```
In [59]:
         from io import BytesI0
         from zipfile import ZipFile
         import pandas as pd
         import numpy as np
         import os
         import datetime
In [60]:
         path = "/home/nacho/Documents/coronavirus/COVID-19_Paper/prediction_d
         ata"
         os.chdir(os.path.join(path))
In [61]: def print_values(df):
             for i in df[df.columns]:
                 #x = df[i].value_counts()
                 x = (df[i].value\_counts()/df[i].count())*100
                 print(i,"\n",x)
                 print()
             print(df.shape)
```

```
In [62]: #CASO 1 - si el paciente contagiado de CoV-2 necesitará hospitalizaci
ón

df = pd.read_csv("df_casol.zip")
    print_values(df)
```

```
SEX0
 0
      50.119343
1 49.880657
Name: SEXO, dtype: float64
TIPO_PACIENTE
 0 81.163394
     18.836606
Name: TIPO_PACIENTE, dtype: float64
EDAD
 30
       2.271840
31 2.223502
29 2.214040
28 2.200095
32 2.169334
117 0.000179
114 0.000179
113 0.000135
119
111
       0.000045
111
      0.000045
Name: EDAD, Length: 122, dtype: float64
EMBARAZ0
      99.367526
 0
      0.632474
Name: EMBARAZO, dtype: float64
DIABETES
      86.635583
     13.364417
Name: DIABETES, dtype: float64
EP0C
 0
      98.886568
      1.113432
Name: EPOC, dtype: float64
ASMA
      97.797573
 0
      2.202427
Name: ASMA, dtype: float64
INMUSUPR
 0
      99.175385
      0.824615
Name: INMUSUPR, dtype: float64
HIPERTENSION
 0
     82.621604
     17.378396
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
      98.44897
 0
1
      1.55103
```

### OBESIDAD

0 85.393414 1 14.606586

Name: OBESIDAD, dtype: float64

RENAL\_CRONICA 0 98.501971 1 1.498029

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 92.710159 1 7.289841

Name: TABAQUISMO, dtype: float64

(2230131, 13)

```
SEX0
0
      59.363903
     40.636097
Name: SEXO, dtype: float64
EDAD
60
       2.519420
61
       2.479646
   2.472937
2.468384
63
56
59
       2.467426
115 0.000240
110 0.000240
109 0.000240
106
       0.000240
118
       0.000240
Name: EDAD, Length: 114, dtype: float64
EMBARAZ0
0
      99.358584
1
      0.641416
Name: EMBARAZO, dtype: float64
DIABETES
     67.527638
 0
     32.472362
Name: DIABETES, dtype: float64
EP0C
0
      96.529598
      3.470402
Name: EPOC, dtype: float64
ASMA
0
      97.989975
      2.010025
Name: ASMA, dtype: float64
INMUSUPR
     97.948045
 0
      2.051955
Name: INMUSUPR, dtype: float64
HIPERTENSION
0
     61.580945
     38.419055
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
0
     95.838585
      4.161415
Name: CARDIOVASCULAR, dtype: float64
OBESIDAD
    78.65861
```

RENAL\_CRONICA 0 94.789605 1 5.210395

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

92.5279987.472002

Name: TABAQUISMO, dtype: float64

hosp\_critica 0 83.091255 1 16.908745

Name: hosp\_critica, dtype: float64

(417358, 13)

```
In [64]: #CASO 2: predecir en base a los descriptores la mortalidad (sin filtr
o)

df = pd.read_csv("df_caso2.zip")
    print_values(df)
```

```
SEX0
 0
      50.119343
1 49.880657
Name: SEXO, dtype: float64
TIPO_PACIENTE
 0 81.163394
     18.836606
Name: TIPO_PACIENTE, dtype: float64
EDAD
 30
       2.271840
31 2.223502
29 2.214040
28 2.200095
32 2.169334
117 0.000179
114 0.000179
113 0.000135
119
111
       0.000045
111
      0.000045
Name: EDAD, Length: 122, dtype: float64
EMBARAZ0
      99.367526
 0
      0.632474
Name: EMBARAZO, dtype: float64
DIABETES
      86.635583
     13.364417
Name: DIABETES, dtype: float64
EP0C
 0
      98.886568
      1.113432
Name: EPOC, dtype: float64
ASMA
      97.797573
 0
      2.202427
Name: ASMA, dtype: float64
INMUSUPR
 0
      99.175385
      0.824615
Name: INMUSUPR, dtype: float64
HIPERTENSION
 0
     82.621604
     17.378396
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
      98.44897
 0
1
      1.55103
```

