UTILISATION & CAPACITY

A high-level analysis of the

NHS

By
Will Burton

Business Context

Through this report, the NHS aims to address issues of capacity and utilisation to keep up with growing demand.

This report will consider two primary business questions as suggested by the NHS:
- **Has there been adequate staff and capacity in the networks?**
- **What was the actual utilisation of resources?**

Further exploration of the business problem can be found in a Five Why's Diagram (Appendix 1**A**) and a SWOT analysis (Appendix 1**B**).

Analytical Approach

I performed the initial administration by extracting the data, backing it up and reviewing the metadata. There are two main insights from this:

- Firstly the data has been already been cleaned.
- Secondly the metadata highlights that the data was not originally designed for analysis, lacking national standards for data entry and resulting in widespread data quality issues. This poses a significant risk to analysis, and there is a need for efforts to establish standard data practices in order to avoid losing hundreds of millions of appointments.

Please refer to (Appendix 2**A**) to see potential exploration for joining datasets which was ultimately not required to answer the question.

Before the initial work begins I must prepare my workstation.
- Pandas and Numpy give me the resources I need to import, sense check and manipulate data
- Seaborn and Matplotlib give me to tools to visualise the data
- Setting the plot, style, layout and palette of visualisations at the start help to ensure consistency and accessibility.

Preparing my workstation

```python
# Import the necessary libraries
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

plt.figure(figsize=(16, 9)) # Set the figure size

sns.set_style("white") # Set the plot style

sns.set_palette('colorblind') # Set the color palette style

plt.tight_layout() # Adjust layout for better spacing

import warnings # Optional - Ignore warnings
warnings.filterwarnings('ignore')
```

```
<Figure size 1600x900 with 0 Axes>
```

The structure of the Jupyter report and the visualisations:

1) Import, sense check and descriptive statistics - Please refer to the Jupyter Notebook for import specifics and (Appendix 2**B** for a standard procedure)
2) Refined data exploration to fill in any knowledge gaps prior to addressing NHS questions
3) Directly addressing NHS questions provided to us
4) Time based analysis of the data
5) Twitter analysis
6) Final visualisations, further questions from the NHS and any additional insights.

Visualisation and insights

It is important to present the entire narrative structure of the process here in the report and subsequent appendices. As such, it is below with the main insights. However only some sections are present in the presentation so they have been highlighted in green:
- I address capacity using the maximum daily capacity metric provided by the NHS (Appendix 3**A**) and explore the date range the appointments take place (Appendix 3**B**).

    o Capacity on average is around 65% but this focuses on a specifically tumultuous time for the NHS.

- I address utilisation of resources thematically by considering:
    o Where are the appointments geographically? (Appendix 3**C**)
        ▪ The analysis reveals effective and appropriate resource utilisation across regions in the UK. However, all levels of granularity show us that spread of appointments is varied and emphasises the relative nature of capacity.

    o Where are the appointments within the NHS? (Appendix 3**D**)
        ▪ 91% of appointments are in general practice

- Analysing data logarithmically reveals a decline in unmapped appointments and a rise in primary care network appointments, hinting at potential government investment.

o What categories are these appointments in ? (Appendix 3**E**)
  - Showed us confirmation of General practice's dominance. However, it also showed us that 13% of appointments suffer from data quality issues.

o When are the appointments taking place? (Appendix 3**F**)
  - Using the metric of under 14 days for the target time before book and appointment showed us that 88% of appointments do reach that target.
  - However the 12% that don't equal 91 million appointments

o Who are the appointment's patients? (Appendix 3**G**)
  - While 91% of patients attended the 4% that didn't cost the NHS around £927 million from January 2020 to June 2022.

o Who are the appointments staff? (Appendix 3**H**)
  - GPs handle 51% of appointments, while Other Practice Staff manage 46%.

o How are these appointments delivered? (Appendix 3**I**)
  - Face to face appointments are mostly made by other practice staff 58%
  - GP's make the most telephone appointments however.

o How are these appointments delivered pt 2? (Appendix 3**J**)
  - Data quality issues affect almost a quarter of all appointments
  - The suggested metric from the BMA shows us only 52% of appointments reach the target duration.
  - 25% percent do not

o Then the requested Twitter analysis (Appendix 3**K**)
  - The Twitter data provided is not directly useful but there is good potential.

Key recommendations:

The NHS capacity debate offers two main paths: reorganise current resources or expand and upgrade capacity.

Capacity

- Capacity on average is around 65% but this focuses on a specifically tumultuous time for the NHS and more data needed to make a more detailed assessment:
    - *Specifically sum of appointments over a larger date range including pandemics to make it fairer*
    - *Maximum capacity limits at several geographical levels of granularity and maximum capacity limits for national categories too.*

- Regional distribution of appointments varies and aligns directly with relative population.
    - *My recommendation is to allocate resources to icb_level to address those relative demands and not rely on one national capacity metric.*

Ultimately to meet population demands, increasing capacity is essential. However, as seen in this report, significant resource restructuring is necessary for more efficient utilization.

General Practice Utilisation

- 91% of appointments are in general practice.
    - Most patients are seen below the recommended 14 day target, however 91 million appointments aren't
    - GP's are responsible for 51% of appointments, Other Practice Staff 46%.
    - 52% of appointments meet the BMA's standard of sub 15 minute duration.
    - Most appointments are attended but still around 31 million aren't.

- In order to make general practice more efficient I recommend the following.
    - *Free up gp's time by decreasing inefficient face to face and increasing access to alternative means of appointments for example telephones, or video/online.*
    - *Allow and encourage other practice staff to continue taking on appropriate appointments where gp's aren't needed.*
    - *Continue to encourage general practice to work through the centralised administration of Primary Care Networks. By pooling resources it allows general practice to work more efficiently at scale.*
    - *Encourage staff to remain within an appropriate time frame for seeing appointments.*

Appointment attendance

- While 91% of patients attend their appointments, the 4% that didn't could be costing the NHS around £927 million in the data period.
    - *The move towards alternative health care professionals, administration and appointment types also benefits patients who missed appointments.*

Twitter

- *Twitter data wasn't useful directly but there is significant potential for insights if used appropriate.*

Appendix 1 - Background Context

1. **A** – *Let's apply the Five Why's approach to problem solving*

Why is the NHS not performing well enough to meet demands?
- There are concerns around the NHS capacity for patients

Why are there are concerns around the NHS capacity for patients?
- There are concerns that there are not enough staff or capacity to meet the demand of patients

Why is there not enough staff or capacity to meet demand for patients
- Because there are not enough resources or the resources are not being used efficiently

Why is there an issue with resourcing?
- Because there isn't enough high level analysis of the available data

Why isn't there enough high level analysis?
- Our team have been brought in to directly address this.

1. **B** – *SWOT analysis*

Strengths:	**Established Healthcare Infrastructure:** The NHS has a trusted, well-established healthcare infrastructure, providing a solid foundation for delivering healthcare services.	**Skilled Workforce:** While there are staffing issues. The NHS has a skilled and experienced workforce.
Weaknesses:	**Lack of High-Level Data Analysis:** Absence of reliable analysis limits the ability to make data driven decisions for resource allocation and optimisation.	**Insufficient Staffing:** The NHS faces routine staff unrest with strikes and increasing turnover rates.
Opportunities:	**Utilisation of External Data (X - Twitter):** Exploring external data sources, such as Twitter (rebranded as X), presents an opportunity to gather additional insights and enhance decision-making.	**Optimising with data driven insights rather increasing staff**: Insights from data can potentially provide opportunities for efficiencies through resources reallocation or technology. Rather than solely relying on expansion.
Threats:	**Risks of Not Meeting Patient Demand:** Failing to address capacity and utilisation issues leads to a probable increase of all current negatives.	**External Pressures on Healthcare System:** Changes in government, pandemics, cost of living all threaten the NHS on a daily basis.

Appendix 2 – Analytical Approach

2. **A** *– Joining*

While joining data sets is possible by aggregating up daily data to monthly it is complex and inefficient. Rather in the future, data provided can be at an icb_level and reporting can be standardised to monthly so as to avoid any issue and allow for more direct comparisons of relevant metrics and KPI's.

I can make two joined tables:
- Joining *actual_duration* and *national_catagories* give me sub_icb level daily data from DEC 2021 - JUNE 2022:
  - Appointment_date (primary key)
  - icb_ons_code
  - sub_icb_location_name
  - service_setting
  - context_type
  - national_category
  - count_of_appointments
  - region_ons_code
  - actual_duration
- Joining all three datasets *national_catagories*, *actual_duration* and *appointments_regional* gives me icb level monthly data from DEC 2021 - JUNE 2022:
  - Icb_ons_code (primary key)
  - appointment_month
  - appointment_status
  - hcp_type
  - appointment_mode
  - time_between_book_and_appointment
  - actual_duration
  - service_setting
  - context_type
  - national_category
  - count_of_appointments

There are limitations and necessary changes that need to be addressed in order to facilitate these joins:
- -Firstly, the short data time frame of actual_duration limits our joined data sets for both daily and monthly to DEC 2021 - JUNE 2022. However, this is the reality and will have to be accepted.
- Second, in order for national_catagories and actual_duration to be joined to the monthly data of appointments_regional we will have to aggregate the data.
  - This will be done via averages of the quantitative data
  - Either the daily data can be aggregated up to monthly or the monthly data can be aggregated down to daily. Both options are valid.
  - Duplicates will potentially be an issue so a focus on avoiding a many to many relationship is essential.

Ultimately joining the data, while interesting, is not necessary and as such, I wanted to focus on efficient actionable insights.

2. **B** – *Import and sense checking routine*

I would normally consider renaming the columns for ease of use as standard practice in the importing routine. However I retain the original column names for two reasons: familiarity with NHS naming conventions among stakeholders and ease of code replication for future or historical data.

First I prepared by work station in Jupyter notebooks setting the crucial initial parameters:

- *import pandas as pd*
- *import numpy as np*
- *import seaborn as sns*
- *import matplotlib.pyplot as plt*

- *plt.figure(figsize=(16, 9))* # Set the figure size

- *sns.set_style("white")* # Set the plot style

- *plt.tight_layout()* # Adjust layout for better spacing

- *import warnings* # Optional - Ignore warnings
- *warnings.filterwarnings('ignore')*

Import and sense check the data. Below is a summary of this process. df = dataframe (of your choice)

- *df = pd.read_csv('csv.csv')* # Import and sense-check 'dataframe as df. print(df.shape)*

- *df.head()* # view the first 5 rows of the DataFrame

- *df.tail()* # view the last 5 rows of the DataFrame

Missing Values checklist:
- *df.isna().sum()* # Determine the sum of missing values
  - process as appropriate

- *df.['column'].isnull().sum()* # number of nulls in one column

Metadata checklist:
- *print(ad.dtypes)* # We start by discovering the data types
  - Change as appropriate
- *df.info()* # Print a report listing all column names, number of null values & the datatype in the column.
  - Useful for basic final review of previous steps and to check that conversion was success.

Descriptive statistics checklist:
- *df.describe().round(2)* # describe the numerical data with basic statistics but limit the decimals
- *ad.describe(include = ['O'])* # describe the descriptive data with basic statistics

- # any necessary graphs to show visualisations
  - E.g. df['column'].value_counts().plot(kind='bar'

Appendix 3 – *Visualisations and Insights* (Charts created for this project are lower resolution images in order to keep file size down, please refer to the Jupyter Notebook for full high resolution images if needed).

3. **A** – *Capacity analysis*

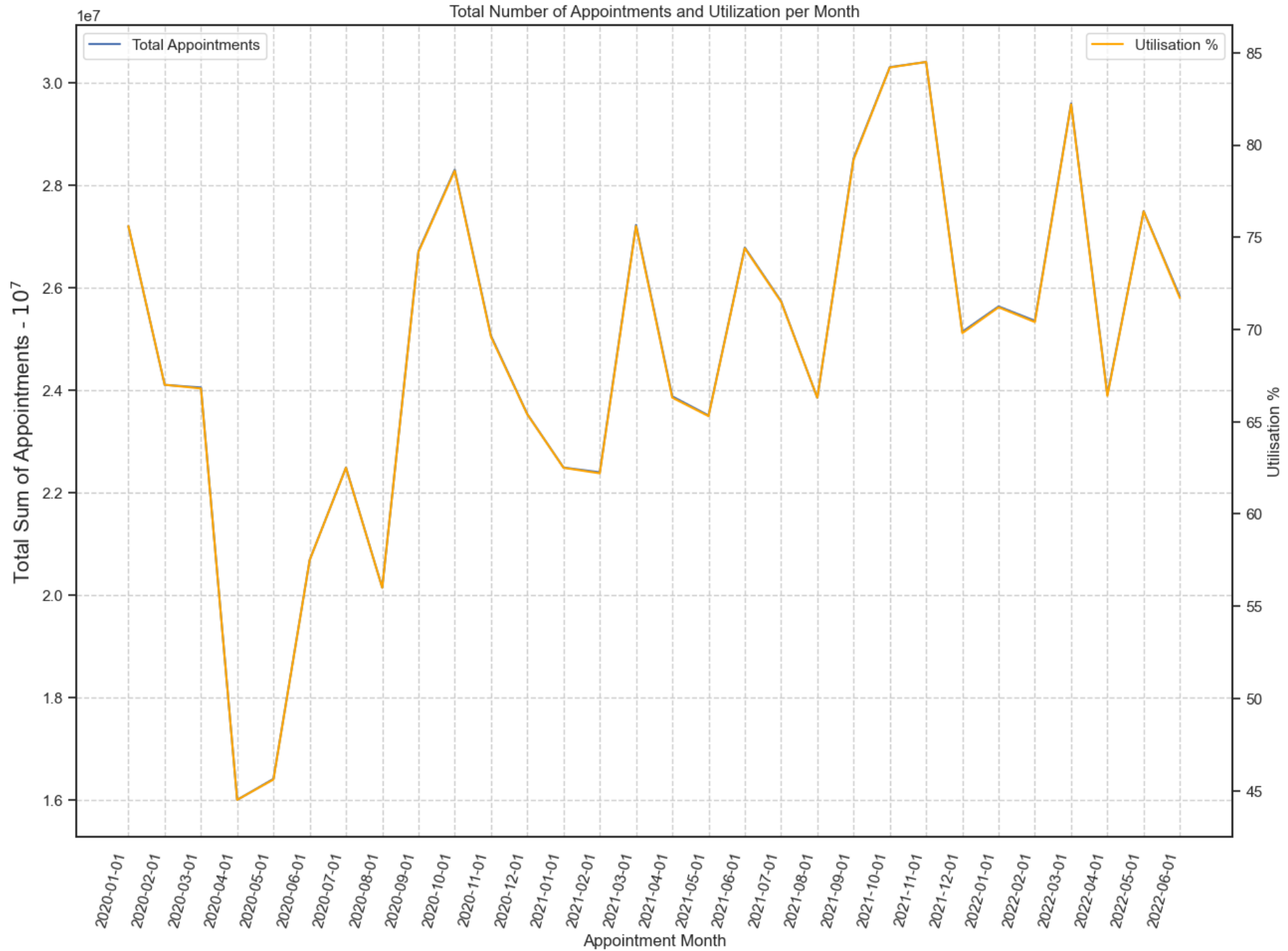Capacity for the NHS across the provided date range is typically at 65%, but that doesn't tell us the whole story
Capacity is a broad concept but it can be defined as - "*the capability, in terms of available resources, that the NHS has to deliver services. It is wrong to think of capacity solely in terms of the number of beds available in acute hospitals. Capacity covers all the equipment (including beds for inpatient services) and the range of appropriately qualified staff that are needed to provide care across all health and care settings.*"(NHS Providers*, (n.d.)* [1])

Using the appointments regional database, I created a new dataframe to aggregate the total number of appointments per month and calculated the utilisation of services based on the NHS guideline of a maximum daily capacity of 1,200,000 appointments.

The utilisation is calculated by dividing the monthly appointments by the product of the assumed average days in a month (30) and the daily capacity, and then multiplying the result by 100 to express it as a percentage. This analysis allows for a quick assessment of how efficiently the appointment system is operating in relation to the recommended daily capacity.

Let's overlay the sum appointments by month with the capacity graph to see a perfect overlap to show that the NHS coped well over the months, never once reaching anything beyond 85% capacity suggesting that in a technical sense, there was adequate capacity and therefore staff for this dataframe's date range.

[1] nhsproviders.org. (n.d.). *Mapping the NHS winter*. [online] Available at: https://nhsproviders.org/mapping-the-nhs-winter/capacity [Accessed 3 Dec. 2023].

Total Number of Appointments and Utilization per Month

* To find the average utilisation I divided the sum of utilisation values by the number of entries.

65% average capacity of course this also suggests that, on average, a further 35% of the NHS resources are not being fully utilised but it is important to remember that there has to be some contingency in place for seasonal variations and pandemics etc.

So It is not as simple as having a fixed idea of what capacity within the NHS should look like. For instance here is an example of an official NHS letter from September 2022 indicating the natural ebb and flow of capacity year round.
"*NHS England has taken action to boost capacity ahead of winter… recognising the fundamental importance of primary care in underpinning NHS services* (NHS. (n.d.) [2])"
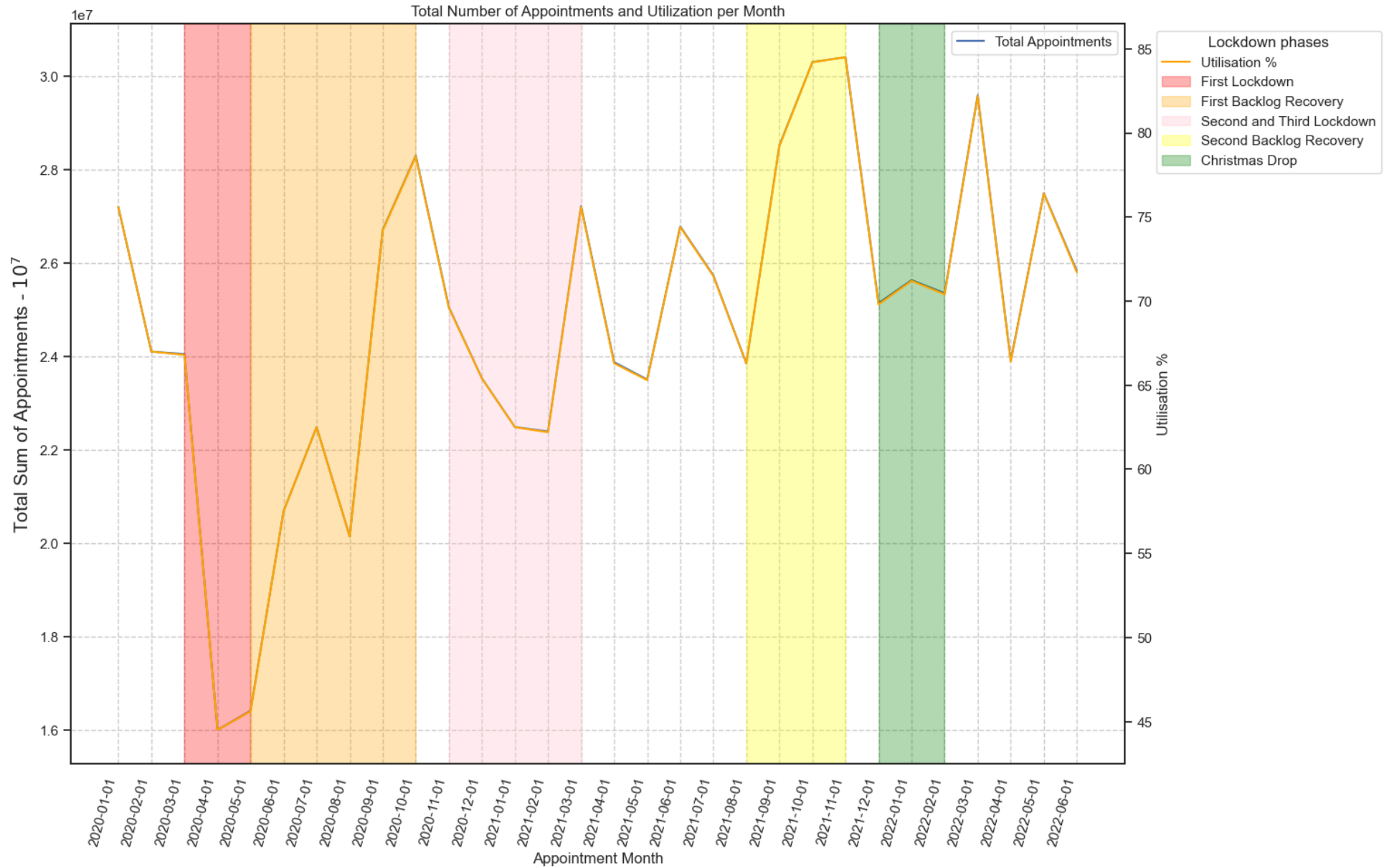
In order to properly answer the capacity question more accurately we would need access to two further data areas:
- First, Sum of Appointments data with a larger date range would give a fairer, more balanced assessment.
    - o This date range should include the period of pandemics and recovery as the NHS can't operate at 100% capacity either.
    - o Capacity needs to be able to account for pandemics and cyclical seasonal peaks and troughs.

