# Exploring Smoking Habits in the UK: Trends, Challenges, and the Role of NHS Cancer Screening

A Data-Driven Public Health Case Study

Ayurma Wagle

15th March 2025

A lungs made out of paper

AI-generated content may be incorrect.

# Introduction

Smoking remains one of the most significant public health concerns in the United Kingdom, with well-documented risks contributing to serious health conditions such as lung cancer, chronic obstructive pulmonary disease (COPD), and cardiovascular diseases. Despite the steady decline in smoking prevalence over the years, the habit continues to disproportionately affect certain groups, particularly those from socio-economically disadvantaged backgrounds.

In 2023, approximately 11.9% of adults aged 18 years and over were current smokers, equating to around 6 million individuals. This marks the lowest proportion of smokers since records began in 2011, when the smoking rate was 20.2% (Office for National Statistics, 2023). The decline in smoking rates can be attributed to various public health interventions, including legislative measures such as the ban on smoking in public spaces, increased taxation on tobacco products, and smoking cessation support services (Action on Smoking and Health, 2023).

Smoking-related illnesses continue to place a substantial burden on the National Health Service (NHS), leading to thousands of preventable deaths each year. Lung cancer, the leading cause of cancer-related mortality in the UK, is primarily attributed to smoking, with around 70% of cases linked to tobacco use (Cancer Research UK, 2023). Recognising the need for early detection, the NHS recently introduced a lung cancer screening programme in selected areas of England. This initiative aims to detect lung cancer at earlier, more treatable stages among high-risk individuals, particularly long-term smokers (Department of Health and Social Care, 2023).

This case study explores smoking trends in the UK by conducting an in-depth analysis of the UK Smoking Data dataset. Using Power BI as the primary analytical tool, this study uncovers patterns in smoking behaviour and assesses their implications in the context of public health initiatives, such as the NHS lung cancer screening programme. The findings aim to provide valuable insights into the effectiveness of current tobacco control policies and highlight areas where further intervention may be needed.

# 2. Methodology

## 2.1. Aim:

The aim of this study is to analyse smoking behaviour in the UK using the UK Smoking Data dataset. It explores how smoking habits vary across different demographic groups and assesses their potential impact on public health. This analysis aligns with the recent NHS lung cancer screening programme, which targets high-risk individuals. By using data analysis and visualisation, this study provides insights that can support public health efforts to reduce smoking-related harm.

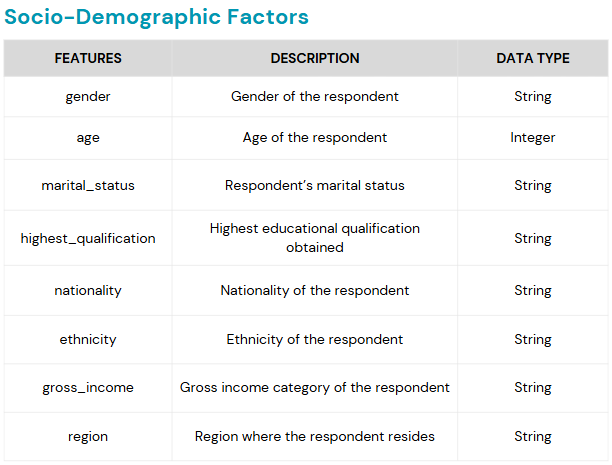
## 2.2. Dataset Description:

This study utilises the UK Smoking Data dataset, which contains 1,691 records detailing smoking habits among different populations in the UK. The dataset includes information such as age, gender, income, and smoking frequency. These factors help in understanding patterns in smoking behaviour and identifying groups that may benefit most from public health interventions.

To ensure a structured analysis, the dataset is divided into two main sections:

**2.2.1. Socio-Demographic Factors:**

Includes variables such as gender, age, marital status, highest qualification, nationality, ethnicity, gross income, and region.



**2.2.2. Smoking Habits:**

Covers smoking status, amount of cigarettes smoked on weekdays and weekends, and the type of tobacco product used.



## 2.3. Tools & Techniques

This study will use the following tools for analysis and visualisation:

* Power BI: Used to create visual reports and interactive dashboards.

