

Data Analysis carried out for National Health Service (NHS), a publicly funded healthcare system in England, to determine reasons behind patient's missing/No-show on GP appointments which is contributing to NHS's overall cost

# Diagnostic Analysis and Presenting Insights using Python

National Health Service (NHS)  
Afshar Sanam

---

## Table of Contents

<b>1. Problem Statement</b> .....	2
<b>2. Analytical Approach</b> .....	3
<b>3. Visualization and Insights</b> .....	4
Service Settings .....	4
Context Type .....	6
National Categories .....	7
Appointment Status .....	8
Question : What is the date range of the provided data sets, and which service settings reported the most appointments for a specific period? .....	9
Q. What is the number of appointments and records per month? .....	10
Q. What monthly and seasonal trends are evident, based on the number of appointments for service settings, context types, and national categories? .....	11
Q. What are the top trending hashtags (#) on Twitter related to healthcare in the UK? .....	12
Q. Were there adequate staff and capacity in the networks? .....	13
Q. What was the actual utilization of resources? .....	14
<b>Patterns and Predictions</b> .....	20
<b>Summary</b> .....	23

## 1. Problem Statement

Data Analysis carried out for National Health Service (NHS), a publicly funded healthcare system in England, to determine reasons behind patient's missing/No-show on GP appointments which is contributing to NHS's overall cost. The reasons for missed appointments need to be analyzed and assessed appropriately instead of penalizing the patients as it would indirectly impact the most vulnerable community. Therefore, the government is seeking for a data-informed approach to deciding how best to handle this approach. The most important Questions to address are:

- a. Has there been adequate staff and capacity in the networks?
- b. What was the actual utilization of resources?

This analysis uses the

- `actual_duration.csv` - Details of appointments made by patients.
- `appointments_regional.csv` - Details on the type of appointments made by patients.
- `national_categories.xls` - Details of the national categories of appointments made by patients.
- `tweets.csv` - Data related to healthcare in the UK scraped from Twitter data sets and
- `metadata_nhs.txt` - Data related to healthcare in the UK scraped from Twitter.

The insights gained from the analysis will inform the NHS, Government & BMA for their decision-making so they can derive the next course of action(s)

## 2. Analytical Approach

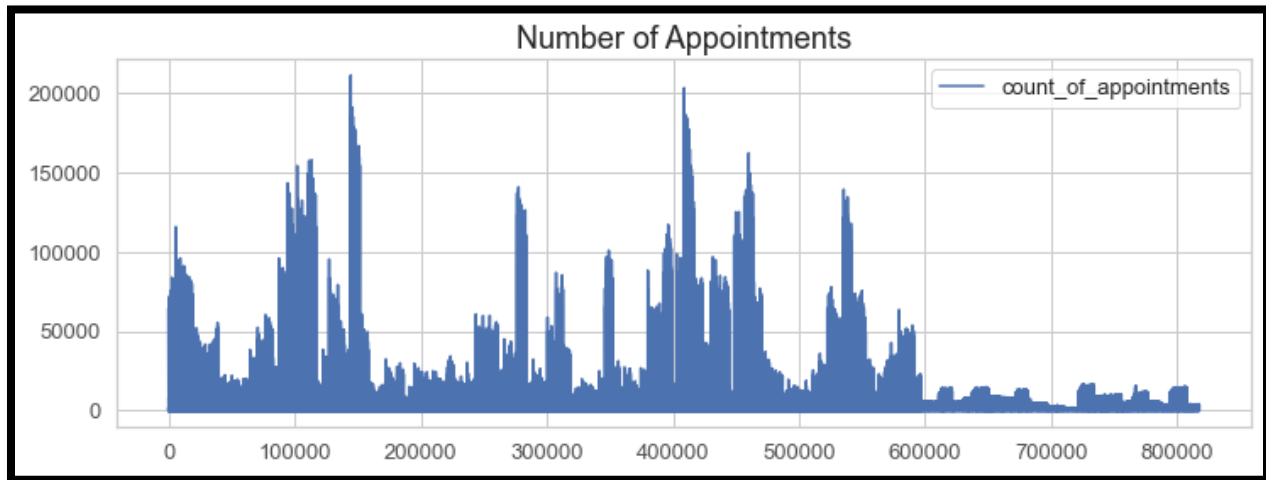
Preparing the workstation by importing all the key libraries (Matplotlib, Seaborn, Numpy, Pandas, warnings, yaml, SafeLoader, twitter, json, plotly.express, plotly.graph\_objs, InteractiveShell etc.) required in cleansing, data wrangling, transformation, visualizing purposes. It was required to filter the warnings to keep the file more readable and had matplotlib inline too and leveraged InteractiveShell to get multiple outputs in the same cell. Applied hacks to make the plots look better, uniform for visualistic representation while using Seaborn libraries.

From data cleansing standpoint, checked for missing values from the data frames by calling isna() method.

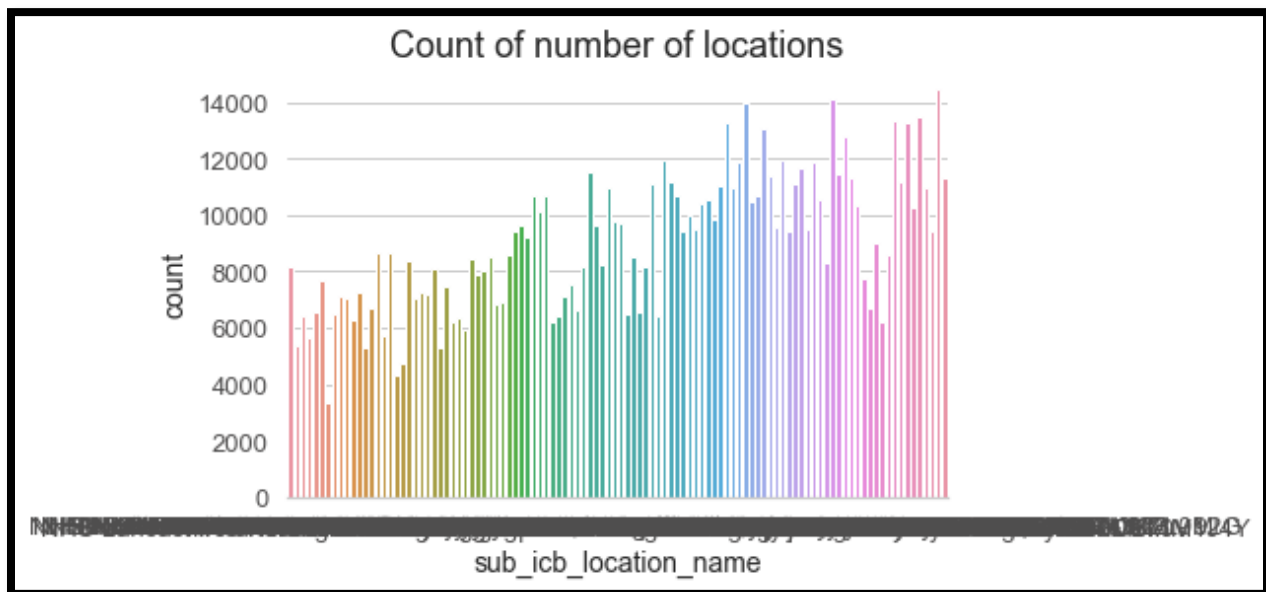
For the graph plots, used different types for better representation of the data. E.g.:

1. Applied Pandas plot for timeseries data representation.
2. For counting the number of locations, service settings, context types, national categories and appointment statuses used CountPlot by varying the x-axis and y-axis to showcase the graph properly.
3. For the date range of the provided data sets, to determine the service settings, national categories, context types based on most appointments found during this date range, leveraged plotly.express and plotly.graph\_objs for plotting the graphs.

### 3. Visualization and Insights



*Time series data of number of appointments*

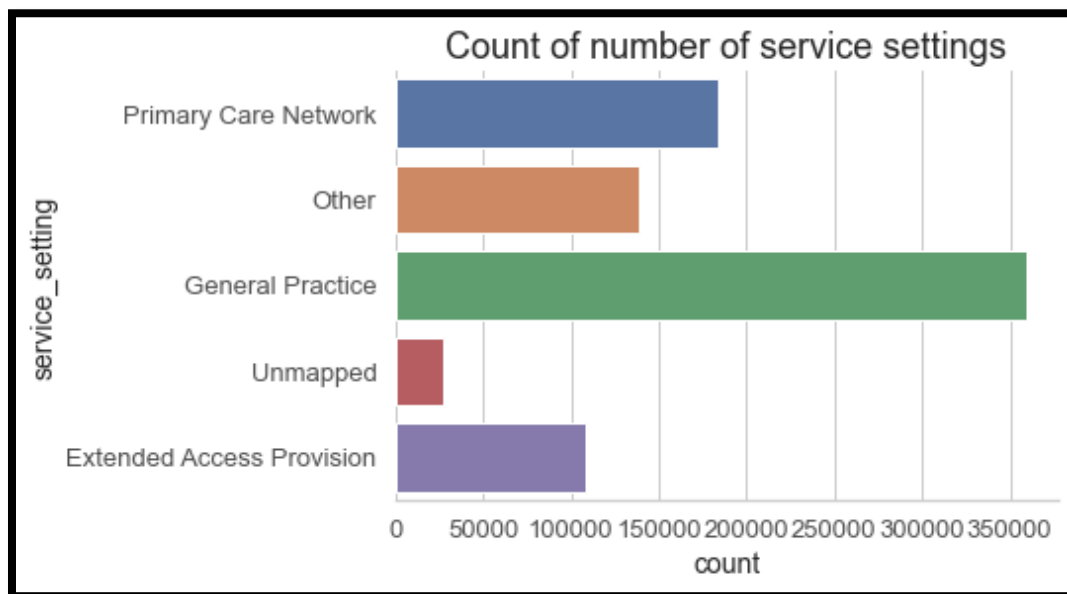


#### Service Settings

The service setting describes the provision of care in which the appointment was delivered. There are four service settings. Where category data is not received, an appointment is assigned to an 'Unmapped' setting.

- Service Setting – Description

- General Practice – Appointments delivered by practice staff activity under the GMS/PMS/APMS contract.
- Primary Care Network
  - Appointments delivered by staff employed or engaged under the Additional Roles Reimbursement Scheme (ARRS). Appointments delivered in fulfilment of the service requirements of the Network Contract Directed Enhanced Service (DES)
  - e.g., Structured Medication Reviews, weekly care home rounds
  - except for Extended Access services, for which the Extended Access Provision service setting should be used.
- Extended Access Provision – Appointments commissioned as part of extended access contractual arrangements.
- Other – Appointments delivered in the practice by another provider.
- Unmapped – No recorded category against an appointment slot type. This could be due to an error receiving the data, or an appointment type has not been mapped.

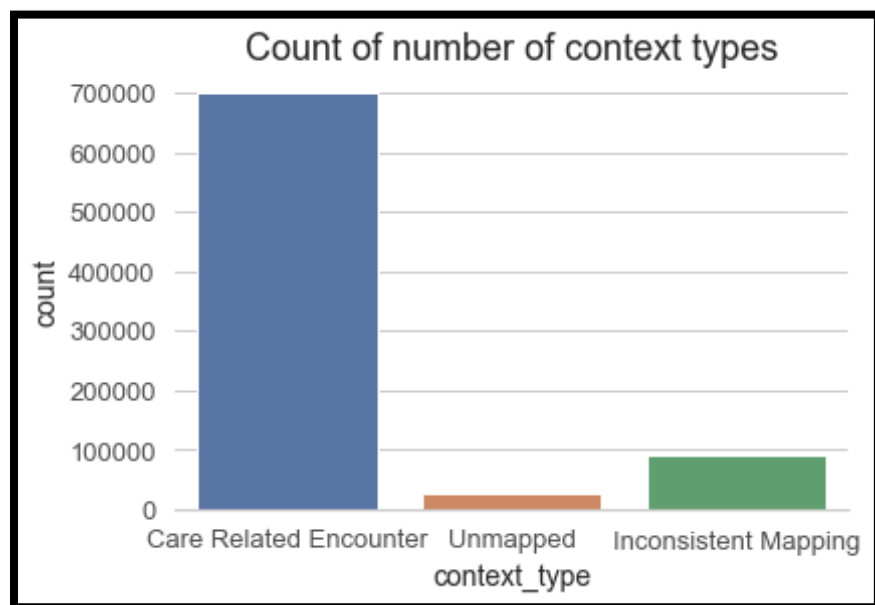


General Practice	359274
Primary Care Network	183790
Other	138789
Extended Access Provision	108122
Unmapped	27419

## Context Type

The National Categories are subdivided into three 'Context Types'. The context describes whether an appointment is an encounter relating to direct patient care, or an activity undertaken as part of patient care where the patient is not involved. Where category data is not received, an appointment is assigned to an 'Unmapped' context type.

