

## **TEAM :**

**Nkem Bamgbose, Zanielle Cabrera, Arundhishaan Kanagaraja**

### **1. Significant Missing Values within multiple Features**

We dropped 15 columns due to the amount of missing values. We made sure that these missing values were not due to survey structure, as dialysis and nutrition questions are optional. However 5 of the columns dropped are from the optional questions related to violence and protest. Due to the inconsistencies of this section, we dropped all the questions of this nature and concluded that the entire hospital infrastructure throughout the country is susceptible to violence and protest due to the crisis.

### **2. Removed Duplicates**

There were 19 rows dropped due to being duplicated, revealed through the timestamps.

### **3. Remove Inconsistencies between Hospital Code and Federal Entity**

There was one entry with the code 'mir001' that has been formatted as the others. There were a few rows in which their hospital code did not correspond to their federal entity; these have been adjusted by changing the region according to the hospital code.

### **4. Change string variables into numbers that represent categories**

In order to analyze and demonstrate in a dashboard, these strings need to be changed to integers.