

# CS5346 2023-24 OTOT Tasks B2

Zhao Shanhe\*

March 28th, 2024

## 1 Introduction

<i>Data Name</i>	<i>Description</i>	<i>Data Type</i>
gender	Gender with levels Female and Male	Discrete
age	Age	Continuous
marital_status	Marital status with levels Divorced, Married, Separated, Single, and Widowed	Discrete
highest_qualification	Highest education level	Discrete
nationality	Nationality with levels British, English, Irish, Scottish, Welsh, Other, Refused, and Unknown	Norminal
ethnicity	Ethnicity with levels Asian, Black, Chinese, Mixed, White, and Refused Unknown	Norminal
gross_income	Gross income with levels Under 2,600, 2,600 to 5,200, 5,200 to 10,400, 10,400 to 15,600, 15,600 to 20,800, 20,800 to 28,600, 28,600 to 36,400, Above 36,400, Refused and Unknown	Discrete
region	Region with levels London, Midlands & East Anglia, Scotland, South East, South West, The North, and Wales	Discrete
smoke	Smoking status with levels No and Yes	Discrete

\*Student ID: A0287084U

<i><b>Data Name</b></i>	<i><b>Description</b></i>	<i><b>Data Type</b></i>
amt_weekends	Number of cigarettes smoked per day on weekends	Continuous
amt_weekdays	Number of cigarettes smoked per day on weekdays	Continuous
type	Type of cigarettes smoked with levels Packets, Hand-Rolled, Both/Mainly Packets and Both/Mainly Hand-Rolled	Discrete

Data source: <https://www.kaggle.com/datasets/mexwell/uk-smoking-data>

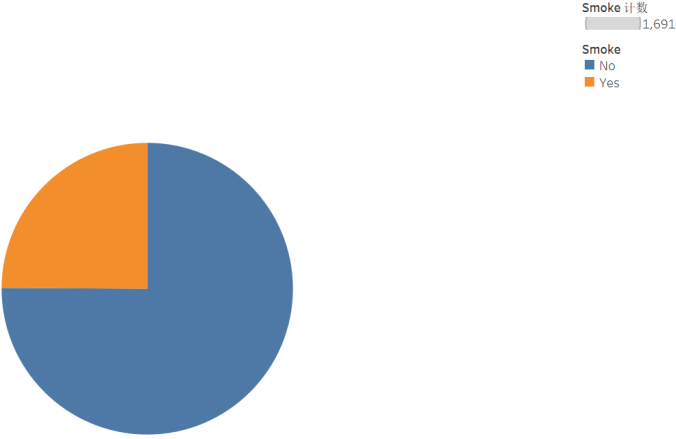
This task is done through **Tableau**.

The purpose of this visualization task is to find out the relationship between the proportion of smokers in the UK and gender, age, nationality, etc.

Task on Github: <https://github.com/Edward-EH-Holmes/CS5346-2023-24-Own-Time-Own-Target-OTOT-Tasks/tree/main/Task%20B/B2>

## 2 Visualization

The proportion of participants who smoked



The proportion of smokers in the survey was 24.90%.

Figure 1: The proportion of participants who smoked (Pie chart)

