NHS Services Analysis: Infrastructure Utilisation and Healthcare Capacity

Introduction

The central goal of this analysis is to transform NHS data into actionable analytical insights to support operational and strategic decision-making. The key NHS stakeholders are seeking to understand trends in healthcare infrastructure and staff resources to guide decisions on whether to **expand or optimise existing structures**.

This analysis addresses two key business questions:

1. Has there been adequate staff and capacity across NHS networks?
2. What is the actual utilisation of NHS healthcare resources?

Our analytical approach focused on **trend detection and exploratory visualisation** to identify inefficiencies and seasonal variations in NHS service delivery over the available time period.

Data Overview

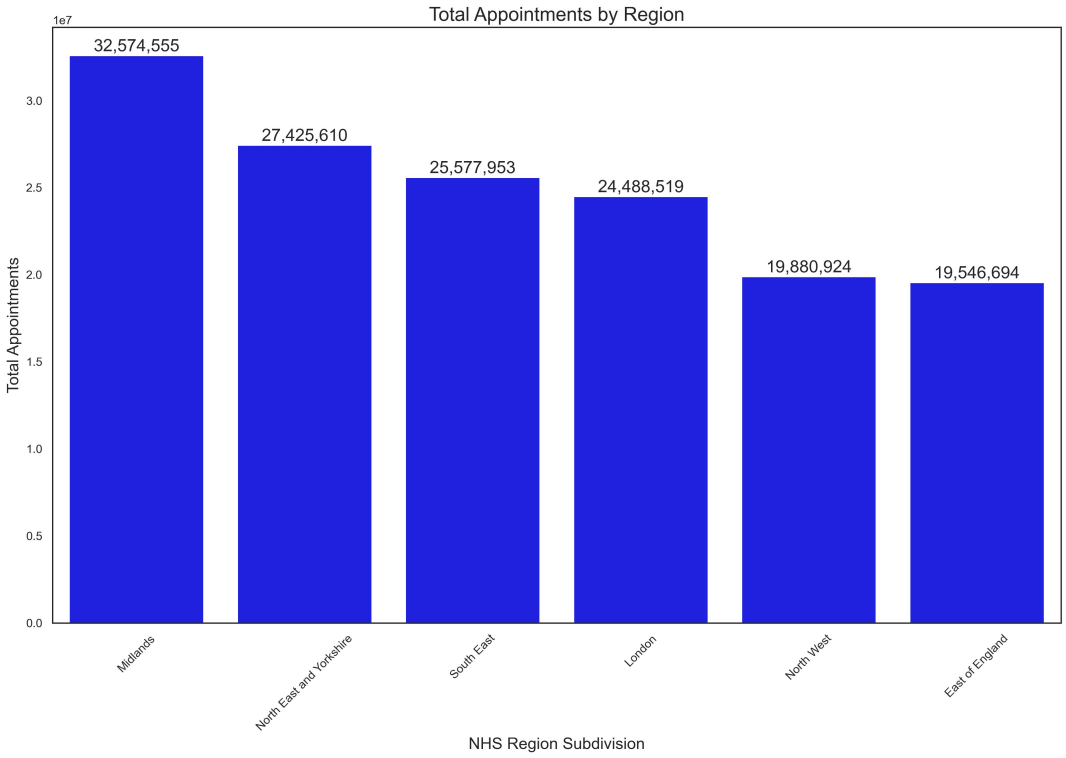
Three main datasets were used in this analysis:

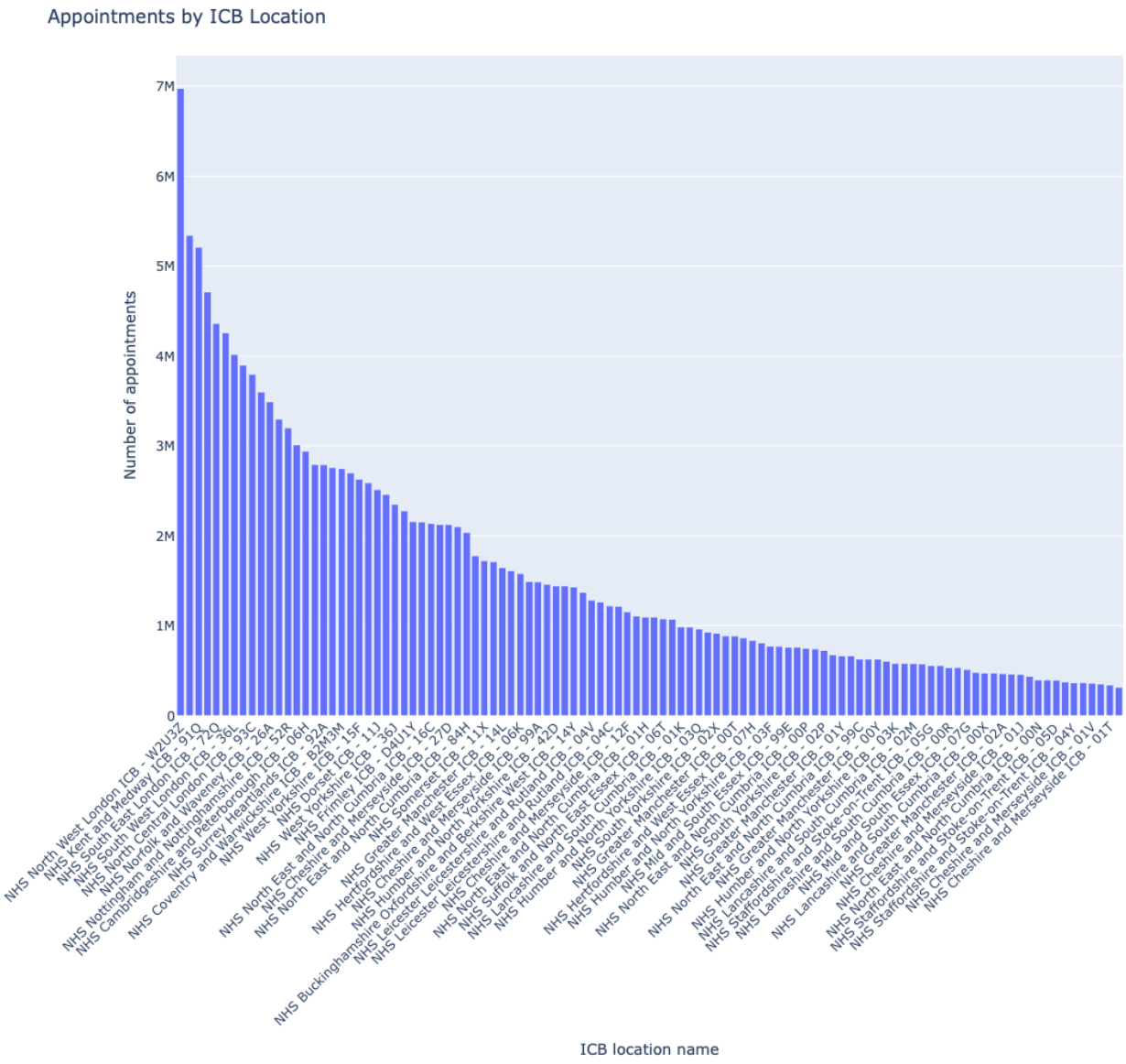
- `actual_duration`
- `appointment_regional`
- `national_categories`

Each dataset contains a *count of appointments* variable, allowing for efficient comparison of trends across features.

An initial import and validation process revealed no missing values; however, a significant number of "**unmapped" or "unknown" entries** were present. This data limitation restricted deeper correlation analyses in some areas.

