

Assignment Activity 3 Summary

First I imported pandas and made DataFrames from the relevant data.

I did the normal checks of header and data types.

For the question of which appointments were scheduled I need to use the DateTime module to handle and format the dates correctly.

Then I used pandas to format 'appointment_date' correctly.

I then used the print and min/ max function to see the earliest and latest dates.

This gives the answer 2021-08-01 00:00:00 2022-06-30 00:00:00.

For the question "Which service setting reported the most appointments in North West London from 1 January to 1 June 2022?" I needed to extract only the NW London information.

I created a subset DataFrame with only those dates:

```
nc_subset = nc[nc['sub_icb_location_name'] == 'NHS North West London ICB - W2U3Z']
```

Then I needed to organise it by service setting, and use the sum and sort_values functions to see the types of service setting BY count of appointments.

Lastly I performed a similar operation for the question

#Which month had the highest number of appointments?

I made another subset of the NC dataframe, which groups by the column 'appointment month'. Then I used the sum function to create a second subset DataFrame with that grouping visible (not just the generic DF object that shows where the information is stored on your computer).

Then I organised that last subset dataframe by "count of appointments" and used the sort_values function, with Ascending = false to give a table of months with the most number of appointments first.