# A Statistical Analysis of Obesity Prevalence and its Association with Socio-Economic and Lifestyle Factors.

Group-13

## 1 Introduction

The prevalence of obesity in Scotland has been a topic of increasing concern over recent years, prompting analysis through the Scottish Health Survey. This report examines trends in obesity prevalence within Scotland from 2013-2016, scrutinizing any notable shifts or patterns. Furthermore, it delves into disparities in obesity rates across various demographic factors including age, gender, socio-economic status, and lifestyle behaviours. By synthesizing data from the survey, this report aims to provide insights into the current landscape of obesity in Scotland, shedding light on potential areas for targeted interventions and public health initiatives. We will be answering the following questions: Has the prevalence of obesity in Scotland changed over the given years of the Scottish Health Survey and are there any differences in obesity by age, gender, socio-economic status, or lifestyle factors?

## 2 Exploratory Analysis

Table 1 displays the obesity rates by year in Scotland from 2013 to 2016. While there doesn't appear to be an obvious overall trend, it is worth noting the jump of 0.46% from 2014 to 2015. Perhaps if the data were sampled over more years we could discern any potential trends.

Table 1: Obesity by Year in Scotland

| Year | Percentage of Population Obese |
|---|---|
| 2013 | 30.30 |
| 2014 | 30.18 |
| 2015 | 30.64 |
| 2016 | 30.53 |

We then look at the prevalence of obesity in each age group, ranging from ages 16 to 99. It appears in Figure 1 that as age increases, obesity rates increase, up to the age of around 65. After this point, obesity rates start to fall, which makes sense, as those who aren't obese will live longer, on average. We also see that at the extreme upper ages, the obesity rates are 0, however there are significantly fewer samples which should be taken into account.

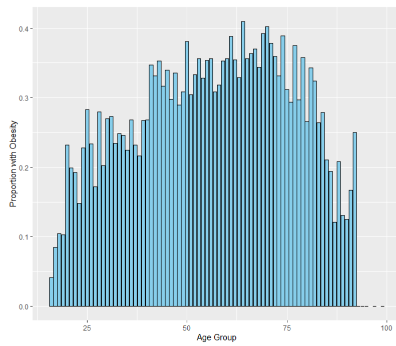


Figure 1: Proportions of Obesity by Age

In Figure 2 and Figure 3, we see the difference in obesity proportions between those who consume the recommended daily vegetable and fruit intake and those who don't, respectively. Intuitively, we see that obesity rates are lower among those who eat enough vegetables (around 2.8%), however there is not much difference between those who eat enough fruit or not, with obesity rates being marginally greater among those who do.

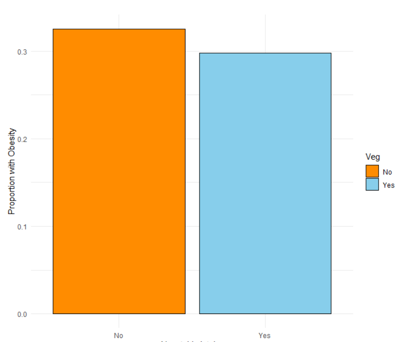


Figure 2: Obesity Proportion by Veg Intake

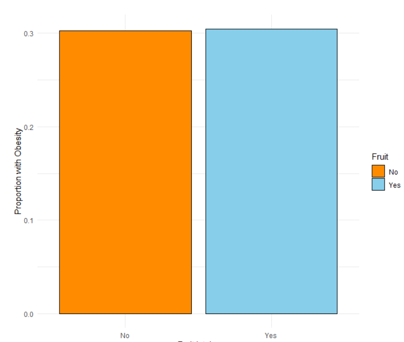


Figure 3: Obesity Proportion by Fruit Intake

Figure 4 displays the differing obesity proportions by the highest level of education the individual received. We can see that no qualifications ranks top, with around 38.2% being obese, while the lowest, those with a degree or higher, is almost 14% lower, at 24.7%.

Lastly, in Figure 5, we can see that obesity rates for males are about 2.3% lower than for females.

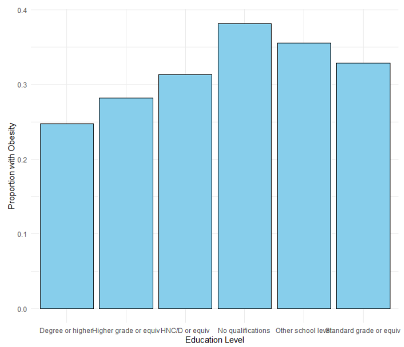


Figure 4: Obesity Proportion by Education

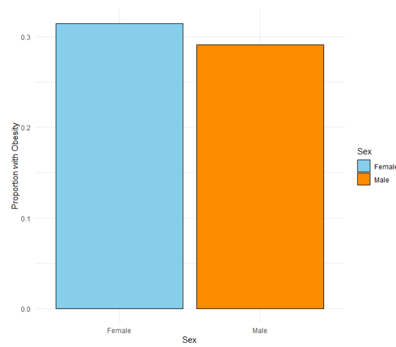


Figure 5: Obesity Proportion by Sex

## 3 Formal Analysis

We always use the logistic regression model while accounting for the binary nature of the response variable.

