



# GroupD

# Covid- Project

---

- ARIADNA PUIGVENTÓS
- XELES RODRIGUEZ
- JAVIER GIL-ANTUÑANO

# Group D

---

**ARIADNA PUIGVENTÓS**



- Visualization Expert
- MIT Computer Engineering

**XELES RODRIGUEZ**



M. Angeles Rodriguez V.

- EDA expert
- CALTEC Computer Science

**JAVIER GIL-ANTUÑANO**



- Intern
- The Bridge Data Science



# World Health Organization

## DEFINITION

*Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus.*

## SYMPTOMS:

- *Fever*
- *Dry cough*
- *Tiredness*
- *Loss of taste or smell.*

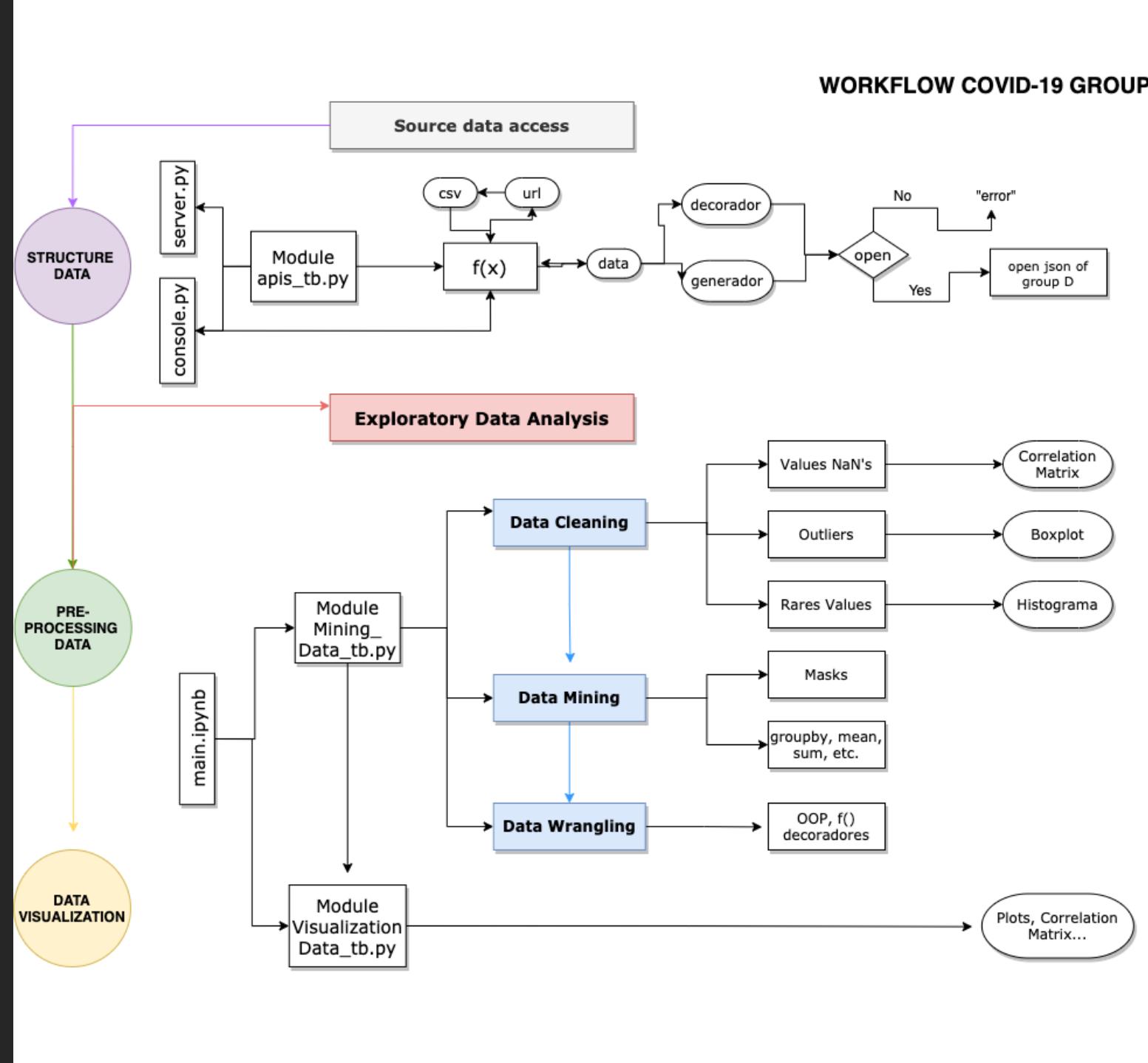
# Index

---

- Workflow
- Pre – Processing
- Visualization
- How to improve

# Workflow

- Start → Suffering/ Hatred
- During → Hair loss/ Stress
- End → Happiness / Relief



# Pre- Processing

---

- What do we have:

```
▶ ▶ M4
from utils.folders_tb import downloader
world = downloader(url='https://covid.ourworldindata.org/data/owid-covid-data.csv')
world
```

