

Analysis of Health Survey for England (HSE) 2019

Candidate Numbers Here

March 07, 2024

Abstract

This report provides an analysis of data related to health, age, socio-economic factors and lifestyle habits in adults (from the age of 16) from the population in England, derived from the Health Survey for England 2019.

Introduction

In the UK, smoking, vaping, and alcohol consumption are widespread, particularly among youngsters. It is crucial to be aware of the consequences and dangers of these habits, and the complications they can cause in later-life. The Office for National Statistics (ONS), regarded as the foremost statistical institute in the UK, is the “go-to” for insights into public health trends. According to their estimations, approximately 14.1% of the adult population aged 18 and above were identified as cigarette smokers (ONS 2019a). Furthermore, their data revealed there were 7,565 deaths attributed to alcohol-specific causes in 2019 (ONS 2019b). These statistics alone warrant a need for an understanding of health-related behaviours and the outcome. Therefore, the aim of our study is to investigate not only the prevalence but also the severity of these habits, exploring different socioeconomic factors that may potentially contribute to each. Additionally, we will investigate the greater implications of these bad habits and how they relate to systolic blood pressure levels throughout the UK population.

Question One

We start by investigating the data

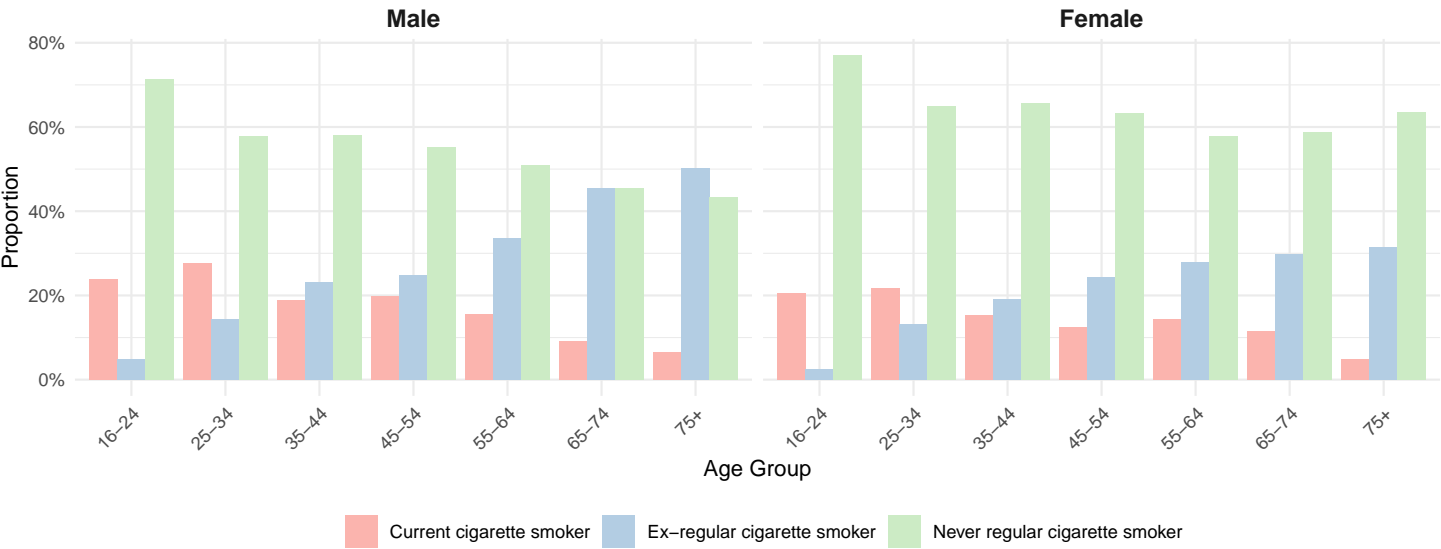


Figure 1: Smoking status proportions by age group and gender

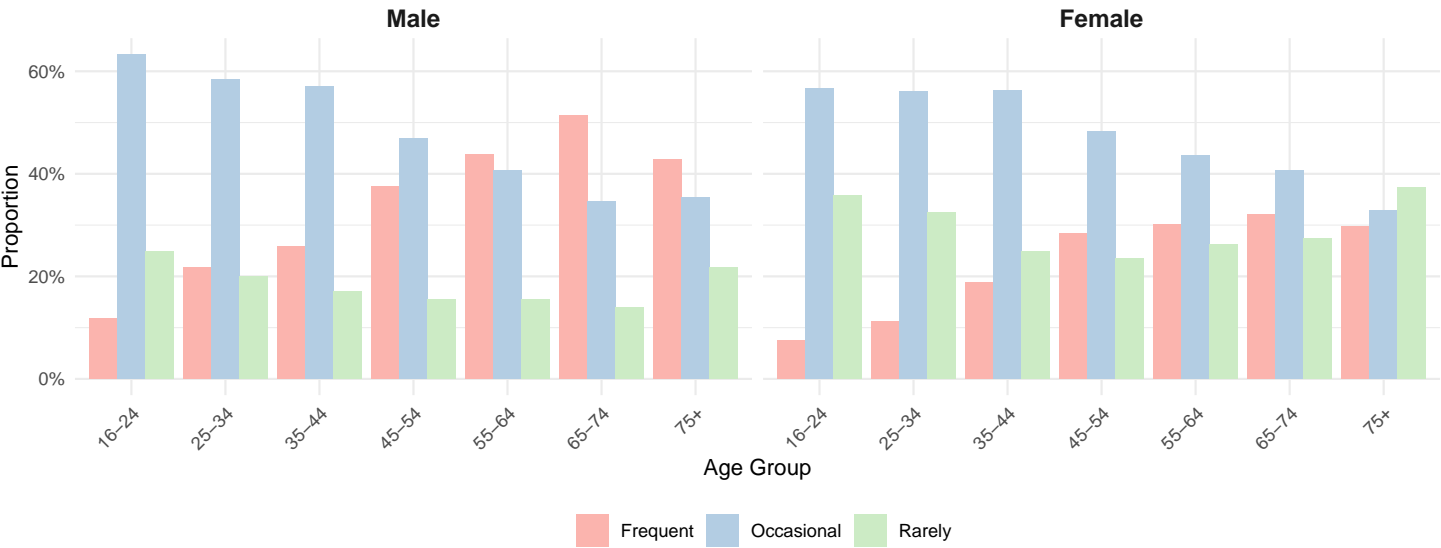


Figure 2: Drinking Status Proportions by age group and gender

Assuming a binomial...

