

CO4 LSE DA201 Data Analytics using Python

# DIAGNOSTIC ANALYSIS REPORT

Elaine Wong • February 2023

## Part 1. Background and purpose

In recent years, the demand for NHS medical services is increasing and the medical staff workforce is decreasing. For the first time in history, tens of thousands of NHS staff took part in strikes in December 2022 to demand fair pay and improved patient safety.

Meanwhile, some patients are hard to make medical appointments and some patients would not attend the booked appointment.

This study aims to identify the potential reasons of the staffing problem and patient no show problem.

## Part 2. Analytical approach

Except for the number of appointments, those datasets only provide categorical variables.

Firstly, to investigate the staffing issue, we would compare the demand side (appointment) and the supply side (staff capacity) by several dimensions:

- Month
- Location
- Patient pathway: hcp type -> mode -> waiting time -> status
- Medical service: service setting -> national category
- Appointment duration

Secondly, to investigate the no-show problem, we would look into the DNA appointments individually.

There is a list of potential factors may cause patients not attend the appointment, it is a good start to find reasons from the history record.

## Part 3. Data ingestion and wrangling

Three datasets have been used for the analysis. There is no unique appointment identity for identifying the appointment record so we could not combine the appointment records between different datasets.

### **Dataset 1: actual duration**

This dataset contains the appointment duration record for the actual **attended** appointments from 1 December 2021 to 30 June 2022 on daily basis per sub-ICB.

### **Dataset 2: regional appointments**

This dataset contains the records of hcp type, appointment mode, waiting time between the booking and appointment, and appointment status for all appointments from January 2020 to June 2022 on monthly basis per icb.

There are some duplicated records which account for 5% of data. By cross-checking with another two datasets, the number of appointments could be aligned and all duplicated records have been keeping to ensure the number of appointments is correct.

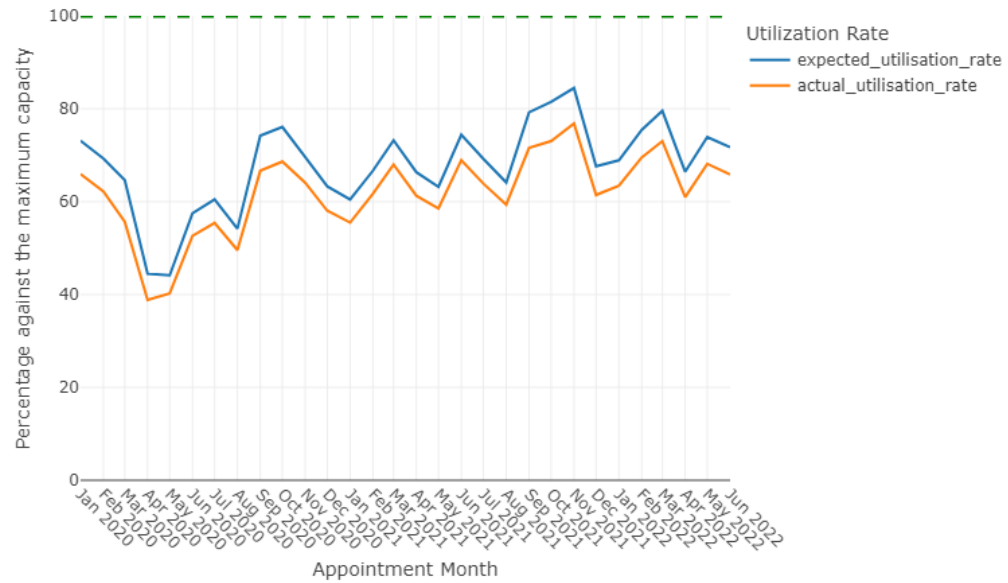
### **Dataset 3: national category**

This dataset contains the records of service settings and national categories for of all appointments from 1 August 2021 to 30 June 2022 on daily basis per sub-ICB.