- Second, Maximum capacity limits ideally at a icb_level and ideally within the actual resource categories themselves.
    - o For instance daily capacity limits, at a national category level, would allow us to address capacity, and utilisation, at a level of detail where resourcing recommendations can be more specific.

That is not something that we can currently address with the data provided but would be keen to in the future.

3. **B** – *Exploring the date range the appointments take place.*

[2] NHS. (n.d.). NHS choices. https://www.england.nhs.uk/long-read/supporting-general-practice-primary-care-networks-and-their-teams-through-winter-and-beyond/

Total Number of Appointments and Utilization per Month

Appointments regional covers the longest period of time from the three provided datasets, January 2020 to June 2022. As such we will use it explore when these appointments did or didn't happen.
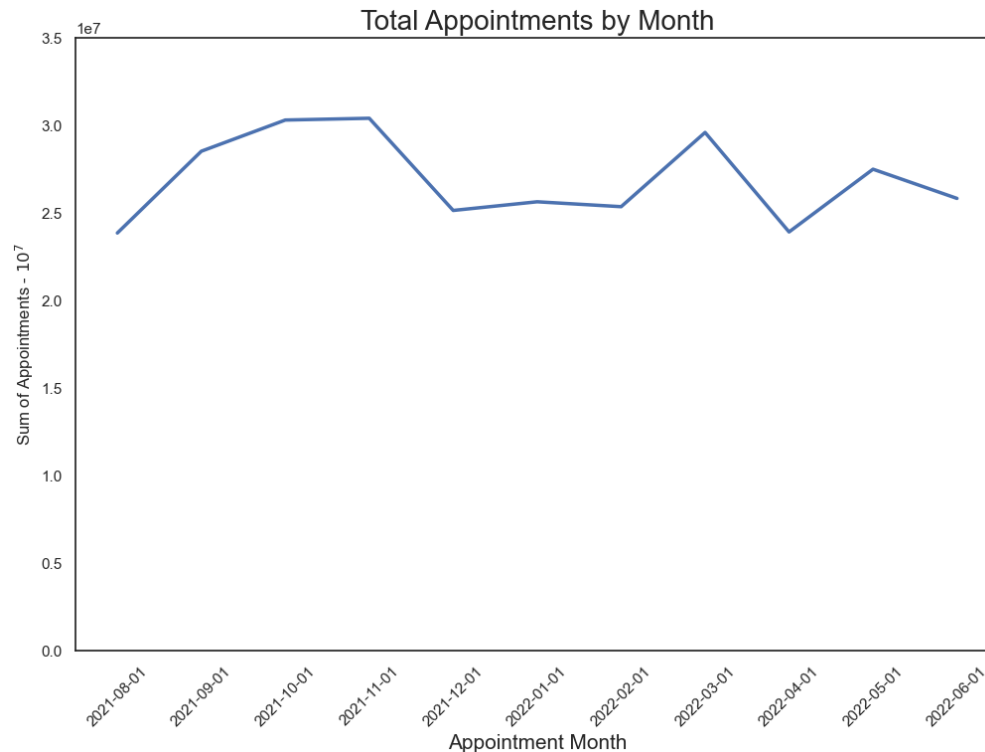
The chart is erratic, but it is possible to suggest some reason as to why there are such peaks and troughs over the provided time period.

Although their respective timelines cover different total time periods, this appointment regional chart aligns well with the overlapping period found in the national_category chart.

You can see this starting at 2021-08, notice the same "W" shape towards the end.

Let's breakdown the suggestions and implications for the peaks and troughs of appointment sums in this period.
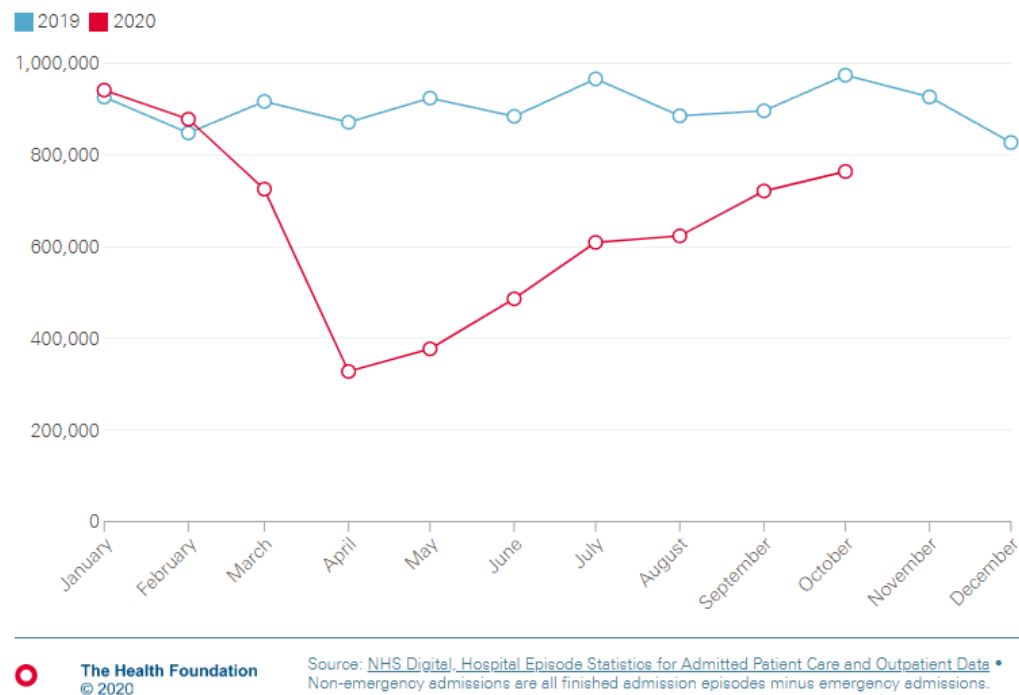
The First England Lockdown was introduced on the 26th of March 2020 and fractured into separate rule on May 11th 2020. Restrictions and government advice to stay isolated saw appointments plummet. "*NHS England instructed hospitals early in the pandemic to postpone non-urgent, planned treatment to free up beds and staff to care for people with serious COVID-19*" (*The Health Foundation.* 2020) [3])

---

[3] Non-covid-19 NHS care during the pandemic. The Health Foundation. 2020, https://www.health.org.uk/news-and-comment/charts-and-infographics/non-covid-19-nhs-care-during-the-pandemic

As such appointments dropped and then created **The First Backlog Recovery,** particularly for planned clinical procedures. It is the view of NHS Providers, the membership organisation for the NHS, that in this time period of October 2020 "*trusts are doing all they can to restore elective care as swiftly as they can, prioritising waiting lists based on clinical need, working collaboratively to help manage the backlog (* NHS Providers*, (n.d.) [4]*)

Figure 5:

The number of non-emergency admissions to hospital

2019 2020

Source: NHS Digital, Hospital Episode Statistics for Admitted Patient Care and Outpatient Data •
Non-emergency admissions are all finished admission episodes minus emergency admissions.

The Health Foundation
© 2020

Indeed, the graph here from The Health Foundation shows a spike in non-emergency admissions to hospital that starts gently in April and ramps up to October. The same trajectory as our graph, including the stutter in July. We will see a similar pattern for the yellow strip coming up, this too might be due to elective care spikes at that time. Were we to have National Category data for 2020 I should imagine that we would have seen a similar spike as the NHS attempted to fight the first backlog.

(*The Health Foundation. 2020* [5])

[4] nhsproviders.org. (n.d.). *Mapping the NHS winter*. [online] Available at: https://nhsproviders.org/mapping-the-nhs-winter/capacity [Accessed 3 Dec. 2023]

[5] Non-covid-19 NHS care during the pandemic. The Health Foundation. 2020, https://www.health.org.uk/news-and-comment/charts-and-infographics/non-covid-19-nhs-care-during-the-pandemic

As Covid cases increase, **The Second England Lockdown** was introduced on the 5th of November 2020 then arguably ended on the 23rd of December 2020 with the exception of residents in London.
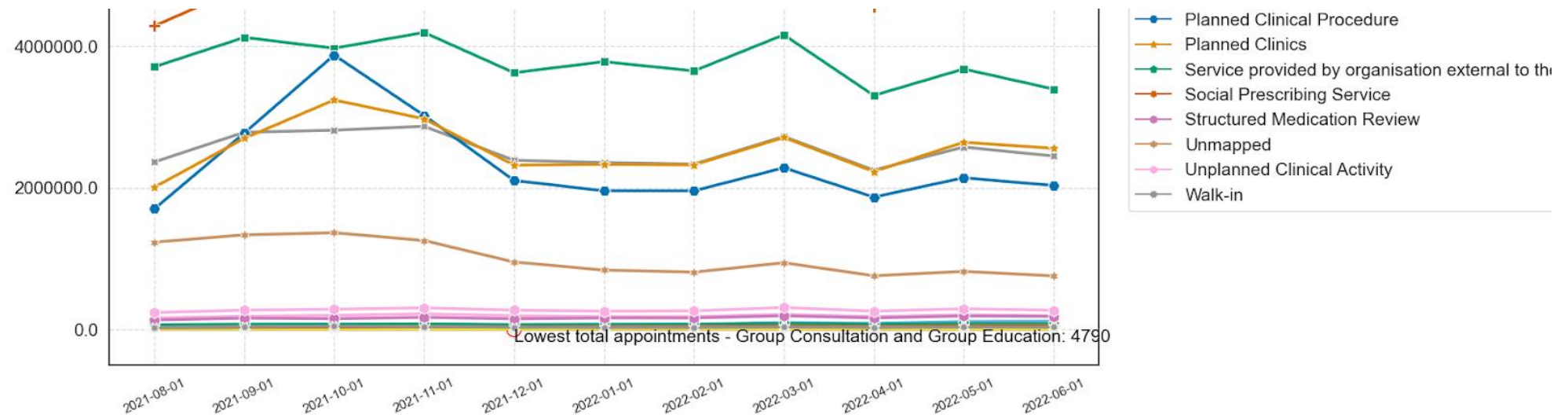
Soon after, **The Third and final England Lockdown** was introduced on the 4th of January 2021 and began to be lifted on the 8th of March 2021. I have chosen the beginning of the timeline for lifting restrictions however the final restrictions were lifted on the 17th of May 2021, which could explain the next dip as the population staggers out of restrictions.

Restrictions once again continued to affect appointments as we see a corresponding dip in the graph. Then we see a natural period of rapid activity as the population books appointments and resumes life as normal

Ultimately this culminated in a big spike around November 2021 which I would suggests is **The Second Backlog Recovery**. Delivery plan for tackling the COVID-19 backlog of elective care document published in February 2022 suggests that "*The wall of protection that our world-leading NHS COVID-19 vaccination programme gave us over spring and summer 2021 meant that hospitals were able to make good progress on recovering non-COVID care including elective treatment… These initiatives and many more besides meant that, despite the ongoing limitations of infection prevention and control measures, by November 2021 elective and diagnostics activity levels in many areas had recovered towards, or above, pre-pandemic levels, with an all-time record number of urgent cancer checks carried out.*" (NHS. (2022) [6])

This can be seen in the national category chart below which shows sum of appointments by month. Notice the sudden rise in planned clinical procedures and the rise in Planned Clinics.

[6] NHS. (2022). Delivery plan for tackling the COVID-19 backlog of elective care. https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2022/02/C1466-letter-delivery-plan-for-tackling-the-covid-19-backlog-of-elective-care.pdf

Then, in December 2021 where there is another significant dip. This dip in appointments is more complex.

The figure below from the Office of National Statistics views the data through General Practice (GP) consultations per person, per year from January 2019 to January 2022 and it follows a similar path to our chart. They too note the "*sharp reduction in consultation rates at the beginning of the pandemic (March 2020) which return to pre-pandemic levels in autumn 2020*" ([Department of Health Social Care](). (2022) [7]).

[7] Department of Health & Social Care. (2022). Direct and indirect health impacts of covid-19 in England: Emerging omicron impacts. GOV.UK. https://www.gov.uk/government/publications/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts#category-c-indirect-impacts-of-covid-19-on-population-health-due-to-living-through-a-pandemic-and-restrictions-1

Figure 17: GP consultations per person per year
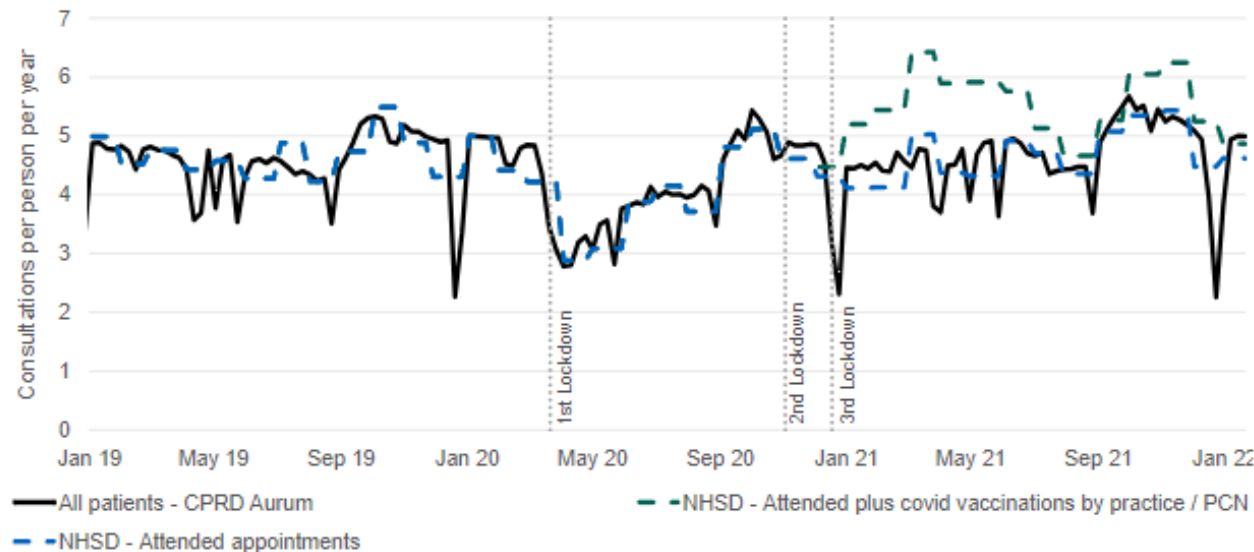
Source: analysis conducted by the Health Foundation's REAL Centre using CPRD Aurum and NHS Digital data. See Annex 3 for background and methodology on the CPRD.

[8] (Department of Health Social Care. (2022))

In spite of the Omicron variant arriving in November 2021 this same article suggests that "*there is no discernible impact of Omicron on consultation rates, with the drop in December 2021 likely reflecting normal seasonal variations.* (Department of Health Social Care. (2022) [9])"

[8] Department of Health & Social Care. (2022). Direct and indirect health impacts of covid-19 in England: Emerging omicron impacts. GOV.UK. https://www.gov.uk/government/publications/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts#category-c-indirect-impacts-of-covid-19-on-population-health-due-to-living-through-a-pandemic-and-restrictions-1

However, seasonal variations clearly play a significant role in appointments, I would suggest that the dip in appointments through to March is affected by Omicron. As the ONS suggests in a report, "*the Omicron BA.1 variant was dominant (20 December 2021 to 1 March 2022)*". ONS. (2023, March 29) [10]). Indeed, "*In the period when the Omicron BA.2 variant was dominant (2 March to 15 June 2022), all English regions peaked at around 7% to 9%, in the 14 days between 23 March and 6 April 2022. ONS. *" (2023, March 29) [11]) With the ONS citing that "*in the week ending 12 March around 1 in 20 people in England were infected with the virus*" ONS. (2023, March 29) [12]).

Finally Plan B restrictions, which started on the 10th December 2020 and ended on the 26th of January, included restrictions such as facemasks, working from home and COVID passes.

These restrictions plus the continuing presence of Omicron would have contributed to a consistent fluctuations and subsequent dip in appointments that extend beyond the natural seasonal dip at that time.

However, due to the uncertainty of conclusions in this time period I am only going to highlight the lockdowns on future maps. In terms of analysis, it seems prudent to focus on the certainties.

So, directly addressing the question "Has there been adequate staff and capacity in the networks?" is complex.

[9] Department of Health & Social Care. (2022). Direct and indirect health impacts of covid-19 in England: Emerging omicron impacts. GOV.UK. https://www.gov.uk/government/publications/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts#category-c-indirect-impacts-of-covid-19-on-population-health-due-to-living-through-a-pandemic-and-restrictions-1
[10] ONS. (2023, March 29). Coronavirus (COVID-19) latest insights: Infections. Coronavirus (COVID-19) latest insights - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/infections#:~:text=During%20the%20period%20when%20the,in%20January%20and%20February%202022.
[11] ONS. (2023, March 29). Coronavirus (COVID-19) latest insights: Infections. Coronavirus (COVID-19) latest insights - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/infections#:~:text=During%20the%20period%20when%20the,in%20January%20and%20February%202022.
[12] ONS. (2023, March 29). Coronavirus (COVID-19) latest insights: Infections. Coronavirus (COVID-19) latest insights - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/infections#:~:text=During%20the%20period%20when%20the,in%20January%20and%20February%202022.

Considering the NHS in England did not rise above 85% capacity during what Siva Anandaciva, Chief Analyst at The King's Fund, has called "*the most extreme pressure seen in recent history*" The King's Fund. (2021, January 14)[13]) you could reasonably argue that yes, there was adequate capacity.

3. **C** - *Where are the appointments geographically?*

The distribution of appointments in regions is in parallel and appropriate with their relative populations of UK suggesting a healthy utilisation.

Then regional appointment sum graph below shows varying distribution of data across regions which implies strongly that capacity is a relative concept and that we can't necessarily draw inferences from the previous NHS wide standard.

- The region with the most appointments is The Midlands which has 19% of total appointments or 32,574,555 appointments
    - o The two regions of the Midlands have a combined population of 10,350,697, 18% percent of the UK population
- The region with the least appointments is The South West which has 11% of total appointments or 18,486,437 appointments.
    - o The South West has a population of 5,701,186, 10% of the uk population.

[13] The King's Fund. (2021, January 14). *"the NHS is now under the most extreme pressure seen in recent history":* https://www.kingsfund.org.uk/press/press-releases/nhs-now-under-most-extreme-pressure-seen-recent-history

Both of these region's populations correlate well with their appointment sum percentage suggesting a healthy balance between the two.

