## **Summary:**

The project deals with the analysis of information about the corona virus in the world,

Data analysis is based on information from Johns Hopkins University,

The data were taken from:

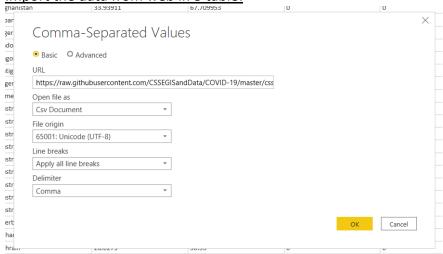
https://github.com/CSSEGISandData/COVID-19/tree/master/csse covid 19 data/csse covid 19 time series

Transformation of the data was performed in 2 different forms on the same information:

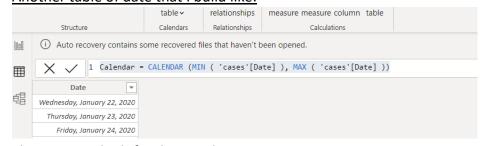
- Power bi
- Ssis

## Details of the project in power bi:

Import the data from web in 3 table:



- The three tables perform a transformation of
  - Unpivoted other columns in the table
  - Then merge three tables to a new table
- Another table of date that I build like:



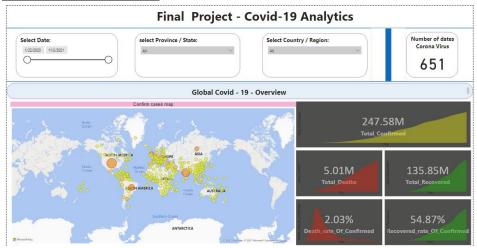
• The measure built for data analysis are:

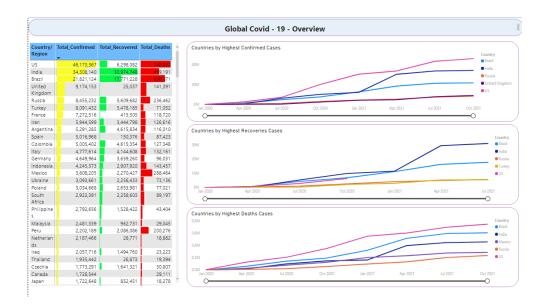
o Daily New Cases

Total\_Confirmed

```
1 Total_Confirmed =
2 VAR sum_confirmed =
3 SUMX( VALUES('cases'[Country/Region]),
4 SUMX( VALUES('cases'[Province/State]),
5 CALCULATE(MAX('cases'[Confirmed]))
6 )
7 )
8 RETURN
9 IF (sum_confirmed = 0, BLANK(), sum_confirmed )
```

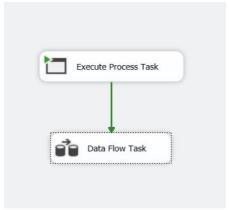
- o Total\_Deaths
- o Total Recovered
- o Recovered\_rate\_Of\_Confirmed
- o Death\_rate\_Of\_Confirmed
- After arranging the data and building the measure the reports built are for a global overview by country on:



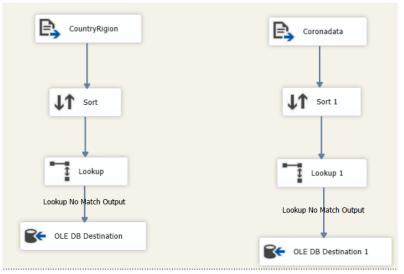


## **Details of the project in ssis:**

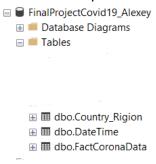
- In ssis folder you can find file:
  - Python for Treatment the data.
  - Csv data file with result after python.
  - Flow chart of the ssis.
- Steps in project execution:
  - Script Import and clean and join 3 table data with python and then pool the data to data flow task



 In the data flow you can find Basic construction of a data warehouse with 2 tables can be found by checking every time new information arrives if it already exists or does not exist in the table:



o In the final step can be fond the new db whit the table:



Thank,