

# COVID-19 on the Time, Mexico and the World

Juan J. Martínez, Ildar Batyrshin, Alexander Gelbukh

Instituto Politécnico Nacional,  
Centro de Investigación en Computación,  
Mexico

{juanjmtzs, batyr1}@gmail.com,  
gelbukh@gelbukh.com

**Abstract:** This paper describes a brief analysis on the COVID-19 over the time Globally and specifically centered in Mexico. In this work, it was created a website and using Google Data Studio for analytical dashboards it is tracked the daily data dynamics of the pandemic, which is collected and represented graphically. For all data collecting it was developed various web scraping scripts mainly based on bash scripting and python which extract data from specific web sites and once the initial inputs are obtained, the transforming process is started making aggregations, key performance indicators, correlations and mappings giving the facility of using that transformed data for future works. Furthermore, it is also used an specific model for studying the changing aspects of the epidemic and this is presented to analyze the rates, model is discussed and it is shown how it can be used to track impactful decisions to reduce the positive cases identified. The data has been collected and treated for study from different sources [1, 2, 3, 4]. Additionally, all the results and final data after transformations is being published on a daily basis in the following sites [5, 6, 7, 8].

**Keywords:** COVID-19, coronavirus, SARS-CoV-2, COVID19 mx, contagion rate.

## 1 Introduction

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe respiratory syndrome coronavirus (SARS-CoV-2). It has being rapidly spread all over the world in only six months. On 31 December 2019 Wuhan Municipal Health Commission, China, reported a cluster of cases of pneumonia in Wuhan, Hubei Province. On January the 4<sup>th</sup> 2020, WHO (World Health Organization) reported on social media that there was a cluster of pneumonia cases – with no deaths – in Wuhan. On the 13<sup>th</sup> of January 2020 officials confirmed a case of COVID-19 in Thailand, the first recorded case outside of China. On 11 March 2020 deeply concerned both by the alarming levels of spread and severity, and by the alarming levels of inaction, WHO made the assessment that COVID-19 could be characterized as a pandemic. Nowadays more than 11.62 million people has being infected all over the world.

The virus is primarily spread between people during close contact, most often via small droplets produced by coughing, sneezing, and talking. The droplets usually fall

to the ground or onto surfaces rather than travelling through air over long distances, although in some cases they may remain airborne for tens of minutes. Less commonly, people may become infected by touching a contaminated surface and then touching their face. It is also known the COVID-19 is most contagious during the first three days after the onset of symptoms, although spread is possible before symptoms appear, and from people who do not show any kind of them.

On this work, many data science techniques are applied and represented visually in order to disclose data analytics for the people to stay informed and compare virus dynamics between countries.

## 2 How Infection Rate is Analysed

To analyse daily behaviour of the pandemic the following model is used where exist an initial formula, which is the Delta of the Infected People on a day  $\Delta I$  that is directly related to three factors:

- **Exposure:** ( $E$ ) describes the contact with the virus not saying at home and being in continuous interactions with people, in other words, not respecting social distancing staying at least 2m apart from others.
- **Contagion probability:** ( $P$ ) describes the compliance of instructions to avoid contagion like washing your hands, using anti-bacterial hand gel, using mask of at least three layers and N95 in case of contact with infected people, avoid touching your face and eyes, etc.
- **Infected people:** ( $I_n$ ) the total confirmed people with the infection on a  $n$  day.

The initial formula is as follows:

$$\Delta I = E * P * I_n .$$

Now,  $\Delta I$  is simply calculated by the difference of those infected on the day  $I_{n+1}$  minus the day  $I_n$ , in other words today's infections subtracting yesterday's infections:

$$\Delta I = I_{n+1} - I_n .$$

Then substitute the new equation in the original one:

$$I_{n+1} - I_n = E * P * I_n .$$

Once with the substitution, process continues isolating the variable  $I_{n+1}$  and factoring the equation:

$$I_{n+1} = E * P * I_n + I_n ,$$

$$I_{n+1} = I_n (E * P + 1) .$$

Now  $(E * P + 1)$  can be named as **Contagion rate** ( $Cr$ ) which is the actual problem factor on the propagation of the virus. Substituting and isolating the new variable the resulted equation is:

$$I_{n+1} = I_n Cr ,$$



**Fig. 1.** Contagion rate in Mexico from 27 of February to 7<sup>th</sup> July 2020.

$$Cr = \frac{I_{n+1}}{I_n}.$$

For knowing that the virus propagation is stopped,  $Cr$  must be one which means that there are no more new infections every day. Applying the proposed model, for the official data of Mexico, the representation looks as shown in Fig. 1.