The data spans **August 2021 to June 2022**, covering **106 locations across 7 NHS regions**. For improved clarity and interpretability, regional-level aggregation was prioritised over location-level analysis, as Integrated Care Board (ICB) data was too granular for effective visualisation.


Total Appointments by Region
1e7
32,574,555
27,425,610
25,577,953
24,488,519
19,880,924
19,546,694
Total Appointments
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Midlands
North East and Yorkshire
South East
London
North West
East of England
NHS Region Subdivision


Appointments by ICB Location
Number of appointments
0
1M
2M
3M
4M
5M
6M
7M
ICB location name

Descriptive and Trend Analysis

The data were analysed to detect **seasonal patterns**, **capacity fluctuations**, and **staffing challenges** across various dimensions, including:

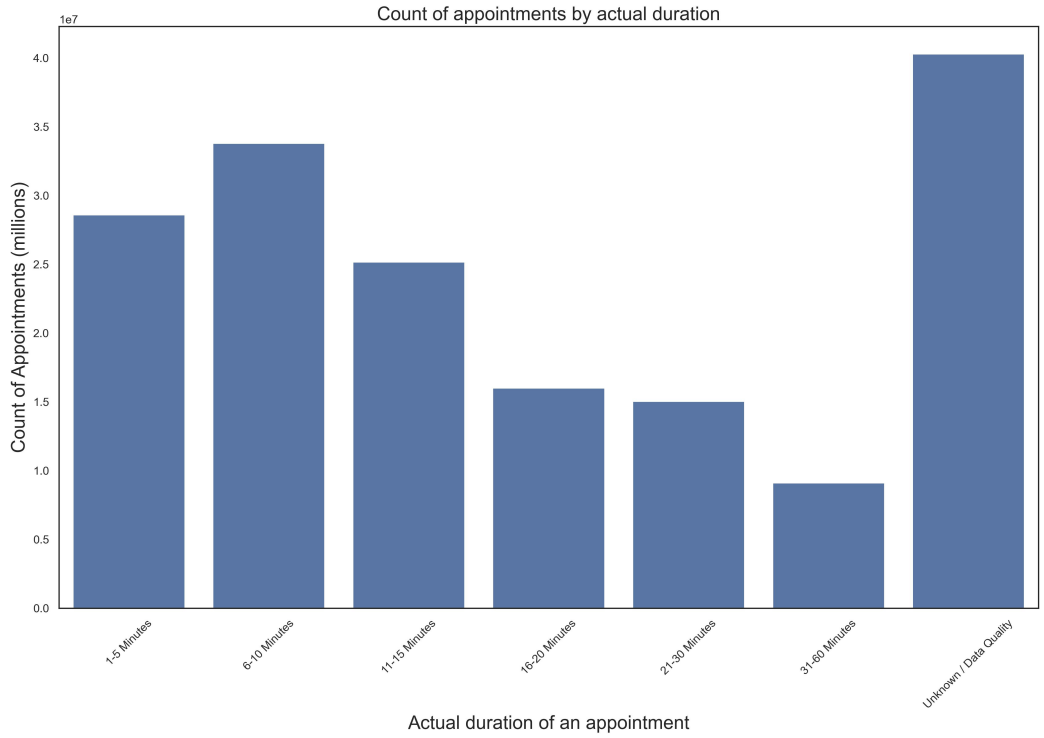
- Healthcare professional type
- Appointment mode (face-to-face, telephone, etc.)
- Attendance rates (attended vs. DNA – did not attend)
- Service context
- NHS service setting and facility utilisation

These focus areas were selected to produce meaningful, reproducible insights and lay the foundation for **regional comparisons** in future analyses.

Data Limitations

Despite the datasets being generally complete and well-structured, several limitations were identified:

- **Unknown or unmapped categories**: Large portions of data in key variables were labelled as 'unknown' or 'unmapped', compromising some analyses and indicating the need for improved data entry protocols.



