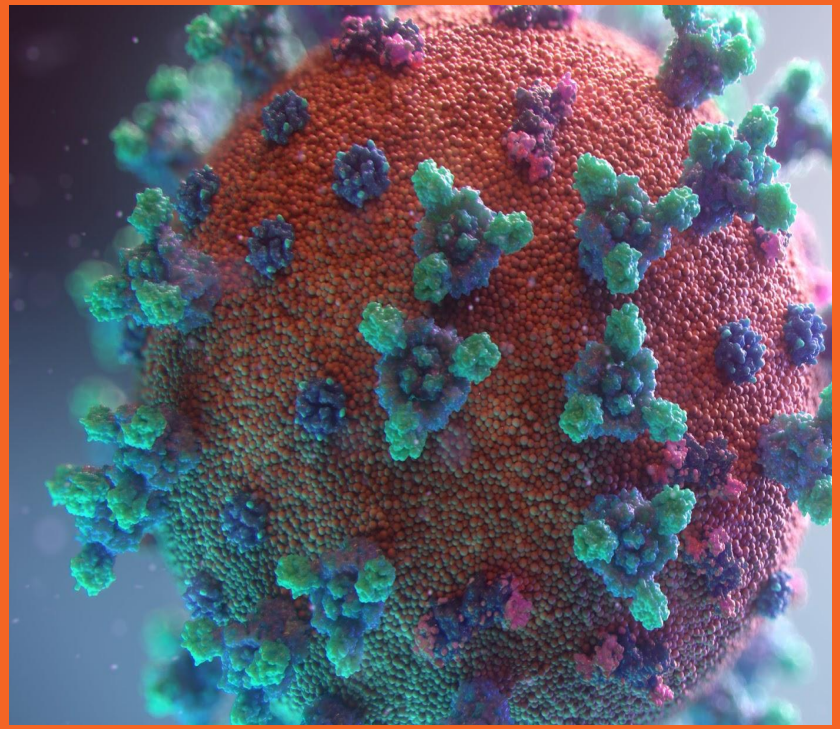


Photo by Fusion Medical Animation on Unsplash

Covid-19 Data Analysis

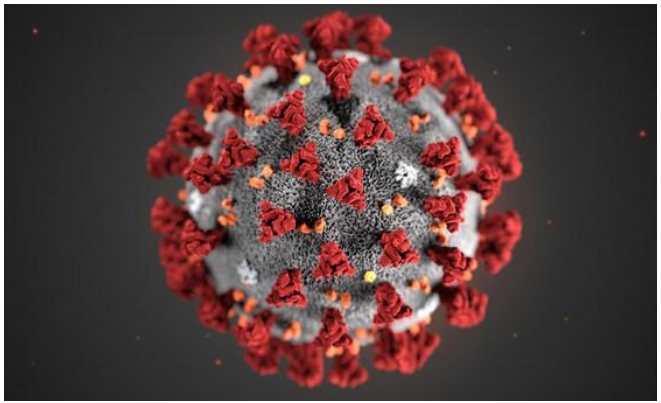
Cassandra Cruz Martínez

Dataset description

The dataset was provided by the Mexican government.

This dataset contains an enormous number of anonymized patient-related information including pre-conditions.

The raw dataset consists of 21 unique features and 1,048,576 unique patients.

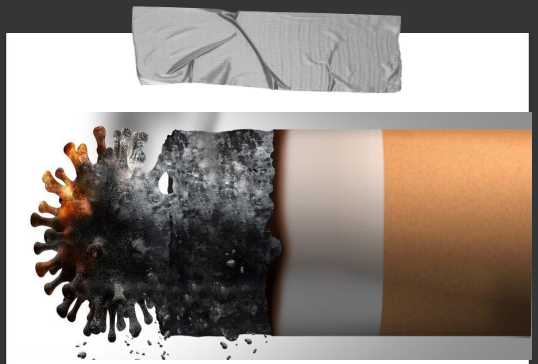


Source: https://www.un.org/es/coronavirus

- Age: of the patient.
- Classification: covid test findings. Values 1-3 mean that the patient was diagnosed with covid in different degrees.
- Patient type: type of care the patient received in the unit.
- Pneumonia: whether the patient already have air sacs inflammation or not.
- Pregnancy: whether the patient is pregnant or not.
- Diabetes: whether the patient has diabetes or not.
- Copd: Indicates whether the patient has Chronic obstructive
- pulmonary disease or not.
- Asthma: whether the patient has asthma or not.
- Inmsupr: whether the patient is immunosuppressed or not.
- Hypertension: whether the patient has hypertension or not.
- Cardiovascular: whether the patient has heart or blood
- vessels related disease.

