VIETNAM NATIONAL UNIVERSITY - HO CHI MINH CITY
UNIVERSITY OF TECHNOLOGY
FACULTY OF COMPUTER SCIENCE AND ENGINEERING

Mathematical Foundation for Data Science

Assignment
COVID-19 Analysis and Prediction

Advisors: PhD. Tran Tuan Anh
PhD. Nguyen Tien Thinh

Group 1: Vu Trung Hieu - 1970215
Nguyen Trong Tinh - 1970219
Vo Tien Dat - 1970214
Nguyen Tan Nam Anh - 1970013

Table of Contents

1. Data Analysis

1.1. Data Source

The data is collected from CSSEGISandData repository which is operated by Johns Hopkins University. They are aggregated from many sources (e.g., WHO, ECDC, US CDC, BNO News), since January 21, 2020. The raw data is listed as below:

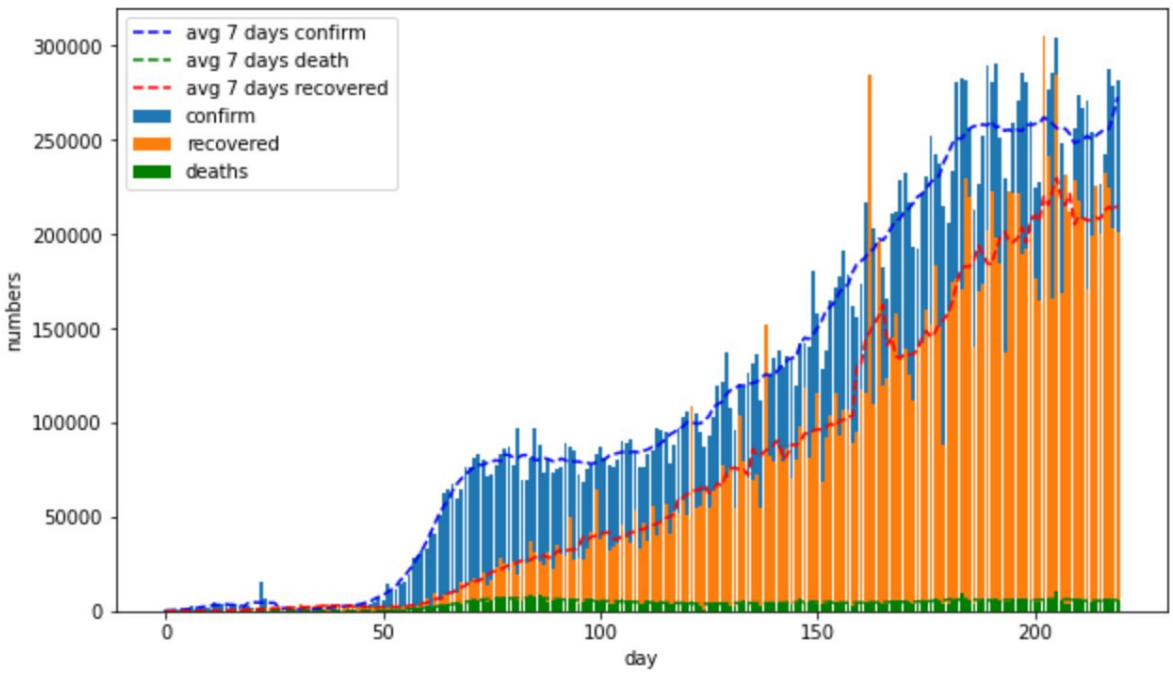
```
  Province/State Country/Region       Lat  ...  8/26/20  8/27/20  8/28/20
0            NaN    Afghanistan  33.93911  ...    38113    38129    38140
1            NaN        Albania  41.15330  ...     8927     9083     9195
2            NaN        Algeria  28.03390  ...    42619    43016    43403
3            NaN        Andorra  42.50630  ...     1098     1098     1124
4            NaN         Angola -11.20270  ...     2332     2415     2471

[5 rows x 224 columns]
```

There are some important attributes in the table: "*Province/State*", "*Country/Region*", "*Days*", they are useful in visualizing data. Currently we have three tables: *confirm table* - number of confirm people, *recovered table* - number of recovered people, *death table* - number of death people.