Mexico and Latin America have been widely affected, on this regard an important methodology taken in Mexico was the National Period of Healthy Distance (In Spanish, Jornada Nacional de Sana Distancia) which started on March 23<sup>rd</sup> 2020. Its main rules consisted in [9]:

- **Basic Prevention:** Frequent hand washing; when sneezing, cover nose and mouth with the internal angle of the elbow; avoid physical greeting; and do not leave home if there are symptoms comparable to the coronavirus.
- **Temporary Suspension of Non-Essential Activities:** The Ministry of Health recommended to Mexican society to suspend the activities considered "not essential", that is, those that do not affect the "substantive activity" of companies, organizations and the government itself as of March 23. Also, those that gather people or that imply that there are constant displacements in public transport. Among the instructions given were:
  - a. Avoid conglomerations of more than 100 people.
  - b. Maximum limitation of public sector activity.
  - c. Interruption of school activity at all educational levels.
  - d. Closure of cultural sites such as museums, cultural centres and archaeological zones.
  - e. Limitation of recreational sites such as activity in cinemas, theatres, restaurants and bars.
- **Reprogramming of Events of Massive Concentration:** Such as concerts or tourist or employer's fairs.
- **Provide Protection and Care for Elderly People:** As they were considered the most vulnerable sector against possible COVID-19 infection.



**Fig. 2.** Contagion rate in Mexico from 27 of February to 23 of March 2020. Contagion rate in Mexico from 23 of March to 7<sup>th</sup> July 2020.

In order to analyse the data in a better way and exemplify our model, a segregation was done over the time series, separating it in two sections for now, before and after 23 of March 2020.

Comparing both periods it is demonstrated a positive result with the National Period Of Healthy Distance strategy implemented by Mexican Government using the proposed model, as in Fig. 2 first chart  $Cr$  is observed a SIDEWAYS TREND while in the second chart is seen a DOWNWARD TREND reducing the infection propagation in the country.

With this support, is suggested to use this model actively as a key indicator of the external factors and its impact on the pandemic dynamics to know if a decision is having good results or wrong ones.

### 3 Current Status in Mexico and the World

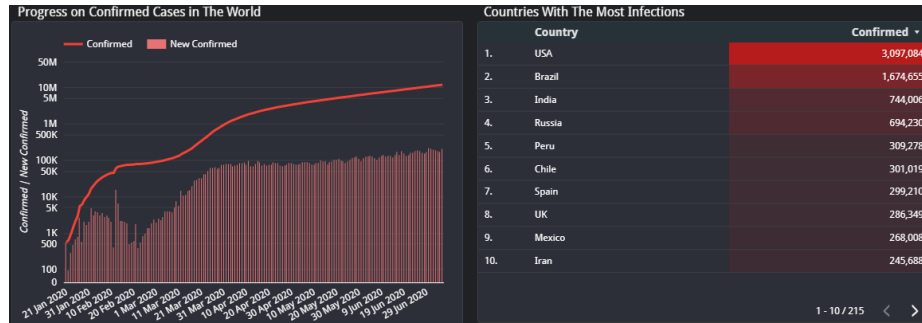
It has passed over 168 days on the world since first detected case and 132 days in Mexico. In each country different decisions have been made by their governments in order to reduce the infection propagation. Here some data statistics are presented.