- **Temporal constraints**: The 11-month window provides valuable insight into short-term and seasonal trends but limits the detection of **long-term structural changes**.

Key Findings

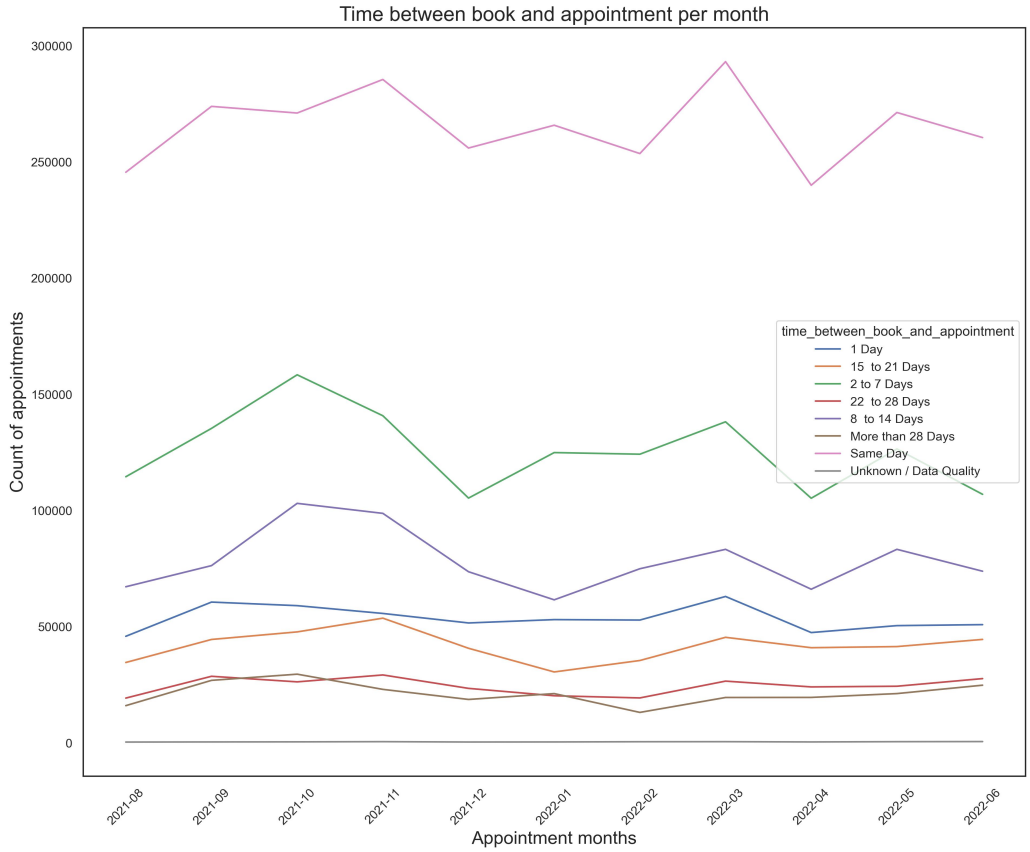
1. **Face-to-Face Appointments Dominate**
   Face-to-face consultations were approximately **double the volume** of telephone appointments, especially during peak periods.

2. **Seasonal Demand Patterns**
   Appointment volumes **peak in autumn and March**, with declines during winter and late spring. This suggests possible links to **holiday periods and seasonal illness** cycles.

3. **Scheduling Inefficiencies**
   A significant number of **DNAs (Did Not Attend)** and **extended wait times (8–14 days)** between booking and appointment indicate **inefficiencies in scheduling** during high-demand periods.

4. **Appointment Duration Distribution**
   Most appointments lasted **1–15 minutes**, though many entries were categorised as 'unknown'. Outliers under 1 minute or over 60 minutes likely reflect **systemic inefficiencies** or **data recording errors**.