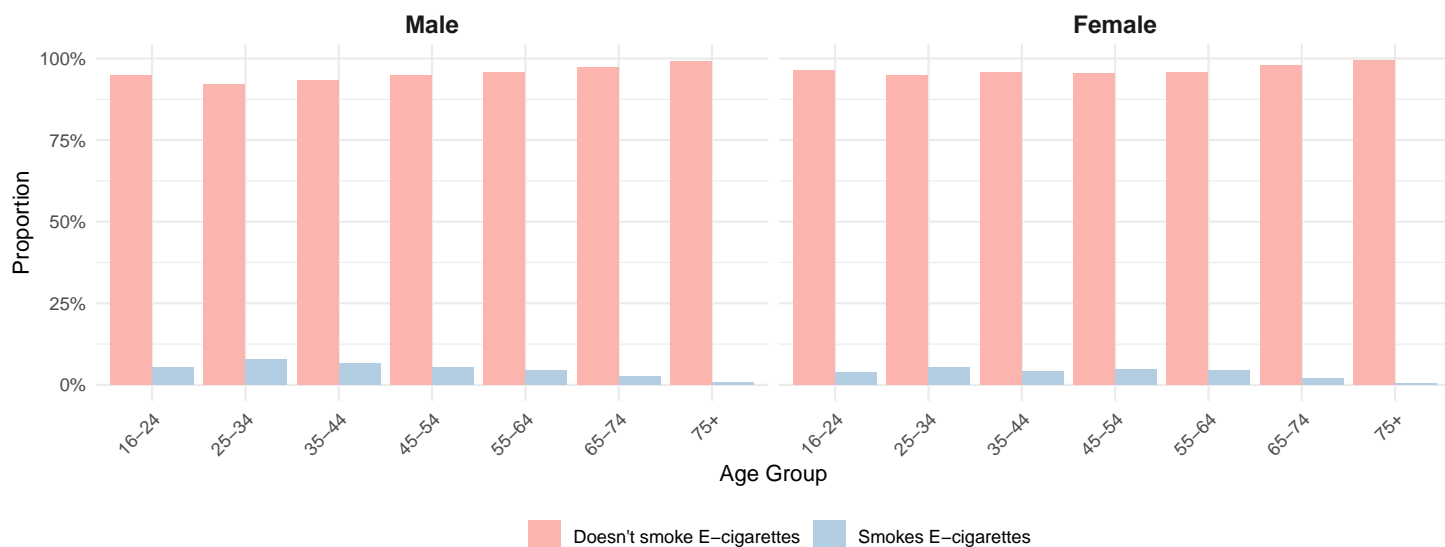


Figure 3: E-cigarette Usage Status Proportions by age group and gender

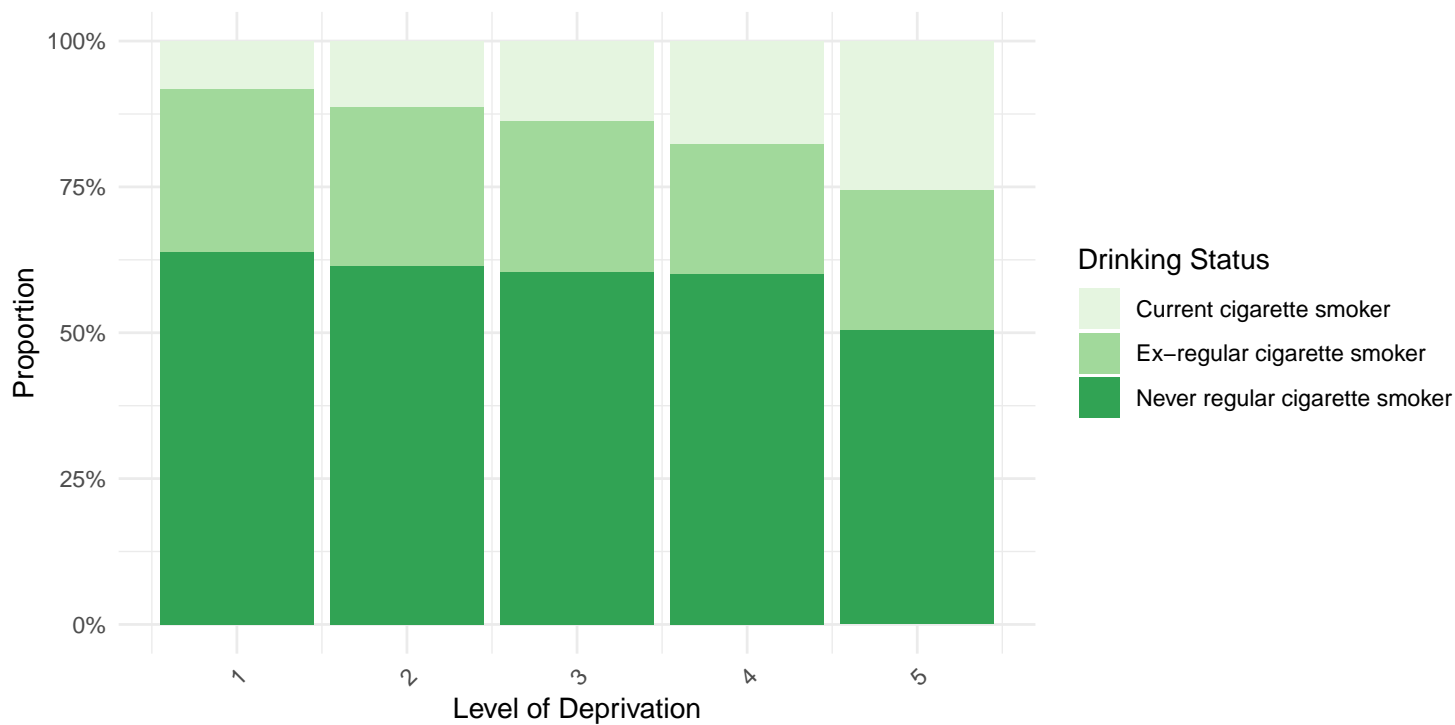


Figure 4: Drinking Status Proportions by Level of Deprivation

Table 1: Estimates and 95% Confidence Intervals for % of Population

Category	Estimate	C.I.
Drinking	80.3%	(79.5%, 81.2%)
Smoking	16.5%	(15.7%, 17.3%)
Smoking E-cigarettes	4.28%	(3.84%, 4.72%)