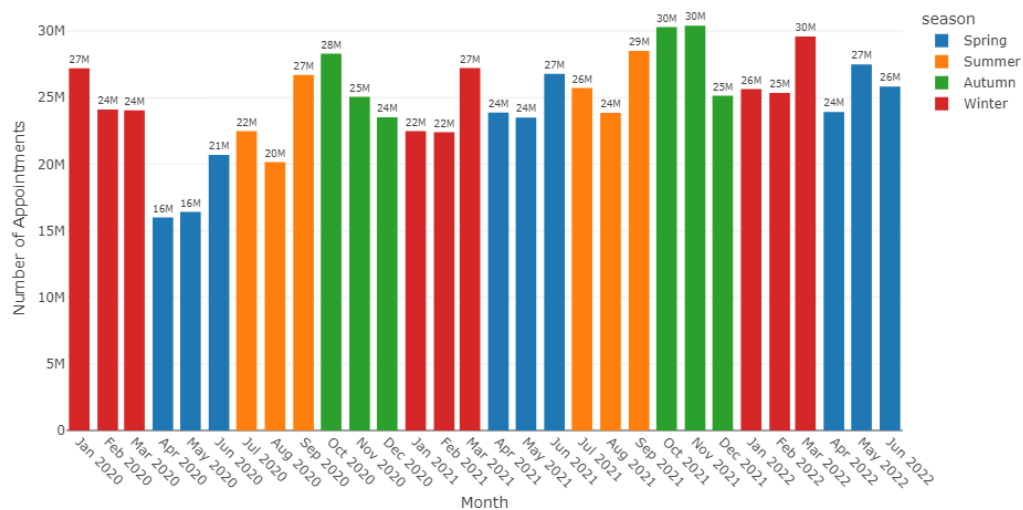
## Part 4. Findings and visualization

- Numbers of appointments per month

Capacity utilization rate per month



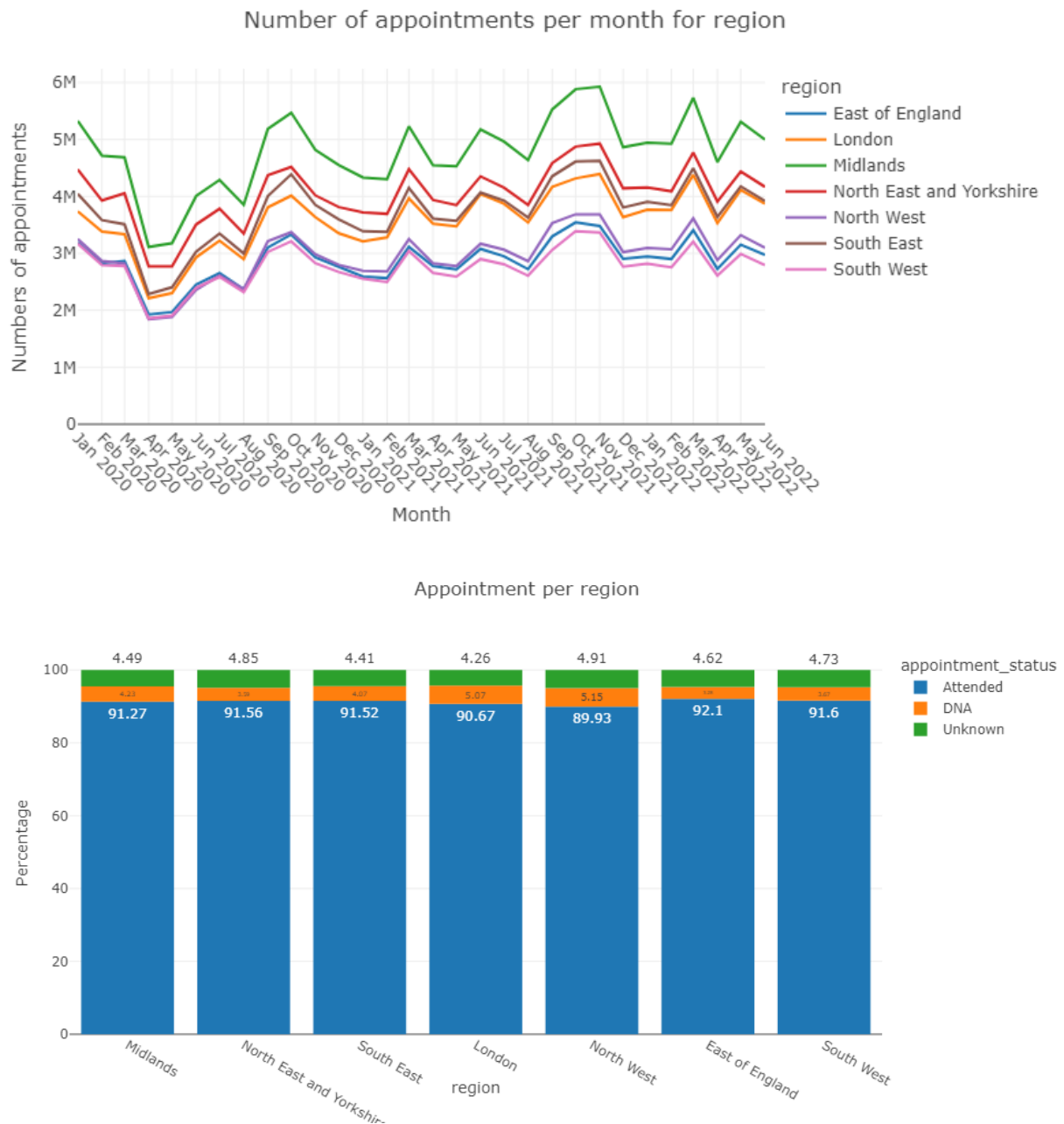
Number of appointments per month



The number of appointments is in an increasing trend after April 2020. Some possible reasons are COVID-19 and the increasing population in UK.

