

NHS Service Utilisation and Missed Appointments: Technical Report

Student: Monica Baracho

Course: LSE DA201 – Diagnostic Analytics with Python

Date: May 2025

Table of Contents

1 Background & Context	2
2 Analytical Approach	2
Data	2
Table 1: Overview of Datasets Used for NHS Appointment Analysis	2
Methods	2
3 Visualisation & Insights	3
3.1 National Trends & System Pressure	3
Figure 1: Monthly Appointments Trends with UK Lockdowns	3
Dashboard 1: Key Performance Indicators Dashboard (August 2021- June 2022)	4
Figure 2: Month-over-Month % Change with 3-Month Avg	5
Figure 3: Appointments Over time by Mode	5
Figure 4: Appointments by Duration (Minutes)	6
Figure 5: Monthly NHS Appointments by Healthcare Professional	6
Figure 6: Appointments by Season: Total and Average	7
Figure 7: Capacity Utilisation Vs Limit	7
3.2 Regional Inequality: Access vs Volume	8
Figure 8: Volume vs Access (Total Appointments vs Per Capita)	8
Figure 9: Utilisation Inequality Chats (overutilised Vs Underutilised)	9
3.3 Missed Appointments – Patterns & Predictors	9
Figure 10: Monthly Missed Appointments	9
Figure 11: Missed Appointments by Booking Delay	10
Figure 12: Missed Appointments per 1,000 Registered Patients	11
Figure 13: Predicting Missing Appointments: Logistic Regression Insights	12
3.4 Public Sentiment	12
Figure 14: Top 10 NHS-related Hashtags	13
Conclusions	13
Recommendations	13
Appendix	14
Appendix A: Average Missed Appointments per Month	14
Appendix B: Monthly Missed Appointments (Volume)	14

1 Background & Context

The NHS is under sustained pressure from an ageing, growing population and tightening budgets. Missed general-practitioner (GP) appointments alone cost millions of pounds each year and squander scarce clinical time. This analysis tackles two business-critical questions:

1. **Capacity:** Have staff and infrastructure kept pace with demand?
2. **Utilisation:** How efficiently are existing resources being used?

We combine internal NHS appointment records (Aug 2021 – Jun 2022) with public Twitter discourse to surface trends, risk factors, and attitudes that inform smarter staffing, booking policies, and patient engagement.

2 Analytical Approach

Data

All three datasets were fully validated prior to analysis. No missing values were found across any column. `appointments_regional.csv` spans 30 months (Jan 2020–Jun 2022), though 21 604 duplicate rows were identified and retained to preserve multi-level appointment entries. `actual_duration.csv` includes 137 793 complete records, with 7 duration bins; ~15% are labelled “Unknown / Data Quality” and excluded from mean-duration calculations. `national_categories.xlsx` covers Aug 2021–Jun 2022, with 42 ICBs and 18 appointment categories. All date fields parsed correctly, and no zero or negative values appeared in `count_of_appointments`.

Table 1: Overview of Datasets Used for NHS Appointment Analysis

Dataset	Description	Timeframe	Source
<code>appointments_regional.csv</code>	Appointment status, mode, HCP	Jan 2020 – Jun 2022	NHS
<code>actual_duration.csv</code>	Duration of appointments	Aug 2021 – Jun 2022	NHS
<code>national_categories.xlsx</code>	Category, context, setting	Aug 2021 – Jun 2022	NHS
<code>icb_locations.csv</code>	Code-name mapping for ICBs	2022	NHS/ONS
<code>tweets.csv</code>	Hashtags & sentiment	2020–2022	Twitter
<code>metadata_nhs.txt</code>	Data dictionary	—	NHS

Methods

- **Cleaning & merge:** Python (Pandas) aligned regional codes, normalised appointment modes, and labelled unknown HCP types.
- **Exploration:** Descriptive stats, seasonal decomposition, and capacity calculations (appointments ÷ 100 % planning limit).
- **Missed-appointment model:** Logistic regression on binary DNA outcome; categorical predictors one-hot encoded. Results expressed as odds ratios.
- **Model validation:** ROC curve, AUC = 0.569.
- **Public sentiment:** Keyword filtering, VADER polarity scoring, and hashtag frequency.

The final notebook, including data preparation, visualisations, and modelling, is available as a rendered HTML file **and** a downloadable PowerPoint presentation: <https://murbaracho.github.io/NHS-dashboard-/>