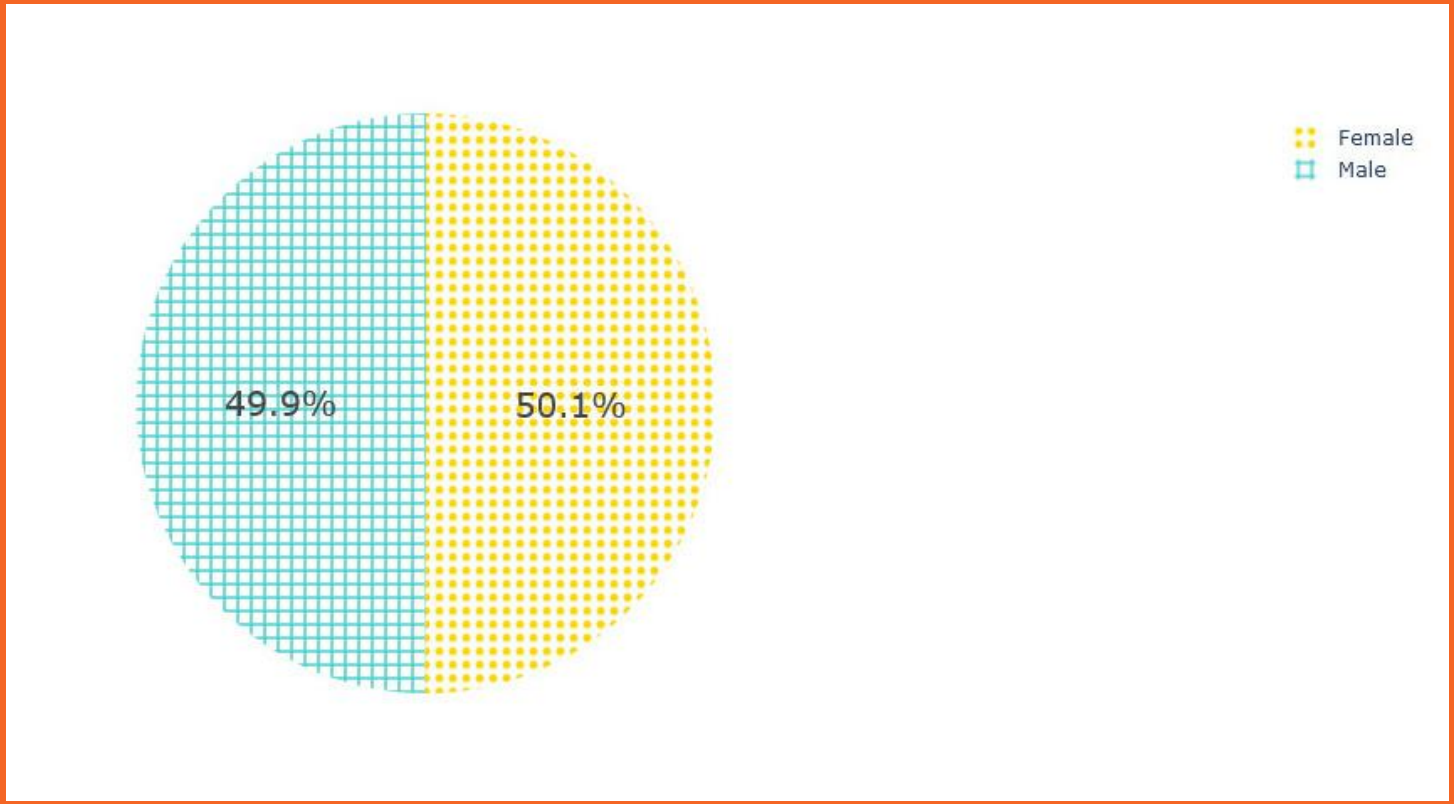
Source: Getty Images / Pollyana Ventura

- Renal chronic: whether the patient has chronic renal disease or not.
- Other disease: whether the patient has other disease or not.
- Obesity: whether the patient is obese or not.
- Tobacco: whether the patient is a tobacco user.
- Usmr: Indicates whether the patient treated medical units of the first, second or third level.
- Medical unit: type of institution of the National Health System that provided the care.
- Intubed: whether the patient was connected to the ventilator.
- Icu: Indicates whether the patient had been admitted to an Intensive Care Unit.
- Date died: If the patient died indicate the date of death

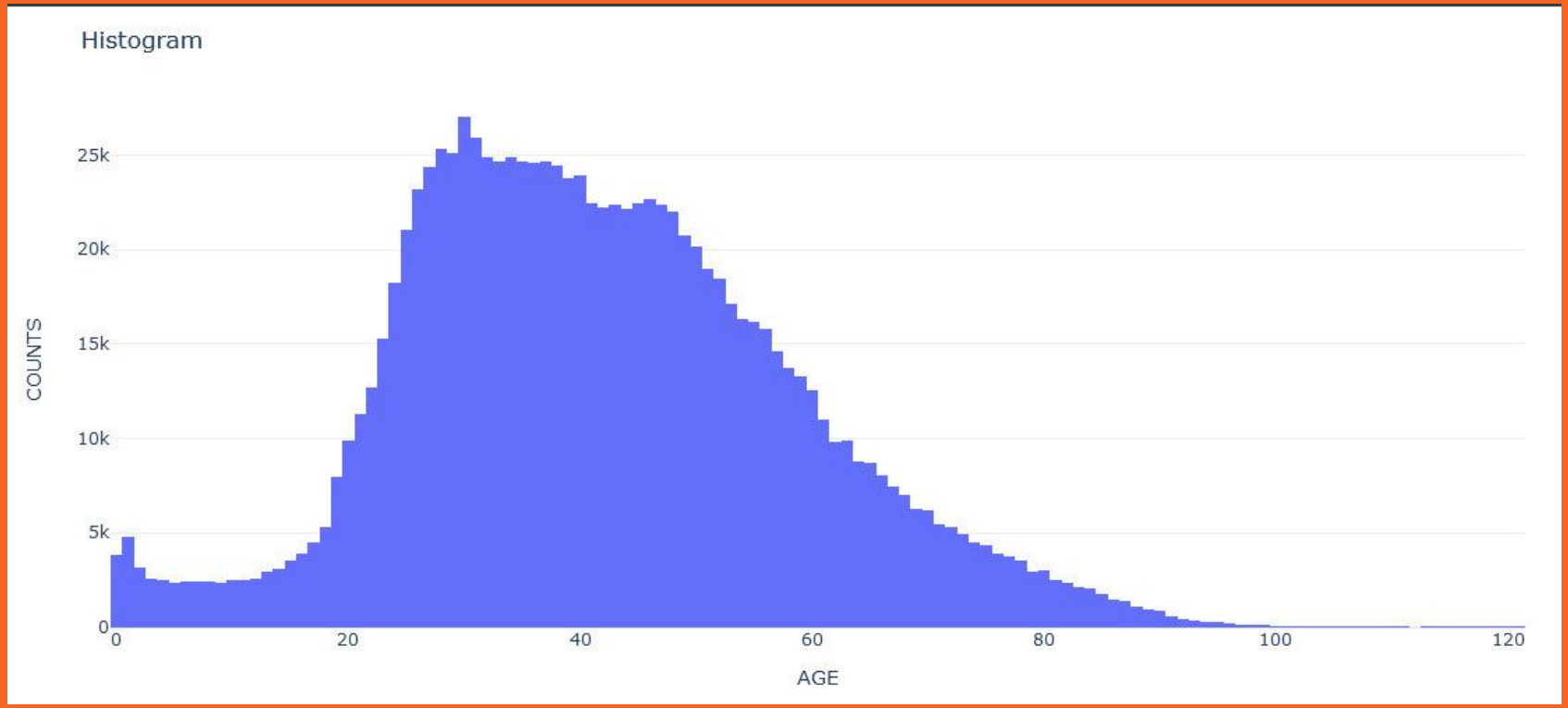


Source: https://www.paho.org/es/temas/prevencion-control-tabaco/ciclo-conferencias-virtuales-sobre-covid-19-tabaco-otros-factores

