

Assignment activity 2: Import and explore the data

Summarise and explain in fewer than 200 words any insights and trends you have identified while answering the questions.

This activity allowed me to familiarise myself with the data and ensure that I am appropriately able to review, extract and analyse data as required.

Analysis approach:

- Values_counts was the primary code used for the initial analysis. There was sum customisation that was required in extracting a particular string.
- There is much for me to improve on the f-string printing but this is a very good way in ensuring the results are accurate and communicates as is.
- Through the analysis, I was able to do have a better grasp on the content than the info overload in the metadata. This allowed me to play with the some of the data and go beyond the original questions to understand some of the trends.

Patterns and Trends:

As far as the trends, we can already see that there is a gap across the databases so there is an improvement opportunity in data collection. Further with the potential changes, there is an opportunity for the second half of the year to implement changes as we already have historical data to compare to.

As already indicated there are a lot of quality issues in the data and that is apparent here. However, looking at the content itself:

1. On appointment status, which we will dwell on in this project: - 91% of the regional appointments show that patients have attended their appointments. there is 5% of Did not Attend (DNA) and 4% Unknown that we have an opportunity to look into further. This is where the unnecessary costs are.
2. On national categories and context types: - From a data sanitation and streamlining perspective, it would be beneficial if the national categories are standardised so that time spent and future prioritisation can be reviewed more accurately. - Inconsistent mapping is at 5th position while Unmapped items is currently 7th (in volume) out of 18 categories. This is a sizeable 13% of transactions that we cannot trace at the moment.

Assignment activity 3: Analyse the data

Summarise and explain in fewer than 200 words how you arrived at the answers to the four questions and what, if any, further questions arose in the process.

Using the datetime function was tricky for me at first and I kept getting errors, but with research and support from our assistant facilitator, I was able to achieve the desired results. After creating numerous trials with single parameters/actions per code/cell I started getting the hang of it and improved my codes to include more parameters in a single code. It took a lot of trial and error but I got there in the end.

I had a bit on an easier time with the groupby requirements in question 3 as there were a lot of hands-on during the module. I did need to read-up on linking datetime to this code. I had to research on the Series.dt.month and Series.dt.year to understand this more.

Analysis approach:

- I used datetime to ensure that the appropriate datatype was available
- Using min and max to review dates
- Loc was used as well when extracting the particular location and coupled with a filter of dates between.
- Groupby was utilised to ensure that specific categories are extracted and the total sum computed to understand trends.

Patterns and Trends:

The trends here are very apparent and support the initial findings. General Practice is the top service setting.

Dates in of appointments are also clear but we have an opportunity to review in the next calendar/cycle to assess the patterns.

Assignment activity 4: Visualise and identify initial trends

Summarise and explain in fewer than 200 words any trends you have identified for both aggregated monthly and seasonal visualisations.

I had a lot of fun doing this assignment despite all the pressures I have had on the person side and with my family life I really had a lot of fun with this. Had I had more leisurely time I would have greatly improved the charts to the nitty gritty improvements.

Overall this is what I really thought was something that made everything cohesive and made all the coding and analysis click,. both for the project/assignment and for myself as a student.

Analysis approach:

- I used datetime to ensure that the appropriate datatype was available
- Using min and max to review dates
- Loc was used as well when extracting the particular location and coupled with a filter of dates between.
- Groupby was utilised to ensure that specific categories are extracted and the total sum computed to understand trends.

