There is any relation between unemployment rates and socio-economic indicators at LSOA level in the London area?

Bogdan Rugina

12 December 2018

## Introduction

#### This study intends to understand the extent of socio-economic indicators as primary education, higher education, income and their relationships to unemployment rate in the London area as of 2011. After pre-analysing the data for possible skewed distributions and correlation between variables, a multiple linear regression is performed in order to test the statistical significance of the findings.The final output of this reasearch will consist on local regression predicting to which extent the above mentioned social indicators influence the unemployment rate in London LSO areas in 2011. The inspiration for this research comes from an international project called “INEQ-CITIES” which main aim was to provide evidences of health inequalties in cities across Europe (<https://www.ucl.ac.uk/ineqcities>). Different pubblications followed the project, interesting in this study, Borrell et al (2014) enphasise “a consistent pattern of socio-economic inqualities in mortality” in Europe emerging from socio-economic indicators such as unemployment and education levels. Similarly in this study the main objective is to identify to which extent education and median income of each Loandon small areas contribute to unmeployment for the year 2011.

#### The first part relates to inferential statistic analysis, in particular a bi-variate linear regression will be perfomed and based on first results other indicators will be inserted one by one into a multiple regression model.

#### The second part will consist on a geographical weighted regression (GWR) perfomed in order to visualize local aspects of the multiple regression,some predictions will be inspected as part of the overall objective.

## Data

#### The material used to develop this research comes interely from the UK Office for National Statistics (OSN <https://www.ons.gov.uk/search>?:uri=search&:uri=search&q=LSOA&page=3) and in particularly refers to the London Lower Super Output Areas (LSOA) which are smallest geographically statistic areas containing data from the offical census of 2011(<https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeography#super-output-area-soa>). The London Datastore website releases as well the ONS based “LSOA Atlas”" which “provides summary of demographic” and other retaled data from population segmentation to housing, land use, employment, etc. In this study the London LSOA Atlas will be the direct source of data for the following analysis.

#### The first set of variables to be extracted from the databased are the “Unemployment Rates” that accordingly to the OSN are calculated as the the total number of unemployed people divided by the active population(<https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/methodologies/aguidetolabourmarketstatistics>). Unemployment rates will represent in this study the dependent value based on which all the indicators will be related.

#### Primary and higher education indicators were calculated respectively as the working aged population between 16-65 that completed the first level of qualification or lower from one side and the working population who obtained level 4 qualifications and above from the other, both were then divided by the total population aged 16-25.

#### Finally the fourth variable resulted from the household median income divided by the working aged population, the median values were chose instead of mean ones, due to the fact that eventual outliers may raise the mean values in the social context as that of London.

#### The following table illustrates the extremes values that will be examined in depth in the next paragraphs.

library(readr)  
social\_indicators <-read\_csv("C:\\Users\\Bogdan\\Documents\\UCL\\MRes\\GIS\\week9\\Assignment3\\LSOA\_data.csv")

## Parsed with column specification:  
## cols(  
## Codes = col\_character(),  
## Persons\_per\_hectare\_2012 = col\_double(),  
## LOG10\_Density = col\_double(),  
## Unemployment\_Rate = col\_double(),  
## Primary\_Educ\_Ind = col\_double(),  
## Higher\_Educ\_Ind = col\_double(),  
## Median\_Housh\_Inc\_Ind = col\_double(),  
## Perce\_United\_Kingdom = col\_double(),  
## Perce\_Not\_United\_Kingdom = col\_double(),  
## Christian = col\_double(),  
## Buddhist = col\_double(),  
## Hindu = col\_double(),  
## Jewish = col\_double(),  
## Muslim = col\_double(),  
## Sikh = col\_double()  
## )

head(social\_indicators[,4:7])

## # A tibble: 6 x 4  
## Unemployment\_Rate Primary\_Educ\_Ind Higher\_Educ\_Ind Median\_Housh\_Inc\_Ind  
## <dbl> <dbl> <dbl> <dbl>  
## 1 9.9 2.02 0.0558 184.  
## 2 8 1.38 0.0766 152.  
## 3 4.9 0.687 0.357 199.  
## 4 8.8 1.61 0.406 140.  
## 5 2.1 0.424 1.05 221.  
## 6 3.6 0.333 0.546 271.