The following analytical methods will be used to interpret the dataset:

* Data Preparation: Cleaning and transforming the dataset to ensure accuracy and consistency before analysis.
* Descriptive Statistics: Summarising key findings, such as smoking prevalence, average cigarette consumption, and demographic distributions.
* Comparative Analysis: Examining variations in smoking behaviour across different demographic groups, such as gender, income levels, and education attainment.

By employing these tools and techniques, this study aims to generate meaningful insights into smoking habits in the UK. The findings contribute to public health discussions and inform future strategies aimed at reducing smoking-related illnesses and supporting prevention efforts.

# Data Preparation

### Data Cleaning and Transformation

To ensure the dataset was suitable for analysis, several transformations were carried out using Power BI and Power Query. These steps helped standardise data, handle missing values, and create meaningful categories, improving the reliability of the findings.

### Standardising Categorical Variables

The dataset contained several categorical variables that required standardisation for consistency.

* The **Nationality** column initially included values such as "Refused" and "Unknown," which did not provide meaningful insights. These responses were changed to "Not Specific" to ensure clarity. Additionally, "British" was changed to "English" to align with standard classifications.
* The **Ethnicity** column originally listed "Chinese" as a separate category, despite it being a subgroup of "Asian." To align with broader demographic classifications, "Chinese" was merged into "Asian", while responses marked as "Refused" were recoded as "Not Specific". These changes improved consistency and ensured more accurate analysis of ethnicity and nationality.
* The **Highest Qualification** column was categorised based on the UK education system, grouping qualifications into meaningful levels for clearer interpretation. This structured approach ensures that the data is logically sorted and can be effectively analysed.

A screen shot of a chart

AI-generated content may be incorrect.

* The **Gross Income** column contained responses such as "Refused" and "Unknown," which were consolidated under "Not Specific" to maintain data integrity. Standardising these responses prevented distortion in income-based insights and allowed for clearer comparisons across income groups. Additionally, the Gross Income categories were arranged in a logical order to improve data readability.

### Handling Missing Data and Non-Smokers

A significant portion of the dataset included missing values in smoking-related questions. Further investigation revealed that **non-smokers had not been asked follow-up questions** about smoking frequency or cigarette consumption, leading to blank responses. To accurately reflect this, all non-smokers were assigned **"N/A"** in smoking-related fields, ensuring a clear distinction between smokers and non-smokers.

Additionally, the Amount of Cigarettes on Weekdays and Weekends columns (amt\_weekdays, amt\_weekends) contained "N/A" values for non-smokers, causing Power BI to interpret these columns as text. To correct this issue, all "N/A" values were replaced with 0, allowing the columns to be recognised as numeric fields. A similar transformation was applied to the Smoking Type column, where non-smokers had "N/A" recorded. These values were replaced with "Non-Smoker", making it easier to differentiate between smoking groups.

### Categorisation of Smoking Levels

To provide meaningful insights into smoking behaviour, smoking intensity was categorised based on guidance from the American Lung Association (American Lung Association, 2024):

A screenshot of a computer screen

AI-generated content may be incorrect.

These transformations ensured that the dataset was structured, complete, and ready for analysis. By improving consistency, handling missing values appropriately, and creating standardised categories, the dataset became a reliable foundation for understanding smoking behaviours in the UK. This structured approach allowed for clearer insights into smoking prevalence across demographic groups, supporting public health efforts in addressing smoking-related harm.

# Data Analysis

To gain insights into smoking habits in the UK, an **Exploratory Data Analysis** was conducted using the cleaned dataset. This analysis examines demographic distributions, socio-economic patterns, and smoking behaviours, helping to identify trends and associations that may influence public health outcomes.

## Exploratory Data Analysis (EDA)

#### Age and Gender Distribution:

The dataset includes a **diverse age range**, with the highest concentration of respondents between **30 and 50 years old.** The number of respondents gradually declines in older age groups, particularly beyond **80 years.**

In terms of gender, the sample consists of **57.07% females (965 individuals)** and **42.93% males (726 individuals)**. Understanding **age and gender** distribution is crucial for identifying at-risk populations and tailoring smoking cessation programmes accordingly.

A close-up of a graph

AI-generated content may be incorrect.

