In [25]:
```
import numpy as np
import pandas as pd

df = pd.read_csv('owid-covid-data.csv')


country_code = df["iso_code"].unique() #getting unique country codes
```

In [51]:
```
# creating a new dataframe with country code as index and dates as columns

new_df = pd.DataFrame(data={"date": new_header}).sort_values(by="date").reset_index(drop=True)

for code in country_code:
    test = df[df["iso_code"] == code][["date","total_cases"]]
    new_df = new_df.merge(test,how="left",on="date")
    new_df = new_df.rename(mapper={"total_cases": code}, axis=1)

new_df = new_df.T
new_df.columns = new_df.iloc[0] #changing column names to the date
new_df = new_df.drop(index="date",axis=0) #dropping the first row which contain column names

new_df.head()
```

Out[51]:

date	2020-01-01	2020-01-02	2020-01-03	2020-01-04	2020-01-05	2020-01-06	2020-01-07	2020-01-08	2020-01-09	2020-01-10	...	2021-03-13	2021-03-14	2021-03-15	2021-03-16
AFG	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	55959	55985	55985	55995
OWID_AFR	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	4.02705e+06	4.03571e+06	4.04412e+06	4.0531e+06
ALB	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	116821	117474	118017	118492
DZA	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	115143	115265	115410	115540
AND	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	11228	11266	11289	11319

5 rows × 447 columns

Time Series data is almost ready. Just have to fill the missing values due to having no Covid-19 cases with 0.

We should also remove rows starting with "OWID" like "OWID_AFR" as they are not countries.

Finally, we should replace the country code index with country names.

In [52]:
```
new_df = new_df.fillna(value=0) #filling NaNs with 0
new_df = new_df[~new_df.index.str.contains(r"OWID")] #dropping rows with OWID

#replacing country code with names
country_name = {}
for code in new_df.index:
    name = df[df["iso_code"] == code]["location"].iloc[0]
    country_name[code] = name
new_df = new_df.rename(country_name,axis=0)
new_df
```

Out[52]:

date	2020-01-01	2020-01-02	2020-01-03	2020-01-04	2020-01-05	2020-01-06	2020-01-07	2020-01-08	2020-01-09	2020-01-10	...	2021-03-13	2021-03-14	2021-03-15	2021-03-16	2021-03-17
Afghanistan	0	0	0	0	0	0	0	0	0	0	...	55959.0	55985.0	55985.0	55995.0	56016.0
Albania	0	0	0	0	0	0	0	0	0	0	...	116821.0	117474.0	118017.0	118492.0	118938.0
Algeria	0	0	0	0	0	0	0	0	0	0	...	115143.0	115265.0	115410.0	115540.0	115688.0
Andorra	0	0	0	0	0	0	0	0	0	0	...	11228.0	11266.0	11289.0	11319.0	11360.0
Angola	0	0	0	0	0	0	0	0	0	0	...	21323.0	21380.0	21407.0	21446.0	21489.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Venezuela	0	0	0	0	0	0	0	0	0	0	...	145379.0	145379.0	146488.0	147028.0	147577.0
Vietnam	0	0	0	0	0	0	0	0	0	0	...	2553.0	2554.0	2557.0	2560.0	2567.0
Yemen	0	0	0	0	0	0	0	0	0	0	...	2771.0	2836.0	2908.0	2969.0	3037.0
Zambia	0	0	0	0	0	0	0	0	0	0	...	84474.0	84797.0	84950.0	85240.0	85502.0
Zimbabwe	0	0	0	0	0	0	0	0	0	0	...	36471.0	36484.0	36504.0	36535.0	36552.0

204 rows × 447 columns

In [ ]:
```
We are now ready to export the dataset.
```

In [53]:
```
new_df.to_csv('new_covid_data2.csv')
```