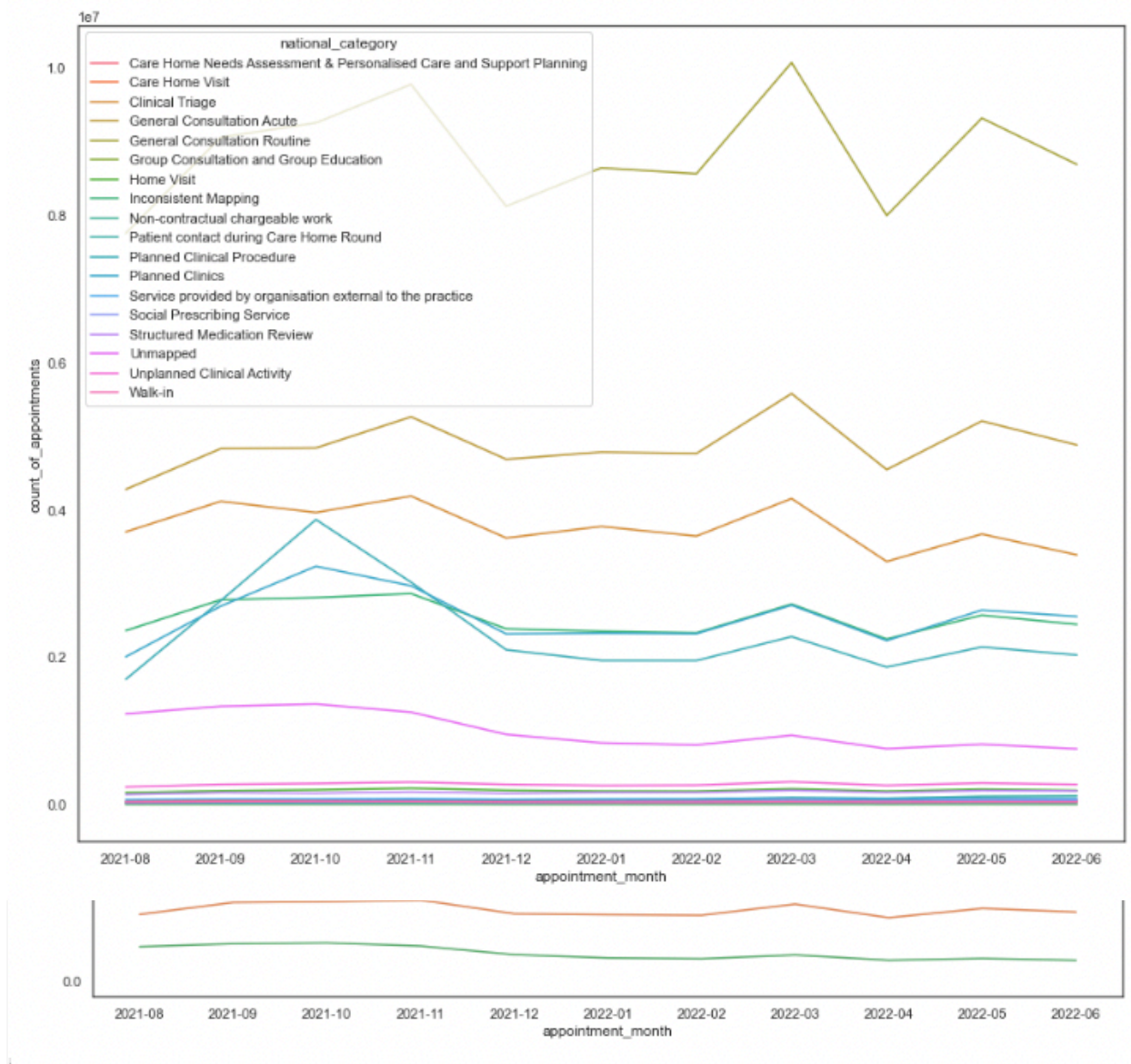
Patterns and Trends:

The trends here are confirmed through visuals that General Practice is the top service setting. The data is influenced my mapping issues and incorrect data collation

Assignment activity 5: Analyse the Twitter data

Summarise in fewer than 200 words any insights you've discovered and anything you would like to explore further. For example, what steps did you follow to improve the visualisations to optimise feedback to your team and the NHS.

There is truly a lot of information we can gather by listening to the voice of the customer- in this case the voice of the patients, their families and those providing care. Databases like this



provide us the signposts on where to look further and streamlines the process of research and review.

Visualising copious amount of data definitely helps in communicating messages better and driving down critical talking points. As a student, this would be one of the things that I would really want to greatly improve on. Not only is it something I enjoy but it is critical in the communication aspect to the stakeholders. The graphs are definitely key to the visualisations but the readability is equally important. Moving forward I would make sure to have all codes handy so that these could be improved with ease.

