

Technical Report: NHS Appointment Trends and Missed Appointments Analysis

1. Background/Context of the Business Scenario

The National Health Service (NHS) in England faces increasing pressure to optimize healthcare resources while managing rising service demand. Missed appointments contribute to financial losses and inefficiencies, necessitating data-driven insights to improve service utilization. The NHS seeks to determine whether existing staff and infrastructure are sufficient and how resources can be allocated effectively. This analysis explores appointment trends across Integrated Care Boards (ICBs), service settings, and appointment modes while identifying patterns in missed appointments. The goal is to provide actionable recommendations that enhance efficiency, reduce appointment no-shows, and inform future resource planning for better healthcare service delivery.

2. Analytical Approach

Data Sources and Preparation

Key Datasets used in Analysis:

- **Appointments Regional (AR dataset):** Contains appointment-level data, including appointment status, mode, and time between booking and appointment.
- **National Categories (NC dataset):** Provides service setting classifications and more details on appointment categorization.
- **ICB Code Reference Dataset:** Downloaded from Office for National Statistics to map ICB codes to their respective names for regional identification.

Other Datasets:

- **Actual Duration (AD dataset):** Details on appointment durations across different regions.
 - Note: This dataset was not utilised as it only contains a short time frame of 7 months. It would not add value significantly unless the NHS is specifically interested in how long different appointment types take.
- **Tweets Dataset:** Contains Twitter data related to NHS and healthcare discussions.
 - Note: Upon analysis, there was no significant finding that contributes to NHS utilisation trends. *Analysis was included in Jupyter Notebook Appendix.*

Data Cleaning and Wrangling

- **Handling Missing and Unmapped Data:** Many records had 'Unknown' or 'Unmapped' values in service settings and context types. These were left unchanged, as they highlight NHS data collection limitations rather than analytical gaps. (Appendix B)
- **Merging Data:** The ICB reference dataset was merged with the AR dataset using a **left join** via the *icb_ons_code* column to improve regional-level analysis.
- **Feature Engineering:**

- **Missed Appointment Percentage:** Calculated as the proportion of 'Did Not Attend' (DNA) status per total appointments.
- **Aggregated Data for Trend Analysis:** Grouped by **appointment month, appointment mode, and time between booking and appointment** etc to identify trends.
- **Comparison of Missed Appointments:** Summed total appointments and DNA occurrences across ICBs to identify patterns in missed appointments.

Exploratory Data Analysis

- **Regional Trends:** Identified the top 5 ICBs with highest DNA rates, but did not analyse sub-ICBs due to data limitations.
- **Missed Appointment Analysis:** Explored the impact of appointment mode and time between booking and appointment on DNA rates.
- **Appointment Mode Breakdown:** Analyzed missed appointment percentages across different appointment modes.
- **Utilization vs. Capacity:** Evaluated monthly appointment trends against the estimated NHS capacity of 1.2 million daily appointments.

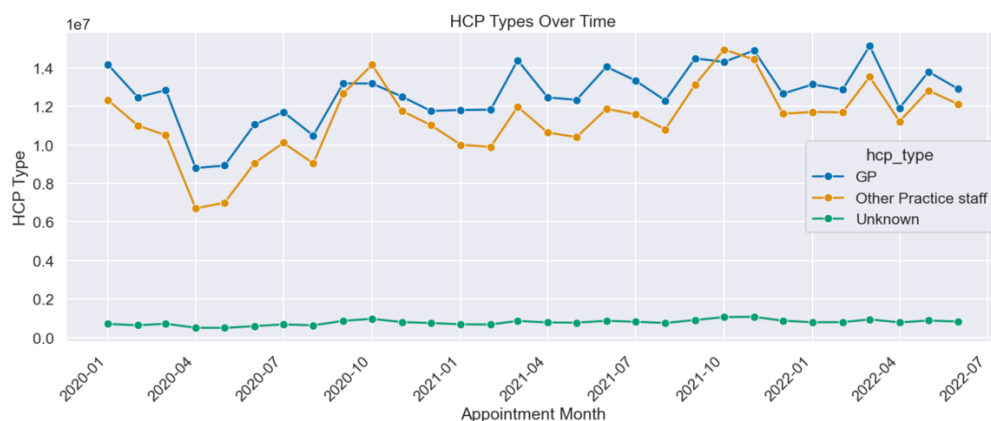
For further information, refer to Appendix (A: Limitations, B: Data Quality Considerations, E: Assumptions, F: Preliminary Findings, G: Additional Findings, H: Additional Visualisations).

3. Visualizations and Insights

Trends in Total Appointments

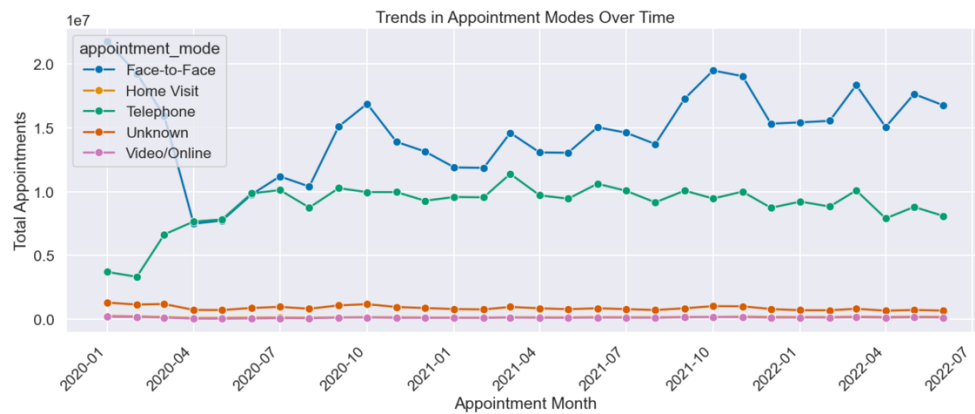
This section examines appointment volumes over time to identify key trends. The significant drop in April 2020 aligns with the first UK COVID-19 lockdown, disrupting healthcare services.