Resourcing should be decided on a regional basis and not a nationwide metric. As Robert King states in an article about region inequalities in the NHS "*it is clear that the East of England, Midlands and areas in the North West are being left further behind, and policymakers can help move the needle by proportionately targeting resources and support to areas with the greatest need.* " (lcp. (2023, March 13).[14])

[14] lcp. (2023, March 13). *Regions being "left further behind" as drive to cut NHS waiting lists has most impact in the South East*. Lane Clark & Peacock LLP.
https://www.lcp.com/media-centre/2023/03/regions-being-left-further-behind-as-drive-to-cut-nhs-waiting-lists-has-most-impact-in-the-south-east

Distribution of Appointment Sum by Region
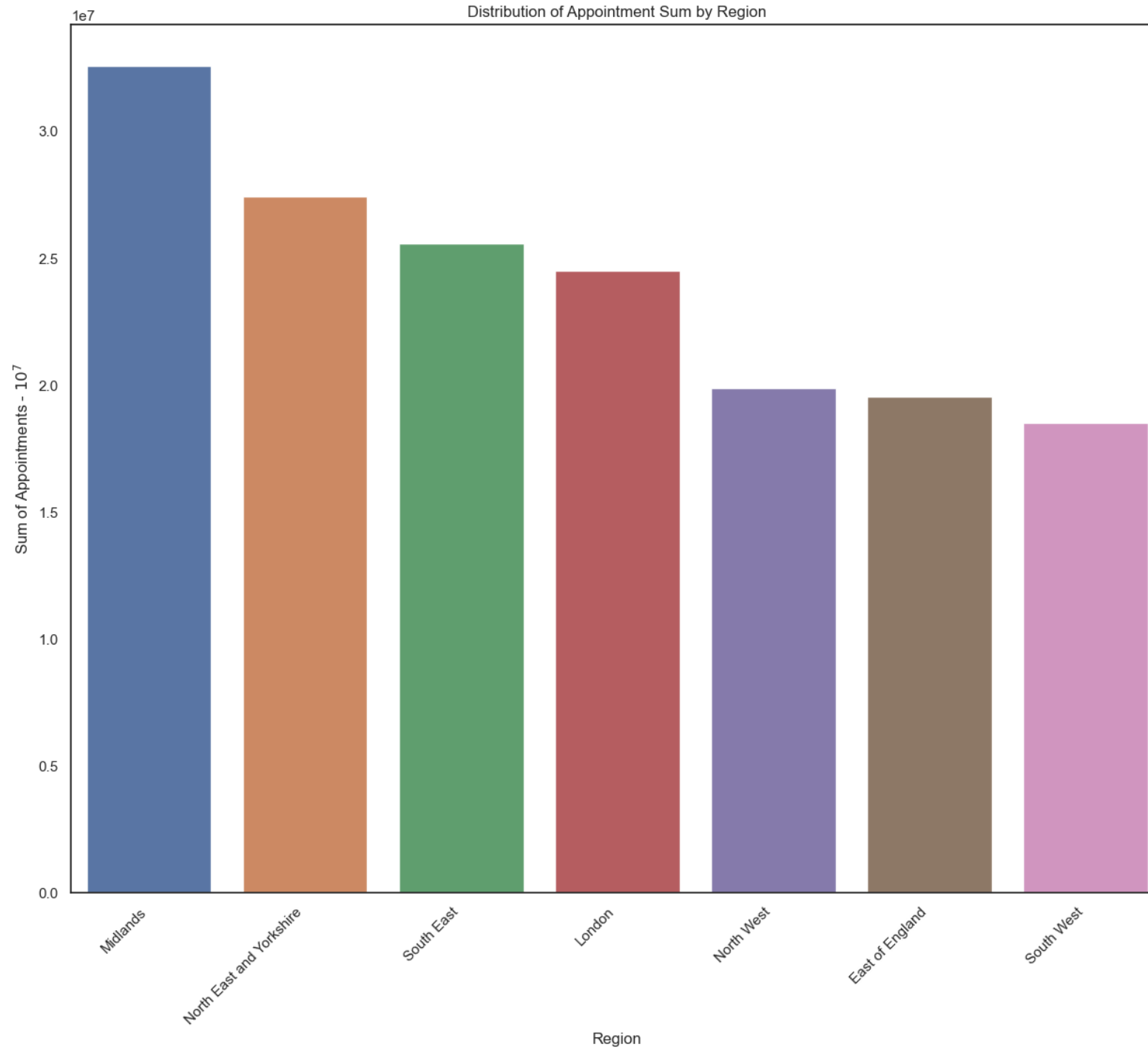
With that in mind, further analysis into the distribution of appointments feels pertinent in order to address any potential resourcing issues.

This is also where we begin to feel the affects of potential data issues. The Midlands might have better reporting of data which could skew results.

The top IBC locations by sum of appointment count below shows us just how varied the icb's are across regions, spanning the length and width of the country.
- The spread of sub_icb's, Kent, London, Birmingham represent several regions.
- The main takeaway from looking at the regions in more detail shows us another shift in representation regionally, making it clear that looking in details also has its merits.
- And like how the regional spread shows us that a national_capactiy metric is problematic, sub_icb's show us how the regional national_capacity metrics would have problems too.

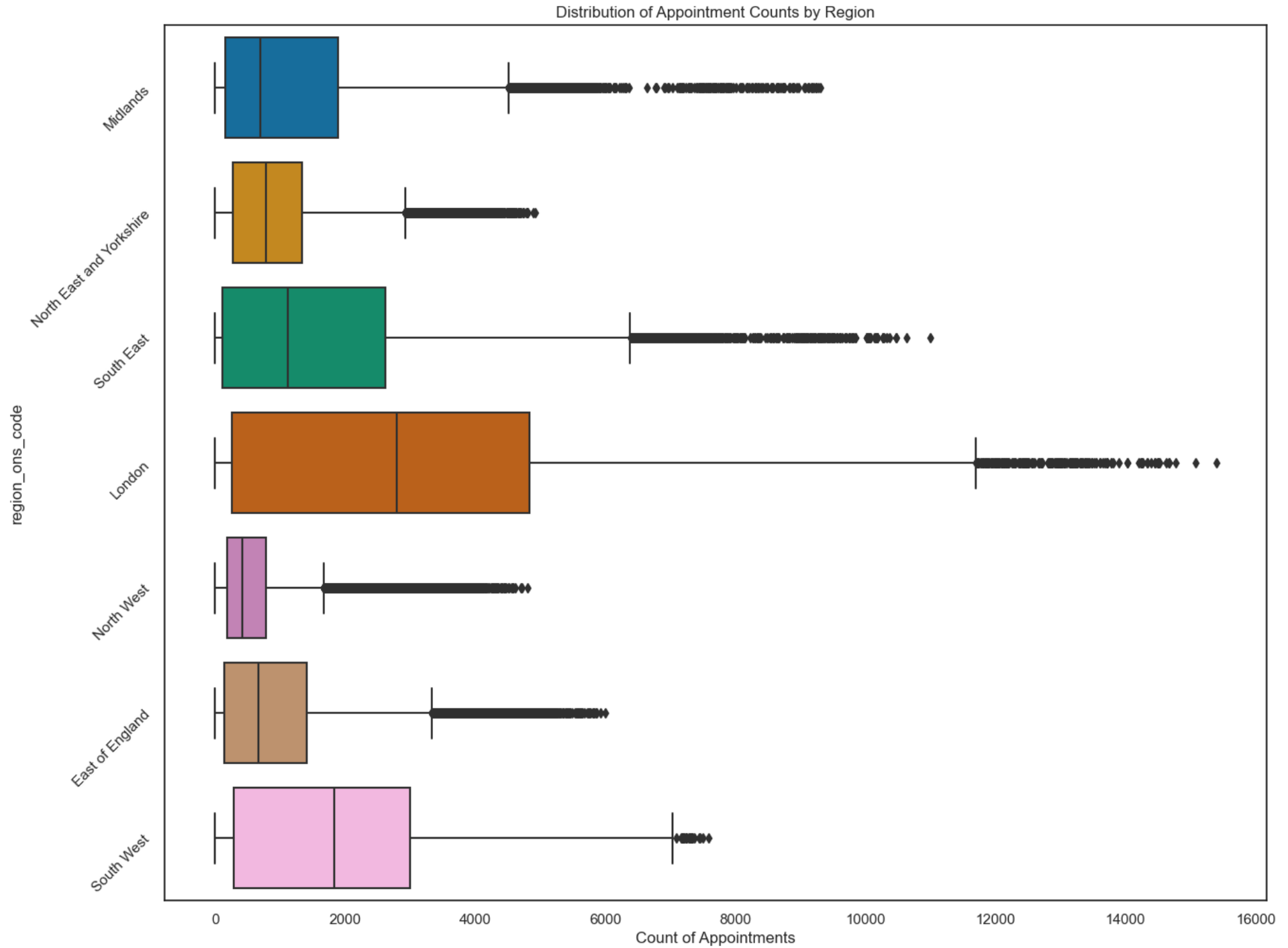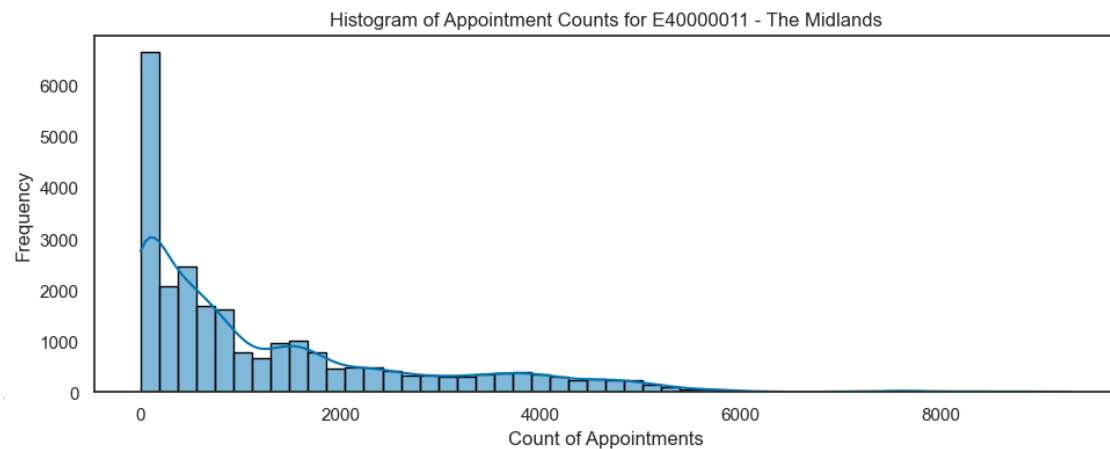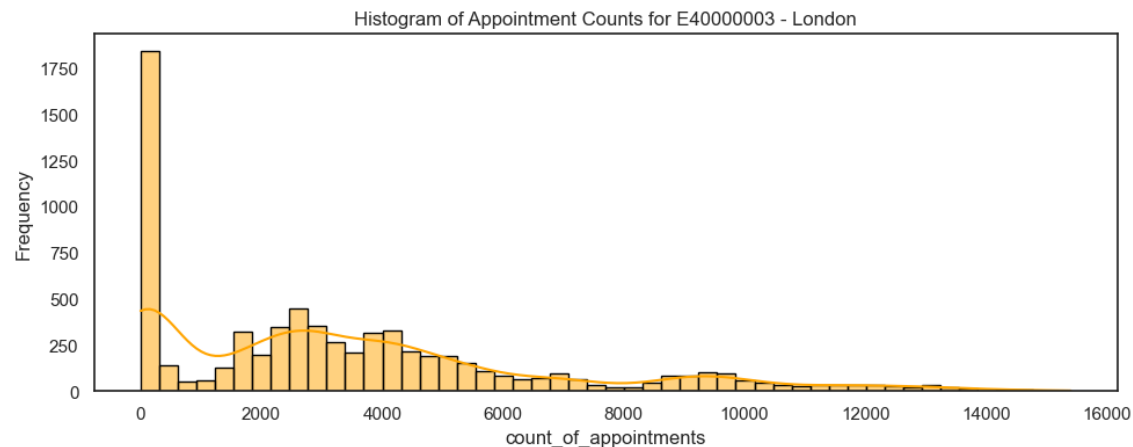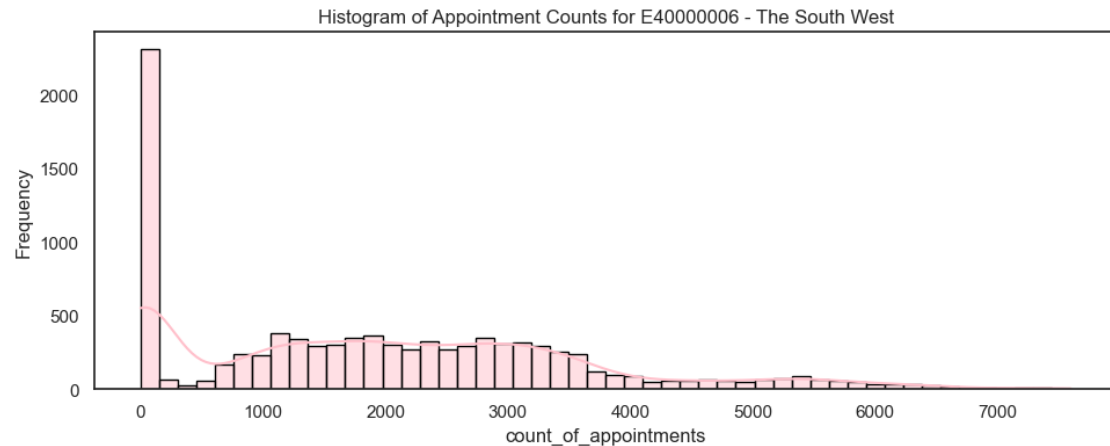Distribution of appointments by sub icb location in E40000011 The Midlands
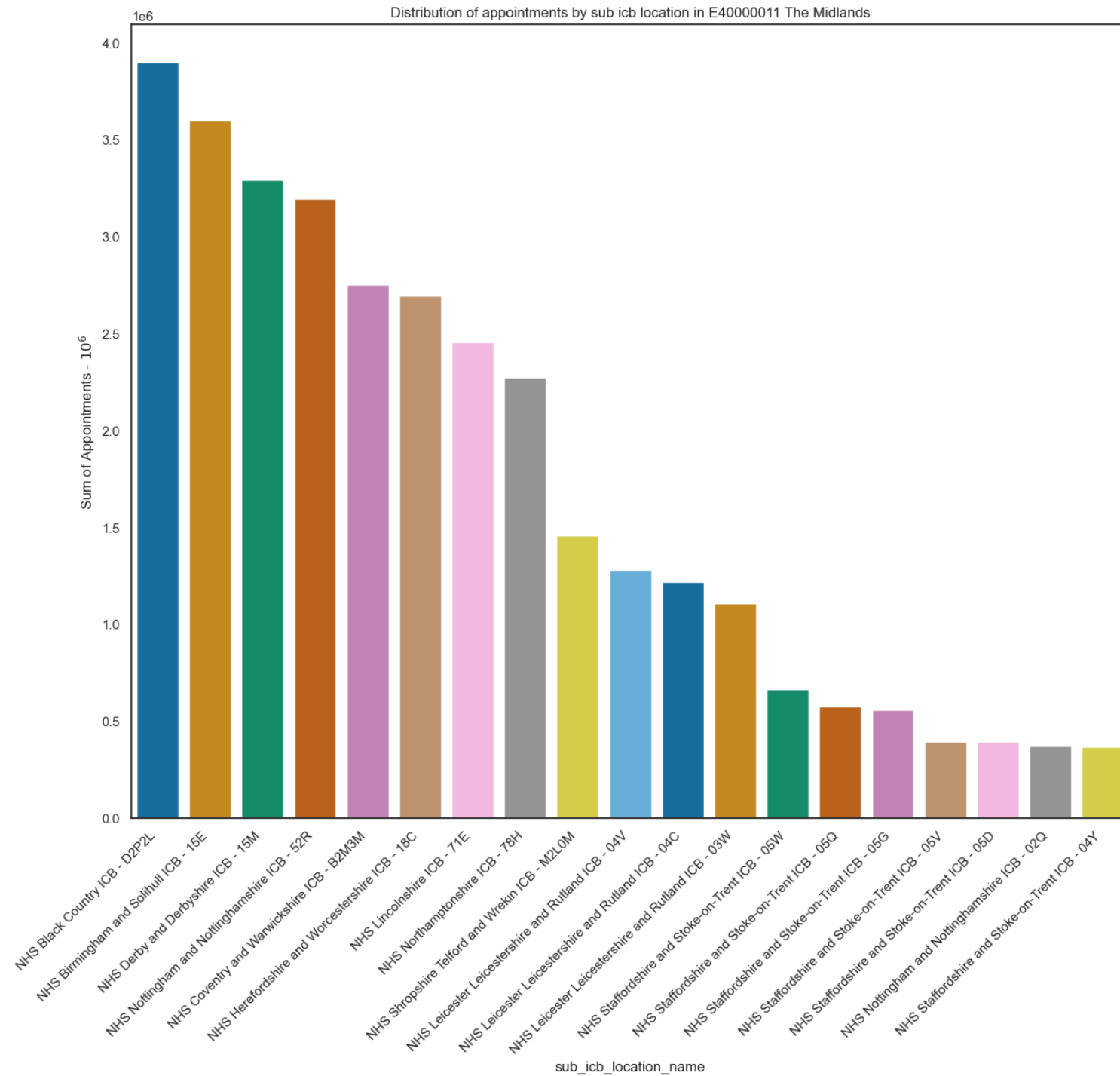
There is perhaps an impracticality to immediately working at a sub_icb level as there are 106 sub icb's and only 42 icb's. Working with resourcing at an icb level to start with would be easier.

Also worth noting potential data quality issues here as data could be affected by something as simple as The Midlands having better data reporting than the North East for instance.

It is interesting to use a boxplot to show the distribution of the appointments in The Midlands.

The figure below shows us that there are many significant outliers beyond the IQR for all charts but The Midlands has a particularly small box considering it's the region with the most total appointments.

Distribution of Appointment Counts by Region

Histogram of Appointment Counts for E40000006 - The South West

Histogram of Appointment Counts for E40000003 - London

Histogram of Appointment Counts for E40000011 - The Midlands

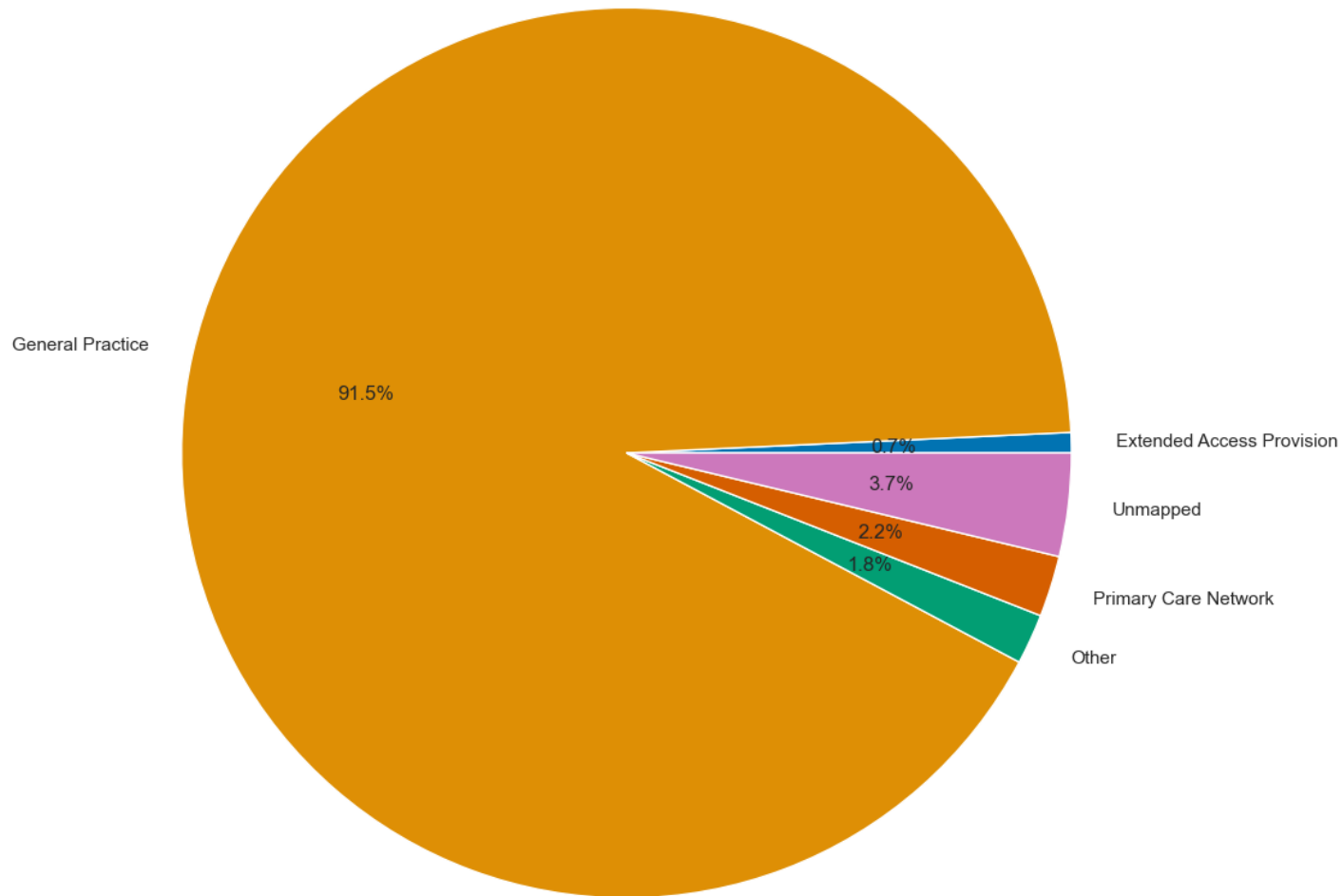This histogram shows a similar story to the boxplot above

Appointments are front loaded

It's a tale of frequency being the key metric for Midlands, that's how it got the region with the highest sum appointments.

- The South West has medium frequency and medium count

- London has low frequency but high count

- The Midlands has very high frequency and medium count, hence it's place as number one.

3. D- Where are the appointments within the NHS?

Distribution of Appointments by Service Setting

General Practice — 91.5%
Extended Access Provision — 0.7%
Unmapped — 3.7%
Primary Care Network — 2.2%
Other — 1.8%

The distribution of service setting by total appointment sum shows us that appointments are taking place in General Practice.

