The business problem

The National health service (NHS) has identified that **missed GP appointments incur a significant cost** which it would like to reduce or eliminate. Each missed GP appointment costs the NHS an average of £30 per person (Source: NHS).

Specifically, the NHS wish to better understand **their network capacity and utilisation of resources.**

Additionally, **seasonality, booking interval, appointment types and locations** will be investigated to uncover correlations with missed appointments.

The intended audience of this report is those with a technical background in the NHS and government.

Analysis approach

For ease of manipulation, I imported the four datasets provided into a jupyter notebook and converted them into dataframes. I utilised the pandas python library to aid me in importing and exploring the datasets, and seaborn and matplotlib libraries for effective visualisations.

My initial review determined that there was no missing data and that the data integrity was sufficient for robust analysis, i.e. no erroneous data. The metadata file also states that the datasets have been cleaned, providing further reassurance.

I ensured that the variable names used were intuitive, especially when cascaded into subsets. For example, every dataframe was prefixed with the appropriate datasource abbreviation: actual_duration as ad, etc.

In order to make the datasets more manageable when exploring the data, I employed the groupby and sort_values functionality. This provided direction for future analysis by highlighting the most relevant criteria to the business problem.

Appointments_regional was the only dataset to have missed appointments data, so was considered the best source to be investigated, despite the date being by month.

1. Network capacity and utilisation of resources

- It had been suggested that the NHS has capacity for 1.2 million appointments per day. However, I have reason to doubt this claim, as appointments per day in practice will vary due to patient demand, appointment length, staff availability, among others. This is a limitation.
- Without knowing capacity, I determined utilisation by calculating the attended appointments divided by the sum of attended and missed appointments.
- By 'unstacking' appointments by status and adding a calculated field, utilisation by month could be established.
- A way to interpret spare capacity in the network is to analyse how many appointments are booked and attended on the *same day*.

2. Seasonality

- By changing the date format, I determined that the date ranges differed by dataset.

3. Booking interval

- By plotting and reviewing a bar-chart, I decided to group appointments booked more than 14 days in advance into a new category to allow for a more meaningful comparison.

4. Locations

- I was able to determine that there are 42 locations (with 106 icb entity codes).
- Merging two subsets allowed me to list known appointments besides not attended appointments, allowing comparisons to be made and the largest contributors to the business problem be understood.
- There are multiple icb_ons_codes per location. Once clarified, these ranked lists may change.

Rank of locations with the highest absolute number of missed appointments:

	index	icb_ons_code	known appointments	not attended appointments	not attended %
0	2	E54000057	32256474	1946114	16.574812
1	0	E54000050	41051227	1676790	24.482032
2	3	E54000008	31540433	1658787	19.014155
3	5	E54000029	22584210	1317005	17.148158
4	4	E54000027	28135234	1286262	21.873642
5	1	E54000054	34342321	1159691	29.613338
6	12	E54000030	19266029	1107753	17.391990
7	17	E54000062	16673410	1023176	16.295740
8	10	E54000032	21441446	1015035	21.123849
9	11	E54000048	20546511	964458	21.303687

- NHS Greater Manchester ICB - 00T (E54000057) has the highest absolute number of missed appointments, but with a relatively low percentage.