We begin by fitting a logistic regression model to explore the relationship between the prevalence of obesity and the year of the Scottish Health Survey. The first model can be represented as follows:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 \cdot \text{Year} + \epsilon_i, \tag{1}$$

where: $p_i$ represents the probability of an individual being classified as obese, $\beta_0$ is the intercept term, $\beta_1$is the coefficient corresponding to the variable Year, $\epsilon_i$ is the error term.

Table 2: Estimates of the first model coefficients

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| (Intercept) | -12.080 | 33.321 | -0.363 | 0.717 |
| Year | 0.005 | 0.016 | 0.338 | 0.736 |

The coefficient of Year is a positive value of 0.005586. This means that as the years increase (as time goes on), the odds of being obese also increase by approximately 1.006%.

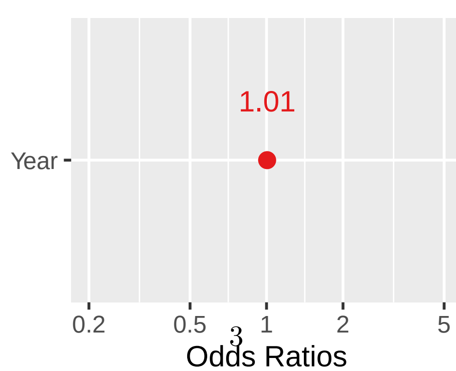


Figure 6: Odds ratios of the year and obesity.

The above figure shows the odds ratios of the year and obesity. From this data it follows that for each year increase between 2013-2016, the odds of an individual being classified as obese increase by a factor of 1.01.

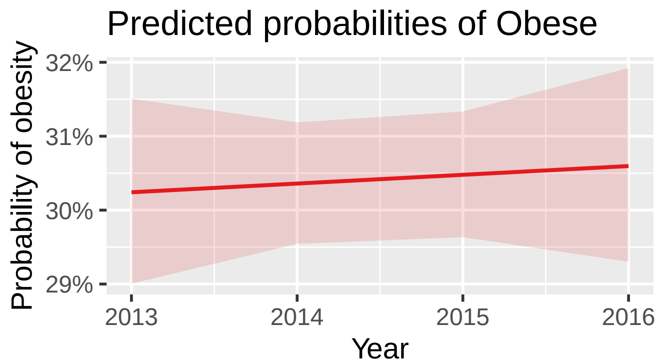


Figure 7: Probability of obesity as time goes on between 2013-2016 in Scotland.

This graph Figure 7 shows the percentage probability of an individual being obese as time increases between 2013-2016. The shaded part of the graph depicts the confidence intervals.The confidence intervals are fairly narrow and range from a gap of roughly 1.5% to 2% indicating that the estimated values are relatively stable. As we can see there is a positive correlation between the percentage probability of an individual being obese and the forwards movement of time.

The second model allows us to examine the relationship between various factors (age, gender, education, vegetable consumption, fruit consumption) and the probability of being classified as obese. It can be written as:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 \cdot \text{Age} + \beta_2 \cdot \text{SexMale} + \beta_3 \cdot \text{EducationHigher grade or equiv}+ \tag{2}$$

$$\beta_4 \cdot \text{EducationHNC/D or equiv} + \beta_5 \cdot \text{EducationNo qualifications}+ \tag{3}$$

$$\beta_6 \cdot \text{EducationOther school level} + \beta_7 \cdot \text{EducationStandard grade or equiv}+ \tag{4}$$

$$\beta_8 \cdot \text{VegYes} + \beta_9 \cdot \text{FruitYes} + \epsilon_i, \tag{5}$$

where: $\beta_0$ is the intercept term, $\beta_1, \beta_2, \ldots, \beta_9$ are the coefficients corresponding to the explanatory variables (Age, SexMale, Education levels, VegYes, and FruitYes) respectively.