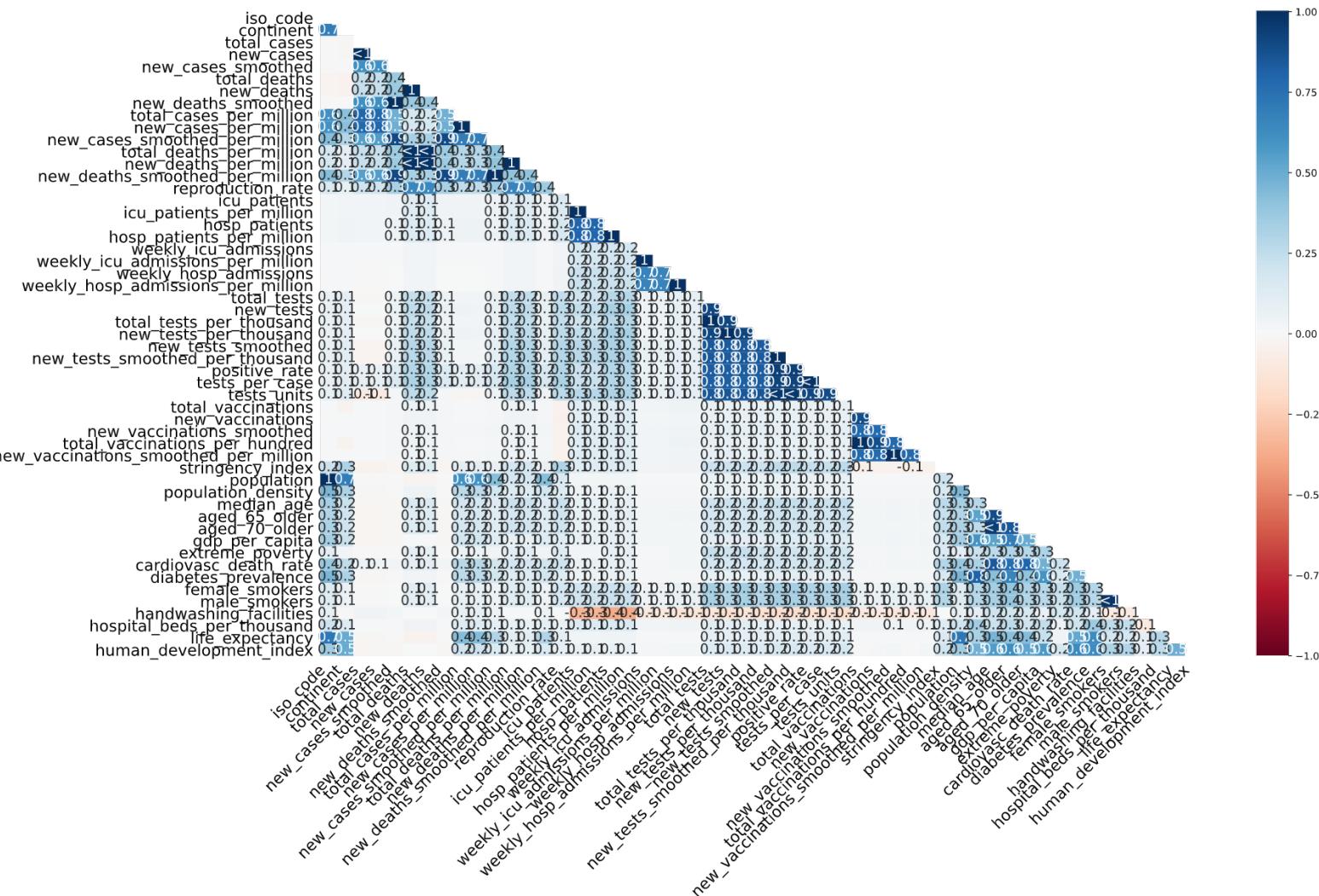
```
def countries(df, code):
    mask = df[df["iso_code"]==code]
    return mask
```

```
▶ ▶ M4
from utils.mining_data_tb import countries
gbr = countries(df=world, code= "GBR")
prt = countries(df=world, code= "PRT")
ven = countries(df=world, code= "VEN")
tur = countries(df=world, code= "TUR")
esp = countries(df=world, code= "ESP")
gbr
```

```
def column_erraser(df, col1=None, col2=None,col3=None,col4=None,col5=None, col6=None, col7=None, col8=None, col9=None, col10=None):
    name = df.drop(df.columns.difference([col1, col2,col3,col4,col5, col6, col7, col8, col9, col10]), 1)
    return name
```

# Pre- Processing

- How bad is it :
- Then we had to get rid of all the NaN values.