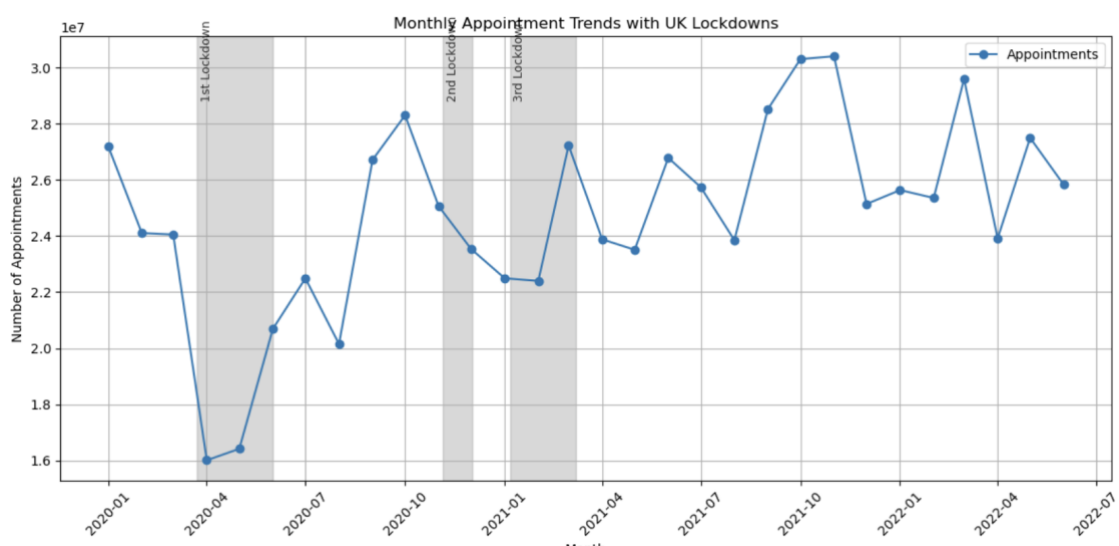
3 Visualisation & Insights

3.1 National Trends & System Pressure

Figure 1 charts monthly appointment volumes against UK lockdowns. Activity collapsed to < 1.6 million in April 2020, rebounded past 3 million by late 2021, and stabilised above pre-pandemic norms. As shown in **Appendix A: Average Missed Appointments per Month**, the average number of missed appointments was significantly lower during lockdown than before or after.

Therefore, this report focuses on the post-lockdown period (from August 2021 onward) to provide a more accurate reflection of ongoing demand and missed appointment behaviour under normalised operating conditions.

Figure 1: Monthly Appointments Trends with UK Lockdowns.



Dashboard 1 summarises Aug 2021 – Jun 2022

- 25.8 million appointments/month (peak 29.1 m in Nov).
- Face-to-face still dominant (61.7 %); GP practices deliver 90.7 % of care.
- **Capacity utilisation averages 115.2 %**, signalling chronic over-stretch.
- DNA rate averages 4.69 %.
- Access inequality: Cornwall delivers 6 201 appointments per 1 000 residents; South East London only 3 850.

Dashboard 1: Key Performance Indicators Dashboard (August 2021- June 2022)



Figure 2 shows large month-over-month swings in appointment demand. December 2021 (-18 %) and April 2022 (-19 %) plunge sharply—holiday closures and post-Omicron illness—while March 2022 and May 2022 rebound > +15 %. The 3-month moving average smooths the saw-tooth pattern, confirming persistent volatility that complicates fixed staffing plans.

Figure 2: Month-over-Month % Change with 3-Month Avg

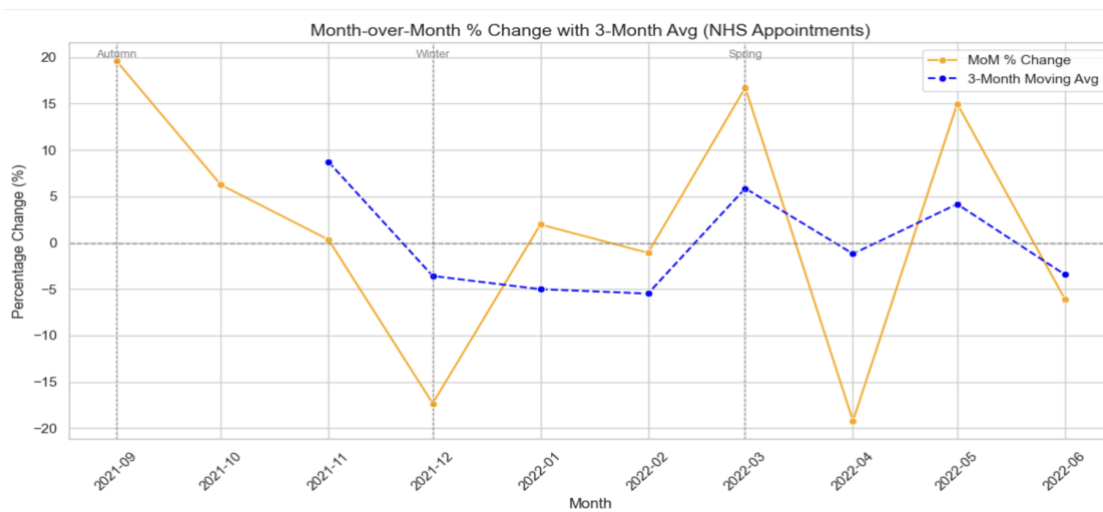


Figure 3 confirms that face-to-face care still dominates—averaging 17 million visits per month and never dipping below 60 % of total activity.

