Course 2 Assignment – Diagnostic Analysis using Python
Case Study: National Health Service (NHS)

Project Overview

The NHS, facing avoidable costs due to missed GP appointments, seeks data-driven insights to understand the reasons behind this issue. As part of a data analyst team, we aim to explore the provided data, identify trends, and answer key questions to aid decision-making.

Our initial analysis includes determining the number of locations, service settings, context types, national categories, and appointment statuses in the data. We also establish the date range and service setting with the most appointments during a specific period. Further, we calculate the total appointments and records per month, revealing possible monthly and seasonal trends.

To provide valuable recommendations, we plan to analyse staff capacity, resource utilization, and Twitter trends related to healthcare. This data-informed approach can help the NHS reduce missed appointments, benefiting both finances and society. Our exploration will lay the foundation for deeper investigations and insightful decision-making to address this critical healthcare challenge.

Section 1: Importing and Exploring Data

The dataset contains appointments from 106 locations, showcasing broad coverage across different regions. The top five locations with the highest number of records are NHS North-West London ICB - W2U3Z, NHS Kent and Medway ICB - 91Q, NHS Devon ICB - 15N, NHS Hampshire and Isle of Wight ICB - D9Y0V, and NHS North-East London ICB - A3A8R, indicating areas with higher appointment demand.

Furthermore, there are five service settings, three context types, and 18 national categories, reflecting diverse healthcare services offered. The dataset also includes appointments categorized into three statuses, allowing monitoring of the booking system's efficiency, and identifying reasons for missed appointments.

These insights provide valuable information for resource allocation, service planning, and improving patient care. By leveraging data-driven strategies, the NHS can optimize resource utilization, reduce missed appointments, and enhance overall healthcare service delivery. Further in-depth analysis and visualization will be beneficial for gaining deeper insights and making informed decisions.

Section 2: Analysing the data

Question 1: Between what dates were appointments scheduled?

To ensure data consistency, the date columns were converted to datetime format and inspected. The results are as follows:

- ad DataFrame: From December 1, 2021, to June 30, 2022.
- ar DataFrame: From January 1, 2020, to June 1, 2022.
- nc DataFrame: From August 1, 2021, to June 30, 2022.

Question 2: Which service setting reported the most appointments in North-West London from 1 January to 1 June 2022?

The number of service settings in North-West London from 1 January to 1 June 2002 is **5.**

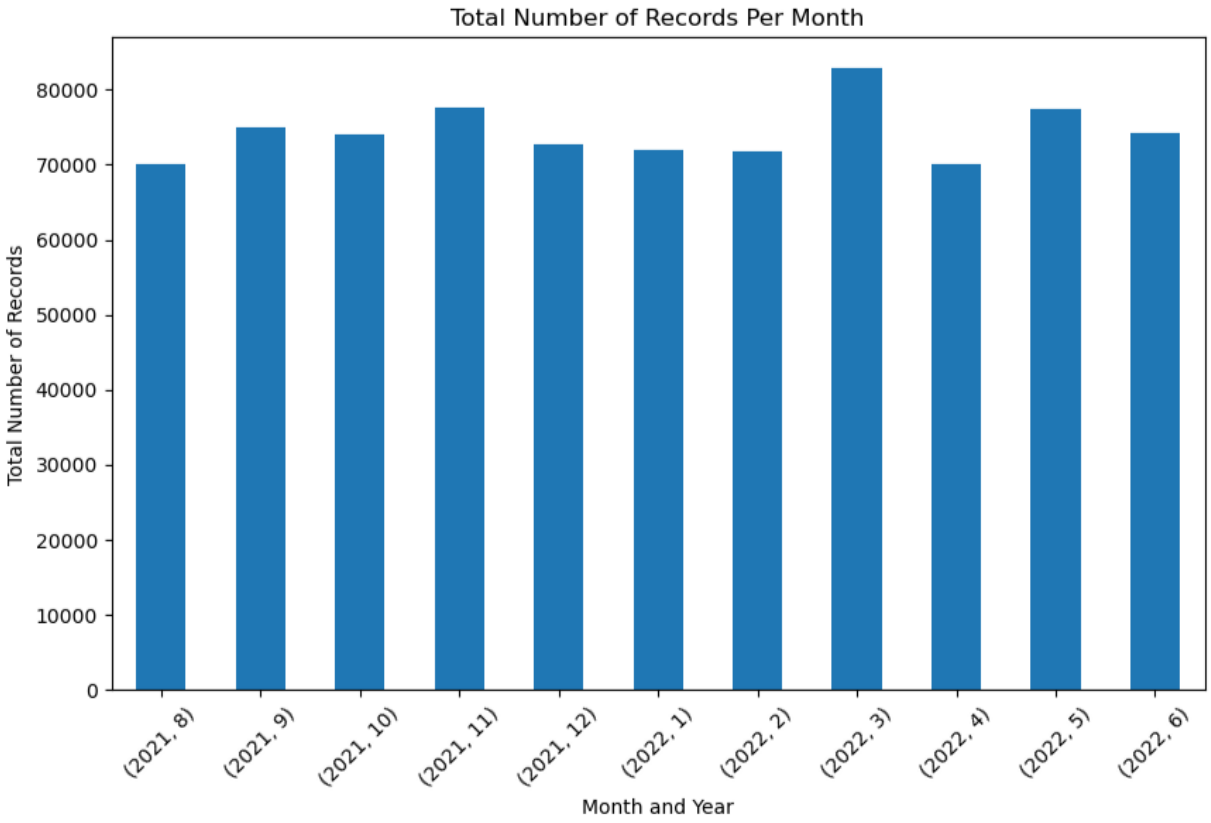
The service setting with the most appointments during this period was **General Practice.**

Question 3: Which month had the highest number of appointments?

The month with the highest number of appointments is **March 2022**.

Question 4: What was the total number of records per month?