The data used in this assignment was obtained from January 21, 2020 to August 24, 2020.

1.2. The World Data

The data of the world are visualized as below:

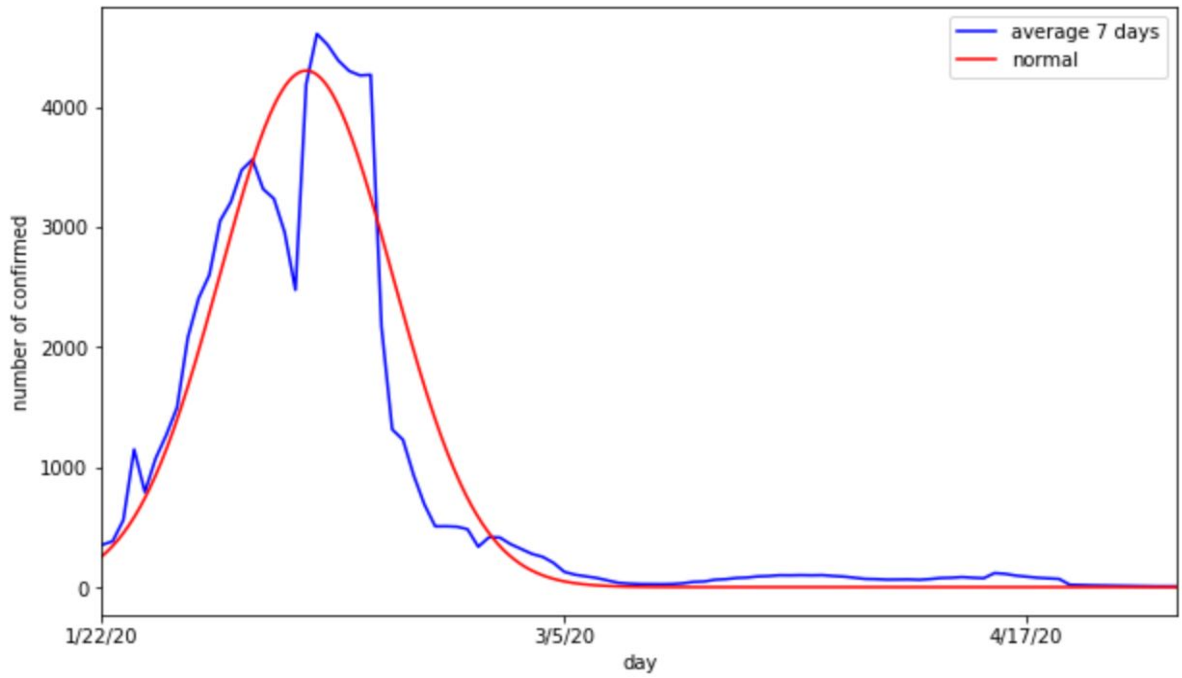


Generally, the number of confirmed, recovered and death cases tend to increase from January 21 to August 24. In that, the confirmed and recovered cases account for the majority. The COVID pandemic is infectious and spreads out highly. The number of dead people is of small proportion, almost they are usually not healthy before. Moreover, there are two stages in above picture:

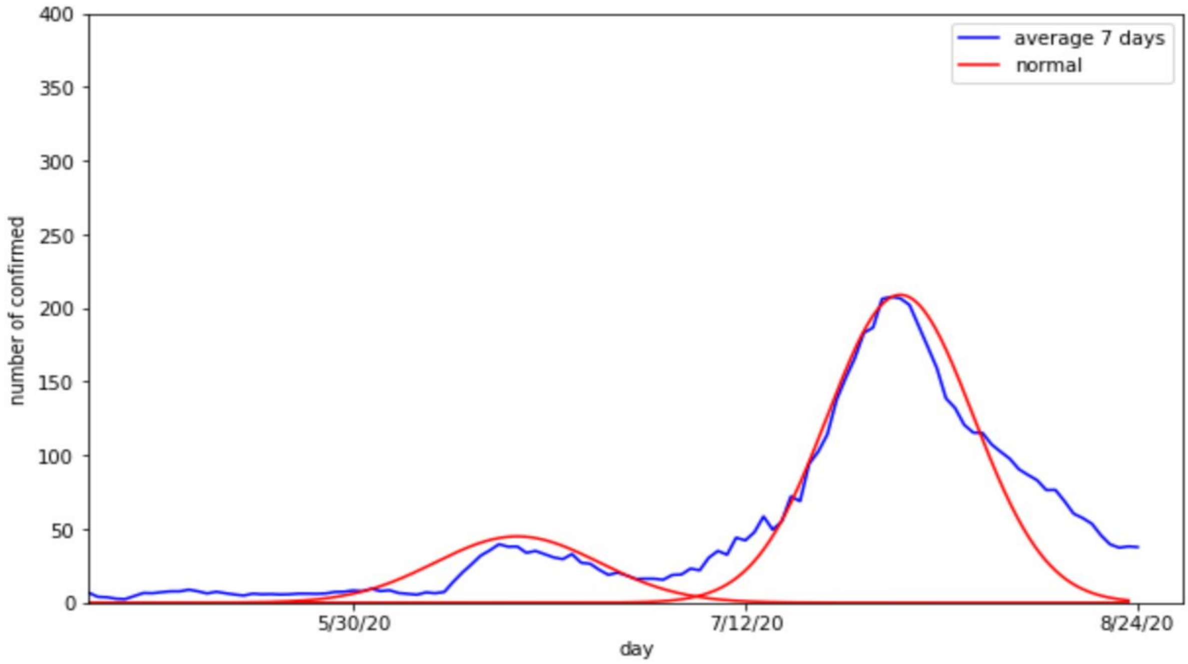
- The first stage: At the end January, there were the first cases in Wuhan, China and spread to many neighboring areas in China, then Asia.
- The second stage (4/2020 - 8/2020): The pandemic breaks out in many countries in Europe and Americas.

1.3. China Data

In China, we also have two stages: The first stage was from January 21 to May 3. At this time, China had the first infectious case at Wuhan and spread quickly in the neighborhood area. We can see in the picture below, at this stage China has more than 1000 new cases/day.
However, the government of China had strong policies to prevent the COVID pandemic. This is the reason why the new cases are decreasing at the beginning of May.



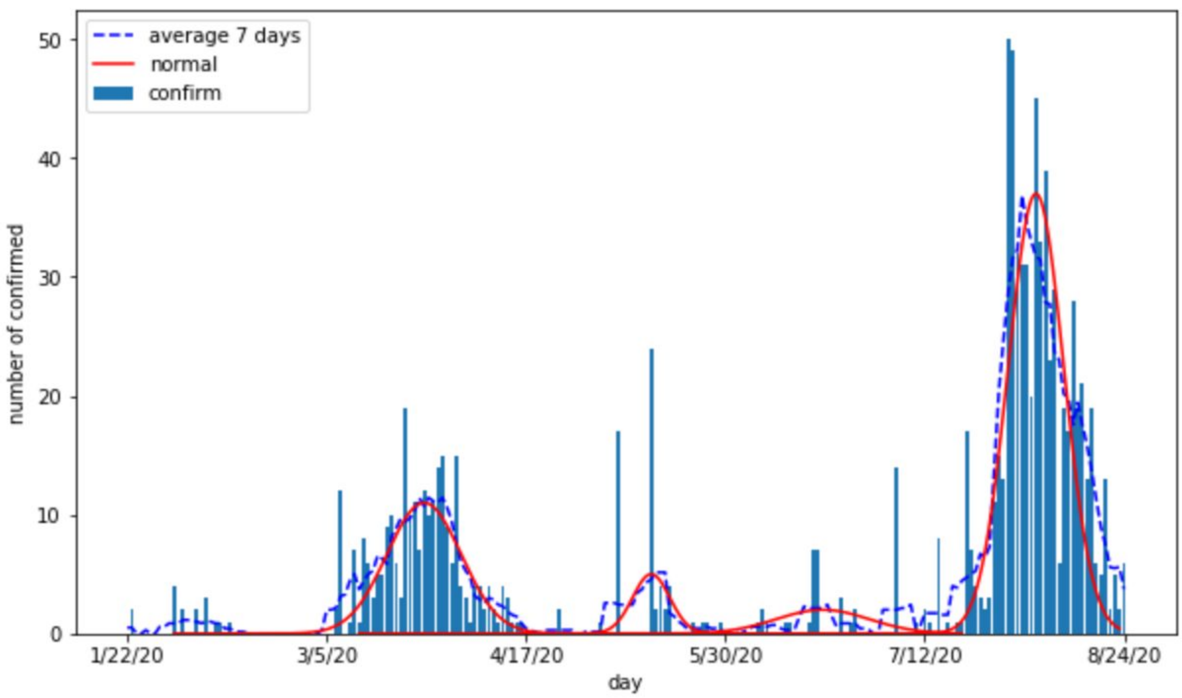
The second stage is from the end of May to August 24. In the long existence time, the covid progresses many different variants, so it is hard to prevent and control. Finally, the pandemic broke out again on May 30.