#### Education Distribution:

A significant proportion of respondents reported having no formal qualifications (34.65%), while GCSE/O Level (18.21%) was the most common completed education. Fewer individuals pursued A Levels, vocational training (ONC/BTEC), or higher education degrees, suggesting that many respondents ended schooling at the secondary level.

Lower educational attainment has been linked to higher smoking prevalence and reduced awareness of smoking-related health risks, making this distribution an important factor in understanding smoking behaviours.

A graph of a number of bars

AI-generated content may be incorrect.

#### Education Distribution by Gender:

The education distribution by gender shows that more women than men have no formal qualifications or GCSE/O Level qualifications. However, at higher education levels (Higher/Sub Degree and Degree), the gender gap narrows, with men slightly outnumbering women.

This suggests that while more women may leave formal education earlier, those who continue to higher education achieve similar levels as men. Understanding education levels in relation to gender can provide insights into the socio-economic factors influencing smoking habits.

A graph of different colored squares

AI-generated content may be incorrect.

#### Income Distribution:

The majority of respondents fall within the £5,200 to £10,400 income range, followed by those earning £10,400 to £15,600 and £2,600 to £5,200, indicating a largely lower to moderate-income population.

As income increases, the number of respondents declines, with the above £36,400 group being the smallest. A notable portion of respondents did not specify their income, which could reflect financial uncertainty, non-traditional income sources, or reluctance to disclose earnings.

Income levels can significantly influence smoking behaviour, as lower-income groups often have higher smoking prevalence due to increased stress levels and reduced access to smoking cessation resources.

A graph with blue and green bars

AI-generated content may be incorrect.

#### Income Distribution by Gender:

The income distribution by gender shows that men and women are fairly represented across lower and middle-income categories, but higher-income groups are dominated by male respondents.

* The highest concentration of respondents falls within the £5,200 to £10,400 range, with more women than men.
* Women also outnumber men in the £2,600 to £5,200 and £10,400 to £15,600 brackets, indicating a higher proportion of women in lower to middle-income groups.
* However, in higher-income brackets (£20,800 and above), men are the majority, suggesting greater male representation in higher-earning roles or industries.

These findings suggest a gendered income disparity that may influence smoking patterns, as financial stability often plays a role in smoking habits and access to cessation support.

A graph of income category by gender

AI-generated content may be incorrect.

#### Ethnic and Nationality Categorisation:

Smoking behaviours and health risks can vary significantly across ethnic groups due to cultural, social, and economic influences.

* The majority of respondents are White (92%), followed by Asian, Black, and Mixed ethnic groups.
* Most respondents identify as English (81%), with smaller proportions identifying as Scottish, Welsh, Irish, and Other.
* The "Not Specified" category includes those who refused to answer or had missing nationality and ethnicity data.

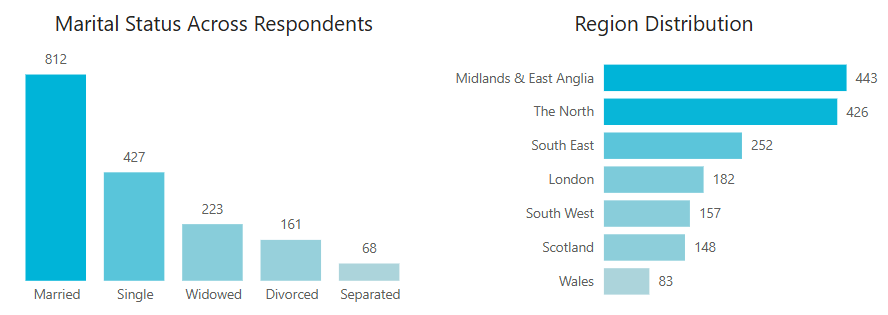
A close-up of a graph

AI-generated content may be incorrect.

Given that this is a UK-based survey, this distribution aligns with national demographic patterns, but it still allows for an exploratory analysis of smoking habits across different ethnic groups and nationalities.

#### Marital Status and Region Distribution:

The survey includes respondents from diverse marital backgrounds, with married individuals making up the largest group (48.04%), while separated respondents account for the smallest proportion (4.02%). Single and widowed individuals also form a significant share of the sample (25.26% and 13.19%, respectively), providing insights into how relationship status may influence smoking behaviour, particularly in relation to social and economic support systems.