#### 3.1 World

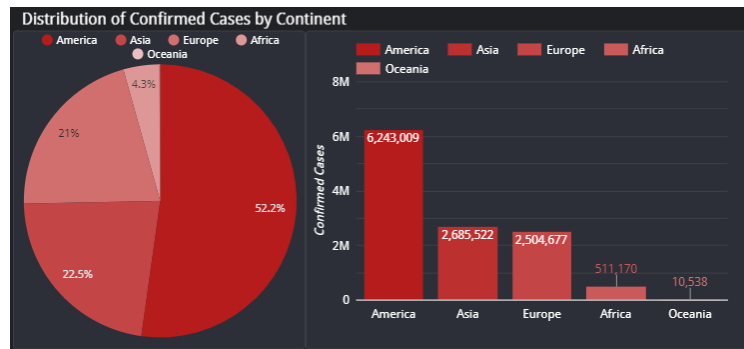
As for July the 7<sup>th</sup> the related numbers for the pandemic globally are:

- 11,829,602-Confirmed Cases,
- 544,163-Deaths,
- 6,447,656-Recovered People,
- ⚠ 1.018-Contagion Rate,
- 💀 1.011-Deaths rate,
- ✅ 1.023-Recovered Rate.

Reviewing the progress on the time for confirmed cases and daily new confirmed cases in the world is seen in Fig. 3 that daily new cases is in a SIDEWAYS TREND and the plateau was reached. Moreover, in the same picture is shown top countries by confirmed cases being USA the most critical country with a difference of around 1.4 million from the second place that is being occupied by Brazil.



**Fig. 3.** Time series chart for confirmed cases and daily confirmed cases in the world. Top Countries by total confirmed cases in the world [6].



**Fig. 4.** Confirmed cases by continent in the world [6].

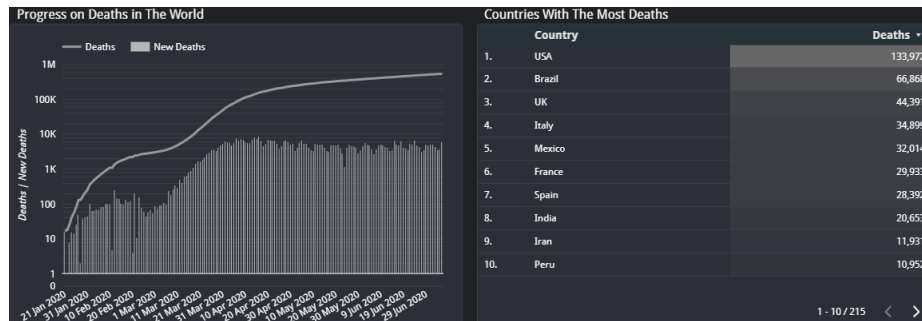
In Fig. 4 is represented graphically the distribution of confirmed cases and in spite of the epidemic started in Asia (Wuhan-China), nowadays America's region has overpassed Asia by 232% having the 52.2% of the total confirmed cases all over the world.

Talking about deaths is identified a DOWNWARD TREND for daily new deaths globally in Fig. 5 reaching the resistance zone by first half of April. Top countries by deaths are represented also in the picture with USA in the first place and a difference of around 67 thousand (200%) from the second place that is being occupied by Brazil. Reviewing Fig. 6, is seen an UPWARD TREND for recovered people from COVID-19 and also the table for top countries by recuperated people is represented.

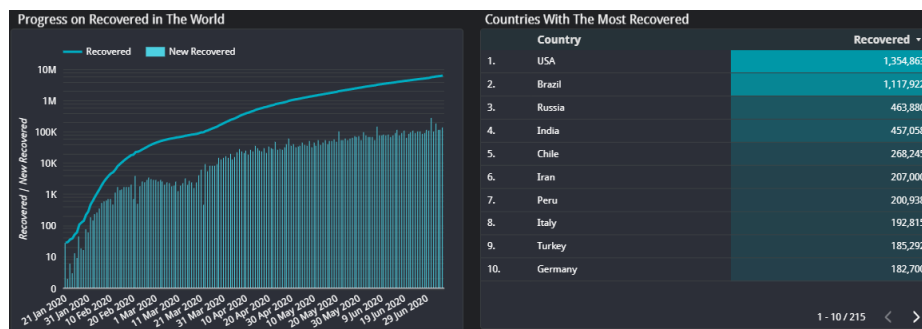
### 3.2 Mexico

On the day 132 (7<sup>th</sup> July 2020), numbers representing the pandemic in Mexico are as follows:

- 268,008-Confirmed Cases,
- ▣ 273,289-Projection (Moving Average Model, 3-day window),



**Fig. 5.** Time series chart for deaths and daily new deaths in the world. Top Countries by total deaths in the world [6].