Figure 3: Appointments Over time by Mode

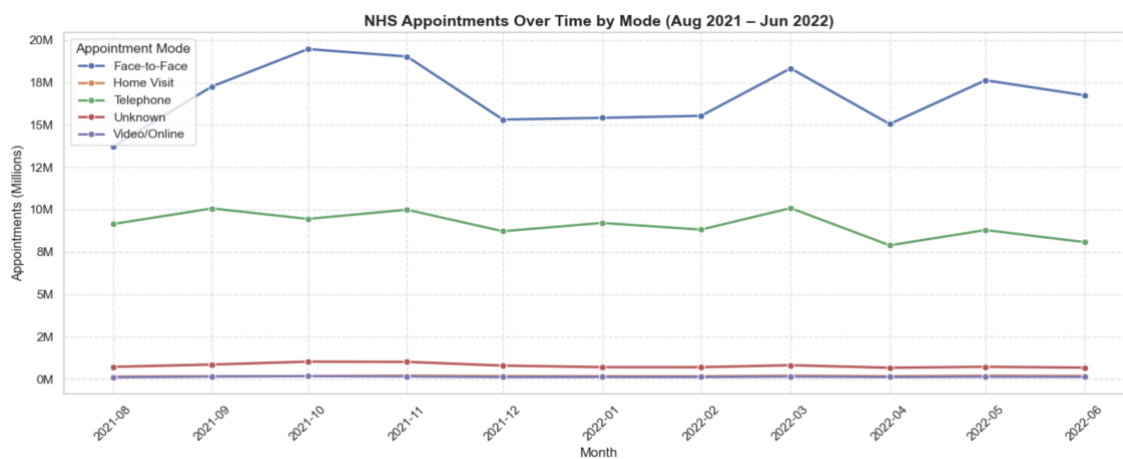


Figure 4 shows that short sessions dominate: 5- to 10-minute slots account for half of all activity, underlining the NHS's heavy reliance on quick GP encounters.

Figure 4: Appointments by Duration (Minutes)

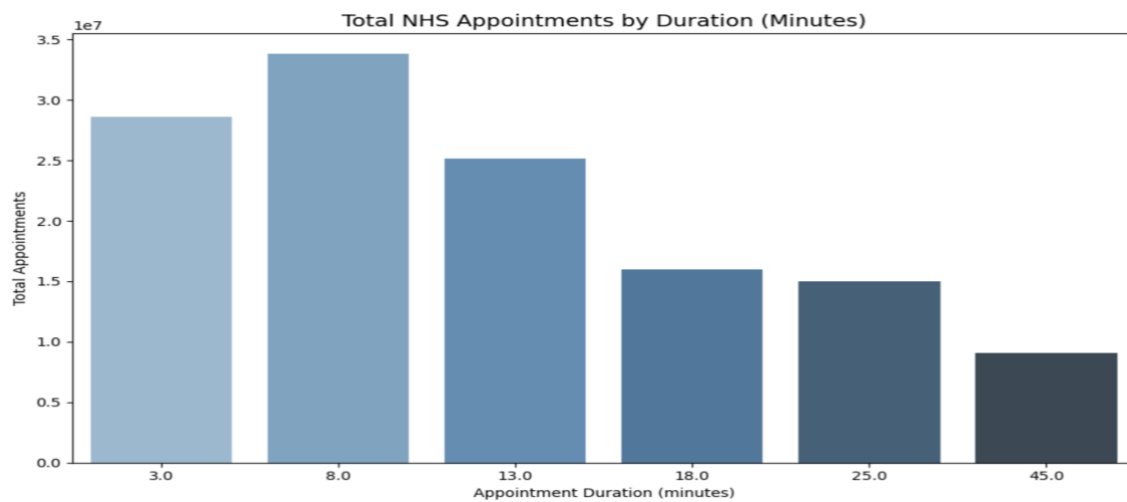
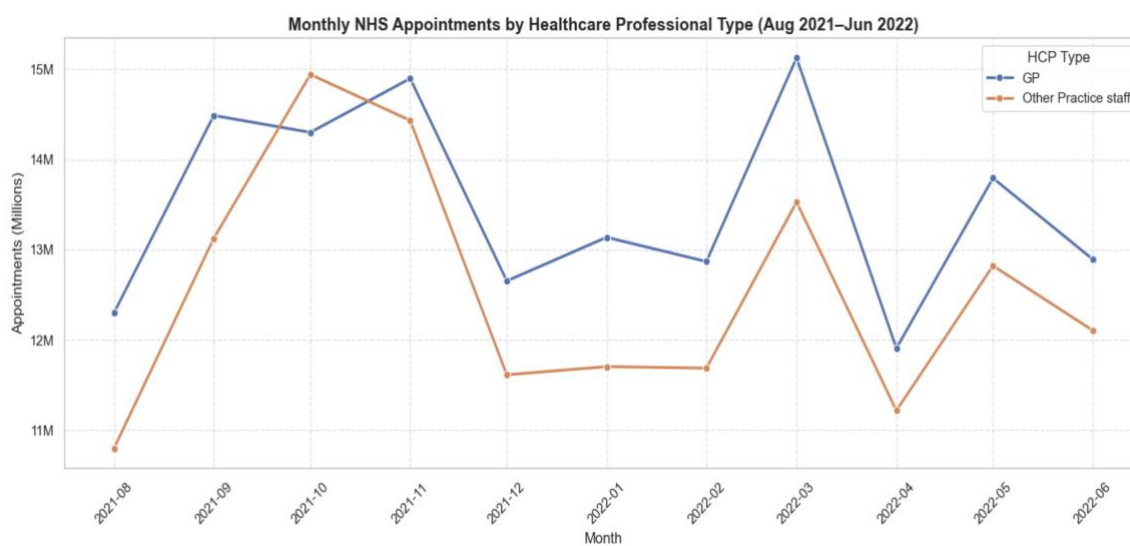


Figure 5 shows that, across the period, GPs consistently handle ~15 million appointments a month, double the load of “Other Practice Staff.” The two curves move in parallel, so allied staff are supplementing—not substituting—GP time. Peak GP spikes (Oct 2021, Mar 2022) coincide with DNA surges, signalling that redistributing routine follow-ups to nurses or pharmacists could smooth those pressure points.

