

# NHS Insights – Project Brief

## Overview

You are part of a team of data analysts that was contracted by the National Health Services (NHS), a publicly funded healthcare system in England. The team has been provided with internal and external data and a number of high-level business questions concerning the utilisation of services, missed appointments, and the potential value of using external data sources such as Twitter (now rebranded as X) . Your role is to refine the business questions to actionable analytic questions based on your review of the available data. You will then plan and perform the exploratory data analysis and present the insights to the respective stakeholders.

The NHS must expand its infrastructure and resources to match its increasing population capacity. For this, it needs to budget correctly. When deciding on budget allotment, the NHS must understand the utilisation trends of each component in its network.

Some stakeholders feel that the NHS's capacity should be increased while others feel that, based on current trends in utilisation, the current capacity is adequate and that efforts to make better use of existing infrastructure and resources are sufficient. Your team has been tasked to explore the available data and offer suggestions and recommendations based on your observations.

## Access the data

Historical data about all appointments at general practitioner (GP) surgeries across the country are publicly available from the NHS. The size and state of the data requires advanced cleaning and wrangling with a powerful computer and processing time to allow any meaningful analysis. Therefore, for this assignment, you have access to three semi-wrangled files. For example, unnecessary columns were dropped to create smaller files. This will ensure that you can perform the analysis needed to complete the assignment.

1. `actual_duration.csv` – Details of appointments made by patients. For example, the regional information, date, duration, and number of appointments pertaining to a certain class.
2. `appointments_regional.csv` – Details on the type of appointments made by patients. For example, regional information, the month of appointment, appointment status, healthcare professional, appointment mode, the time between booking and the appointment, as well as the number of appointments pertaining to a certain class.
3. `national_categories.xlsx` – Details of the national categories of appointments made by patients. For example, the regional information, date of appointment, service setting, type of context, national category, and the number of appointments pertaining to a certain class.

4. metadata\_nhs.txt – Details of the data set, data quality, and reference.
5. tweets.csv – Data related to healthcare in the UK scraped from Twitter.