1. Healthcare Professional (HCP) Types over Time



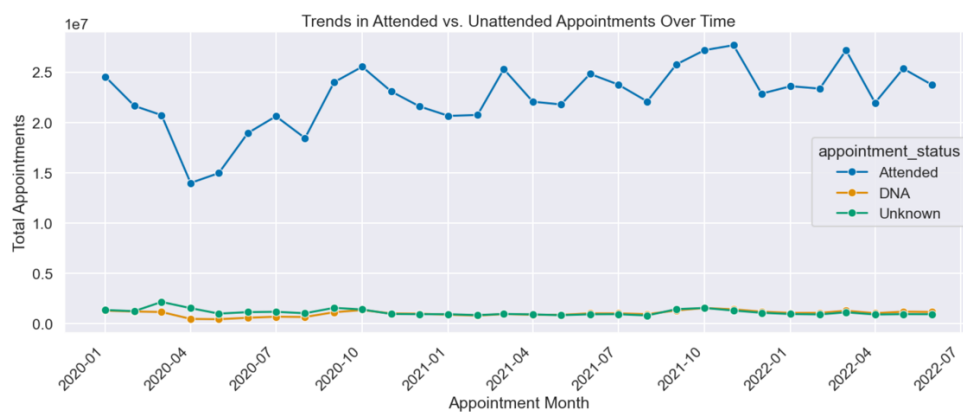
Insight: GPs handle the highest appointment volume, highlighting strong demand for GP-led care. Balanced staffing is needed to prevent overload and optimize NHS capacity.

2. Appointment Modes over Time



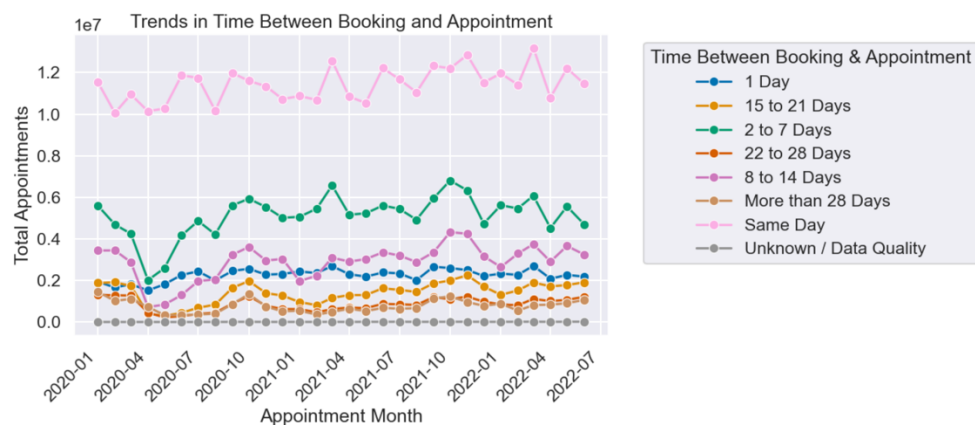
Insight: Patients prefer face-to-face appointments despite the availability of remote care. Low adoption of video/online consultations suggests accessibility barriers need to be addressed.

3. Attended vs Unattended Appointments Over Time



Insight: Attended appointments vastly outnumber missed (DNA) appointments, but even a small percentage of DNAs translates to significant inefficiencies given the scale of the NHS.

4. Time Between Booking and Appointment over Time

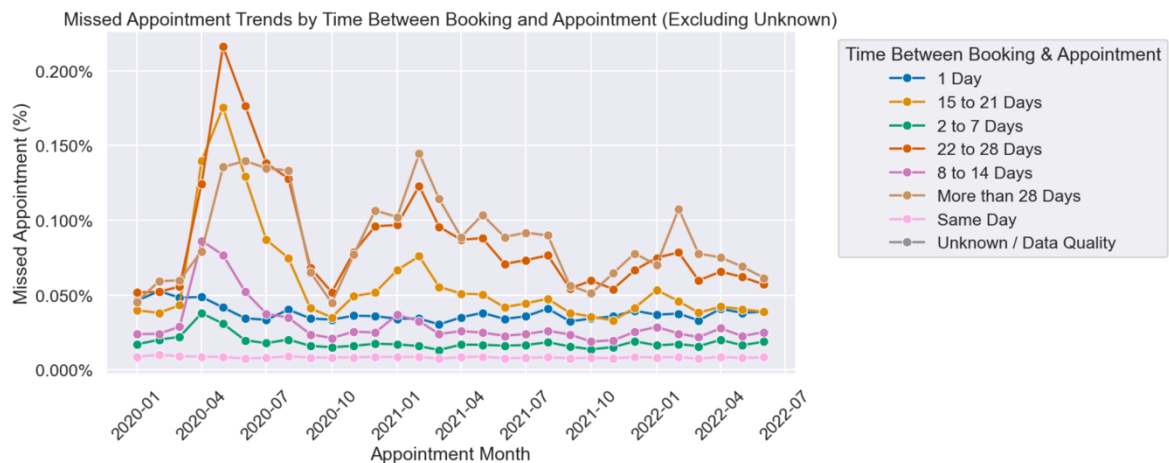


Insight: Patients favor same-day bookings, but long wait times (more than 8 days) indicate scheduling inefficiencies. NHS should balance urgent care with better long-term planning.