- Context Type – Description
- Care Related Encounter – Categories within this context type involve the patient and can be any modality, e.g., face-to-face appointments, telephone consultations, video and online consultations, home visits etc.
- Inconsistent Mapping – Appointment types that have been mapped but not to a Care Related Encounter are classed as Inconsistent Mapping. Appointments under this context type conflict with the description of an appointment. Work is ongoing to improve the consistency of mapping.
- Unmapped – No record of a category against an appointment slot type. This could be due to an error receiving the data, or an appointment type has not been mapped.



Care Related Encounter	700481
Inconsistent Mapping	89494
Unmapped	27419

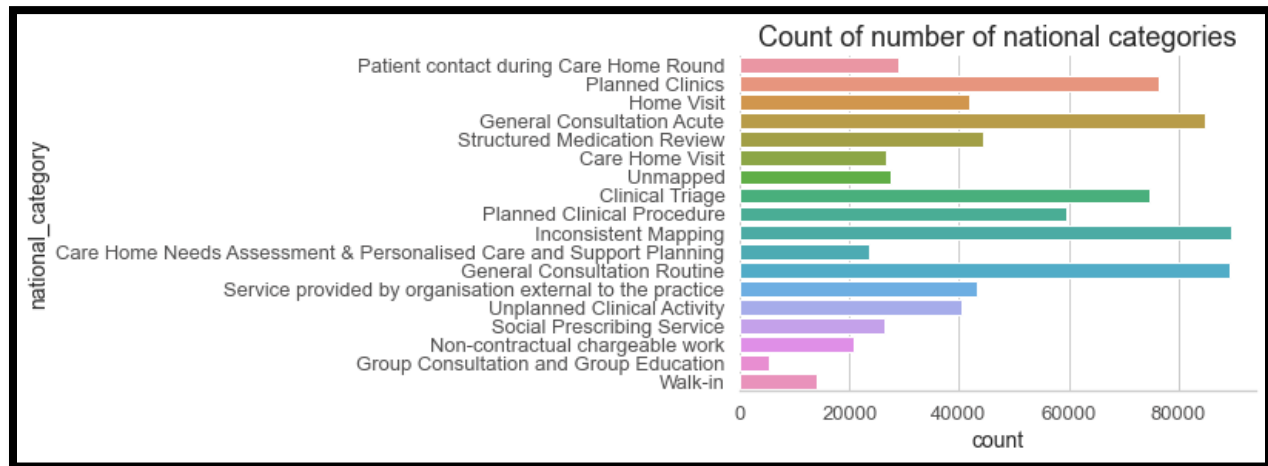
## National Categories

Within each context type, there is an option for practices to choose a 'Does Not Fit' category. This is used only when a practice is unable to place an appointment type within one of the predefined categories. These appointments are assigned to the 'Inconsistent Mapping' context type with further analysis planned to establish whether new categories in future releases are needed, or additional assistance and guidance is required for mapping.

- Context Type
  - National Category
- Care Related Encounter
  - General Consultation Acute
  - General Consultation Routine
  - Planned Clinics
  - Planned Clinical Procedure
  - Unplanned Clinical Activity
  - Walk-in, Clinical Triage, Home Visit, Care Home Visit
  - Group Consultation and Group Education
  - Structured Medication Review
  - Patient contact during Care Home Round
  - Care Home Needs Assessment & Personalized Care and Support Planning
  - Social Prescribing Service
  - Service provided by organization external to the practice.
  - Non-contractual chargeable work



- Inconsistent Mapping
  - Care Related Encounter but does not fit into any other category.
  - Appointment types not mapped to a Care Related Encounter



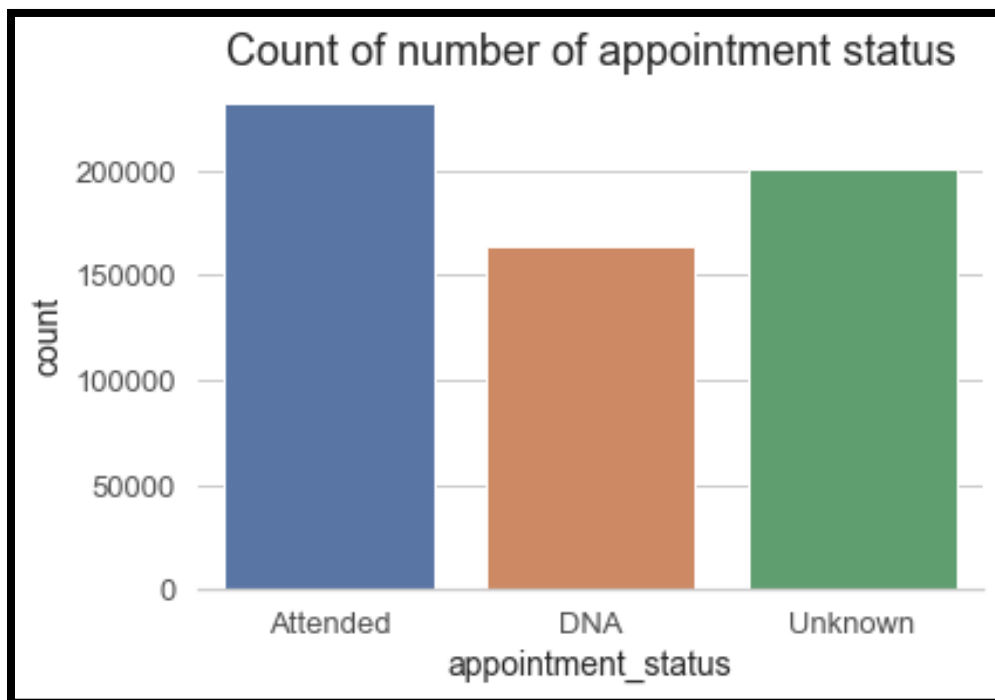
Inconsistent Mapping	89494
General Consultation Routine	89329
General Consultation Acute	84874
Planned Clinics	76429
Clinical Triage	74539
Planned Clinical Procedure	59631
Structured Medication Review	44467
Service provided by organisation external to the practice	43095
Home Visit	41850
Unplanned Clinical Activity	40415
Patient contact during Care Home Round	28795
Unmapped	27419
Care Home Visit	26644
Social Prescribing Service	26492
Care Home Needs Assessment & Personalized Care and Support Planning	23505
Non-contractual chargeable work	20896
Walk-in	14179
Group Consultation and Group Education	5341

## Appointment Status

The appointment status shows whether the appointment is available, booked, has been attended by the patient (or not), or has been cancelled. This status changes over time, for example when a patient makes a booking, the appointment status changes from 'Available' to 'Booked'. This release reports the final status of each appointment.