### OBESIDAD

0 85.393414 1 14.606586

Name: OBESIDAD, dtype: float64

RENAL\_CRONICA 0 98.501971 1 1.498029

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 92.710159 1 7.289841

Name: TABAQUISMO, dtype: float64

# BOOL\_DEF

0 90.829552 1 9.170448

Name: BOOL\_DEF, dtype: float64

(2230131, 14)

```
SEX0
0
      59.363903
     40.636097
Name: SEXO, dtype: float64
INTUBADO
0 86.908362
     13.091638
Name: INTUBADO, dtype: float64
EDAD
60
       2.519420
   2.479646
2.472937
2.468384
61
63
56
59 2.467426
115 0.000240
110 0.000240
109 0.000240
106
       0.000240
118
       0.000240
Name: EDAD, Length: 114, dtype: float64
EMBARAZ0
     99.358584
0
      0.641416
Name: EMBARAZO, dtype: float64
DIABETES
     67.527638
     32.472362
Name: DIABETES, dtype: float64
EP0C
0
      96.529598
      3.470402
Name: EPOC, dtype: float64
ASMA
      97.989975
0
      2.010025
Name: ASMA, dtype: float64
INMUSUPR
 0
      97.948045
      2.051955
Name: INMUSUPR, dtype: float64
HIPERTENSION
0
     61.580945
     38.419055
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
     95.838585
1
      4.161415
```

### OBESIDAD

0 78.65861 1 21.34139

Name: OBESIDAD, dtype: float64

RENAL\_CRONICA 0 94.789605 1 5.210395

Name: RENAL\_CRONICA, dtype: float64

### TABAQUISMO

0 92.527998 1 7.472002

Name: TABAQUISMO, dtype: float64

### UCI

0 92.037052 1 7.962948

Name: UCI, dtype: float64

### BOOL DEF

0 55.602145 1 44.397855

Name: BOOL\_DEF, dtype: float64

(417358, 15)

```
In [65]:
         #CASO 3.1: Mortalidad de los contagiagos DESPUES de INTUBADO, UCI (con
         filtro)
         INTUBADO
         1
         100.0
         Name: INTUBADO, dtype: float64
         UCI
         1
         100.0
         Name: UCI, dtype: float64
         TIPO_PACIENTE
         1
         100.0
         Name: TIPO_PACIENTE, dtype: float64
         df = pd.read_csv("df_caso_3_1.zip")
         print_values(df)
```

```
SEX0
 0
      65.531989
     34.468011
Name: SEXO, dtype: float64
EDAD
 60
        2.901231
61 2.750968
67 2.658499
65 2.641160
66
   2.623822
7 0.023117
101 0.011559
98 0.011559
       0.005779
100
      0.005779
116
Name: EDAD, Length: 102, dtype: float64
EMBARAZ0
 0
      99.341155
      0.658845
Name: EMBARAZO, dtype: float64
DIABETES
      66.346876
 0
     33.653124
Name: DIABETES, dtype: float64
EP0C
 0
      96.971623
      3.028377
Name: EPOC, dtype: float64
ASMA
 0
      98.092816
      1.907184
Name: ASMA, dtype: float64
INMUSUPR
     97.468647
 0
      2.531353
Name: INMUSUPR, dtype: float64
HIPERTENSION
    61.544241
 0
     38.455759
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
 0
      95.133792
      4.866208
Name: CARDIOVASCULAR, dtype: float64
OBESIDAD
   71.941282
```

RENAL\_CRONICA 0 95.688609 1 4.311391

Name: RENAL\_CRONICA, dtype: float64

TABAQUISMO

0 92.140091 1 7.859909

Name: TABAQUISMO, dtype: float64

BOOL\_DEF

1 76.420274 0 23.579726

Name: BOOL\_DEF, dtype: float64

(17303, 13)

```
SEX0
 0
     63.807903
     36.192097
Name: SEXO, dtype: float64
EDAD
 63
       2.931972
61
       2.888047
       2.869745
65
       2.834971
67
       2.794707
100
       0.010981
99
      0.007321
101
    0.003660
116
       0.001830
108
       0.001830
Name: EDAD, Length: 104, dtype: float64
EMBARAZ0
 0
      99.685207
1
      0.314793
Name: EMBARAZO, dtype: float64
DIABETES
     64.710189
 0
     35.289811
Name: DIABETES, dtype: float64
EP0C
 0
      96.271894
      3.728106
Name: EPOC, dtype: float64
ASMA
 0
      97.977635
      2.022365
Name: ASMA, dtype: float64
INMUSUPR
     97.604275
 0
      2.395725
Name: INMUSUPR, dtype: float64
HIPERTENSION
 0
     57.788393
     42.211607
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
 0
      95.29823
      4.70177
Name: CARDIOVASCULAR, dtype: float64
OBESIDAD
   73.776972
```

