In the past decade, how have budgets allocations and different local authority types intervened in the pattern of childhood obesity in the UK?

Introduction

Childhood obesity harms children's present and future health. It raises the risk of various ailments in children and has a detrimental impact on their quality of life and self-esteem (Croker et al. 2012). The data source from the Department of Health provides the cases number of childhood obesity, total population, and allocation of the total budget in six categories, combined with areas, region, and local authority type.

This paper will explore the effectiveness of the government intervention and the impacts of local authority types on patterns of childhood obesity in the UK by developing a multiple linear model. My analysis would separate into four parts: Data Processing & Data Selection, Method, Results, Conclusion.

Data Processing

The dataset needed to be manipulated to focus on answering the research question, meaning we would remove gender, local authority area, and region data. The figure below shows the frequency of local authority types, indicating the 'other_London' is the only one row in the dataset. Thus, we would ignore the data of 'other_London' in the analysis (Boyd et al.2012).

Fig1. Histogram of local authority

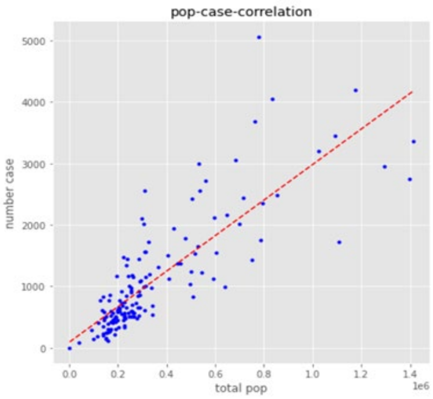

Fig2. A linear regression between total population and the number of cases

The number of obesity cases could not be directly used as a response variable since the number of obesity cases is easily influenced by the total population. The following figure indicates that as the entire population increases, so will the number of cases. In order to solve the impact of the total population on the number of obesity cases, we use the difference value between the ratio of cases to the population from 2008 to 2018(). We will use CRCP(1) as the dependent variable in the multiple linear regression model.

$$CRCP = \frac{2018\ cases\ number}{2018\ total\ poplation} - \frac{2008\ cases\ number}{2008\ total\ poplation} \quad (1)$$

Regarding the independent variable, six variables about the budget will become independent variables, respectively 'clean air', 'clean environment', 'health training,' 'school awareness," media awareness," sub counseling'. These six variables would be natural logarithm so that the scale of each variable would be comparable with the left-hand variable. In the research, the different local authority types would also be included in the independent variables to explore the governmental structure differences. Therefore, the local authority type is the categorical variable and should be converted into numerical data before fitting into the model. The definitions of the variables are shown in the table below (Table.1).

Before establishing a multiple linear regression model, outliers in the independent variables need to be removed. We could use the method of Tukey's fences to detect the outliers (Fig3). Three rows are removed because of an exception in the value after the logarithm.

Variable	*Definition*
crcp	the difference value between the ratio of cases to population in 2008 and 2018
clean_air_log	log(clean air)
clean_environ_log	log(clean_environ)
health_training_log	log(health_training)
school_awareness_log	log(school_awareness)
media_awareness_log	log(media_awareness)
sub_counselling_log	log(sub_counselling)
local_authority_type	categorical variable

Table.1 the definition of variable

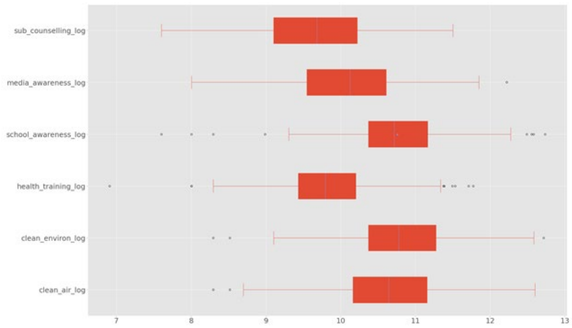

Fig3 box plot of the independent variable

Method

After identifying the variables and removing the outliers, the variables need to be tested for multicollinearity to ensure that the model for multiple linear regression is sufficiently simple. The variance inflation factor (VIF) of independent variables is calculated and drop the independent variable which the value of its VIF is bigger than 5. From the result, we found that the categorical variables exist the multicollinearity, so the categorical variables of 'local_authority_type_london_borough' should be dropped.

Then we could build a linear regression model to access the relationship between the budget, local authority type, and CRCP. The formula of the model is as follows.

$$\begin{aligned} CRCP = \alpha + \beta 1 * clean\_air\_log + \beta 2 * clean\_environ\_log + \beta 3 \\ * health\_training\_log + \beta 4 * school\_awareness\_log + \beta 5 \\ * media\_awareness\_log + \beta 6 * sub\_counselling\_log + \beta 7 \\ * Categorical\ (local\_authority\_type) + \varepsilon \end{aligned}$$

Result

From the table below, the multiple linear regression model's R-squared is 0.659, and the adjusted R-squared is 0.637, which indicates that the budgets allocations and different local authority types could explain the 60% variance of CRCR.