Table 3: Estimates of the second model coefficients

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| (Intercept) | -1.522 | 0.079 | -19.236 | < 2e-16 |
| Age | 0.011 | 0.001 | 9.234 | < 2e-16 |
| SexMale | -0.124 | 0.037 | -3.308 | 0.00094 |

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| Ed Higher grade or equiv | 0.226 | 0.059 | 3.800 | 0.00015 |
| Ed HNC/D or equiv | 0.358 | 0.063 | 5.631 | 1.79e-08 |
| Ed No qualifications | 0.462 | 0.058 | 7.904 | 2.70e-15 |
| Ed Other school level | 0.312 | 0.086 | 3.616 | 0.00029 |
| Ed Standard grade or equiv | 0.421 | 0.055 | 7.637 | 2.23e-14 |
| VegYes | -0.103 | 0.045 | -2.279 | 0.02265 |
| FruitYes | 0.003 | 0.042 | 0.073 | 0.94209 |

We see that the coefficient for Age is positive (0.011004), indicating that with each unit increase in age, the log odds of being classified as obese increase by approximately 1.1%. The coefficient for males is negative (-0.124389), suggesting that males have lower odds of being classified as obese compared to females.

Regarding education, individuals with higher grades or equivalent education, HNC/D or equivalent education, no qualifications, other school level education, and standard grade or equivalent education all have positive coefficients. This suggests that individuals with these educational backgrounds have higher odds of being classified as obese compared to Degree or higher level education.

Interestingly, the coefficient for consuming recommended daily vegetables (Veg Yes) is negative (-0.103471), indicating that individuals who consume recommended daily vegetable intake have lower odds of being classified as obese. However, the coefficient for consuming recommended daily fruit intake (Fruit Yes) is not statistically significant (p-value = 0.942087), suggesting that there is no significant association between fruit intake and obesity status.

We interpret the odds ratios as follows from Figure 8: For each year increase in the individuals' age, their odds of being classified as obese increase (by a factor of 1.01). Men's odds of being classified as obese were 0.88 times those of women. For educational levels, compared to individuals with a degree or higher education, those with higher grades or equivalent education had 1.25 times the odds of being classified as obese. Similarly, individuals with HNC/D or equivalent education had 1.43 times the odds, those with no qualifications had 1.59 times the odds, those with other school level education had 1.37 times the odds, and those with standard grade or equivalent education had 1.52 times the odds of being classified as obese.

Regarding lifestyle factors, individuals who consume recommended daily vegetables had 0.90 times the odds of being classified as obese compared to those who do not consume recommended daily vegetables. However, there was no significant association found between consuming recommended daily fruit intake and obesity status.

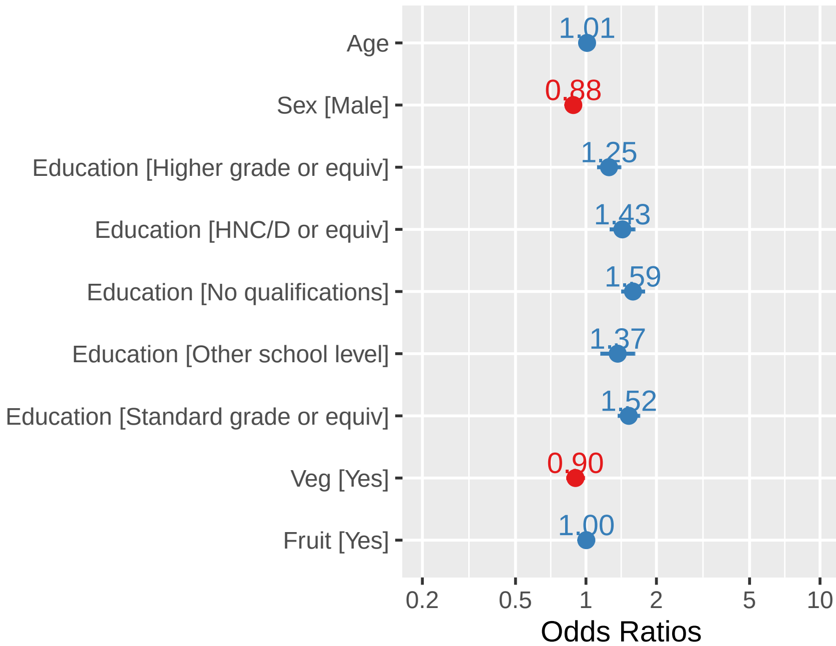


Figure 8: Odds of obesity classification with age, gender, socio-economic status or lifestyle factors.

## 4 Conclusions

In conclusion, this statistical analysis illuminates the prevalence of obesity in Scotland and its correlation with socio-economic and lifestyle factors, drawn from data spanning 2013 to 2016 from the Scottish Health Survey. Key findings reveal significant associations with age, gender, education, and vegetable consumption. Older age correlates with higher obesity rates, increasing by approximately 1.1% per unit age, while males exhibit lower odds than females (0.88 times). Lower educational attainment and reduced vegetable intake also heighten obesity risk. These results highlight obesity's multifaceted nature, shaped by demographic and lifestyle influences. Understanding these links informs targeted interventions to combat obesity, address disparities and promote healthy behaviours like vegetable consumption. Ongoing surveillance and comprehensive interventions are vital for addressing this pressing public health challenge and fostering a healthier future for Scotland.