Compare to the estimated maximum capacity, the number of appointments usually does not reach 80% of the capacity.

- Regions

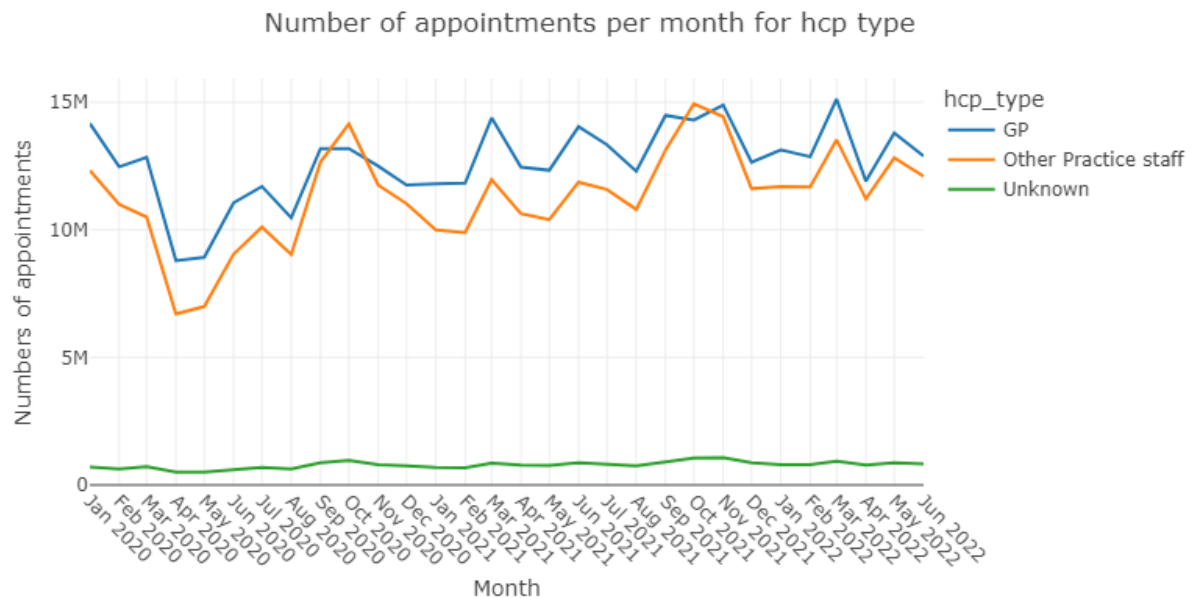


There is no abnormal trend for different regions.

However, the appointment difference between regions is huge.

It should be related with the population size in the region.

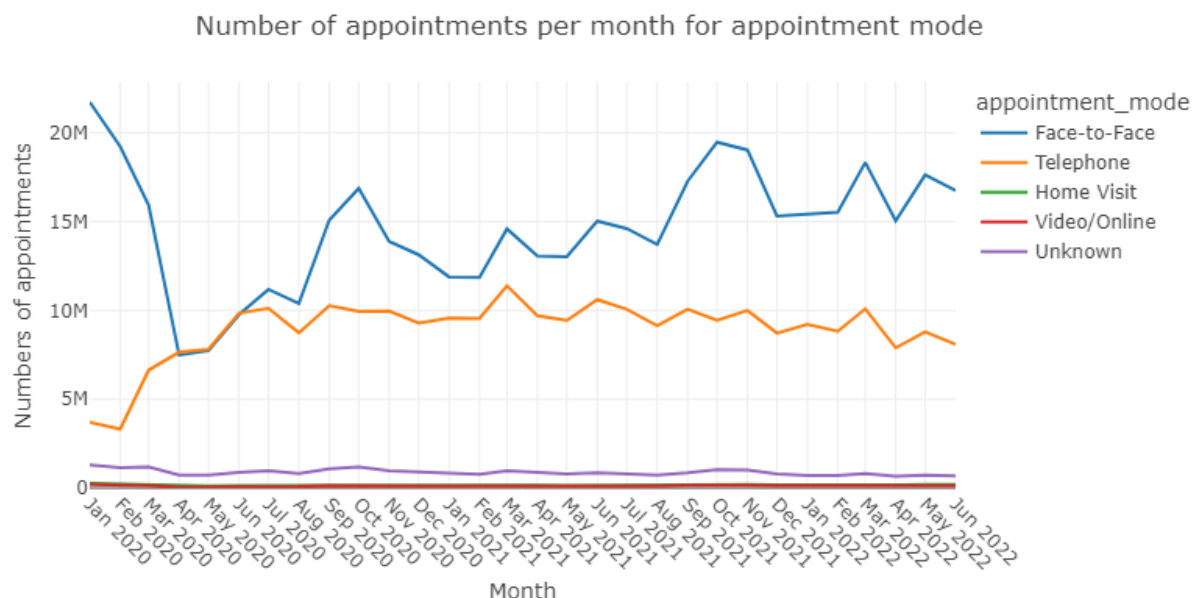
- HCP types



The numbers of appointments for GP and other practice staff are quite similar, but the fluctuation of the appointments of other practice staff is larger than GP appointments.