Figure 5: Monthly NHS Appointments by Healthcare Professional



Seasonal patterns (Figure 6) show that NHS activity peaked in Autumn 2021, with over 89 million appointments—an average of nearly 29.8 million per month. This confirms strong post-lockdown recovery, driven by backlog clearance and rising demand. Spring 2022 followed closely with similarly high totals, while Winter 2021–22 saw a slight drop, likely reflecting seasonal disruptions (e.g. holidays, illness spikes).

Figure 6: Appointments by Season: Total and Average

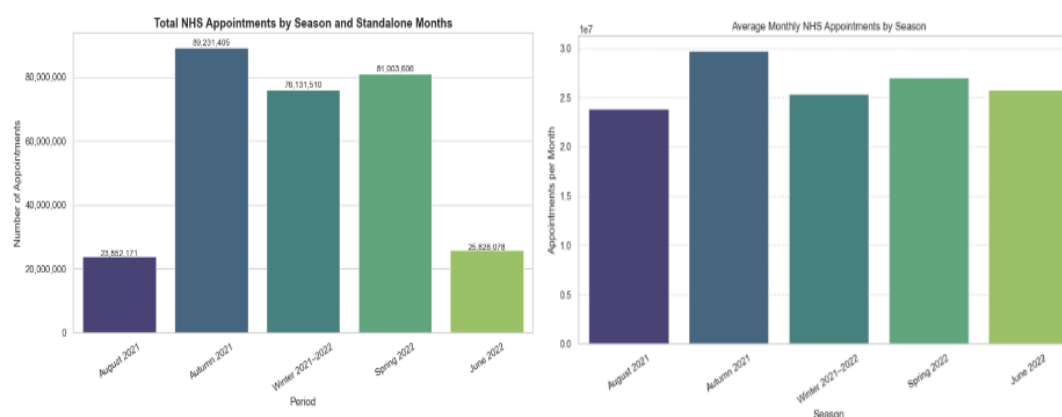
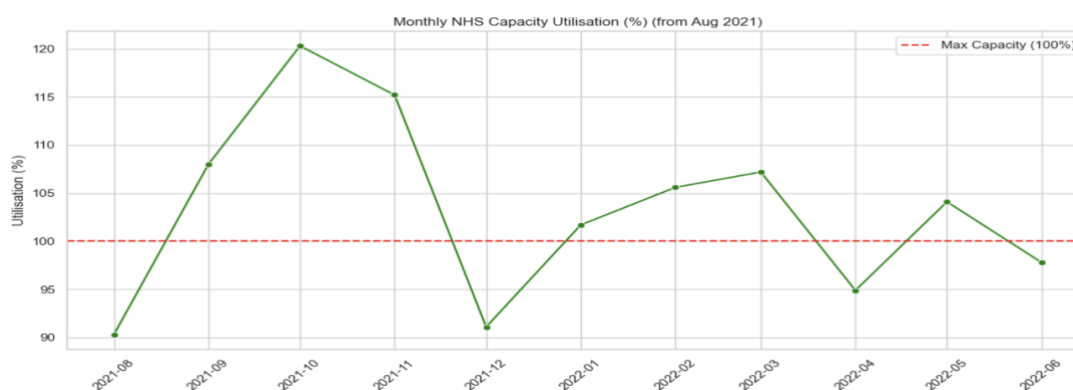


Figure 7 shows utilisation repeatedly breaching the 100 % planning ceiling—especially in October (120 %) and November (115 %). Even “quiet” months flirt with full capacity, leaving scant slack for shocks.

Figure 7: Capacity Utilisation Vs Limit



Implication: Temporary surges are no longer exceptional; structural overload is now routine. Dynamic rostering and surge buffers are essential.

3.2 Regional Inequality: Access vs Volume

Figure 8 contrasts raw volume with per-capita access. High-density ICBs (e.g., Manchester, London) command the largest totals but underperform on access, while rural areas (Cornwall, Herefordshire) top per-capita charts.

Figure 8: Volume vs Access (Total Appointments vs Per Capita)