To explore this further let's consider breaking down the distribution of service settings by month. This will give us an idea of whether there is any specific time variation to the data.

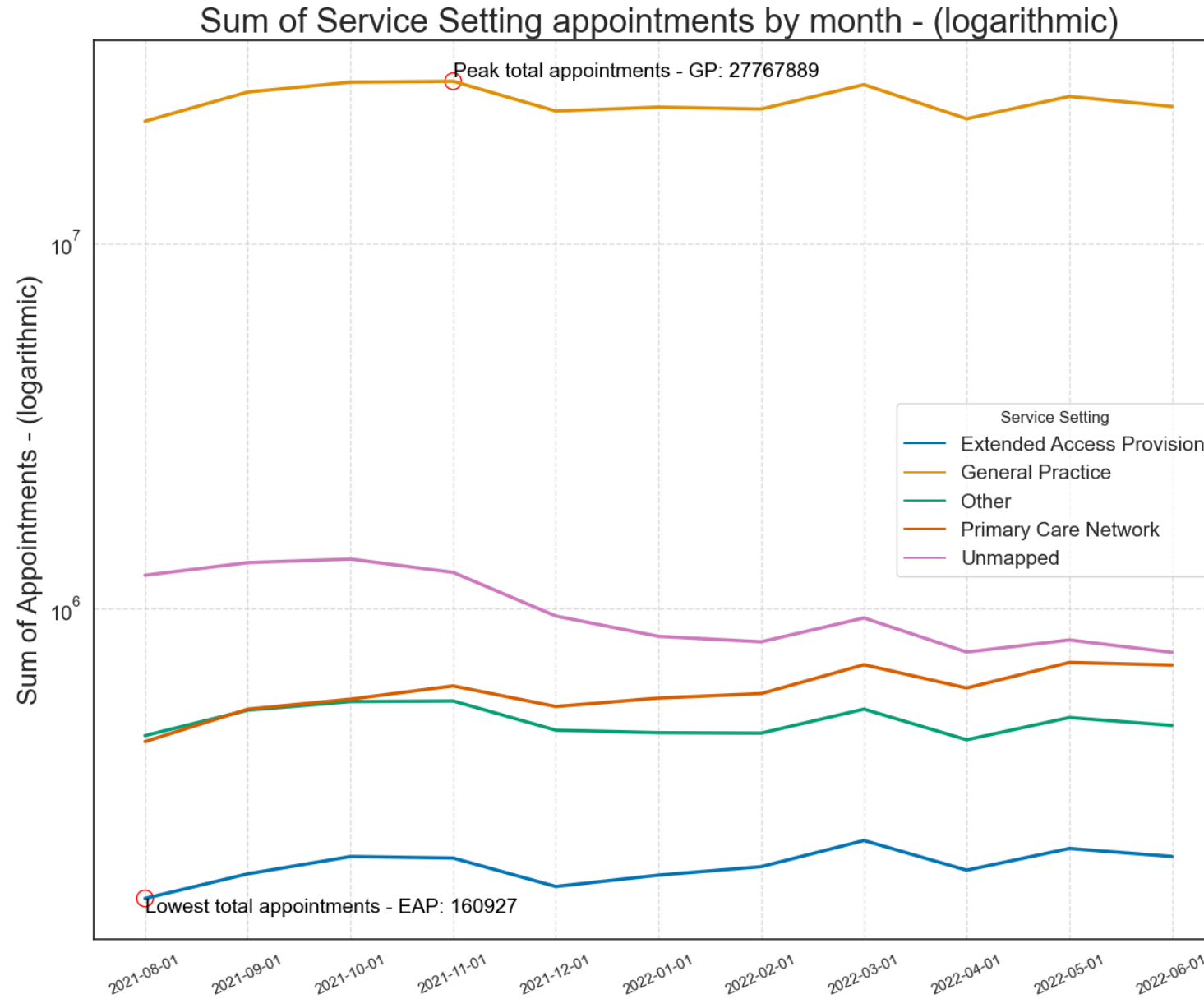## Sum of Service Setting appointments by month
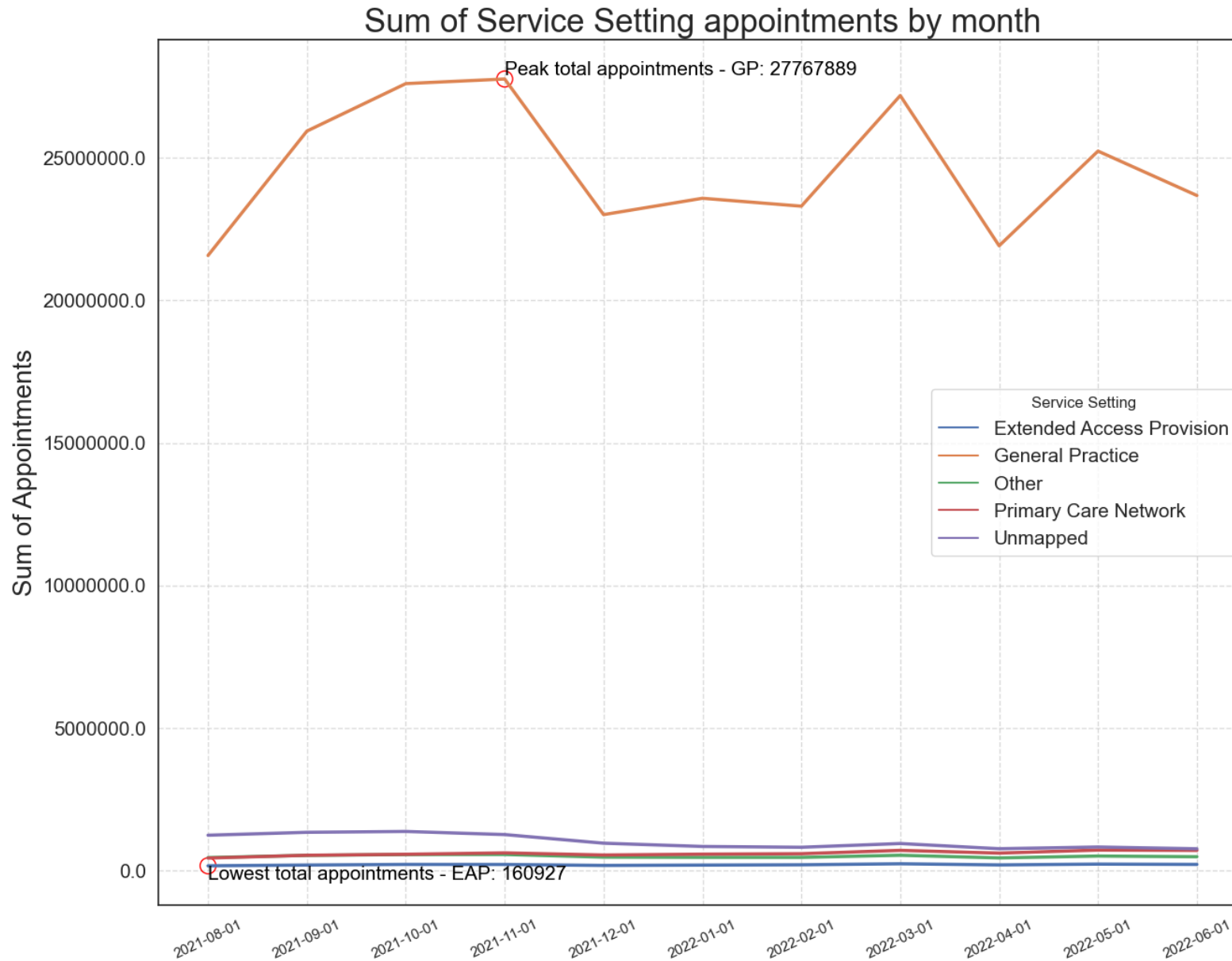
This graph shows the same sum of service setting appointments per month.

Which confirms that the biggest service setting utilisation is in General Practice coming in at 27,767,889 million total appointments in the data set.

The lowest total appointments is Extended Access Provision with only 160,927 total appointments in the data set.

Extended Access Provision is the ability to see your GP for routine appointments outside of normal hours. Either early in the morning, evening or at weekends.

Given the gulf between service settings in the data, the lower half of the chart is hard to read so let's expand the remaining clustered settings by looking at them logarithmically.

Sum of Service Setting appointments by month - (logarithmic)

Peak total appointments - GP: 27767889

Service Setting
— Extended Access Provision
— General Practice
— Other
— Primary Care Network
— Unmapped

Lowest total appointments - EAP: 160927

Making the plot logarithmic compresses the vertical scale, making it easier to see details in data that vary widely.

Imagine a graph where the numbers on the side go 1, 2, 3, 4... and so on. Now, switch to a logarithmic scale, and the numbers might be 1, 10, 100, 1000. This is helpful when the data has big differences between values.

In our plot where the months or service settings have a lot more appointments than others, you can still see the patterns and differences clearly, making the information more accessible and easier to interpret.
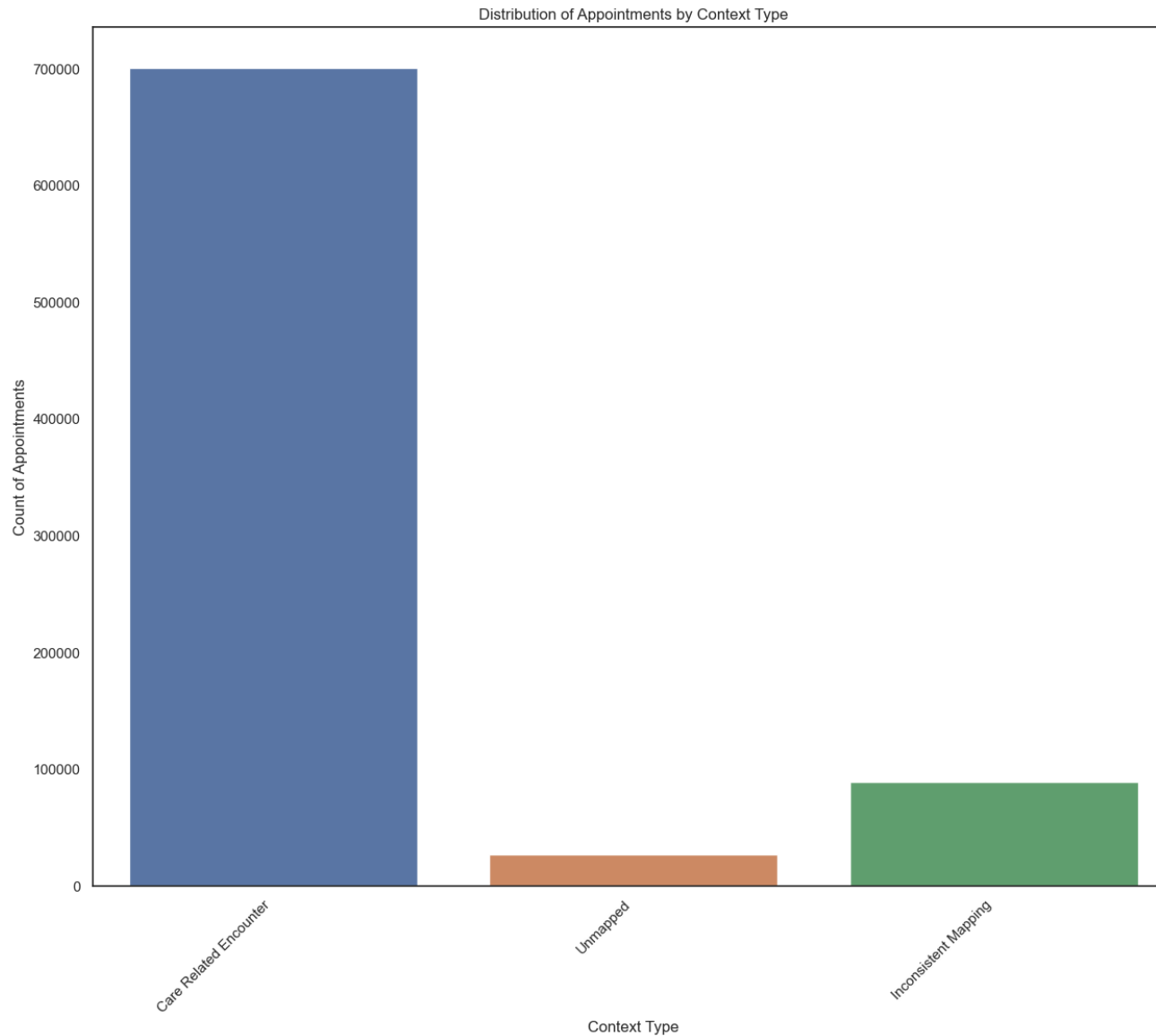
By changing the scale we see two key insights:
- -First is that the unmapped data in purple is trending down, which is a positive sign in terms of data quality and reporting
- Second is that Primary Care Network appointments are trending up which is consistent with the investment and attention being paid in particular to the PCN's by the NHS in recent years.
  - *"Funding through voluntary incentives such as the IIF"* (NHS. (n.d.) [15]) *is the principal way in which NHS England will be promoting PCN service improvement goals from the Long Term Plan"* (NHS. (n.d.) [16])
  - In simple monetary terms. The "IIF will be worth £150m to PCNs for 2021/22 and £225m for 2022/23. In addition, we can also confirm new funding for PCN leadership and management support, of £43m in 2021/22."

Still, while relevant and notable, the main insight still remains that the vast majority of appointments utilise the general practice setting.

.

[15] *"The Investment and Impact Fund (IIF) is an incentive scheme focussed on supporting PCNs to deliver high quality care to their population, and the delivery of the priority objectives articulated in the NHS Long Term Plan and in Investment and Evolution; a five-year GP contract framework."* NHS. (n.d.). Investment and Impact Fund. NHS choices. https://www.england.nhs.uk/primary-care/primary-care-networks/network-contract-des/iif/
[16] NHS. (n.d.). Investment and Impact Fund. NHS choices. https://www.england.nhs.uk/primary-care/primary-care-networks/network-contract-des/iif/
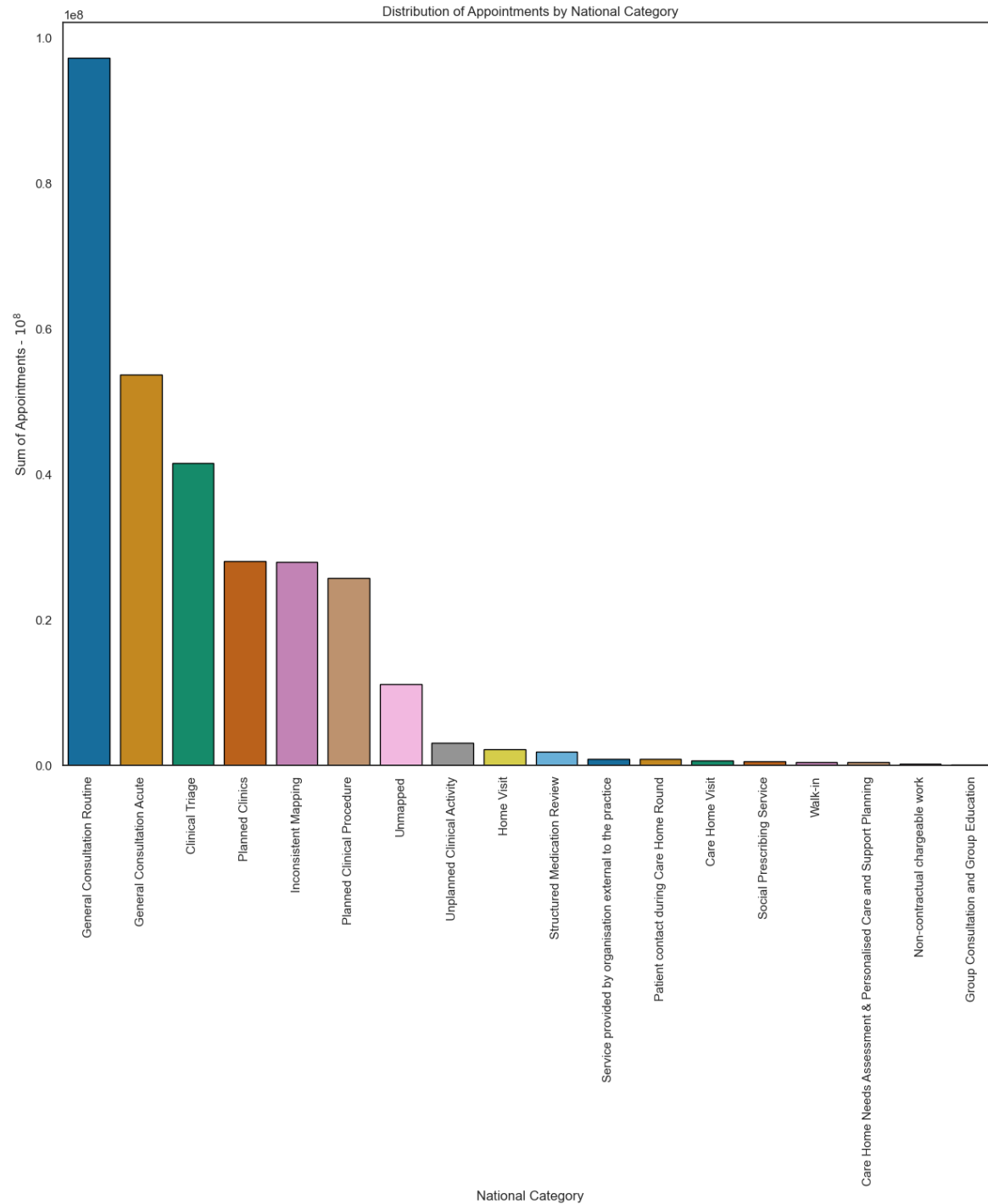
3. **E** - *What categories are the appointments in?*

Distribution of Appointments by Context Type

Context type by sum of appointments reveals, as expected, that most appointments are care related encounters, 26,282,778 appointments.

Categories within this context type involve the patient and can be any modality
Another key takeaway here is about data quality, which looks small but will be significant when further unpacked.

Let's now unpack this data into its component parts by looking at national_category.

Distribution of Appointments by National Category

National category by Sum of appointments shows us explicitly where utilisation is actually taking place within those care related encounters

Both general appointment categories take the top two positions with the most appointments in the dataset. Worth noting that Service provided by organisation external to the practice is very small 0% or 852,133. A good statistic to show just how much of an offering the NHS has and how much it retains complete direct control of most of it's services.
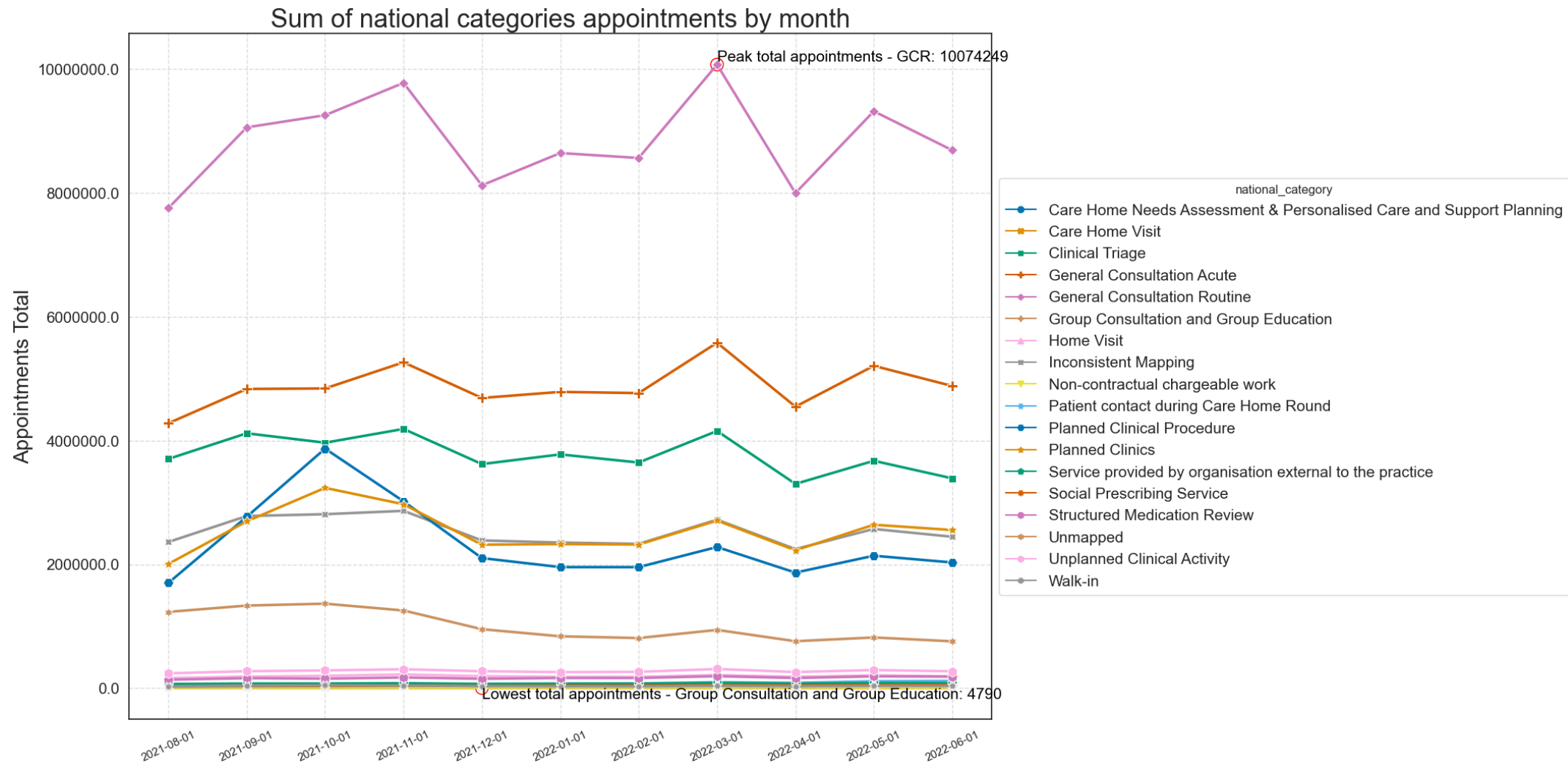
Inconsistent mapping is more obviously seen here in this chart and accounts for 9% of total appointments. Together with the result from unmapped, this equals around 13% of appointments reported suffer from data quality issues.

