

NATIONAL HEALTHCARE SERVICE

BACKGROUND

In this report, we have delved into the rich datasets provided by the National Health Services (NHS) to address critical questions surrounding infrastructure utilization and patient engagement. Our exploration spans the realms of locations, service settings, and national categories, shedding light on trends, capacities, and potential resource optimization strategies. Furthermore, we investigate the impact of missed appointments, leverage Twitter data to identify key trends and hashtags, and evaluate the influence of appointment modes on patient morale. Our analysis aims to provide actionable insights, empowering the NHS with data-informed strategies for enhanced capacity planning, resource utilization, and patient engagement.

ANALYTICAL APPROACH

The data analysis journey commenced with the importation of diverse datasets – `actual_duration`, `appointments_regional`, `national_categories`, and `tweets`. In Python, I employed the pandas library to load the CSV and Excel files into DataFrames, providing a convenient structure for exploration and analysis. Preliminary inspection revealed the need for uniform date formats, necessitating the conversion of relevant columns to datetime objects for consistency. Addressing data quality, I conducted a thorough examination for missing values and outliers.

The alignment of datasets facilitated holistic analysis, allowing to answer key questions. Merging `actual_duration` and `appointments_regional` datasets on common identifiers empowered to gauge actual resource utilization against regional variations. The combination of `national_categories` with temporal information enabled the exploration of monthly and seasonal trends in service settings and national categories.

The analysis of appointment modes and their impact on patient morale involved statistical examination and visualizations. Leveraging Seaborn and Matplotlib, I crafted informative plots to unveil patterns in patient engagement and appointment modalities, paving the way for nuanced recommendations to the NHS.

Throughout the analysis, my decisions were driven by the overarching goal of answering the business questions posed by the NHS. The methodology prioritized data cleanliness, harmonization, and contextual relevance, ensuring that the derived insights are both robust and actionable for stakeholders.

Visualisation and insights :

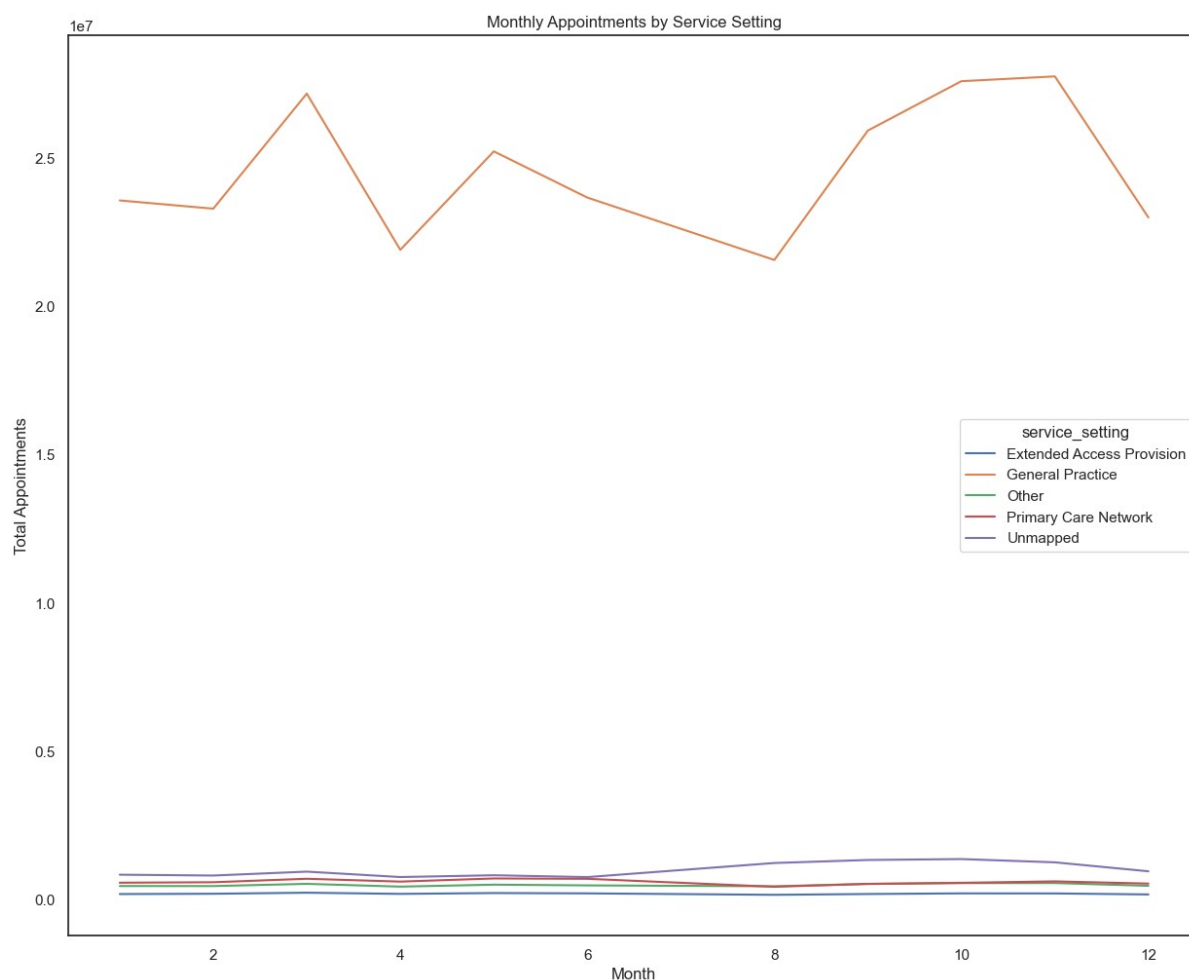
In presenting visualizations to our business stakeholders, we strategically chose lineplots and boxplots as they offer a nuanced perspective on the intricate dynamics within the NHS datasets, aligning with the overarching business objectives.