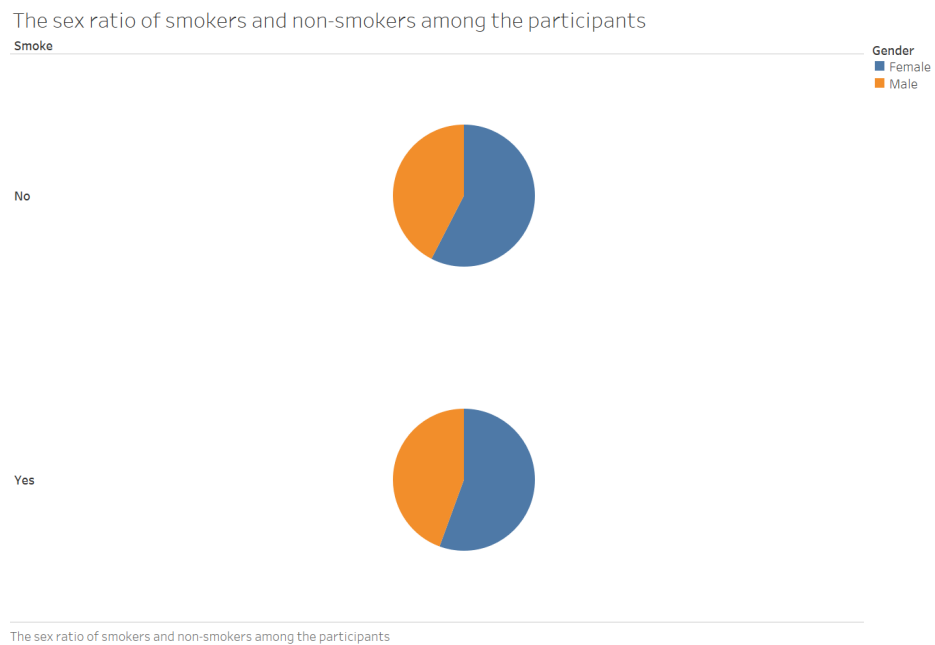


Figure 2: The sex ratio of smokers and non-smokers among the participants (Pie chart)

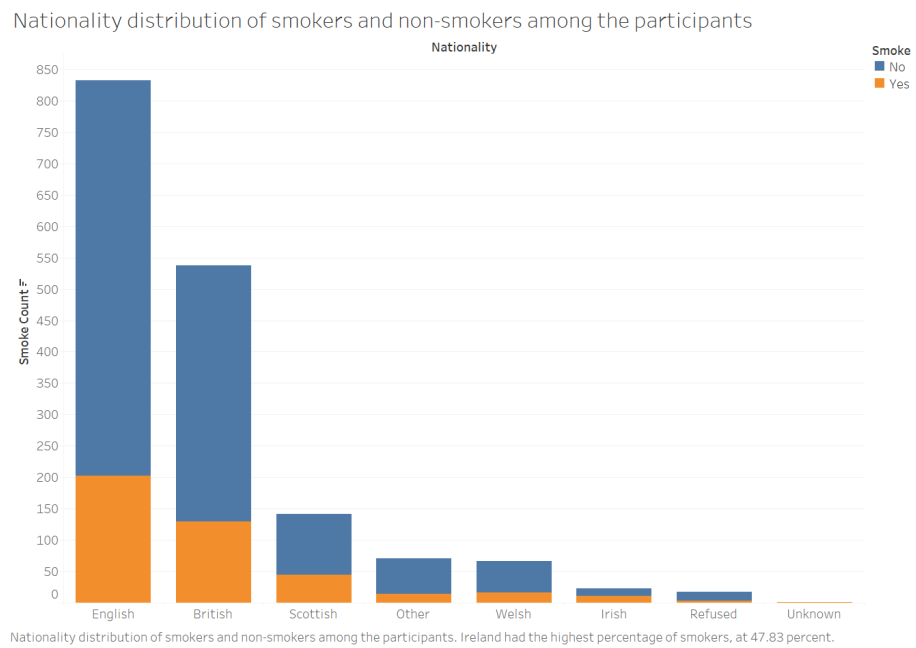


Figure 3: Nationality distribution of smokers and non-smokers among the participants (Stacked Bar chart)

Plots of average smoking per working day, average smoking per weekend versus income among smokers

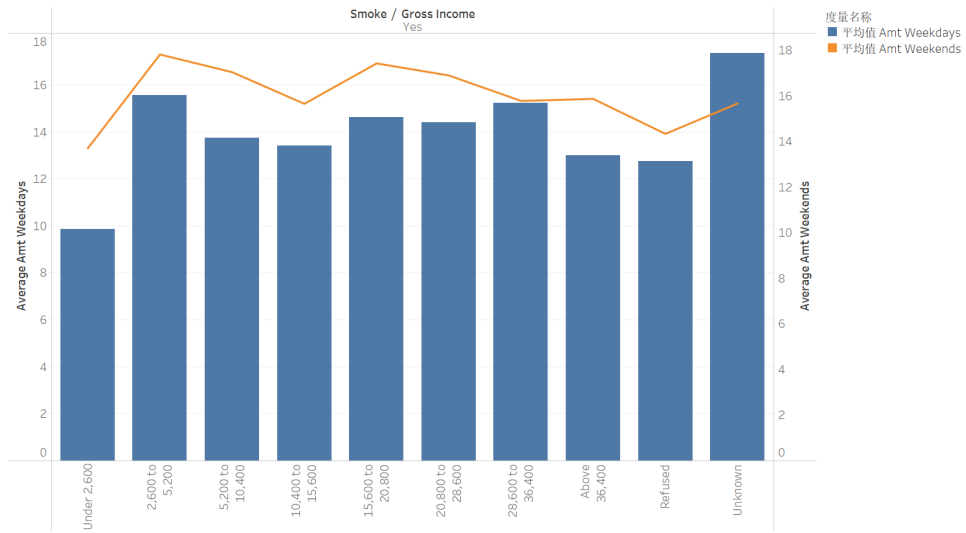


Figure 4: Plots of average smoking per working day, average smoking per weekend versus income among smokers (Line-Bar chart)

