MASTER THESIS

In Order to Obtain the

PROFESSIONAL MASTER

in

Geographic Information Systems (GIS)

And Data Science

Presented and defended by:

Rassil Ghassan Salloum

On Tuesday, November 2, 2021

Title:

Covid 19 Predictions For Herd Immunity in Lebanon

Supervisors:

Mrs. Gretta Antoine Kelzi

Reviewers:

Dr. Kifah Raafat Tout

Mrs. Manal Zuheir Sayed

Contents

ACKNOWLEDGEMENTS

First and foremost, I would like to praise and thank God, the almighty, who has granted countless blessing, knowledge, and opportunity to me, so that I have been finally able to accomplish this thesis.

Apart from the efforts of me, the success of this thesis depends largely on the encouragement and guidelines of many others. I take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this thesis.

The internship opportunity I had with Esri Lebanon and Khatib & Alami Company was a great chance for learning and professional development. I would like to show my greatest appreciation to Pr. Kifah Tout for giving me this opportunity and for his support throughout the year.

Special thanks to my supervisor Mrs. Gretta Kelzi , to Ms. Nour Nasser and to Mr. Tarek Hajar for all their help and advice with this master, it was a pleasure to me working with them all.

And of course I would like to thanks all the doctors for all the effort and knowledge the gave us during this year despite the difficult conditions the country is facing.

Finally, I must express my very profound gratitude to my parents and to my best friend Hanine Serhan for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

ABSTRACT

Do people want to be vaccinated against COVID-19? Herd immunity is dependent on individuals' willingness to be vaccinated since vaccination is not mandatory.

Our main goal is the study of herd immunity and predict when to achieve it. Herd immunity occurs when a large portion of a community (about 90%) becomes immune to a disease, making the spread of disease from person to person unlikely. As a result, the whole community becomes protected. This prediction is related to the study of vaccination over Lebanon, that's why we have to know all information about vaccination. From here we have implemented a clear dashboard using ArcGIS Dashboard that contains a web map presenting the vaccine process in each district in Lebanon, charts includes the total number of vaccine delivered over Lebanon and in each district, the total number of vaccine delivered by gender, nationality, vaccine type, age, etc.

In addition of the mentioned information above, the dashboard contains the predicted month for herd immunity, this last is reached by using a machine learning model (linear regression) applied on the data related to the cumulative percentage of vaccine from the first vaccination day. So the model predict when we will reach 90% of vaccinated people.

Unfortunately, despite the importance of achieving herd immunity, there are many people who do not know what "herd immunity" means. For this reason we have developed a story map (web app) that let the people know what is herd immunity and how we can achieve it , encourage people to get the vaccine and focus on its importance in our context.

Keywords: COVID-19, herd immunity, vaccination, dashboard, ArcGIS, machine learning model, linear regression, data, story map, web app.

Chapter I: INTRODUCTION

The COVID-19 pandemic is considered as the most crucial global health calamity of the century and the greatest challenge that the humankind faced since the 2nd World War. It crisis comes on the heels of grave economic and political crises that have hit the world wide at the end of 2019, and it led to huge losses at various levels. (Chen, August 2020) It has rapidly spread around the world, posing enormous health, economic, environmental and social challenges to the entire human population.

How can we limit the spread of this virus?

A. Background

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has become a worldwide pandemic since it emerged in December 2019. The COVID-19 has resulted in unprecedented social and economic disruptions, with a profound worldwide impact on public health, lifestyle, and food security.

After the discovery of a vaccine for this virus, all studies in the world became about the vaccine and how to achieve herd immunity.

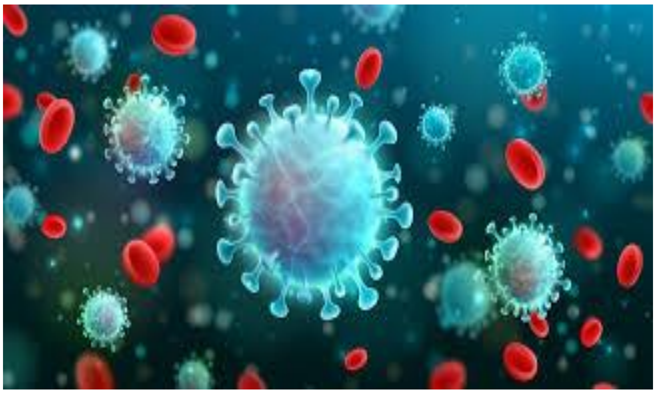


Figure 1:SARS-CoV-2

Comprehensive vaccination coverage is one of the most effective ways to stop a pandemic and reopen communities.

In the months following the emergence of SARS-CoV-2, “herd immunity” was frequently cited as the long-term destination of the COVID-19 pandemic. (Gypsyamber D'Souza, 13 September 2021)

Till this day, more than 5 million deaths worldwide due to this virus, and about 8 490 deaths in Lebanon. This number is not normal, and we have to stop the spread of the virus by achieving herd immunity.

As vaccination has rolled out, variants have emerged, and as cases surge once again, we are learning more about the nuances of SARS-CoV-2 infection and what short- and long-term immunity to this virus may look like.

With this changing perspective, how should we be thinking about herd immunity?

B. Overview

Herd immunity occurs when a large portion of a community (about 90%) becomes immune to a disease, making the spread of disease from person to person unlikely. As a result, the whole community becomes protected. (Gu, 2020)

When most of a population is immune to an infectious disease, this provides indirect protection—also called population immunity, herd immunity, or herd protection—to those who are not immune to the disease. It even protects those who cannot be vaccinated, like newborns or people who are allergic to a vaccine. The percentage of people who need to have protection to achieve population immunity varies by disease.

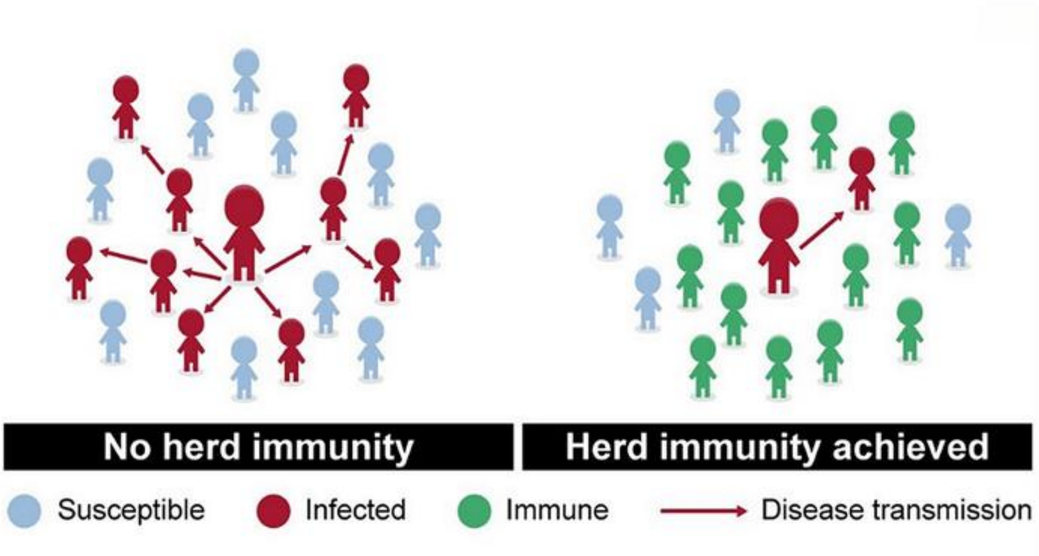


Figure 2: Graphic explaining herd immunity, where immunity to a disease in a population — whether through natural infection or preferably, by widespread vaccination — reaches a high-enough percentage to stem the spread of an outbreak. (Newman, 25, August, 2020)

For example, currently over 90% of all children in the U.S. are vaccinated against measles, mumps, and rubella by their second birthday. This level of vaccination provides protection to the population as a whole—even to those who aren't vaccinated—by decreasing viral circulation and the chance someone who is unvaccinated will encounter the virus. If a person with measles were to come to the U.S., for example, nine out every 10 people that person could infect would be immune, making it very hard for measles to spread in the population. As a result, even though we still see localized outbreaks of measles in the U.S., those outbreaks generally die down without starting a nationwide epidemic.

The same idea works for any infectious agent, including coronavirus. The hope is that the population can develop a high enough level of immunity to keep spread low.

The more contagious an infection is, the higher the proportion of the population that needs immunity before infection rates start to decline. But this percentage isn't a "magic threshold" that we need to cross—and it's not just dependent on the level of population immunity. Both viral evolution and changes in how people interact with each other can bring this number up or down as well. But even below any "herd immunity threshold," immunity in the population (for example, from vaccination) can still have a positive effect by reducing the total number of infections that happen.

Therefore, people most never hesitate to take the vaccine, because they will save people who are unable to receive it and will speed up the process of eliminating the virus.
To facilitate the process of herd immunity, we can start by studying it at the level of regions, districts or governorates leading to its achievement at the level of the entire country.

This thesis provide prediction for the date in which herd immunity will be achieve in Lebanon. From computer science background diving into GIS and Data Science world, this prediction was an application of machine learning assets focusing in patterns inside using data. From GIS perspective an important data analysis was applied on special data available in the collected data in order to bring location data into our analysis and mapping each vaccinated case to its district since as mentioned before our study was applied at the level of Lebanese district . (ESRI)

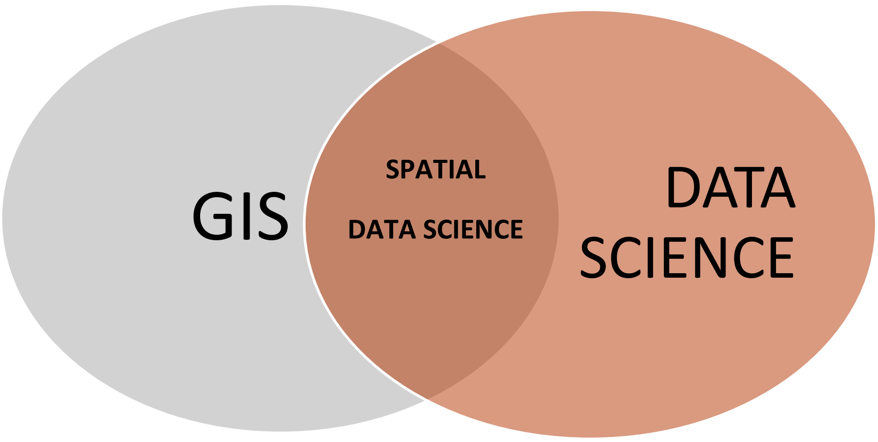


Figure 3:Gis+Data Science

Chapter II: Related Work

- **Path to normality – covid-19 vaccine projection By: Youyang Gu**
  With the availability of the COVID-19 vaccine, this study estimate of the path to COVID-19 herd immunity / normality in the United States. Immunity against the SARS-CoV-2 virus comes from two sources: vaccination and natural infection. On this page, they provide the latest COVID-19 vaccine projections and current vaccination progress using artificial intelligence to accurately forecast infection, deaths, and recovery timelines of the covid 19/ coronavirus pandemic in the Us and globally. (Gu, 2020)
- **Rethinking Herd Immunity and the Covid-19 Response End Game By Dr. Makary**
  A article from Dr. Makary's university described herd immunity thusly: "When most of a population is immune to an infectious disease, this provides indirect protection—or population immunity (also called herd immunity or herd protection)—to those who are not immune to the disease". The article gave examples of measles, mumps, polio, and chickenpox as diseases for which herd immunity exists in the US, while noting that outbreaks can still occur in pockets of unvaccinated people, as occurred with measles during an outbreak in Disney World in 2019. (Gypsyamber D'Souza, 13 September 2021)

Below we present a list of links to helpful vaccination data sources and dashboards. While we do not directly use them in our modeling, we highly recommend giving them a look.

Vaccine Rollout Tracker

- Bloomberg
- Our World in Data
- New York Times

Vaccine Development Tracker

- New York Times
- Milken Institute

Chapter III: MATERIALS AND METHODS

As we mentioned before , due to the COVID-19 pandemic we have to do an effort to reduce this disaster and predict when we will reach the herd immunity.
As a brief summary, our scenario is to use machine learning to predict herd immunity and bring location data into our analysis using ArcGIS, to show the distribution of vaccination over Lebanon and knowing which area is the first to achieve herd immunity. In addition to present all data related to this study in a dashboard to make everything clear.
And finally we create a web application (story map) that let people knows about herd immunity and focus on the importance of taking the vaccine to achieve it.
In this chapter we address the different steps taken to perform the creation of this studies by detailing how we have implement this and which software are used.

A. DATA PREPARATION

Data preparation is the first step in data analytics projects and can include many discrete tasks such as determine what information you want, determine the data collection method, loading data or data ingestion, data fusion, data cleaning, etc. (lotame, 2019)

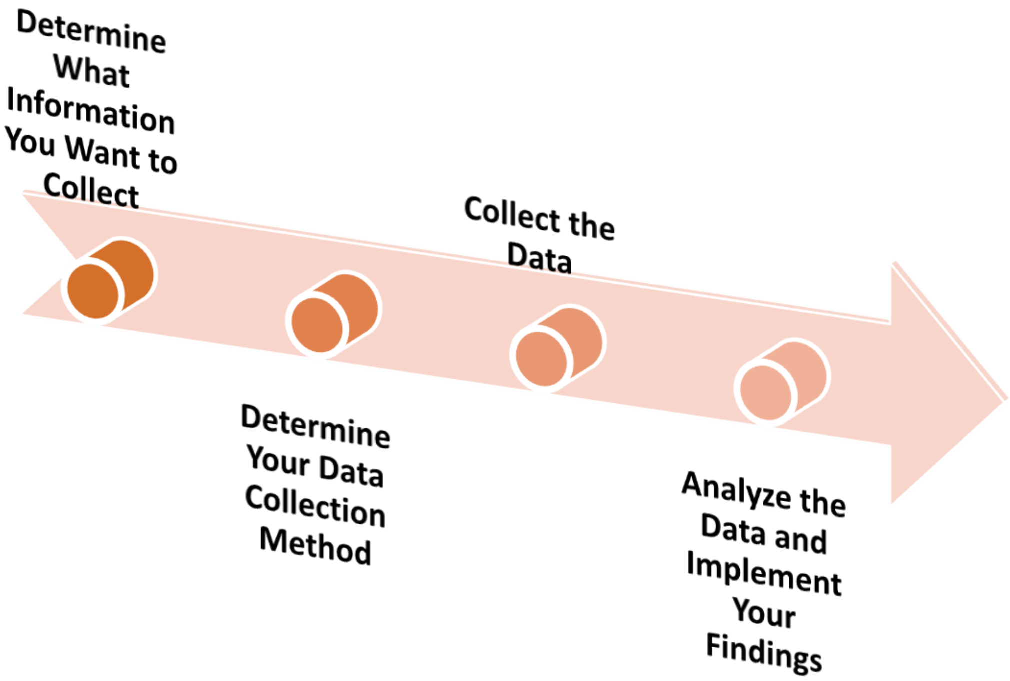


Figure 4: Data preparation process

∇ Determine What Information You Want to Collect:

The first thing we need to do is choose what details we want to collect. we'll need to decide what topics the information will cover, from where and how much data we need. Our project will determine the answers to these questions.

In our project the data must be real and precise, for this reason the data was collected from MOPH (Ministry of public health) in Lebanon:

1. To do the prediction part related to when achieve herd immunity we need to have:
    - The daily number of people vaccinated to predict from them the date we will achieve 90% of vaccine (herd immunity) using machine learning model.
2. To implement the dashboard we need:
    - All data related to vaccination (location, gender, age, nationality, vaccine type, vaccine date)

∇ Determine Your Data Collection Method and collect data:

At this step, we will choose the data collection method that will make up the core of your data-gathering strategy. To select the right collection method, you'll need to consider the type of information you want to collect.

Because there is a large amount of information and it must be accurate the only way to get the data related to the vaccination is : MOPH (Ministry of public health)

Once we have finalized the collection of data, we have to store and organize this data.

∇ Analyze the Data and Implement Your Findings:

Once we've all data, it's time to analyze it and organize our findings. The analysis phase is crucial because it turns raw data into valuable insights that you can use into the model or for the GIS .

B. PREDICTION MODEL

A part of our project is related to machine learning model prediction so what is machine learning and how we can use it in our project?

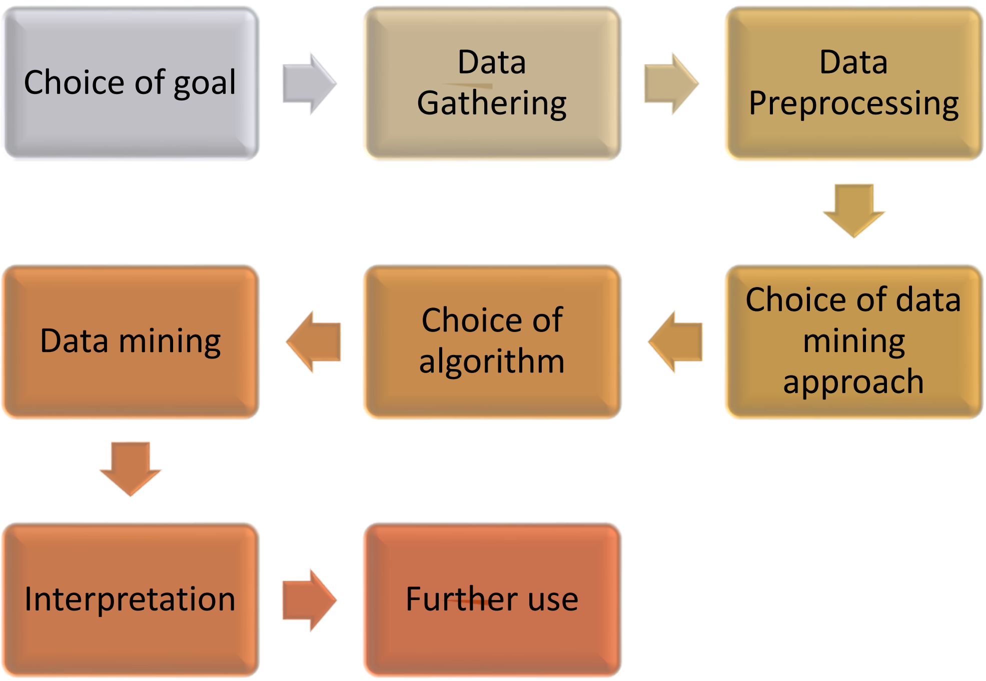


Figure 5: Adapted knowledge discovery process, containing the taken steps from gathering the data to interpreting the results and using them for further tasks.

i. Overview

We talked about the section related to the preparation of data before. Now we should talk about :

- Choice of data mining approach
- Choice of algorithm
- Data mining

1. Data Mining

Data mining is often used as a synonym for the KDD. In fact, data mining represents the part of the KDD-process, in which the appropriate approach and algorithm is chosen and applied to the data set. Therefore, it is a central part of the knowledge discovery in databases process. Data mining consists of taking any form of data and applying analysis algorithms to it in order to reveal patterns or models within the data set and using these structures to classify the data into different classes (labels).

It comprises a number of research fields such as database systems, statistics and pattern recognition. Data mining tasks are differentiated depending on the knowledge the algorithm has about the existing classes in the data set.

- **Supervised learning** includes every task in which the algorithm has access to input and output values. Input values are defined as the external information that the algorithm is allowed to use, such as attribute values and meta data, while output values, are the specific labels of the class attribute. This means that the structure of the data is already known and the goal of these programs is to assign new data to the correct classes.
- In contrast to supervised learning, **unsupervised learning** includes all tasks that have no access to output values and therefore try to find structures within the data by creating classes on their own.

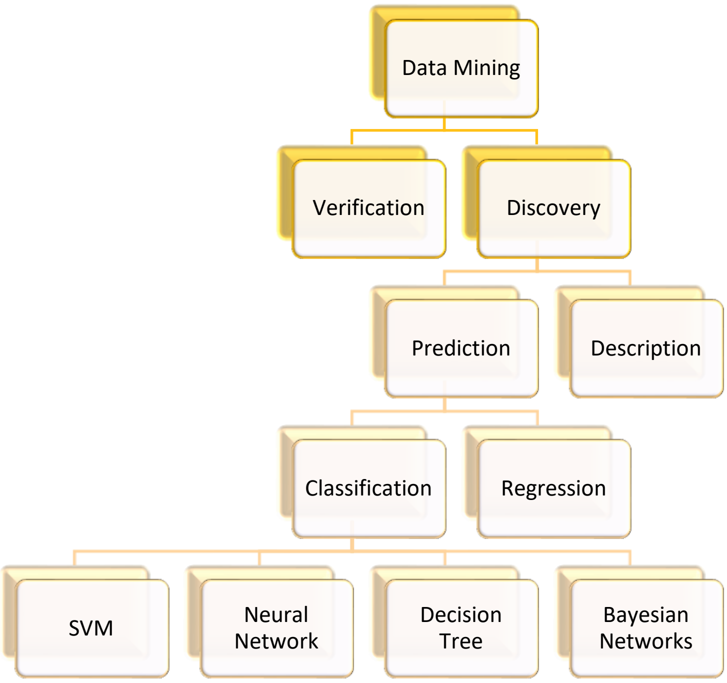


Figure 6: Data mining taxonomy based on

Since the proposed task is a prediction problem, which it started with a dataset with a known dependent variable, train our model (with independent variables) then apply it later to predict real number so this work will focus on **supervised** learning methods.

Furthermore, data mining can be differentiated into two main objectives: verification and discovery. While verification tries to prove the user's hypothesis, discovery looks for yet unknown patterns within the data. The discovery step splits up into description, where the system finds patterns in order to present the data in an understandable format and **prediction**, where the system tries to predict the future outcomes of data from patterns. The subgroup prediction can further be distinguished into classification and regression tasks. While classification tasks have fixed labels and each data record has one of these labels as its class attribute value, regression tasks have continuous values as output . This work focuses on algorithms that try to predict when to achieve herd immunity. Therefore, the problem consists of continuous data as output (Number of people vaccinated) and belongs to the discovery-prediction sector of data mining.

So we deduce that our case is related to regression model.

2. Machine Learning

In context of the data mining step, it is important to choose the correct approach for tackling the task appropriately. This is often done using machine learning methods.

A major difference between humans and computers has been for a long time that a human beings tend to automatically improve their way of tackling a problem. Humans learn from previous mistakes and try to solve them by correcting them or looking for new approaches to address the problem.

Traditional computer programs do not look at the outcome of their tasks and are therefore unable to improve their behavior.

The field of machine learning addresses this exact problem and involves the creation of computer programs that are able to learn and therefore improve their performances by gathering more data and experience.

The first scientist to create a self-learning program was A. Samuel in 1952, who created a program that became better at playing the game checkers with the number of games played .

In 1967, the first pattern recognition program was able to detect patterns in data by comparing new data to known data and finding similarities between them. Since the 1990's machine learning is used in data mining areas, adaptive software systems as well as text and language learning fields.

As an example: A computer program that gathers data concerning the customers of an e-commerce shop and creates better personalized advertisements out of these pieces of information has the ability to acquire new knowledge and comes close to being artificial intelligence.

Furthermore, machine learning systems are normally classified by their underlying learning strategies, which are often identified by the amount of inference the computer program is able to perform :

- **Rote Learning** describes the strategy that all traditional computer programs use. They do not perform any kind of inference and all their knowledge has to be directly implemented by the programmer, since the application is not able to draw any conclusions or transformations from the given information.

- **Learning from Instruction** encompasses all computer programs that are able to transform information from the given input language to an internal language. Although the knowledge on how to effectively perform this transformation is still given by the programmer, this requires little forms of inference from the side of the computer program. Therefore, this defines a separate level of learning system compared to rote learning.

- In contrast to Learning from Instruction, **Learning by Analogy** tries to develop new skills that are almost similar to existing skills and therefore easy to adopt, by performing transformations on known information. This system requires the ability of creating mutations and combinations of a dynamic knowledge set. It creates new functionalities, which were unknown to the original computer program and therefore requires a lot of inference.

- **Learning from Examples** is nowadays one of the most commonly used learning strategies as it provides the most flexibility and enables computer programs to develop completely unknown skills or find unknown structures and patterns in data . Learning from examples is a technique that is often used in regression and data mining tasks to predict the independent variable of new data entries. In this work, the proposed study questions will be tackled with strategies and algorithms that belong to this category.

Following, the most common machine learning systems will be briefly described:

3. Regression

A regression model is used to **investigate the relationship between two or more variables and estimate one variable based on the others**. Its divided into linear and non linear regression.

In our project we will use linear regression .

Linear Regression is a machine learning algorithm based on **supervised learning**. It performs a **regression task**. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on the kind of relationship between dependent and independent variables, they are considering and the number of independent variables being used. (My accounting course, 2016)

Linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x). So, this regression technique finds out a linear relationship between x (input) and y(output). Hence, the name is Linear Regression.

The regression line is the best fit line for our model.

While training the model we are given :

x: input training data (univariate – one input variable(parameter))

y: labels to data (supervised learning)

When training the model – it fits the best line to predict the value of y for a given value of x.

The model gets the best regression fit line by finding the best θ_1 and θ_2 values.

θ_1: intercept

θ_2: coefficient of x

Once we find the best θ_1 and θ_2 values, we get the best fit line. So when we are finally using our model for prediction, it will predict the value of y for the input value of x.

Software used: Anaconda Navigator/ Jupyter Lab

ii. Implementation

After we chose the data mining approach and the algorithm we have to apply it using **python** language.

So as we mentioned above, linear regression performs the task to predict a dependent variable value (y) based on a given independent variable (x).

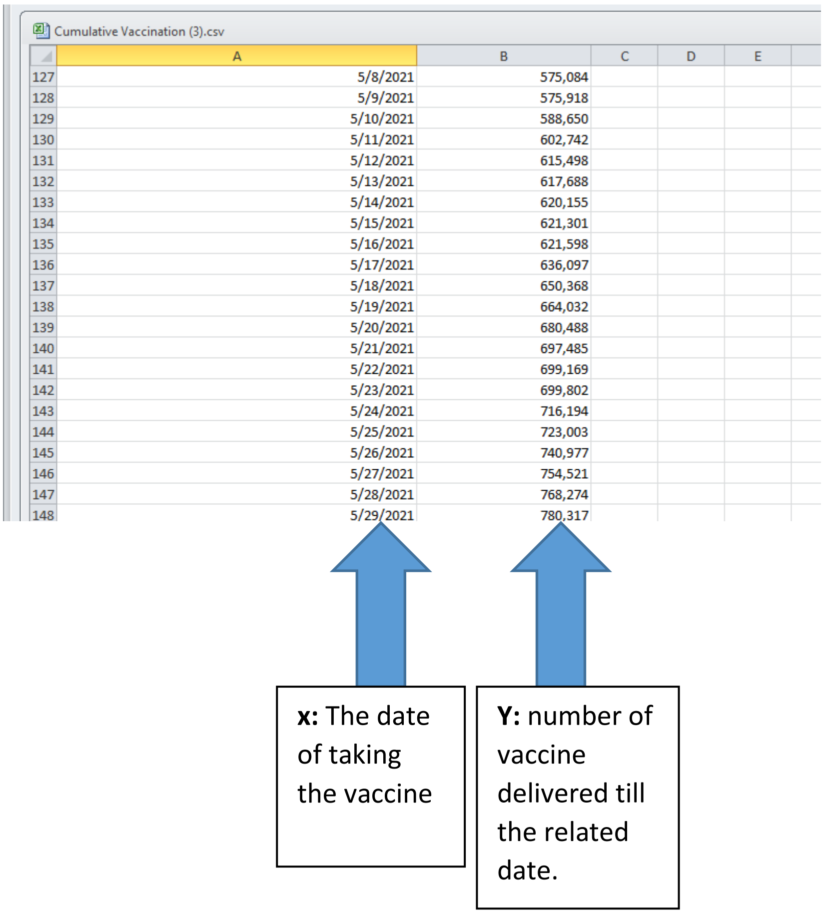
Our study is to predict when we reach 90% of people vaccinated, so we have to know the cumulative number of vaccine delivered and predict the future number to deduce if we have achieve this percentage.

While training the model we are given :

x: The date of taking the vaccine

y: cumulative number of vaccine

Figure 7:Tables of data used for the model



Cumulative Vaccination (3).csv

	A	B	C	D	E
127	5/8/2021	575,084			
128	5/9/2021	575,918			
129	5/10/2021	588,650			
130	5/11/2021	602,742			
131	5/12/2021	615,498			
132	5/13/2021	617,688			
133	5/14/2021	620,155			
134	5/15/2021	621,301			
135	5/16/2021	621,598			
136	5/17/2021	636,097			
137	5/18/2021	650,368			
138	5/19/2021	664,032			
139	5/20/2021	680,488			
140	5/21/2021	697,485			
141	5/22/2021	699,169			
142	5/23/2021	699,802			
143	5/24/2021	716,194			
144	5/25/2021	723,003			
145	5/26/2021	740,977			
146	5/27/2021	754,521			
147	5/28/2021	768,274			
148	5/29/2021	780,317			

Steps of applying the Linear Regression Model:

- **The Data Set Used:**

  X: The date of taking the vaccine

  y**:** number of vaccine delivered till the related date.

- **The Libraries We Will Use in This Tutorial**

  The first library that we need to import is pandas, which is a portmanteau of "panel data" and is the most popular Python library for working with tabular data.

  Next, we'll need to import Numpy, which is a popular library for numerical computing. Numpy is known for its Numpy Array data structure as well as its useful methods reshape, arrange, and append .

  Next we need to import matplotlib, which is Python's most popular library for data visualization.

- **Importing the Data Set**

  We import the data needed (x and y)

- **Splitting our Data Set into Training Data and Test Data**

  Scikit-learn makes it very easy to divide our data set into training data and test data in a predefined function.

- **Building and Training the Model**

  The first thing we need to do is import the LinearRegression estimator from scikit-learn.

  Next, we need to create an instance of the Linear Regression Python object to use it to fit and predict

- **Making Predictions From Our Model**

Making prediction and compare the predicted and the original data.

- Making Predictions From Our Model

```
[22]: predictions = model.predict(x_test.reshape(-1, 1))
      plt.scatter(y_test, predictions)
```

```
[22]: <matplotlib.collections.PathCollection at 0x24ada013b00>
```

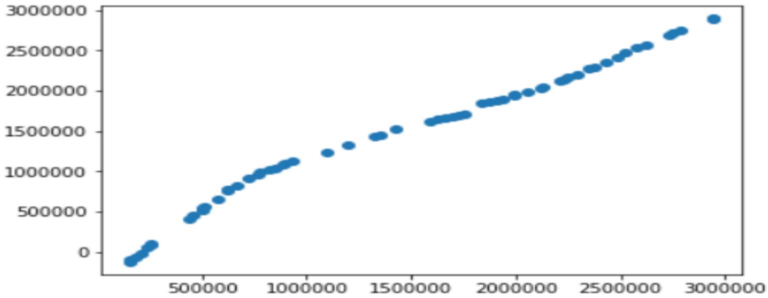


Figure 8: Comparion between predicted and original data

```
[29]: predictions2 = model.predict(x_data_new.reshape(-1, 1))
      plt.plot(x_data_new, y_data_new, 'ro', label='data')
      plt.plot(x_data_new,predictions2, linewidth=3.0, label='fit')
      plt.legend(loc='best')
      plt.ylabel('Nb_Vaccinated')
      plt.xlabel('Day')
      plt.show()
```

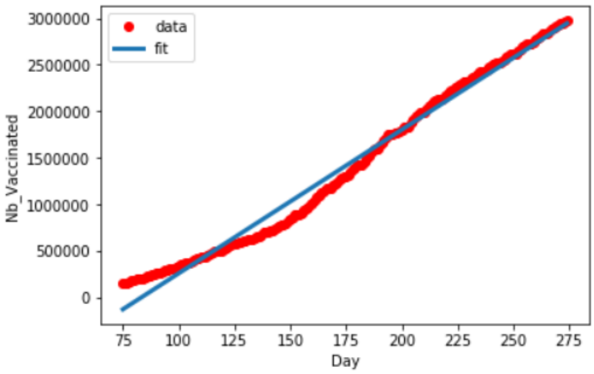


Figure 9:graph contain y_data and y_predicted

- **Testing the Performance of our Model**

  By calculating:

  - Mean Absolute Error (MAE)
  - Mean Squared Error (MSE)
  - Root Mean Squared Error (RMSE)

C. WEB MAP

Web maps are online maps created with ArcGIS that provide a way to work and interact with geographic content organized as layers.

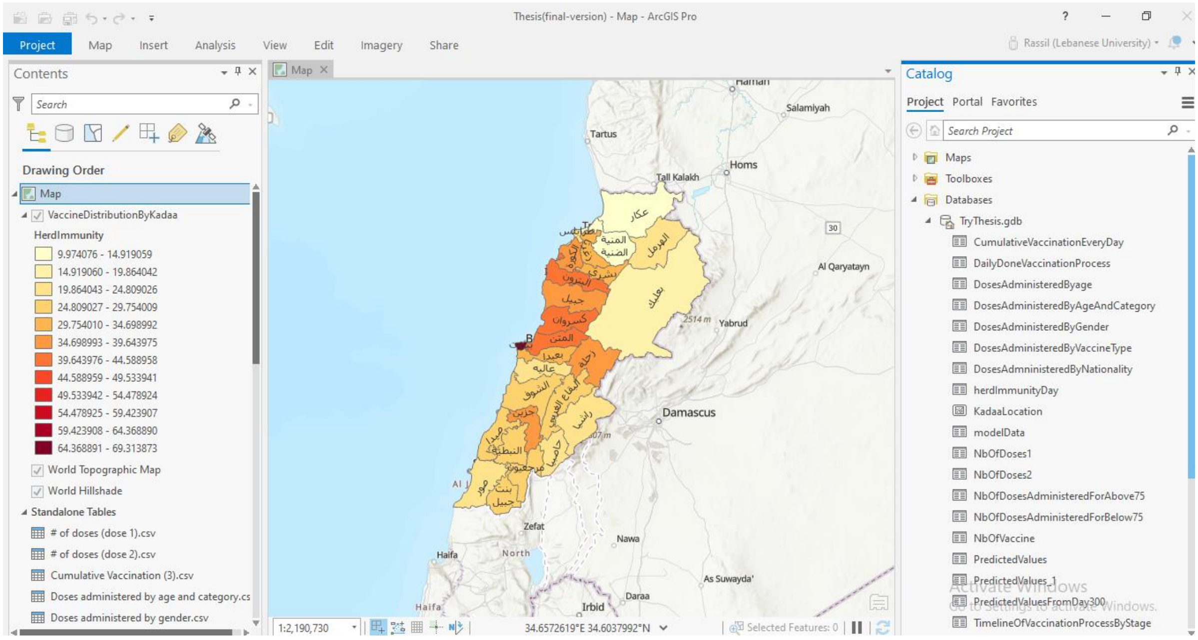


Figure 10:web map in ArcGIS Pro

i. Overview

An ArcGIS web map is an interactive display of geographic information, you can use it to tell stories and answer questions.

Maps contain a basemap, a set of data layers (many of which include interactive pop-up windows with information about the data), an extent, and navigation tools to pan and zoom. In general, the basemap and layers are hosted and shared through ArcGIS Online. However, maps can also contain layers added directly to the map and layers and basemaps referenced externally. Many maps also contain scaled symbols and other smart styling that reveal data and patterns as you interact with it.

Maps can be opened in standard web browsers, mobile devices, and desktop map viewers. They can be shared through links or embedded in websites, and used to create map-based web apps .

Software used: ArcGIS Pro, ArcGIS online

Web Map link: https://arcg.is/1O8HKe0

ii. Implementation

Using ArcGIS Pro, we chose a suitable basemap, after that we have to add the tables (csv files) of data (collected from MOPH) to this map to organize and filter them based in our needs.

This last step is applied using some Geoprocessing tools, some of used tools:

- **Summarizing data in a table**: Sometimes the attribute information you have about map features is not organized the way you want. This tools allow us to "Group By" the table.
  In our project, we have the data of vaccination in each area, we have to group by district because our study was applied in each district in Lebanon.

- **Join field:** Joins the contents of a table to another table based on a common attribute field. The input table is updated to contain the fields from the join table. You can select which fields from the join table will be added to the input table.

  In our project we use this tools to join the tables containing data about vaccination to a feature class containing the polygon coordinates for each district, to link this data to the map.

- **Calculate field** :Calculates the values of a field for a feature class, feature layer, or raster.
  In our project we use it to calculate the vaccine frequency based on the number of vaccine administered in the district and the population.

After organize the tables and the features class we have to symbolize and label the map as needed.

In our case we need to visualize the difference of vaccination process in each district, that's why we have used a graduated colors as symbology.

Finally we publish the web map to ArcGIS online to use it in the created dashboard and story map.

D. DASHBOARD

ArcGIS Dashboards enables users to convey information by presenting location-based analytics using intuitive and interactive data visualizations on a single screen.

i. Overview

A dashboard is a view of geographic information and data that allows you to monitor events, make decisions, inform others, and see trends. Dashboards are designed to display multiple visualizations that work together on a single screen. They offer a comprehensive view of your data and provide key insights for at-a-glance decision-making. Similar to web maps and web layers, dashboards are part of the ArcGIS geoinformation model.

Software used: ArcGIS Dashboard

Dashboard link: https://www.arcgis.com/apps/dashboards/65298d6c7eb441e79504cd8d43c9d4dd

ii. Implementation

Our dashboard shows the process of vaccination over Lebanon and containing :

- A map presenting the percentage of vaccine (herd immunity) in each district and shows the difference of percentages between one district and another by using the graduated colors symbology.

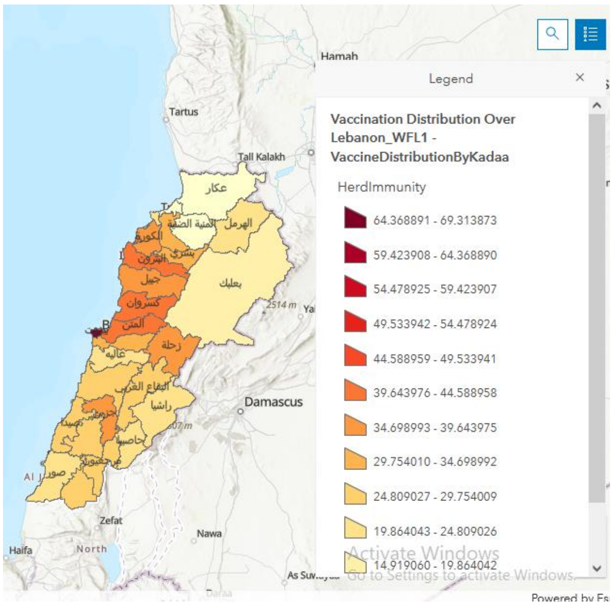


Figure 11:Map presented in the dashboard shows the herd immunity progress in each district

- Total number of administered doses (Doses 1 and Doses 2) in each district

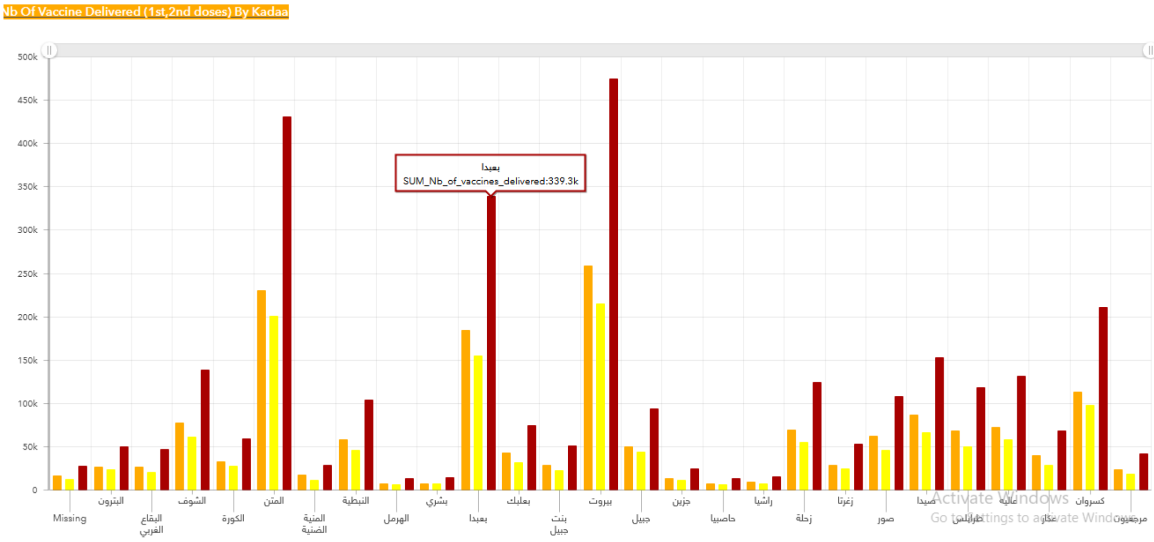


Figure 12: the number of total vaccine , doses1, doses2 in each district

- Total number of administered doses (Doses 1 and Doses 2) over Lebanon

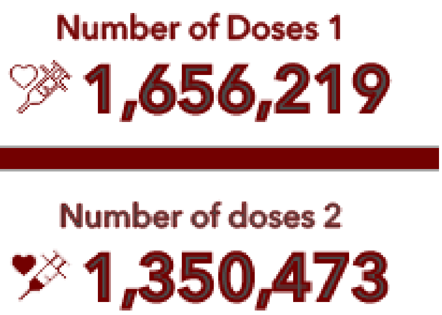


Figure 13: Number of Doses1 , Doses2 in Lebanon

- Number of administered doses by gender.

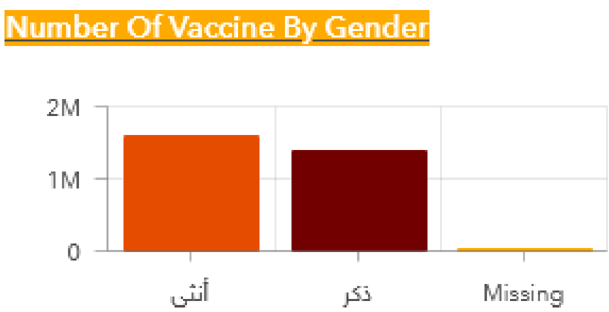


Figure 14: Number of vaccine by gender

- Number of administered doses by vaccine type.

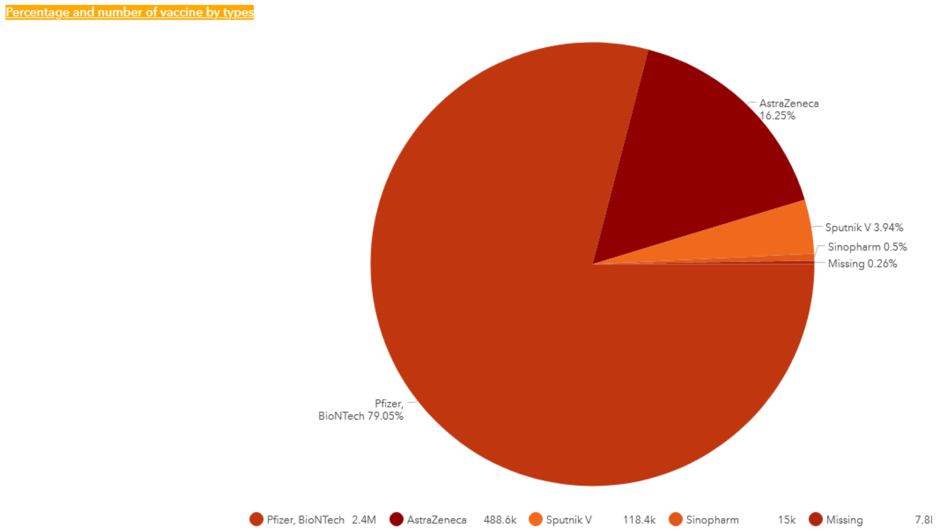


Figure 15:Number of vaccine by vaccine type

- Number of administered doses by nationality.

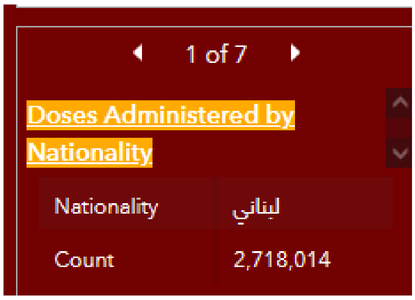


Figure 16: Number of gender by nationality

- Number of administered doses by age.

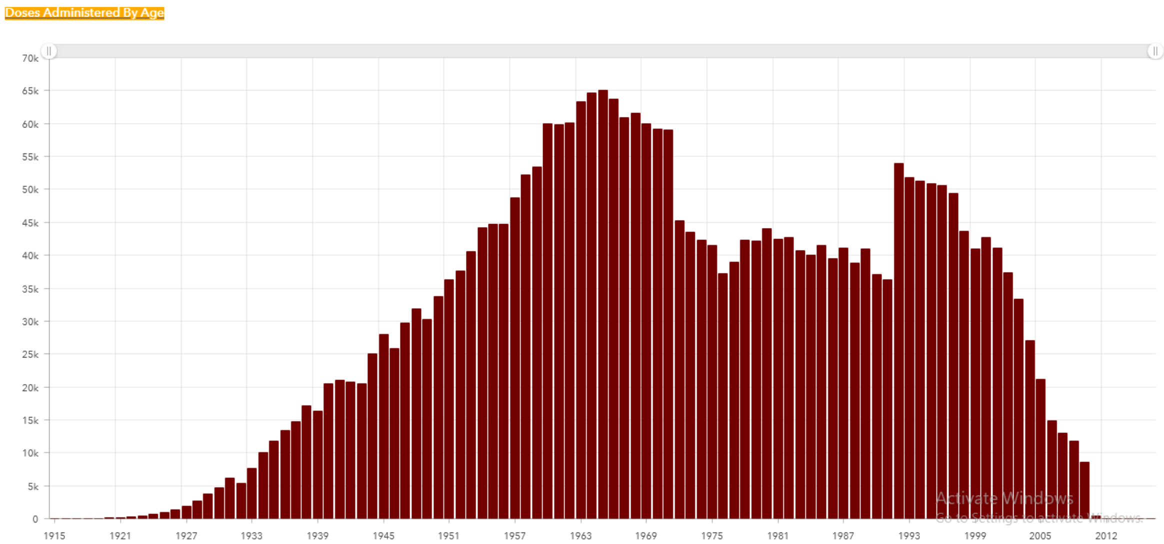


Figure 17: Number of vaccine by age

- Cumulative vaccine process from the first vaccination day.

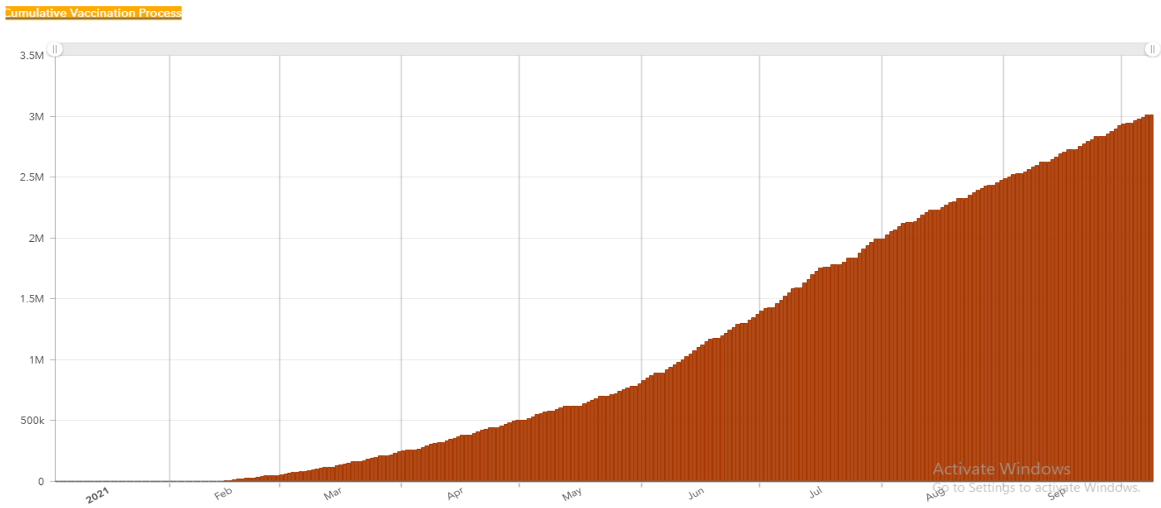


Figure 18: Cumulative vaccine process

- Daily number of vaccine.

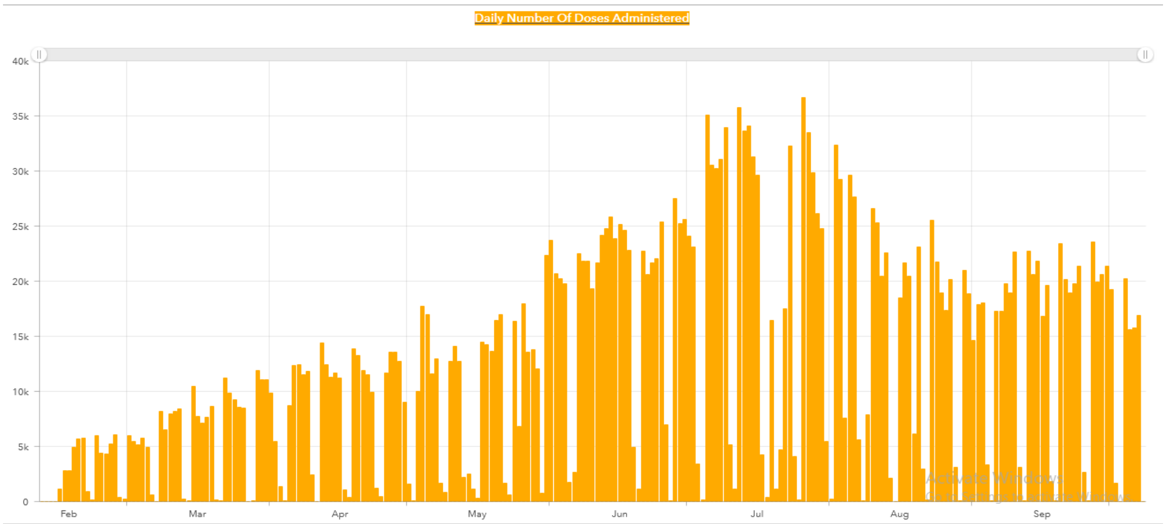


Figure 19:Daily number of vaccine in Lebanon

- Herd immunity predicted month, and a link for a story map explained the herd immunity

Figure 20: Herd Immunity month and link to the story map

E. STORY MAP

i. Overview

ArcGIS StoryMaps is a story authoring web-based application that enables you to share your maps in the context of narrative text and other multimedia content.

Stories are commonly used across ArcGIS to share maps with supporting text and media.

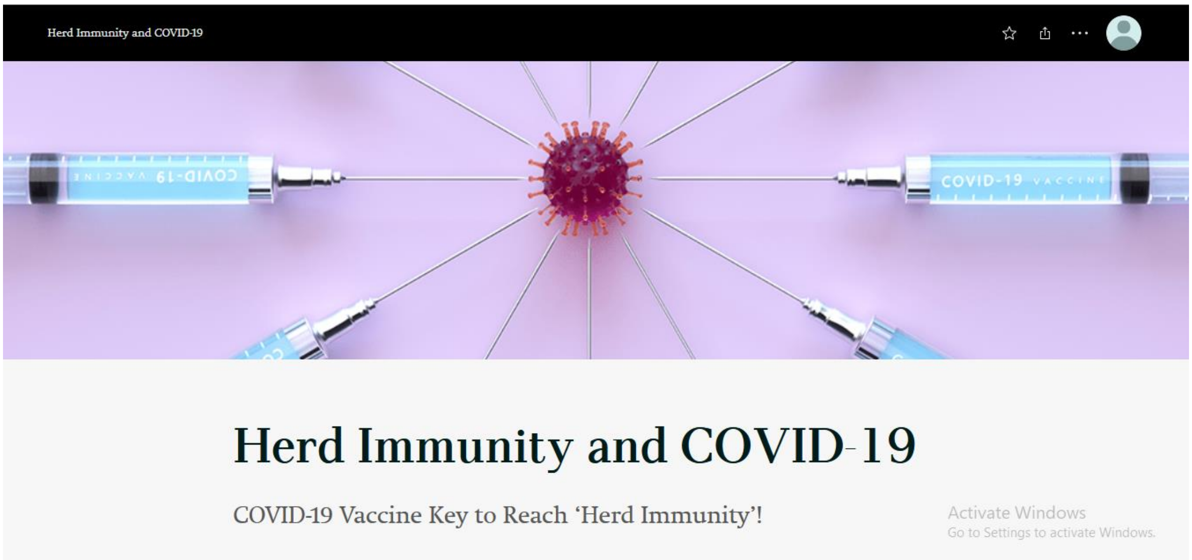
Software used: ArcGIS Story Map

Story Map link:https://storymaps.arcgis.com/stories/d6c4f1732ec444b9a733b9b9548823c2

ii. Implementation

In our project, the Story Map aims to introduce the herd Immunity for those ignorant it and focus on the importance of take the vaccine to achieve herd Immunity and encourage people not vaccinated to take the vaccine since it contains:

- Link to the vaccine registration site (for people not vaccination)
- Clear and brief description about herd immunity and how we can reach it.
- YouTube video and illustration about herd Immunity
- The web application contains the map that shows the vaccine distribution in each district to show the process of vaccine in district whose they are in (with vaccine frequency).



Chapter IV: RESULTS

As a recap, our project objectives are:

1. Create a model that predict when we achieve herd immunity
2. Create a dashboard containing the result of prediction and present all vaccination data in a clear view.
3. Create a Story Map about herd immunity in each district in Lebanon and contains description (text, video, image) about herd immunity to let people know what is herd immunity.

So what are the results of these?

1. Prediction Model Results

Our model take as input the vaccination day, and predict the number of vaccine in this day. Our object is to predict when we will achieve herd Immunity.

We have mentioned before that herd immunity occurs when a large portion of a community (about 90%) take the vaccine.

So first we have to do a study to know the number of vaccine related to the 90% of population.

Lebanon population : 4 842 400

90% of population: 4 358 160

Since everybody must take the 2 doses so the total number of vaccine delivered to achieve her immunity is: 8 716 320

So our mission is to find the day when this number is achieved.

After many tries we deduce that the after 650 day of the first vaccination day, we will reach this number of vaccine.

First vaccination day: 3 January 2021

After 650 days of the mentioned date means in October 2022.

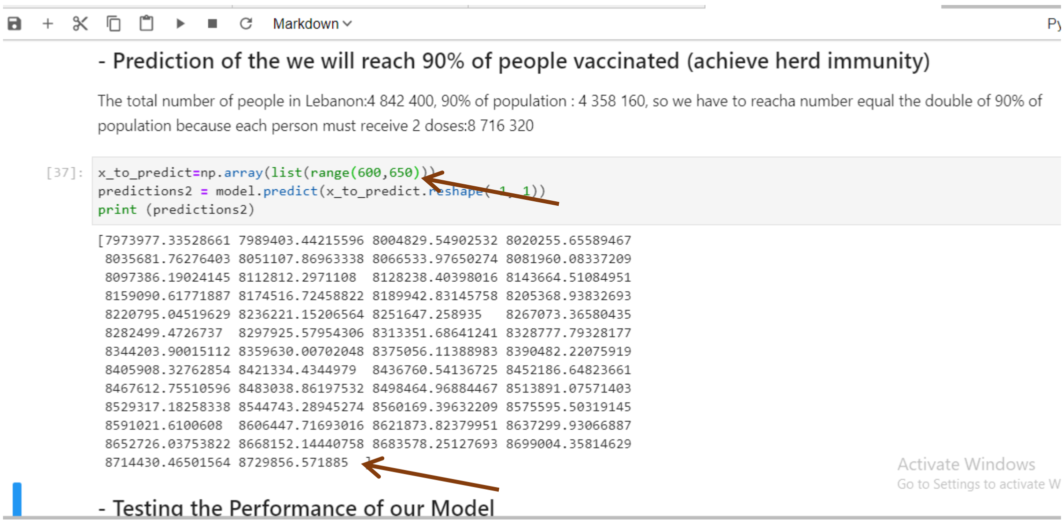


Figure 21:Prediction Model Results

2. Dashboard Results

This Dashboard facilitates the study of herd immunity by divide each data in different categories and present them in a clear way.

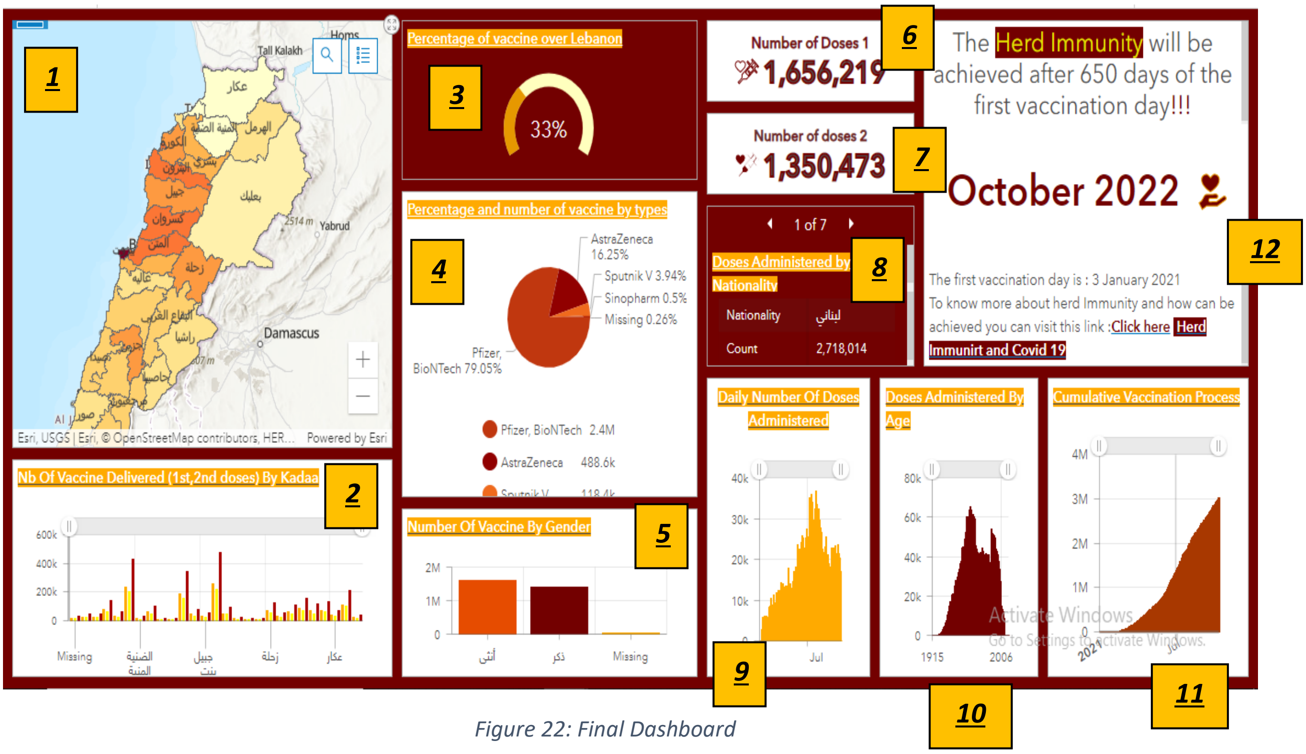


Figure 22: Final Dashboard

Through this dashboard we can know:

1) Distribution of vaccine in each district and number of doses delivered (1st and 2nd doses) in each area by clicking in this area and reading the pop-ups.
2) Number of total vaccine delivered , doses 1 and 2 in all districts of Lebanon.
3) Average percentage of vaccination process.
4) Percentage and number of vaccine delivered by vaccine type. We deduce that about 80% of people take the Pfizer type.
5) Number of vaccine delivered by gender.
6) Total number of doses 1 delivered.
7) Total number of doses 2 delivered.
8) Number of vaccine delivered by nationality.
9) Daily number of doses administered.
10) Number of vaccine delivered by age.
11) Cumulative vaccination process
12) The result of the prediction model and a link to the story map explaining the herd immunity.

3. Story Map Results

Throw the Story Map anybody can:

- Register for the vaccine if he haven't taken it yet.
- Understand what is herd immunity by reading the text description, illustration and videos that make this description more clear.
- Understand the role of vaccination to achieve herd immunity
- Learn the importance of taking the vaccine to protect himself and others, whose cannot take it.
- Know the process of vaccine in each district and visualize where the herd immunity will be achieved first by clicking in the web map presented in the story map.

Figure 23:Story Map

Chapter V: DISCUSSION

There should be no doubt about it: Covid-19 vaccines are saving lives!
Despite initial worries that the current vaccines may be less effective against the Delta variant, analyses suggest that both the AstraZeneca and the Pfizer-BioNTech jabs reduce hospitalization rates by 92-96%. As many health practitioners have repeated, the risks of severe side effects from a vaccine are tiny in comparison to the risk of the disease itself.
Yet a sizeable number of people are still reluctant to get the shots.
Some people may object to getting a COVID-19 vaccine because of religious objections, fears about the possible risks or skepticism about the benefits. If the proportion of vaccinated people in a community is below the herd immunity threshold, a contagious disease could continue to spread. (Robson, 23rd July 2021)

The result is becoming something of a culture war on social media, with many online commentators claiming that the vaccine hesitant are simply ignorant or selfish. But psychologists who specialize in medical decision-making argue these choices are often the result of many complicating factors that need to be addressed sensitively, if we are to have any hope of reaching population-level immunity.

Figure 24: People and vaccine

Over all, we must be optimistic that everything will be fine and we all have to take the vaccine to achieve herd immunity in the predicted date.

Chapter VI: CONCLUSION

'Herd immunity', also known as 'population immunity', is the indirect protection from an infectious disease that happens when a population is immune either through vaccination.

Herd immunity against COVID-19 should be achieved by protecting people through vaccination, not by exposing them to the pathogen that causes the disease.

Vaccines train our immune systems to create proteins that fight disease, known as 'antibodies', just as would happen when we are exposed to a disease but – crucially – vaccines work without making us sick.

Vaccinated people are protected from getting the disease in question and passing on the pathogen, breaking any chains of transmission.

To safely achieve herd immunity against COVID-19, a substantial proportion of a population would need to be vaccinated, lowering the overall amount of virus able to spread in the whole population. One of the aims with working towards herd immunity is to keep vulnerable groups who cannot get vaccinated (e.g. due to health conditions like allergic reactions to the vaccine) safe and protected from the disease.

The number of people who are immune to the coronavirus is increasing every day. This includes people getting vaccinated and, unfortunately, a lot of people getting COVID-19. At some point—hopefully soon —enough people in our population will have SARS-CoV-2 immunity. With higher immunity levels, and continuing some of the behavioral changes that reduce transmission levels of infection will fall again. (World Health Organization, 2020)

So achieving herd immunity with safe and effective vaccines makes diseases rarer and saves lives.

For this reason, people should not neglect vaccination, and the development of vaccine (daily number of vaccine) must be maintained to reach herd immunity in the predicted month : October 2021

Bibliography

My accounting course. (2016). Retrieved from Regression Model: https://www.myaccountingcourse.com/accounting-dictionary/regression-model

lotame. (2019, may 13). Retrieved from https://www.lotame.com/what-are-the-methods-of-data-collection/

World Health Organization. (2020, december 31). Retrieved from https://www.who.int/news-room/q-a-detail/herd-immunity-lockdowns-and-covid-19

ArcGIS. (n.d.). Retrieved from https://en.wikipedia.org/wiki/ArcGIS

Chen, J. (August 2020). Science of The Total Environment. *ELSEVIER*.

Esri. (n.d.). Retrieved from https://en.wikipedia.org/wiki/Esri

ESRI. (n.d.). Retrieved from https://www.esri.com/en-us/arcgis/products/spatial-analytics-data-science/overview

Gu, Y. (2020). *Path to Normality - COVID-19 Vaccine Projections.* US.

Gypsyamber D'Souza, D. D. (13 September 2021). Rethinking Herd Immunity and the Covid-19 Response End Game.

Howard, J. (25 july 2021). “We’ll Have Herd Immunity by April”: Reflections on a Failed Prediction.

Newman, M. E. (25, August, 2020). COVID-19 Story Tip: Herd Immunity Is a Dangerous Strategy for Fighting COVID-19, Says Johns Hopkins Expert.

Robson, D. (23rd July 2021). Why some people don't want a Covid-19 vaccine. 1-10.

LIST OF ABBREVIATIONS

SARS-CoV-2:severe acute respiratory syndrome coronavirus 2

GIS: Geographic Information Systems

MOPH: Ministry Of Public Health

KDD: Knowledge discovery in databases

LIST OF FIGURES

Output :

Dashboard : https://www.arcgis.com/apps/dashboards/65298d6c7eb441e79504cd8d43c9d4dd

Story Map : https://storymaps.arcgis.com/stories/d6c4f1732ec444b9a733b9b9548823c2