Optimistic Prediction Model For the COVID-19 Coronavirus Pandemic based on the Reported Data Analysis

Khalid S. Aloufi

Department of Computer Engineering
College of Computer Science and Engineering
Taibah University
Medina

Elsevier Inc[a,b], Global Customer Service[b,]*

[a]1600 John F Kennedy Boulevard, Philadelphia
[b]360 Park Avenue South, New York

Abstract

The world is facing new challenges every day. however, with the pandemic spread over the world, a new challenge is different. The pandemic challenge is taking all the challenges to concentrate and increase in one time. Although different sectors are facing consequences, the most important sectors, health and economy are the most affected. When the pandemic start, it is not known when it will last for different health and economic planning. Therefore, it is very important for decision makers and the public to know the prediction and expectations of the future of these challenges to successfully goes over it. In this work, the current situation is analyzed. Then, an expectation model is developed based on the statistics of the pandemic using a growth rate model based on the exponential and logarithmic increase rate. Based on the available open data about the pandemic spread, the model successfully can predict the future expectations. The model expects the time and the maximum number of cases of the pandemic. The model uses the equilibrium point as the day the of cases decreases. The model can be used for planning and development of strategies to overcome the challenges.

Keywords: IR, COVID-19 Coronavirus Pandemic

1. Introduction

In different time in history, humanity have faced great health challenges that affects the life in general. Different pandemic challenges spread across the world in history, such as the Smallpox, the Spanish flu, Tuberculosis or Plague [?][?][?] [?][?] [?]. The effects of pandemic on health sector is the greatest

*Corresponding author
Email address: support@elsevier.com (Global Customer Service)
URL: www.elsevier.com (Elsevier Inc)
1

challenge and also, consequently, affects economy and social life. Human has daily activity that are highly changing during the pandemic. The daily life requirement is consequently done with caution until the pandemic end. However, when is that is only known when the pandemic is known to end. At that time, the people will be able to go back to the daily normal life activities. However, it is more convenient to expect the time when life is going to go back to normal.

In late 2019, the corona virus names COVID-19 has spread all over the world. all aspect of life has suddenly become difficult to do. Travel between countries have suddenly stops. Governments around the world are having challenges to provide the daily needed supplies of food and other consumer goods. Also, governments are having challenges in the continuity of the economy stability by paying part of the salaries of employees of the private sector. Some travelers with transit flight are living in an airport because they get in a transit flight through a country that require a visa and the continuing flight is cancelled. Some people went to the food shop and did not find their required wanted food. Therefore, planning of the requirement of people in such circumstances are very important. The pandemic requires and expectation model that is used to plan the trade activity and as well as travel activity to decrease the amount of challenges people are facing.

As well as the governments and private sectors require some expectation model so therefore the planning become reasonable and therefore can decide f the activity required. For example, a factor needs to know for how long it is expected that the factory needs to work in it low or high performance. Also, a medicine company will need to expect the production size required for a specific medicine, the range required. as well as other factors as well. The statistics of corona virus are gathered from the around the world are are announced publicity. The world for the first time is collecting the data daily, which is helping the world to work together to face the pandemic [?] [?] [?].

The power of open data is shown very well in this pandemic [?].This work gets the data of the daily statistics and develop the expectation model of the pandemic of COVID-19. Then, the prediction is required for decision makers to follow and develop plans for dealing with the pandemic. An example trajectory of cases over daily records of an example territory is shown in figure 1. The green line is the actual current cases. The possible future cases scenarios is shown in red. The figure shows different number of cases over time with different increase ratio. Also, the figure shows different maximum number of cases. Each expected scenario has different peak number of cases and different time to get to that peak. Therefore, a model is required to measure the expected peak and the number of days to get to that peak.

In this work, based on the actual statistics of each country a model is developed that can expect the number of cases any country will get to and the number of days to get to that number.