# Pre- Processing

---

- Then get the data to do what you want:

```
def get_redundant_pairs(df):
    '''Get diagonal and lower triangular pairs of correlation matrix'''
    pairs_to_drop = set()
    cols = df.columns
    for i in range(0, df.shape[1]):
        for j in range(0, i+1):
            pairs_to_drop.add((cols[i], cols[j]))
    return list(pairs_to_drop)

def get_top_abs_correlations(df, n=5):
    au_corr = df.corr().abs().unstack()
    labels_to_drop = get_redundant_pairs(df)
    au_corr = au_corr.drop(labels=labels_to_drop).sort_values(ascending=False)
    return au_corr[0:n]
```

new_cases	new_cases_smoothed	0.990950
total_cases	total_deaths	0.977766
new_cases_smoothed	total_deaths	0.976873
icu_patients	hosp_patients	0.975158
new_tests	new_tests_smoothed	0.975068
new_deaths	new_deaths_smoothed	0.973488
total_cases	new_cases_smoothed	0.970599
	dtype: float64	

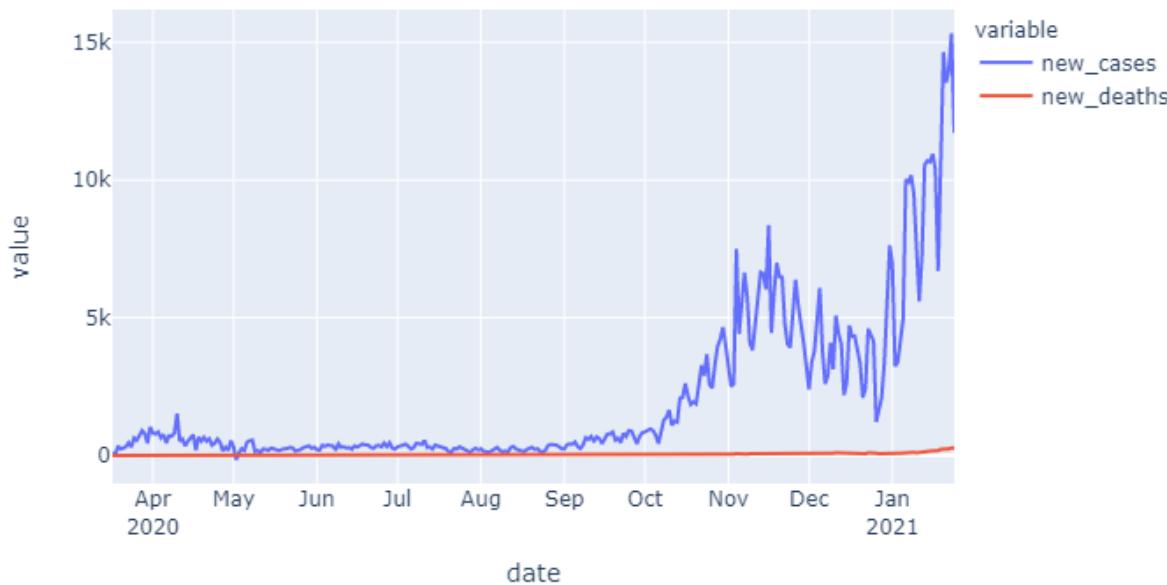
```
[MacBook-MacBook-Pro-de-Ariadna:argparse ariadnapuigventos$ python3 console.py -j 18
{'j': 18}
[MacBook-MacBook-Pro-de-Ariadna:argparse ariadnapuigventos$ python3 console.py -j 18
{'j': 18}
{'t_d_averages': [[['2020-03-03', 1.0], ['2020-03-04', 2.0], ['2020-03-05', 3.0], ['2020-03-06', 3.0], ['2020-03-07', 6.0], ['2020-03-08', 9.5], ['2020-03-09', 15.5], ['2020-03-10', 21.0], ['2020-03-11', 30.5], ['2020-03-12', 32.0], ['2020-03-13', 71.5], ['2020-03-14', 112.0], ['2020-03-15', 166.0], ['2020-03-16', 204.0], ['2020-03-17', 154.5], ['2020-03-18', 185.75], ['2020-03-19', 249.75], ['2020-03-20', 312.0], ['2020-03-21', 412.25], ['2020-03-22', 526.25], ['2020-03-23', 684.0], ['2020-03-24', 849.5], ['2020-03-25', 1113.25], ['2020-03-26', 1346.5], ['2020-03-27', 1296.2], ['2020-03-28', 1531.6], ['2020-03-29', 1746.8], ['2020-03-30', 2016.0], ['2020-03-31', 2259.6], ['2020-04-01', 2596.8], ['2020-04-02', 2941.0], ['2020-04-03', 3280.0], ['2020-04-04', 3600.4], ['2020-04-05', 3879.8], ['2020-04-06', 4152.0], ['2020-04-07', 4535.8], ['2020-04-08', 4916.4], ['2020-04-09', 5295.8], ['2020-04-10', 5672.0], ['2020-04-11', 5971.8], ['2020-04-12', 6250.0], ['2020-04-13', 6530.2], ['2020-04-14', 6834.0], ['2020-04-15', 7170.6], ['2020-04-16', 7531.0], ['2020-04-17', 7882.0], ['2020-04-18', 8142.0], ['2020-04-19', 8341.4], ['2020-04-20', 8565.2], ['2020-04-21', 8925.4], ['2020-04-22', 9211.8], ['2020-04-23', 9466.6], ['2020-04-24', 9772.2], ['2020-04-25', 10037.4], ['2020-04-26', 10192.2], ['2020-04-27', 10347.0], ['2020-04-28', 10623.8], ['2020-04-29', 10891.4], ['2020-04-30', 11095.2], ['2020-05-01', 11254.4], ['2020-05-02', 11501.6], ['2020-05-03', 11601.2], ['2020-05-04', 11705.4], ['2020-05-05', 11901.4], ['2020-05-06', 12095.4], ['2020-05-07', 12244.2], ['2020-05-08', 12417.4], ['2020-05-09', 12520.6], ['2020-05-10', 12603.8], ['2020-05-11', 12678.8], ['2020-05-12', 12851.4], ['2020-05-13', 12991.8], ['2020-05-14', 13118.6], ['2020-05-15', 13227.2], ['2020-05-16', 13341.0], ['2020-05-17', 13366.2], ['2020-05-18', 13433.4], ['2020-05-19', 13556.2], ['2020-05-20', 13651.8], ['2020-05-21', 13725.0], ['2020-05-22', 13928.6], ['2020-05-23', 13991.6], ['2020-05-24', 14091.4], ['2020-05-25', 13737.2], ['2020-05-26', 13828.2], ['2020-05-27', 13922.2], ['2020-05-28', 13999.8], ['2020-05-29', 14064.0], ['2020-05-30', 14103.4], ['2020-05-31', 14123.6], ['2020-06-01', 14148.8], ['2020-06-02', 14205.8], ['2020-06-03', 14264.2], ['2020-06-04', 14297.0], ['2020-06-05', 14354.4], ['2020-06-06', 14389.6], ['2020-06-07', 14406.2], ['2020-06-08', 14420.6], ['2020-06-09', 14465.2], ['2020-06-10', 14502.4], ['2020-06-11', 