Trends in Missed Appointments (DNA Rates)

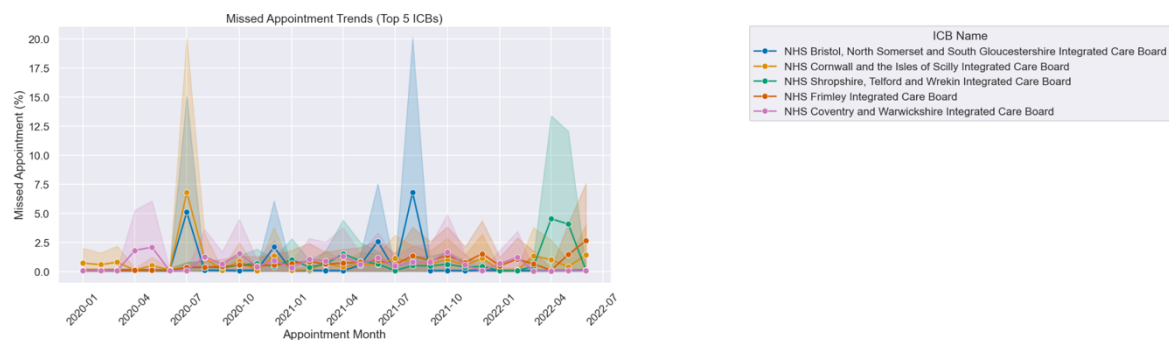
This section explores missed appointment trends as a percentage of total appointments, helping to identify inefficiencies. *Additional details can be found in Appendix C and D.*

1. By Time Between Booking and Appointment



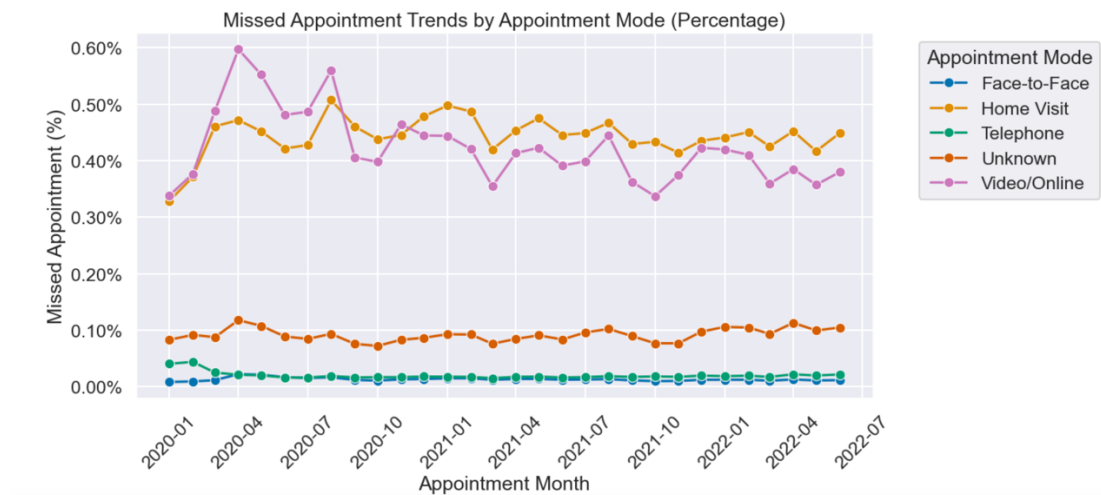
Insight: Longer lead times lead to higher DNA rates as patients disengage over time. Predictive analytics can help identify high-risk no-shows and optimize scheduling.

By Region (Top 5 ICBs with highest DNA rates)



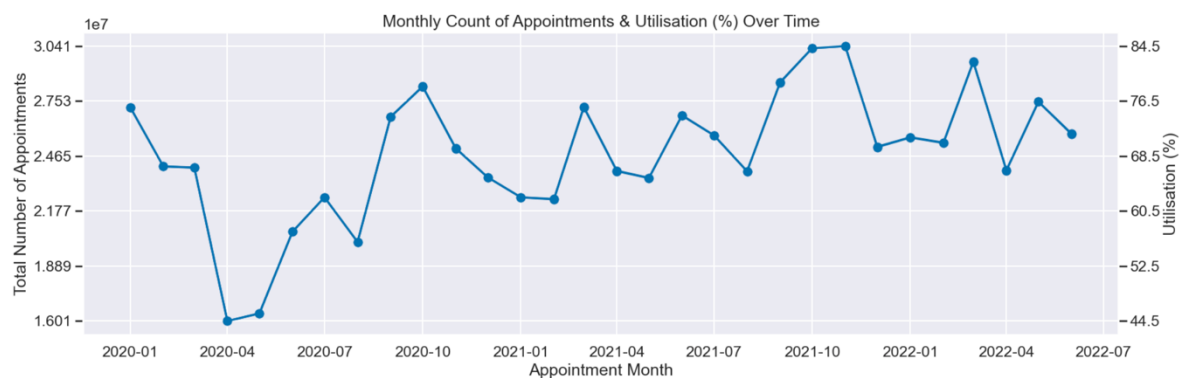
Insight: While the top 5 ICBs with the highest DNA rates show regional disparities, overall missed percentages remain low, suggesting localized challenges rather than system-wide issues.

2. By Appointment Mode



Insight: Video/Online and Home Visits have the highest DNA rates likely due to tech barriers and patient disengagement. This strains capacity and increases wait times for others.

Utilization vs. Capacity



Insight: Despite fluctuations in appointment volume, NHS utilization remains under 85%, indicating no consistent resource overuse (*Healthcare systems aim for 85-90% utilisation for efficiency*). Instead of expanding capacity, efforts should focus on optimizing appointment scheduling and reducing missed appointments.

