How does allocation of funding to London local authorities affect obesity cases during 2008 and 2018?

Introduction

Today, almost one in five Year 6 children in the UK is found to be obese, and sadly this number is not dropping during the past few years (NHS, 2018). The origins of childhood obesity stem from various aspects, including lifestyle issues, genetic and environmental causes. The government has been taking considerable forms of actions to tackle this problem from its root, from sugar reduction to advertising and promotions. This study investigates how government has been allocating their funding to local authorities in London, and the performance of local authorities using said funding.

Part 1:

What affects the amount of funding allocated to local authorities in London?

Data

The data used in this study contains population, obesity cases, total budget, and allocation of funding for local authorities across London in 2018. An illustration of data employed is shown in Table 1:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | local\_authority\_area | | --- | | cases\_total | pop\_total | Obese density | total\_budget | clean\_air\_p |  |  |  |
| 0 |  |  |  |  |  |  |  |  |  |
| 1 |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |
| 13 |  |  |  |  |  |  |  |  |  |
| 16 |  |  |  |  |  |  |  |  |  |

Table 1 illustration of data used, list of column names includes:

Names of local authority area; total obesity cases in each area; total population; obesity density (obesity cases divided by population); total budget allocated; percent of budget spent on improving air quality, cleaner environment, health training, raising school awareness, media awareness and subsiding counselling.

Only one observation is considered an outlier due to its relatively small scale, and therefore dropped from the dataset, City of London data. Its population is below the average population in London boroughs by 97%, having this data in the linear regression plot of obesity cases in 2018 vs total budget spent lowers the regression coefficient from 0.437 to 0.349.

Methodology

Three approaches were taken to investigate the criteria of funding:

1. A linear line was fitted using scipy.stats.linregress() between total budget spent and obesity density.
2. A more outlier-robust linear approach, called Random Sample Consensus (RANSAC) was used. This method compliments the ordinary least squares methods by adding detections of outliers and accord them to have no influence on the parameters of the model (Fischler and Bolles, 1981). In sklearn.linear\_model.RANSACRegressor(), outliers are classified as those whose residual exceed the median absolute deviation of dependent variables (Pedregosa, F. *et al.*, 2011).
3. Finally, a non-linear approach was taken. Polynomial regression of various degrees was fitted to the data using numpy.polyfit() in Python.

Results

1. Logically one might assume that the more obesity cases discovered per unit population, the more funding would be allocated. The linear line fitted to total budget spent vs obesity density yields a coefficient of determination of 0.002, which means that only 0.2% of the variance of dependent variable (total budget) is explained by independent variable (obesity density). Log transformations were also applied to the variables, associated parameters are shown in Table ????

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | X,y | Log x, y | Logy ,x | Logx, logy |
| slope | 3e-9 | 0.001 | 3e-7 | 0.057 |
| constant | -0.002 | -0.005 | 0.47 | -0.259 |
| R2 | 0.037 | 0.032 | 0.29 | 0.29 |
| Pearson correlation coefficient |  |  |  |  |
| P value | 0.29 | 0.32 | 0.001 | 0.002 |
| Relationship implied | linear | exponential | exponential | Power |

Table 2: coefficients related to the relationship between total budget spent and obesity density

As the maximum value of R2 being under 0.3 in Table ???? , this indicates that the amount of funding allocated to local authorities is extremely weakly correlated with the obesity cases per unit population by the power law, thus indicating that this relationship is meaningless.

1. The lack of a meaningful linear relationship could be induced by number of outliers involved. The outliers in this case originate from changes in human behaviour, the government may identify some boroughs need more attention due to other reasons that cannot be demonstrated from statistical measures.

1. Selected degrees of 2 to 50 of polynomial regression and associated r squared values are displayed in the following table.

table

Polynomial graphs

It can be spotted that as the degree of polynomial fitted increases, r squared increases. One can deliberately dial the degree up and manipulate r squared to the range accepted, but this would not be a valid approach logically. If so, the model yielded would be highly influenced by the noise in the dataset, rather than demonstrating the data’s underlying real trend. Therefore, there is no appropriate polynomial relationship that can be fitted.

Part 2:

The relationship between areas of funding spent and obesity cases.

Conclusion

Hypo 1:

Logically, the obesity problem should be alleviated with more budget spent on tackling it. The drawback of this hypothesis is the lack of a reference level, which arises from an existing annual increase of obesity density along with population expansion. The reference level of obesity density increase could be calculated from different budget spent in previous years and associated obesity densities. Consequently, the

Another approach to examine the criteria of allocating funding is to consider the relationship of total budget with population and number of obesity cases respectively.

In this approach, the decision of outliers are entirely in the

Bib

Fischler, M. A. and Bolles, R. C. (1981). ‘Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography’. *Communications of the ACM*, 24 (6), pp. 381–395. doi: 10.1145/358669.358692.

NHS. (2018). *National Child Measurement Programme, England 2018/19 School Year.* Available at: https://digital.nhs.uk/data-and-information/publications/statistical/national-child-measurement-programme/2018-19-school-year.

Pedregosa, F., Varoquaux, G., Gramfort, A. and Michel, V. (2011). ‘Scikit-learn: Machine Learning in Python’. *Journal of Machine Learning Research*, (12), pp. 2825–2830.