14522.4], ['2020-06-12', 14551.8], ['2020-06-13', 14577.6], ['2020-06-14', 14587.2], ['2020-06-15', 14597.4], ['2020-06-16', 14625.4], ['2020-06-17', 14651.6], ['2020-06-18', 14669.4], ['2020-06-19', 14927.6], ['2020-06-20', 14948.4], ['2020-06-21', 14959.8], ['2020-06-22', 14968.6], ['2020-06-23', 14994.2], ['2020-06-24', 15018.0], ['2020-06-25', 15044.0], ['2020-06-26', 15066.4], ['2020-06-27', 15079.8], ['2020-06-28', 15090.4], ['2020-06-29', 15100.4], ['2020-06-30', 15118.4], ['2020-07-01', 15144.6], ['2020-07-02', 15159.2], ['2020-07-03', 15178.8], ['2020-07-04', 15191.2], ['2020-07-05', 15201.2], ['2020-07-06', 15209.0], ['2020-07-07', 15226.8], ['2020-07-08', 15244.6], ['2020-07-09', 15259.0], ['2020-07-10', 15271.8], ['2020-07-11', 15281.4], ['2020-07-12', 15289.0], ['2020-07-13', 15296.6], ['2020-07-14', 15311.8], ['2020-07-15', 15323.6], ['2020-07-16', 15334.6], ['2020-07-17', 15345.4], ['2020-07-18', 15351.6], ['2020-07-19', 15358.4], ['2020-07-20', 15365.4], ['2020-07-21', 15376.4], ['2020-07-22', 15385.8], ['2020-07-23', 15393.4], ['2020-07-24', 15406.2], ['2020-07-25', 15414.0], ['2020-07-26', 15420.0], ['2020-07-27', 15425.6], ['2020-07-28', 15434.8], ['2020-07-29', 15447.0], ['2020-07-30', 15451.2], ['2020-07-31', 15461.8], ['2020-08-01', 15469.6], ['2020-08-02', 15475.4], ['2020-08-03', 15486.0], ['2020-08-04', 15500.0], ['2020-08-05', 15508.6], ['2020-08-06', 15517.2], ['2020-08-07', 15525.0], ['2020-08-08', 15531.0], ['2020-08-09', 15537.8], ['2020-08-10', 15560.6], ['2020-08-11', 15569.4], ['2020-08-12', 15579.0], ['2020-08-13', 15595.6], ['2020-08-14', 15606.4], ['2020-08-15', 15613.8], ['2020-08-16', 15620.2], ['2020-08-17', 15632.6], ['2020-08-18', 15646.6], ['2020-08-19', 15681.6], ['2020-08-20', 15691.8], ['2020-08-21', 15703.6], ['2020-08-22', 15713.2], ['2020-08-23', 15719.8], ['2020-08-24', 15733.6], ['2020-08-25', 15754.0], ['2020-08-26', 15772.8], ['2020-08-27', 15787.2], ['2020-08-28', 15802.0], ['2020-08-29', 15814.6], ['2020-08-30', 15824.6], ['2020-08-31', 15852.2], ['2020-09-01', 15875.2], ['2020-09-02', 15896.6], ['2020-09-03', 15918.4], ['2020-09-04', 15970.6], ['2020-09-05', 15986.8], ['2020-09-06', 15999.8], ['2020-09-07', 16033.6], ['2020-09-08', 16068.2], ['2020-09-09', 16089.8], ['2020-09-10', 16120.6], ['2020-09-11', 16144.8], ['2020-09-12', 16159.0], ['2020-09-13', 16174.4], ['2020-09-14', 16211.6], ['2020-09-15', 16264.0], ['2020-09-16', 16330.8], ['2020-09-17', 16384.4], ['2020-09-18', 16423.4], ['2020-09-19', 16445.2], ['2020-09-20', 16465.2], ['2020-09-21', 16517.8], ['2020-09-22', 16589.2], ['2020-09-23', 16639.6], ['2020-09-24', 16681.2], ['2020-09-25', 16728.4], ['2020-09-26', 16753.0], ['2020-09-27', 16773.0], ['2020-09-28', 16826.8], ['2020-09-29', 16857.4], ['2020-09-30', 16963.6], ['2020-10-01', 17027.8], ['2020-10-02', 17079.0], ['2020-10-03', 17104.2], ['2020-10-04', 17125.0], ['2020-10-05', 17171.6], ['2020-10-06', 17254.2], ['2020-10-07', 17297.4], ['2020-10-08', 17353.0], ['2020-10-09', 17433.2], ['2020-10-10', 17463.0], ['2020-10-11', 17491.6], ['2020-10-12', 17556.4], ['2020-10-13', 17617.8], ['2020-10-14', 17700.6], ['2020-10-15', 17772.8], ['2020-10-16', 17864.2], ['2020-10-17', 17912.2], ['2020-10-18', 17944.8], ['2020-10-19', 18022.6], ['2020-10-20', 18134.4], ['2020-10-21', 18221.8], ['2020-10-22', 18309.2], ['2020-10-23', 18422.2], ['2020-10-24', 18476.2], ['2020-10-25', 18525.2], ['2020-10-26', 18622.6], ['2020-10-27', 18770.8], ['2020-10-28', 18887.4], ['2020-10-29', 19000.0], ['2020-10-30', 19127.0], ['2020-10-31', 