1.4. Vietnam Data

We also have two stages in Vietnam: The first stage is at the beginning of May. At this time, Vietnam welcomed many tourists who came back from the country having pandemic. Moreover, Vietnam is the neighboring country of China. These are the reasons why the pandemic came to Vietnam and spread to many provinces in Vietnam, especially Hanoi, Ho Chi Minh and Da Nang.

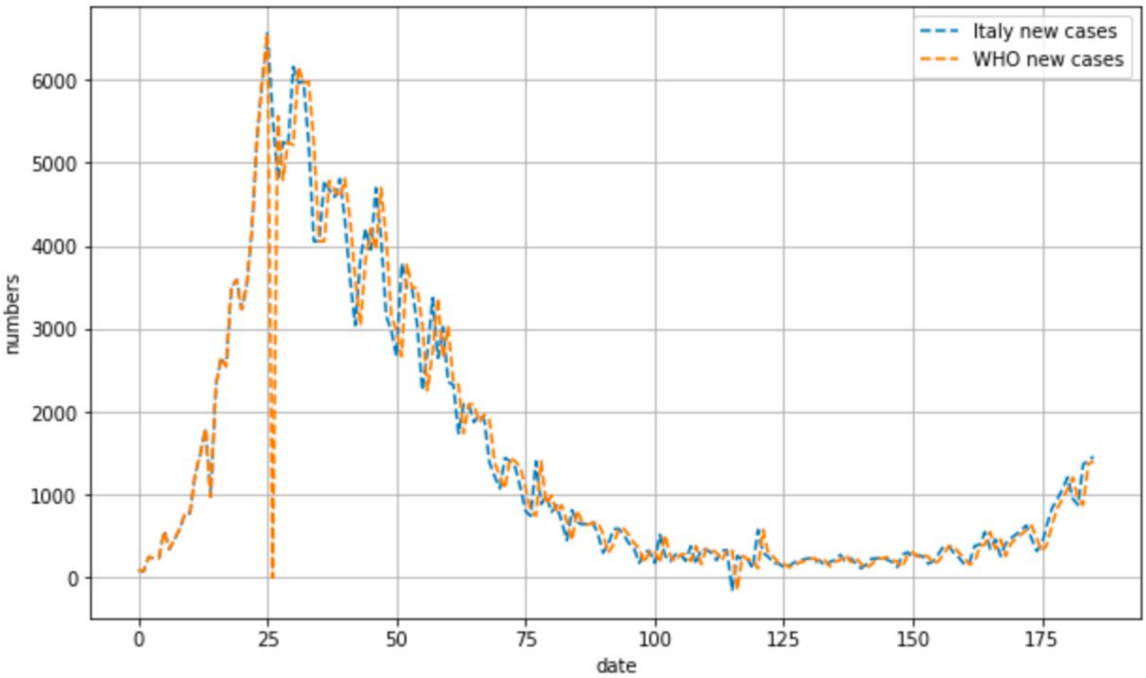
In Vietnam, the government also applied many strong policies to prevent the COVID. The lockdown and social distancing is one of the efficient methods to prevent spreading, so the new cases decrease quickly after 15 days break out.



After the first stage, Vietnam opened and welcomed the tourists from many countries in the world. Additionally, there are many variants of COVID appear and relax social distancing. The pandemic broke out again on July 12, but Vietnam controlled quickly.