Table2 OLS Regression Results

OLS Regression Results			
Dep. Variable:	average-obesity-rates-change	**R-squared:**	0.659
Model:	OLS	**Adj. R-squared:**	0.637
Method:	Least Squares	**F-statistic:**	29.63
Date:	Sat, 13 Nov 2021	**Prob (F-statistic):**	3.60E-28
Time:	15:51:29	**Log-Likelihood:**	81.384
No. Observations:	148	**AIC:**	-142.8
Df Residuals:	138	**BIC:**	-112.8
Df Model:	9		
Covariance Type:	nonrobust		

Further analysis of the regression results reveals the statistical significance of different independent variables(Table.3). Firstly, we will explain the importance of the six independent variables of budget allocation in the multiple linear regression model. When all the variables are fixed, expected the 'clean_air_log,' an increase of the 'clean_air_log' by 1 unit, the predicted CRCP will decrease 2.47%. Similarly, when the 'clean_environ_log' and 'school_awareness_log' grow by one unit, the dependent variable could predict to reduce by 1.49% and 9.58%. On the contrary, 'health_training_log,' 'media_awareness_log,' and' sub_counselling_log,' increase by one unit, the other variables remain unchanged, and the obesity rate is predicted to increase by 1.65%, 9.88%, and 5.66%, respectively. Although all budgets are designed to mitigate the growth of CRCP, some parts of the budget do not correlate significantly with CRCP. Concerning the difference of local authority type, 'non_metropolitan_county' is more effective in controlling CRCP than other government types when budget allocations are at the same level.

Table.3 OLS Regression Results

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
const	0.4435	0.335	1.323	0.188	-0.219	1.106
local_authority_type_metropolitan_borough	-0.2095	0.036	-5.872	0	-0.28	-0.139
local_authority_type_non_metropolitan_county	-0.5361	0.053	-10.085	0	-0.641	-0.431
local_authority_type_unitary_authority	-0.4351	0.033	-13.174	0	-0.5	-0.37
clean_air_log	-0.0274	0.021	-1.301	0.195	-0.069	0.014
clean_environ_log	-0.0149	0.02	-0.746	0.457	-0.054	0.025
health_training_log	0.0165	0.021	0.778	0.438	-0.025	0.058
school_awareness_log	-0.0958	0.022	-4.412	0	-0.139	-0.053
media_awareness_log	0.0988	0.022	4.467	0	0.055	0.143
sub_counselling_log	0.0566	0.024	2.309	0.022	0.008	0.105

Omnibus:	10.616	**Durbin-Watson:**	1.776
Prob(Omnibus):	0.005	**Jarque-Bera (JB):**	20.972
Skew:	0.243	**Prob(JB):**	2.79E-05
Kurtosis:	4.779	**Cond. No.**	719

From the Residuals vs. Fits plot (Fig 4), it is clear to observe that the residuals "bounce randomly" around the residual = 0 line, implying that the linear assumption is reasonable. And the residuals roughly form a "horizontal band" around the residual = 0 line, which means that the variances of the error terms are equal. Moreover, none of the residuals' stand out from the underlying random pattern of residuals, indicating the absence of outliers. Thus, they all confirm the regression model's assumptions.

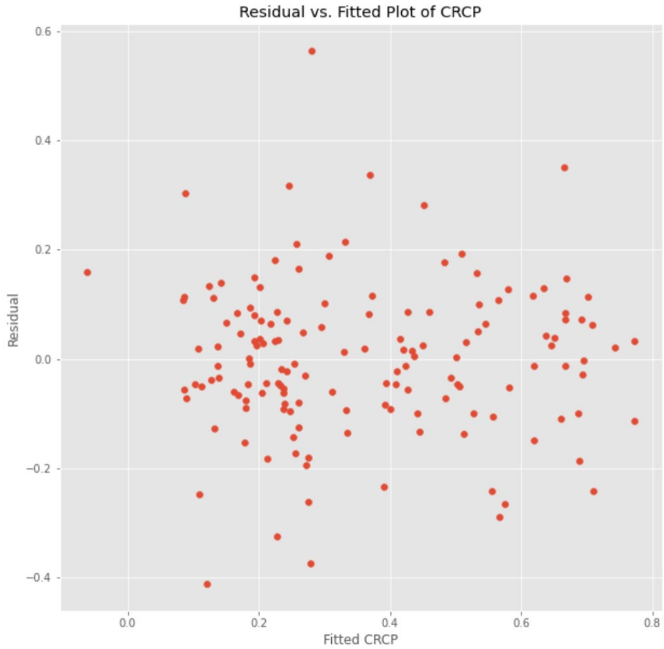

Fig4 Residual vs Fitted Plot of CRCP

Conclusion

The above analysis shows that not all the government's budgetary allocations are useful to reduce obesity rates. Budgeting for school awareness could go a long way towards reducing the rate of childhood obesity in the UK. A budget for clean air and a clean environment also has a positive effect on decreasing the rate of childhood obesity. Surprisingly, some budgets could be useless obesity rates, especially media awareness and counseling. Thus, the government's budget should be directed more towards school awareness and the environment. The government should further evaluate the effectiveness of using funds in other areas and reform the strategy for the benefit of funds accordingly. Meanwhile, it is also essential to consider the impact of local government types of differences. The limitation of the study is the lack of data on the child population, as child obesity rates should be based more convincingly on the number of children in the population than on the total population (Stamatakis et al.2010).

(Word count:998)

References

Boyd, T., Docken, G. and Ruggiero, J., 2016. Outliers in data envelopment analysis. Journal of Centrum Cathedra. doi: 10.1108/JCC-09-2016-0010

Croker, H., Viner, R.M., Nicholls, D., Haroun, D., Chadwick, P., Edwards, C., Wells, J.C. and Wardle, J., 2012. Family-based behavioural treatment of childhood obesity in a UK National Health Service setting: randomized controlled trial. International Journal of Obesity, 36(1), pp.16-26.doi: ijo2011182

Stamatakis, E., Wardle, J. and Cole, T.J., 2010. Childhood obesity and overweight prevalence trends in England: evidence for growing socioeconomic disparities. International journal of obesity, 34(1), pp.41-47.doi: ijo2009217

GitHub: Https://github.com/lijianqiao-UCL-scua/QM-coursewokr1.git