19216.0], ['2020-11-01', 19271.2], ['2020-11-02', 19399.6], ['2020-11-03', 19552.2], ['2020-11-04', 20003.2], ['2020-11-05', 20178.8], ['2020-11-06', 20346.6], ['2020-11-07', 20457.2], ['2020-11-08', 20515.2], ['2020-11-09', 20687.0], ['2020-11-10', 20824.0], ['2020-11-11', 21129.6], ['2020-11-12', 21347.2], ['2020-11-13', 21517.0], ['2020-11-14', 21639.6], ['2020-11-15', 21706.8], ['2020-11-16', 21884.2], ['2020-11-17', 22128.4], ['2020-11-18', 22343.8], ['2020-11-19', 22533.6], ['2020-11-20', 22742.4], ['2020-11-21', 22850.4], ['2020-11-22', 22973.0], ['2020-11-23', 23092.0], ['2020-11-24', 23211.6], ['2020-11-25', 23330.2], ['2020-11-26', 23449.8], ['2020-11-27', 23568.4], ['2020-11-28', 23687.0], ['2020-11-29', 23806.6], ['2020-11-30', 23925.2], ['2020-12-01', 24044.8], ['2020-12-02', 24163.4], ['2020-12-03', 24282.0], ['2020-12-04', 24401.6], ['2020-12-05', 24520.2], ['2020-12-06', 24639.8], ['2020-12-07', 24758.4], ['2020-12-08', 24877.0], ['2020-12-09', 25000.0], ['2020-12-10', 25120.0], ['2020-12-11', 25240.0], ['2020-12-12', 25360.0], ['2020-12-13', 25480.0], ['2020-12-14', 25600.0], ['2020-12-15', 25720.0], ['2020-12-16', 25840.0], ['2020-12-17', 25960.0], ['2020-12-18', 26080.0], ['2020-12-19', 26200.0], ['2020-12-20', 26320.0], ['2020-12-21', 26440.0], ['2020-12-22', 26560.0], ['2020-12-23', 26680.0], ['2020-12-24', 26800.0], ['2020-12-25', 26920.0], ['2020-12-26', 27040.0], ['2020-12-27', 27160.0], ['2020-12-28', 27280.0], ['2020-12-29', 27400.0], ['2020-12-30', 27520.0], ['2020-12-31', 27640.0], ['2021-01-01', 27760.0], ['2021-01-02', 27880.0], ['2021-01-03', 28000.0], ['2021-01-04', 28120.0], ['2021-01-05', 28240.0], ['2021-01-06', 28360.0], ['2021-01-07', 28480.0], ['2021-01-08', 28600.0], ['2021-01-09', 28720.0], ['2021-01-10', 28840.0], ['2021-01-11', 28960.0], ['2021-01-12', 29080.0], ['2021-01-13', 29200.0], ['2021-01-14', 29320.0], ['2021-01-15', 29440.0], ['2021-01-16', 29560.0], ['2021-01-17', 29680.0], ['2021-01-18', 29800.0], ['2021-01-19', 29920.0], ['2021-01-20', 30040.0], ['2021-01-21', 30160.0], ['2021-01-22', 30280.0], ['2021-01-23', 30400.0], ['2021-01-24', 30520.0], ['2021-01-25', 30640.0], ['2021-01-26', 30760.0], ['2021-01-27', 30880.0], ['2021-01-28', 31000.0], ['2021-01-29', 31120.0], ['2021-01-30', 31240.0], ['2021-01-31', 31360.0], ['2021-02-01', 31480.0], ['2021-02-02', 31600.0], ['2021-02-03', 31720.0], ['2021-02-04', 31840.0], ['2021-02-05', 31960.0], ['2021-02-06', 32080.0], ['2021-02-07', 32200.0], ['2021-02-08', 32320.0], ['2021-02-09', 32440.0], ['2021-02-10', 32560.0], ['2021-02-11', 32680.0], ['2021-02-12', 32800.0], ['2021-02-13', 32920.0], ['2021-02-14', 33040.0], ['2021-02-15', 33160.0], ['2021-02-16', 33280.0], ['2021-02-17', 33400.0], ['2021-02-18', 33520.0], ['2021-02-19', 33640.0], ['2021-02-20', 33760.0], ['2021-02-21', 33880.0], ['2021-02-22', 34000.0], ['2021-02-23', 34120.0], ['2021-02-24', 34240.0], ['2021-02-25', 34360.0], ['2021-02-26', 34480.0], ['2021-02-27', 34600.0], ['2021-02-28', 34720.0], ['2021-02-29', 34840.0], ['2021-02-30', 34960.0], ['2021-02-31', 35080.0], ['2021-03-01', 35200.0], ['2021-03-02', 35320.0], ['2021-03-03', 35440.0], ['2021-03-04', 35560.0], ['2021-03-05', 35680.0], ['2021-03-06', 35800.0], ['2021-03-07', 35920.0], ['2021-03-08', 36040.0], ['2021-03-09', 36160.0], ['2021-03-10', 36280.0], ['2021-03-11', 36400.0], ['2021-03-12', 36520.0], ['2021-03-13', 36640.0], ['2021-03-14', 36760.0], ['2021-03-15', 36880.0], ['2021-03-16', 37000.0], ['2021-03-17', 37120.0], ['2021-03-18', 37240.0], ['2021-03-19', 37360.0], ['2021-03-20', 