Topic
Did the government spend too much on environmental factors which has less effectiveness on childhood obesity throughout the past ten years?

Introduction
In the recent years, the UK government has been putting tremendous efforts in coping with childhood obesity. They spent approximately £5.1 billion on obesity supports in 2014/15 (GOV.UK, 2017). Budgets covered widely from environmental to psychological elements, such as greater air quality and raising children awareness. It raises a doubt that whether the government spent appropriately.

Research question
In this paper, it is going to investigate whether the government spend too much on environmental factors which has less effectiveness on childhood obesity throughout the past ten years.

Factors will be categorized by:
1) Environmental factors ("clean_air"; "clean_environ")
2) Educational factors, which is more psychological
("health_training"; "school_awareness"; "media_awareness"; "sub_counselling")

In order to reach the answer for this question, the main objectives are to compare the environmental budget and total budget distributions, as well as the relationship between cases and the environmental budget over the past ten years.

Presentation of Data
The data is obtained from the Department of Health at GOV.UK. The children data observed is aged 4-5 and 10-11. 95% confidence intervals should be considered when conducting a direct comparison. The data could have a potential error in the collection, collation and interpretation of the data. Date was released from the year 2014.

Table 1 demonstrates the incremental columns added on the original dataset. In which, the average spending for both two factors are the environmental/ educational spending in 2018 minus 2008 divided by ten years. The "average increased cases" is the cases in 2018 minus 2008 divided by ten years, measuring how many cases inflated throughout the past.

	environ_budget	educat_budget	average_environ_budget	average_educat_budget	average_increased_cases
0	43000	96000	4300	9600	17.0
1	67000	153000	6700	15300	23.4
2	86000	74000	8600	7400	16.7
3	75000	86000	7500	8600	3.8
4	89000	41000	8900	4100	5.6
...	...	...	...	...	...
147	116000	93000	11600	9300	42.7
148	61000	48000	6100	4800	5.0
149	45000	156000	4500	15600	28.6
150	240000	230000	24000	23000	51.9
151	44000	96000	4400	9600	24.4

(Table 2)

Methodology

The processing and visualisation of data are produced by the *Python* programming language. All the analysis is reproducible and publicly shared on the *Github* hyperlink provided at the end of this paper.

In order to discover the research topics and reach the objectives, the analysis is going to be separated into three main parts with different statistical approaches.

Firstly, it is to produce a histogram for observing how the total environmental budget and total budget are distributed. It is to provide an insight of the underlying frequency distribution of these subsets, indicating how the total budget interrelates with only two environmental elements. The result would help to determine whether the environmental spending is part of the "huge economic costs". Secondly, by using Spearman's rank correlation, to test the relationship between average cases and the average environmental spending over the past ten years. Furthermore, the third method is going to use single linear regression to model the relationship between these two variables by fitting a linear line. The total environmental spending is selected as independence variable and total cases in 2018 as dependent variable. It is because it aims to discover how much the accumulated environmental spending affected the cases recorded in the latest year from the data.

Results and Discussion

For the first objective, *Figure 1a* and *1b* demonstrate the distribution of the total budget and total environmental budget. The dashed line indicates the mean for the specific subset. However, due to the distributions for two data are not normal, the mean does not provide a useful information. Besides, we can see that there are some outliers on the right side in *Figure 1b*, which have a significant dispersion on the right. Thus, the outliers are being removed and to compare with the total budget (*Figure 1c*).