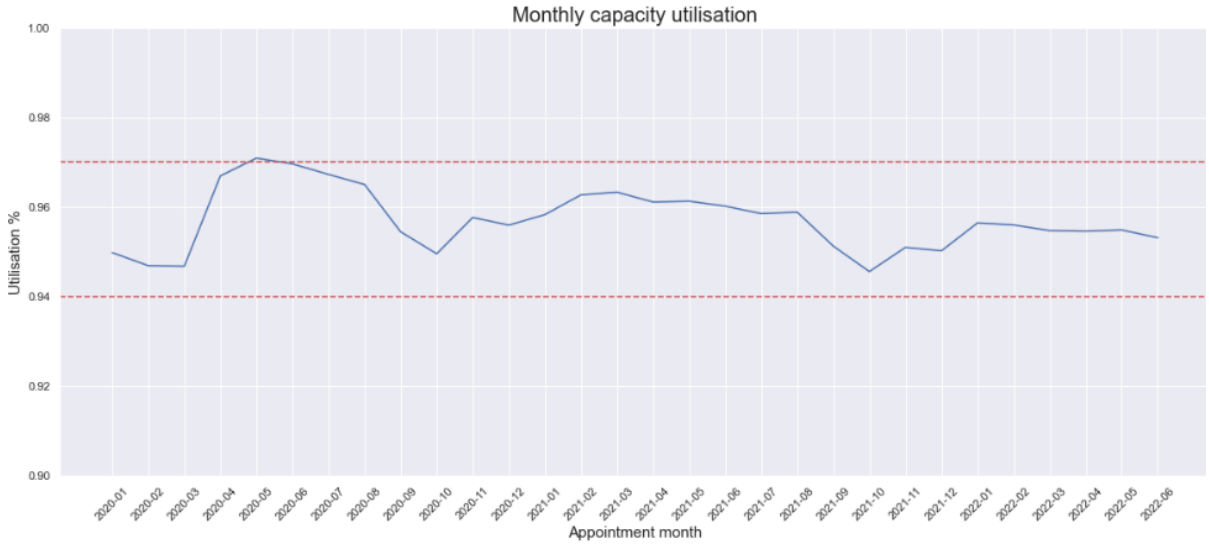
Rank of locations with the highest proportion of missed appointments:

	index	icb_ons_code	known appointments	not attended appointments	not attended %
0	24	E54000026	13530234	383465	35.284143
1	36	E54000059	9542996	295150	32.332699
2	26	E54000040	12933306	404831	31.947420
3	34	E54000041	10562706	333381	31.683587
4	23	E54000058	13980063	450691	31.019175
5	1	E54000054	34342321	1159691	29.613338
6	22	E54000022	14720364	498990	29.500319
7	25	E54000023	13088676	443964	29.481390
8	28	E54000056	12361323	422986	29.223953
9	39	E54000043	8765625	302390	28.987814

- NHS Mid and South Essex ICB - 06Q (E54000026) has the highest percentage of missed appointments.