RENAL\_CRONICA 0 94.73453 1 5.26547

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 91.659803 1 8.340197

Name: TABAQUISMO, dtype: float64

### UCI

0 68.332144 1 31.667856

Name: UCI, dtype: float64

# BOOL\_DEF

1 83.017625 0 16.982375

Name: BOOL\_DEF, dtype: float64

(54639, 14)

```
SEX0
 0
     62.878378
     37.121622
Name: SEXO, dtype: float64
INTUBADO
 1 52.064151
     47.935849
Name: INTUBADO, dtype: float64
EDAD
 60
       2.867545
61 2.647891
58 2.641873
65 2.494433
   2.491424
55
   0.018054
0.015045
98
99
101
    0.015045
       0.006018
103
116
       0.003009
Name: EDAD, Length: 104, dtype: float64
EMBARAZ0
      99.103328
 0
      0.896672
Name: EMBARAZO, dtype: float64
DIABETES
     66.507191
     33.492809
Name: DIABETES, dtype: float64
EP0C
 0
      96.762352
      3.237648
Name: EPOC, dtype: float64
ASMA
      97.996028
 0
      2.003972
Name: ASMA, dtype: float64
INMUSUPR
 0
      97.634952
      2.365048
Name: INMUSUPR, dtype: float64
HIPERTENSION
 0
     62.117109
     37.882891
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
      95.284949
1
      4.715051
```

OBESIDAD

0 73.289402 1 26.710598

Name: OBESIDAD, dtype: float64

RENAL\_CRONICA 0 96.046218 1 3.953782

Name: RENAL\_CRONICA, dtype: float64

TABAQUISMO

0 92.751399 1 7.248601

Name: TABAQUISMO, dtype: float64

BOOL\_DEF

1 56.312812 0 43.687188

Name: BOOL\_DEF, dtype: float64

(33234, 14)

```
SEX0
 0
      59.366202
     40.633798
Name: SEXO, dtype: float64
NEUMONIA
 1 65.745186
     34.254814
Name: NEUMONIA, dtype: float64
EDAD
 60
       2.519364
61 2.480036
63 2.473562
59 2.467327
   2.466608
56
115 0.000240
110 0.000240
109 0.000240
       0.000240
106
118
       0.000240
Name: EDAD, Length: 114, dtype: float64
EMBARAZ0
      99.358289
 0
      0.641711
Name: EMBARAZO, dtype: float64
DIABETES
     67.521163
     32.478837
Name: DIABETES, dtype: float64
EP0C
 0
      96.528381
      3.471619
Name: EPOC, dtype: float64
ASMA
      97.989976
 0
      2.010024
Name: ASMA, dtype: float64
INMUSUPR
 0
      97.947531
      2.052469
Name: INMUSUPR, dtype: float64
HIPERTENSION
 0
     61.575981
     38.424019
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
      95.838469
1
      4.161531
```

### OBESIDAD

0 78.658545 1 21.341455

Name: OBESIDAD, dtype: float64

RENAL\_CRONICA 0 94.788374 1 5.211626

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 92.528477 1 7.471523

Name: TABAQUISMO, dtype: float64

### UCI

0 92.046474 1 7.953526

Name: UCI, dtype: float64

(417010, 14)

```
SEX0
0
     60.443549
     39.556451
Name: SEXO, dtype: float64
EDAD
60
       2.628060
63
       2.591255
     2.576679
2.561739
65
61
       2.561739
59
       2.541697
116 0.000364
110 0.000364
108
    0.000364
104
       0.000364
118
       0.000364
Name: EDAD, Length: 110, dtype: float64
EMBARAZ0
0
      99.660014
1
      0.339986
Name: EMBARAZO, dtype: float64
DIABETES
 0
      66.272506
     33.727494
Name: DIABETES, dtype: float64
EP0C
0
      96.454379
      3.545621
Name: EPOC, dtype: float64
ASMA
 0
      98.094183
      1.905817
Name: ASMA, dtype: float64
INMUSUPR
     97.97466
 0
      2.02534
Name: INMUSUPR, dtype: float64
HIPERTENSION
0
     60.533191
     39.466809
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
0
     95.783881
      4.216119
Name: CARDIOVASCULAR, dtype: float64
OBESIDAD
    77.413701
```