After removing the outliers, the environmental spending is positively skewed meanwhile the total budget is approximately skewed with a long tail end to the right. Both categories record similar highest frequency in the range below £ 200,000, with 2.17e-05 in Chi-Squared test. It is surprise that the frequency of only two elements from the total budget would distribute closely with the total. Also, since the Chi-Squared is less than 0.05, we can conclude that that the level of total budget is linked to the proportion of the environmental spending.
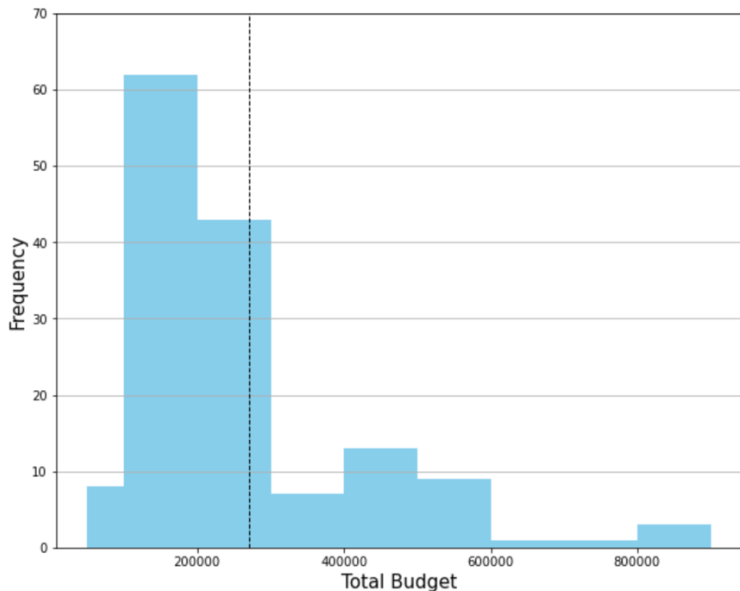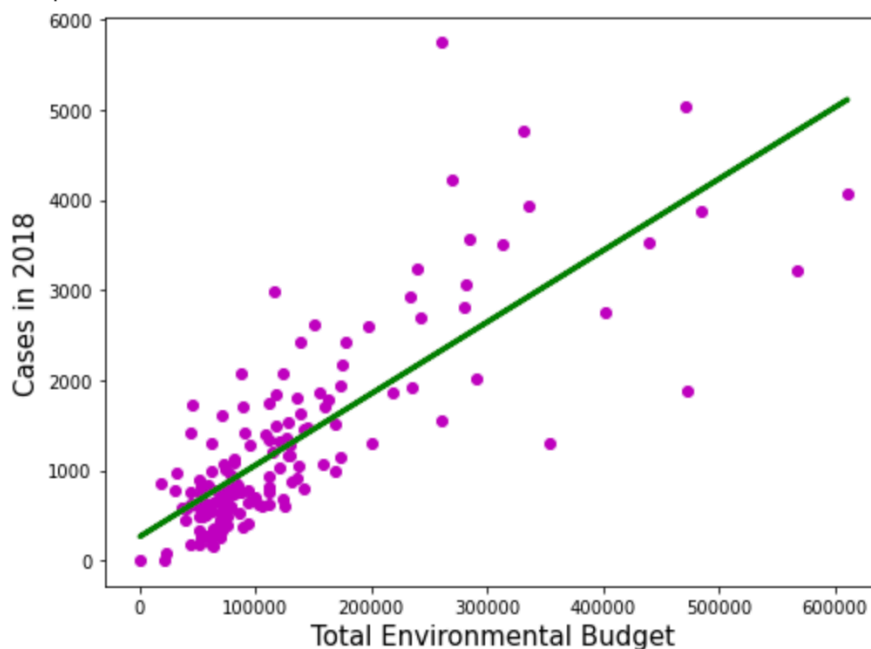