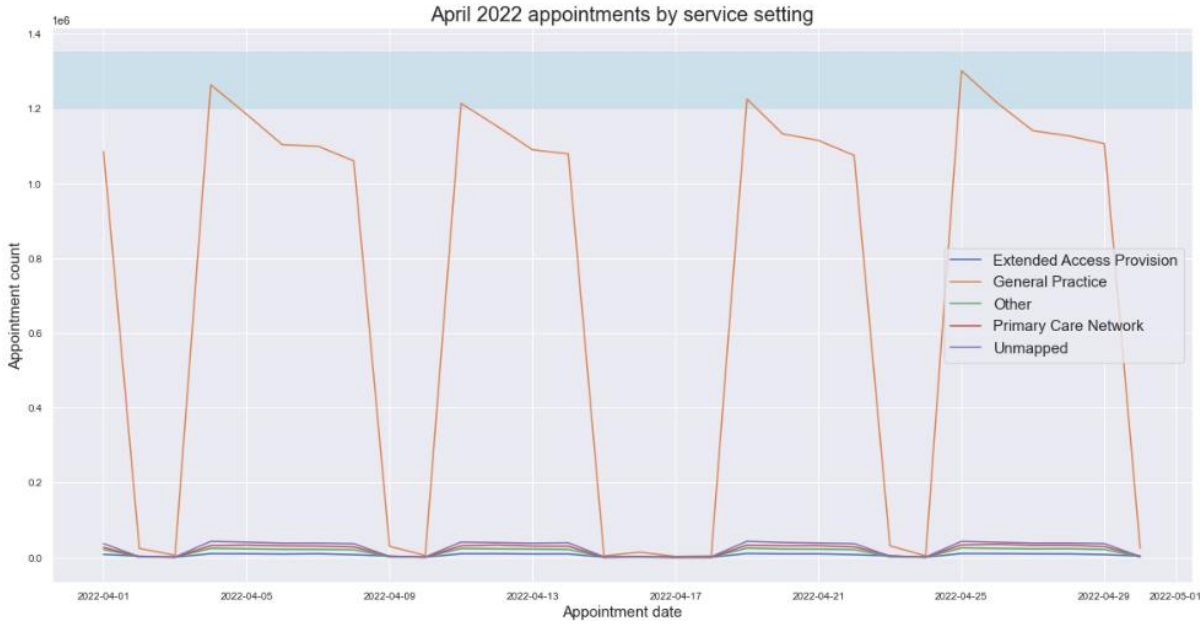
Visualisations & insights

1. Network capacity and utilisation of *known* data is consistently very high



A line-plot best demonstrates the relative consistency of all known appointments that were attended, as they remained between 94-97% on average. Therefore, missed appointments equate to a very low percentage of the total.

2. Seasonality by week is prominent

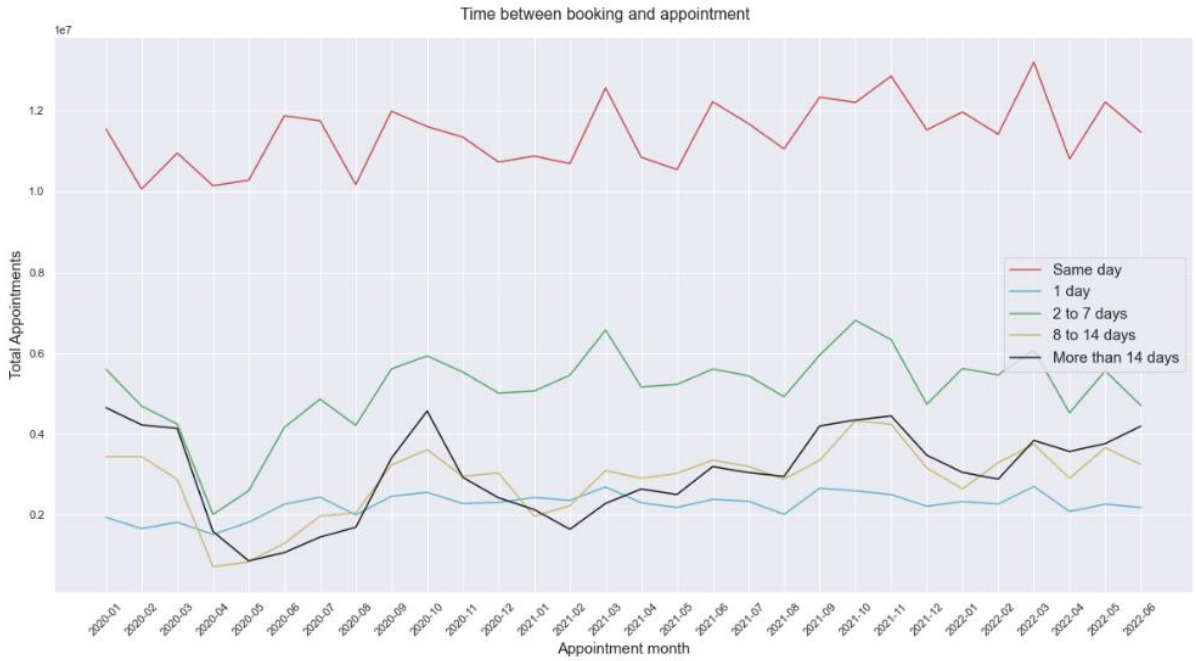


Line-plot demonstrating the weekly seasonality of appointments:

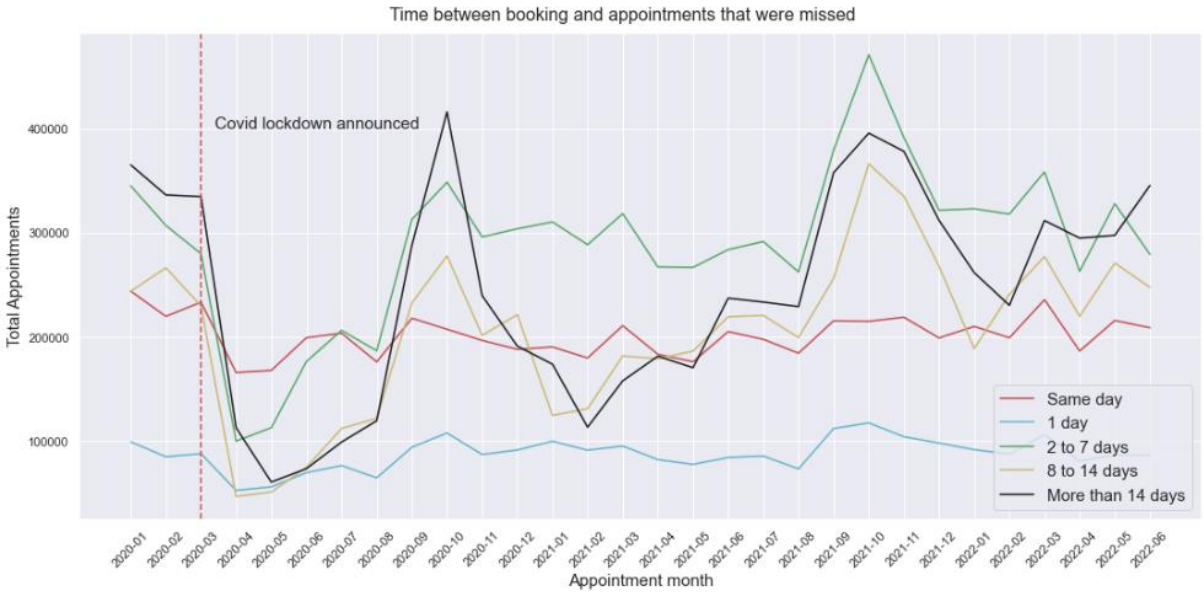
- The blue band highlights the peaks that correspond to Mondays.
- Wednesday to Friday sees a relatively consistent number.
- Weekends correspond to a drastic reduction in appointments.

This is a snapshot of April 2022 to demonstrate the seasonality shape, but the dominance of 'General Practice' and weekly seasonality are consistent across the entire dataset.

