**Medical Laboratory Database Analysis**

* Tools used (software, Python packages)
  + Jupyter Notebook software to work on this project, along with the following libraries:
  + sqlite3: to interact with the databases and query data.
  + pandas: for data manipulation, analysis, and handling structured data.
  + seaborn: to create statistical data visualizations to explore patterns and relationships in the data.
  + matplotlib: to create various types of plots and charts to visualize data.
  + folium: to create interactive maps.
  + <https://www.mockaroo.com/> to generate fake data for the database.
  + https://elevenlabs.io/ to convert text to speech.
  + PowerPoint for presentation
* Questions asked:
  + Who is the EMPOLYEE (PHLEBOTOMIST) OF THE YEAR?
  + How was his performance this year?
  + Which state had the highest number of uninsured patients?
  + Which state had the highest number of COVID-19 cases in 2023?
  + What is the infection rate between different genders?
  + What is the infection rate among different age groups?
* Insights that were discovered
  + Fritz Leather was the Phlebotomist of Year 2023.
  + His highest performance occurred in March, while the lowest performances were observed in February and November.
  + Texas has the highest number of uninsured patients.
  + California had the highest number of COVID cases followed by Texas and then New York.
  + Men had a higher infection rate of COVID compared to women.
  + Seniors aged over 65 experienced the highest incidence of COVID-19 infections.
* Recommendations:

Use the power of data and insights in driving tangible improvements within the organization and enhancing the healthcare system for patients.

* Future work:

I would like to continue working on this data to analyze the peak infection period in each state as I find that this will be helpful for healthcare providers to prepare for that period and provide vaccinations prior to that, also public can use it to take the proper precautions like wear masks and social distancing during that period.