**Fig. 6.** Time series chart for recovered people and daily new recovered cases in the world. Top Countries by total recovered people in the world [6].

- 2,331,670-Sentinel Model,
- 77,703-Suspected,
- 322,826-Negative Cases,
- 32,014-Deaths,
- 209,437-Recovered,
- 1.024-Contagion Rate,
- 1.029-Deaths Rate.

Sentinel Model was the Mexican government's strategy to monitor the behaviour of the pandemic for not identified cases in the country. It is an epidemiological surveillance program and is used in several countries to monitor diseases such as seasonal influenza. Nowadays is being taken last reported factor by the Mexican Government of 8.7x and continued the tracking and calculation on a daily basis.

About daily forecasting, it is used a Moving Average method over 3-day window, in Fig. 7 is shown the tracking over predictions vs real confirmed cases with a mean error in the last month of  $\pm 0.003\%$ .

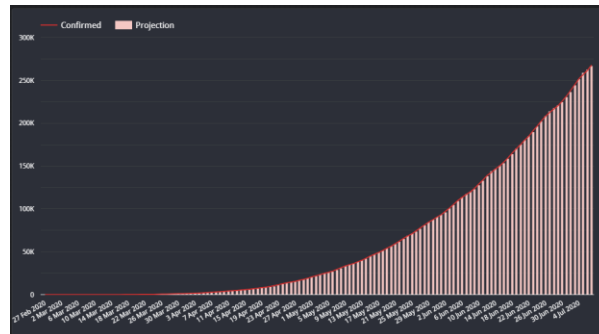


Fig. 7. Confirmed cases compared with daily predictions in Mexico.

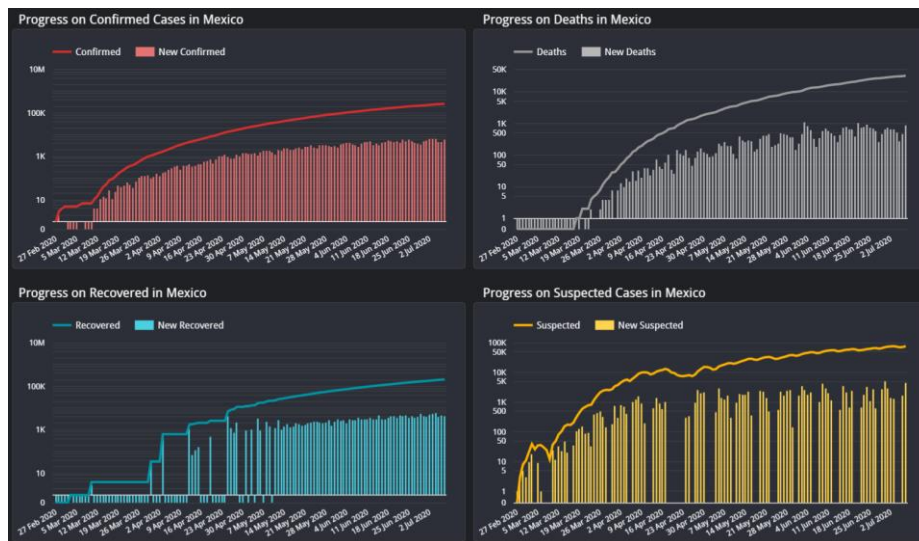
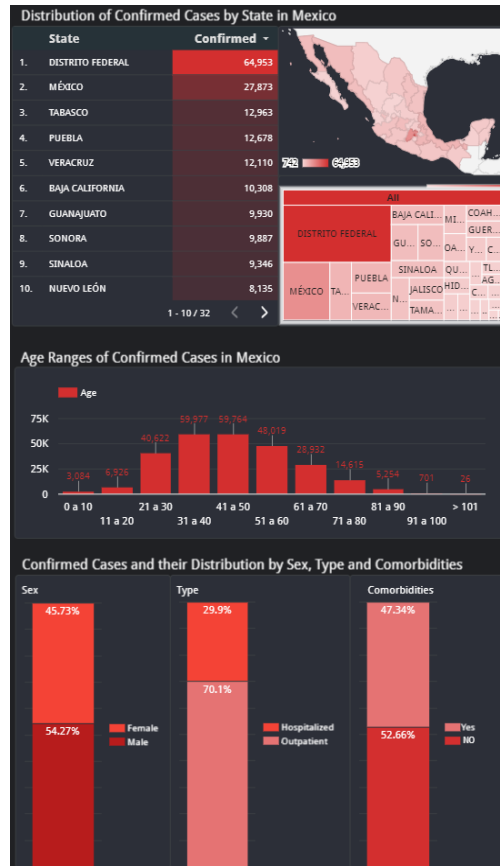