We can see that in Oct 2020 and Oct 2021, the appointments of other practice staff had been more than GP appointments, and it is keep increasing.

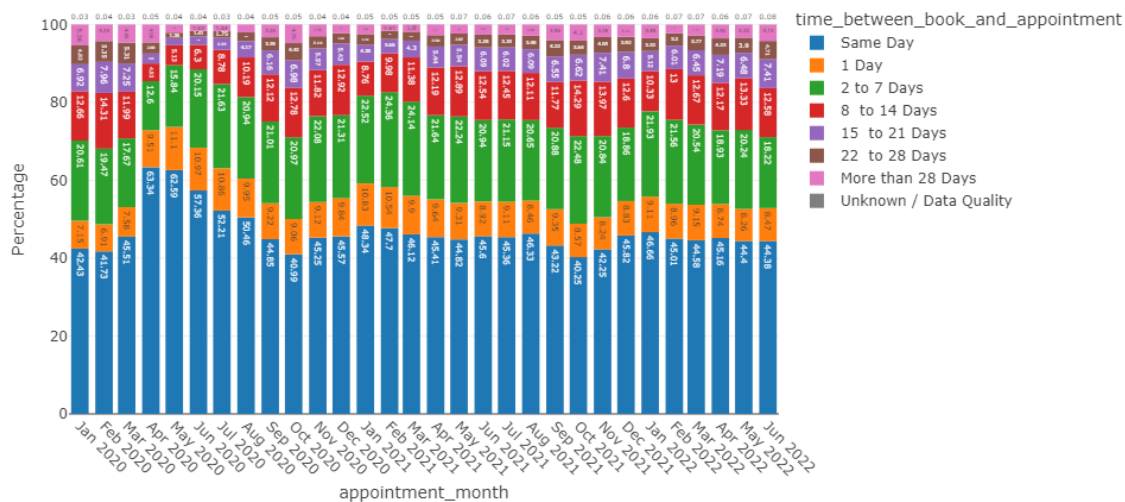
- Appointment modes



The telephone appointment is decreasing, video/online appointment has a trend to increase.

- Waiting time between booking and appointment

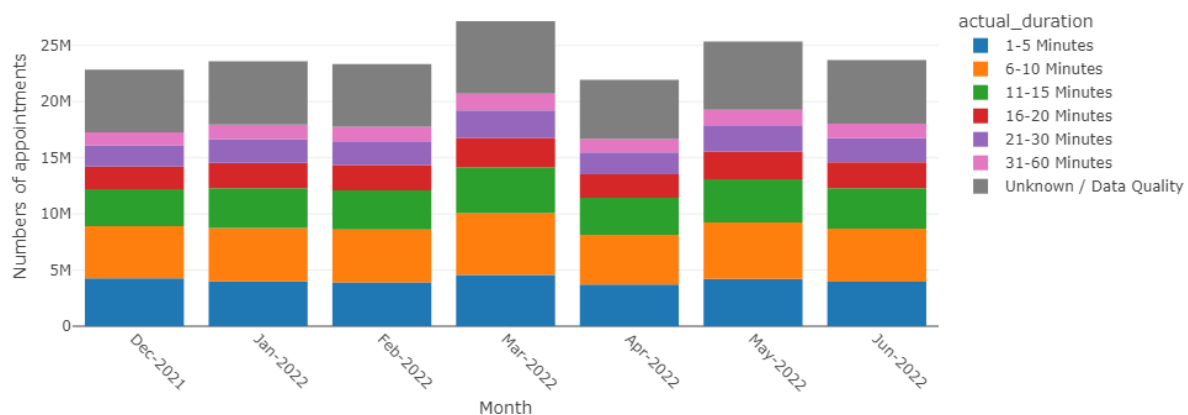
Appointments per waiting time



The proportion of longer waiting time appointments is increasing and it implies that service efficiency may be becoming lower.