4. Patterns and Predictions

Key Findings (*Appendix D*)

1. Missed Appointments Are Highest in Remote and Non-Traditional Settings
2. Longer Appointment Lead Times Increase DNA Rates
3. Capacity Is Sufficient, But Inefficiencies Persist

Recommendations

1. Improve Patient Engagement for Video/Online Appointments to Reduce DNA Rates

KPI: Decrease missed appointment rate for Video/Online consultations by 15% over 6 months. (Appendix C)

- **Implement Automated Reminders:** Send SMS or email reminders 24-48 hours before video or online appointments to reduce no-shows.
- **Introduce Confirmation Incentives:** Require patients to confirm their attendance via SMS or an NHS app. Those who confirm and attend consistently could receive priority scheduling for future appointments.
- **Enhance Access to Online Consultations:** Ensure that technical issues (e.g., unstable connections) are minimized by providing clear instructions and technical support.

2. Optimize Scheduling to Reduce Missed Appointments

KPI: Reduce missed appointment rate for appointments booked more than 7 days in advance to below 0.05%.

- **Increase Same-Day and Next-Day Appointments:** Data shows shorter booking times lead to lower DNA rates. NHS should allocate more slots for immediate or next-day consultations.
- **Use Predictive Analytics for Scheduling:** Identify high-risk missed appointment patients (e.g., those with a history of DNAs) and adjust scheduling strategies to reduce missed visits.

3. Improve Data Collection and Standardization for More Reliable Analysis

KPI: Reduce 'Unknown' entries in *appointment_mode* and *hcp_type* fields by 30% in 12 months.

- **Mandate Drop-Down Selection for Appointment Modes:** Standardize appointment classification to reduce errors and inconsistent mappings.
- **Flag 'Unknown' and 'Unmapped' Data for Review:** NHS systems should trigger alerts when excessive "Unknown" values appear in records, prompting manual review.
- **Develop a Centralized Data Entry System:** A single, unified NHS platform should ensure all regions follow the same data recording practices, reducing discrepancies.

Conclusion

The NHS has sufficient infrastructure, but inefficiencies in scheduling and patient adherence lead to avoidable financial losses and underutilized capacity. DNA rates are highest in remote consultations and long booking lead times, indicating a need for stronger patient engagement strategies. Addressing scheduling inefficiencies and improving data standardization will allow NHS to optimize service delivery without unnecessary capacity expansion.

Appendix

Appendix A: Limitations

1. **Regional Comparisons are Limited Due to Data Constraints**
 - The **large number of ICBs and sub-ICBs** makes meaningful regional analysis challenging.
 - The **AR dataset only contains ICB-level data** with attendance rates but lacks sub-ICB details, restricting deeper insights.
 - Other datasets provide **sub-ICB details but no attendance rates**, preventing a granular analysis of missed appointments at a lower level.
 - As a result, regional analysis was only conducted to identify the **top 5 ICBs with the highest DNA rates**, as this was the only feasible comparison given the available data.
2. **Visualization Choices are Limited**
 - **Line charts were used for all visualizations** as they effectively show trends over time.
 - **Stacked bar charts** were considered but deemed impractical due to the large number of categories, which would have made interpretation difficult.
 - While line charts allow for trend analysis, they may **not fully capture proportional differences** in missed appointments across categories.
3. **Data Quality Issues and Standardization Challenges**
 - The NHS datasets contain **inconsistent classifications**, missing values, and "Unknown" categories that reduce analytical accuracy.
 - The **lack of a unified NHS data entry system** results in discrepancies across regions, limiting the reliability of some insights.

Appendix B: Data Quality Considerations: High proportion of missing and unmapped data

A significant proportion of appointment records contain 'Unmapped' or 'Inconsistent Mapping' values. These likely stem from variations in data entry across NHS systems rather than actual missing appointments. While this does not affect high-level trends, it may limit the accuracy of more granular analyses, such as specific service settings or regional insights.

Given the large scale of the dataset and the lack of a reliable way to accurately reclassify these entries, no adjustments have been made. Instead, these inconsistencies are acknowledged as a limitation, and recommendations for improving NHS data quality are provided.

Appendix C: Financial Implications of Missed Appointments

Financial Impact of Did Not Attend (DNA) Rates:

- **Volume of Appointments:** The NHS offers approximately 1.4 million GP appointments every working day
- **Cost per Missed Appointment:** Each missed GP appointment costs the NHS an average of £30

- **Annual Financial Loss:** In 2019, approximately 7.2 million GP appointments were missed annually in England, costing the NHS around £216 million.
- **Potential Savings:** By reducing DNAs by 15%, the NHS could potentially recover approximately £32.4 million annually.

Caveat: This calculation assumes that the cost per GP appointment remains constant and that all recovered appointments would have otherwise gone unused. However, actual savings may vary depending on operational efficiency, patient demand, and the NHS's ability to effectively reallocate these slots. Factors such as administrative costs, scheduling constraints, and workforce capacity could influence the overall financial impact.

Appendix D: 5 Whys Root Cause Analysis of Missed Appointments (DNA Rates)

Step 1: Why do patients miss appointments?

- Patients forget about their appointments or are unaware of them.

Step 2: Why do they forget or remain unaware?

- They do not receive reminders, or the reminders they receive are ineffective.

Step 3: Why are reminders ineffective or not received?

- Reminder methods vary across regions, appointment types and patient demographics, leading to inconsistency.

Step 4: Why is there inconsistency in reminder methods?

- The NHS lacks a standardised, automated reminder system across all Integrated Care Boards (ICBs) and service settings.

Step 5: Why does the NHS not have a standardised reminder system?

- Differences in data management systems and funding constraints across regions create challenges in implementing a uniform approach.

Findings from the 5 Whys Analysis

- **Reminder Inconsistencies Drive DNA Rates:** Patients in different ICBs may receive different levels of engagement, leading to avoidable no-shows.
- **Longer Booking Lead Times Increase Missed Appointments:** When appointments are scheduled weeks or months in advance without follow-up reminders, patients are more likely to forget.
- **Regional Variations in Communication:** Some NHS regions use SMS/email reminders, while others rely on letters or no reminders at all.
- **Lack of a Centralized NHS Reminder System:** Without a nationwide automated system, appointment adherence remains inconsistent across service settings.