1.5. WHO vs Italy

Currently, the data from WHO and Italy is the same, but there were some mistakes in records from WHO (be like as below). There is a zeros point where WHO can not collect data.



2. Time-Dependent SIR Model

2.1. Introduction to SIR Model

The COVID-19 infectious disease is an epidemic. To model an epidemic, suitable groups (also known as compartments) are defined to cover the entire population of a country. The population N is assumed to be constant because:

- The epidemic has a (relatively) short time scale, so the new births can be neglected.
- The number of deaths is small as compared with the entire population.
- Travel restrictions are enforced.

The population N is divided into mutually exclusive groups: susceptible (S), infected (I), and recovered (R) or dead as below:

In this model, the deaths due to COVID-19 are assumed in the recovered group. Every group is assumed to have the same characteristics. Susceptible people can contract the virus and be infected. The recovered group is immune. Mathematically, the population N, which does not vary in time, is the sum of the groups S, I, R, which vary in time. With the assumptions, the SIR model is expressed by three ordinary differential equations (ODEs).

$$\frac{dS}{dt} = -\beta \frac{I}{N} S$$
$$\frac{dI}{dt} = \beta \frac{I}{N} S - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

I/N is the probability to come into contact with an infected individual; β is the average number of contacts per person per unit of time weighed by the transmissibility (contact rate); γ is the average number of recovered/dead people per unit of time (recovery/death rate). The basic reproduction ratio, R0=β/γ, is the expected number of secondary infections from a single infection entering a population where all members are susceptible. If R0 > 1, the number of infected increases. If R0 < 1, the disease does not grow on average.

2.2. Time-Dependent Model

Consider β and γ changes over time:

$$\frac{dS(t)}{dt} = -\beta(t) \frac{I(t)}{N} S(t)$$
$$\frac{dI(t)}{dt} = \beta(t) \frac{I(t)}{N} S(t) - \gamma(t) I(t)$$
$$\frac{dR(t)}{dt} = \gamma(t) I(t)$$

Due to the COVID-19 data is updated in days, the above equations can be revised into discrete time difference equations:

$$S(t+1) - S(t) = -\beta(t)\frac{I(t)}{N}S(t)$$
$$I(t+1) - I(t) = \beta(t)\frac{I(t)}{N}S(t) - \gamma(t)I(t)$$
$$R(t+1) - R(t) = \gamma(t)I(t)$$

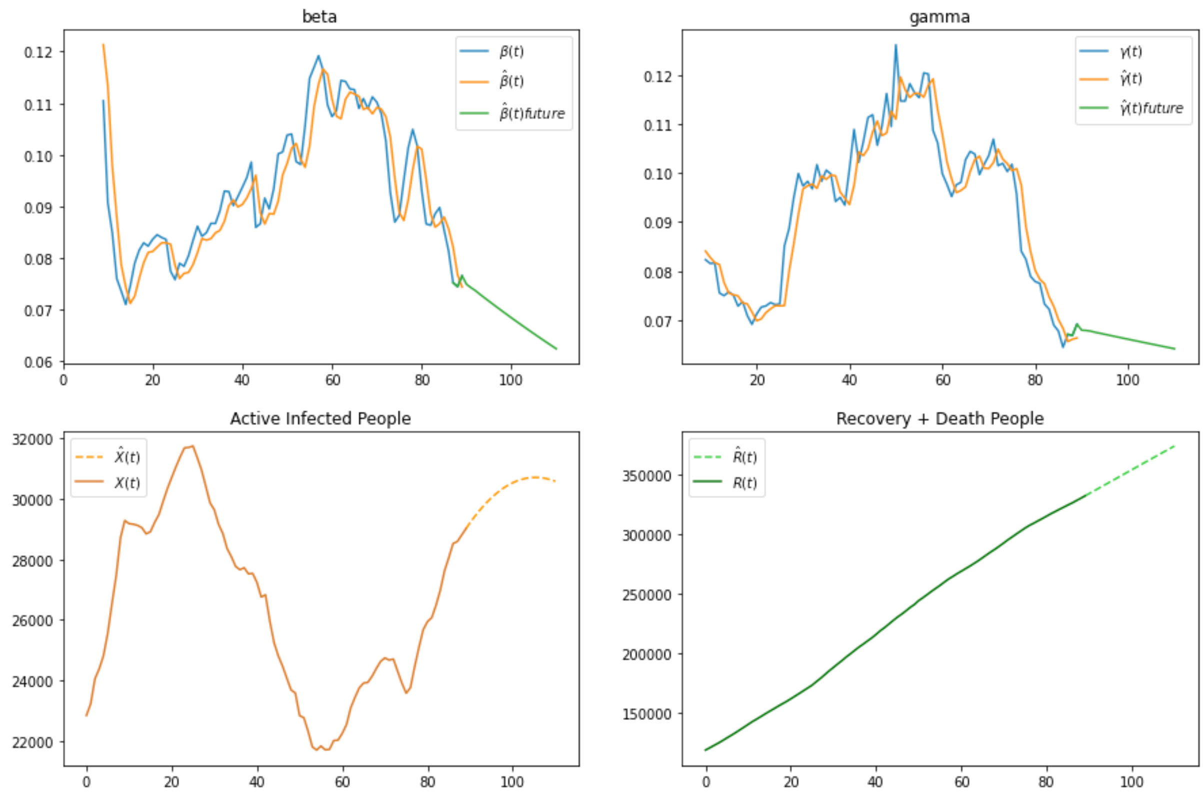
In COVID-19 data, we can obtain β(t) and γ(t) from S(t), I(t), and R(t) (inferred from the above discrete time difference equations):