Data set population



Age histogram

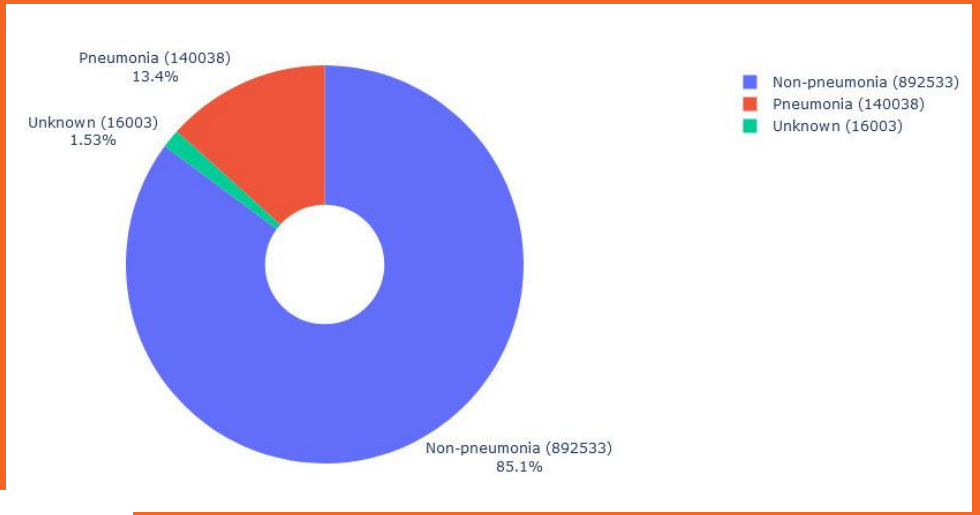


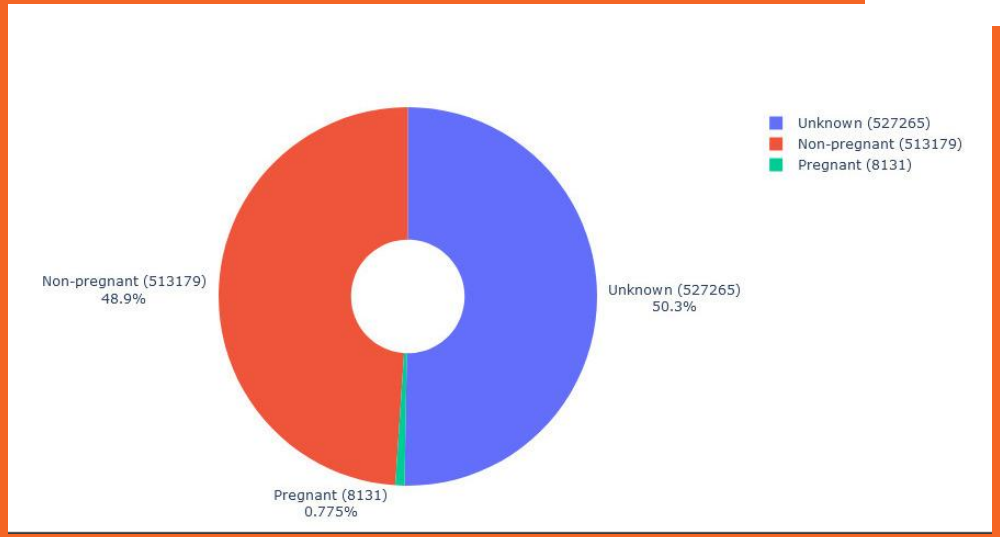
2021-2022
Covid - 19

Most common diseases of patients (Mexican population)