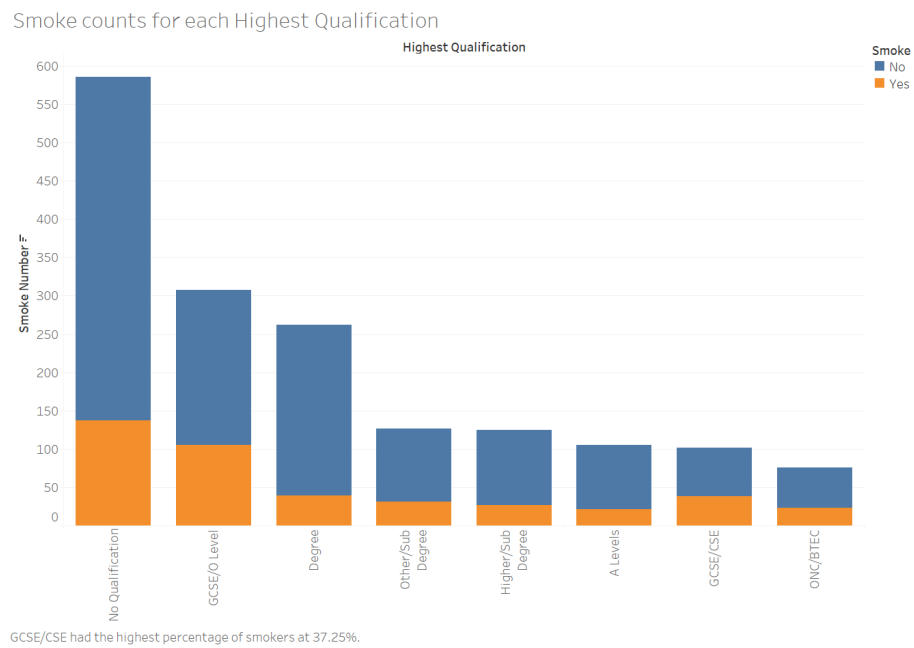


Figure 5: Smoke counts for each Highest Qualification (Stacked Bar chart)

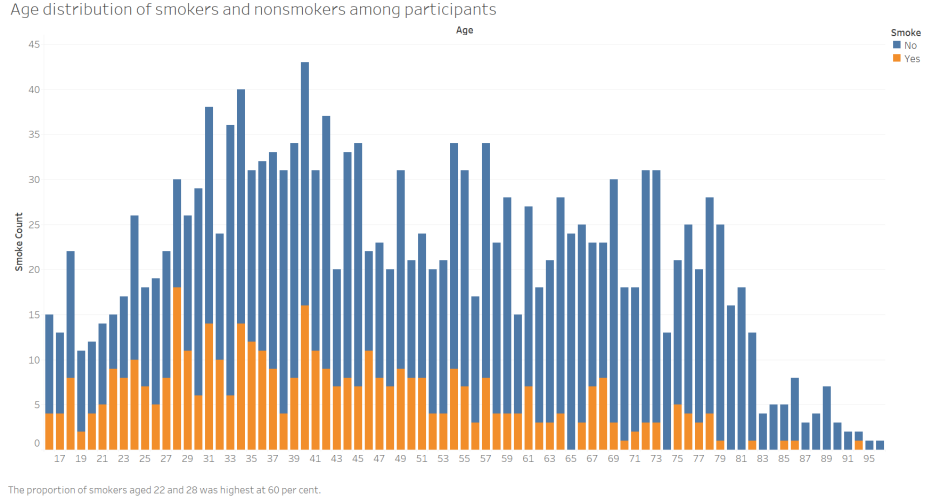


Figure 6: Age distribution of smokers and nonsmokers among participants (Stacked Bar chart)

### 3 Conclusion

From this, we can draw the following conclusions:

- From Figure 3, there is a strong correlation between the percentage of smokers and nationality. Excluding those who did not want to disclose their nationality, Irish smokers had the highest percentage of Irish smokers at 47.83%. British smokers had the highest proportion of British respondents at 23.97%.
- From Figure 5, there is a strong correlation between the percentage of smokers and highest qualification. Among those surveyed, GCSE/CSE had the highest proportion of smokers at 37.25%. In contrast, those with degrees had the lowest percentage of smokers at 14.89%.
- From Figure 6, there is a strong correlation between the percentage of smokers and age. And smokers tend to be younger In the survey results, 22 and 28 years old smokers accounted for the highest proportion, 60%.