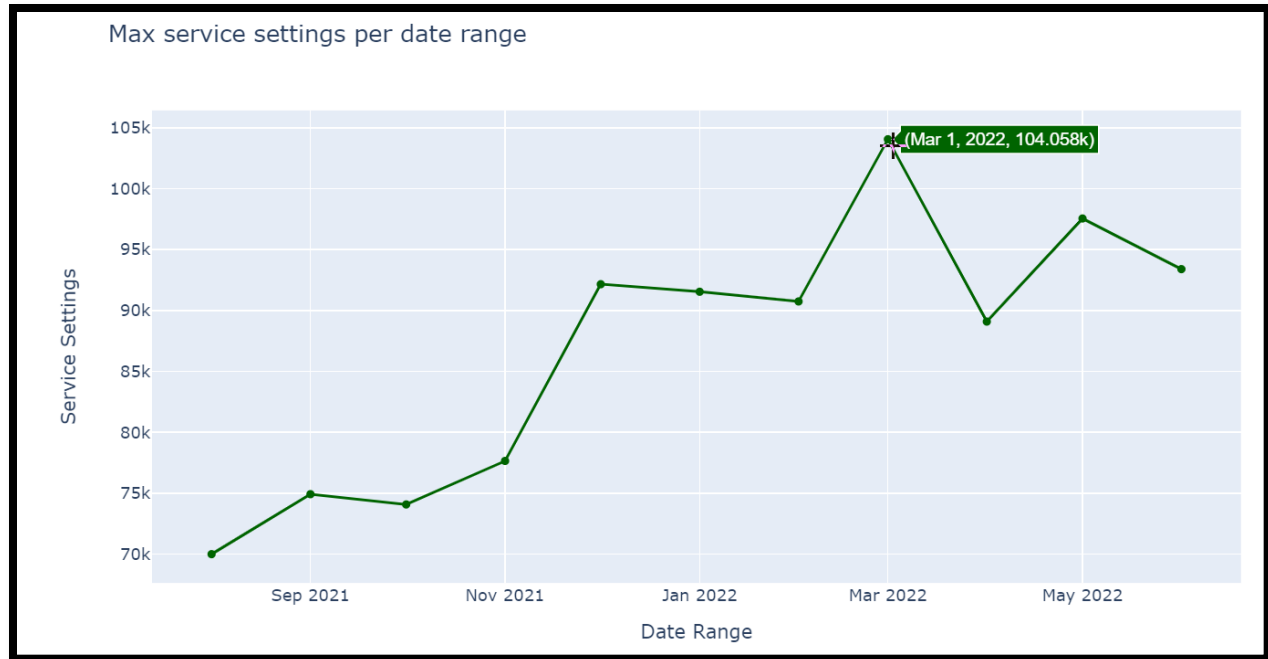
In most cases, at the time of the appointment a 'Booked' status will be changed to either 'Attended' or 'DNA' (Did Not Attend). In some cases (3%–6% of monthly appointments), the final status of an appointment remains as 'Booked'. It is not known from the data whether or not these appointments actually attended. For this reason, the status of these appointments is reported as 'Unknown'.

Due to an issue with the data collection, DNA appointments were not captured correctly after June 2018 and are under-reported until and including November 2018 for all practices using the TPP SystemOne system.



Attended	232137
Unknown	201324
DNA	163360

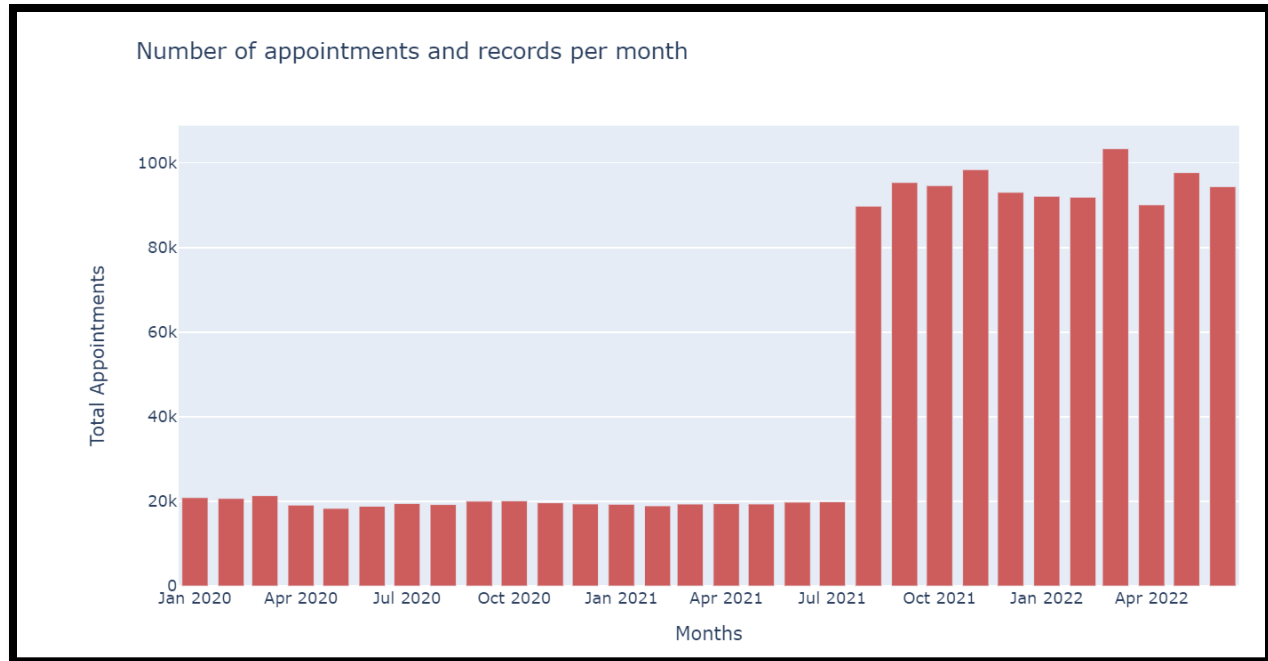
Question : What is the date range of the provided data sets, and which service settings reported the most appointments for a specific period?