The total number of records per month revealed varying appointment patterns. The graph for this is as follows:

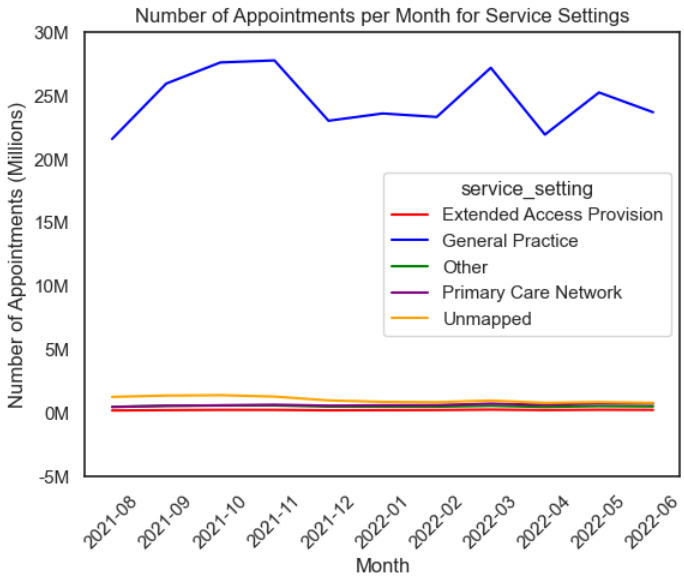


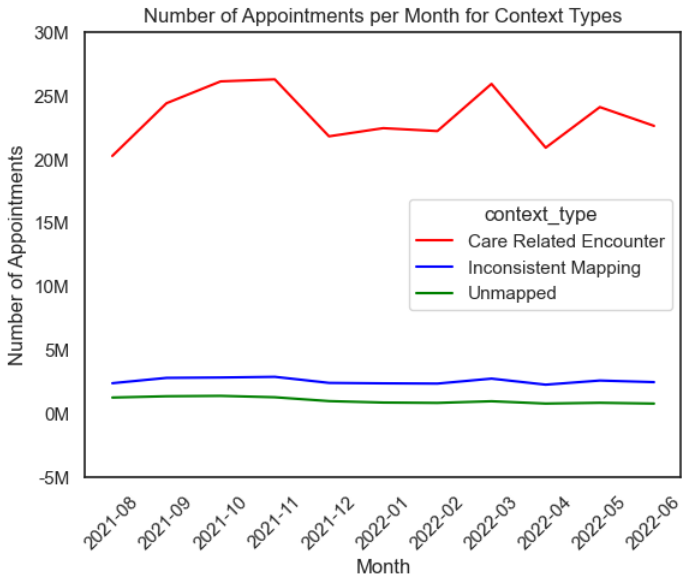
The analysis of the data from the nc DataFrame reveals that the highest number of appointments occurred in March 2022, totalling 82,822 appointments. There was a consistent level of appointments from August to December 2021, but in 2022, March stood out with significantly higher appointment numbers.

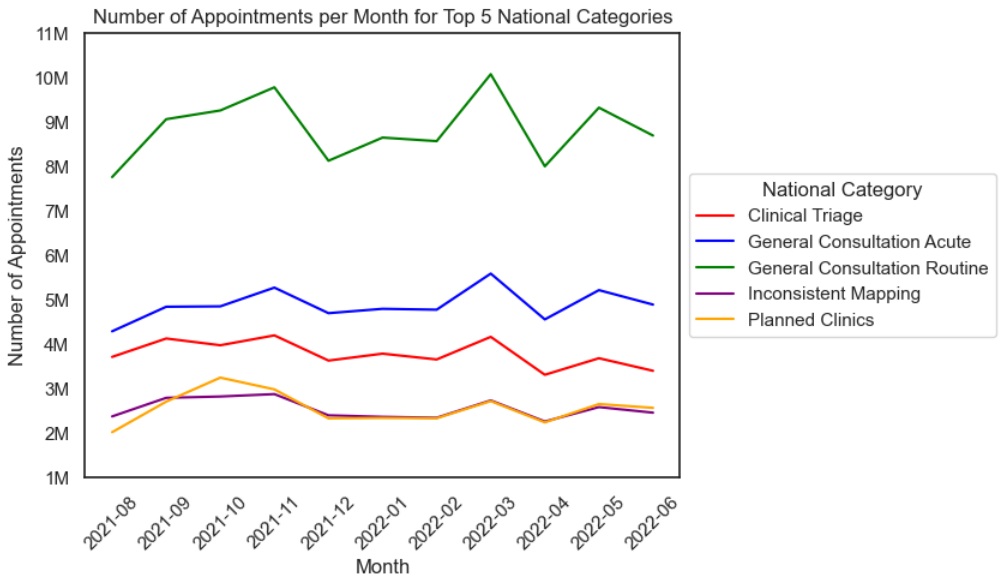
Section 3: Visualising and identifying initial trends

Objective 1: Create three visualisations indicating the number of appointments per month for service settings, context types, and national categories.

- The nc_ss graph displays the number of appointments for different service settings from August 2021 to June 2022. General practice had the highest number of appointments, exceeding 27 million, followed by Other and Primary Care Network.
- The nc_ct table displays the number of appointments for different context types from August 2021 to June 2022. Care Related Encounter had the highest number of appointments, exceeding 25 million, followed by Inconsistent Mapping and Unmapped.
- The nc_nc table displays the number of appointments for different national category types from August 2021 to June 2022. For the ease of visualisation and analysis, only the top 5 national categories are displayed. General Consultation Routine had the highest number of appointments, exceeding 10 million in March 2022. In order of number of appointments this was generally followed month to month by General Consultation and Clinical Triage in 2nd and 3rd place respectively.



