COVID19

May 20, 2023

[2]:
```
from matplotlib import pyplot as plot
import numpy as np
import pandas as pd
import scipy.stats as stats
from sklearn.preprocessing import StandardScaler, MinMaxScaler
```

[6]:
```
document = pd.read_csv(r'C:\Users\bonin\Downloads\covid_data.csv')
```

[17]:
```
(row, columns) = document.shape
```

[37]:
```
document = document.replace("male", 1)
document = document.replace("female", 2)
document = document.replace("unknown", 3)
```

[40]:
```
document = document.replace("NonICU", 0)
document = document.replace("ICU", 1)
```

[42]:
```
document.head()
```

[42]:

	Sample	Age	Sex	Severity	A1BG	A1CF	A2M	A2ML1	A3GALT2	A4GALT	…	\
0	C1	39	1	0	0.49	0.00	0.21	0.04	0.07	0.0	…	
1	C2	63	1	0	0.29	0.00	0.14	0.00	0.00	0.0	…	
2	C3	33	1	0	0.26	0.00	0.03	0.02	0.00	0.0	…	
3	C4	49	1	0	0.45	0.01	0.09	0.07	0.00	0.0	…	
4	C5	49	1	0	0.17	0.00	0.00	0.05	0.07	0.0	…	

	ZWILCH	ZWINT	ZXDA	ZXDB	ZXDC	ZYG11A	ZYG11B	ZYX	ZZEF1	ZZZ3
0	2.84	4.22	0.95	1.63	15.51	0.06	8.17	363.01	19.17	6.05
1	3.55	12.15	0.60	1.15	15.62	0.14	8.20	399.80	15.72	4.12
2	1.34	2.79	0.18	0.32	17.67	0.28	3.62	430.35	13.95	1.81
3	3.71	5.87	1.40	2.21	15.61	0.27	7.88	209.25	14.78	7.15
4	1.44	4.46	0.28	0.55	9.34	0.07	5.96	272.91	8.69	2.70

[5 rows x 19476 columns]

[43]:
```
numeric_cols = document.select_dtypes(include=['int64', 'float64']).columns
```