Lineplots were instrumental in conveying temporal trends, providing stakeholders with a comprehensive view of service utilization over time. These plots facilitated the identification of monthly and seasonal patterns in appointments, enabling stakeholders to discern peak periods, resource-demand fluctuations, and potential areas for capacity optimization. The choice of lineplots was deliberate, offering a clear depiction of trends and

enabling stakeholders to make informed decisions regarding resource allocation and infrastructure planning. Boxplots, on the other hand, were selected for their efficacy in visualizing the distribution of key metrics, such as the number of appointments and time intervals. These visualizations offered a succinct representation of the spread, central tendencies, and potential outliers within the datasets. The rationale behind our choice of visualizations was rooted in their ability to communicate complex information in a visually accessible manner. The visualizations served as a bridge between raw data and actionable insights, enabling stakeholders to make informed decisions aligned with the NHS's objectives.

Patterns and predictions

This comprehensive analysis of the National Health Services (NHS) datasets has revealed compelling patterns and predictions crucial for strategic decision-making. Firstly, the data underscores that visiting General Practitioners (GPs) stands out as the most prevalent mode of appointment among patients. This insight is pivotal for resource allocation and planning,



indicating the necessity for the NHS to focus on bolstering GP services to meet the prevalent demand

Secondly, the preference for face-to-face appointments emerges prominently in the data. This finding aligns with the essential need for direct, in-person healthcare interactions,

highlighting the significance of maintaining robust infrastructure to support such preferences. Tailoring services to accommodate and enhance face-to-face interactions may contribute to improved patient satisfaction and engagement.

```

Most popular appointment mode

r['appointment_month']=pd.to_datetime(ar['appointment_month'])

in_date = ar['appointment_month'].min()
ax_date = ar['appointment_month'].max()

rint("Minimum Date in nc DataFrame:", min_date)
rint("Maximum Date in nc DataFrame:", max_date)

tart_date = '2022-01-01'
nd_date = '2022-06-01'
iltered_ar = ar[(ar['appointment_month'] >= start_date) & (ar['appointment_month'] <= end_date)]

Determine the most popular service setting
ost_popular_appointment_mode = filtered_ar['appointment_mode'].value_counts().idxmax()

Print the result
rint("Most Popular appointment mode for NHS North West London from 1 January to 1 June 2022:", most_popular_appointment_mode)

View the output.

```

Minimum Date in nc DataFrame: 2020-01-01 00:00:00
 Maximum Date in nc DataFrame: 2022-06-01 00:00:00
 Most Popular appointment mode for NHS North West London from 1 January to 1 June 2022: Face-to-Face

Lastly, the analysis strongly suggests that the NHS must consider increasing its overall capacity. The consistently high demand for appointments across various settings signals the need for proactive measures to expand infrastructure and resources.

	appointment_month	count_of_appointments	utilisation
0	2020-01-01	27199296	9.066432e+05
1	2020-02-01	24104621	8.034874e+05
2	2020-03-01	24053468	8.017823e+05
3	2020-04-01	16007881	5.335960e+05
4	2020-05-01	16417212	5.472404e+05
5	2020-06-01	20690805	6.896935e+05
6	2020-07-01	22491437	7.497146e+05
7	2020-08-01	20150520	6.716840e+05
8	2020-09-01	26714255	8.904752e+05
9	2020-10-01	28301932	9.433977e+05
10	2020-11-01	25061602	8.353867e+05
11	2020-12-01	23535936	7.845312e+05
12	2021-01-01	22492069	7.497356e+05
13	2021-02-01	22399569	7.466523e+05
14	2021-03-01	27225424	9.075141e+05
15	2021-04-01	23879932	7.959977e+05
16	2021-05-01	23508395	7.836132e+05
17	2021-06-01	26784182	8.928061e+05
18	2021-07-01	25739219	8.579740e+05
19	2021-08-01	23852171	7.950724e+05
20	2021-09-01	28522501	9.507500e+05
21	2021-10-01	30303834	1.010128e+06
22	2021-11-01	30405070	1.013502e+06
23	2021-12-01	25140776	8.380259e+05
24	2022-01-01	25635474	8.545158e+05
25	2022-02-01	25355260	8.451753e+05
26	2022-03-01	29595038	9.865013e+05
27	2022-04-01	23913060	7.971020e+05
28	2022-05-01	27495508	9.165169e+05
29	2022-06-01	25828078	8.609359e+05