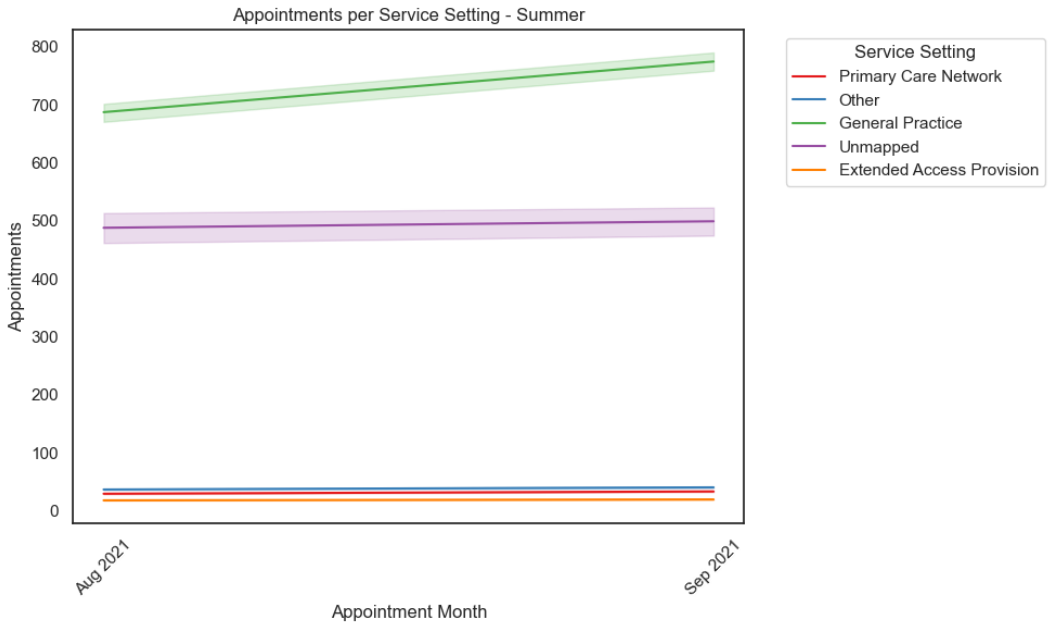


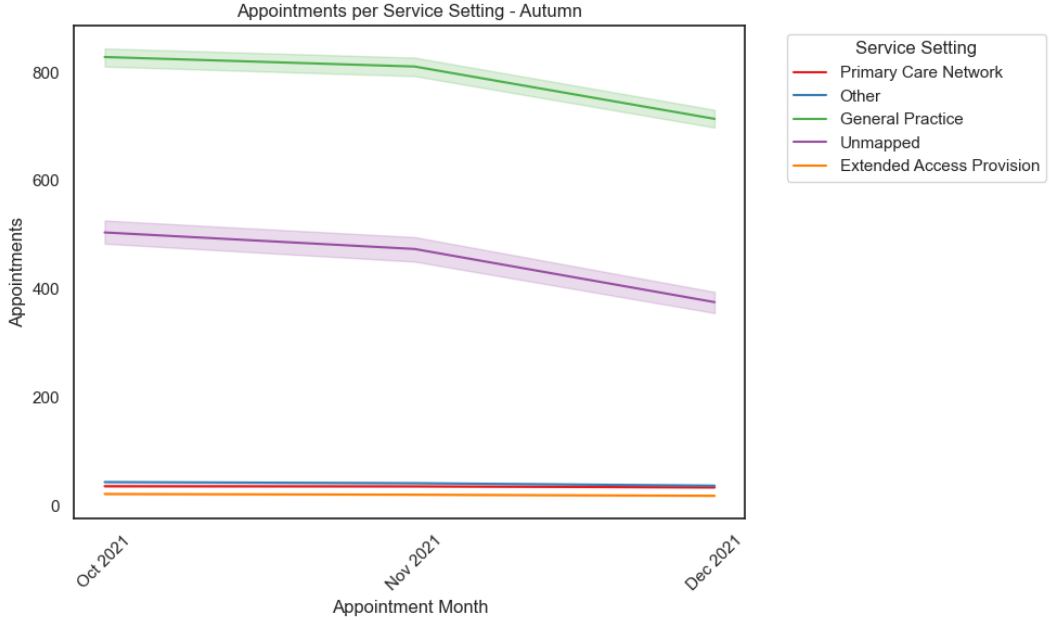


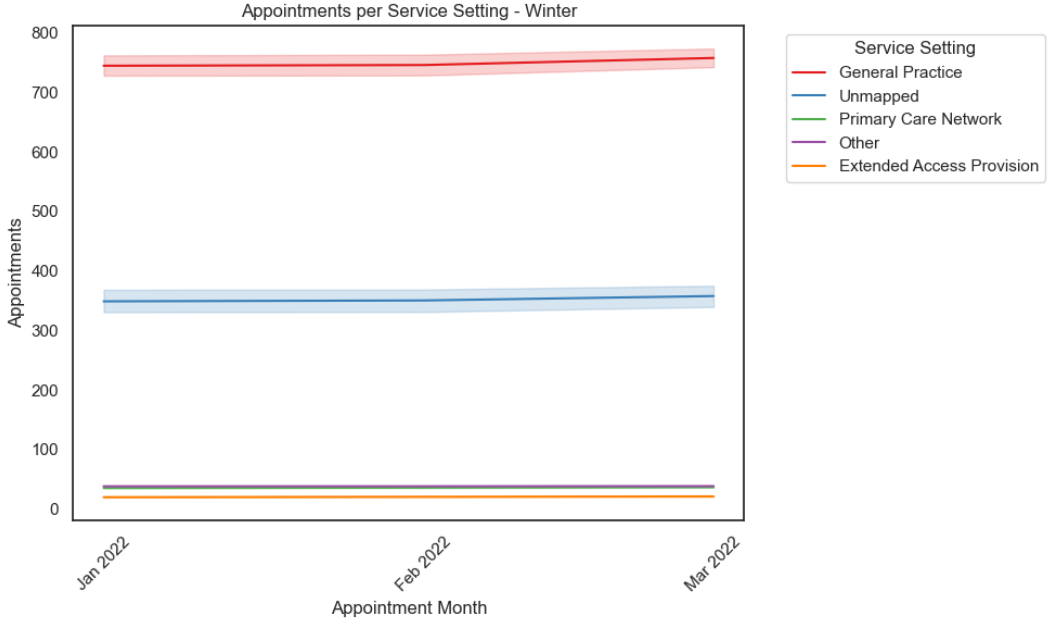
Objective 2: Create four visualisations indicating the number of appointments for service setting per season. The seasons are summer (August 2021), autumn (October 2021), winter (January 2022), and spring (April 2022).