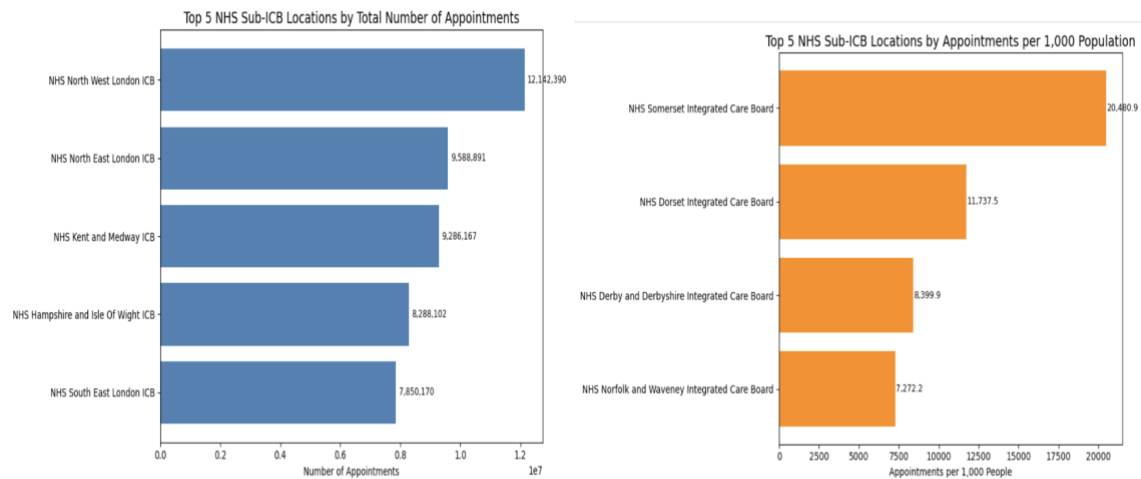
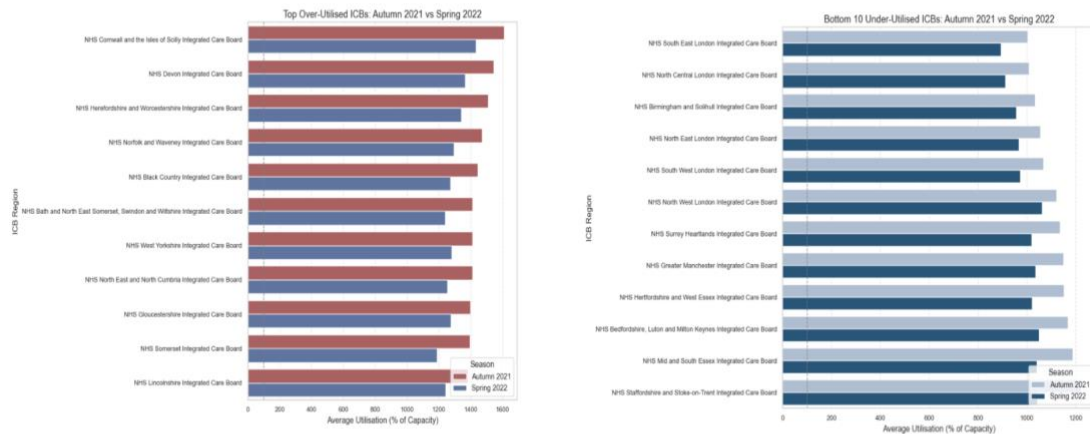


Figure 9 maps utilisation extremes: Cornwall operates at 130-150 %; South East London lingers below 100 %. Such mismatches suggest opportunity for cross-ICB redistribution of staff or telehealth capacity.

Policy cue: Equity demands directing extra resource to population-heavy but under-served urban ICBs.

Figure 9: Utilisation Inequality Chats (overutilised Vs Underutilised)



3.3 Missed Appointments – Patterns & Predictors

The chart in **Figure 10** displays the monthly DNA (Did Not Attend) rate from August 2021 to June 2022. The DNA rate peaked at **5.18% in October 2021**, following a steady increase from August. Afterwards, it declined to a low of **4.2% in January 2022**. From February to June 2022, the rate began to rise again, ending at **4.5%**.

Although these fluctuations appear relatively modest, their **impact is amplified when considering the actual volume of missed appointments** — shown in [Appendix B: Monthly Missed Appointments \(Volume\)](#). For example, in October 2021 alone, over **1.5 million** appointments were missed. According to the NHS England, with each slot costing an estimated **£30–£40**, even a 1% increase in the DNA rate can translate to **millions of pounds in losses**. This highlights the importance of tracking both percentage trends and absolute volumes to inform policy decisions.

Figure 10: Monthly Missed Appointments

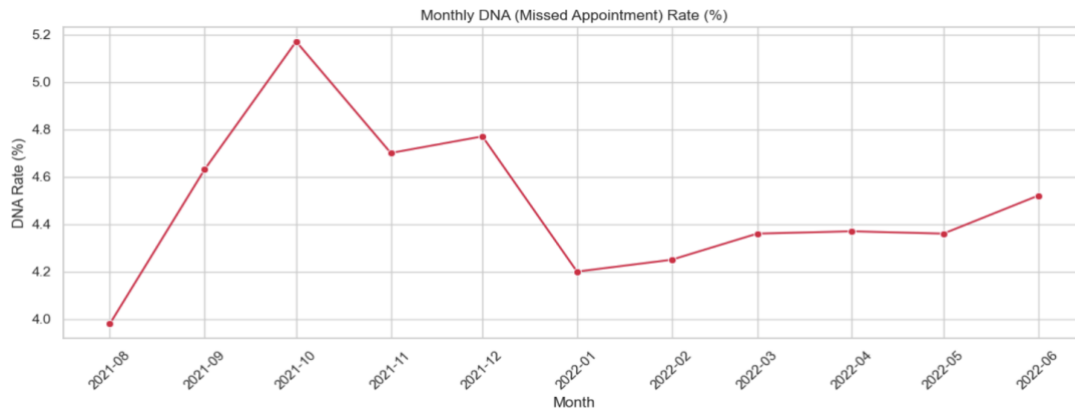
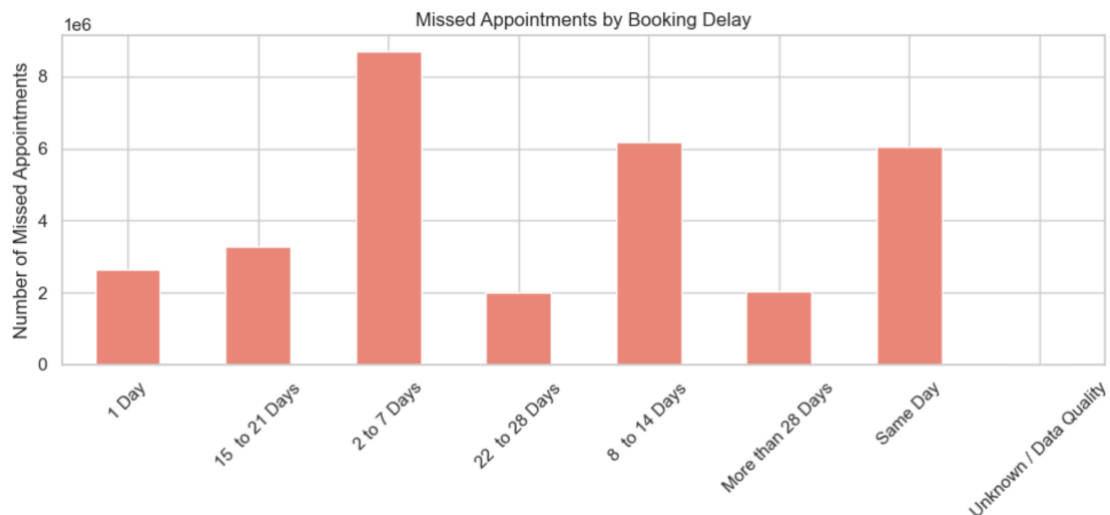