$$\beta(t) = \frac{N}{I(t)S(t)}(S(t) - S(t+1))$$
$$\gamma(t) = (R(t+1) - R(t))\frac{1}{I(t)}$$

We will use machine learning (linear regression) to learn two curves β(t) and γ(t). Then, use the learned model to predict the β and γ in the future. Finally, use the discrete time difference equations to calculate S(t), I(t) and R(t).

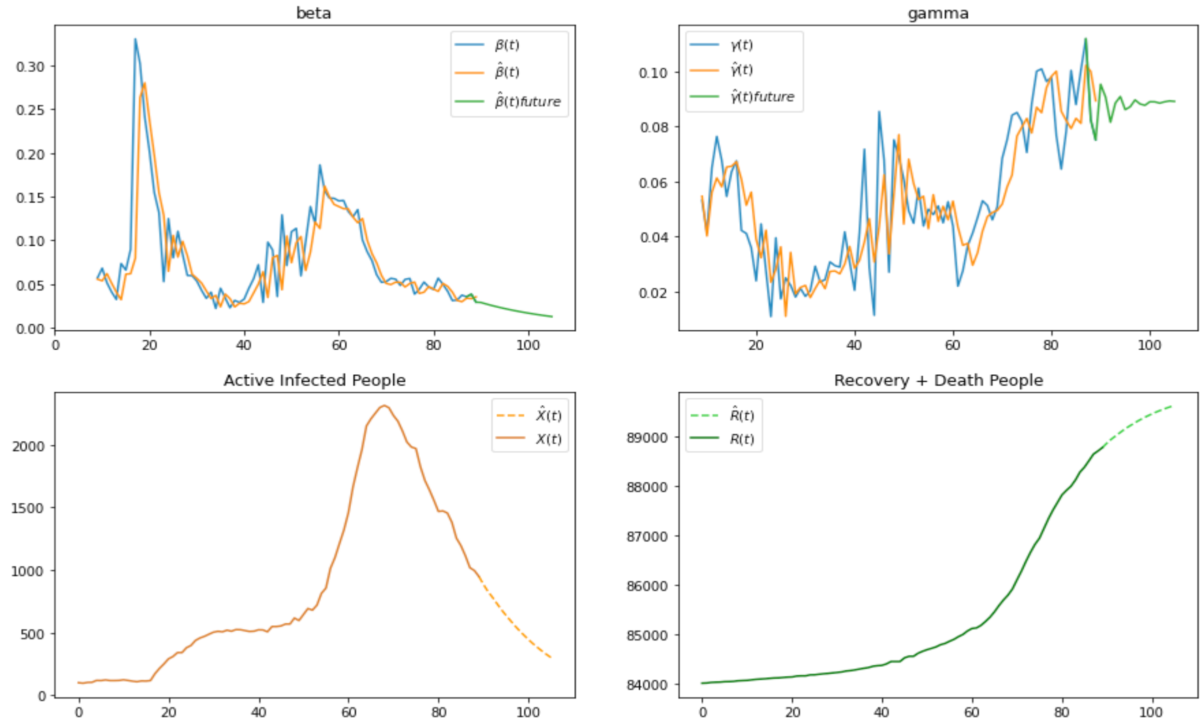