(Figure 1a)
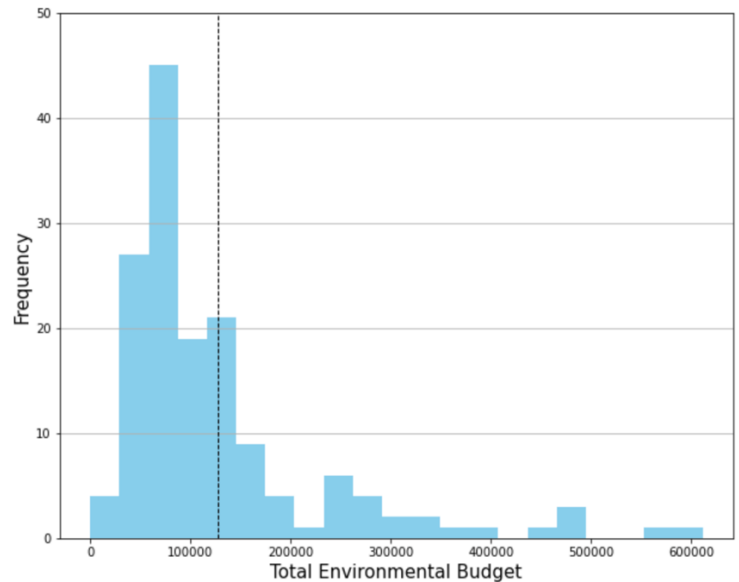
(Figure 1b)
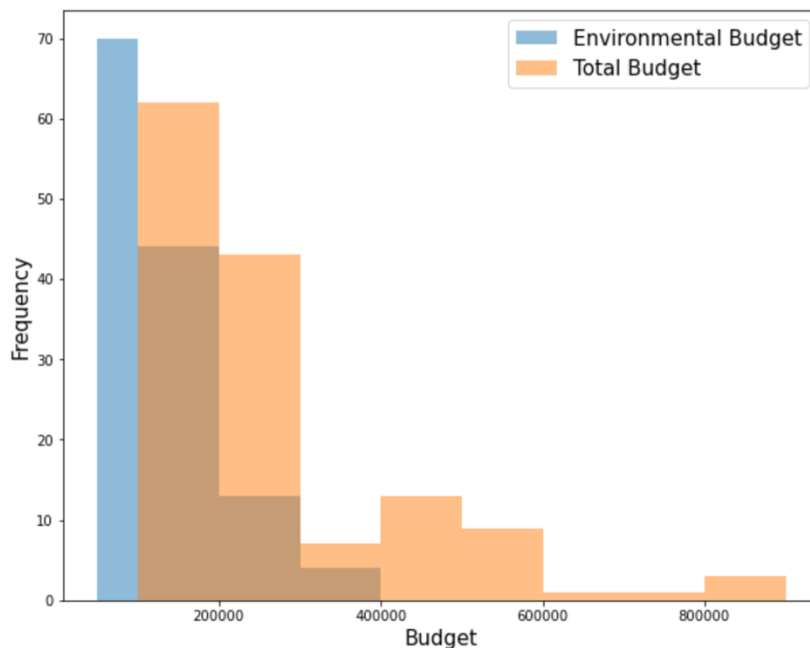
(Figure 1c)

For the second objective, by using the Spearman's rank correlation coefficient, the result is shown as below:

Spearmans correlation coefficient: 0.714

Samples are correlated (reject H0) p=0.000

Since it is result is between 0 and 1, it reflects that, with 95% confidence, the relationship between average increase cases and average environmental spending over the past ten years is positively correlated. In other words, it could mean that the high increase in average cases comes with the high amount of accumulated environmental spending. However, the evidence of correlation is still insufficient to demonstrate the effectiveness towards childhood obesity issue.

After ensuring the correlation, for elaborating the second method, *Figure 2* describes the linear relationship between the total environmental spending and the total cases in the latest year from the data. The R-Squared statistic of 0.6367 is positive and relatively high, suggesting it is a reasonable model. The calculated slope means that the total environmental budget increases slightly by 0.0079 when a case rises by one. Although there is a positive correlation between these two variables, it does not provide sufficient statistical evidence that the amount of environmental budget causes a significant impact towards the obesity cases.

(Figure 2)

Limitations

The analysis above assumes all other variables do not affect the environmental spending. The meaning could vary under different situations, for instance, such budget is determined by the level of population.

Conclusion

To sum up, statistical evidence indicate that the government has been spending excessively and ineffectively on environmental factors towards dealing with childhood obesity. Two out of six elements from the total budget contribute closely to the most frequent amount of total budget. Despite the marginal increase of environmental budget was minimal in rising the cases, budget spent should not have a negative outcome. In the future, the budget is recommended to reallocate and spend it on the areas which actually help to reduce the childhood obesity cases.

Reference

Github: https://github.com/wingochau/QM_Coursework1_2020.git

GOV.UK (2017). *Childhood obesity: a plan for action*. Available at: https://www.gov.uk/government/publications/childhood-obesity-a-plan-for-action/childhood-obesity-a-plan-for-action [Accessed: 15 November 2020].