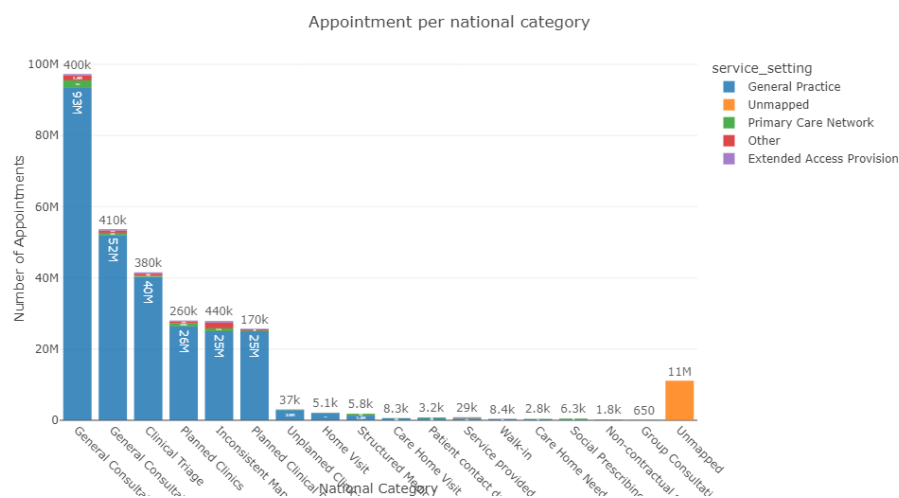
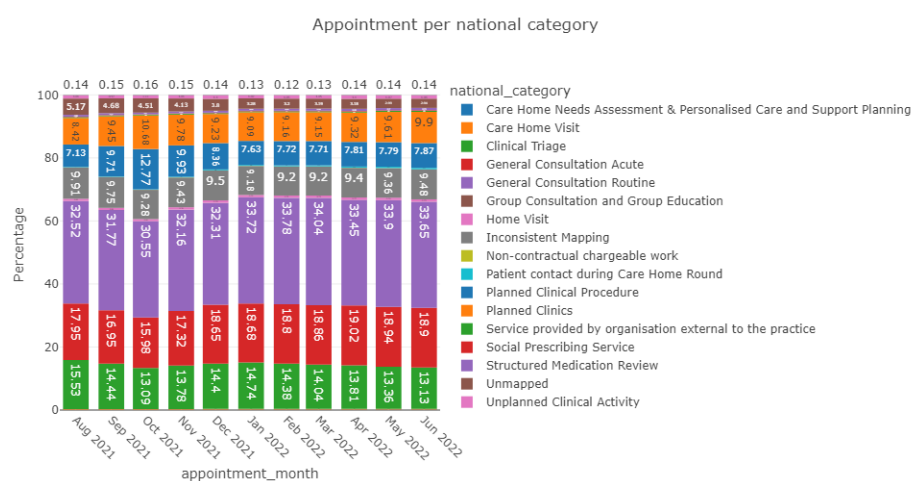
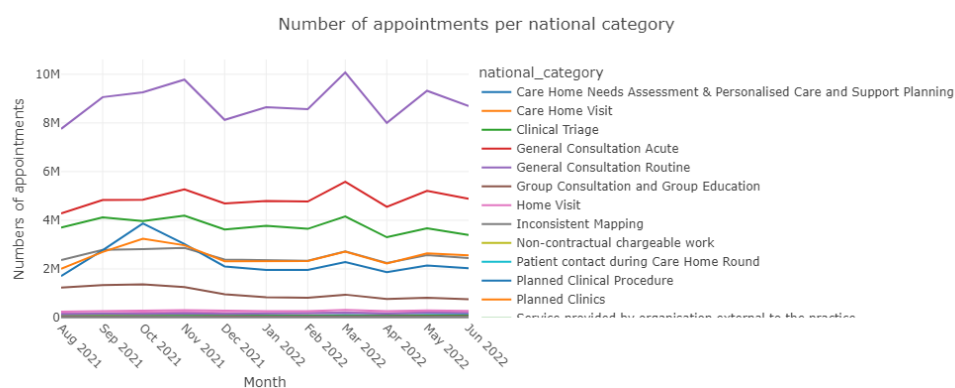
- Actual Duration

Number of appointments per actual duration



There is no huge fluctuation for appointment duration.

## ● National Category & service setting

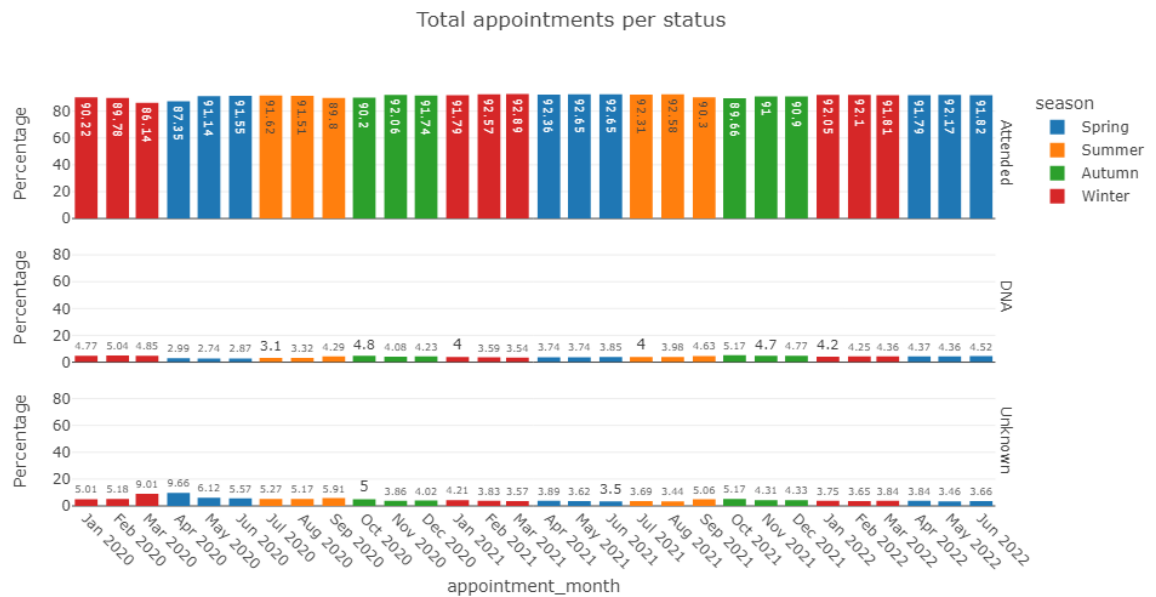


General Consultation Routine is the main national category among all service settings. And the General Consultation Acute is the second.

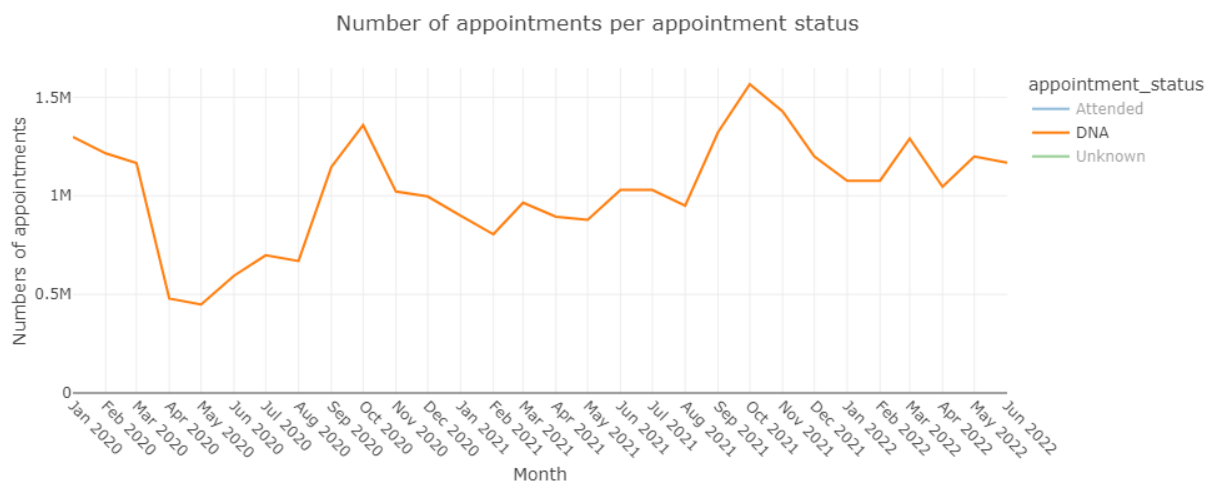
But we can see the planned clinical procedure has been decreased.



## ● Appointment status

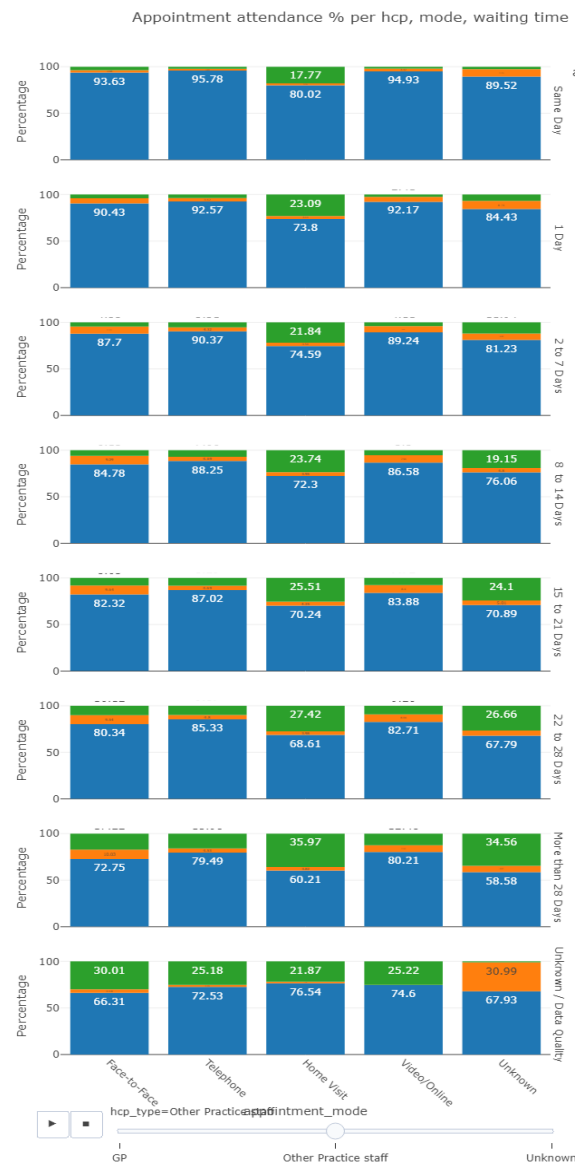
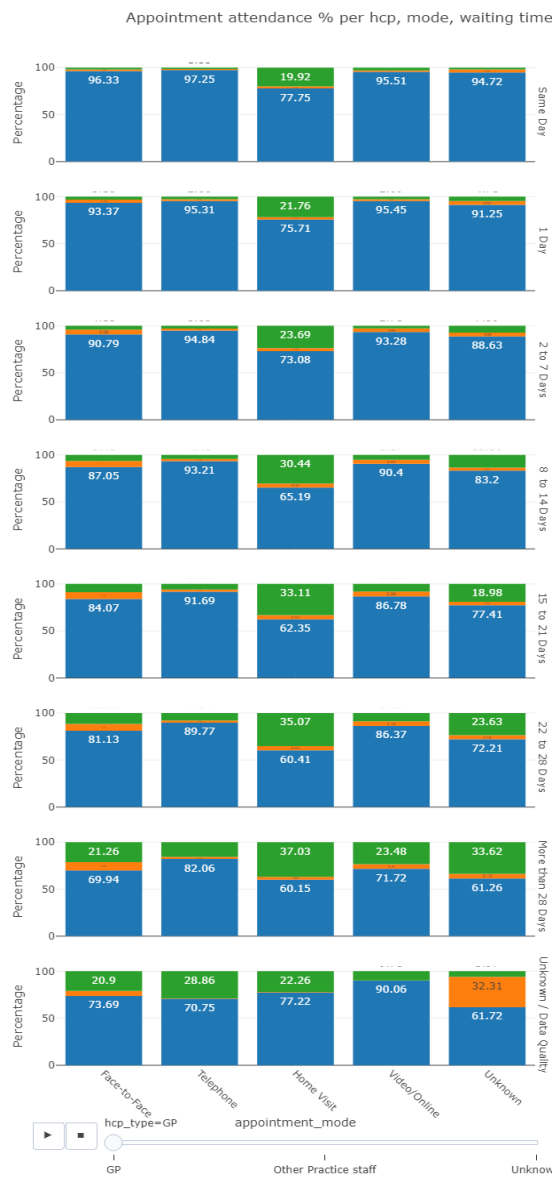


The attendance rate did not have a large fluctuation in the period.



However, the number of unattended appointments is in increasing trend and it has reached 1.5 million in Oct 2021. It caused a lot of resource waste and affected other patients to make appointments.

## ● Appointment per cluster

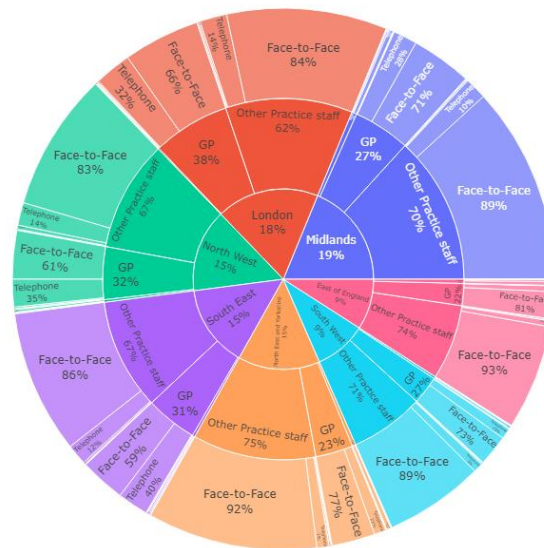


Since the patients were not making the same appointments, we can group the patients into clusters of different pathways, which can understand the specific group of appointments. The above diagrams shows the matrix of divided all appointments by 3 hcp types, 5 appointment modes, and 7 waiting time period.

We can see home visit is the mode with the highest DNA rate for all waiting period, while the the home visit DNA rate of GP appointment is higher than the the home visit DNA rate of Other Practise staff appointment.

It is designed for comparing the specific service and it can also add the region filter for compare the attendance rate in different regions.