It's very hard to commit to an "acceptable" threshold of errors in data reporting and as such the NHS doesn't provide a hard number. However the local commissioning data quality guidance lists the reasons why data quality is vital:

- -"*Acceptable data quality is crucial to operational and transactional processes and to the reliability of business analytics / business intelligence reporting,*
- *High quality information leads to improved decision making which in turn results in better patient care, wellbeing and safety. There are potentially serious consequences if information is not correct, secure and up to date,*
- *-Management information produced from patient data is essential for the efficient running of the trust and to maximise utilisation of resources for the benefit of patients and staff,*
- *Poor data quality puts organisations at significant risk of: damaging stakeholder trust; weakening frontline service delivery; incurring financial loss; and poor value for money* " (NHS. (2016). [17])
- With this in mind, 13% or, 38,971,612 appointments, feels like an unacceptable loss of data.

Let's explore these national categories further by plotting them over the appointments regional time period.

[17] NHS. (2016). Data Quality - NHS England. https://www.england.nhs.uk/wp-content/uploads/2016/03/local-commissioning-data-quality-guidance.pdf

Sum of national categories appointments by month

Sum of national categories appointments by month shows us explicitly the utilisation of categories by month and for the most part there doesn't seem to be a notable specific monthly deviation in appointment numbers by category from the general monthly sum appointment list

This is with the notable exception of Planned Clinical Procedures which went up from 2021-08 to 2021-10 almost doubling the amount of appointments before heading down
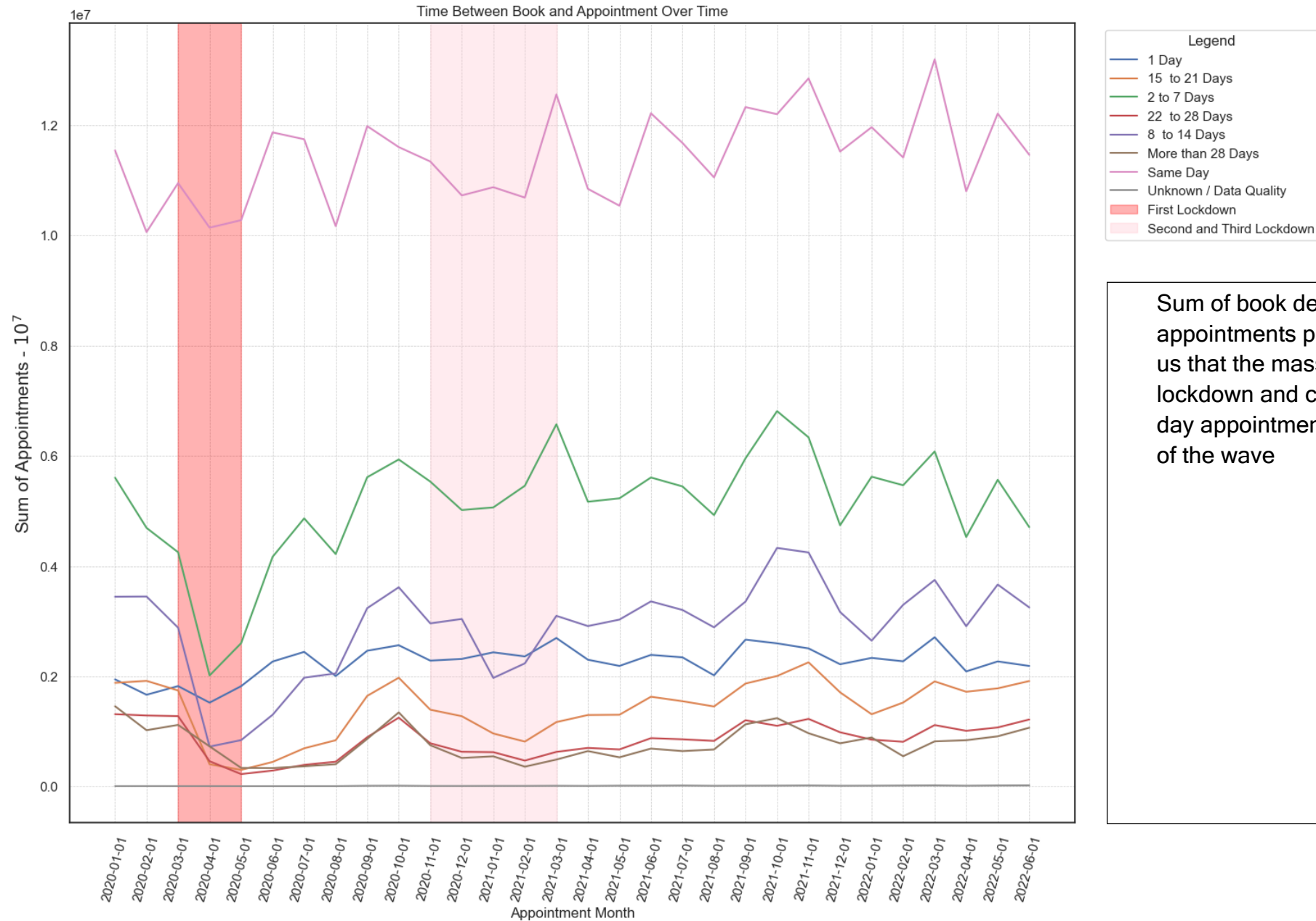
One possible explanation according to Professor Stephen Powis, National Medical Director for NHS England is that all of this is could be down to *"additional resources made available to us to recover services which were inevitably disrupted during the pandemic, and we are continuing to tackle the COVID backlog."* NHS. (2021[18]) A point I will refer back to often to justify my covid backlog recovery theory.

Beyond that, it's clear to see again that General Consultation Routine makes the majority of cases at 10,074,249 overall appointments.

General Practice's dominance continues to suggest that it plays a vital role as the "front door" of the NHS. But many patients still find it hard to reliably get a foot in that front door, so let's look at… The When

3. F - *When are the appointments taking place?*

[18] NHS. (2021). NHS choices. https://www.england.nhs.uk/2021/08/nhs-cuts-waiting-times-despite-busy-summer-period/

Time Between Book and Appointment Over Time

Sum of book delay appointments per month shows us that the massive effect of the lockdown and clearly how same day appointments is at the top of the wave

Bar Chart of Recommended 14 Days Wait Time
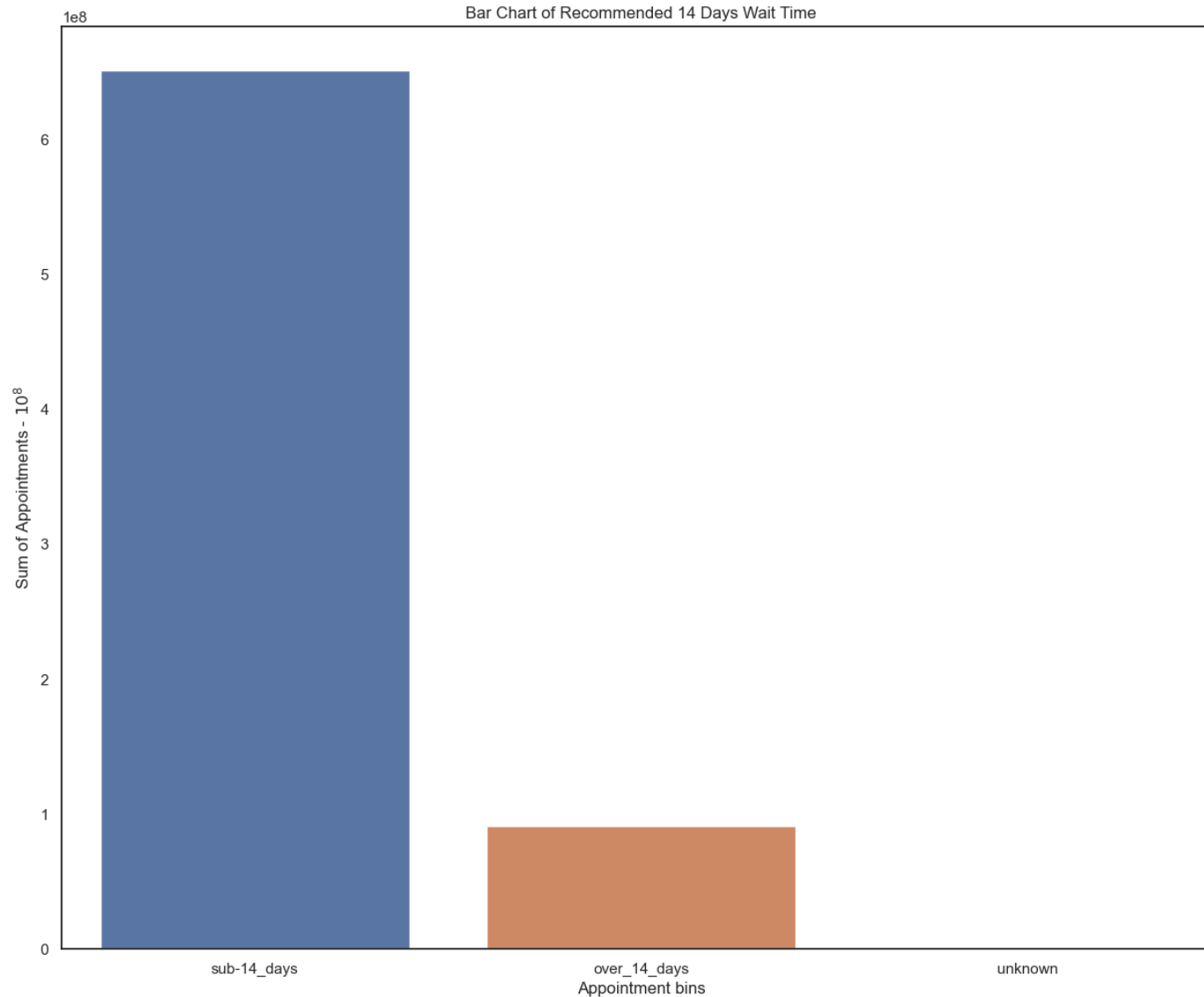
Plotting an exact target for NHS time between book and appointment is difficult and it is hard to find a figure suggesting one. However A good idea of the Government's target suggested by Therese Coffey when she was Health Secretary. She suggested that "*No one should wait more than two weeks to see GP* ." As such I will use this as my metric for analysing how many GP appointments fall within this category and compare the two groups whilst including Unknown data.

Stats from this graph:
- The majority of appointments are sub_14_days. That's 651,104,318 total or 88%.
- Over_14_days is 91,298,102 total appointments or 12%
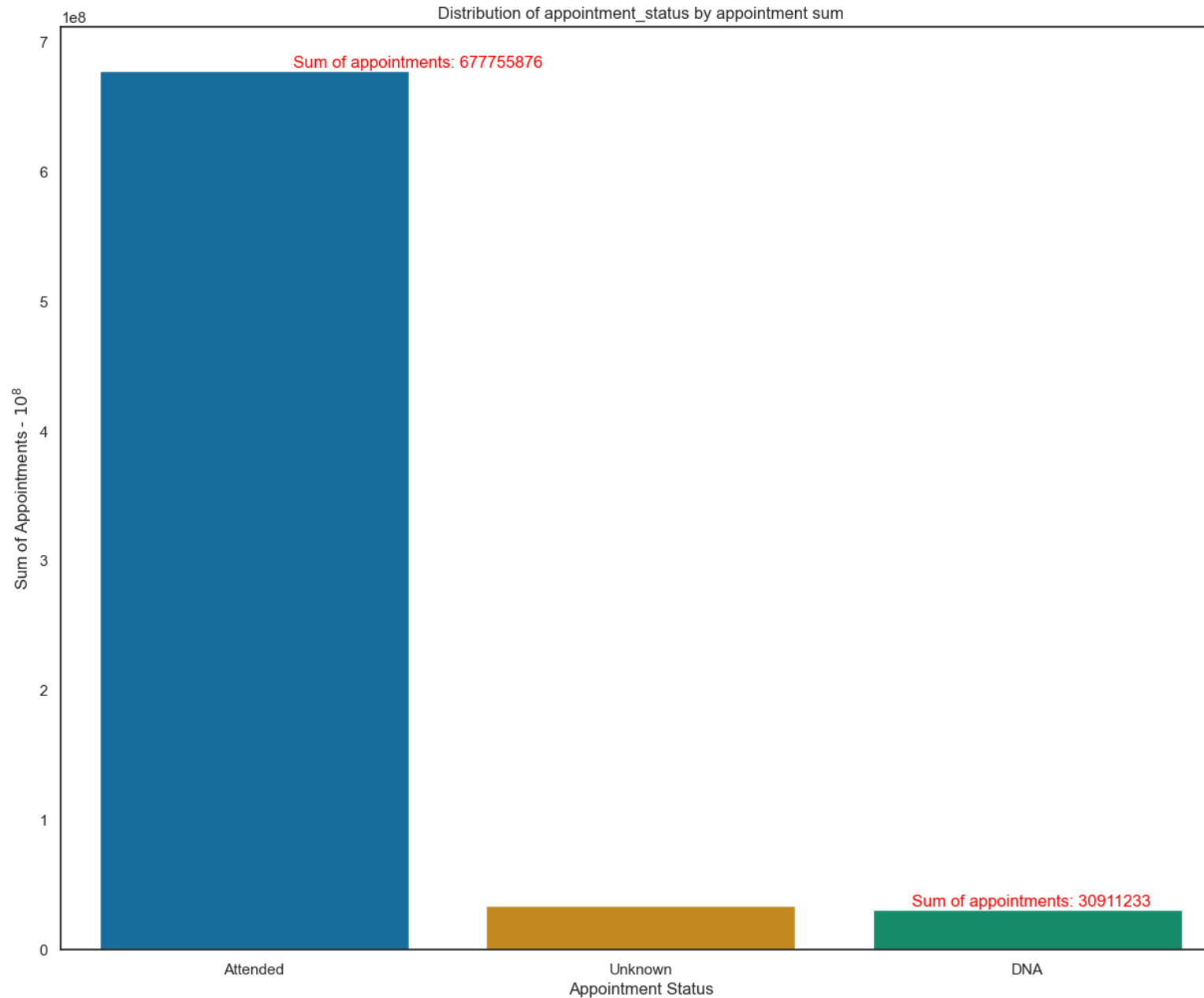- And Unknown is only 402,105 or 0%

* Therese Coffey "*No one should wait more than two weeks to see GP .*" ([Guardian News and Media](). (2022, September 21) [19]).

So the NHS is operating quite efficiently according to the Government's earlier suggested figure of sub 14 days. In fact, Same Day appointments count for 46% of all total appointments.

Nevertheless, 91,298,102 is not a small number of appointments and even though 3% of appointments are more than More than 28 Days that is still 23,050,987 appointments.

3. **G** - *Who are the appointment's patients?*

[19] Guardian News and Media. (2022, September 21). No one should wait more than two weeks to see GP, Coffey to say. The Guardian. https://www.theguardian.com/society/2022/sep/21/no-one-should-wait-more-than-two-weeks-to-see-gp-coffey-to-say

Distribution of appointment_status by appointment sum

Sum of appointments: 677755876

Sum of appointments: 30911233

Sum of Appointments - 10⁸

Appointment Status

Attended   Unknown   DNA

Appointment status over time by sum of appointments by month shows us that vast majority of appointments attended by patients.
- 91% attended, 4% did not attend and 5% are unknown.
- Again it is worth noting that 4% feels small but that is 30,911,233 not attended appointments.

In 2019 a report on the cost of missing appointments calculated the following.

- *"Each appointment costs an average of £30, putting the total cost to the NHS at more than £216 million pounds on top of the disruption for staff and fellow patients that would pay for:*
  - *The annual salary of 2,325 full time GPs*
  - *224, 640 cataract operations*
  - *58,320 hip replacement operations*
  - *216,000 drug treatment courses for Alzheimer's*
 - *The annual salary of 8, 424 full time community nurses"*. (NHS. (2019) [20])

If we apply that £30 missed appointment metric to our data, then 30,911,233 not attended appointments over the 18 months period that the dataset covers costs the NHS £ 927,336,990

An eye watering figure but not impossible as the dataset ar covers a long time period and from April 2020 to February 2021 the NHS lost £288 million because of COVID alone. (Person, O. (2023, March 29) [21])

3. **H** - *Who are the appointment's staff?*

[20] NHS. (2019). NHS choices. https://www.england.nhs.uk/2019/01/missed-gp-appointments-costing-nhs-millions/
[21] Person, O. (2023, March 29). Coronavirus (COVID-19) latest insights: Infections. Coronavirus (COVID-19) latest insights - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/infections#:~:text=During%20the%20period%20when%20the,in%20January%20and%20February%202022.

Distribution of Appointment Mode by appointment sum
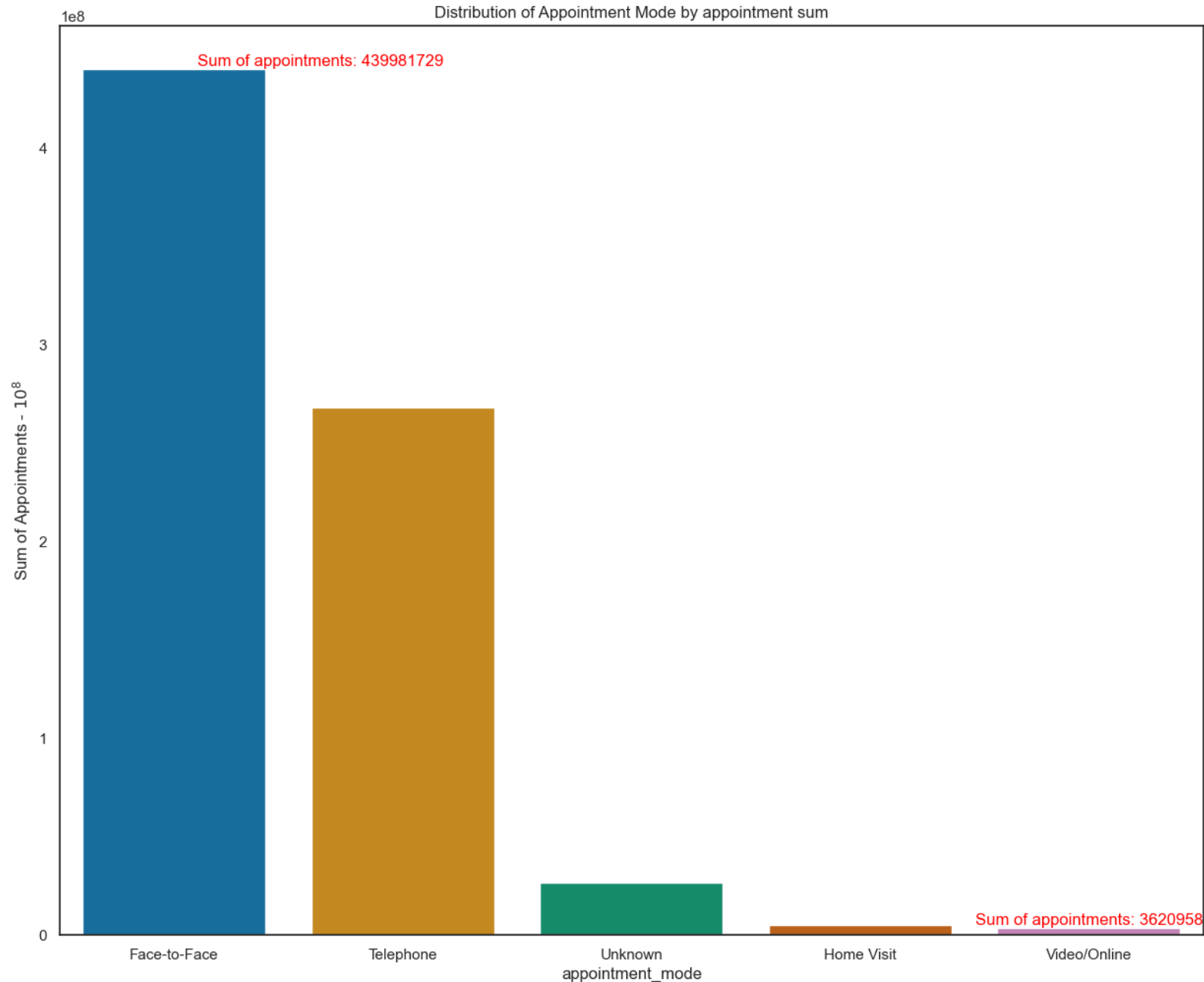
HCP by sum of appointments shows us that in terms of utilisation, 'other practice staff' are almost equal to GP's

The sum of healthcare professional types appointments over time shows us that GP's are responsible for 51% of appointments, Other Practice Staff 46% and Unknown 3%.