Regionally, the highest representation is from Midlands & East Anglia (26.21%) and The North (25.20%), while Wales has the lowest proportion (4.91%). Other regions, including the South East (14.91%) and London (10.77%) contribute a moderate share.

#### Smoking Habit and Product Use Distribution:

The survey reveals that the majority of respondents do not smoke (75.1%), while 24.9% identify as smokers. This suggests that most individuals in the dataset are either non-smokers or former smokers. The quarter of respondents who smoke provides a significant group for further analysis of factors influencing smoking behaviour, such as income, education, and regional differences.

A blue rectangles with black text

AI-generated content may be incorrect.

Among smokers, the majority (70.6%) prefer packaged cigarettes, while 17.1% opt for hand-rolled cigarettes, which are often a more affordable alternative. A smaller proportion use both types, with 10% smoking a mix of packaged and hand-rolled cigarettes, and 2.4% relying mainly on hand-rolled options.

Understanding smoking product preferences provides insight into cost-related smoking choices and consumption patterns.

#### Amount of Consumption during Weekdays:

On weekdays, most smokers fall into the Light Smoker category (210 individuals), meaning they smoke fewer than 15 cigarettes per day. A smaller but still significant group, 144 individuals are Moderate Smokers, consuming between 15 and 24 cigarettes daily. The smallest group are Heavy Smokers, smoking 24 or more cigarettes per day.

The histogram confirms this pattern, showing that weekday smoking is most common between 5 and 20 cigarettes per day, with fewer individuals smoking at higher levels.

A graph of smoking distribution

AI-generated content may be incorrect.

#### Amount of Consumption during Weekends:

On weekends, smoking habits shift slightly. The number of Moderate Smokers increases to 175, meaning more people smoke between 15 and 24 cigarettes per day compared to weekdays. The Light Smoker category (153 individuals) remains significant, while Heavy Smokers (77 individuals) also increase.

The histogram reflects this, showing a higher concentration of people smoking 15 to 25 cigarettes per day.

A graph of smoking distribution

AI-generated content may be incorrect.

This suggests that some smokers increase their cigarette consumption on weekends, possibly due to more free time, social settings, or reduced work-related restrictions.

#### Smoking Patterns by Age: Weekday vs. Weekend Trends:

Cigarette consumption varies across age groups, with distinct patterns emerging on weekdays and weekends. The data shows that younger smokers (under 30) and middle-aged individuals (40-50) tend to smoke more on weekends than weekdays, likely influenced by social settings and free time. The trend stabilizes in older age groups, where weekday and weekend smoking habits become similar before declining in individuals over 60 years.

A graph of smoking patterns

AI-generated content may be incorrect.

The line chart clearly illustrates these changes, with smoking activity peaking in the 20s and 40s age groups, followed by a steady decline as age increases. This trend suggests that factors such as lifestyle, work schedules, and growing health awareness play a role in shaping smoking habits over time.

## Conclusion:

This analysis highlights the socio-economic and demographic factors associated with smoking behaviours.

* Smoking prevalence is higher in lower-income groups and among individuals with fewer formal qualifications.
* Men are more represented in higher-income brackets, while women have a slightly higher smoking prevalence.
* Weekend smoking is more common, particularly among younger and middle-aged adults.

These insights contribute to understanding smoking habits in the UK and provide valuable input for public health policies.

# **Conclusion and Public Health Implications**

This study analysed smoking habits in the UK using a dataset of 1,691 respondents. The findings indicate that smoking prevalence is higher among individuals with lower educational attainment and lower incomes, highlighting the role of socio-economic factors in smoking behaviour.

Among smokers, the majority prefer packaged cigarettes, while a smaller proportion opt for hand-rolled tobacco. The analysis also revealed that weekend smoking is more frequent than weekday smoking, particularly among younger and middle-aged adults, suggesting that lifestyle and social factors influence smoking patterns.

### Implications for Public Health

The findings align with data from the Office for National Statistics (ONS), which reports that while smoking rates have declined in the UK, they remain disproportionately high among individuals from lower socio-economic backgrounds (ONS, 2023).