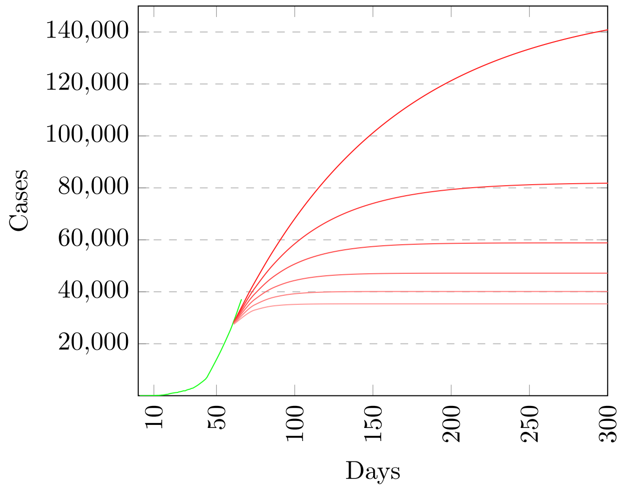

Figure 1: The Trajectory of Cases of a Territory

2. COVID-19 and Open Data

Different organizations around the world are reporting daily or even instantly reporting the data of COVID-19. The statistics of corona virus are gathered from the around the world are are announced publicity. The world for the first time is collecting the data daily, which is helping the world to work together to face the pandemic [?] [?] [?]. Open data is one of the initiatives of world wide web consortium. Open Data i the publication of data from different sectors, governments or private and make it public for the general use which increases the services provided [?]. The Open Data Institute (ODI) is one of the initiatives to support the idea of open data, started in 2012 [?]. Different governments are publishing the open data on the web. During the COVID-19 pandemic, health authorities around the world and the World Health Organization (WHO) with continuing arrangements governments and different health organization, data are collected daily or even hourly. Data sharing increase the data collection accuracy and increase the efficiency of the different parties working together to fight the pandemic.

In this work the open data published by different countries, is used and published by different organization. Also, the data are used by researcher in different majors and subjects. One of the god things about the data of the COVID-19 is that the reported data from the whole world is available on source. As mentioned earlier different sectors are collecting the data and publish it as well. One of the mentioned sources are world meter and European Centre for Disease Prevention and Control [?] [?].

3. Proposed model

In this work a model is developed using the open data published about COVID-19. After retrieving the data, the data goes through a processing for computational processing. For this reason, the data is available in a format available for direct processing [?]. The main source file is in MS XLSX format. Java is used to process the data. Different packages are available with Java to process data [?] [?]. The data is then transformed to text file, where the data can be processed to get plot using PDF-LATEX engine. other methods are possible depending on the publication required. On the collected data, the current known information about the cases in any country is the number of the daily cases and reporting day. Also, the number of death cases is reported as well as recovery number of cases.

From this information, the developed model can predict the situation of the cases in a country and the expected day of steady situation given the assumption that the cases are started to decrease based on the recovery ratios reported from around the world. Consequently, a mathematical prediction model is presented to predict the short range and long range of COVID-19 coronavirus pandemic. Then, the predicted data is saved in a text file, where the data can be processed again in the same method mentioned. The model is used to generate visualization of the data in text format where the data can then be processing and presented in any required format. The model is developed using a logistic function as a process that starts with an exponential growth then logarithmic growth defined by the function in equation 1. With the following equation the data can be processed and then generate the correlation of the data and the degree of increase or the decreases of the data. The covariance of the data and the correlation are computed using the equation 2 3. Using the 1, the predicted trajectory of the pandemic for each country or the whole world can be computed [?][?][?]. For equation 1, the parameter used are e,x_0,k,L. e is the natural logarithm base, known as Euler's number. x_0 is the x value of the sigmoid midpoint. this parameter is $SteadyDay/2$ of each country, which is computed using the equations 2 and 3. L is the curve's maximum value and this parameter is $SteadyNumber$ computed by the equation 4. k is the logistic growth rate or steepness of the curve.[1]

Algorithm 1 shows the algorithm to generate the predicted trajectory of any country. Then the algorithm is used to generate the steady date and steady number of cases for each country or the whole world. **Definition** :Let $X = (X_t : t = 1, 2, 3, ...)$ be wide-sense stationary stochastic process with mean $E[X_t]$, variance $E[(X_t - E[X_t])^2]$.