5. **Utilisation Coefficient Analysis**
   A utilisation coefficient was developed to compare actual appointment usage against a benchmark of **1.2 million daily appointments**. The strong alignment between the daily appointment count (blue line) and utilisation coefficient (red line) suggests the benchmark is valid, highlighting periods of over- or under-utilisation across facilities.

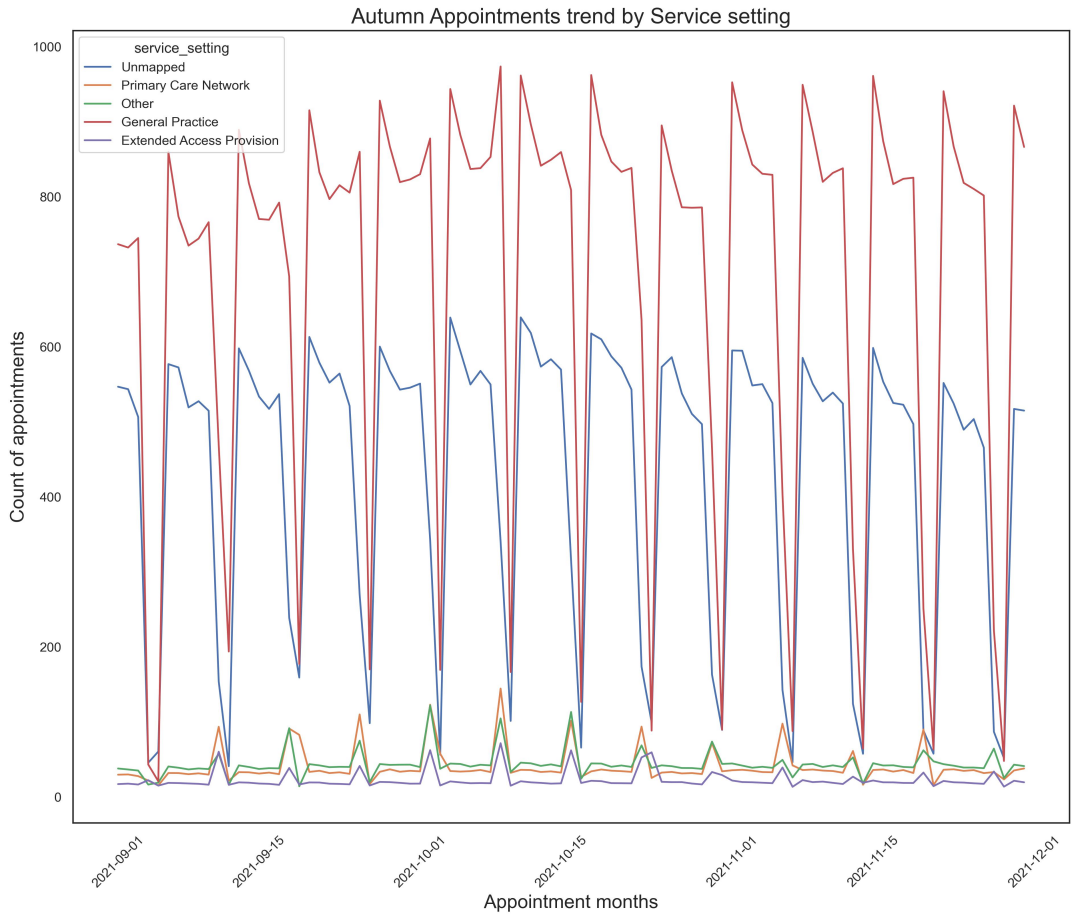


Visualisation Approach

To ensure interpretability, the following visualisation strategies were used:

- **Line charts** for temporal and seasonal trends (appointment volume, professional type, service setting).
- **Histograms** for appointment duration distribution.
- **Dual-axis line charts** for comparing appointment counts with utilisation coefficients.

Each chart was refined for clarity with labelled axes, legends, and concise titles to ensure insights could be **grasped at a glance**.