Smoking continues to be one of the leading causes of preventable disease, contributing to lung cancer, cardiovascular disease, and respiratory illnesses (NHS, 2023). Despite progress in reducing smoking prevalence, health disparities persist, with disadvantaged groups experiencing a higher burden of smoking-related illnesses (Public Health England, 2023).

This highlights the need for continued public health interventions, particularly those targeting vulnerable groups through education, smoking cessation support, and policy enforcement.

### Potential Impact of NHS Lung Cancer Screening

Lung cancer remains one of the most significant health risks associated with smoking, with early detection being key to improving survival rates. Recognising this, the NHS has introduced a nationwide lung cancer screening programme to detect lung cancer at earlier, more treatable stages. This initiative is specifically aimed at individuals aged 55 to 74 who are current or former smokers, as they are at the highest risk of developing the disease(NHS England, 2024).

The programme is being rolled out across England, with an ambitious goal of achieving 40% coverage of the eligible population by 2025 and full coverage by March 2030. The phased implementation ensures that resources are effectively allocated to reach those at the highest risk. Since its inception, over one million individuals have been invited for lung health checks, leading to the detection of more than 5,500 cases of lung cancer. Importantly, more than 75% of these cancers were diagnosed at stage 1 or 2, where treatment options and survival rates are significantly improved (UK National Screening Committee, 2025).

One of the programme’s defining features is its accessibility strategy, which includes the use of mobile scanning units situated in community locations such as supermarket car parks and sports stadiums. This approach reduces barriers to participation, particularly for individuals in underserved areas who may otherwise struggle to access specialist healthcare facilities.

Findings from this case study highlight higher smoking prevalence among individuals from lower-income backgrounds and those with lower educational attainment. These groups are often at greater risk of developing smoking-related illnesses due to long-term exposure, increased consumption rates, and limited access to healthcare. As such, this initiative is expected to be particularly beneficial to these at-risk populations by providing targeted screening and early intervention opportunities

By focusing on early detection and intervention, the NHS lung cancer screening programme has the potential to save thousands of lives, reduce the burden of smoking-related illnesses on the NHS, and contribute to the broader goal of improving public health outcomes in the UK.

### Final Thoughts

This case study highlights the importance of data analysis in understanding smoking behaviours and public health trends. By identifying patterns in smoking habits, policymakers and healthcare professionals can develop targeted interventions that promote smoking cessation, reduce smoking-related diseases, and improve overall health outcomes.

The findings reinforce the need for continued investment in education, prevention campaigns, and accessible screening programmes. These efforts are essential to address health inequalities, support national smoking reduction strategies, and ultimately reduce the burden of smoking-related illnesses on the NHS and wider society.

# References:

Pixabay (2024). AI-generated cigarette smoke lungs illustration. Available at: https://pixabay.com/illustrations/ai-generated-cigarrette-smoke-lungs-8811081/ [Accessed 15 March 2025].

Action on Smoking and Health. (2023). Smoking statistics. Available at: <https://ash.org.uk/resources/view/smoking-statistics>

American Lung Association (2024). Tobacco Trends Brief: Smoking Intensity Categories. Available at: www.lung.org [Accessed 15 Mar. 2025].

Cancer Research UK. (2023). Lung cancer statistics. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer

Department of Health and Social Care. (2023). New lung cancer screening roll out to detect cancer sooner. Available at: <https://www.gov.uk/government/news/new-lung-cancer-screening-roll-out-to-detect-cancer-sooner>

NHS England (2024). Lung cancer screening roll-out to detect cancer sooner. Available at: https://www.england.nhs.uk/news/2024/lung-cancer-screening/ [Accessed 10 March 2024].

Office for National Statistics (ONS) (2023). Adult smoking habits in the UK: 2023. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandlifeexpectancies/bulletins/adultsmokinghabitsingreatbritain/2023 [Accessed 10 March 2024].

Public Health England (2023). Health inequalities and smoking prevalence. Available at: https://www.gov.uk/government/publications/smoking-and-health-inequalities [Accessed 10 March 2024].

NHS (2023). Smoking and your health. Available at: https://www.nhs.uk/live-well/quit-smoking/why-quit-smoking/ [Accessed 10 March 2024].

NHS England (2024). Thousands of cancers caught early through NHS lung checks. Available at: www.england.nhs.uk