The 5 Whys analysis highlights that NHS DNA rates are primarily driven by inconsistent patient reminders and long booking lead times. Standardizing appointment reminders, automating follow-ups, and leveraging AI-driven scheduling can significantly improve attendance rates, leading to better resource utilization and reduced financial losses for the NHS.

Appendix E: Assumptions

1. "Unknown" or "Unmapped" values are due to inconsistencies in data entry, not actual missing appointments.

NHS Digital acknowledges that data in general practice appointments contains quality issues due to variations in recording methods across practices.

2. Assumption: Regional data inconsistencies are due to NHS reporting variations, not actual service issues.

NHS Digital points out that data quality issues stem from variations in working methods and recording between different practices and Primary Care Networks (PCNs).

3. Assumption: Despite missing data, general trends in appointment volume and utilization are still valid.

NHS England emphasizes the importance of improving data quality to better reflect actual activity, suggesting that current data, despite its issues, is used to understand trends.

4. Assumption: Historical data is representative of current and future NHS trends. However, external factors such as policy changes may impact future predictions. Trends should be interpreted with consideration of these factors.

The Health Foundation discusses how better use of routine health data can address current challenges, implying that historical data is valuable for understanding and predicting trends.

Appendix F: Preliminary Findings

Question 1: How many locations are there in the data set?

- Answer: The NHS dataset is divided into 7 regions, 42 ICB locations, and 106 sub-ICB locations.

Question 2: What are the five locations with the highest number of appointments?

- Answer: The top 5 sub-ICB locations are North West London, North East London, Kent and Medway, Hampshire and Isle of Wight, and South East London.

Question 3: How many service settings, context types, national categories, and appointment statuses are there?

- Answer: There are 5 service settings, with General Practice being the most commonly available. While GPs handle most appointments, PCN also plays an important role. Extended Access Provision is also important, meaning many patients need care outside standard working hours.
- There are 3 context types, with Care Related Encounter being the most commonly available.
- There are 18 national categories in the dataset. However, a majority of records indicate Inconsistent Mapping.
 - A large portion of records are not mapped properly, reducing data reliability. The General Consultation Routine and Acute are the most common appointment types, indicating that they are core NHS services.
 - Home Visits and Care Home Services indicate demand for at-home or elderly patient care.
 - Walk-in Appointments are low compared to scheduled services, suggesting that patients rely more on scheduled appointments. This also shows a good sign of system efficiency.
 - Group Consultation and Group Education is the least used category, indicating that group healthcare programs are not widely adopted.
- There are 3 types of appointment status. The number of unknown records and attended records are almost the same, indicating that there is a high proportion of Unknown data.
 - A high number of appointments have an Unknown status, which could indicate data quality issues or missing updates. Over 160k appointments were missed, leading to potential inefficiencies and financial losses for the NHS.

Appendix G: Additional Findings

Question 1: Between what dates were appointments scheduled?

- Answer:
 - In ad, appointments run from 1 Dec 2021 to 30 Jun 2022 (7 months)
 - In nc, appointments run from 1 August 2021 to 30 Jun 2022 (11 months)
 - In ar, appointments run from Jan 2020 to June 2022 (30 months)

Question 2: Which service setting was the most popular for NHS North West London from 1 January to 1 June 2022?

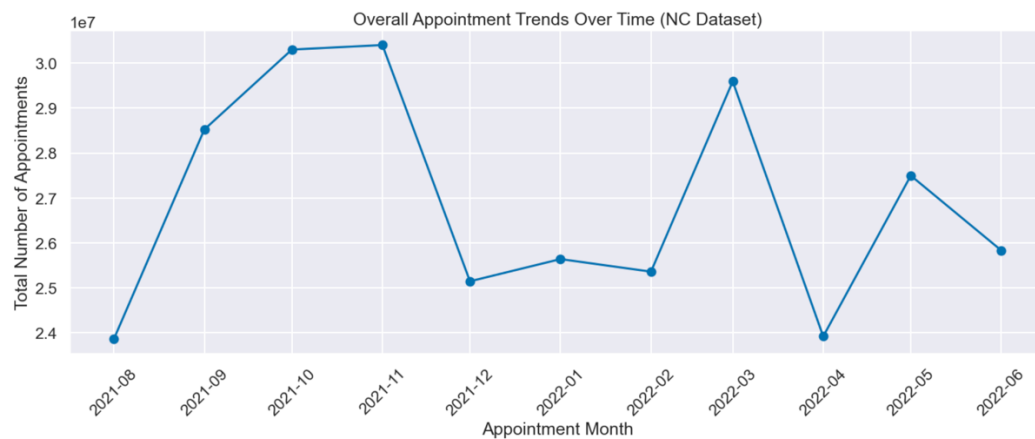
- Answer: General Practice was the most popular service setting for NHS North West London from 1 January to 1 June 2022.