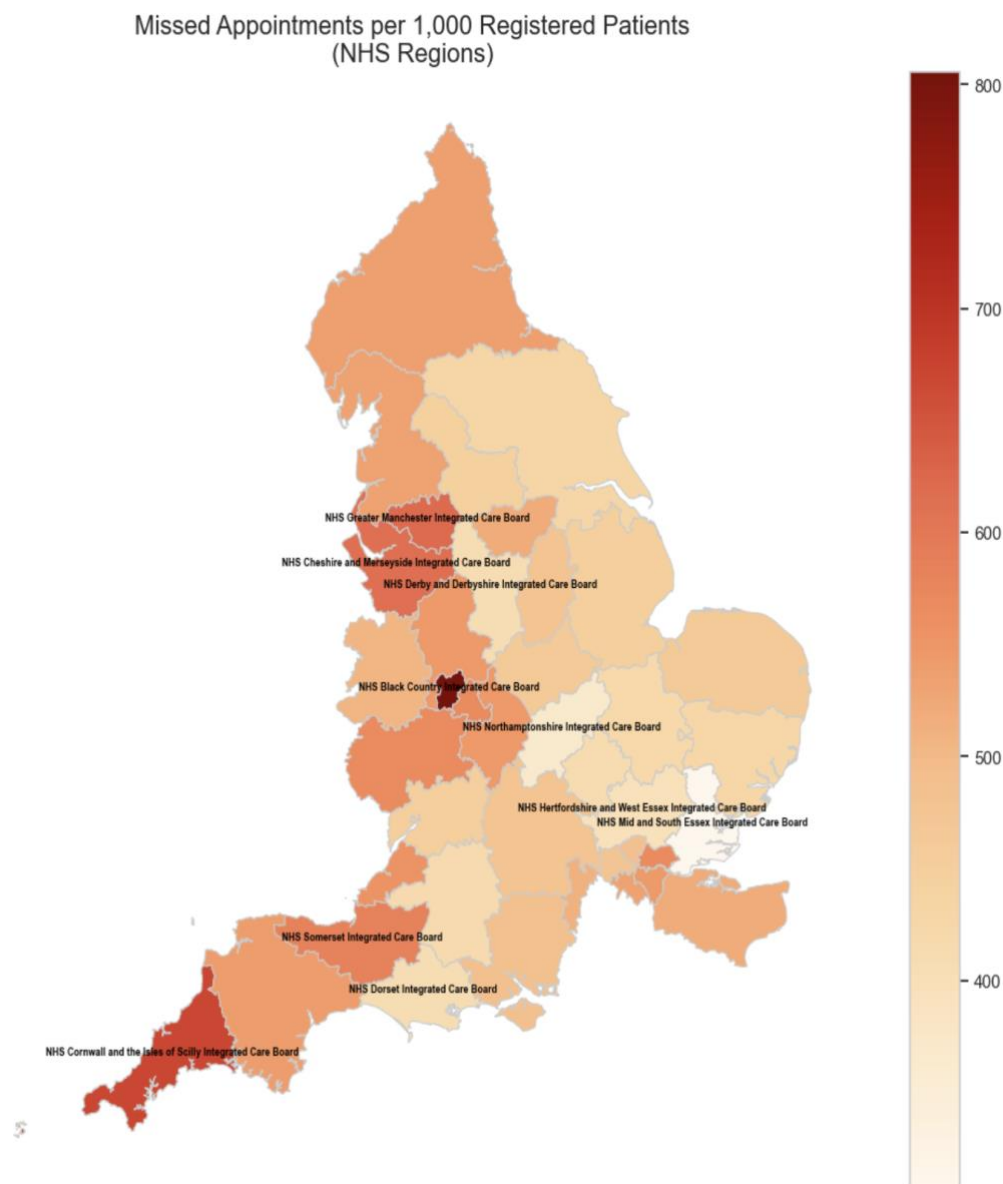
Figure 11 confirms lead-time risk: same-day and 2-7-day bookings generate the most no-shows; 8-14 days also elevate risk, likely through forgetfulness.

Figure11: Missed Appointments by Booking Delay



Map (Figure 12) highlights DNA hot-spots for targeted interventions.