$$f(x) = L/(1 + e^{-k(x-x_0)}) \tag{1}$$

$$cov(E[X_1], E[X_2]) = \sum_{i=1}^{N}((X_{1i} - E[X_1])(X_{2i} - E[X_2])) \tag{2}$$

$$corr(E[X_1], E[X_2]) = \frac{cov(E[X_1], E[X_2])]}{E[(X - E[X_1])]E[(X - E[X_2])]}, where - 1 < corr(E[X_1], E[X_2]) < 1 \tag{3}$$

$$p(X) = E[X] + E[(X - E[X])] - (E[X] + E[(X - E[X])]) * .33 \tag{4}$$

Data: [?]

Result: The COVID-19 prediction data of a country

initialization;

get the list of reported data. the sample is the last *sampleSize* of the data in the list.

while *corr > 1* **do**

  for $i = 0;\ i < sampleSize;\ i = i + 1$ **do**

    read sample;

    set Regression;

  end

  Compute Corr using equation 3.

  Compute predicted values using equation 4

  add the predicted value to the list.

end

Algorithm 1: Prediction Data Setup Algorithms

The situation with the pandemic is either Stable, Increasing, Highly Increasing, Decreasing or Highly Decreasing. The situation is stable if either there were no cases reported or there are no more new cases. Each country is in one of the cases shown in the table 1.

Situation
Stable
Increasing
Highly Increasing
Decreasing
Highly Decreasing

Table 1: situation

The Pandemic preliminaries Data of different Countries are used in the expectation model. For each country the daily reported data are used to generate data with specifying the Country name, Daily Mean, Max in a Day, Situation, Steady Day and Steady Number. Country is the country name listed with its associated data. Daily Mean is the mean number or the average number of daily reported cases by different organizations as mentioned earlier. Max in a Day is the value of the maximum number of reported cases in a day during the data collection period. Situation, Steady Day and Steady Number.

4. Results and analysis

Figures 2,3 and 4 shows the reported daily cases of one of the countries, Saudi Arabia. The number of cases is increasing with 1452.14 mean number of daily reported cases. The first case was reported in 03-03-2020. In the first plot the daily total reported cases are shown in blue line. Th red line shows the expected trajectory of the number of cases and the time it will get to that time. The second plot shows the reported cases on daily bases. The most left plot shows the daily reported death cases. Based on each country reported data the model is predicting different trajectory. The model is predicting the country future cases based on the trajectory of cases and data published.

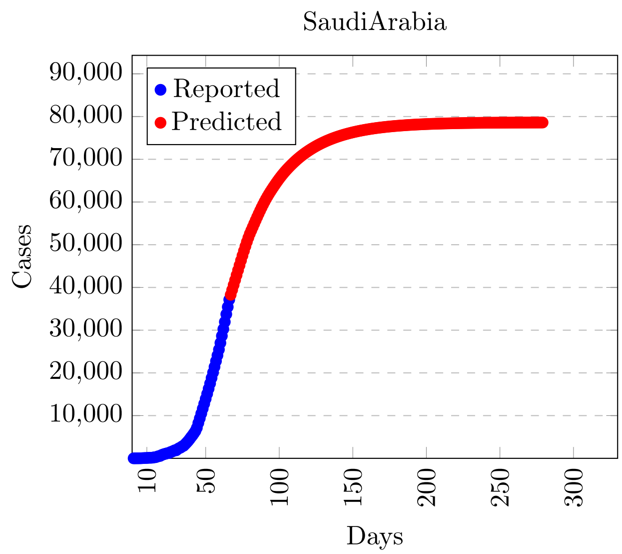

Figure 2: SaudiArabia Reported and Predicted

For the COVID-19 in the world, the number of cases is decreasing with 79314.57 mean number of daily reported cases. The first case was reported in 31-12-2019. Figure 8 shows the trajectory of cases and the expectation of the maximum number the cases will reach. The expectation is that it will get to 6000000 in about 250 days since it starts. However, the pandemic will exist for as long as 400 days according to the expected model. Figure 9 shows the daily reported cases from around the world. The figure shows that the range of daily cases is between 60000 and 100000.

Figure 7 shows the plots of all countries from around the world. The blue line shows the reported data. Each country has different reporting date since the pandemic is expected to reach countries in different dates. The red line shows the expected cases growth. Figure 5 shows the maximum and mean number of cases reported per day in increasing order. Figure 6 shows the steady day expected of each country from the day the pandemic was first reported.

The model is used to simulate the pandemic trajectory using the equation presented earlier 4. With

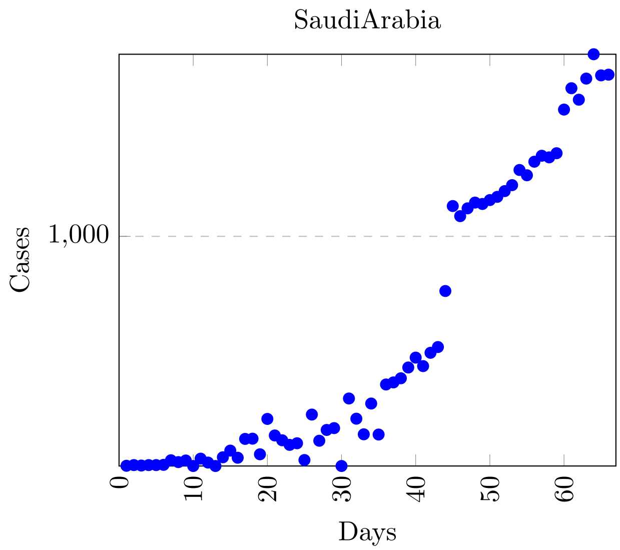

Figure 3: SaudiArabia Daily cases

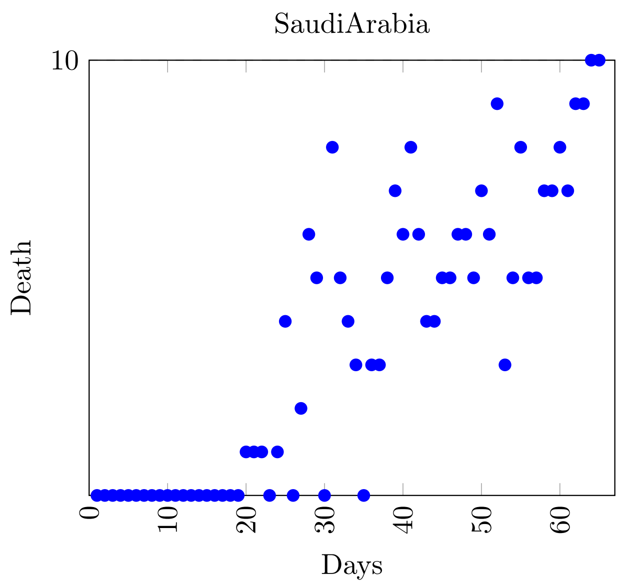

Figure 4: SaudiArabia Redported Death cases

the right parameters the model can expect the growth rate of cases. For example, for a country where the expected number of cases will reach 50000 and the steady day midpoint is 100, the trajectory is shown in figure 10. Figure 12 shows the initial dates of countries first reporting about the pandemic. Most countries are reporting between the 29-2-2020 and 9-4-2020, which is about three months after the pandemic has started. Figure 13 shows the percentage of countries in either decreasing or increasing number of cases according to the reported statistics. Figure 11 shows the pandemic expected growth rate based on the expected maximum

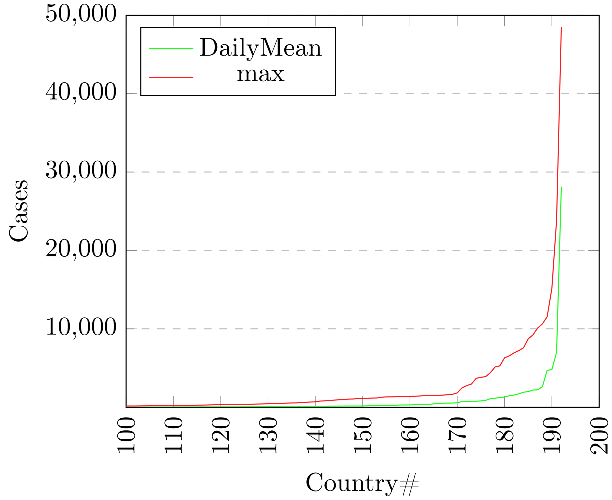

Figure 5: Maximum and Mean number of cases for countres

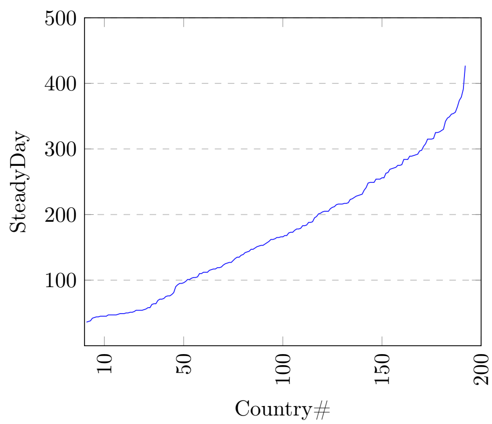

Figure 6: Steady Day of countries

number of 6 million cases and the midpoint is in 150 days.

5. Recomendaions

This research recommends in using the expectation model presented to be used in the planning associated with the pandemic. The Model can be used to predict the expected number of cases and in how many days. The model can be used to expect the maximum and the minimum number of cases and in how many days. The model can be used as optimistic model as well as pessimistic model. Using the model as an expectation model in both scenarios is very important. The model can be used to know the domain and range of expectation. The domain is the pandemic and its consequences and associated challenges with economy, education and social activities. The range of the time and places the pandemic will get to and for how long.

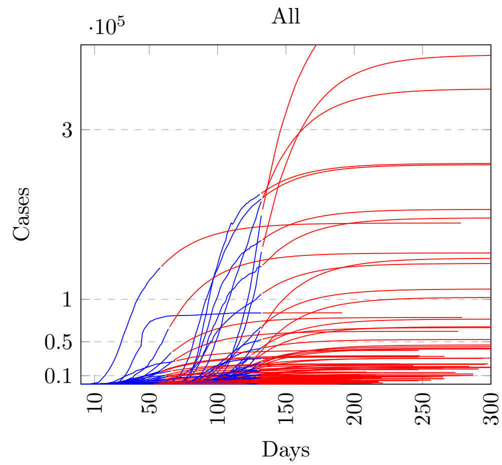

Figure 7: All countries starting reported and predicted cases

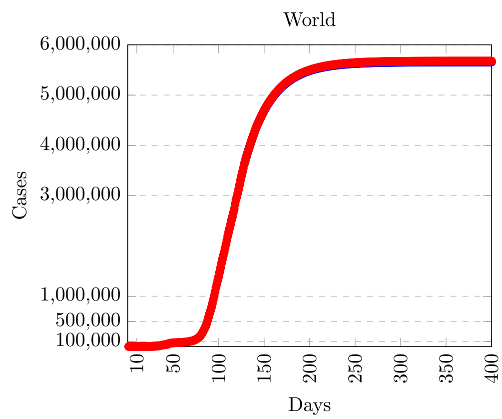

Figure 8: World Cases Reported and Expected

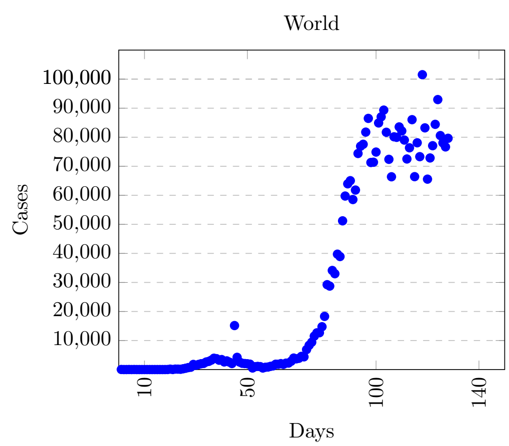

Figure 9: SteadyDay

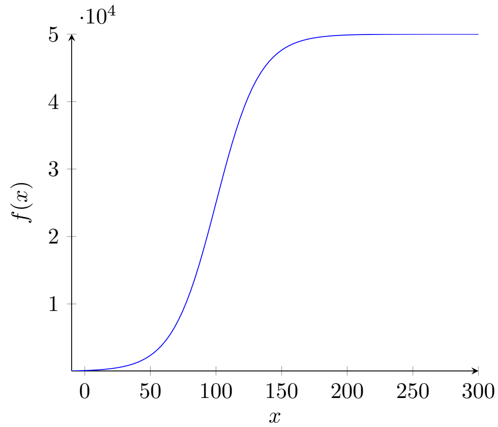

Figure 10: Growth Rate of 50000 cases

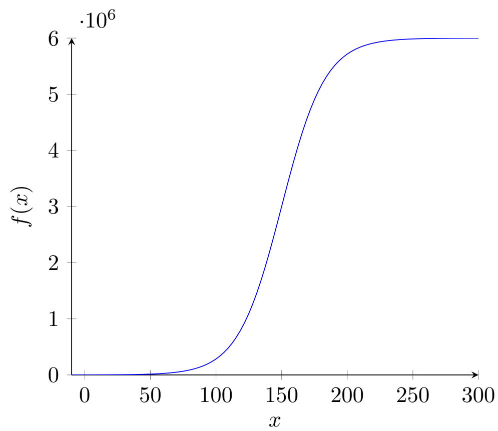

Figure 11: Growth Rate of 6000000 cases

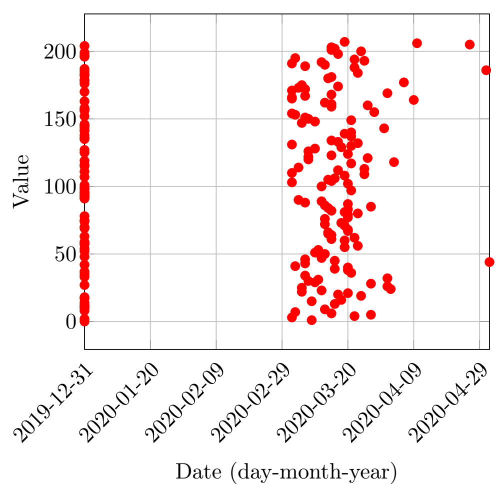

Figure 12: Initial Reporting Date for countries

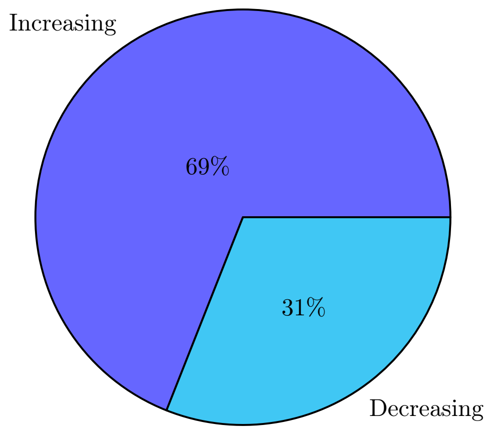

Figure 13: Increasing and Decreasing %

Future work should consider analyzing the data after the pandemic finish and develop mathematical models that help in fighting such pandemic in earlier time in more advanced regulations.

6. Conclusion

In this work, the data published about the pandemic has shown a proof of the power of data and open data. The date can be used and processed to show different effects of the pandemic. In this work the data is processed and analyzed. Then, the data is analyzed to show the percentage of countries with decreasing or increasing number of cases, Using the mathematical computation of mean, variance, covariance and correlation of the data, the situation of the pandemic can be known in different countries. Then, the data is used to know the scenario expected. The data has shown that the growth rate scenario is the best mathematical model to represent the data. After that, the model is developed. The model can successfully represent the pandemic trajectory in a territory.

References

[1] Wikipedia contributors, Smallpox — Wikipedia, the free encyclopedia, `https://en.wikipedia.org/w/index.php?title=Smallpox&oldid=955501571`, [Online; accessed 9-May-2020] (2020).

[2] Wikipedia contributors, Tuberculosis — Wikipedia, the free encyclopedia, `https://en.wikipedia.org/w/index.php?title=Tuberculosis&oldid=955501145`, [Online; accessed 9-May-2020] (2020).

[3] Wikipedia contributors, Plague (disease) — Wikipedia, the free encyclopedia, `https://en.wikipedia.org/w/index.php?title=Plague_(disease)&oldid=952549003`, [Online; accessed 9-May-2020] (2020).

[4] Wikipedia contributors, Spanish flu — Wikipedia, the free encyclopedia, https://en.wikipedia.org/w/index.php?title=Spanish_flu&oldid=955737989, [Online; accessed 9-May-2020] (2020).

[5] Wikipedia contributors, Coronavirus disease 2019 — Wikipedia, the free encyclopedia, https://en.wikipedia.org/w/index.php?title=Coronavirus_disease_2019&oldid=955735118, [Online; accessed 9-May-2020] (2020).

[6] Wikipedia contributors, Pandemic — Wikipedia, the free encyclopedia, [Online; accessed 9-May-2020] (2020).
URL https://en.wikipedia.org/w/index.php?title=Pandemic&oldid=955705211

[7] Worldometers.info, Covid-19 coronavirus pandemic, https://www.worldometers.info/coronavirus/, place of publication: Dover, Delaware, U.S.A. (5 May, 2020).

[8] E. O.-O. Max Roser, Hannah Ritchie, J. Hasell, Coronavirus pandemic (covid-19), Our World in DataHttps://ourworldindata.org/coronavirus.

[9] E. C. for Disease Prevention, Control, the geographic distribution of covid-19 cases worldwide, https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide (4 May, 2020).

[10] Wikipedia contributors, Open data — Wikipedia, the free encyclopedia, https://en.wikipedia.org/w/index.php?title=Open_data&oldid=953018662, [Online; accessed 10-May-2020] (2020).

[11] Wikipedia contributors, Open data institute — Wikipedia, the free encyclopedia, https://en.wikipedia.org/w/index.php?title=Open_Data_Institute&oldid=953330894, [Online; accessed 10-May-2020] (2020).

[12] K. Aloufi, Covid-19, Mendeley Data, v1doi:10.17632/zk32frw6p5.1.

[13] B. D. Hahn, K. M. Malan, Acknowledgements, in: B. D. Hahn, K. M. Malan (Eds.), Essential Java for Scientists and Engineers, Butterworth-Heinemann, Oxford, 2002, p. xvi. doi:https://doi.org/10.1016/B978-075065991-8/50001-0.
URL http://www.sciencedirect.com/science/article/pii/B9780750659918500010

[14] Preface, in: B. D. Hahn, K. M. Malan (Eds.), Essential Java for Scientists and Engineers, Butterworth-Heinemann, Oxford, 2002, pp. xiii – xv. doi:https://doi.org/10.1016/B978-075065991-8/50000-9.
URL http://www.sciencedirect.com/science/article/pii/B9780750659918500009

[15] Wikipedia contributors, Logistic function — Wikipedia, the free encyclopedia, https://en.wikipedia.org/w/index.php?title=Logistic_function&oldid=955315778, [Online; accessed 10-May-2020] (2020).

[16] J. Peng, K. Lee, G. Ingersoll, An introduction to logistic regression analysis and reporting, Journal of Educational Research - J EDUC RES 96 (2002) 3–14. doi:10.1080/00220670209598786.

[17] G. Rzadkowski, I. Głażewska, K. Sawińska, Logistic function as a tool of planning, Foundations of Management 6. doi:10.1515/fman-2015-0004.