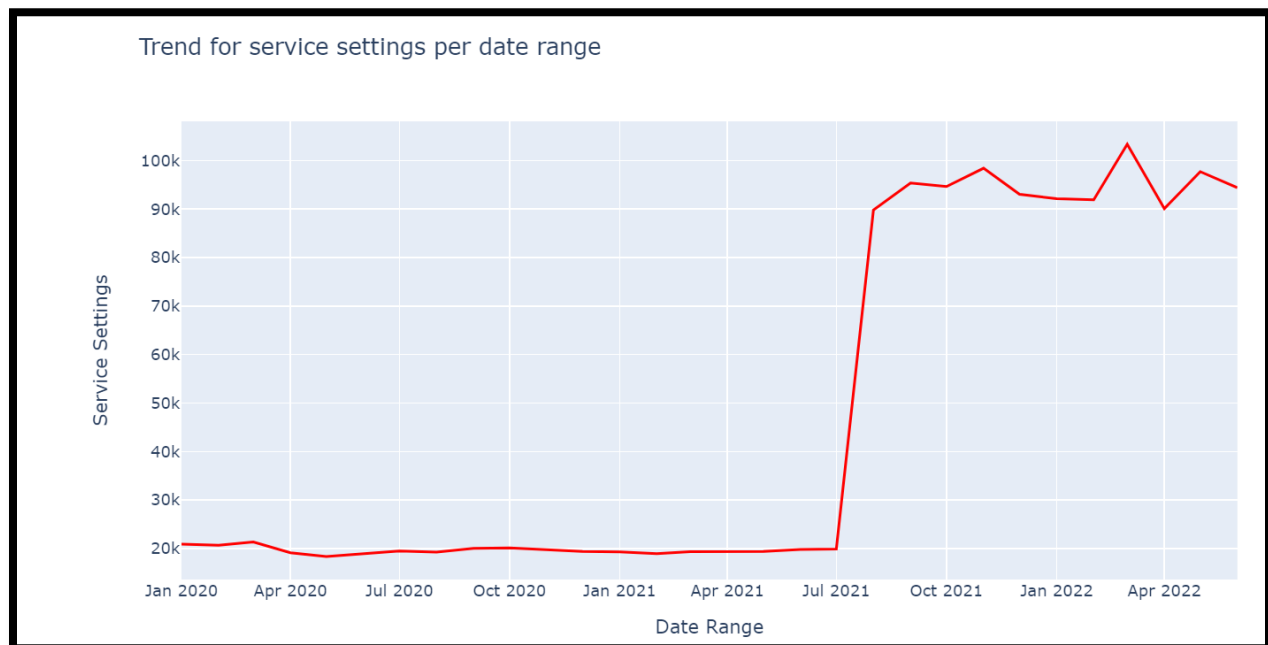
#### Appointment\_Date Service\_Setting

7	2022-03	104058
9	2022-05	97553
10	2022-06	93395
4	2021-12	92158
5	2022-01	91539
6	2022-02	90743
8	2022-04	89090
3	2021-11	77652
1	2021-09	74922
2	2021-10	74078
0	2021-08	69999

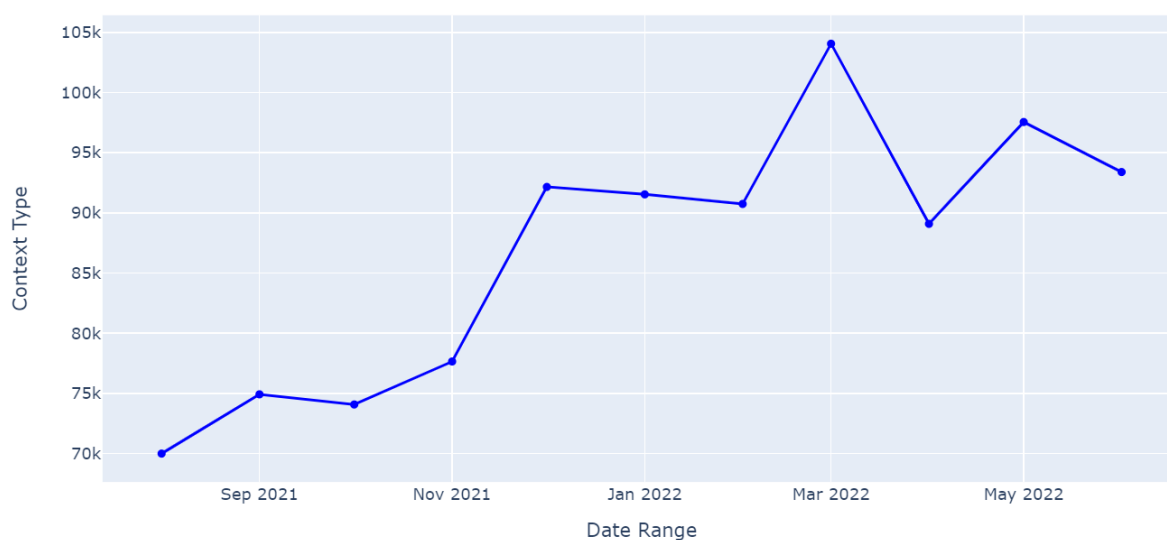
Q. What is the number of appointments and records per month?