Fig. 8. Time series charts for confirmed cases and daily confirmed cases in Mexico. Time series chart for deaths and daily new deaths in Mexico. Time series chart for suspected cases and daily new suspected cases in Mexico. Time series chart for recovered cases and daily new recovered cases in Mexico [6].

In Fig. 8 are represented firstly the progress of the pandemic over the time for confirmed cases and daily new confirmed cases since 27<sup>th</sup> of February when the first case was reported can be seen in where it is identified that plateau and resistance zone have been reached for daily new cases. Secondly, the progress of the deaths over the time in Mexico and daily new deaths. Thirdly, another time series that is being tracked is the progress of suspected cases and daily new suspected people with COVID-19. Finally, the progress of the recovered cases over the time and daily new recovered people that resulted positive with COVID-19 is shown in the same figure.

As can be seen the recovered time series chart presents an anormal behaviour on the first months as the Mexican Government changed the definition of recovered cases, naming active cases just the ones that resulted positive in the last 14 days. However,



**Fig. 9.** Confirmed cases distribution by states and top states with confirmed infected people in Mexico. Confirmed cases distribution by ages in Mexico. Confirmed cases distribution by gender, type of treatment and with or without comorbidities [6].

apart from that there is not certain information on the number of recovered people in Mexico. Taking those official numbers means that 78.14% of the infected people in Mexico is already recovered which turns Mexico into one of the countries with most recovered ratio in the world, in contrast USA that is the most critical country reports only 43.57% of recovered people. Based on those numbers, it is difficult to trust on the Mexican Government numbers.

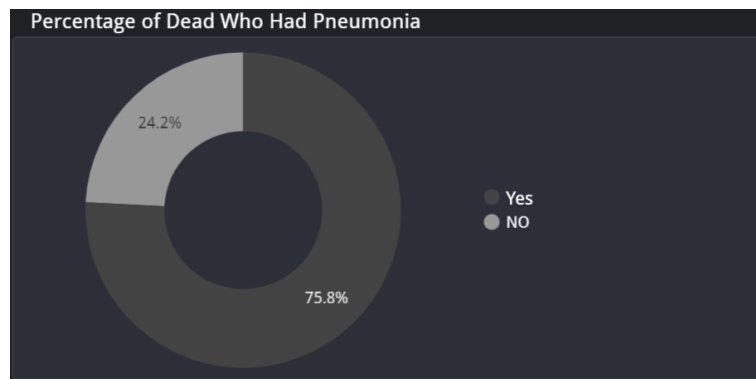
### 3.2.1 Confirmed Cases Segmentation

The actual concentration of the pandemic per State in Mexico is represented in Fig. 9. As it is seen, Mexico City is the state that on the time has had the majority of confirmed cases with 24.23% followed by Mexico State with 10.40% of the total confirmed cases in the country. Also represented in the picture, age ranges for confirmed cases has the major concentration of infected people between 31-50 years old. Also, the distribution



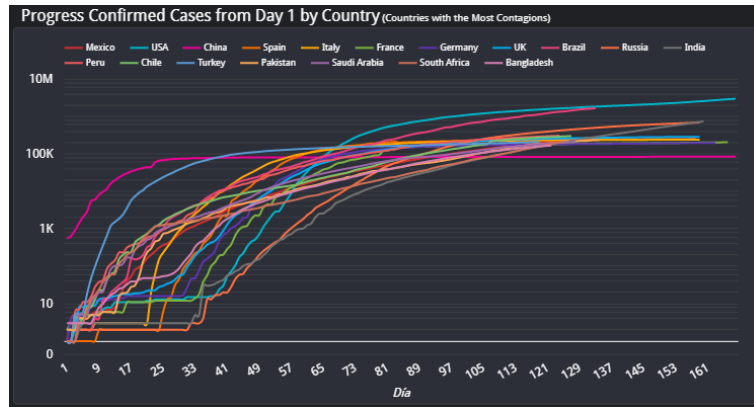


**Fig. 10.** Deaths distribution by states and top states with deaths in Mexico. Deaths distribution by ages in Mexico. Deaths distribution by gender and top comorbidities in Mexico [6].