From the Metadata: (*Healthcare professional type (HCP type) was incorrectly extracted for some practices from October 2017 onwards. This has led to potential incorrect reporting of some HCP Types, including nurses. The only HCP type currently collected with high enough consistency for publication is GP. Therefore, all other HCP types are reported as 'Other Practice Staff' throughout this release.*)

For further reference :
The occupations grouped as HCP type 'GP' are:
GP registrar, Locum GP and Principal GP

The occupations grouped as HCP type 'other practice staff' are:
Acupuncturist, Chiropodist, Community Psychiatric Nurse, Counsellor, Dispenser, District Nurse, Health Visitor, Interpreter/Link Worker, Osteopath, Other Practice Staff, Physiotherapist and Practice Nurse

It would be very useful, in terms of exploring utilisation to have the data available for each one of the Other Practice Staff so that we might see a spread of utilisation on a more granular level to in turn address capacity.

This could be used to further explore these principles in conjunction with other charts to see which regions, for instance, need which Other Practice Staff the most in which months.

Healthcare Professional Types Over Time
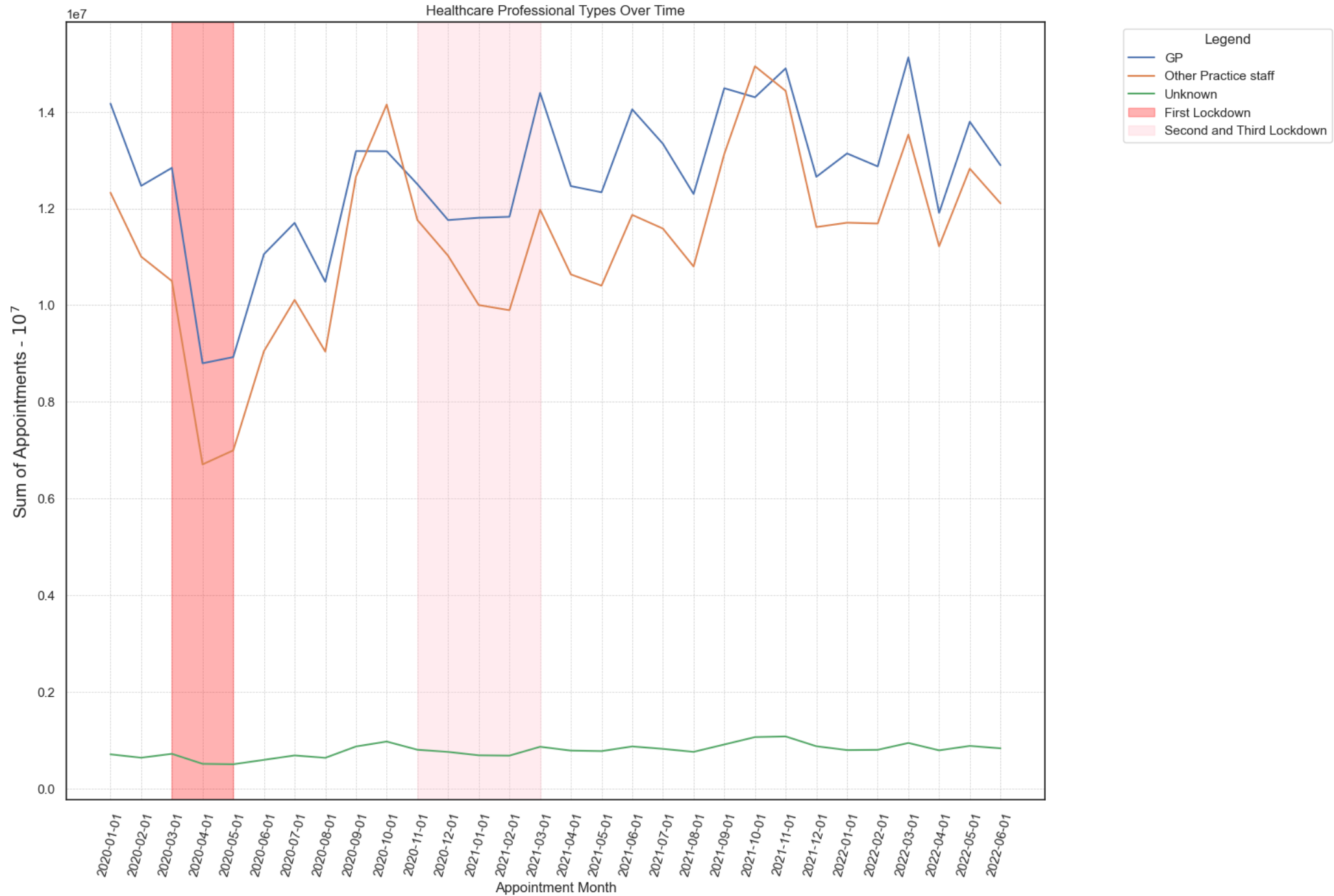
Healthcare professionals sum of appointments by time show us that other Practice Staff are working mostly in parallel with each other with GP's responding to the same dips and the same peaks. Notably there are two large spikes where Other Practice Staff come from far below GP's to overtake them around 2020-10 and 2021-10.

Healthcare professional analysis is important because it is there is real potential to free up resources. In an article published by the NHS Pickering Medical Practice, North managed to reduce their GP appointment waiting times by 47% and here were a few of their insights:
-   "*GPs were collectively spending up to eight hours per week on medication tasks, most of these could have been dealt with by a clinical pharmacist.*" (NHS. (n.d.) [22])
-   "*30% of GP interruptions were inappropriate or could have waited, a further 40% of interruptions could have been dealt with by a more junior member of the clinical team. (NHS. (n.d.) [23])*
-   Overall, initiatives like these, this helped to reduce GP appointment waiting time from 19 to 10 days and provided many individual examples of efficiencies.
        For instance "*The new clinical pharmacist has helped release six hours of GP time per week across the practice. This represents a 90% drop in medication tasks that were previously sent to GPs and are now dealt with by the pharmacist.*" (NHS. (n.d.) [24])
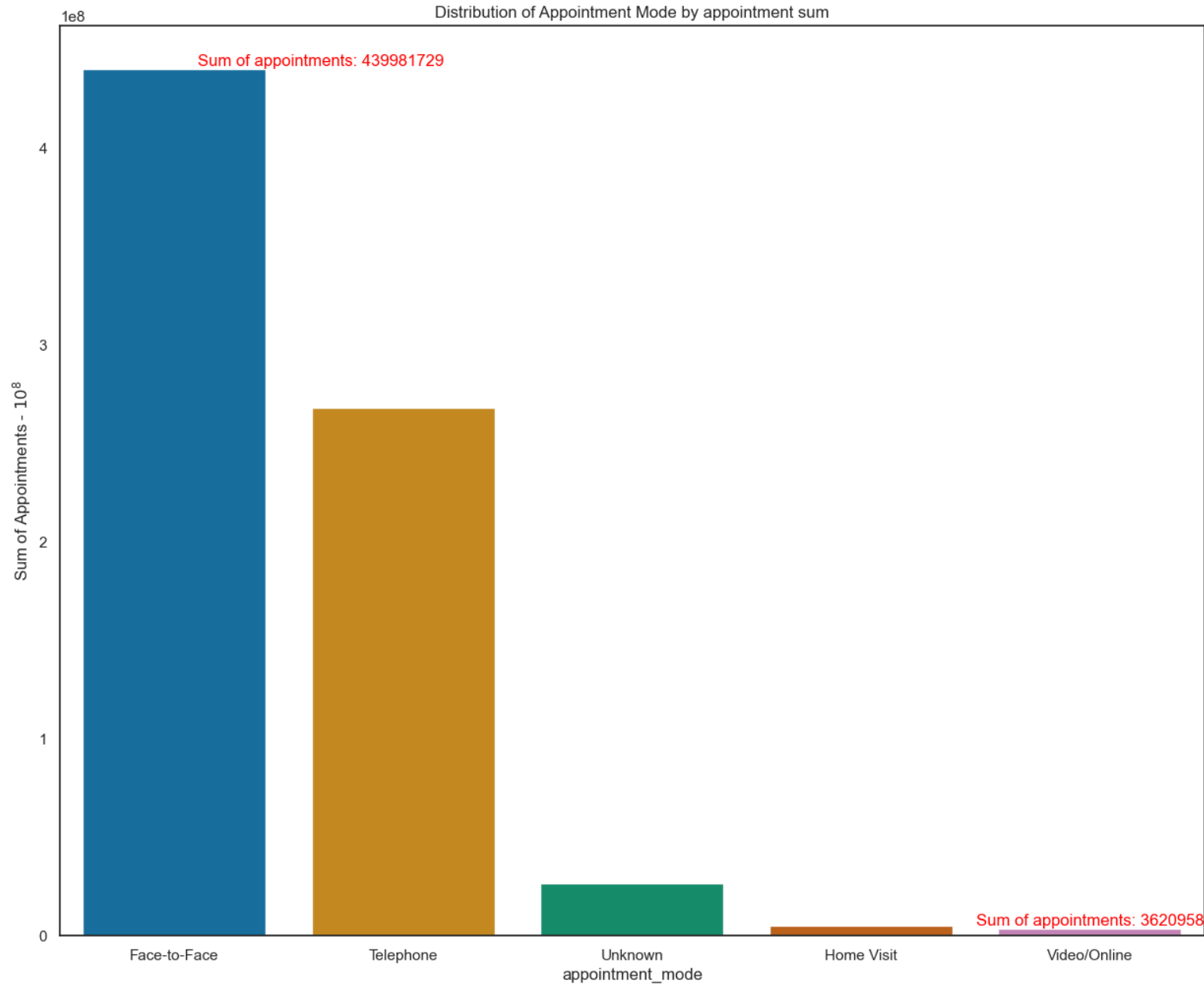
It is clear that other practice staff have a vital role to play in freeing up GP's to offer more in depth and complex services to the public. There has to be a flexibility to the utilisation of the NHS.

3. I - *How are these appointments delivered?*

[22] NHS. (n.d.). NHS choices. https://www.england.nhs.uk/gp/case-studies/routine-gp-appointment-waiting-times-reduced-by-47-pickering-medical-practice-north/
[23] NHS. (n.d.). NHS choices. https://www.england.nhs.uk/gp/case-studies/routine-gp-appointment-waiting-times-reduced-by-47-pickering-medical-practice-north/
[24] NHS. (n.d.). NHS choices. https://www.england.nhs.uk/gp/case-studies/routine-gp-appointment-waiting-times-reduced-by-47-pickering-medical-practice-north/

Distribution of Appointment Mode by appointment sum

Appointment_mode by sum of appointment shows us face to face is the most used appointment mode at 59% compared to telephone's 36%.

Noteworthy again is the Unknown data at just 4% however, 4% is 26,451,305 appointments so no small number

Video/Online comes at 0% around 3,620,958.00 appointments.

Appointment mode trends over time
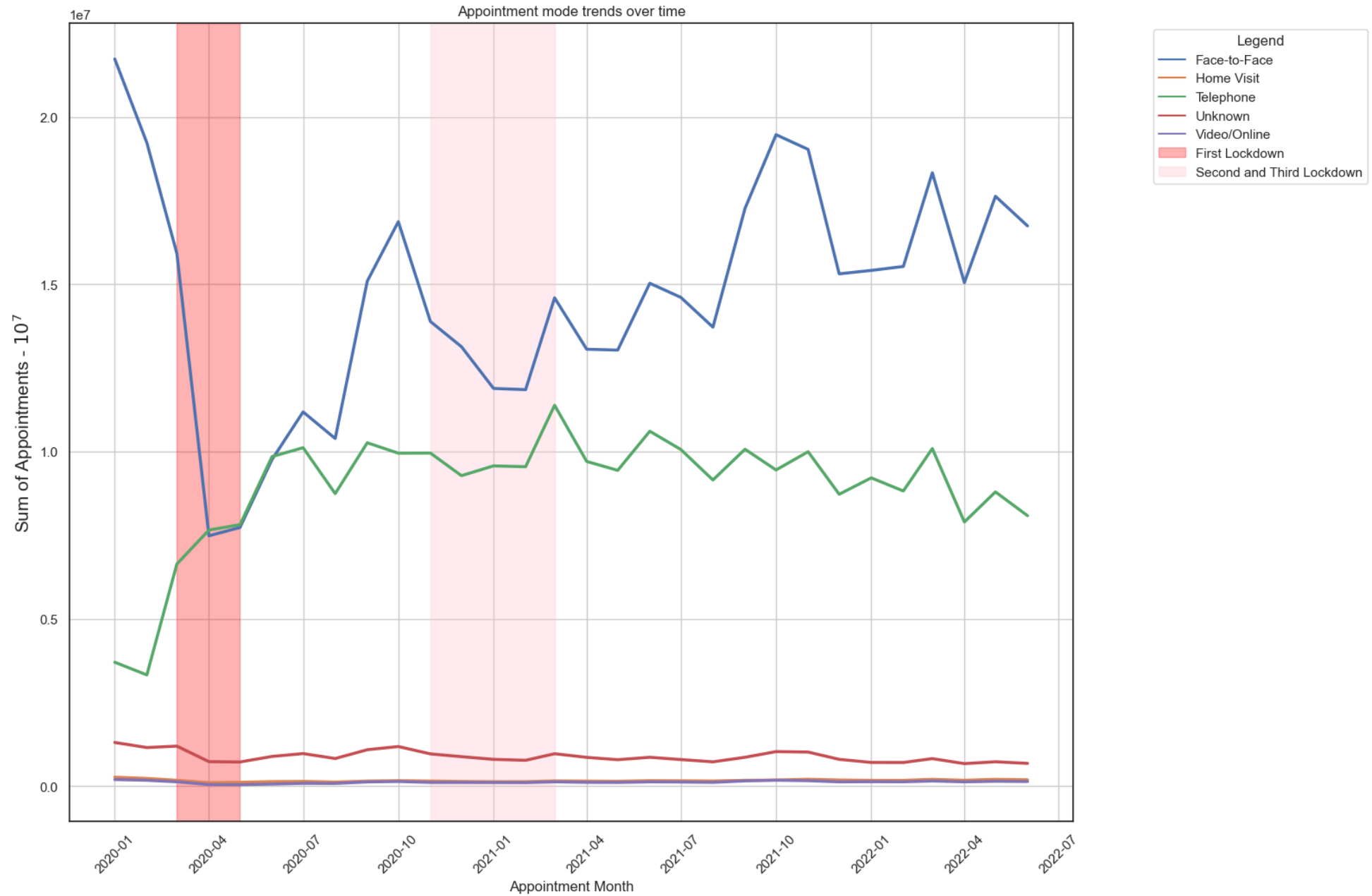
Sum of appointment mode appointments per month - shows us that while there is some fluctuation, face to face comes out on top

Lockdown changes the dynamic of appointments considerably with a sharp decrease in Face-To-Face appointments because of restrictions being met by a large rise in Telephone appointments.

These telephone appointments seem to lose momentum in our dataset after a while as Face-To-Face begins to show a more positive trend, which telephone does not mirror. This perhaps shows a public preference towards Face-To-Face appointments

Notably, Video/Online remains unutilised even with so many months of lockdown.
Returning to Pickering Medical Practice, North where their GP appointment waiting times were reduced by 47% we can gain further insights still:
- Specifically "*only 3% of all appointments were by telephone however, feedback from the whole team suggested this could be higher as some appointments they were dealing with did not require face to face. As telephone appointments are shorter than face to face they would, therefore, be more efficient.*"(NHS. (n.d.) [25])
- That efficiency was measured and proved to be quite significant. "*For every one face to face appointment a GP can deal with two patient calls, freeing up capacity to deal with more patients… There has been a 12% increase in telephone appointments which has reduced the number of patients seeing a GP face to face by 8%.*"(NHS. (n.d.) [26])

If we hypothetically take the calculation that 1 Face-To-Face appointment = 2 Telephone appointments and apply it to our data:
- We can see that if GP's did only telephone appointments they could roughly do another 180 million appointments within the given date range. Of course it's not as simple as that but it does provide a very rough indication of how much more resources can be freed up for GPs

[25] NHS. (n.d.). NHS choices. https://www.england.nhs.uk/gp/case-studies/routine-gp-appointment-waiting-times-reduced-by-47-pickering-medical-practice-north/
[26] NHS. (n.d.). NHS choices. https://www.england.nhs.uk/gp/case-studies/routine-gp-appointment-waiting-times-reduced-by-47-pickering-medical-practice-north/

Figure 23: Doctors' views on 'smarter' working practices.

"In the past two years have you considered or implemented any of the following to adjust your working practices, in order to try and alleviate pressure on workload and capacity?"

Working practice	Considered and implemented	Considered but not implemented	Neither considered nor implemented
Online access to imaging results	66%	15%	19%
Online access to blood test results	75%	12%	14%
Reviewed working practices to gain efficiencies	59%	29%	13%
Allocated tasks previously only undertaken by doctors to other healthcare staff	53%	28%	19%
Increased number of telephone consultations over face-to-face	52%	23%	25%
Used online technology to gain efficiencies in administration	51%	28%	21%
Implemented telephone triage	46%	29%	25%
Used online technology to gain efficiencies in service delivery	44%	32%	24%

Source: General Medical Council (GMC). The state of medical education and practice in the UK. December 2018.

This chart from the General Medical Council shows us the gradual adoption of modes that aren't face-to-face and also links to the further inclusion of other healthcare staff.

There are clearly opportunities and demand for digitising healthcare and the NHS has acknowledged this with a progressive plan in place.

NHS. (2019) [27])

[27] NHS. (2019). NHS Long Term Plan. https://www.longtermplan.nhs.uk/wp-content/uploads/2019/01/nhs-long-term-plan-june-2019.pdf

The NHS acknowledges this in a wider sense too "*by 2028, outcomes for individuals, communities and whole populations are being improved by digital tools including:*
- *risk-based screening*
- *digital health checks*
- *access to online peer support*
- *virtual consultations*
- *at-home diagnostics.*"   Department of Health and Social Care. (2022, June 29) [28])

The NHS can't lose sight of this plan and must keep pushing forward, particularly with online appointments which are at 0%. With each new generation being born into a greater understanding and adoption of technology, the NHS must do the same. Which will not only serve to meet the newer generations at an appropriate level of technology but will also optimise the NHS and make it more efficient as we have seen.

[28] Department of Health and Social Care. (2022, June 29). A plan for Digital Health and Social Care. GOV.UK. https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care

Appointment Mode and HCP Type
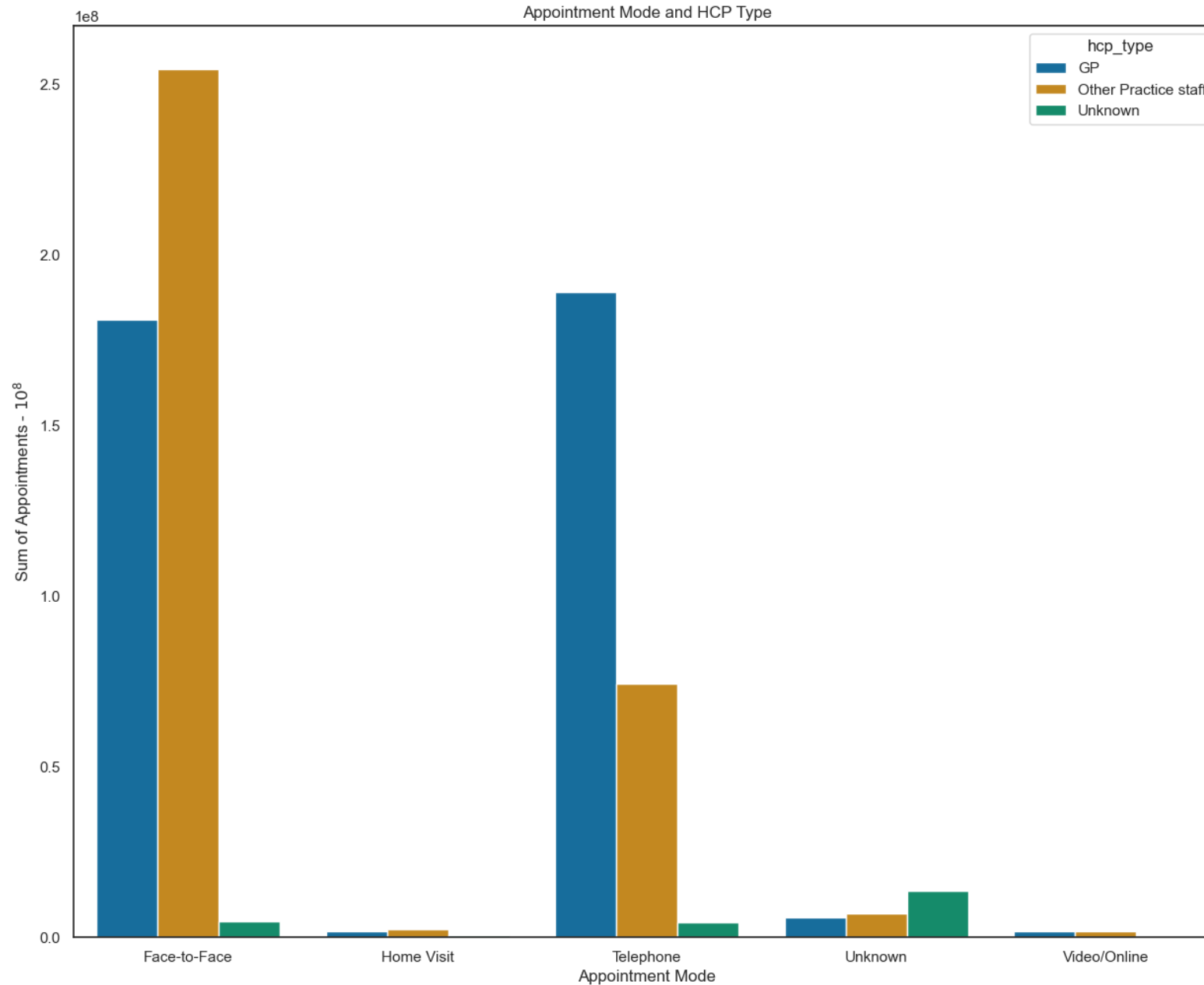
By comparing the healthcare professional type with the appointment mode we can see that different practitioners have different roles within then networks.

This furthers out previous analysis suggesting that Other Practice Staff can shoulder most of the Face-To-Face appointments 58% allowing GP's 41% to be most efficient with other modes such as telephone appointments.

3. **J** - *How pt 2 are these appointments delivered?*

Distribution of Actual Durations by appointment sum
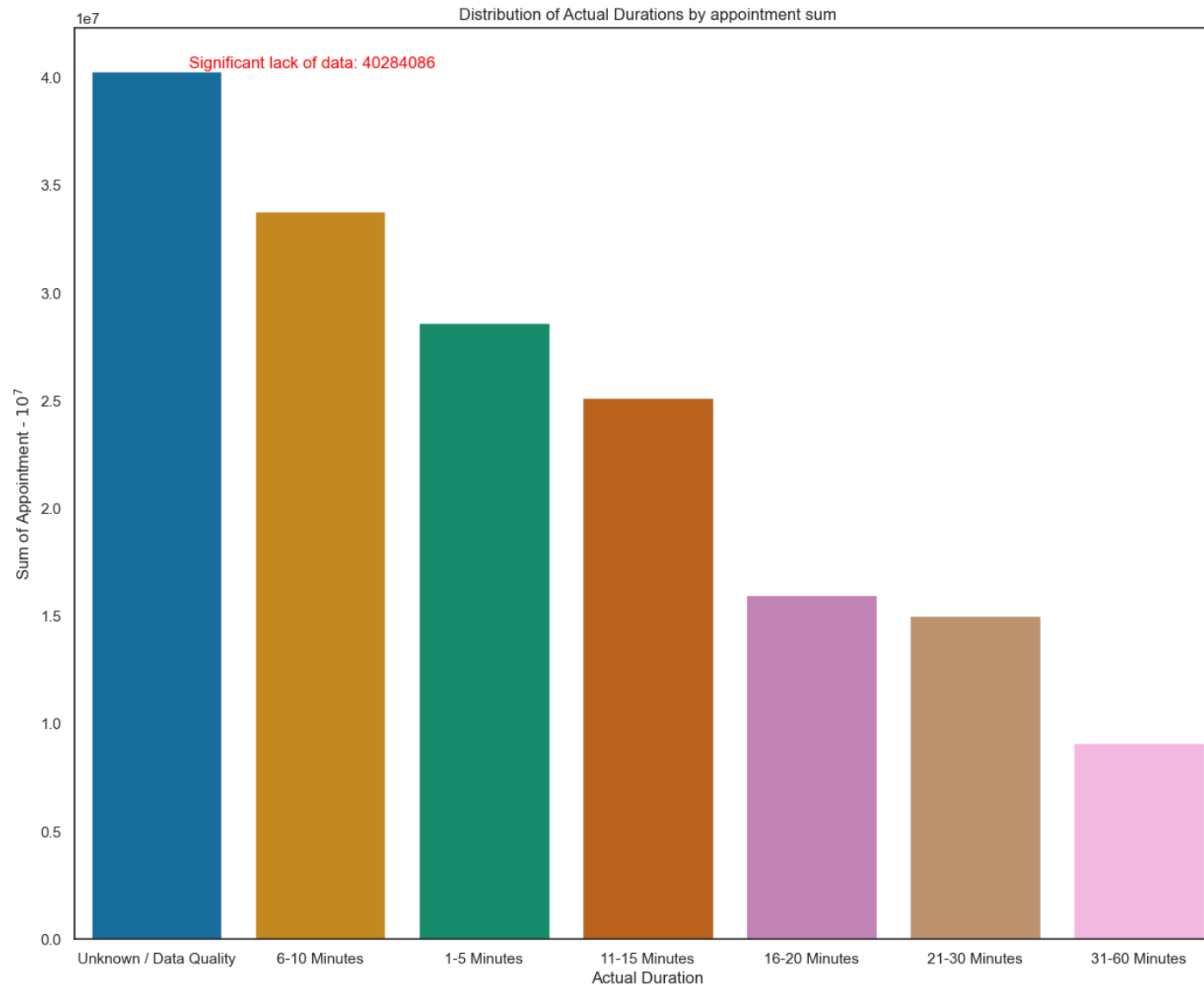
Actual Duration by sum of appointment shows a problematic amount of data quality issues which should be mentioned first.

That's 24% of appointments misreported/unreported because of data quality, 40,284,086 appointments total.

The metadata allows us to see that "*Any appointments with a null duration or a duration of less than one minute or greater than 60 minutes have been grouped into an 'Unknown / Data Quality Issue' category within this publication.*"
- This could be a significant contributing factor

Bar Chart of recommended 15 mins appointment time
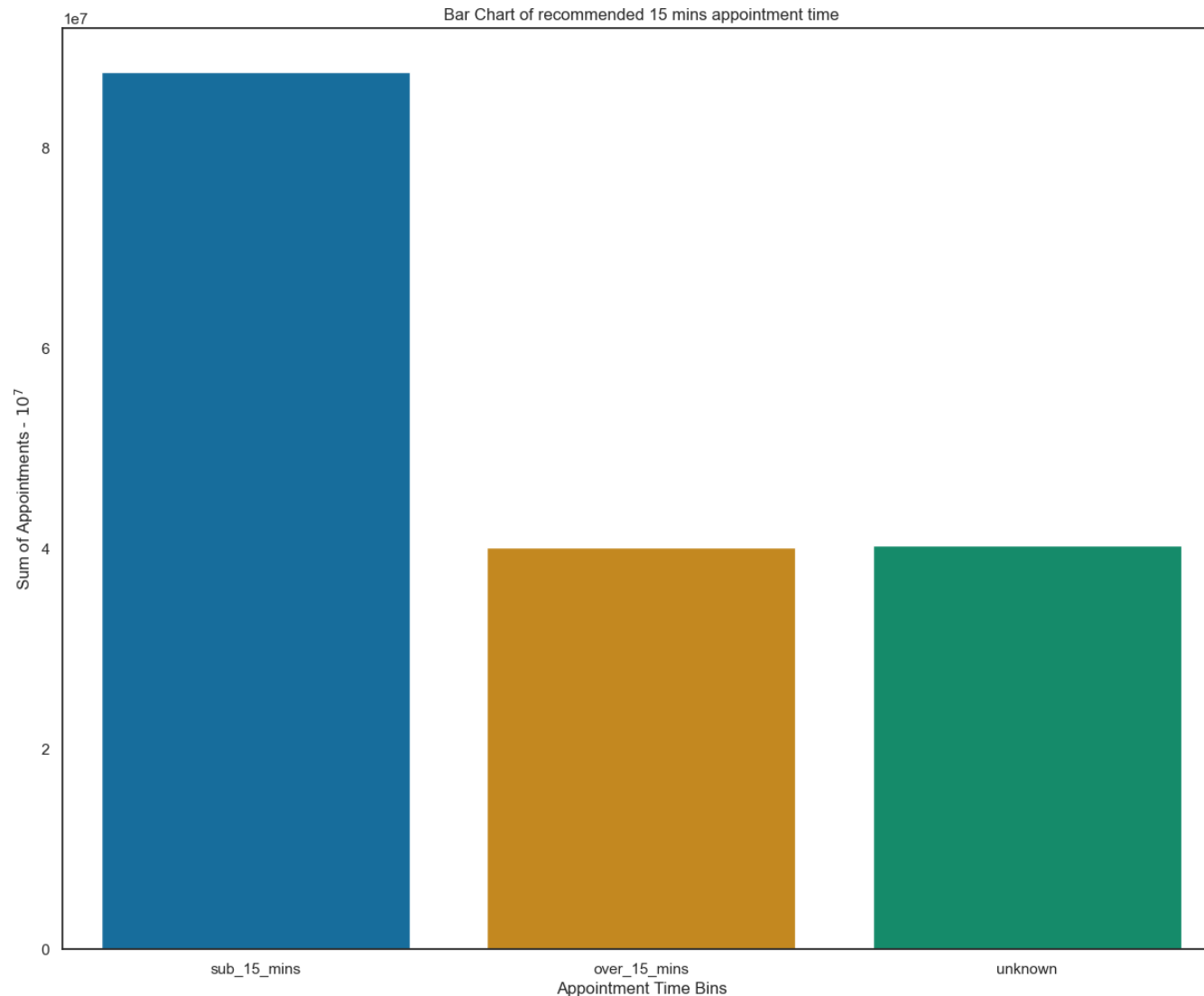
Ideally it would be good to have a standard metric for appointment duration provided but that is a difficult task. However the British Medical Association has a suggested figure of 15 minutes . As such, this will be our metric.

Just over half, 52% of appointments meet the BMA's standard of sub 15 minutes, that's 87,562,562.00 appointments Appointments over 15 minutes is 24%, 40,134,044.00 appointments. Very similar to the unknown column.

"*By extending appointments to 15 minutes, practices can reduce the need for repeated consultations with patients whilst still preserving quality of care and patient satisfaction.*" ([British Medical Association](#) (2023) [29])

[29] here Pushing back on inappropriate workload - British Medical Association. (2023). https://www.bma.org.uk/advice-and-support/gp-practices/managing-workload/pushing-

If GP's are spending too much time on patients during appointments then that can lead to inefficiencies within the staffing network.

As we have seen with the types of Health Care Professional, there are ways to maximise efficiencies by freeing up GP's from appointments they don't need to be at or might take too long if Face-To-Face.

There is an argument as well that Time_Between_Book_And _Appointment is an indicator of capacity too but there are so many variables that go into the metric: quality of data reporting varying between icb's, season peaks and troughs etc.

3. **K** – *Twitter analysis*

The infodemiology, *"the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy"* , Gohil, S., Vuik, S; Darzi, A. (2018, April 23) [30]). that surrounds social media and twitter should not be an ignored resource. Dr Felix Greaves from Imperial College London's School of Public Health argues in a study for Imperial that "*Hospitals should be listening to what patients say on Twitter. It provides a new route for the NHS to understand what patients experience in hospital, and for hospitals to improve the care they provide.*" Davenport, F. (2014, May 2) [31])

Worth noting that "*The volume of tweets sent to hospitals, almost 200 000 in a year, shows that there is a substantial body of information to be listened to. For the median hospital, this represents two to three tweets a day—which is feasible to monitor and respond to.*" Gohil, S., Vuik, S; Darzi, A. (2018, April 23) [32])

back-on-inappropriate-workload
[30] Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). Sentiment analysis of Health Care Tweets: Review of the methods used. JMIR public health and surveillance. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/
[31] Davenport, F. (2014, May 2). Patients' twitter messages give useful insights into NHS hospital care: Imperial News: Imperial College London. Imperial News. https://www.imperial.ac.uk/news/148087/patients-twitter-messages-give-useful-insights/
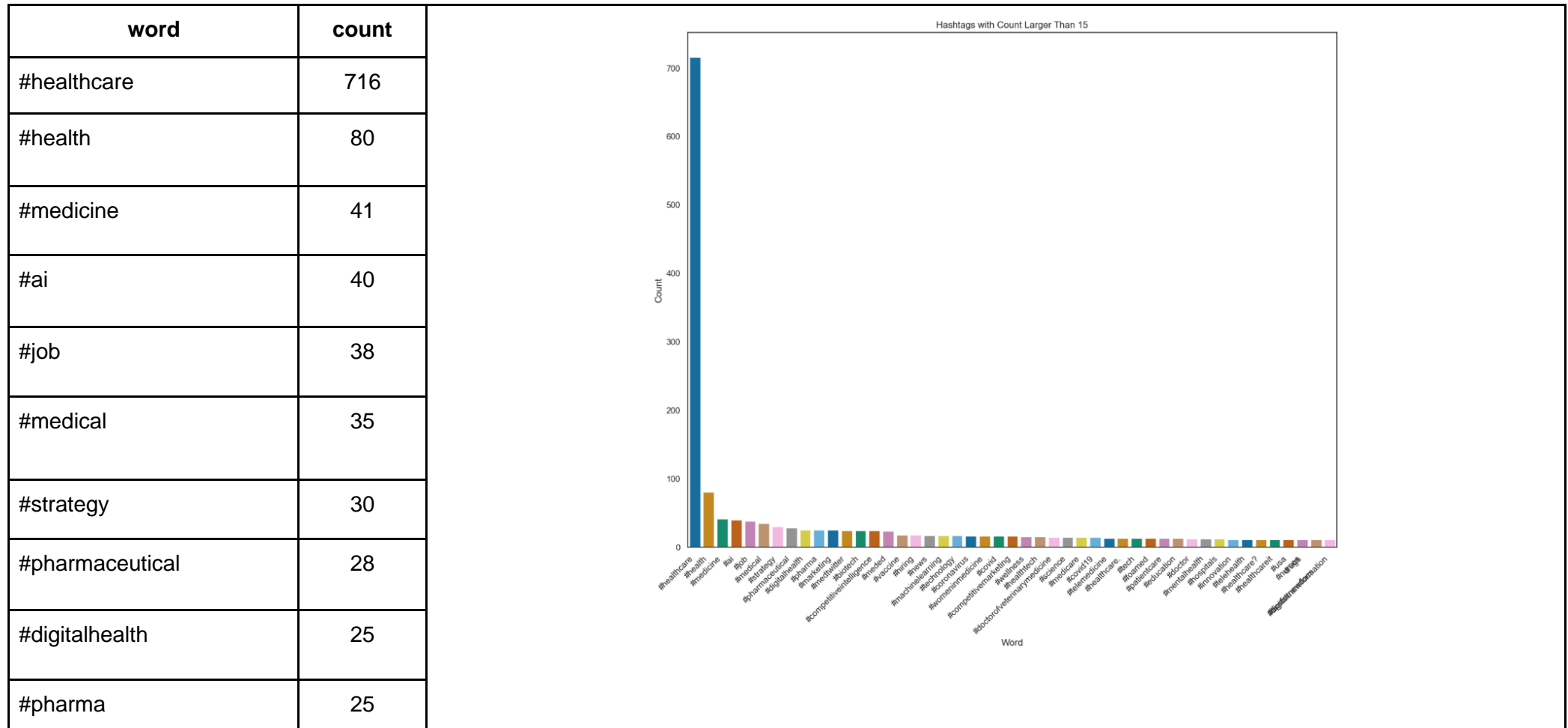[32] Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). Sentiment analysis of Health Care Tweets: Review of the methods used. JMIR public health and surveillance. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/

How can the NHS utilise tweets to provide feedback to stakeholders

With the data provided to us from Twitter, the most effective way for our team to provide feedback is an analysis of hashtags. By scraping through the available tweets we have filtered the hashtags and ranked them in order of count. Here are the top 10 hashtags that have a count larger than 10 in a table on the left.

word	count
#healthcare	716
#health	80
#medicine	41
#ai	40
#job	38
#medical	35
#strategy	30
#pharmaceutical	28
#digitalhealth	25
#pharma	25

