"Healthcare provider fraud detection analysis" from Kaggle ( <https://www.kaggle.com/rohitrox/healthcare-provider-fraud-detection-analysis> ).

We are going to take the multiple CSVs from the dataset, process the data, look for insights (noted below), and then package and add to a database. After, we will submit to github and deploy on a website (Github Pages).

We are looking at 2 different insights:

1. What are the most common Procedures and Diagnosis codes

2. Which states/localities have the highest number of frauds