~1.5M patients

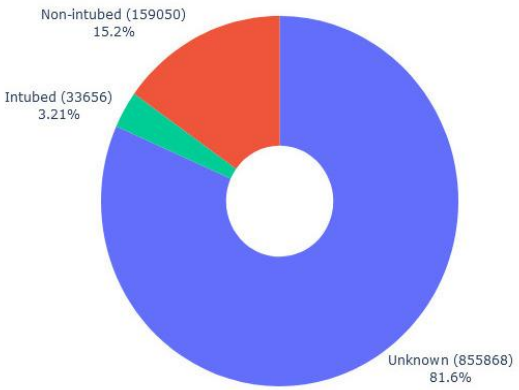
Pneumonia

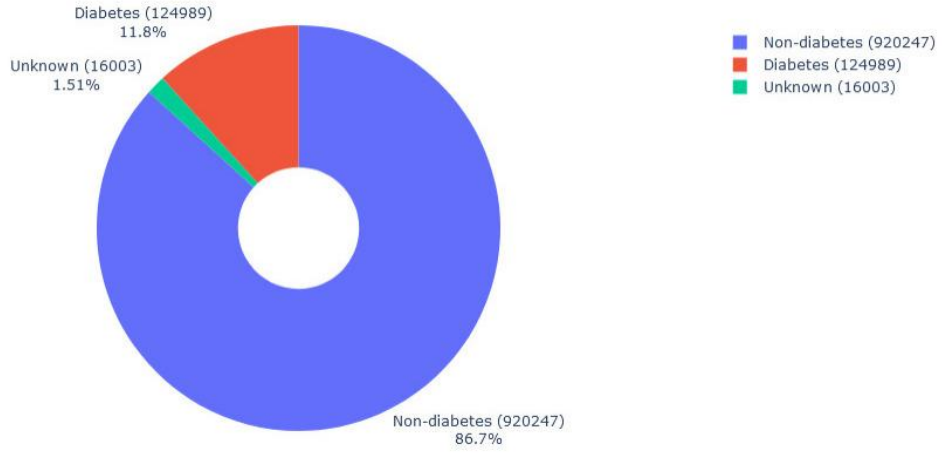




Pregnant women

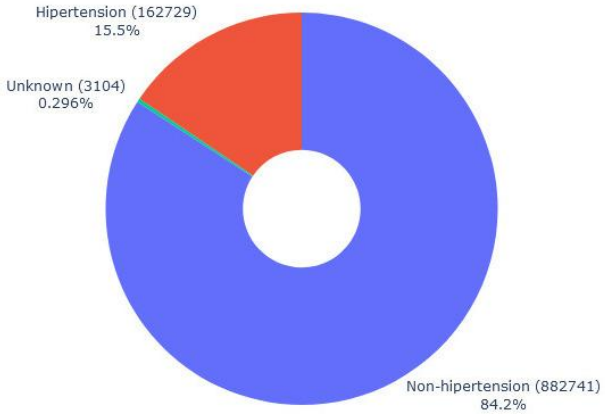
Intubed

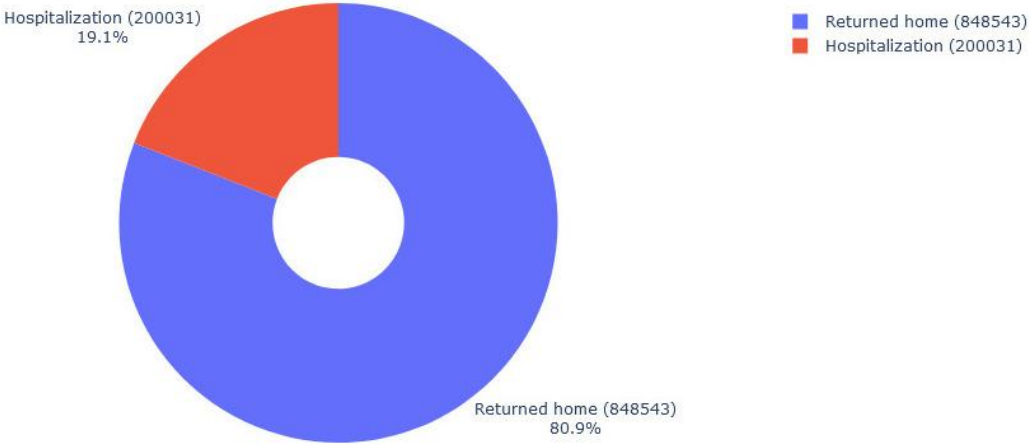




Diabetes

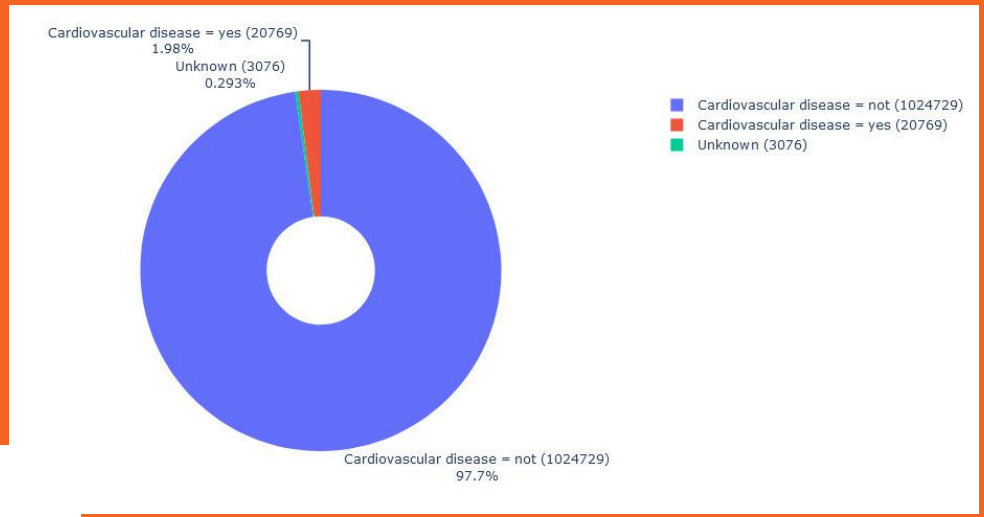
Hipertension

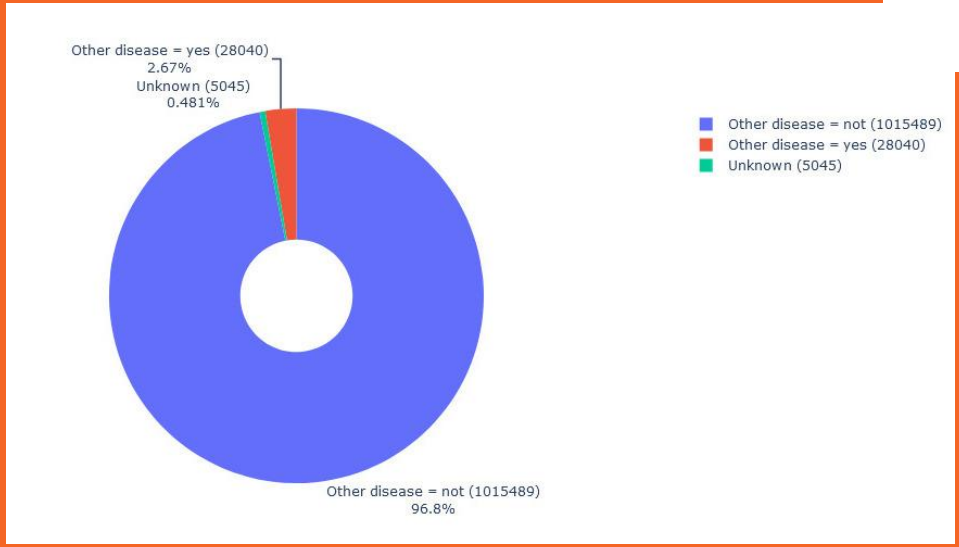




Patient type

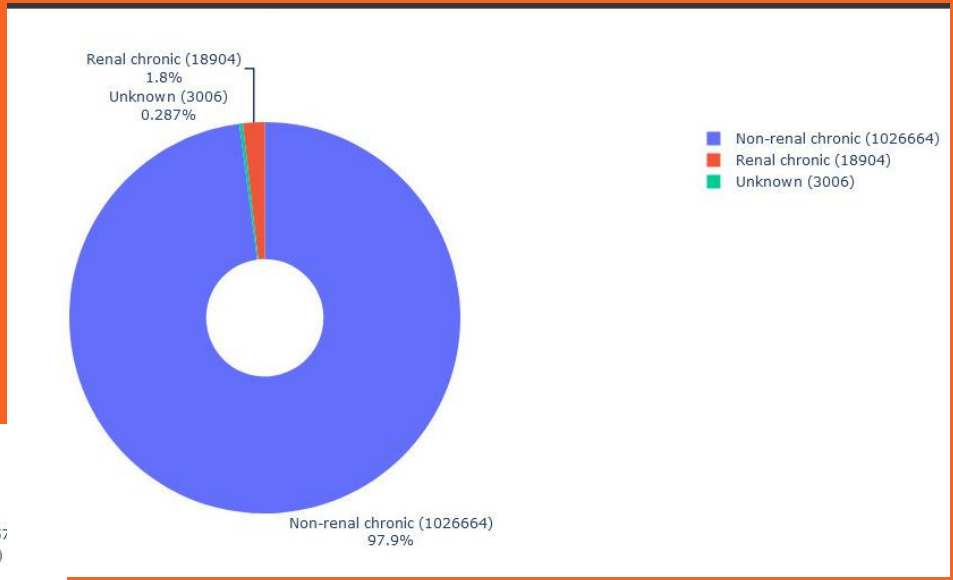
The patient has heart or blood vessels related disease