37480.0], ['2021-03-21', 37600.0], ['2021-03-22', 37720.0], ['2021-03-23', 37840.0], ['2021-03-24', 37960.0], ['2021-03-25', 38080.0], ['2021-03-26', 38200.0], ['2021-03-27', 38320.0], ['2021-03-28', 38440.0], ['2021-03-29', 38560.0], ['2021-03-30', 38680.0], ['2021-03-31', 38800.0], ['2021-04-01', 38920.0], ['2021-04-02', 39040.0], ['2021-04-03', 39160.0], ['2021-04-04', 39280.0], ['2021-04-05', 39400.0], ['2021-04-06', 39520.0], ['2021-04-07', 39640.0], ['2021-04-08', 39760.0], ['2021-04-09', 39880.0], ['2021-04-10', 40000.0], ['2021-04-11', 40120.0], ['2021-04-12', 40240.0], ['2021-04-13', 40360.0], ['2021-04-14', 40480.0], ['2021-04-15', 40600.0], ['2021-04-16', 40720.0], ['2021-04-17', 40840.0], ['2021-04-18', 40960.0], ['2021-04-19', 41080.0], ['2021-04-20', 41200.0], ['2021-04-21', 41320.0], ['2021-04-22', 41440.0], ['2021-04-23', 41560.0], ['2021-04-24', 41680.0], ['2021-04-25', 41800.0], ['2021-04-26', 41920.0], ['2021-04-27', 42040.0], ['2021-04-28', 42160.0], ['2021-04-29', 42280.0], ['2021-04-30', 42400.0], ['2021-05-01', 42520.0], ['2021-05-02', 42640.0], ['2021-05-03', 42760.0], ['2021-05-04', 42880.0], ['2021-05-05', 43000.0], ['2021-05-06', 43120.0], ['2021-05-07', 43240.0], ['2021-05-08', 43360.0], ['2021-05-09', 43480.0], ['2021-05-10', 43600.0], ['2021-05-11', 43720.0], ['2021-05-12', 43840.0], ['2021-05-13', 43960.0], ['2021-05-14', 44080.0], ['2021-05-15', 44200.0], ['2021-05-16', 44320.0], ['2021-05-17', 44440.0], ['2021-05-18', 44560.0], ['2021-05-19', 44680.0], ['2021-05-20', 44800.0], ['2021-05-21', 44920.0], ['2021-05-22', 45040.0], ['2021-05-23', 45160.0], ['2021-05-24', 45280.0], ['2021-05-25', 45400.0], ['2021-05-26', 45520.0], ['2021-05-27', 45640.0], ['2021-05-28', 45760.0], ['2021-05-29', 45880.0], ['2021-05-30', 45900.0], ['2021-05-31', 45920.0], ['2021-06-01', 45940.0], ['2021-06-02', 45960.0], ['2021-06-03', 45980.0], ['2021-06-04', 46000.0], ['2021-06-05', 46020.0], ['2021-06-06', 46040.0], ['2021-06-07', 46060.0], ['2021-06-08', 46080.0], ['2021-06-09', 46100.0], ['2021-06-10', 46120.0], ['2021-06-11', 46140.0], ['2021-06-12', 46160.0], ['2021-06-13', 46180.0], ['2021-06-14', 46200.0], ['2021-06-15', 46220.0], ['2021-06-16', 46240.0], ['2021-06-17', 46260.0], ['2021-06-18', 46280.0], ['2021-06-19', 46300.0], ['2021-06-20', 46320.0], ['2021-06-21', 46340.0], ['2021-06-22', 46360.0], ['2021-06-23', 46380.0], ['2021-06-24', 46400.0], ['2021-06-25', 46420.0], ['2021-06-26', 46440.0], ['2021-06-27', 46460.0], ['2021-06-28', 46480.0], ['2021-06-29', 46500.0], ['2021-06-30', 46520.0], ['2021-07-01', 46540.0], ['2021-07-02', 46560.0], ['2021-07-03', 46580.0], ['2021-07-04', 46600.0], ['2021-07-05', 46620.0], ['2021-07-06', 46640.0], ['2021-07-07', 46660.0], ['2021-07-08', 46680.0], ['2021-07-09', 46700.0], ['2021-07-10', 46720.0], ['2021-07-11', 46740.0], ['2021-07-12', 46760.0], ['2021-07-13', 46780.0], ['2021-07-14', 46800.0], ['2021-07-15', 46820.0], ['2021-07-16', 46840.0], ['2021-07-17', 46860.0], ['2021-07-18', 46880.0], ['2021-07-19', 46900.0], ['2021-07-20', 46920.0], ['2021-07-21', 46940.0], ['2021-07-22', 46960.0], ['2021-07-23', 46980.0], ['2021-07-24', 47000.0], ['2021-07-25', 47020.0], ['2021-07-26', 47040.0], ['2021-07-27', 47060.0], ['2021-07-28', 47080.0], ['2021-07-29', 47100.0], ['2021-07-30', 47120.0], ['2021-07-31', 47140.0], ['2021-08-01', 47160.0], ['2021-08-02', 47180.0], ['2021-08-03', 47200.0], ['2021-08-04', 47220.0], ['2021-08-05', 47240.0], ['2021-08-06', 47260.0], ['2021-08-07', 47280.0], ['2021-08-08', 47300.0], ['2021-08-09',
```