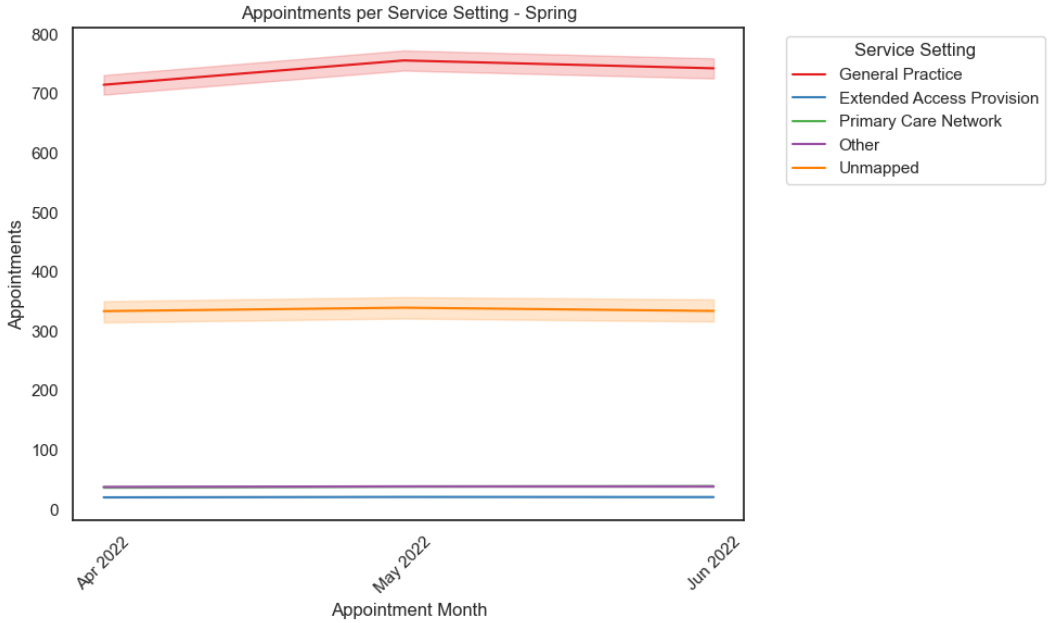
Throughout all seasons, General Practice consistently had the highest number of appointments, indicating its crucial role in healthcare services. The number of appointments for Extended Access Provision and Primary Care Network showed noticeable fluctuations across different seasons. Appointments peaked during Winter (January, February, and March 2022) and started to decline in Spring (April, May, and June 2022).

The data helps the NHS identify seasonal trends and patterns in appointments for various service settings. Understanding these trends can enable better resource allocation, staff planning, and scheduling to accommodate the varying demand during different seasons. Furthermore, it could lead to targeted interventions and strategies to encourage patient attendance during specific months, thereby optimizing healthcare services and reducing missed appointments.


Appointments per Service Setting - Summer
Appointments
800
700
600
500
400
300
200
100
0
Aug 2021
Sep 2021
Appointment Month
Service Setting
Primary Care Network
Other
General Practice
Unmapped
Extended Access Provision


Appointments per Service Setting - Autumn
Appointments
800
600
400
200
0
Oct 2021
Nov 2021
Dec 2021
Appointment Month
Service Setting
Primary Care Network
Other
General Practice
Unmapped
Extended Access Provision

