Background/context of the business scenario (100 words):

Briefly describe the business scenario (context) and the business problem you aim to solve.

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 NHS incurs significant, potentially avoidable, costs when patients miss general practitioner (GP) appointments.

The reasons for missed appointments need to be better understood as explained by The British Medical Association (BMA) chair Professor Philip Banfield:

While it is frustrating when patients do not attend, the reasons why this happens should be investigated rather than simply resorting to punishing them. Financially penalising patients inevitably impact the poorest and most vulnerable in the community (GP Practice News 2022).

Therefore, reducing or eliminating missed appointments would be beneficial financially as well as socially. The government needs a data-informed approach to deciding how best to handle this problem. At this stage of the project the two main questions posed by the NHS are:

* Has there been adequate staff and capacity in the networks?
* What was the actual utilisation of resources?

Business problem – the National health service (NHS) is publicly funded

The NHS was put under enormous strain during the pandemic. Even now that life in the UK is somewhat returning to normality, the effects of COVID-19 are set to have an impact on our healthcare services for years to come due to the backlog of delays.

Members of the public have started to realise, more so than ever, how important it is to do their bit to keep our NHS functioning at its best and not waste funds or resources. One thing that everyone can do to avoid wasting NHS funds is to ensure that they attend their GP appointment when booked in, or cancel ahead of time if they cannot attend.

Each missed GP appointment costs the NHS an average of £30 per person (Source: [NHS](https://www.england.nhs.uk/2019/01/missed-gp-appointments-costing-nhs-millions/)).

Through analysing NHS data, we have been able to rank the areas in the UK which have the highest proportion of missed GP appointments, and which locations have cost the NHS the most money as a result.

Stoke-on-Trent came out as the area with the highest percentage of missed appointments in the country, with over 7% of all appointments unattended in the past year. As a result, this has cost the local NHS an estimated £2,935,020 per year – the equivalent annual salary of 73 additional doctors or 144 nurses.

Whilst Stoke-on-Trent has the highest ratio of missed GP appointments out of all the areas surveyed, North East London actually has the highest cost to its local NHS trust as a result of missed general practice appointments. This area’s appointments bill stacks up to an enormous **£16,927,260** a year, the equivalent of the yearly salaries of 420 doctors, 420 nurses or 833 healthcare assistants.

However, missed appointments are not always due to carelessness. According to a [British Journal of General Practice report](https://bjgp.org/content/71/707/e406), one of the main reasons for missed appointments is financial difficulties, with patients unable to get to their appointments or being unable to take the time off work to attend appointments.

<https://www.patientclaimline.com/article/which-areas-in-the-uk-have-the-most-missed-gp-appointments-and-how-much-does-it-cost-the-nhs/#:~:text=Each%20missed%20GP%20appointment%20costs,most%20money%20as%20a%20result>.

**Same Day Appointments**

**45.0% of appointments in February took place on the same day that they were booked.**

<https://digital.nhs.uk/data-and-information/publications/statistical/appointments-in-general-practice/february-2022>

However, steps were taken to inform patients of the cost to the NHS of missing their appointments after [a study by Imperial College London showed that fewer appointments would be missed](http://www.imperial.ac.uk/patient-safety-translational-research-centre/our-work/behavioral-interventions-/) if people were given this information.

<https://www.theguardian.com/society/2018/jan/02/patients-missing-their-appointments-cost-the-nhs-1bn-last-year>

By sharing appointments, more patients could be treated more quickly, reducing waiting times, saving costs, yet raising standards of care.

They have been tried by GPs in Edinburgh, Sheffield and Newcastle, following the lead of doctors in the US and Australia. As a surgeon, I can see the potential benefits in bringing together patients undergoing the same procedure for pre- and post-surgical care.

Shared appointments are not appropriate for all patients or all conditions. They should always be offered, never imposed, and patients would always retain the option of a one-to-one consultation, if that was what they preferred. There might, however, be trade offs. Patients might be offered a one-to-one consultation in four weeks or a shared appointment in 48 hours.

<https://www.theguardian.com/healthcare-network/2017/apr/18/shared-medical-appointments-help-nhs-patients>

Analytical approach (350 words):

Describe the approach taken to import, clean, and analyse data in Python.

Include a detailed and insightful description of the processes you used and the decisions you made during analysis, such as the choice of libraries, functions, and variables.

Ensure that the description of the steps taken to prepare data for analysis is clear, well organised, and relevant to the given scenario.

Approach

I was provided four datasets to analyse:

* actual\_duration.csv
* appointments\_regional.csv
* national\_categories.xlsx
* tweets.csv

For ease of manipulation, I imported all four into a jupyter notebook, so that I could perform my analysis using pythonic code.

My initial review determined that there was no missing data and that the data integrity was sound, i.e. no erroneous data.

**Actual\_duration.csv:**

Table

Description automatically generated

The definition of the appointment duration differs by practice. Therefore, any conclusions need to bear this in mind.

Unknown/Data Quality covers appointments of zero, less than 1 minute or more than 60 minutes.

By definition, an appointment having a duration dictates that it was attended. The main purpose of this analysis is to investigate appointments that were not attended. Therefore limited conclusions can be made from this dataset.

Visualisation and insights (350 words): Describe the rationale for the selected visualisations, and ensure that the interpretations of the visualisation outputs are detailed, insightful, and relevant to the business objectives.

Patterns and predictions (200 words): Clearly articulate any patterns, trends, or insights you discovered. Ensure they relate to the business scenario. Include recommendations on any areas for further exploration.

Recommendations:

* Create procedure for data entry standards -> ensures quality and aids analysis
* Standardise appointment management practices -> if possible, understand the different methods and identify efficiencies or correlation with desired outcome, e.g. costs of missed appointments is beneath acceptable threshold, or % of missed appointments is the lowest.
* Standardise systems -> metadata states multiple issues have been caused by particular systems, indicating a range of systems are used. One robust system would increase resilience and therefore data quality, e.g. standardised inputs/headings, definition of appt length
* I would want to attempt to quantify the mismapped data as it is believed to be a small proportion. I would want to test this hypothesis.