Twitter Data and Sentiment Insights

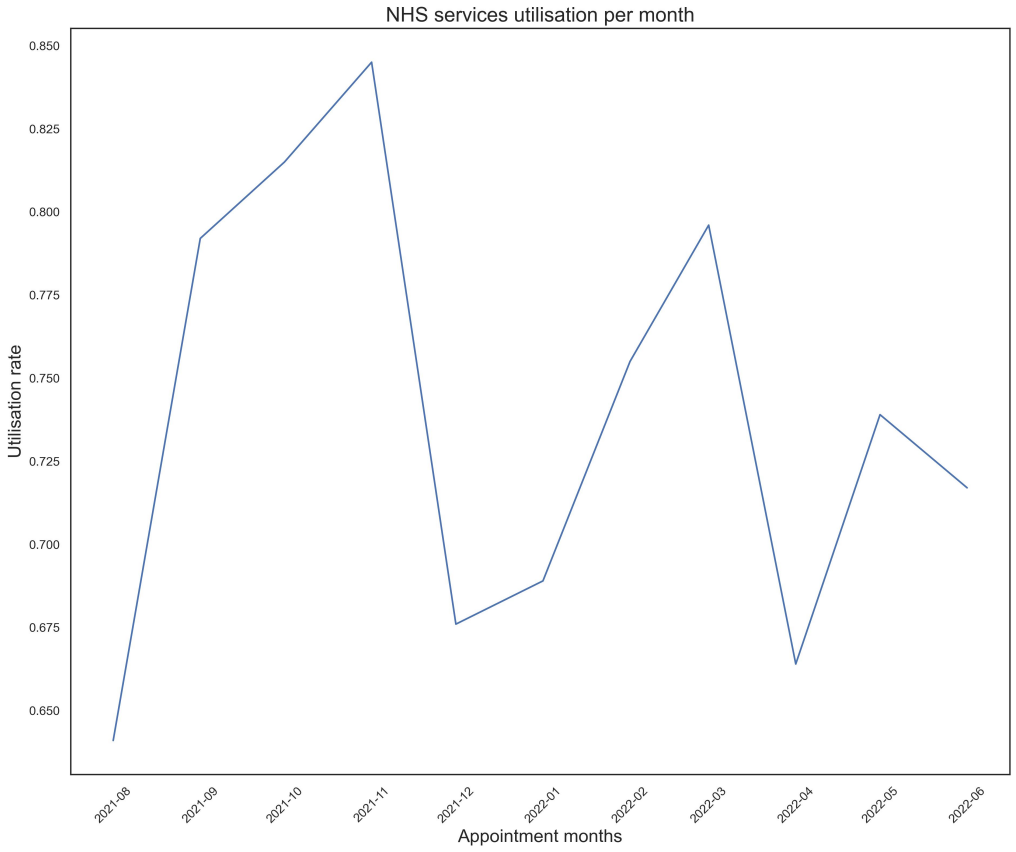
Social media sentiment analysis using Twitter data was performed to understand **public perceptions** of NHS services, including satisfaction, concerns about waiting times, and service accessibility.
The analysis revealed **no significant negative sentiment or widespread complaints** during the analysed period. Moreover, the nature of social sharing (likes and retweets) indicates a potential **multiplier effect** in the spread of NHS-related messages.

It is important to note, however, that social media sentiment often shows an **asymmetry of response** — individuals are generally **more inclined to express dissatisfaction when services perform poorly** than to comment when experiences are positive or routine. Therefore, the absence of negative sentiment should not necessarily be interpreted as strong approval, but rather as an **indication of general stability and the absence of major service issues** during the observation period.