DNA Appointment Distribution - region, HCP type, appointment mode, Waiting time



### ● No show appointment

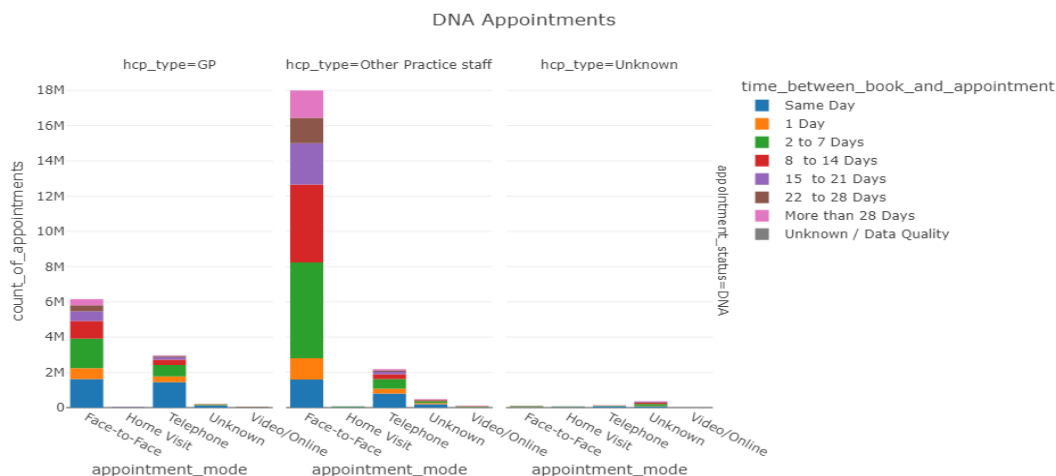
In order to understand the no show problem, we should know more about the patients who did not attend the appointments.

For region, Midland and London account for around 40% of DNA appointments.

For hcp type, there are 67% DNA appointment belongs to appointments of Other Practice Staff.

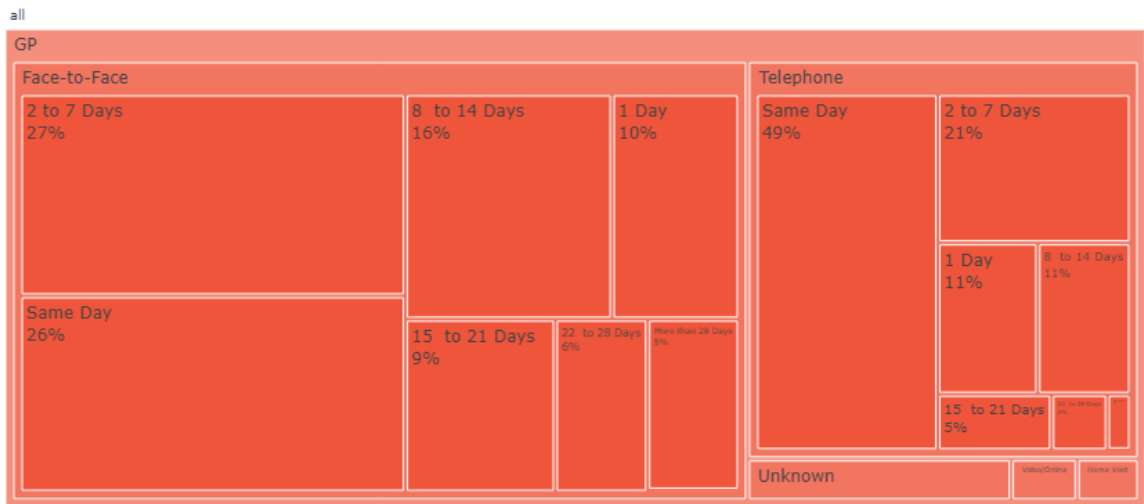
The sunburst chart is an interactive chart that enables users to click and see the detail information of each layer in the jupyter notebook.

For the static report, a bar chart is also good to present the total number of DNA appointments per hcp type, appointment mode and the bar also show the waiting time proportion.



For mode, face-to-face appointments account for the most of DNA appointments. For waiting period, 2-7 days waiting period account for the most of DNA appointments.

DNA appointments per month, appointment mode, waiting time



DNA appointments per month, appointment mode, waiting time



The above treemaps show the proportion of DNA appointments for GP appointments and other practice staff appointments, each rectangle has an area proportional to the total number of DNA appointments (GP/Other) it represents.

For GP appointments, the top 2 areas are same-day & 2-7 days waiting period face-to-face appointments.

For other practice staff appointments, the top 2 areas are 2-7 & 8-14 days of waiting period face-to-face appointments.

By comparing 2 treemaps, we can compare the proportion but cannot interpret the difference of values.

For example, the framed (green line) area proportions look similar in 2 treemaps but it did not imply the number of appointments.

## Part 5. Recommendations

For staffing issues, the actual operation did not meet the expected capacity level. Therefore, NHS should increase the ratio between workforce and population in each service location or department. Besides, it can increase the online/video usage and lower the home visit service.

For the no-show problem, face-to-face appointment is the key contributor. NHS can allow over-booking schedules, and provide more walk-in quota for same-day appointments.