Hashtags with Count Larger Than 15

This graph above of tweets with a count larger than 10 is not designed for careful viewing rather just to illustrate just how large the gulf between #healthcare and the rest of the hashtags is.
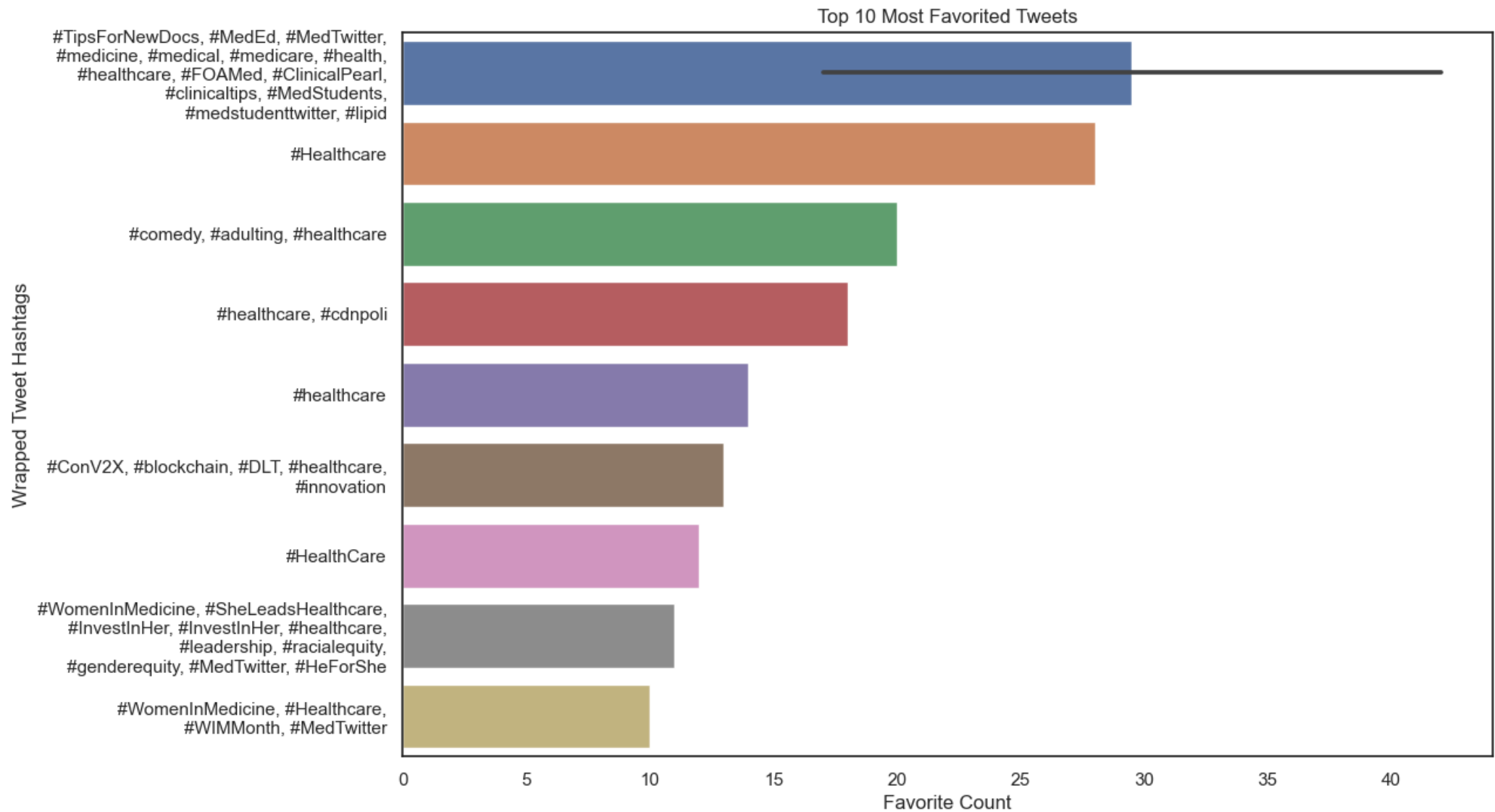
Hashtags with Count Larger Than 15
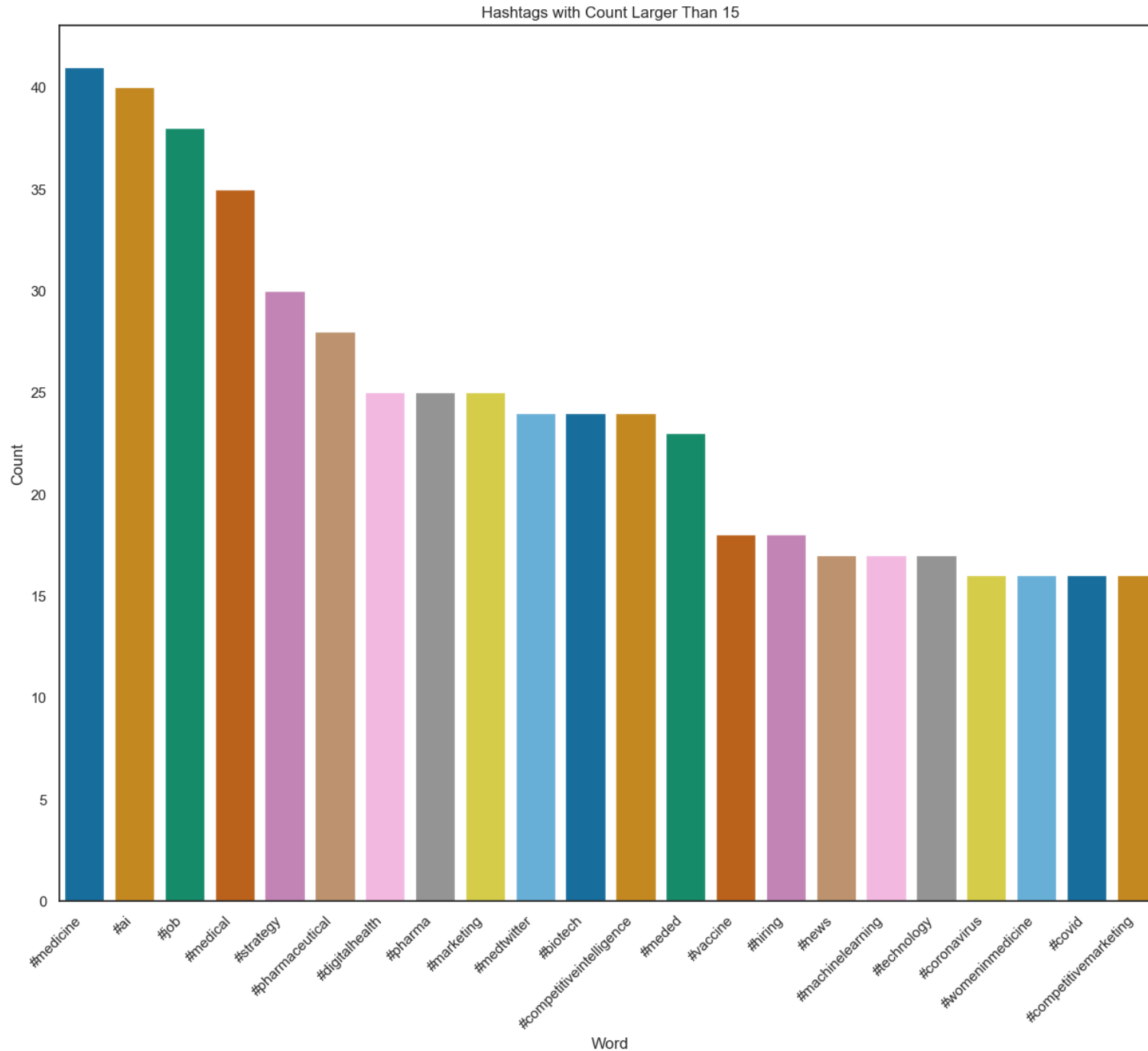
In this graph, tweets have been filtered to ones that have a count of over 15 and we exclude the two over represented hashtags of "#healthcare" and "#health" .

This allows us to visualise the data and any correlations much more clearly.

This graph allows us to see the most common hashtags in relevant tweets and what topics might resonate with relevant users. Key takeaways include:
- A focus on **#technology** with #ai, #biotech, #machinelearning and #digitalhealth
- Hashtags relating to **#job**(s) appears too with #marketing, #hiring and #womeninmedicine
- **#covid**, #vaccine and #coronavirus surprisingly haven't received as much engagement as the focus on technology
- The more general hashtag **#Healthcare** dominates - with #health, #medicine, #medical, #pharma, #pharmaceutical

It is useful to dive deeper into engagement to see what the most engaged tweets look like and whether there is further insight to be had. In this instance, we will use most favourites as the metric for most engaged.

Top 10 Most Favorited Tweets

Below is a table for clearer viewing.

Tweets	Summary	Hashtags	Favourites
"Lipid-Lowering Drugs #TipsForNewDocs #MedEd #MedTwitter #medicine #medical #medicare #health #healthcare #FOAMed #ClinicalPearl #clinicaltips #MedStudents #medstudenttwitter #lipid Credit: @medics_AbuSaif https://t.co/biwq6A1yVD"	Link is to an image highlighting the roles that statin,, pcsk9 inhibitors, nicotinic acid and fibrates play in lipid-lowering drugs. Abu Saif is a medical student form Saudi Arabia. Who illustrates summaries of medical themes in Arabic and English	#TipsForNewDocs, #MedEd, #MedTwitter, #medicine, #medical, #medicare, #health, #healthcare, #FOAMed, #ClinicalPearl, #clinicaltips, #MedStudents, #medstudenttwitter, #lipid	42
"You ready for $JCO @_JennyCo ♡ #Healthcare data powered by @Conste11ation $DAG ◌"	Jennyco is a cryptocurrency coin Conste11ation is a crypto economy company that aims ot link business to blockchain technology	#Healthcare	28
"How health insurance works 😂 #comedy #adulting #healthcare https://t.co/ciksdeoAkb"	1 minute comedy short about American health care insurance hosted on Twitter	#comedy, #adulting, #healthcare	20

Do these tweets add value to the overall project

Here in lies the limitations of using hashtag analysis of Twitter on this project. These three tweets as individual pieces of information don't individually reveal anything useful to the NHS and don't address the two main business questions:

- Are there adequate staff and capacity in the networks?
- What are the actual utilisation of resources?

Beyond that there are real ethical concerns with how we can use Tweets, especially using the full unedited body of a tweet. A very detailed report on using Twitter in research by UCL (UCL. (2020, June 18) [33]) goes into good detail on the subject,

As the report argues, seeking consent from Twitter and every individual involved is impractical. However, "*researchers must inform Twitter of their intentions at the outset (and if these change) in order that it can approve the proposed work*." (UCL. (2020, June 18) [34])

In order to circumvent individual consent I would ensure that the data collected does not contain personally identifiable information (PII) and that the NHS should respect user privacy and adhere to relevant data protection laws, such as GDPR in Europe.

In real terms this means exclusively just scraping Tweets for keywords or hashtags from the "tweet_full_text" and "tweet_entities_hashtags" columns while only collecting impersonal location and time related data.

As Twitter writes on their own website "*aggregate analysis of Twitter content that does not store any personal data (for example, user IDs, usernames, and other identifiers) is permitted, provided that the analysis also complies with applicable laws and all parts of the Developer Agreement and Policy.*" (Twitter (n.d.) [35])

Even with these restrictions on the data in place it's "*worth noting that social media data is almost impossible to anonymise because internet search engines can resolve the content back to its source (and thus the identity of the participant).*" (UCL. (2020, June 18) [36]) Never the less, with strict restrictions and proper adherence to the rules and regulations of GDPR I believe there is value to be had from Twitter scraping.

So, with a macro lens, a wider analysis of hashtags has allowed us to highlight several areas of interest and engagement among twitter users as we saw with the impersonal hashtag value counts analysis.

But these tweets in particular do not add value to the project.

---

[33] Ucl. (2020, June 18). Research registration guidance. Data Protection. https://www.ucl.ac.uk/data-protection/guidance-staff-students-and-researchers/research/research-registration-guidance
[34] P. 6, (2020, June 18). Research registration guidance. Data Protection. https://www.ucl.ac.uk/data-protection/guidance-staff-students-and-researchers/research/research-registration-guidance
[35] Twitter. (n.d.). More on restricted use cases - twitter developers | twitter developer platform. https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases
[36] *P. 7*, (2020, June 18). Research registration guidance. Data Protection. https://www.ucl.ac.uk/data-protection/guidance-staff-students-and-researchers/research/research-registration-guidance

How do we use tweets to add value to the project?

By using the same wider scraping of hashtags and by compiling several hashtags into appropriate themes, the NHS can create relevant categories to analyse. Remember the categories we have seen already from our analysis:
- The technology theme
- The job theme
- The Covid19 theme
- The healthcare theme

A recent study that analysed tweets sent to hospitals within the English NHS in order to gauge patients experience of healthcare[37] discovered useful and relevant insights.

In terms of content analysis the team behind the study "*identified six key themes from the tweets examined: (1) quality (2) fundraising activities, (3) health information, (4) organisational or practical information about the hospital, (5) promotional messages and (6) messages to patients receiving care. Within these themes, some were divided into further categories.*" Gohil, S., Vuik, S.; Darzi, A. (2018, April 23) *38*)

Our current project focuses on the two main questions:
- whether there were adequate staff and capacity in the networks.
- what the actual utilisation of resources was.

As such let's explore sentiment analysis around "capacity". We have chosen the keywords = ['wait', 'slow', 'fast', 'staff', 'quick', 'NHS'] but this is only for demonstration purposes and entirely customisable. Please check the attached Jupyter notebook for the code and the ability to customise searches and download a .csv if needed.

---

[37] The Design of the paper- A mixed methods study including a quantitative analysis of all 198 499 tweets sent to English hospitals over a year and a qualitative directed content analysis of 1000 random tweets. Twitter sentiment and conventional quality metrics were compared using Spearman's rank correlation coefficient
[38] Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). Sentiment analysis of Health Care Tweets: Review of the methods used. JMIR public health and surveillance. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/

Hashtag(s)
1. #healthcare, #Healthcare, #speechpathology 2. #healthcare 3. #fitness, #beuty, #healthcare, #fitnesslife, #beauty, #beautytips, #beaut 4. #Healthcare, #Birmingham

Here's a sample of hashtags scraped from the Tweets using the previously mentioned keywords. A quick run through of the exported .csv shows us that there really is a need to focus on providing relevant Twitter data both in terms of time and geography as many of the provided Tweets are based in North America

However, this targeting searching and sorting the Tweets to relevant locations is feasible. Twitter. (n.d.) [39]) The initial .csv file provided with the Tweets does not have this level of detail available but this would be possible to obtain.

One other method of filtering would be to obtain Tweets that use NHS trust's and facilities directly tagged in their Tweets. This would also ensure relevance.

By monitoring these themes of keywords and hashtags within a certain geography we can analysing the reach, frequency and engagement of relevant tweets to the NHS. Please see Table 2 of the study for examples of these tweets and sub categories. Gohil, S., Vuik, S.; Darzi, A. (2018, April 23) [40])

Conclusions

The main area of data collection and analysis that has consistent value and could improve future efforts is through further sentiment analysis.

---

[39] Twitter. (n.d.-a). Filtering tweets by location | docs | twitter developer platform. https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location
[40] Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). Sentiment analysis of Health Care Tweets: Review of the methods used. JMIR public health and surveillance. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/

Sentiment analysis can provide insights into public perception. The study referenced earlier that found 6 main themes of 1000 tweets directed at NHS England facilities discovered that "*of the 1000 tweets, 47% were positive, 47% were neutral and 6% were negative.*" Gohil, S., Vuik, S.; Darzi, A. (2018, April 23) [41])

Within the theme of healthcare quality, which amounted to 11% of those examined, "*77% were positive, 21% were negative and 2% were neutral.*" Gohil, S., Vuik, S.; Darzi, A. (2018, April 23) [42])

This rings true with this specific example from St George's Healthcare here. "*We use Twitter to glean a better understanding of the public opinion of our services and the NHS as a whole,*" said Colin Wren at St George's. "*The media love a negative story, and often paint the NHS as a cruel run-down institution where targets come first and the patient comes last. Our social media intelligence doesn't support this view – to date, for St George's Healthcare, positive mentions outweigh the negative by 300%.* " Whitelaw, B. (2011, February 16) [43])

By systematically compiling and categorising sentiment data received through tweets. The NHS can analyse positive, negative and neutral sentiments which can guide improvements in communication and address concerns effectively.

It is important to note however that there are several important limitations. As the same study goes on suggest.

"*We are uncertain about the completeness of the information from the Twitter API as there may be filtering methods in place that we are not aware of…*

There are demonstrable biases in the people who use Twitter. Its usage has been shown to be larger in younger groups and minority ethnic populations. However, these groups have often been hard to engage in other participatory healthcare activity, so this might represent an opportunity for wider population engagement with healthcare services….

[41] Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). Sentiment analysis of Health Care Tweets: Review of the methods used. JMIR public health and surveillance. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/
[42] Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). Sentiment analysis of Health Care Tweets: Review of the methods used. JMIR public health and surveillance. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/
[43] Whitelaw, B. (2011, February 16). @NHS: How the NHS uses Twitter. The Guardian. https://www.theguardian.com/healthcare-network/2011/feb/16/nhs-twitter-use-tweets-communication-healthcare

Our analysis is limited to situations where people deliberately targeted messages at hospitals. Many people discussing their experiences using social networks might not do this. People might also be concerned about risk of censure if they make critical comments, and may not identify specific hospitals in these tweets." Gohil, S., Vuik, S.; Darzi, A. (2018, April 23) [44])

In conclusion, there is an opportunity for the NHS to utilise Twitter and gain vital insights into users of both the platform and the National Health Service. Through cleaned, ethical and geographically relevant keyword/hashtag analysis, we can track engagement across users in order to discover insights about capacity and utilisation of services.

However it is good to be aware of limitations. As we have seen from the Tweets provided in the project, higher engagement doesn't mean more reliable information, Twitter's API and sample bias will affect the quality of results and there are serious ethical implications to navigate.

[44] Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). Sentiment analysis of Health Care Tweets: Review of the methods used. JMIR public health and surveillance. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/

Bibliography:

Davenport, F. (2014, May 2). *Patients' twitter messages give useful insights into NHS hospital care: Imperial News: Imperial College London*. Imperial News.
https://www.imperial.ac.uk/news/148087/patients-twitter-messages-give-useful-insights/

Department of Health & Social Care. (2022a). *Direct and indirect health impacts of covid-19 in England: Emerging omicron impacts*. GOV.UK.
https://www.gov.uk/government/publications/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts#category-c-indirect-impacts-of-covid-19-on-population-health-due-to-living-through-a-pandemic-and-restrictions-1

Department of Health & Social Care. (2022b). *Direct and indirect health impacts of covid-19 in England: Emerging omicron impacts*. GOV.UK.
https://www.gov.uk/government/publications/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts/direct-and-indirect-health-impacts-of-covid-19-in-england-emerging-omicron-impacts#category-c-indirect-impacts-of-covid-19-on-population-health-due-to-living-through-a-pandemic-and-restrictions-1

Department of Health and Social Care. (2022, June 29). *A plan for Digital Health and Social Care*. GOV.UK. https://www.gov.uk/government/publications/a-plan-for-digital-health-and-social-care

Gohil, S., Vuik, S., & Darzi, A. (2018, April 23). *Sentiment analysis of Health Care Tweets: Review of the methods used*. JMIR public health and surveillance.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5938573/

Guardian News and Media. (2022, September 21). *No one should wait more than two weeks to see GP, Coffey to say*. The Guardian.
https://www.theguardian.com/society/2022/sep/21/no-one-should-wait-more-than-two-weeks-to-see-gp-coffey-to-say

The King's Fund. (2021, January 14). *"the NHS is now under the most extreme pressure seen in recent history":* https://www.kingsfund.org.uk/press/press-releases/nhs-now-under-most-extreme-pressure-seen-recent-history

lcp. (2023, March 13). *Regions being "left further behind" as drive to cut NHS waiting lists has most impact in the South East*. Lane Clark & Peacock LLP.
https://www.lcp.com/media-centre/2023/03/regions-being-left-further-behind-as-drive-to-cut-nhs-waiting-lists-has-most-impact-in-the-south-east

NHS. (2016). Data Quality - NHS England. https://www.england.nhs.uk/wp-content/uploads/2016/03/local-commissioning-data-quality-guidance.pdf

NHS. (2019a). NHS choices. https://www.england.nhs.uk/2019/01/missed-gp-appointments-costing-nhs-millions/

NHS. (2019b). NHS Long Term Plan. https://www.longtermplan.nhs.uk/wp-content/uploads/2019/01/nhs-long-term-plan-june-2019.pdf

NHS. (2021). NHS choices. https://www.england.nhs.uk/2021/08/nhs-cuts-waiting-times-despite-busy-summer-period/

NHS. (2022, February). Delivery plan for tackling the COVID-19 backlog of elective care. https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2022/02/C1466-letter-delivery-plan-for-tackling-the-covid-19-backlog-of-elective-care.pdf

NHS. (n.d.). *Investment and Impact Fund.* NHS choices. https://www.england.nhs.uk/primary-care/primary-care-networks/network-contract-des/iif/

NHS. (n.d.-a). NHS choices. https://www.england.nhs.uk/gp/case-studies/routine-gp-appointment-waiting-times-reduced-by-47-pickering-medical-practice-north/

NHS. (n.d.-b). NHS choices. https://www.england.nhs.uk/long-read/supporting-general-practice-primary-care-networks-and-their-teams-through-winter-and-beyond/

Non-covid-19 NHS care during the pandemic. The Health Foundation. (n.d.). https://www.health.org.uk/news-and-comment/charts-and-infographics/non-covid-19-nhs-care-during-the-pandemic

NHS. (2022). Delivery plan for tackling the COVID-19 backlog of elective care. https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2022/02/C1466-letter-delivery-plan-for-tackling-the-covid-19-backlog-of-elective-care.pdf

Palfreyman, A. (2022, October 27). *The cost of missed GP appointments*. Envisage Coda. https://www.envisagecoda.co.uk/cost-missed-gp-appointments/

Person, O. (2023, March 29). *Coronavirus (COVID-19) latest insights: Infections*. Coronavirus (COVID-19) latest insights - Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/infections#:~:text=During%20the%20period%20when%20the,in%20January%20and%20February%202022.

Pushing back on inappropriate workload - British Medical Association. (2023). https://www.bma.org.uk/advice-and-support/gp-practices/managing-workload/pushing-back-on-inappropriate-workload

Twitter. (n.d.-a). *Filtering tweets by location | docs | twitter developer platform*. https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location

Twitter. (n.d.-b). *More on restricted use cases – twitter developers | twitter developer platform*. https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases

Ucl. (2020, June 18). *Research registration guidance.* Data Protection. https://www.ucl.ac.uk/data-protection/guidance-staff-students-and-researchers/research/research-registration-guidance

Whitelaw, B. (2011, February 16). *@NHS: How the NHS uses Twitter*. The Guardian. https://www.theguardian.com/healthcare-network/2011/feb/16/nhs-twitter-use-tweets-communication-healthcare