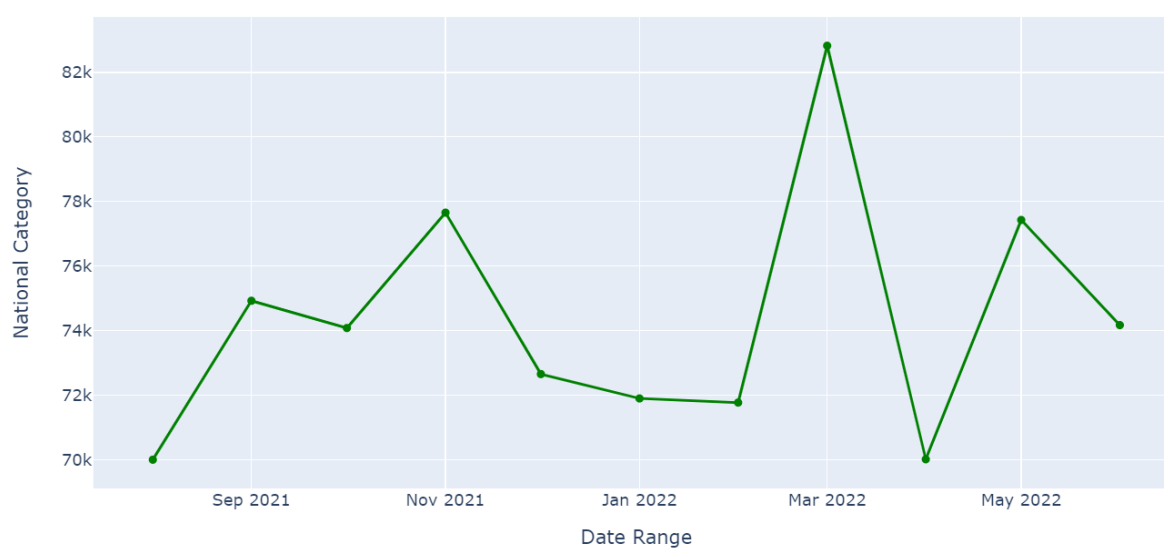
Q. What monthly and seasonal trends are evident, based on the number of appointments for service settings, context types, and national categories?



Trend for Context Type per date range



Trend for National Category per Date Range



Q. What are the top trending hashtags (#) on Twitter related to healthcare in the UK?



tweet\_output.txt

### Q. Were there adequate staff and capacity in the networks?

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.

Blank or missing HCP types are due to the use of generic lists for GPs (for example visits and prescription requests) or staff not set up correctly on the system (for example medical students and team leaders).

Where HCP type is included, it has been grouped in the following way. The occupations grouped as HCP type 'GP' are:

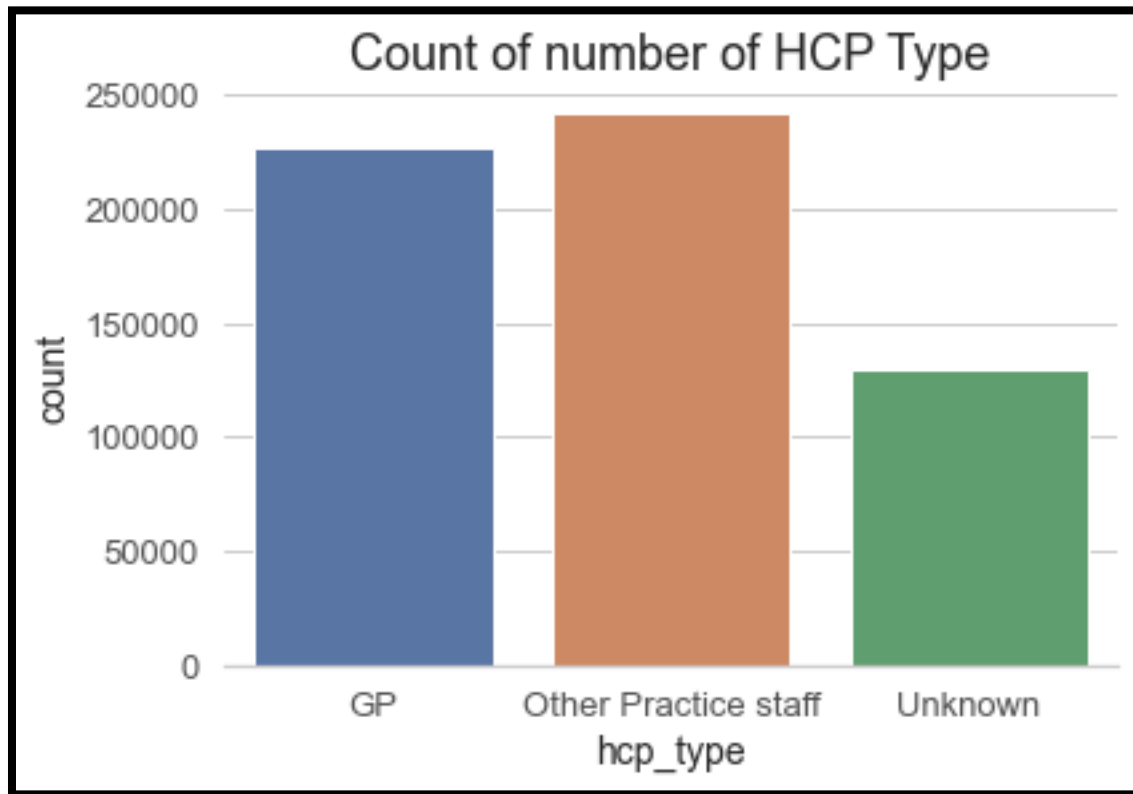
- GP registrar
- Locum GP
- Principal GP

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
- Practice Nurse

If HCP type is recorded as NULL/Missing Data, it will be classified as Unknown. It is possible that a small proportion of GP appointments are classified within Other Practice Staff due to incorrect mappings with

data suppliers, for example, Trainee GP or GP Partner. For information on how this differs from SDS Role Group go to the SDS Role Group section.



Other Practice staff	241557
GP	226036
Unknown	129228

Q. What was the actual utilization of resources?

*Check for the Appointment Mode, Actual Duration of the appointment, National Category and Time from booking to appointment.*