RENAL\_CRONICA 0 94.921708 1 5.078292

Name: RENAL\_CRONICA, dtype: float64

TABAQUISMO

0 92.35669 1 7.64331

Name: TABAQUISMO, dtype: float64

UCI

0 89.556998 1 10.443002

Name: UCI, dtype: float64

(274423, 13)

```
In [19]: #CASO 6: necesidad de ventilador ANTES de DIAGNOSTICO de neumonia e I
    CU (sin filtro)
    TIPO_PACIENTE
    1
    100.0
    Name: TIPO_PACIENTE, dtype: float64
    '''
    df = pd.read_csv("df_caso6.zip")
    print_values(df)
```

```
SEX0
0
     59.366262
    40.633738
Name: SEXO, dtype: float64
INTUBADO
0 86.912076
     13.087924
Name: INTUBADO, dtype: float64
EDAD
60
       2.519316
     2.479989
61
   2.473514
2.467280
63
59
   2.466800
56
115 0.000240
110 0.000240
109 0.000240
106
       0.000240
118
       0.000240
Name: EDAD, Length: 114, dtype: float64
EMBARAZ0
     99.358301
0
      0.641699
Name: EMBARAZO, dtype: float64
DIABETES
     67.521066
     32.478934
Name: DIABETES, dtype: float64
EP0C
0
      96.528447
      3.471553
Name: EPOC, dtype: float64
ASMA
      97.990015
0
      2.009985
Name: ASMA, dtype: float64
INMUSUPR
 0
     97.947571
      2.052429
Name: INMUSUPR, dtype: float64
HIPERTENSION
0
     61.57552
     38.42448
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
 0
     95.838549
1
      4.161451
```

### OBESIDAD

0 78.658715 1 21.341285

Name: OBESIDAD, dtype: float64

RENAL\_CRONICA 0 94.788474 1 5.211526

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 92.52838 1 7.47162

Name: TABAQUISMO, dtype: float64

(417018, 13)

```
SEX0
0
     59.363903
    40.636097
Name: SEXO, dtype: float64
INTUBADO
0 86.908362
     13.091638
Name: INTUBADO, dtype: float64
NEUMONIA
1 65.752424
0
     34.247576
Name: NEUMONIA, dtype: float64
EDAD
60
       2.519420
61
     2.479646
   2.479040
63
   2.468384
2.467426
56
59
115 0.000240
110 0.000240
109
      0.000240
106
       0.000240
118
       0.000240
Name: EDAD, Length: 114, dtype: float64
EMBARAZ0
     99.358584
      0.641416
Name: EMBARAZO, dtype: float64
DIABETES
0
     67.527638
     32.472362
Name: DIABETES, dtype: float64
EP0C
      96.529598
 0
      3.470402
Name: EPOC, dtype: float64
ASMA
 0
      97.989975
      2.010025
Name: ASMA, dtype: float64
INMUSUPR
     97.948045
0
      2.051955
Name: INMUSUPR, dtype: float64
HIPERTENSION
0
   61.580945
```