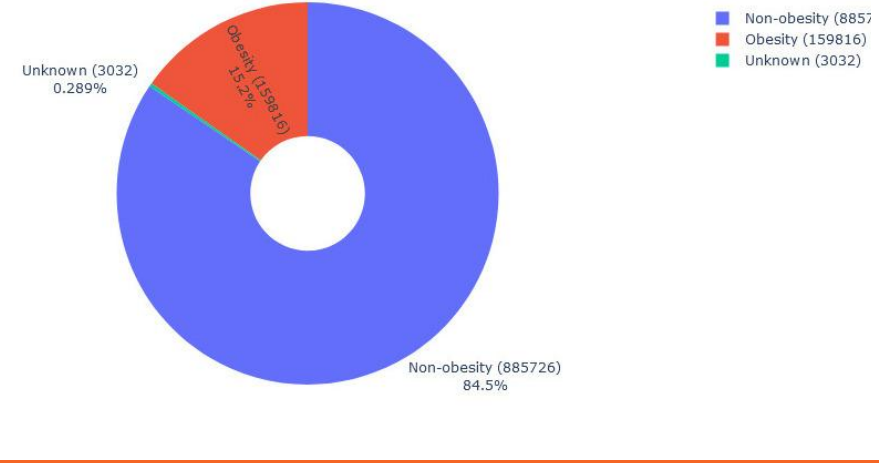




Other disease

Renal chronic

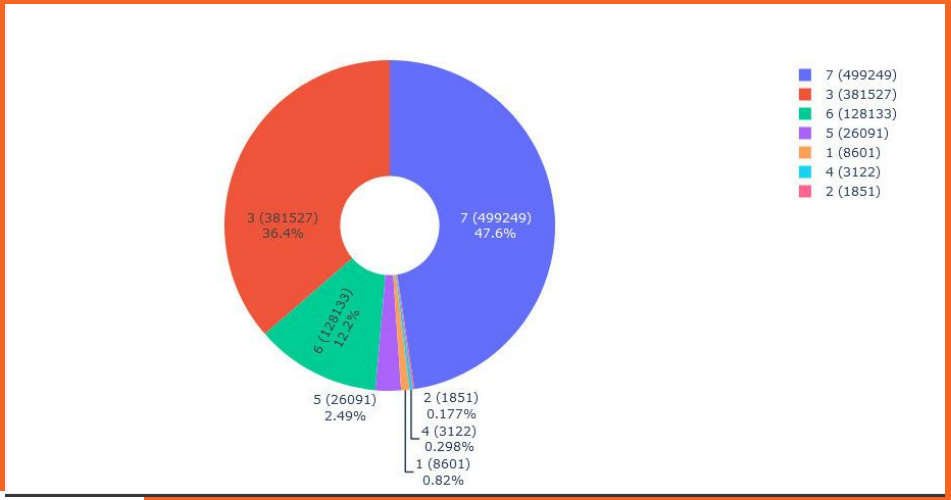


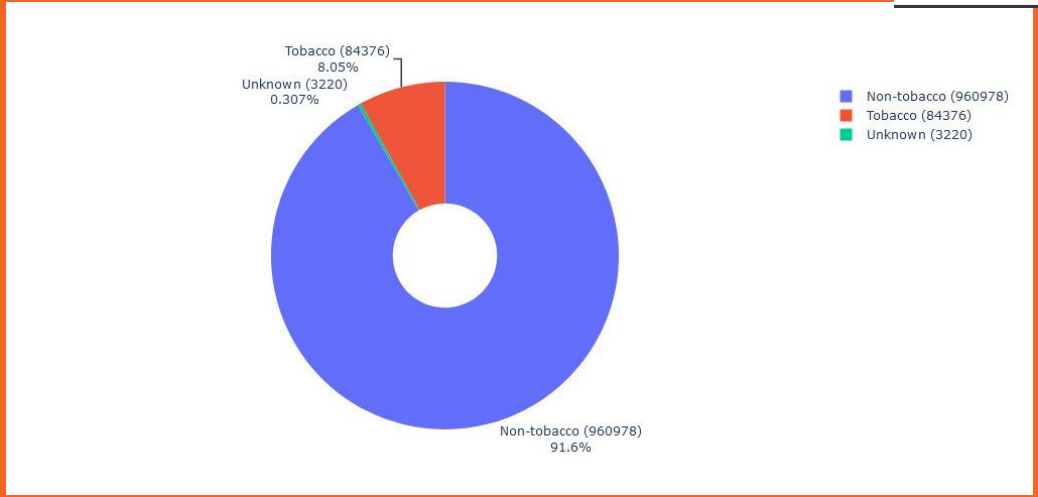


Obesity

Classification

Covid test findings: 1-3 mean the patient was diagnosed with covid in different degrees.
4 or higher means test is inconclusive.

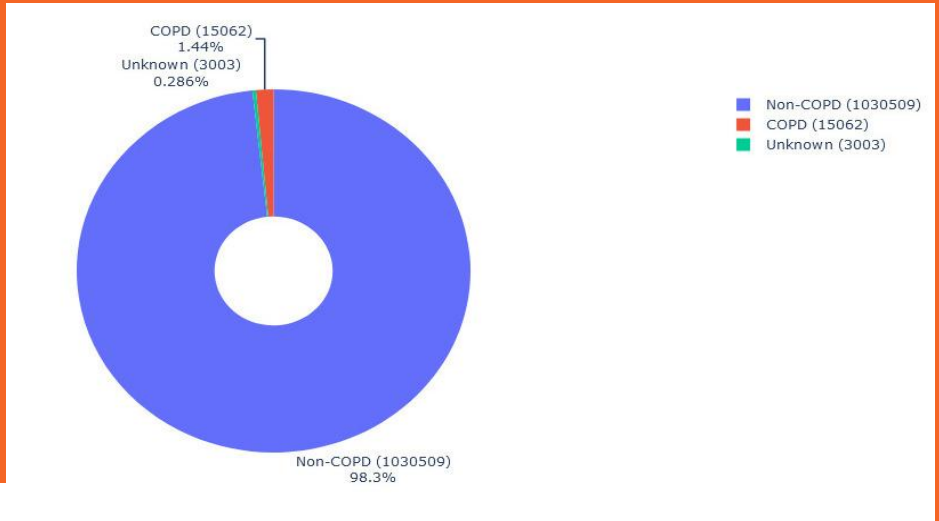


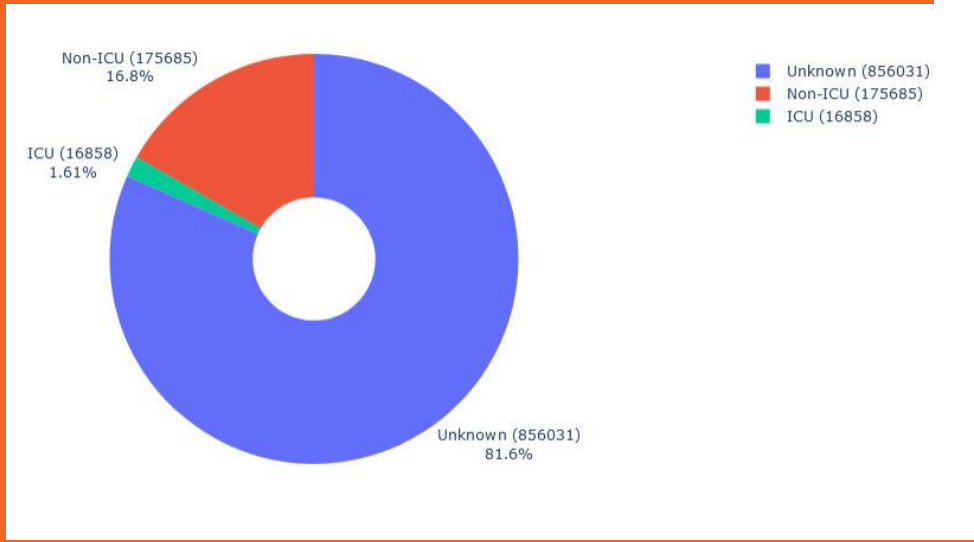


Tobacco

copd

Pulmonary disease or not

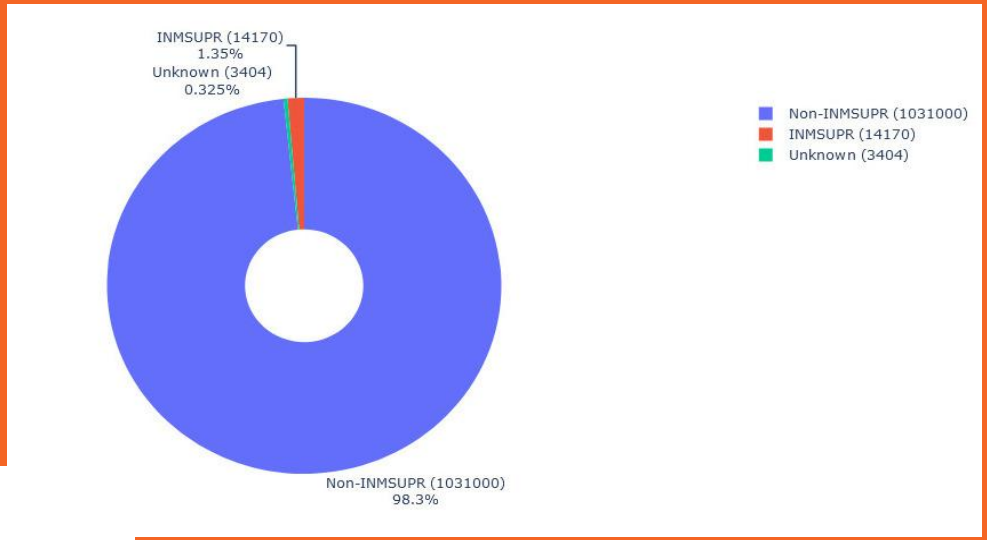


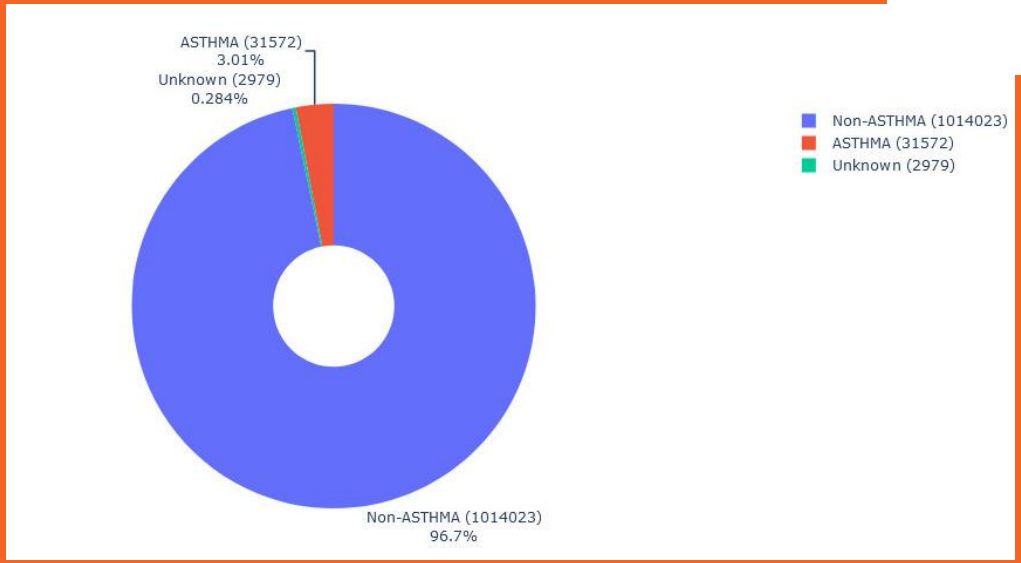


ICU

Patient had been admitted to an Intensive Care Unit

inmsup

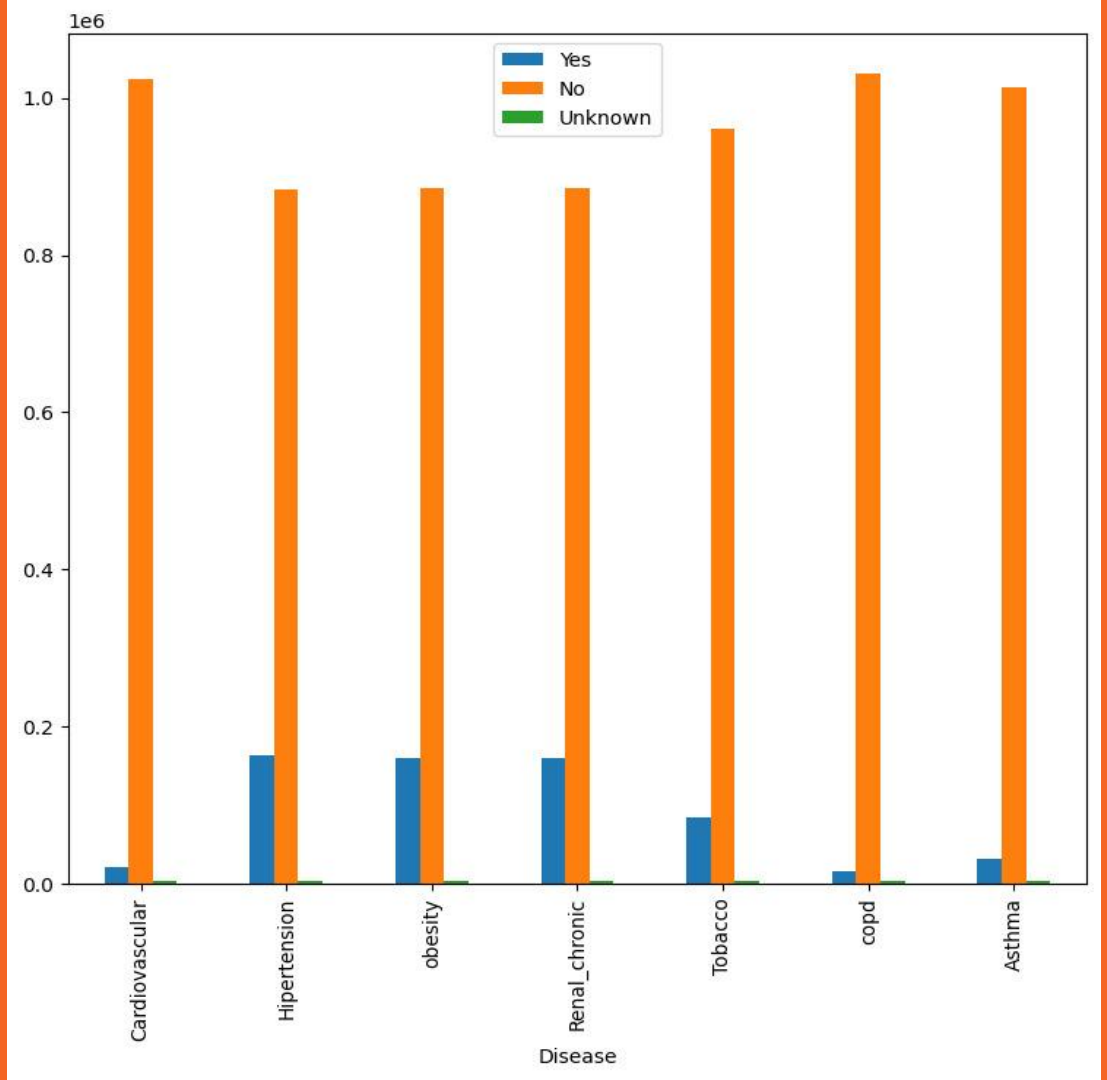


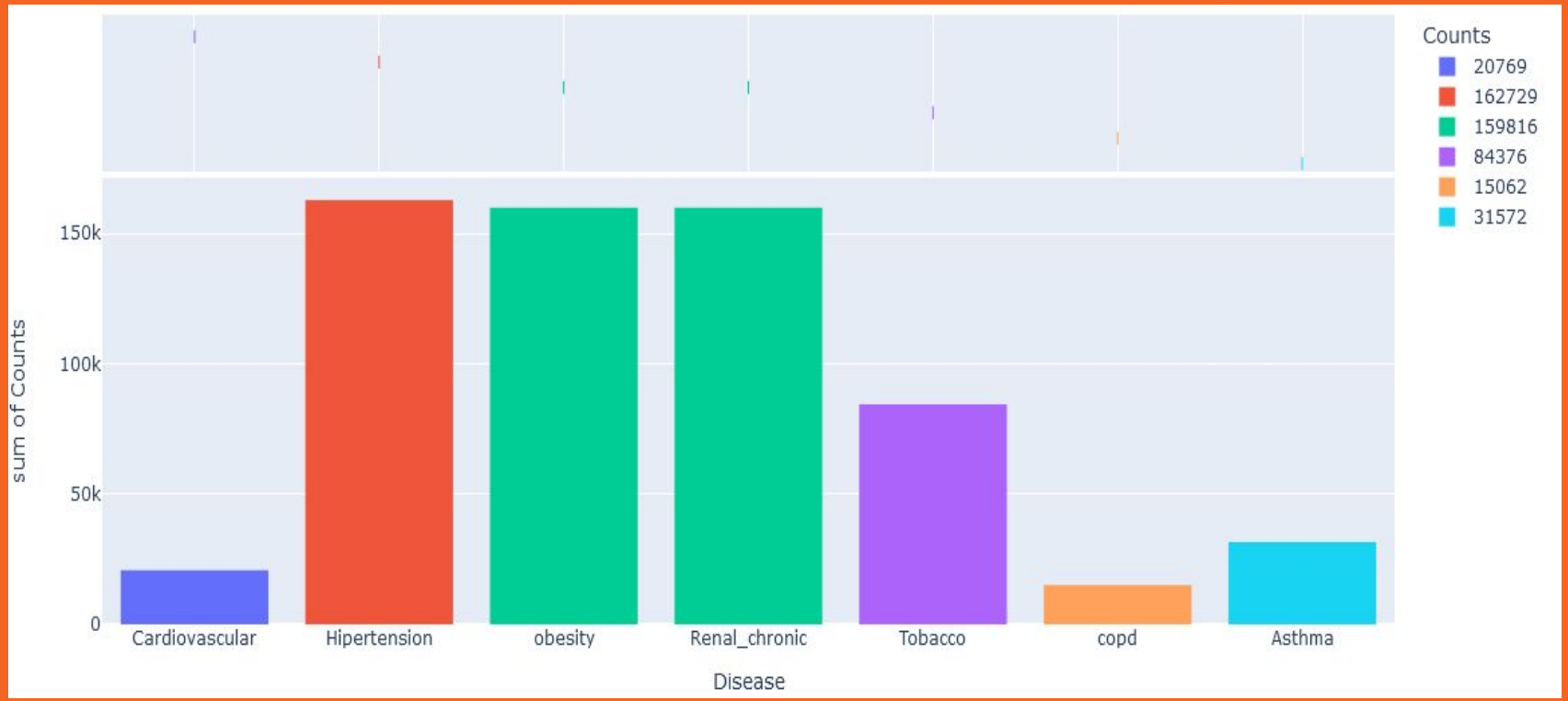


Asthma

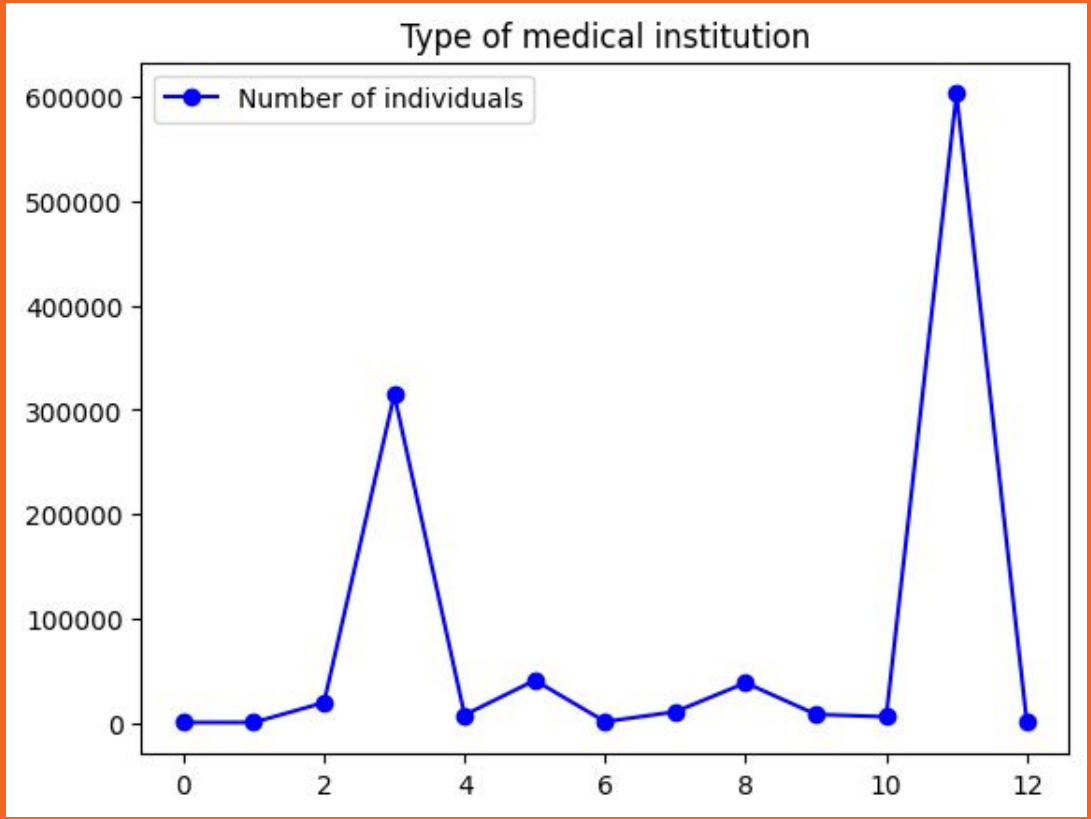
Number of people with a medical condition studied in the dataset

-Condition
-Non-condition
-Unknown





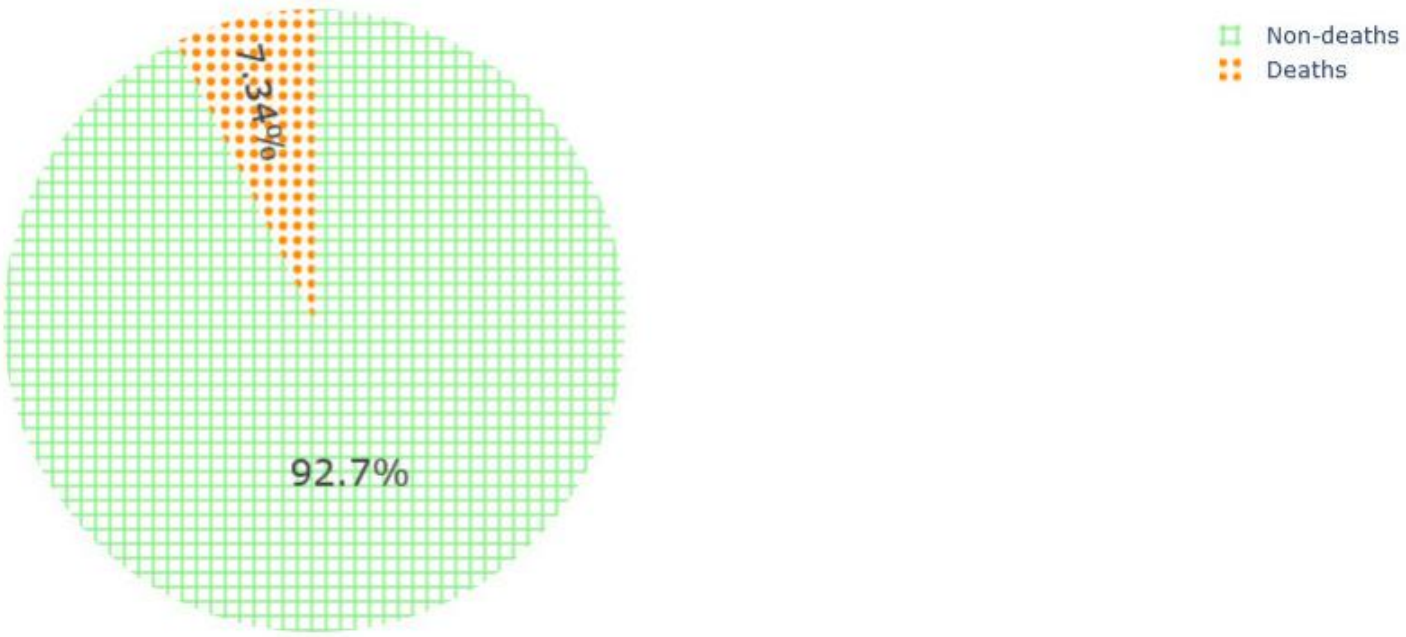
Number of people with a medical condition



Type of institution of the National Health System

Type of hospital: inconclusive.
It can be: Health centers, General Hospital, etc.

Percentage of deaths



In a nutshell

- ➔ Pregnant women with diabetes: 64224
- ➔ Patients who use tobacco and have pneumonia: 12393
- ➔ Patient has kidney damage and obesity: 3734
- ➔ Patients with asthma who were intubated: 697
- ➔ patients with immunosuppressed system sent home: 7149/hospitalized: 7021
- ➔ Number of men and women with hypertension: 81778/80951
- ➔ Cardiovascular problems and age: 50-60 years old, on average 400 people in this population.

Interesting data that can be further explored:

1

Main medical conditions related to deaths.

2

How many deaths there are compared to the number of cases?

Covid-19

2

Exploration of the degree of classification of patients, with respect to their age.

Other variables outside this dataset

- Patient physical activity
- Covid symptoms
- Vaccines (dose/type)

- Medical equipment in hospitals
- Case study by region

Thanks