# Visualization

---

- After that you show the fruits of your labor:

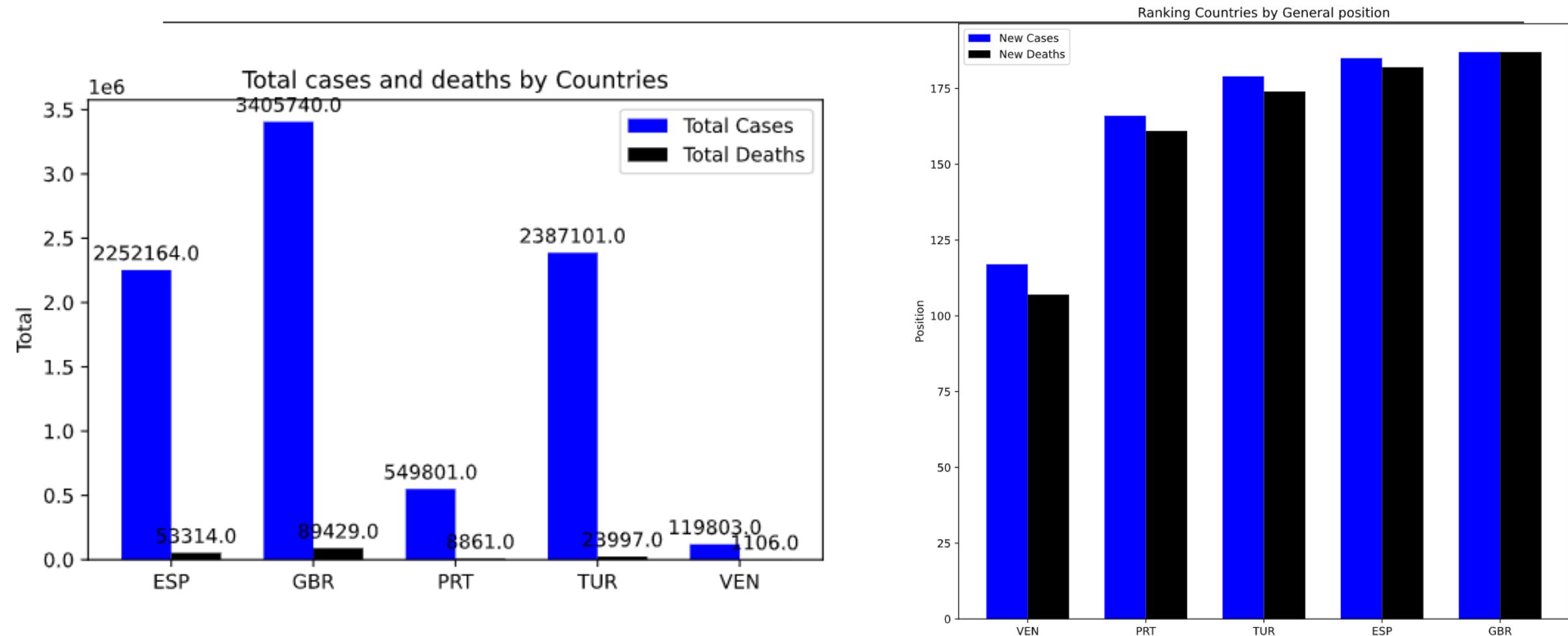
PRT covid deaths



TUR covid deaths



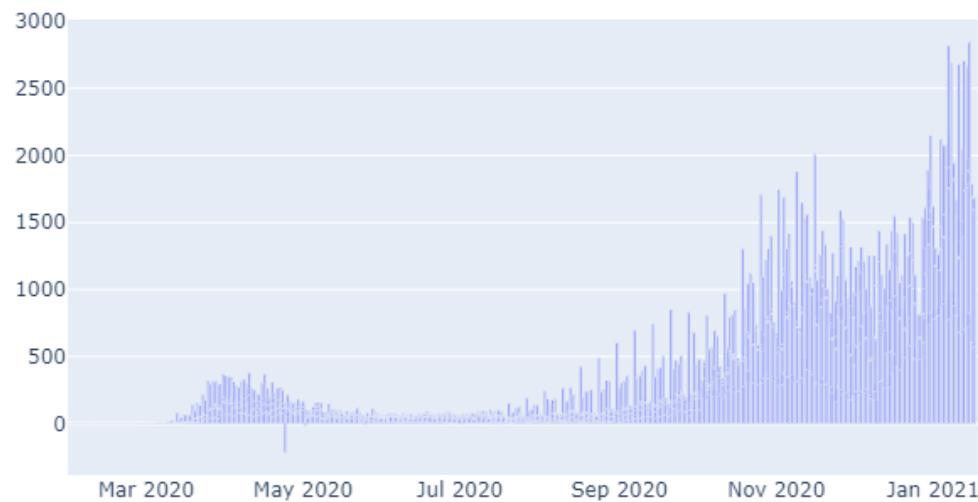
# Visualization



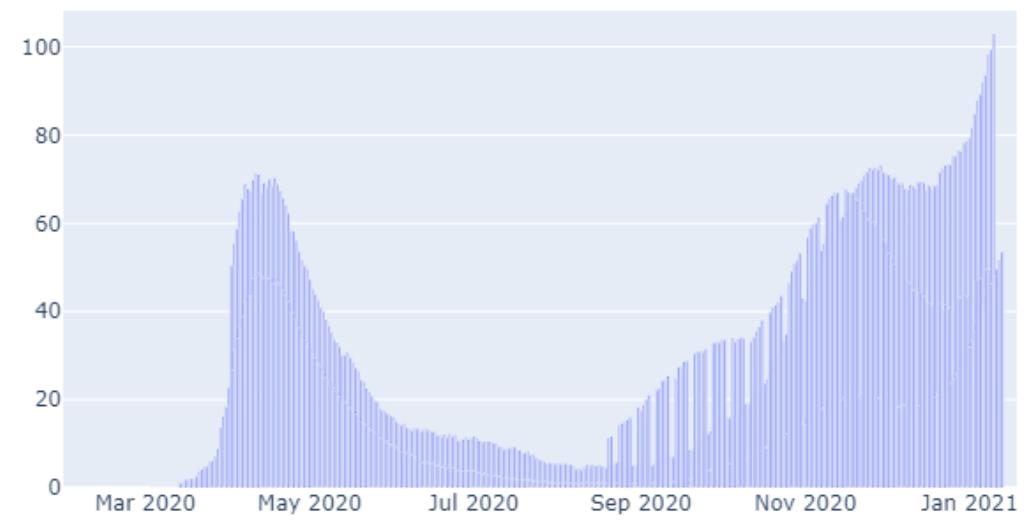
# Visualization

---

new\_cases\_per\_million