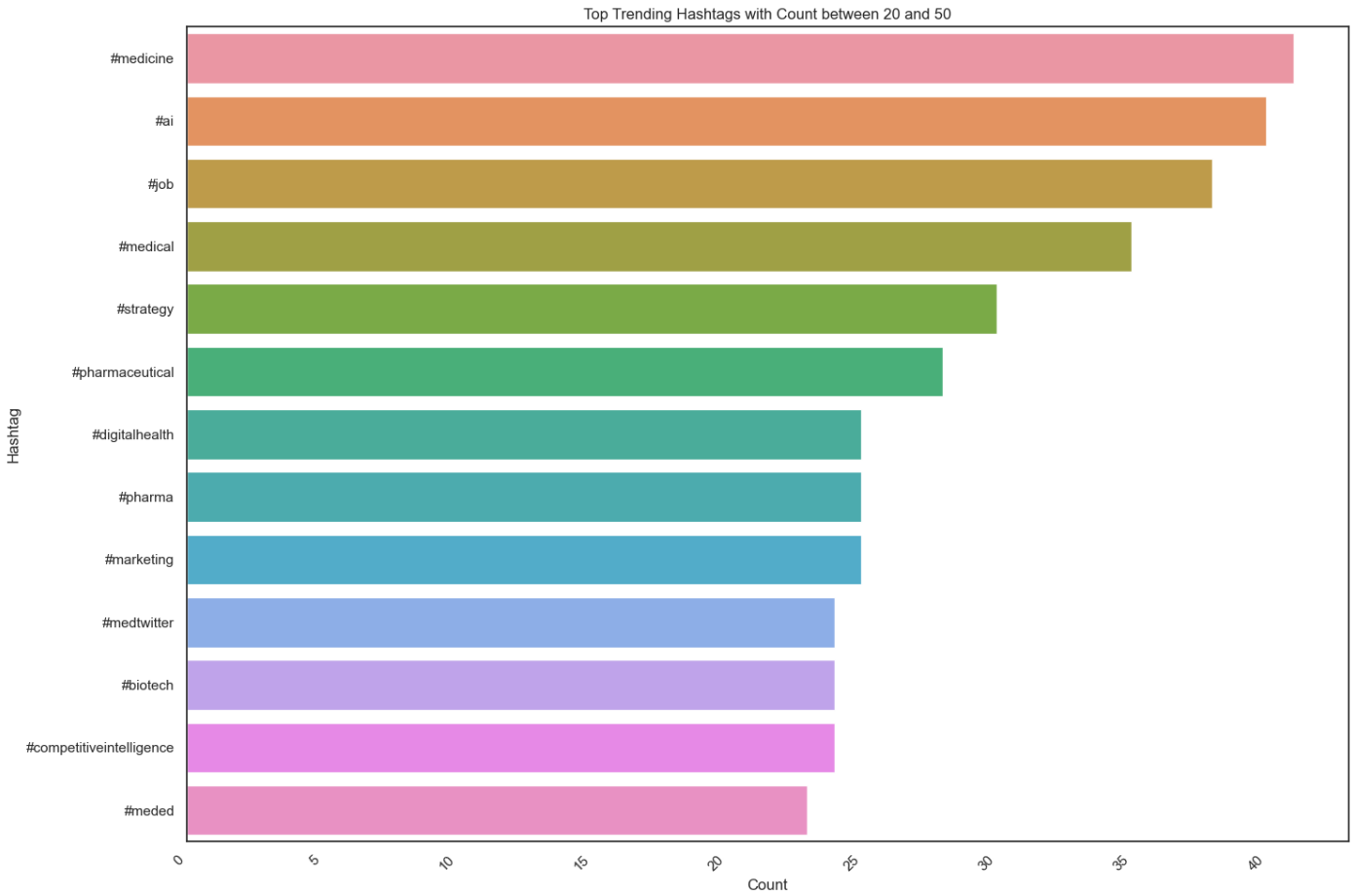

Appointments per Service Setting - Winter
Appointments
800
700
600
500
400
300
200
100
0
Jan 2022
Feb 2022
Mar 2022
Appointment Month
Service Setting
General Practice
Unmapped
Primary Care Network
Other
Extended Access Provision


Appointments per Service Setting - Spring
Appointments
800
700
600
500
400
300
200
100
0
Apr 2022
May 2022
Jun 2022
Appointment Month
Service Setting
General Practice
Extended Access Provision
Primary Care Network
Other
Unmapped

Section 4: Analysing NHS-related Twitter data

The objective of this analysis is to identify and review the top trending hashtags related to healthcare in the UK using the tweets.csv data received from the NHS. After preparing the data and exploring the hashtags, a bar plot is created to visualize the most frequently used hashtags.

Overrepresented hashtags are removed, and the final plot displays the remaining popular hashtags. Overrepresented hashtags in this case were those with a count over 50. The graph below shows the top trending hashtags with a count between 20 and 50. Utilizing tweets can provide valuable feedback to stakeholders and the NHS. The insights from the analysis can help in understanding the healthcare-related discussions on Twitter and can aid in making data-driven decisions to improve healthcare services.

