Massachusetts General Hospital Project – Some Observations

After having worked on this set of synthetic patient data related to Massachusetts General Hospital with different technologies, I've noticed that the output values may end up varying between the different technologies despite the attempt to answer the same four data-related questions.

With respect to the patients admitted or readmitted over time, Power BI Desktop, Microsoft SQL Server, and Tableau Public all omit the months when no patients were admitted or readmitted according to the dataset. However, Looker includes them, so the charts or tables that present this information will be a bit different.

Moreover, when cycling through each year of the slicer/filter feature set up for each technology, the revised outputs ended up being sometimes different among the technologies. This could be a matter of how "internal math" is done by each technology with each one possibly rounding values at different points of the calculations. This could also be a matter of how the technologies each handle datetime data types, or how each technology handles the joining of more than one file/table.

When performing analyses with Python and the pandas library, I was able to create the same output as what had been created using SQL. However, for the final analysis of the average cost per visit, I decided to round the value to two decimal places in order to try to recreate the output generated with Looker, and fortunately the numbers generated by both Python and Looker did end up matching.

From what I can see, the information output by both SQL and Tableau Public appear to consistently match each other, while the information output by both Power BI Desktop and Looker will have their own deviations from the other two technologies. Moreover, I could recreate the SQL output using Python and the pandas library. Therefore, I'd be inclined to utilize SQL, Tableau, or Python to handle calculations or visualizations while resorting to Power Bi or Looker only as circumstances require (e.g., an employer requires me to use Power BI or Looker, or a job a posting mentions those technologies thus compelling me to practice and showcase data analysis or visualizations with those technologies).