2.3. Result

2.3.1. Iran



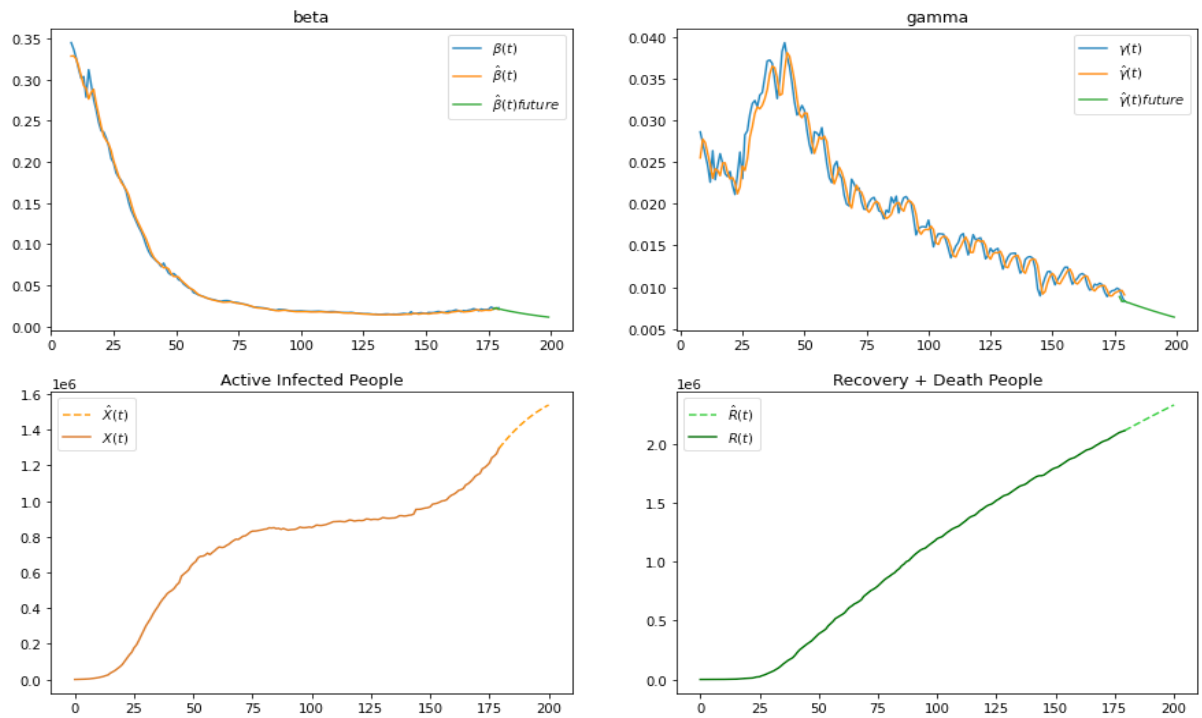
From 60th day, the β is high, and the γ is low, so the active infected people increase dramatically. However, the trend of β is decreasing while the γ is decreasing with lower rate. The predicted values show that they will keep that trend. So, the infected group will come to a peak then start to fall off.