### 1. Appointment Mode

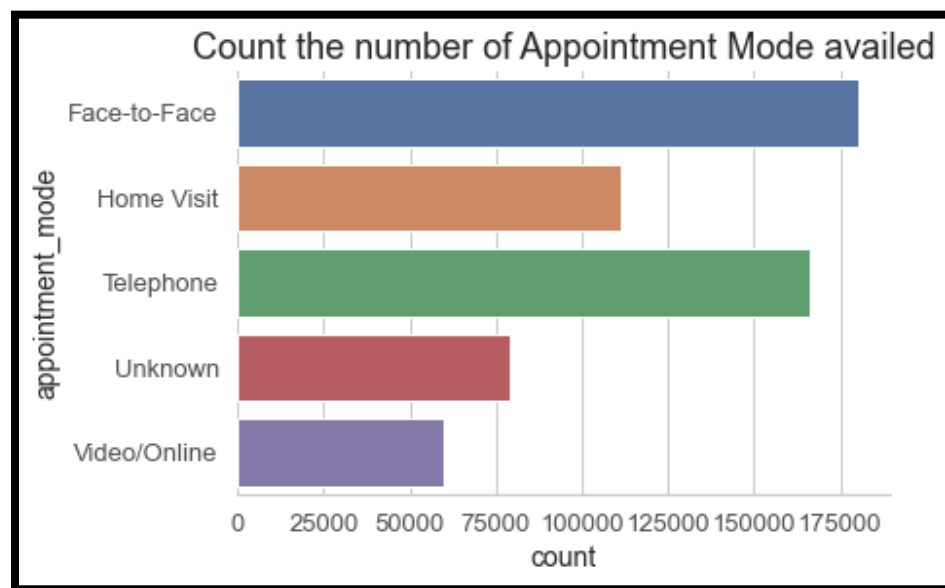
The mode of the appointment shows the setting of the consultation. For example, face to face, telephone, or home visits. This is set locally by the practices so it may not represent the actual care setting of the appointment. For example, some video conference appointments may be logged by the practice as face to face.

Most face-to-face time is booked as individual appointment time, typically with one patient attending each time slot. By contrast, many telephone triage and home visits appear as one long blocked period and are not booked (e.g., triage and home visits) to individual patients.

Some home visits and telephone triage sessions are booked in to early-morning sessions (6 am to 6:20 am) for each GP as untimed appointments, but the consultations happen in the afternoon. They are booked in to early-morning sessions so that they appear on the top of the appointment book screen for each GP. At some practices, home visits are logged at reception and not in the practice system.

Unless home visits and telephone triage are logged as individual appointments and booked to a patient, they will not appear in this publication. Appointments marked as online, video or video conference are shown as 'Online / Video'. This may or may not include a video element. Non-video based online consultations such as live chat or Voice over Internet Protocol (VoIP) and video-based appointments are all included in this category.

Practices using the Cegedim GP system are unable to supply appointment mode data. Consequently, the proportion of appointments with an 'Unknown' appointment mode is higher in releases from July 2019 onwards when Cegedim practices were included in the publication.



Face-to-Face	180410
Telephone	166483
Home Visit	111194
Unknown	79147
Video/Online	59587

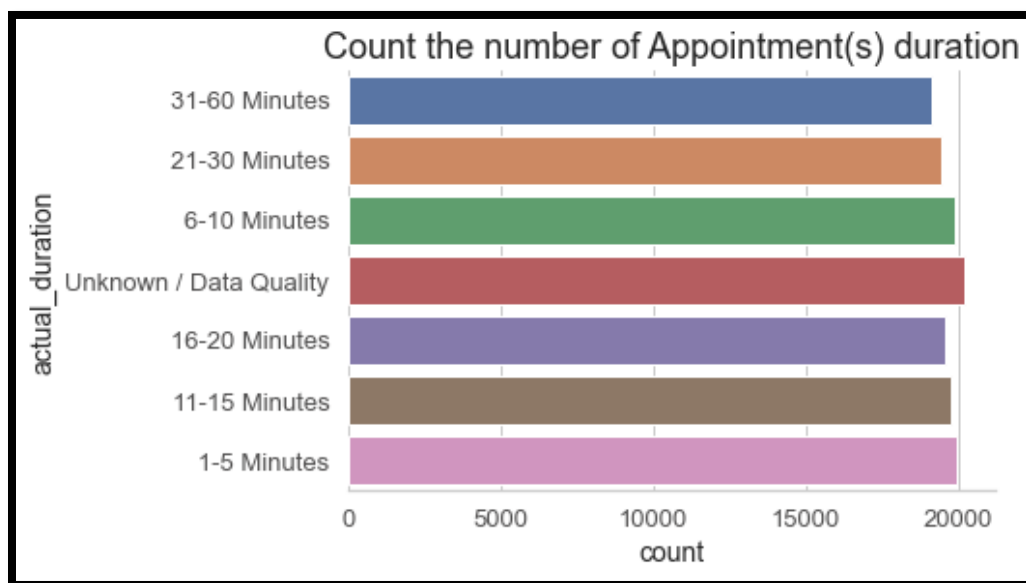


## 2. Actual Duration

Actual duration data was first included in the April 2022 release with data from December 2021. Actual duration is the length in minutes of the consultation that took place.

This field is recorded differently depending on the practice's system supplier. For example, EMIS records the actual duration as the time between the patient being called into their appointment and the time the clinician records that the consultation is complete. TPP can record the start time as either when the consultation is started, when a clinician calls for a patient, or when the appointment status is changed to 'in progress'. The end time is recorded as when the appointment is marked as finished with the actual duration being the difference between these two times. This leads to variations in data quality that NHS Digital is continuing to work with system suppliers to address.

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.



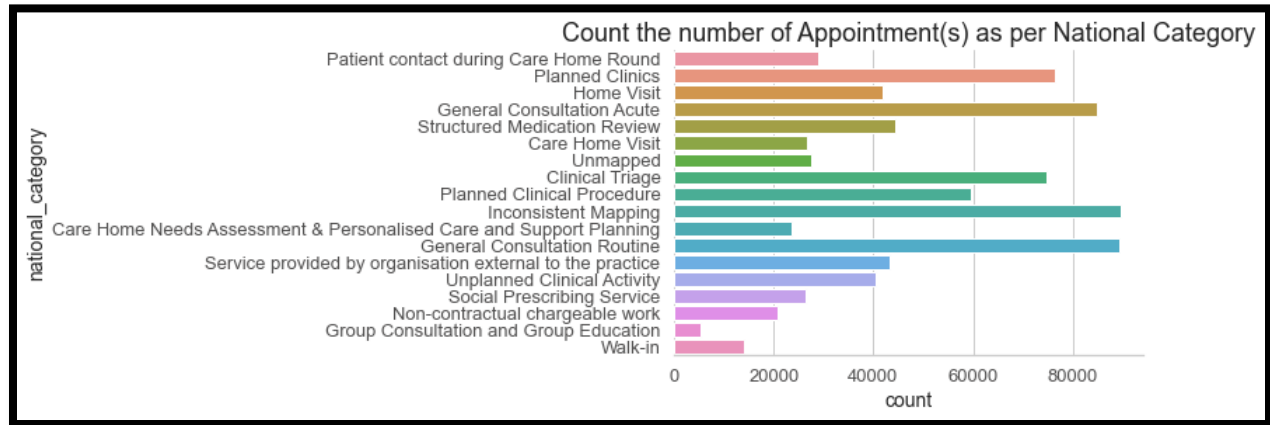
Unknown / Data Quality	20161
1-5 Minutes	19909
6-10 Minutes	19902
11-15 Minutes	19738
16-20 Minutes	19534
21-30 Minutes	19452
31-60 Minutes	19097

### 3. *National Category*

Within each context type, there is an option for practices to choose a 'Does Not Fit' category. This is used only when a practice is unable to place an appointment type within one of the predefined categories. These appointments are assigned to the 'Inconsistent Mapping' context type with further analysis planned to establish whether new categories in future releases are needed, or additional assistance and guidance is required for mapping.

- Context Type: National Category
- Care Related Encounter
  - General Consultation Acute
  - General Consultation Routine
  - Planned Clinics
  - Planned Clinical Procedure
  - Unplanned Clinical Activity
  - Walk-in, Clinical Triage, Home Visit, Care Home Visit
  - Group Consultation and Group Education
  - Structured Medication Review
  - Patient contact during Care Home Round
  - Care Home Needs Assessment & Personalized Care and Support Planning
  - Social Prescribing Service
  - Service provided by organization external to the practice.
  - Non-contractual chargeable work
- Inconsistent Mapping

- Care Related Encounter but does not fit into any other category.
- Appointment types not mapped to a Care Related Encounter



Inconsistent Mapping	89494
General Consultation Routine	89329
General Consultation Acute	84874
Planned Clinics	76429
Clinical Triage	74539
Planned Clinical Procedure	59631
Structured Medication Review	44467
Service provided by organisation external to the practice	43095
Home Visit	41850
Unplanned Clinical Activity	40415
Patient contact during Care Home Round	28795
Unmapped	27419
Care Home Visit	26644
Social Prescribing Service	26492
Care Home Needs Assessment & Personalized Care and Support Planning	23505
Non-contractual chargeable work	20896
Walk-in	14179
Group Consultation and Group Education	5341

#### 4. Time from booking to appointment

Time elapsed from when the booking was made to when the appointment happened is presented in this release. There are several factors that drive this variable, including:

- appointment availability at the practice
- patient availability
- urgency of the appointment
- GP advice
- regular/repeat appointments that may be booked in advance.



Same Day	95502
2 to 7 Days	92409
1 Day	88957
8 to 14 Days	82698
15 to 21 Days	73666
22 to 28 Days	68755
More than 28 Days	65147
Unknown / Data Quality	29687

## Patterns and Predictions

### 1. Count on Number of Locations:

NHS North West London ICB - W2U3Z	14491
NHS Kent and Medway ICB - 91Q	14121
NHS Devon ICB - 15N	14009
NHS Hampshire and Isle Of Wight ICB - D9Y0V	13499
NHS North East London ICB - A3A8R	13321

**2. Service Setting:** *The service setting describes the provision of care in which the appointment was delivered.*

- 27419 appointments were assigned to an 'Unmapped' setting since the category data is not received.
- Additional cost to NHS due to Primary Care Network (183790), Extended Access Provision(108122), Other (138789) as they are delivered via various schemes (ARRS, DES, Commissions, Practice by another provider etc.)

### 3. Context Type:

There are around **89494 "Inconsistent Mapping"** and **27419** are **"Unmapped"** which needs to be improved/reviewed for the consistency of the mapping.

### 4. National Categories:

Here there is a significant number of **"Inconsistent Mapping" 89494** which is due to **"Does Not Fit"** category. This is used only when a practice is unable to place an appointment type within one of the predefined categories.

**5. Appointment Status:**

There is a significant number of **DNA** (Did Not Attend): **163,360** and **Unknown: 201,324**.

Totaling to **364,684**, which is higher than the Attended number: 232,137.

**6. 2022/March reported highest number of appointments delivered: 104058**

**7. HCP (Healthcare Professional Type):** Higher number of Non-GP and Unknown category staff delivered healthcare support around 370,785 v/s actual GP's (226,036)!

GP: 226036

**Unknown: 129228**

**Non-GP (Other Practice staff) : 241557**

Acupuncturist

Chiropodist

Community Psychiatric Nurse

Counsellor

Dispenser

District Nurse

Health Visitor

Interpreter/Link Worker

Osteopath

Other Practice Staff

Physiotherapist

Practice Nurse

**8. Appointment mode:**

Face-to-Face      180410

Telephone          166483

Home Visit	111194
<b>Unknown</b>	<b>79147</b>
Video/Online	59587

**9. Actual Duration (Consultation):**

<b>Unknown / Data Quality</b>	<b>20161 (&lt;1-min or &gt;60-min)</b>
1-5 Minutes	19909
6-10 Minutes	19902
11-15 Minutes	19738
16-20 Minutes	19534
21-30 Minutes	19452
31-60 Minutes	19097

**10. Time from booking to appointment:**

Same Day	95502
2 to 7 Days	92409
1 Day	88957
8 to 14 Days	82698
15 to 21 Days	73666
22 to 28 Days	68755
<b>More than 28 Days</b>	<b>65147</b>
<b>Unknown / Data Quality</b>	<b>29687</b>

## Summary

The reason for patient's missing/No-show on GP appointments are due to multiple factors:

- Lack of availability of actual GP's
- Healthcare supports were provided by Non-GP Health Professionals like Acupuncturist, Chiropodist, Community Psychiatric Nurse, Counsellor, Dispenser, District Nurse, Health Visitor, Interpreter/Link Worker, Osteopath, Other Practice Staff, Physiotherapist, Practice Nurse
- Patients have a long wait time for the appointment from the time of booking.
- Consultations are of shorter duration (<15-mins)
- Appointment mode is not 100% Face-to-Face.
- One of the main reasons is due to the COVID-19 (pandemic), patients were not willing to go to the hospitals in-person from Jan-2020 until end of 2022 and were resorting to more of virtual-based consultation over Face-to-Face consultation.
- NHS is incurring more cost due to Extended Access Provision, delivered via various schemes like ARRS, DES, Commissions, Practices by other providers etc. which are all added cost.