# Independent Data Science Analysis

Tapia-Tellez Jose Medardo

August 28, 2021

## 1 Region and Domain

**Region and domain category that the data sets are about.**

- **Region:** Mexico City, Mexico
- **Domain:** SARS-COV2

## 2 Research Question

**Question about the domain category and region that we identified as being interesting.**

- How have the SARS-COV2 positive cases and hospitalization cases due to SARS-COV2 change over the course of the pandemics in Mexico City?

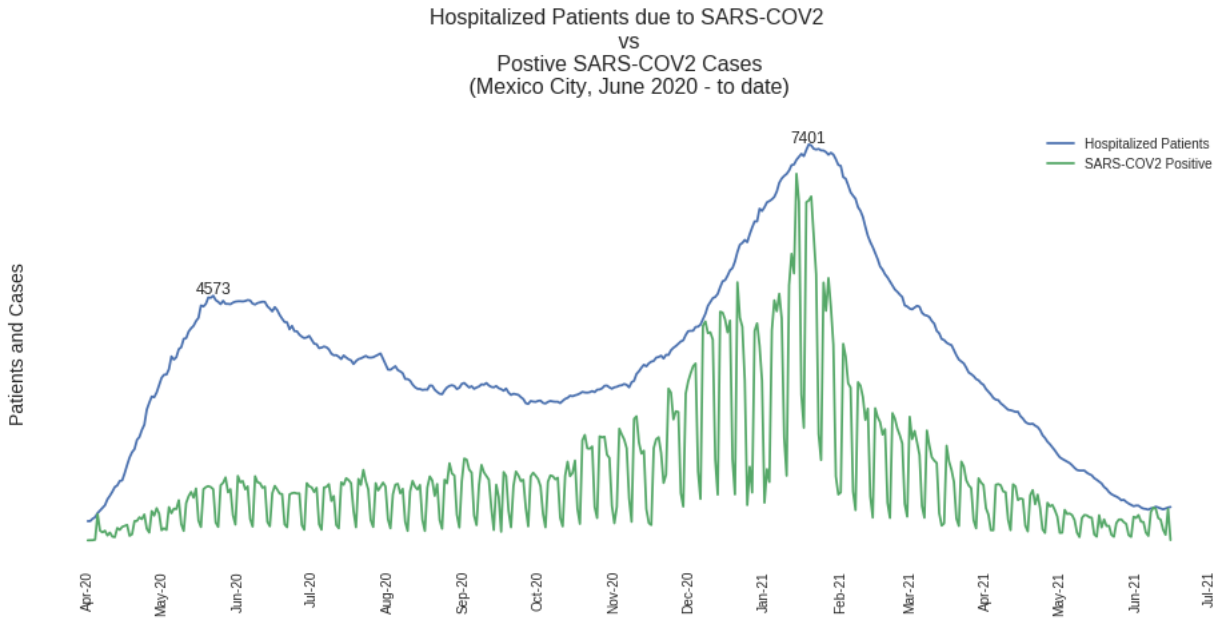
## 3 Links

**Links to publicly access the datasets.**

- https://datos.cdmx.gob.mx/dataset/total-de-pruebas-total-de-positivos-y-tasa-de-positividad/resource/ae2cd1aed-45a1-8ee4-3d0b0852ae4b
- https://datos.cdmx.gob.mx/dataset/personas-hospitalizadas-en-hospitales-de-zmvm/resource/8b29f1ab-6245-42f1-878b-78e9a4b02374

## 4 Image

**Image which addresses the research question we stated. In addition to addressing the question, this visual follows Cairo's principles of truthfulness, functionality, beauty, and unsightliness.**




# 5 Discussion

**Written justification of how your visualization addresses our stated research question.**

This visualization was concerned with answering the question of how have the SARS-COV2 positive cases and hospitalization cases due to SARS-COV2 compare over the course of the pandemics in Mexico City. We first obtained the data from the portal of open data that belongs to the government of Mexico City. We then obtained two data sets: hospitalized people due to SARS-COV2, and the second data set icluded information regarding total tests, positives and positivity rate. With this sets we then worked with the attributes we were interested (hospitalized and positives), we got rid of outliers (first month of the pandemic where the data didn't help for the visualization), and finally merged both of the datasets. We finally graph the merged datasets and improve the graph by taking into account Cairo's principles.

Regarding on how our visualization answers the stated question, fist of all, we can see a clear comparison between SARS-COV2 positive cases and hospitalized patients due to SARS-COV2. The timeline in the bottom, gives us the continuity necessary in order to understand the evolution of the two previous mentioned characteristics through all of the pandemic. It is important to point out that the graph also helps to visualize the first and second wave of the pandemic in Mexico city, this is shown through two maximums indicated with a number on top of them. Some things that come to our attention are that the SARS-COV2 test has an electrocardiogram behavior, and that there's always more hospitalized people than positive SARS-COV2 cases, the latter certainly questioning the amount of tests done by the government or the veracity of this data since we should see a greater amount of people being infected and not being in the hospital. It was certainly an entertaining exercise which ended up in being a good learning experience.