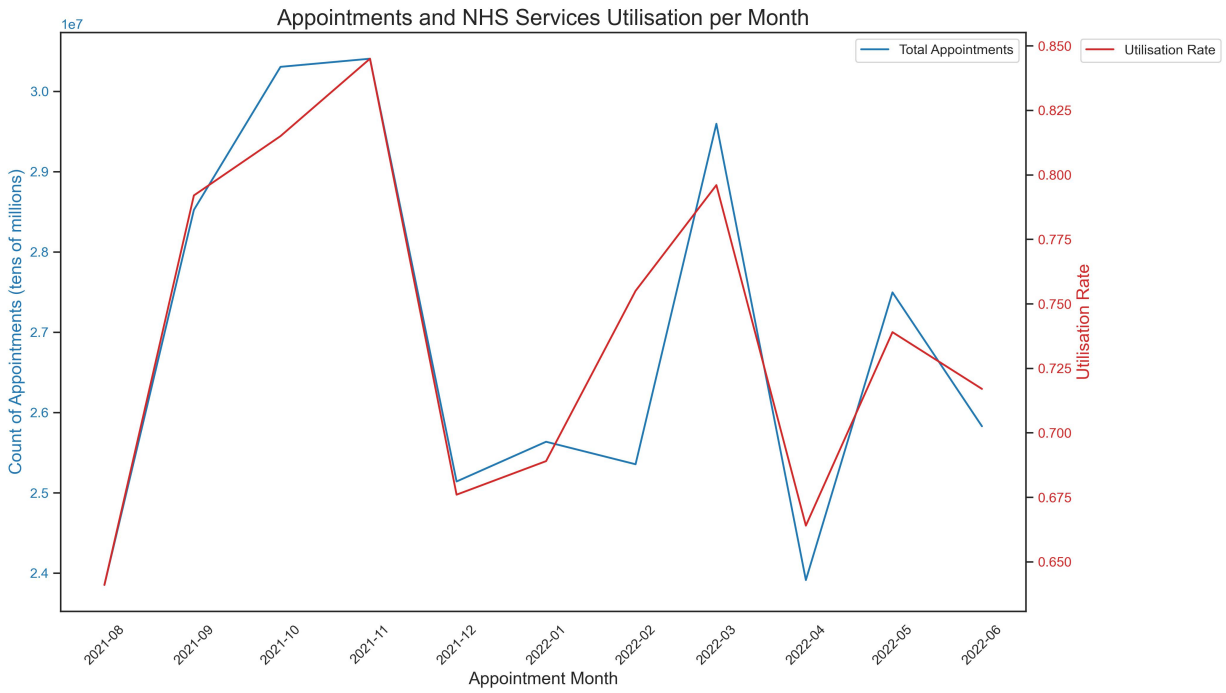
Summary of Insights

This analysis highlights critical aspects of NHS resource allocation and utilisation:

- **Seasonal Peaks:** Demand is highest in **autumn and March**, requiring proactive staff and resource planning.



- **In-Person Care Dominance:** Despite digital alternatives, face-to-face consultations remain the **core delivery mode**.
- **Unsatisfied Demand:** During peak months, appointment data indicate **unmet patient demand**, suggesting capacity shortfalls.
- **Systemic Data Issues:** Extreme appointment durations (<1 or >60 minutes) and unmapped categories reduce analytic precision and signal **data entry or classification issues**.



Recommendations

1. **Implement Dynamic Staffing Models**
   Use forecast-based scheduling to redeploy staff dynamically during peak demand periods.

2. **Enhance Patient Communication Tools**
   Reduce DNAs through digital reminders, rescheduling tools, or SMS rebooking options.

3. **Expand Remote Care Options**
   Encourage telephone and video consultations to alleviate pressure on in-person services where appropriate.

4. **Improve Data Quality and Consistency**
   Strengthen data entry validation, particularly for appointment duration and status fields, to improve analytical accuracy.

5. **Regional Deep-Dive Analysis**
   Conduct follow-up analyses at the **regional or ICB level** to identify location-specific improvement opportunities.