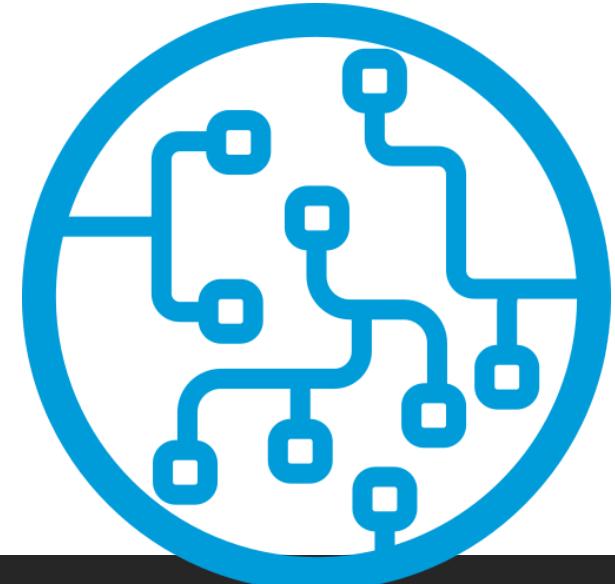
icu\_patients\_per\_million



# How to improve

---

- Achieve an A+
- Study de effectiveness of the vaccines.
- Use machine learning.
- Analyze how COVID-19 developed in different areas of the world.



Thank you for your  
time.

---

ANY QUESTIONS?

