

A doctor in a white coat and stethoscope is using a tablet. Overlaid on the image is a collage of medical and scientific icons, including a bar chart, a pie chart, a DNA helix, a brain, a globe, and various medical symbols like a heart, a cross, and a microscope.

In this project, you will play the role of an analytics consultant for a hospital. You have been asked to develop a high-level KPI report for the management team, based on a subset of patient records. The aim of the report is to give stakeholders visibility into the hospital's recent performance, and to answer the following questions:

- The dashboard needs to adapt to new data over time, but the CEO has asked you to summarize all the information you can draw from the sample provided.

About the Dataset:

Synthetic data on approximately 1,000 patients between 2011 and 2022, including information on patient demographics, insurance coverage, medical encounters and procedures.

You'll find it attached to the file.

Instructions:

- Understand your customer's needs and requirements.
- Define a workflow for your data analysis project.
- Understand the data and its components. A data dictionary is provided.
- Perform data cleansing and pre-processing (you can use python scripts and be sure to take data updates into account).
- Data modeling and visualization (preferably Power BI)
- As an additional solution, create an interactive data application that helps doctors obtain all the information on a patient's health using streamlit.

Output:

- Processed data file
- Data rows where anomalies were found (add a column describing the anomalies)
- Scripts used for data pre-processing and application
- Report file