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

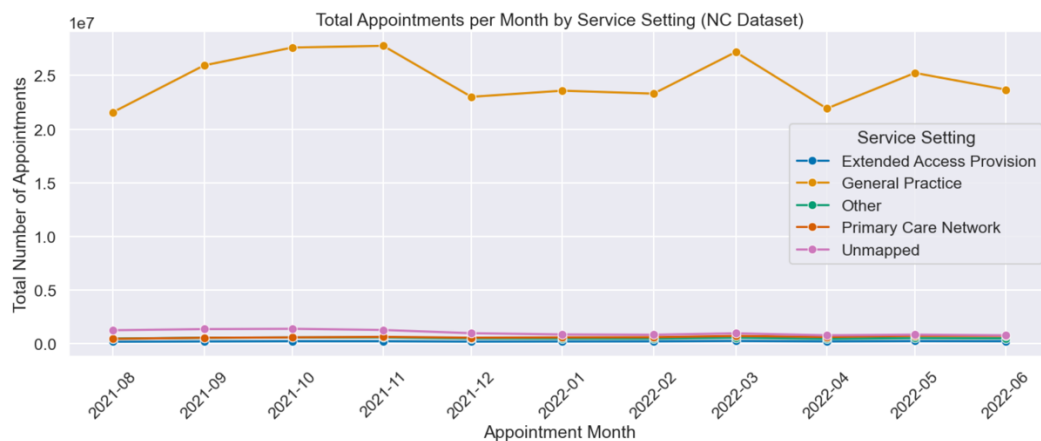
- Answer: November 2021 had the highest number of appointments, which is consistent in nc and ar datasets. This is likely due to the Covid 19 Booster Vaccination Campaign by the NHS, which invited many people to book their booster jab.
- The ad dataset only covers 7 months and is thus not suitable for monthly appointment trend.