Figure 12: Missed Appointments per 1,000 Registered Patients

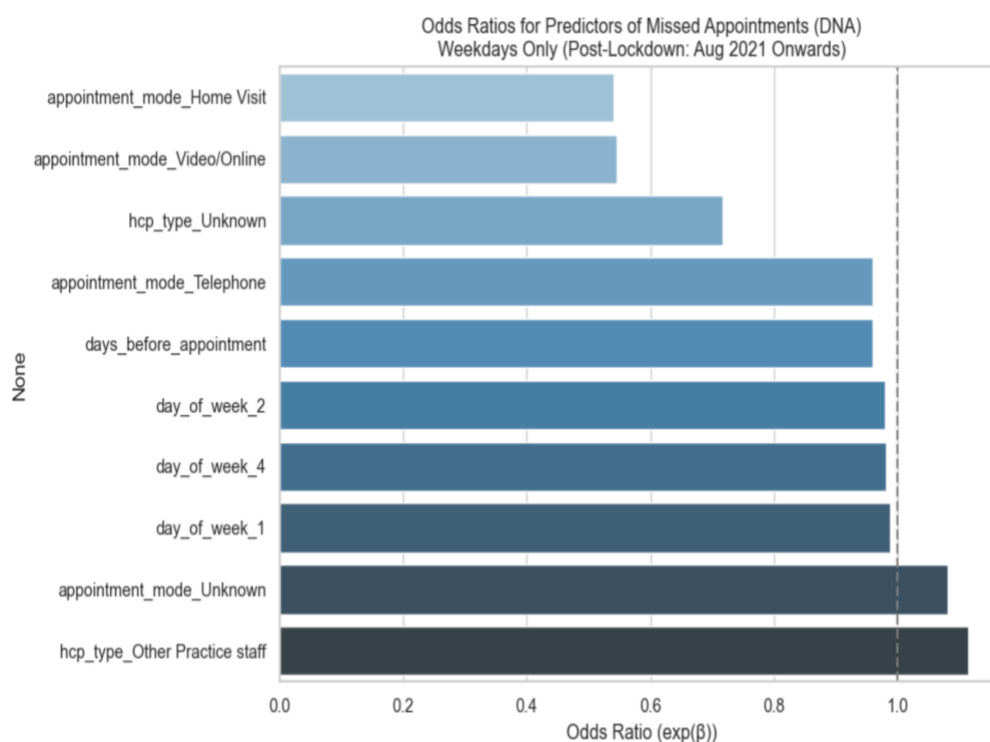


Logistic Regression (Figure 13)

- **Short-notice bookings (0-1 day)** sharply raise no-show odds.
- **Remote formats (telephone, video)** carry higher risk than in-person.
- **Non-GP staff** and **Mondays** show modestly higher DNA likelihood.
- Home visits and video/online *reduce* odds after other factors controlled.

Model accuracy: AUC 0.569 is only marginally better than random, reflecting limited features (no age, comorbidities) and data quality challenges. Thus, use results for **system-level guidance**, not individual prediction.

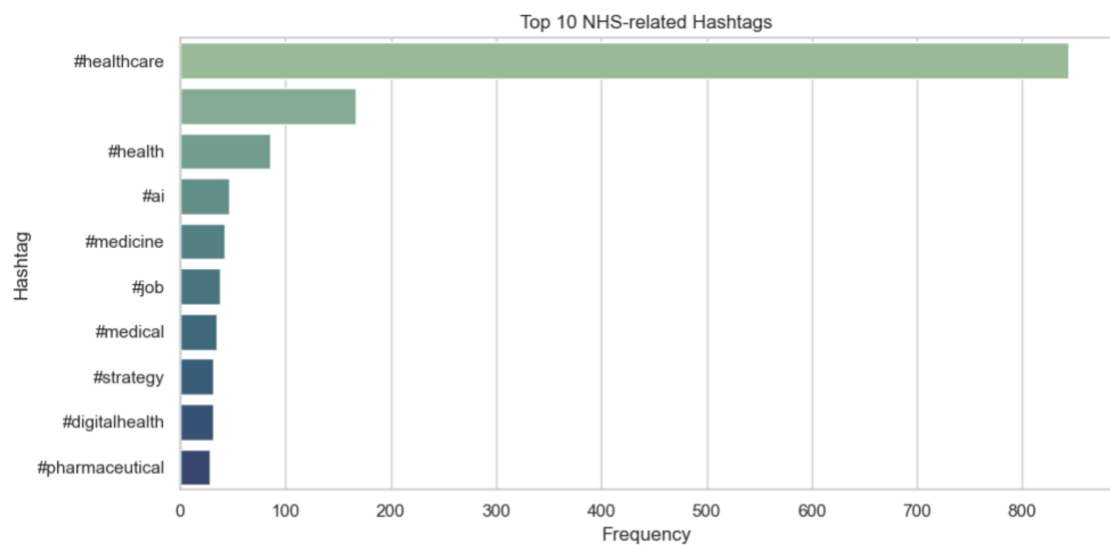
Figure 13: Predicting Missing Appointments: Logistic Regression Insights



3.4 Public Sentiment

Figure 14 shows that social media discussion around the NHS is dominated by the generic **#healthcare** tag, but a sizeable secondary cluster focuses on **digital transformation (#ai, #digitalhealth)** and **workforce issues (#job)**. Monitoring these three pillars gives an early-warning system for sentiment shifts and a high-impact channel for targeted messaging.

Figure 14: Top 10 NHS-related Hashtags



Conclusions

- **Demand:** Autumn and spring peaks sustain > 25 m appointments/month.
- **Mode:** Face-to-face remains > 60 % despite digital growth.
- **Capacity:** System runs > 110 % most months; regional imbalance persists.
- **DNA:** Short-notice and remote bookings drive no-shows (~5 %).
- **Public mood:** Generally neutral/positive but reacts quickly to crises—an actionable feedback loop.