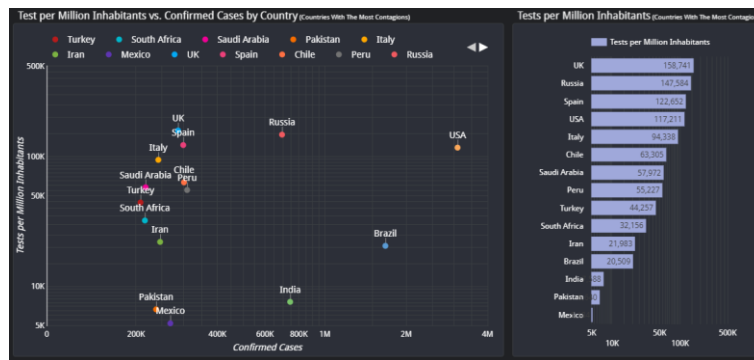


**Fig. 11.** Deaths distribution by pneumonia presence in Mexico [6].

of the infected people divided by gender, type of treatment (Hospitalized/Outpatient) and percentage of people having any comorbidity is illustrated in the same figure.



**Fig. 12.** Confirmed cases progression in Mexico and top countries with most confirmed cases taking day one in each country when the first reported case [6].



**Fig. 13.** Comparative for tests per million inhabitants vs. confirmed cases in Mexico and countries with more than 100,000 cases and tests/million. Comparative for tests per million inhabitants in Mexico and countries with more than 100,000 Cases and tests/million [6].

### 3.2.2 Death Cases Segmentation

Concentration of the deaths due to the pandemic per state in Mexico is represented in Fig. 10, being Mexico City the State that on the time has had most deaths followed by Mexico State. The age distribution for the death people due to the infection has the concentration between 51-70 years old which represents 51.14% of the total deaths in the country.

In Mexico, the majority of the death people has been males covering the 65.58% of the total and the segmentation by comorbidity can be seen also in the Fig. 10 being hypertension the one that most of deadly cases presented apart from COVID-19. However, 27.59% of the cases did not have any other chronic disease during their infection.

Another interesting statistic is that 75.8% of the deaths presented pneumonia as illustrated in Fig. 11.

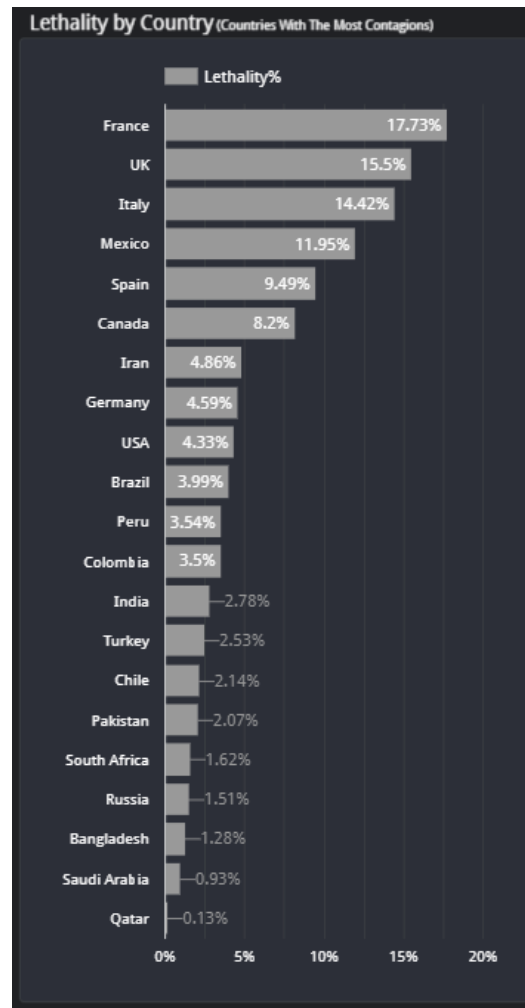


**Fig. 14.** Confirmed cases progression in Mexico and top countries with most confirmed cases. Deaths progression in Mexico and top countries with most confirmed cases. Recovered people progression in Mexico and top countries with most confirmed cases [6].