Appendix H: Additional Visualisations

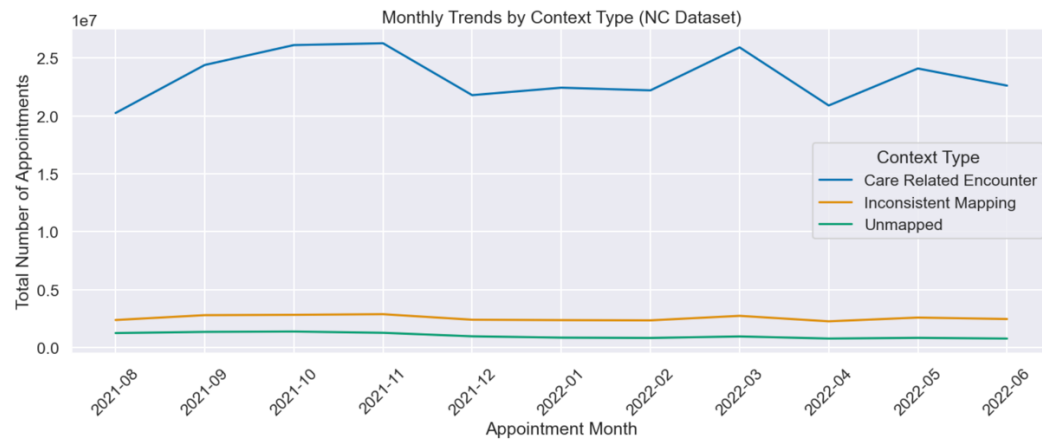
Monthly Trends of Appointments



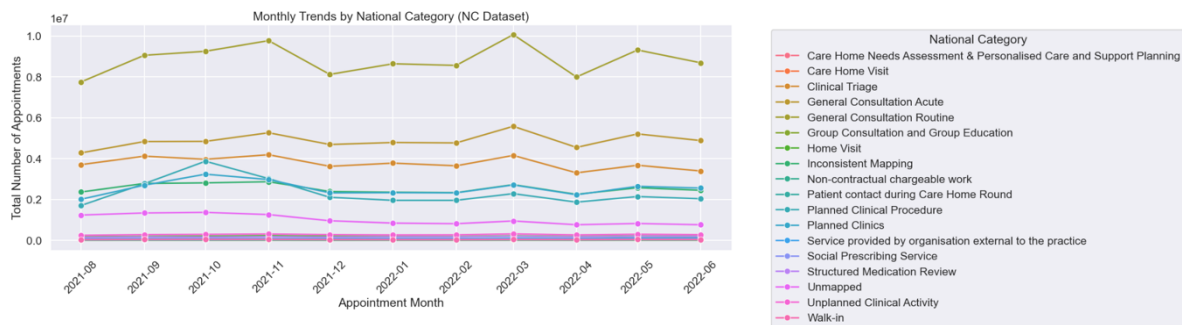
Monthly Trends by Service Setting



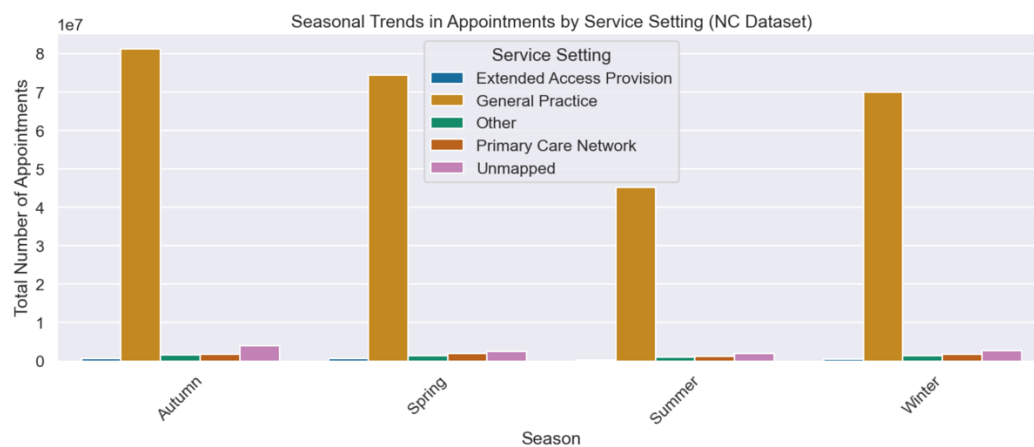
Monthly Trends by Context Type



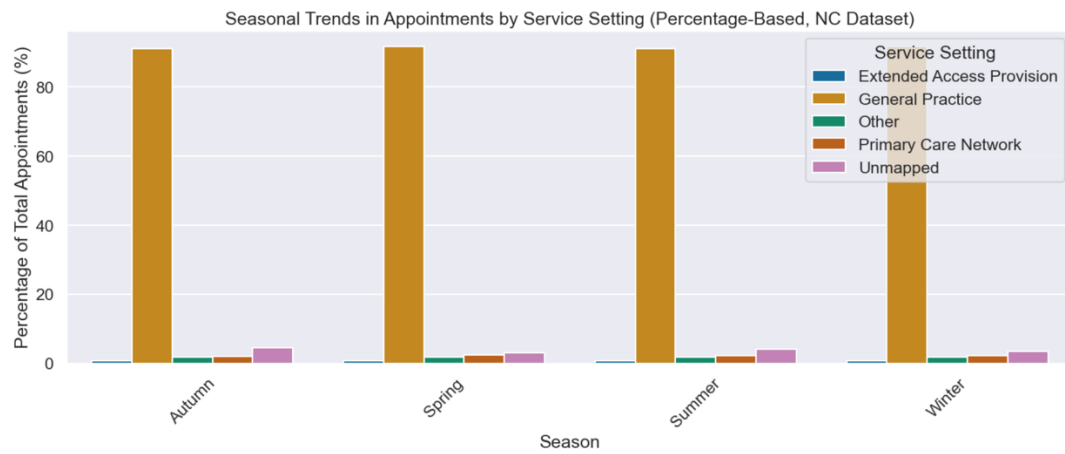
Monthly Trend by National Category



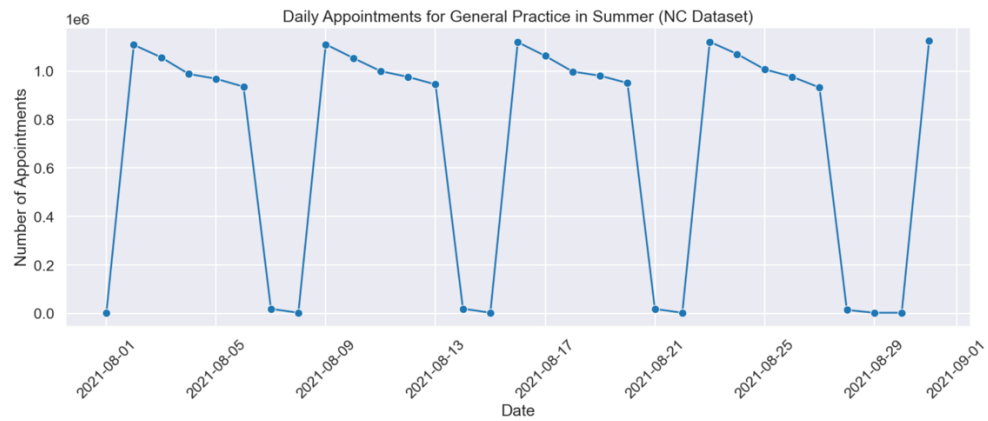
Seasonal Trends of Number of Appointments for Service Setting



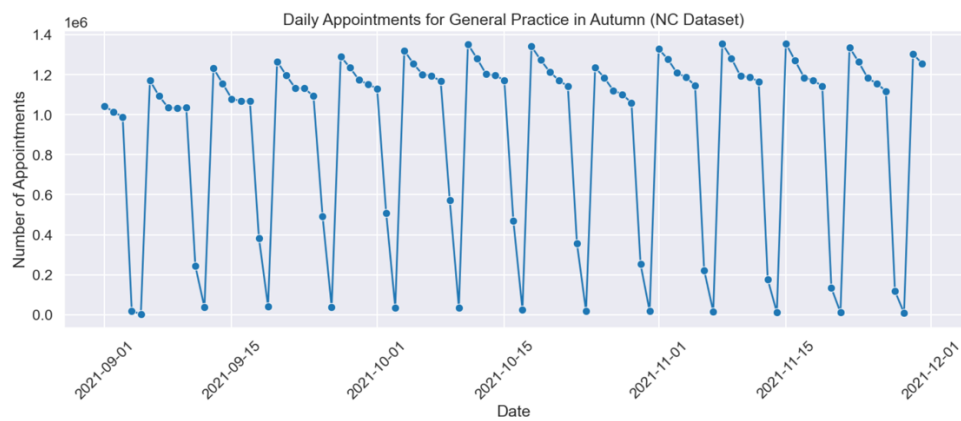
Percentage Distribution of Service Settings by Season



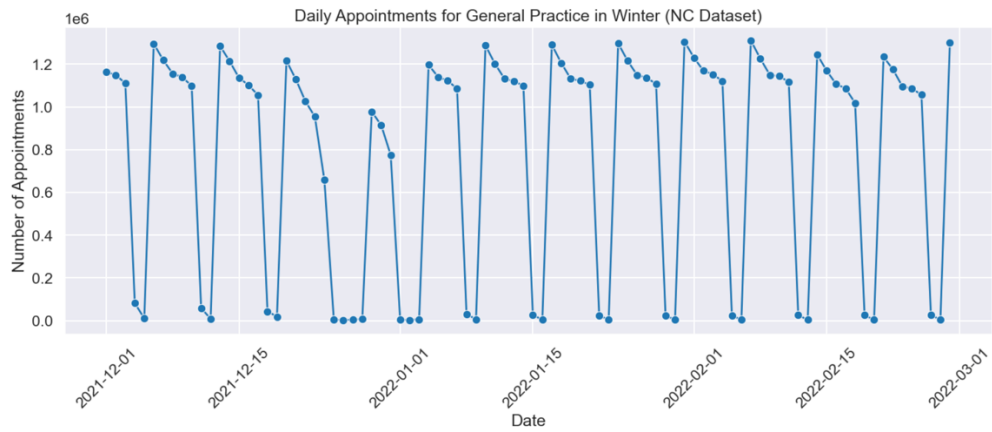
Summer (June to August 2021):



