Assignment 3 Written Submission Darshan Panesar

Dataset: **COVID-19 cases in hospital and ICU, by Ontario Health (OH) region**

**Spanning 2020-2024**

Link: <https://data.ontario.ca/dataset/covid-19-cases-in-hospital-and-icu-by-ontario-health-region>

Visualization 1:

What software did you use to create your data visualization?

I used MS Excel to generate a multi-line, line plot.

Who is your intended audience?

My intended audience is again healthcare professionals, Ontario Health, hospitals, Toronto Public Health, and researchers.

What information or message are you trying to convey with your visualization?

I am trying to convey the trends across 2024 for patients in the ICU (with subgroups) who have tested positive for COVID, specifically for the Toronto Health Area.

What aspects of design did you consider when making your visualization? How did you apply them? With what elements of your plots?

I considered, how the data would be best presented over time for this reason I chose a line plot. What data to include: the data spanned a large range from 2020-2024. I felt the most recent would be the most interesting or relevant. Given the severity and seriousness of Covid in relation to the ICU I chose to present that data. I wanted to make it clear, easy on the eyes, with a legend so the data was easily understood. Though the finer data points may be a bit hard to isolate the trends are clear and there are several dates present for each month to show a reliable trend. I chose to have a y axis range that best captured the full range of data without biasing the viewpoint. Lastly it was recommended I add more gridlines to make data points easier to discern.

How did you ensure that your data visualizations are reproducible? If the tool you used to make your data visualization is not reproducible, how will this impact your data visualization?

The visualization is very clear, the data is also clearly available in the background of the excel file as the graph is embedded. I also added the data link within the chart so that users may acquire the original data. In addition, the steps to filter the data for Toronto and 2024 are also straightforward.

How did you ensure that your data visualization is accessible?

The colours for each line are bold and I tried to make them different enough to easily follow them. Overall, the trends are clear.

Who are the individuals and communities who might be impacted by your visualization?

Patients are the first that come to mind. Maybe the visualization can be used to identify specific months, seasons, or group-based trends which are key for both ICU patient and covid patient care. Overall, my target audience and policy makers may benefit from this visualization to not only understand the trends of covid and ICU but implement strategies that improve patient care.

How did you choose which features of your chosen dataset to include or exclude from your visualization?

I chose only the most recent year 2024 as I felt this would be the most relevant close data for this visualization. I chose the most populus area and one that we fall under for U of T as I felt this was a reasonable decision to make the visualization more relevant.

What ‘underwater labour’ contributed to your final data visualization product?

Not much.

Visualization 2:

What software did you use to create your data visualization?

I used Python to make an interactive bar chart with error bars using plotly.

Who is your intended audience?

My intended audience is again healthcare professionals, Ontario Health, hospitals, Toronto Public Health, and researchers and even the general public whose interested in trends for covid hospitalizations and covid within the ICU within Toronto over the last year.

What information or message are you trying to convey with your visualization?

I am trying to convey the monthly average of covid Hospitalizations and ICU patients in Toronto over 2024.

What aspects of design did you consider when making your visualization? How did you apply them? With what elements of your plots?

I again considered, how the data would be best presented over time but also summarized per month which led me to a bar chart. I wanted the graph to convey more information therefore the plotly style was perfect as it is interactive. The most recent year 2024 and monthly averages would be the most interesting or relevant. Proving both general hospitalizations and more serious ICU patients with covid would provide both a general picture and more critical information. In terms of my visualization, I wanted to make it clear, easy on the eyes, with a legend so the data was easily understood. I chose averaging the data as it would give a good summary without compromising information. Adding both plots side by side would give an extra point of comparison. Lastly, I added error bars to add more depth the data that was visualized and give the viewer more information as I was providing averages.

How did you ensure that your data visualizations are reproducible? If the tool you used to make your data visualization is not reproducible, how will this impact your data visualization?

The visualization is very clear and interactive. The python code is heavily annotated with links to the dataset and resources.

How did you ensure that your data visualization is accessible?

The colours for each line are bold with the plot featuring palatable information (not too much) and its fully interactive.

Who are the individuals and communities who might be impacted by your visualization?

Overall, my target audience and policy makers may benefit from this visualization to not only understand the trends of covid and ICU but implement strategies that improve patient care. The general public can also benefit but easily interpreting and understanding covid trends within hospital patients.

How did you choose which features of your chosen dataset to include or exclude from your visualization?

I chose only the most recent year 2024 as I felt this would be the most relevant close data for this visualization. I chose the most populus area and one that we fall under for U of T as I felt this was a reasonable decision to make the visualization more relevant and less overwhelming. I averaged the data per month but provided standard error of the mean bars to accurately represent the data distribution.

What ‘underwater labour’ contributed to your final data visualization product?

A ton of reading and research went into making this visualization reading plotly documentation, looking at stack overflow for coding information, and I used some generative ai with errors in my code as well.