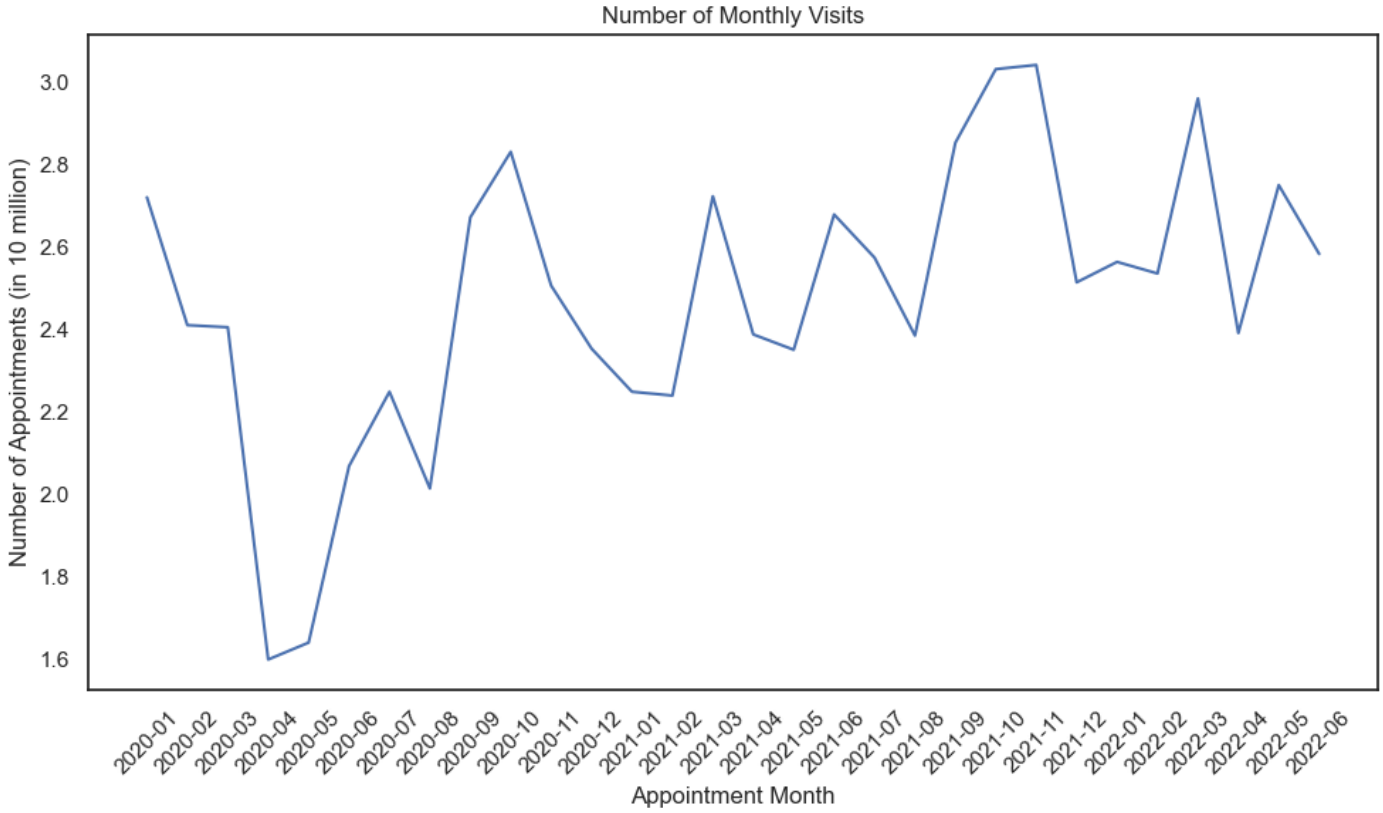


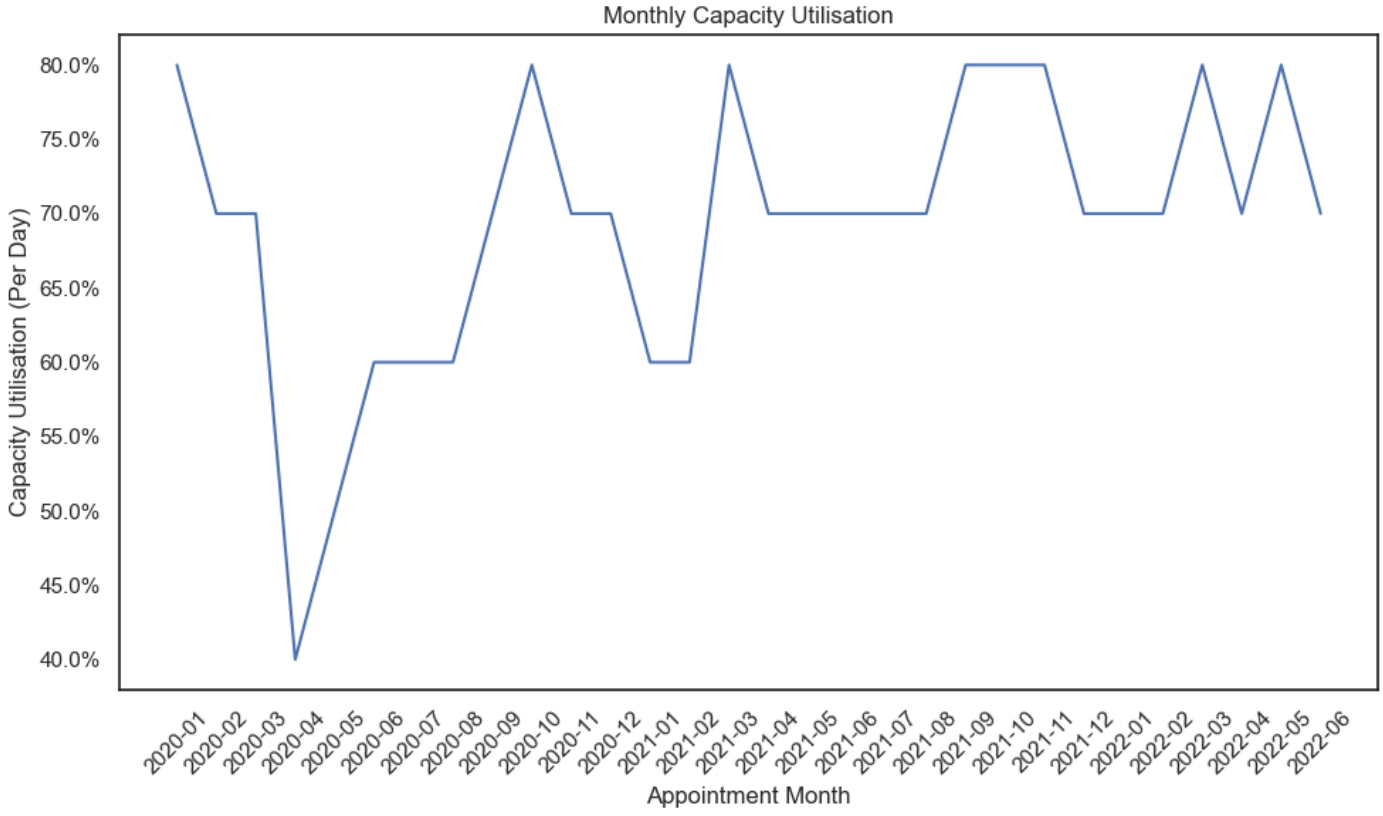
The top trending hashtags related to healthcare in the UK are #medicine (41), #ai (40), and #job (38), among others. From the top trending hashtags in the healthcare domain on Twitter, the NHS can gain valuable insights to inform their initiatives and strategies. The high count of #medicine and #medical hashtags suggest a strong interest in medical topics among Twitter users, indicating a potential focus for public health awareness campaigns or information dissemination. The presence of hashtags like #ai and #digitalhealth highlights the growing importance of technology and artificial intelligence in healthcare, encouraging the NHS to explore innovative solutions to improve patient care and efficiency. Additionally, hashtags like #job and #marketing could indicate trends in healthcare job opportunities and workforce demands, aiding the NHS in addressing staffing needs and recruitment strategies. Analysing these hashtags helps the NHS stay attuned to public interests, challenges, and opportunities in the healthcare landscape.

Section 5: Making recommendations

Question 1: Should the NHS start looking at increasing staff levels?

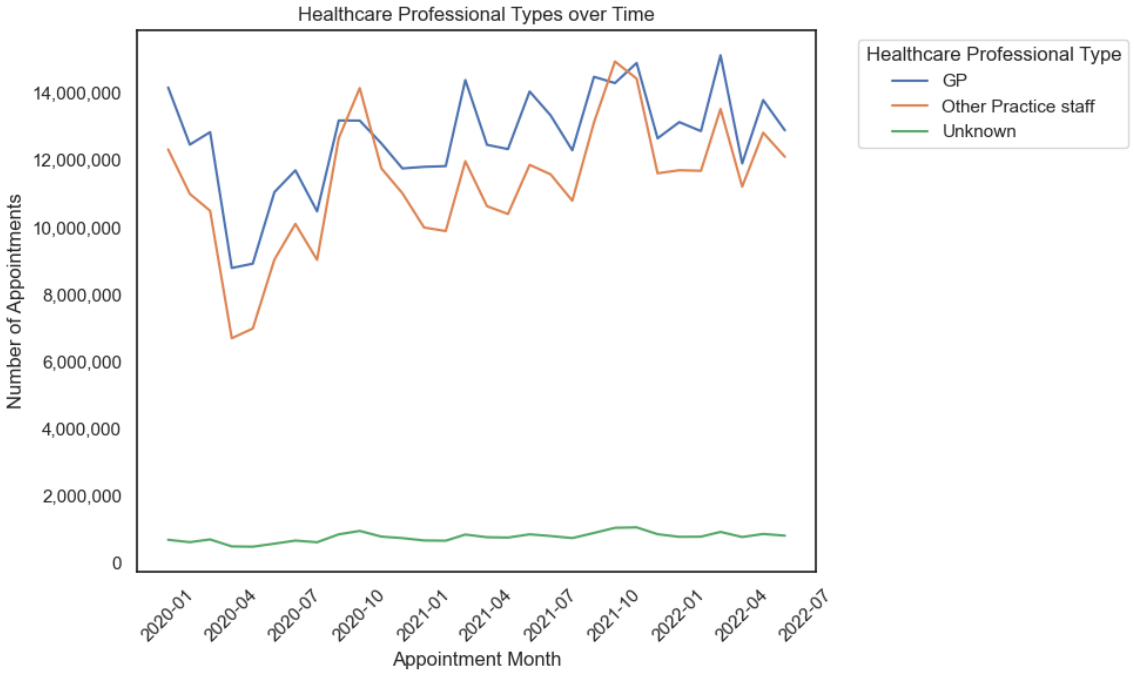
Based on the graphs below, the NHS should consider increasing staff levels. Monthly appointment counts have consistently risen, with many months nearing the NHS's daily capacity. The average month utilisation, ranging from 40% to 80%, indicates the need for additional resources to manage patient demands effectively.





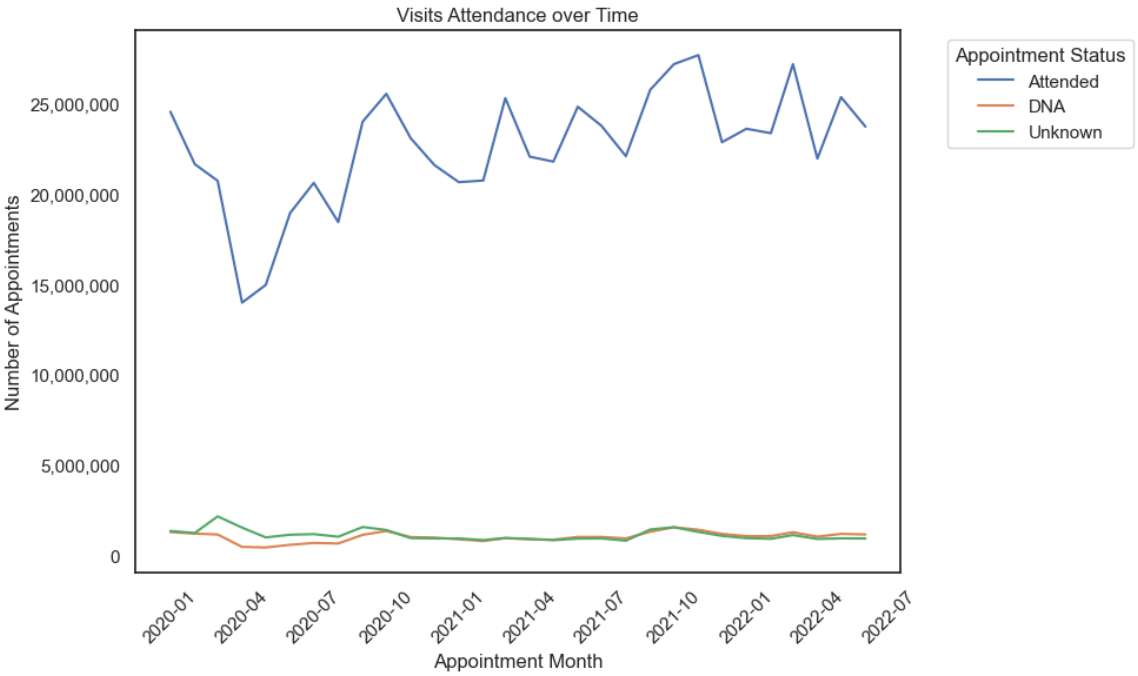
Question 2: How do the healthcare professional types differ over time?

The line plot illustrates the trends in different healthcare professionals types over time. GP and Other Practice Staff appointments show a steady increase from April 2020 to July 2022, while the Unknown category remains relatively constant at around 1,000,000 appointments.