3. Booking intervals offer an insight into missed appointments



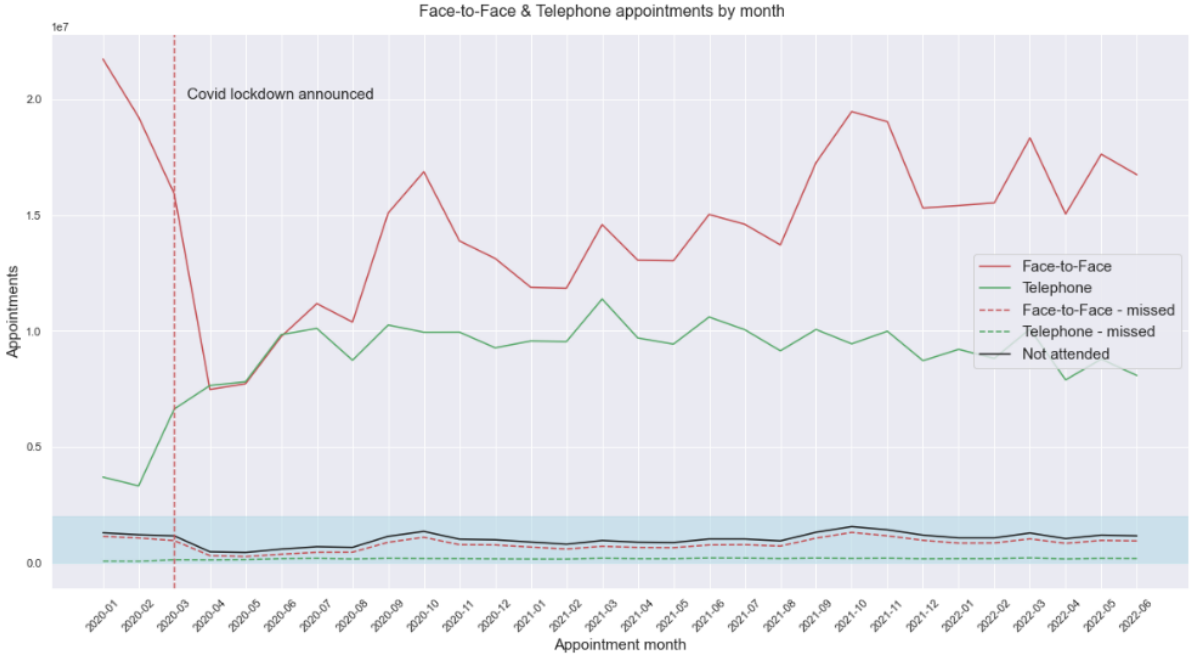
- The majority (55%) of all appointments were booked for the same day or next day.
- When we compare this to the line-plot of missed appointments (below) and when they were booked, we see that same day or next day appointments are <u>not</u> the main contributing factor towards the missed appointments.
- Intriguingly, there are a significant number of missed appointments that were booked the same day.



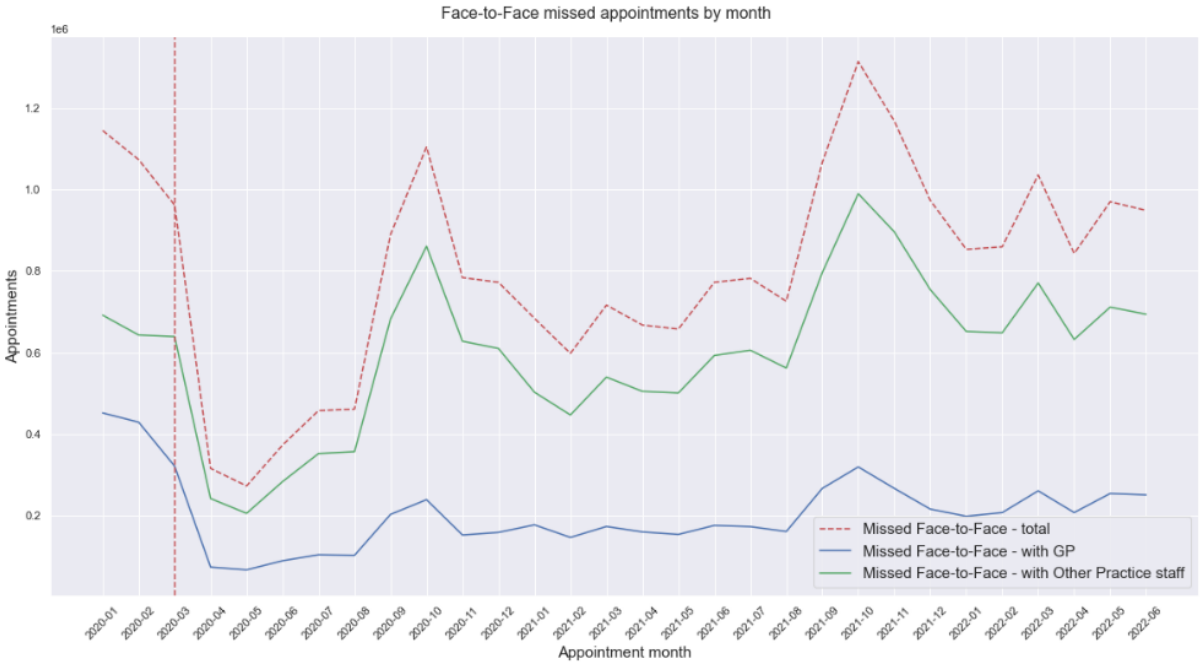
This line-chart demonstrates the volatility of missed appointment by booking intervals.

