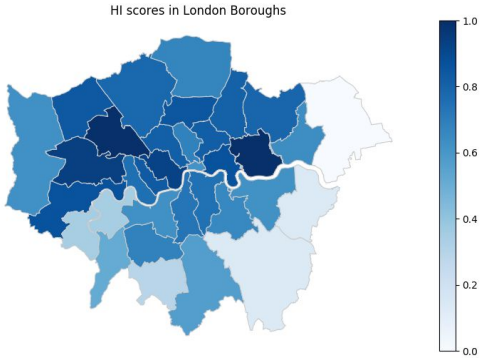
[3] Reibel, Michael, and Moira Regelson. "Quantifying neighborhood racial and ethnic transition clusters in multiethnic cities." *Urban Geography* 28.4 (2007): 361-376.

station is situated, we are able to calculate the density (unit: per hectare) of bus stops and railway stations of each borough.

As we are comparing 2011 diversity data with physical infrastructure data from 2017 and 2019, a key assumption here is that the infrastructure data is comparable to that in 2011. In analysing the statistical relationship between diversity and connectivity, we use polynomial regressions, as elaborated further in the next section.

Results & Discussion

Exploratory Analysis



We first visualise the spatial distribution of each variable using choropleths. As seen from Figures 1b, c and d, transport connectivity diminishes moving away from Central London. Spatial patterns for the HI score, as seen in Figure 1a is less conclusive.

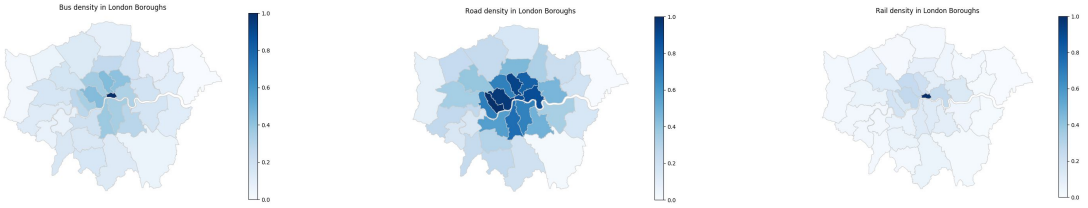
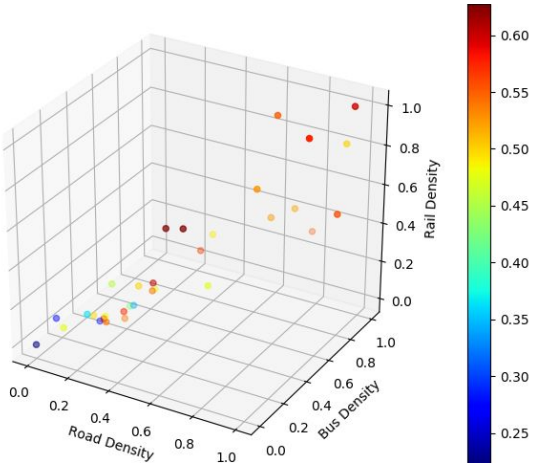


Fig 1: Choropleths showing distribution of a. (top) HI score b. (bottom left) Bus stop density c. (bottom centre) Road density d. (bottom right) Railway station density



We further visualise the relationship between the four variables, after normalisation, using a 3D scatter plot, as shown in Figure 2. HI scores of each borough are denoted by the colour chart. We observe that there is a positive correlation between the densities of roads, bus stop and railway stations. Boroughs which rank low in diversity are also more certain to have lower connectivity densities.

Fig 2: 3D scatter plot visualisation

Polynomial Regression

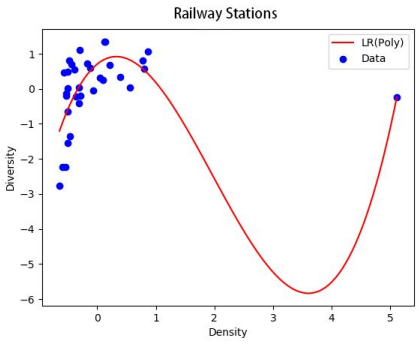


Fig 3: Visualisation of regression model with outlier

As can be seen from Figure 3 on the left, there is one point distributed quite differently from other boroughs (City of London). To fit a better model for the whole dataset, this point is removed for subsequent analysis.

After removing the outlier, the relationship between spatial connectivity and the diversity level in a borough can be somewhat distinguished in polynomial regression. In general, the diversity level would increase rapidly with the density of each feature when the density is not high (Figure 4). With the density increasing to higher level, the growth rate of diversity becomes slower and even presents a sense of decrease.

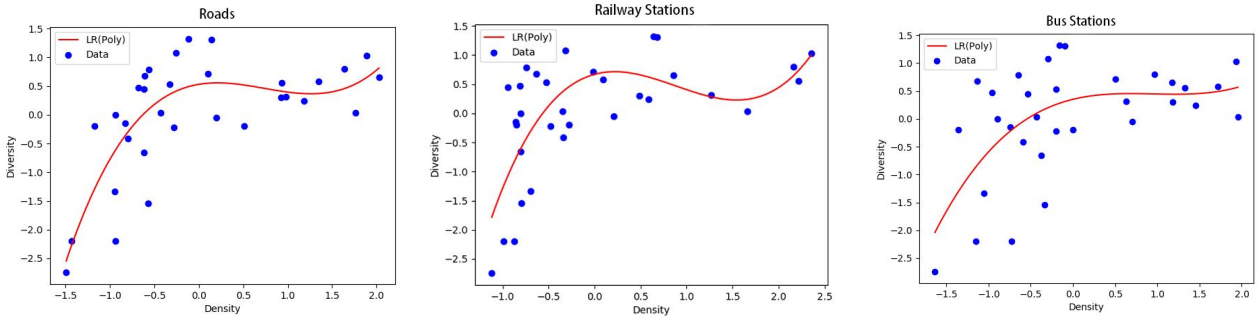


Fig 4: Visualisation of regression models for the correlation between three connectivity features (the data has been normalised, unit: per hectare) and the diversity level of a borough: a. (left) the density of roads b. (centre) the density of railway stations c. (right) the the density of bus stations

Here are the results of the 3 degree polynomial regression models we fitted for the three connectivity features: density of roads, bus stations and railway stations.

Roads

$$Y = 0.531 + 0.259 * x - 0.721 * x^2 + 0.327 * x^3$$

Bus Stations

$$Y = 0.351 + 0.371 * x + 0.424 * x^2 + 0.149 * x^3$$

Railway Stations

$$Y = 0.670 + 0.428 * x - 1.112 * x^2 + 0.420 * x^3$$

In conclusion, cultural diversity appears to exhibit some extent of positive trend with connectivity indices, across all density measures, initially. This case study has shown that there is a possible relationship between diversity and connectivity measures. A direction of future study could be to examine the rate of change in diversity in relation to change in physical connectivity. Other measures of connectivity can also be considered - these include distance measures such as metric distances or travelling time, as well as intersection of roads.