Name: HIPERTENSION, dtype: float64

### CARDIOVASCULAR

0 95.838585 1 4.161415

Name: CARDIOVASCULAR, dtype: float64

### OBESIDAD

0 78.65861 1 21.34139

Name: OBESIDAD, dtype: float64

# ${\tt RENAL\_CRONICA}$

0 94.789605 1 5.210395

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 92.527998 1 7.472002

Name: TABAQUISMO, dtype: float64

### UCI

0 92.037052 1 7.962948

Name: UCI, dtype: float64

(417358, 15)

```
In [70]: #CASO 7.1: necesidad de ventilador DESPUES de DIAGNOSTICO de neumonia
         e ICU (con filtro)}
         TIPO_PACIENTE
         1
         100.0
         Name: TIPO_PACIENTE, dtype: float64
         UCI
         1
         100.0
         Name: UCI, dtype: float64
         NEUMONIA
         1
         100.0
         Name: NEUMONIA, dtype: float64
         df = pd.read_csv("df_caso_7_1.zip")
         print_values(df)
```

```
SEX0
 0
     63.776258
     36.223742
Name: SEXO, dtype: float64
INTUBADO
 1 56.741573
     43.258427
Name: INTUBADO, dtype: float64
EDAD
 60
       2.948566
61 2.679880
58 2.658943
55 2.578687
62
       2.557750
       0.013958
5
7
     0.013958
     0.010468
99
       0.003489
103
116
       0.003489
Name: EDAD, Length: 104, dtype: float64
EMBARAZ0
      99.340498
 0
      0.659502
Name: EMBARAZO, dtype: float64
DIABETES
     65.629144
     34.370856
Name: DIABETES, dtype: float64
EP0C
 0
      96.636192
      3.363808
Name: EPOC, dtype: float64
ASMA
      98.018005
 0
      1.981995
Name: ASMA, dtype: float64
INMUSUPR
 0
      97.634168
      2.365832
Name: INMUSUPR, dtype: float64
HIPERTENSION
 0
     61.281318
     38.718682
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
      95.254379
1
      4.745621
```

### OBESIDAD

0 72.318375 1 27.681625

Name: OBESIDAD, dtype: float64

RENAL\_CRONICA 0 95.987159 1 4.012841

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 92.490753 1 7.509247

Name: TABAQUISMO, dtype: float64

(28658, 13)

```
SEX0
0
     62.878378
    37.121622
Name: SEXO, dtype: float64
INTUBADO
1 52.064151
     47.935849
Name: INTUBADO, dtype: float64
NEUMONIA
1 86.230968
0
     13.769032
Name: NEUMONIA, dtype: float64
EDAD
60
       2.867545
61
      2.647891
   2.641873
58
   2.494433
2.491424
65
55
   0.018054
0.015045
98
99
101
       0.015045
       0.006018
103
116
       0.003009
Name: EDAD, Length: 104, dtype: float64
EMBARAZ0
     99.103328
      0.896672
Name: EMBARAZO, dtype: float64
DIABETES
0
     66.507191
     33.492809
Name: DIABETES, dtype: float64
EP0C
      96.762352
 0
      3.237648
Name: EPOC, dtype: float64
ASMA
 0
      97.996028
      2.003972
Name: ASMA, dtype: float64
INMUSUPR
     97.634952
0
      2.365048
Name: INMUSUPR, dtype: float64
HIPERTENSION
0
   62.117109
1
     37.882891
```

Name: HIPERTENSION, dtype: float64

### CARDIOVASCULAR

0 95.284949 1 4.715051

Name: CARDIOVASCULAR, dtype: float64

### OBESIDAD

0 73.289402 1 26.710598

Name: OBESIDAD, dtype: float64

# ${\tt RENAL\_CRONICA}$

0 96.046218 1 3.953782

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

0 92.751399 1 7.248601

Name: TABAQUISMO, dtype: float64

(33234, 14)

```
SEX0
0
     60.443549
     39.556451
Name: SEXO, dtype: float64
INTUBADO
0 83.196379
     16.803621
Name: INTUBADO, dtype: float64
EDAD
60
       2.628060
63
       2.591255
65
      2.576679
61
      2.561739
59
   2.541697
116 0.000364
110 0.000364
108 0.000364
104
       0.000364
118
       0.000364
Name: EDAD, Length: 110, dtype: float64
EMBARAZ0
0
     99.660014
      0.339986
Name: EMBARAZO, dtype: float64
DIABETES
     66.272506
     33.727494
Name: DIABETES, dtype: float64
EP0C
0
      96.454379
      3.545621
Name: EPOC, dtype: float64
ASMA
 0
      98.094183
1
      1.905817
Name: ASMA, dtype: float64
INMUSUPR
 0
     97.97466
      2.02534
Name: INMUSUPR, dtype: float64
HIPERTENSION
0
     60.533191
     39.466809
Name: HIPERTENSION, dtype: float64
CARDIOVASCULAR
     95.783881
1
      4.216119
```

```
Name: CARDIOVASCULAR, dtype: float64

OBESIDAD
0 77.413701
1 22.586299

Name: OBESIDAD, dtype: float64
```

RENAL\_CRONICA 0 94.921708 1 5.078292

Name: RENAL\_CRONICA, dtype: float64

# TABAQUISMO

92.35669 1 7.64331

Name: TABAQUISMO, dtype: float64

### UCI

0 89.556998 1 10.443002

Name: UCI, dtype: float64

(274423, 14)

In [ ]: