

Calculating the average age of deceased patients

Introduction

In this notebook, I will be calculating the average age of deceased patients from COVID-19. For this, I am using a database from kaggle (link on the README and on the previous notebook).

Author: Arthur Gatignol (arthur.gatignol@gmail.com)

Version: 0.1

Getting the data

For this, I am opening the CSV file using pandas.

In [4]:

```
import pandas as pd
df = pd.read_csv('patient.csv')
```

Calculating the average age of patients

Before calculating the average age of deceased patients, we need to calculate the average age of all the patients

In [5]:

```
count = 0
total = 0
for i in df['birth_year']:
    if i > 0:
        total = total + i
        count = count + 1
average = total/count
print(average)
```

1965.8102189781023

As seen in the above calculations, the average birth date of COVID-19 patients (for the total of 137 disclosed birth dates) is 1966 (rounded up), which is equivalent to 54 years old.

Finding the youngest and oldest patients infected by COVID-19

In the cell below I am finding out the youngest and oldest patients infected by COVID-19. This has been done by setting an initial maximum value of 999,999 and an initial minimum value of -999,999. Then, I run a 'for loop' that will go through the 'birth_date' column and compare the value in the cell to the initial 2 values. At the end of the loop, I outputted the two values in a proper sentence.

In [6]:

```
minimum = 999999
maximum = -999999
for i in df['birth_year']:
    if i > 0:
        if i > maximum:
            maximum = i
        elif i < minimum:
            minimum = i
print(f"The youngest person infected by COVID-19 was: {int(2020-maximum)} year old")
print(f"The oldest person infected by COVID-19 was: {int(2020-minimum)} years old")
```

The youngest person infected by COVID-19 was: 1 year old
The oldest person infected by COVID-19 was: 98 years old

Calculating the average age of deceased patients

Now that we have the average age of patients infected by COVID-19, we can start calculating the average age of deceased patients.

In [21]:

```
# Here, I am setting two variables that
# will be used to calculate the average
total = 0
count = 0

# Defining 'column' as the birth year
# column so that I can iterate through
# it later on
column = df['birth_year']

# Running a loop that goes through the
# status column
for i, n in enumerate(df['status']):

    # If the status of that patient is
    # deceased, run the following steps
    if n == 'deceased':

        # If the birth year is more than 1
        # run the following steps
        if column[i] > 1:

            # Increment the total birth years
            # with the new birth year
            total = total + column[i]

            # Increment count by 1 to have the
            # total number of deceased patients
            count = count + 1

# Calculating the average
average = total/count

# Printing the average
print(f'The average birth year of deceased patients: {round(average)}')
```

The average birth year of deceased patients: 1940.0

From the above calculations, we can see that the average birth year of deceased patients was 1940. That means that the average age of deceased patients is 80 years old.

Conclusion

In conclusion, we can see from all the above calculations that the average age of COVID-19 patients is **54 years old**, and the average age of **deceased** patients is **80 years old**. This is due to an immune system more prone to diseases.