- There are two distinct peaks of missed appointments in 2020-10 and 2021-10.
- Missed appointments booked for the next day remain relatively low.
- The height of the 'more than 14 days' category indicates that booking appointments way in advance gives no better reassurance of attendance.

4. Certain Appointment types are more likely to be missed

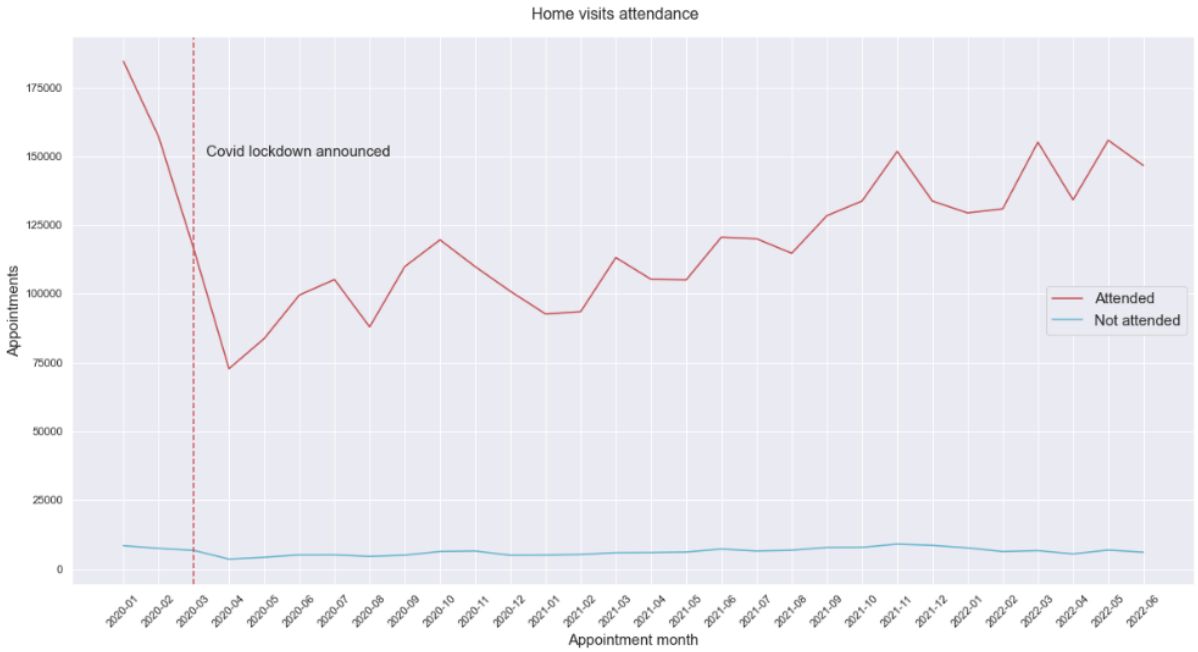


- This plot is focussed on face-to-face and telephone appointments which are prominent in the dataset, with associated missed appointments by type included (dashed lines).
- This demonstrates that the main drivers of total missed appointments are the face-to-face type (shaded blue bar), indicating further analysis should be focussed here.

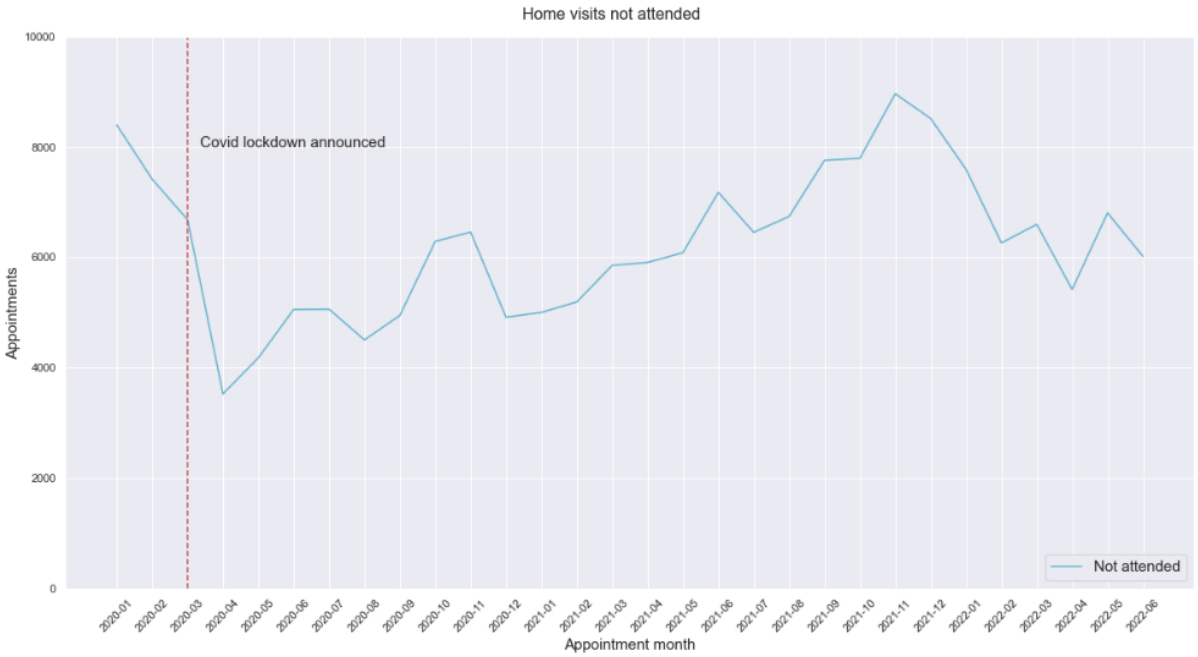


- You can see that the majority of missed face-to-face appointments are with ‘other practice staff’.

5. Missed Home visits are in proportion with other types



- Since March 2020, home visits have been steadily increasing with non-attendance seemingly remaining quite flat.
- However, when missed home visits are plotted on their own (below), you can see the shape of the line correlates fairly well with attended visits (above).
- Therefore, we can conclude that missed home visits are proportionate to total visits.
- Perhaps this is surprising as the NHS staff are travelling to the patient. Considering that home visits only account for 1% of appointments, further investigation should not be prioritised.



Patterns & trends

1. Network capacity and utilisation of resources needs further investigation

- Limited conclusions can be drawn regarding utilisation of resources.
- A reasonable level of spare capacity is required for urgent appointments.
- To enable a better understanding, capacity limits could be calculated by practice, knowing staff availability.

2. Weekly seasonality is evident but further work required

- Knowing that appointment numbers spike on Mondays and drop at weekends suggests that more weekend resource could be required. This would help weekday workers attend appointments who would otherwise need to take time off. However, this may result in greater NHS staff costs.
- Linking seasonality to missed appointments is the next logical step to investigate.

3. An appointment booked the same or next day's maximises attendance likelihood

- Perhaps there is an overconfidence in patients' availability when booking more than 1 day ahead.
- Appointments being missed is on the gradual increase.
- Therefore, any remedial actions should be implemented promptly.

4. The cost of different appointment types should be taken into consideration

- With the average GP salary range being £65,000-£98,000 it can be assumed that they earn more than most other practice staff, therefore any missed GP appointments cost more.
- Further iterations of analysis should look to apply better cost estimates to missed appointments.
- It could be assumed that patients are more likely to miss less urgent appointments with, say, a physiotherapist, rather than with a GP.

5. Missed appointments significantly vary by location

- Any practical steps to minimise missed appointments should focus on the locations identified as having the greatest absolute and greatest relative proportion of missed appointments.