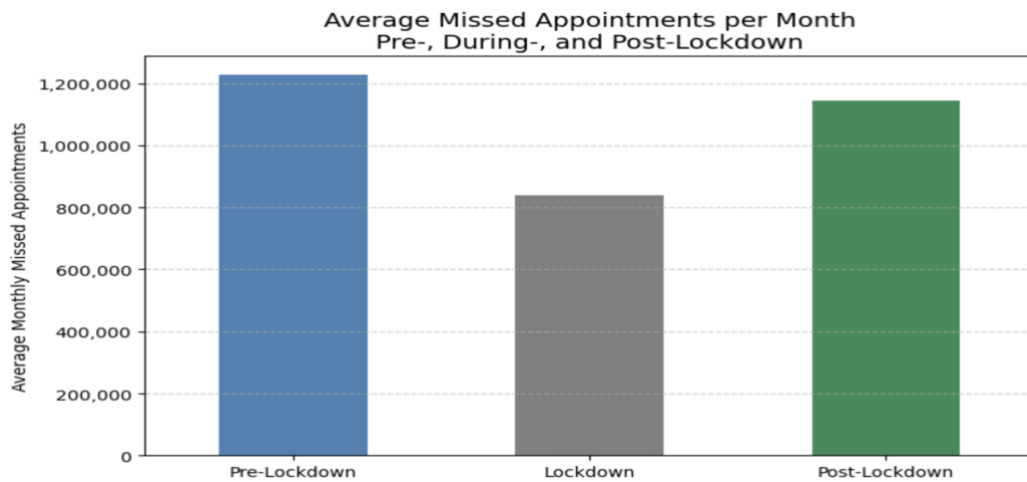
Recommendations

1. **Smart reminders:** SMS or app nudges for bookings ≤ 7 days; confirm remote slots proactively.
2. **Flexible staffing:** Dynamic rosters and locum pools ahead of expected seasonal dips.
3. **Regional rebalancing:** Direct funding and staff to high-demand, low-access ICBs; leverage tele-consult hubs.
4. **Data uplift:** Standardise capture of patient demographics and appointment metadata to boost predictive accuracy.
5. **Sentiment monitoring:** Track hashtags and polarity weekly; align communication campaigns to defuse emerging concerns.

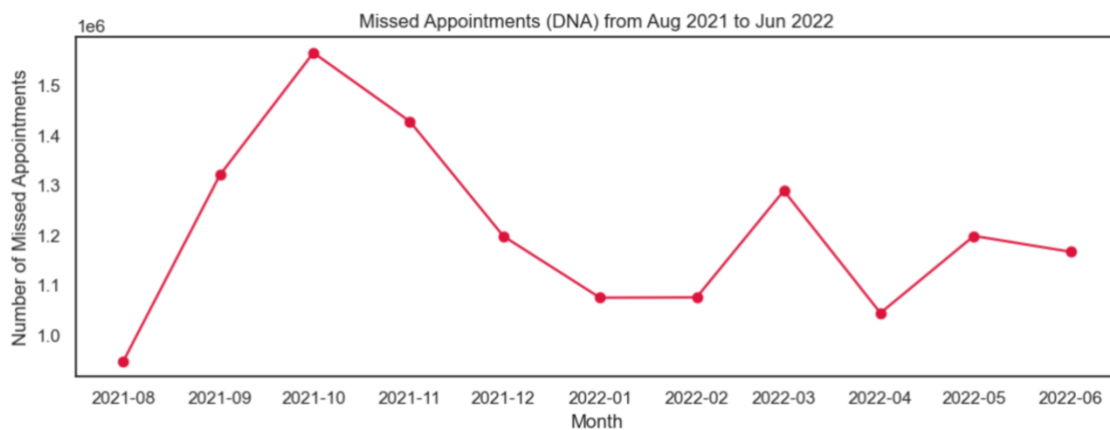
Implementing these actions can trim DNA losses, ease regional inequities, and protect frontline staff from chronic overload while keeping patient voices central to service redesign.

Appendix

Appendix A: Average Missed Appointments per Month



Appendix B: Monthly Missed Appointments (Volume)



REFERENCES

Florence Road Surgery (2022) *Missed appointments cost the NHS millions*. Available at: <https://www.florenceroadsurgery.nhs.uk/2022/05/30/missed-appointments-cost> (Accessed: 20 May 2025).

Hutto, C. and Gilbert, E. (2014) ‘VADER: A parsimonious rule-based model for sentiment analysis of social media text’, *Proceedings of the 8th International Conference on Weblogs and Social Media (ICWSM)*, pp. 216–225.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2021) *An introduction to statistical learning*. 2nd edn. New York: Springer.

NHS Digital (2023) *Appointments in General Practice*. Available at: <https://digital.nhs.uk/data-and-information/publications/statistical/appointments-in-general-practice> (Accessed: 26 May 2025).

NHS Digital (2023) *Appointments in General Practice: Metadata*. Supplied with dataset `metadata_nhs.txt`.

NHS England (2023) *Integrated Care Boards: data, boundaries, and planning guidance*. Available at: <https://www.england.nhs.uk/integratedcare/> (Accessed: 20 May 2025).

Office for National Statistics (ONS) (2023) *Population estimates by Integrated Care Board (ICB), England*. Available at: <https://www.ons.gov.uk/> (Accessed: 22 May 2025).