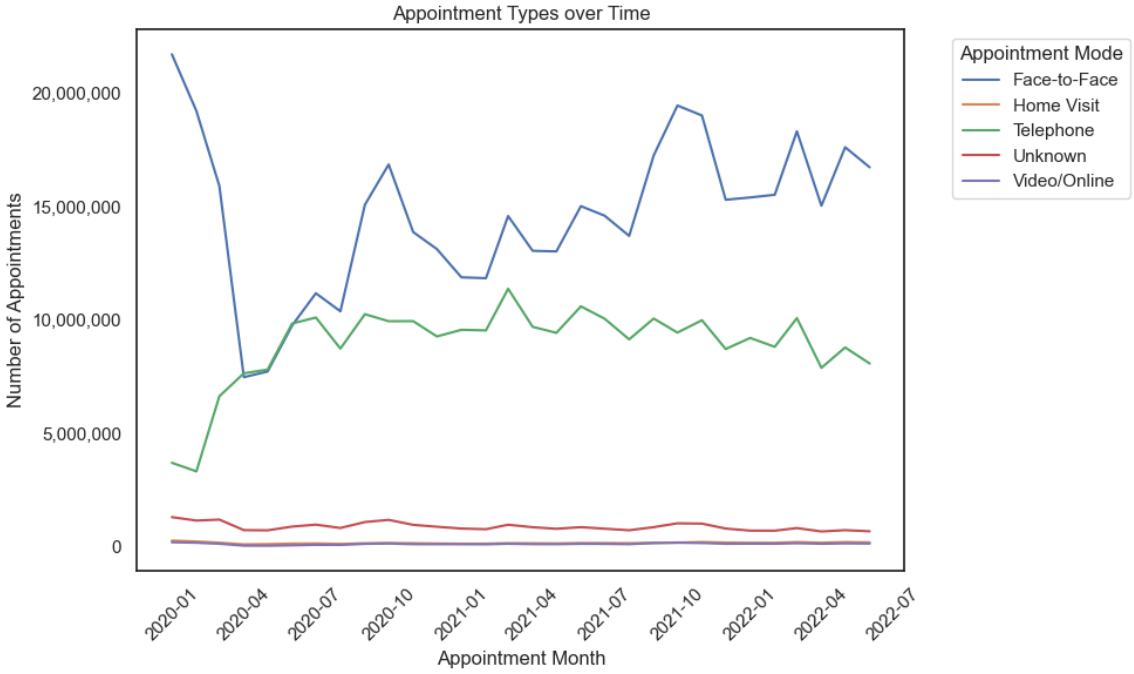
Question 3: Are there significant changes in whether or not visits are attended?

The line plot shows that DNA and Unknown appointments haven't varied much over time and stayed consistently below 3,000,000 appointments. However, attended appointments initially dropped off January 202 to April 2020 which could have been due to COVID restrictions being in place. After April 2020, the number of Attended appointments has steadily increased.



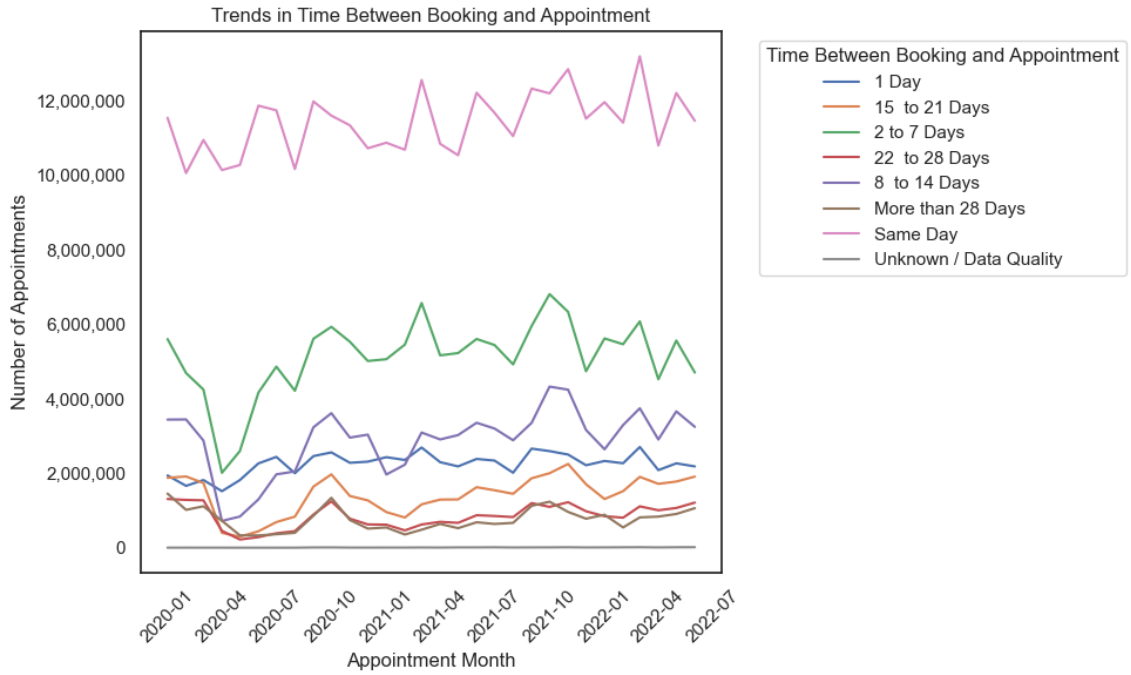
Question 4: Are there changes in terms of appointment type and the busiest months?

The graph shows how appointment types have evolved over time from January 2020 to June 2022. Face-to-Face appointments have consistently been the most common except in April 2020 where Home Visits were slightly higher. There was a big drop off in Face-to-Face appointments but an increase in Home Visits in January 2020 most likely due to the spread of COVID-19.



Question 5: Are there any trends between booking an appointment?

The graph shows how the time between booking and appointment has changed from January 2020. The number of Same Day appointments has stayed consistently the highest, fluctuating around 11 million appointments. This is followed by the 2 to 7 days bracket which has fluctuated around the 5 million appointment mark



Question 6: How do the various service settings compare?

The table below shows the comparison of service settings by number of appointments. There is also a graph showing the data but having excluded GP visits as these were by far the highest and made the graph hard to read. GP visits consistently have the highest appointment volumes, surpassing all other service settings. This highlights the central role of the General Practice in healthcare services, being the most sought-after option for patients. Among the other service settings, Primary Care Network, Extended Access Provision, and Other exhibit fluctuations in appointments, indicating varying demand patterns or service availability. The Unmapped category's high appointment volumes raise questions about its nature and relevance, warranting further investigation. By understanding these trends and acknowledging the significant number of GP visits, the NHS can strategically allocate resources and enhance service planning to meet patient demands effectively. Further exploration of the Unmapped category remains important to ascertain its role and optimize healthcare services.