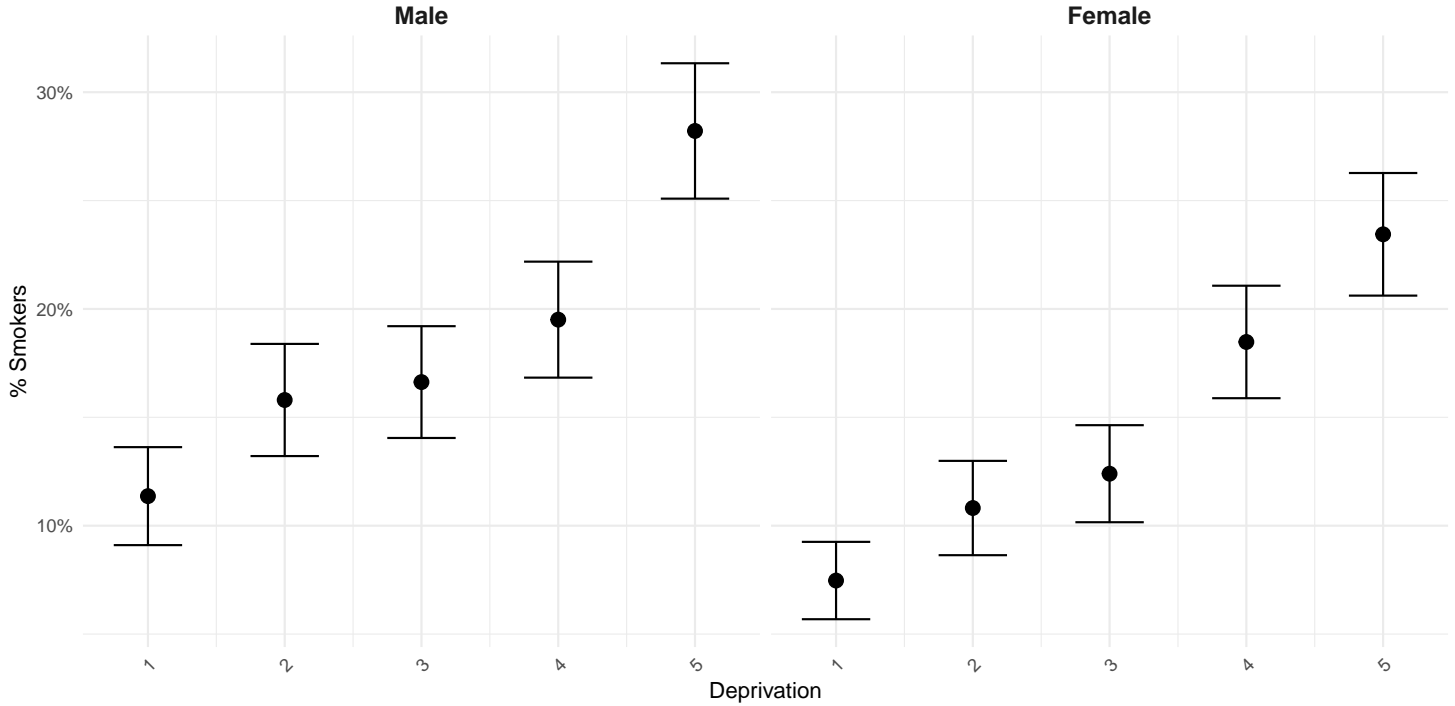


Figure 5: Estimation of the prevalence of smokers by deprivation

Question Two

Explanation of why we split into test/train dataset.

This leaves us with 6563 observations in our training dataset.

Table 2: Missing values in the training dataset

Variable	Missing Values	% Missing
omsysval	3217	49%
BMIVal	1203	18.3%
dnoft_19	1187	18.1%
cigdyal_19	45	0.686%
cigsta3_19	44	0.67%
NDPNow_19	43	0.655%
d7many3_19	40	0.609%
drinkYN_19	39	0.594%
topqual2	38	0.579%
origin2	24	0.366%
marstatD	1	0.0152%

Creating first model on train data

References

ONS. 2019a. “Adult smoking habits in the UK: 2019.” <https://shorturl.at/qQW27>.

———. 2019b. “Alcohol-specific deaths in the UK: registered in 2019.” <https://shorturl.at/gqxY1>.