### 3.3 Mexico in Contrast with the World

As part of the developed research it is analysed confirmed cases progression in Mexico and top countries with most confirmed cases taking day one in each country as the first reported case Fig. 12. What is seen is that despite Mexico did not have same velocity for increasing the total cases, other countries reached their plateau in less days.

Globally, one of the most discussed topics has been number of test taken per millions of inhabitants and the conclusion obtained based on the illustration Fig. 13 is that there is not a strong correlation between number of test/million and confirmed cases, total confirmed cases may increase if more tests are made, however, it must be considered



**Fig. 15.** Lethality in Mexico and top countries with most confirmed cases [6].

the possibility of doing it based on the total population of each country, for example Monaco is the territory with more tests per million but with a population of only 39,227. On this regard, Mexico has the last place on tests per million of inhabitants compared with countries with more than 100,000 cases and more tests/million as represented in Fig. 13.

Another visualisation that is presented was making comparisons between Mexico and top countries with most COVID-19 contagions on confirmed, deaths and recovered cases as shown in Fig. 14.

As the final analysis for this work, in Fig. 15 it is identified Mexico as one of the countries with more lethality compared with the countries that have most confirmed cases in the world occupying the 6<sup>th</sup> place with 11.95%.

## 4 Conclusions and Future Work

Here is presented a basic mathematic model for tracking contagion rate, a simple forecasting model for daily rebalancing and most work was based on presenting visualisations on the pandemic over the time and current state in Mexico and the World. In most of the countries plateau has been reached and downward trends are being identified for confirmed cases and deaths. However, a regrowth may happen in coming days once activities are restarted.

For future research is planned to work over a longer time forecasting and countries classification depending on a multi-factor clustering, with the advanced work made here, the research path will continue taking the transformed data and improving the dashboards.

Moreover, specific work is being done to compare contagion, deaths and recovered curves on a Moving Average of 14 days to avoid peaks on data due to tracking methods in each country, on this path it will be easier to identify real trends and know over real data when a country has overpassed the pandemic or still is going upwards. Additionally, is intended to have countries classifications on different aspects and look for answers on why some territories made it better dealing with the pandemic or maybe there is hidden data in other ones explaining the reason of high lethality rates.

## 5 Conclusions

Some interesting related works that have been developed by different organizations like CONACYT [13], Scriby [12], Youyang Gu [11] and Institute for Health Metrics [10] and Evaluation, have similar work over data statistics on the pandemic and also different forecasting techniques using machine learning. These are carefully tracking data related to the spread of COVID-19 in the world encouraged by what is being seen in some areas and concerned about what is seen in others. As world move forward, is needed to continue making decisions based on the science, data and facts related to the specific conditions in our communities. Those researches are committed to providing accurate, reliable reports to the public and the information presented is updated daily and is dependent on reporting by numerous agencies across the world.

## References

1. Gobierno de México: Datos Abiertos-Dirección General de Epidemiología (2020)
2. Gobierno de México: Coronavirus (COVID-19)-Comunicado Técnico Diario (2020)
3. Johns Hopkins University: COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) (2020)
4. Worldometer: Covid-19 coronavirus pandemic (2020)
5. Martínez, J.J.: @JuanJMtzS. Retrieved from Twitter (2020)
6. Martínez, J.J.: BI-COVID19. Bi-covid19.hiketech (2020)
7. Martínez, J.J.: BI-COVID19. GitHub (2020)
8. Martínez, J.J.: Juan J. Martínez LinkedIn (2020)

*Juan J. Martínez, Ildar Batyrshin, Alexander Gelbukh*

9. Gobierno de México: Jornada Nacional de Sana Distancia (2020)
10. Institute for Health Metrics and Evaluation: COVID-19 Projections. Healthdata (2020)
11. Gu, Y.: COVID-19 Projections Using Machine Learning. Covid19-projections (2020)
12. Scriby Inc.: COVID-19 Map. Coronavirus app (2020)
13. CONACYT: Covid-19 México. Gobierno de México (2020)