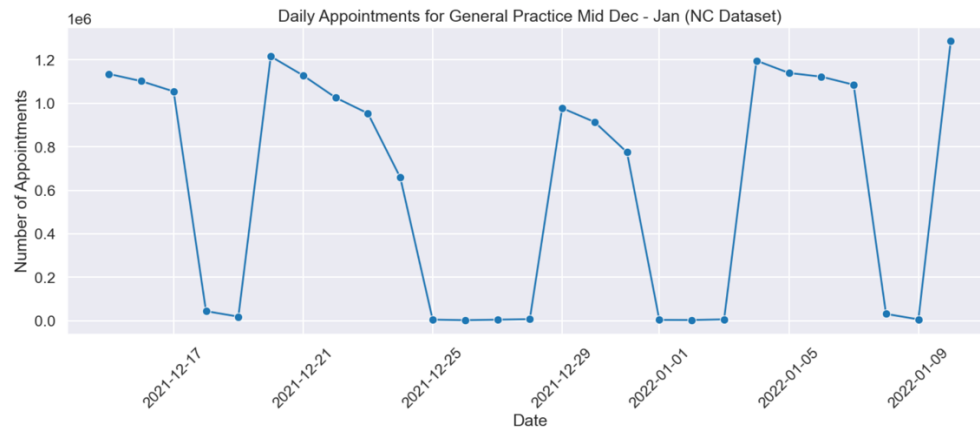
Autumn (September to November 2021):



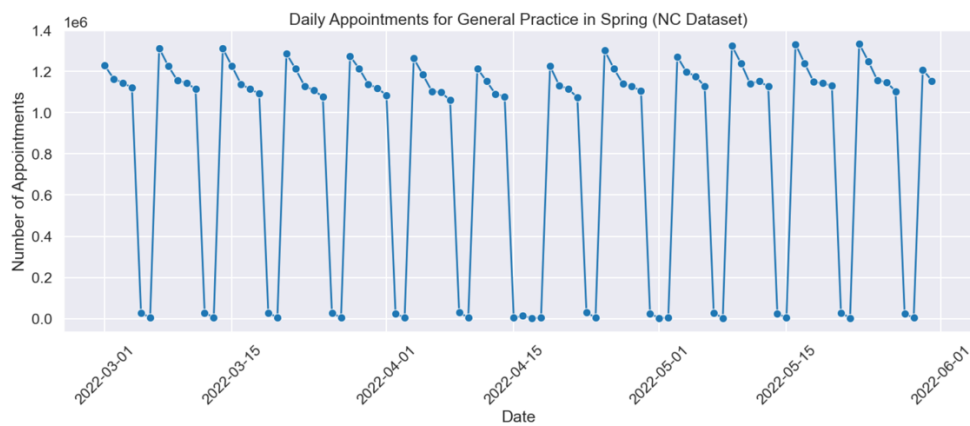
Winter (December to February 2022):



Winter Closeup on Mid Dec - Jan



Spring (March to May 2022):



Appendix I: References

- **ONS Code Dataset:** Obtained from the UK Government's Geoportal
 - <https://geoportal.statistics.gov.uk/documents/25ba241a775e4a9db8e5c721ee73d85d/about>
- **Financial Impact of Did Not Attend (DNA) Rates:**
 - **Volume of Appointments:** The NHS offers approximately 1.4 million GP appointments every working day (<https://www.england.nhs.uk/2024/04/millions-more-gp-appointments-in-march-than-before-pandemic/>)
 - **Cost per Missed Appointment:** Each missed GP appointment costs the NHS an average of £30 (<https://www.patientclaimline.com/article/which-areas-in-the-uk-have-the-most-missed-gp-appointments-and-how-much-does-it-cost-the-nhs/>)
 - **Annual Financial Loss:** In 2019, approximately 7.2 million GP appointments were missed annually in England, costing the NHS around £216 million. (<https://pmc.ncbi.nlm.nih.gov/articles/PMC9976814/>)
- **5 Whys Analysis Section:**
 - <https://phescreeing.blog.gov.uk/2019/12/19/help-develop-principles-text-reminders/>
 - <https://bmjopen.bmj.com/content/6/10/e012116?>
 - <https://elht.nhs.uk/patients/patient-reminder-service>
 - <https://www.uclh.nhs.uk/patients-and-visitors/outpatients/sms-appointment-reminder-service>
- **Assumptions Section:**
 - Source for Assumption 1-3: <https://digital.nhs.uk/data-and-information/publications/statistical/appointments-in-general-practice/appointments-in-general-practice-supporting-information>
 - Source for Assumption 4: <https://www.health.org.uk/reports-and-analysis/briefings/how-better-use-of-data-can-help-address-key-challenges-facing-the>
- **Appointment Spikes:**
 - Nov 2021: Due to the NHS COVID-19 booster jab campaign
 - <https://www.england.nhs.uk/2021/11/record-breaking-week-for-nhs-covid-19-booster-jabs/>
 - Mar 2022: Linked to the NHS Spring COVID-19 booster jab campaign
 - <https://www.england.nhs.uk/2022/03/nhs-covid-19-vaccine-programme-delivers-first-spring-boosters/>
 - May 2022: Corresponds with NHS Equality, Diversity and Human Rights Week
 - <https://www.nhsemployers.org/articles/equality-diversity-and-human-rights-week-2024> (Page from 2022 is no longer available, but as this event occurs annually, the 2024 page is referenced)
- **Healthcare Utilisation Benchmark:**

- Research indicating that healthcare systems generally aim for 85-90% utilization for efficiency while maintaining flexibility for surges:
<https://pmc.ncbi.nlm.nih.gov/articles/PMC7060560/>