2.3.2. China



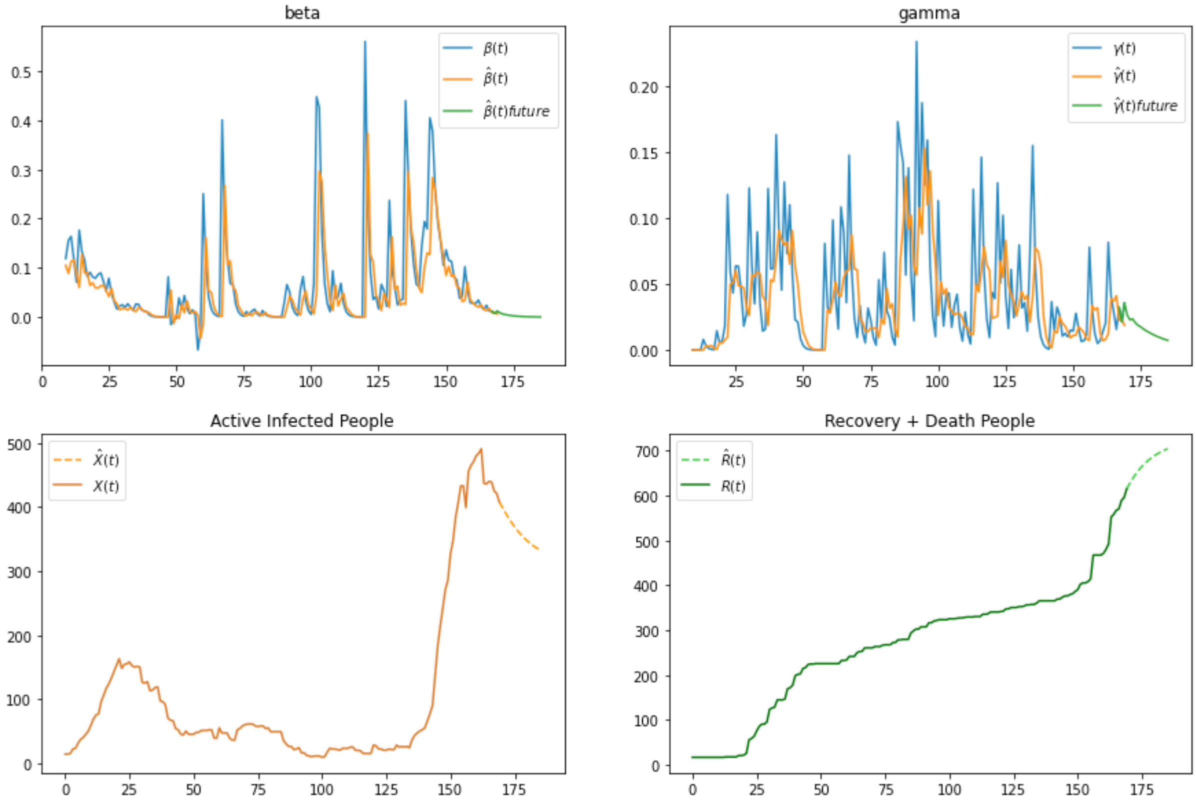
We can see from the graph, the trend of β is decreasing and γ is going stable. So the infected people keep declining.

2.3.3. Europe



The trend of β is stable but the γ is decreasing. So the prediction values of infected people are increasing.

2.3.4. Vietnam



Vietnam has a small number of infected people. The parameters β and γ calculated from S, I and R, have a lot of noises. It is difficult to learn the trend of β and γ. The prediction values may not be good.

2.4. Conclusion

In this section, the report shows the mathematical model for COVID-19. The time-dependent SIR model is more adaptive than traditional static SIR models. It provides the information about contact rate and recovery (or death) rate for each stage of COVID-19. This method helps to predict the trend of these parameters and the infected with recovered (dead) people.

3. Linear Model

3.1. Polynomial Linear Model

A linear regression model has the form:

$f(\theta,x) \coloneqq \theta_0+\theta_1 x_1+ \ldots +\theta_n x_n$

It learns to set its parameters, denoted as θ, so that it can best describe the actual data, by minimizing the loss function:

$\sum(y^{(i)} - f(\theta, x^{(i)}))^2$

The number of confirmed cases, deaths and recoveries over time is a nondecreasing function y = f(x), where y is the number of cumulative cases at day number x. If we use x as the only feature for the linear regression model, it will only produce a straight 2d line as the prediction, which is not appropriate to predict most of the countries' data.

Using intuition, we observed that the cumulative number of cases can be described by a polynomial function (easier than for daily data). A polynomial function has the form:

$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$

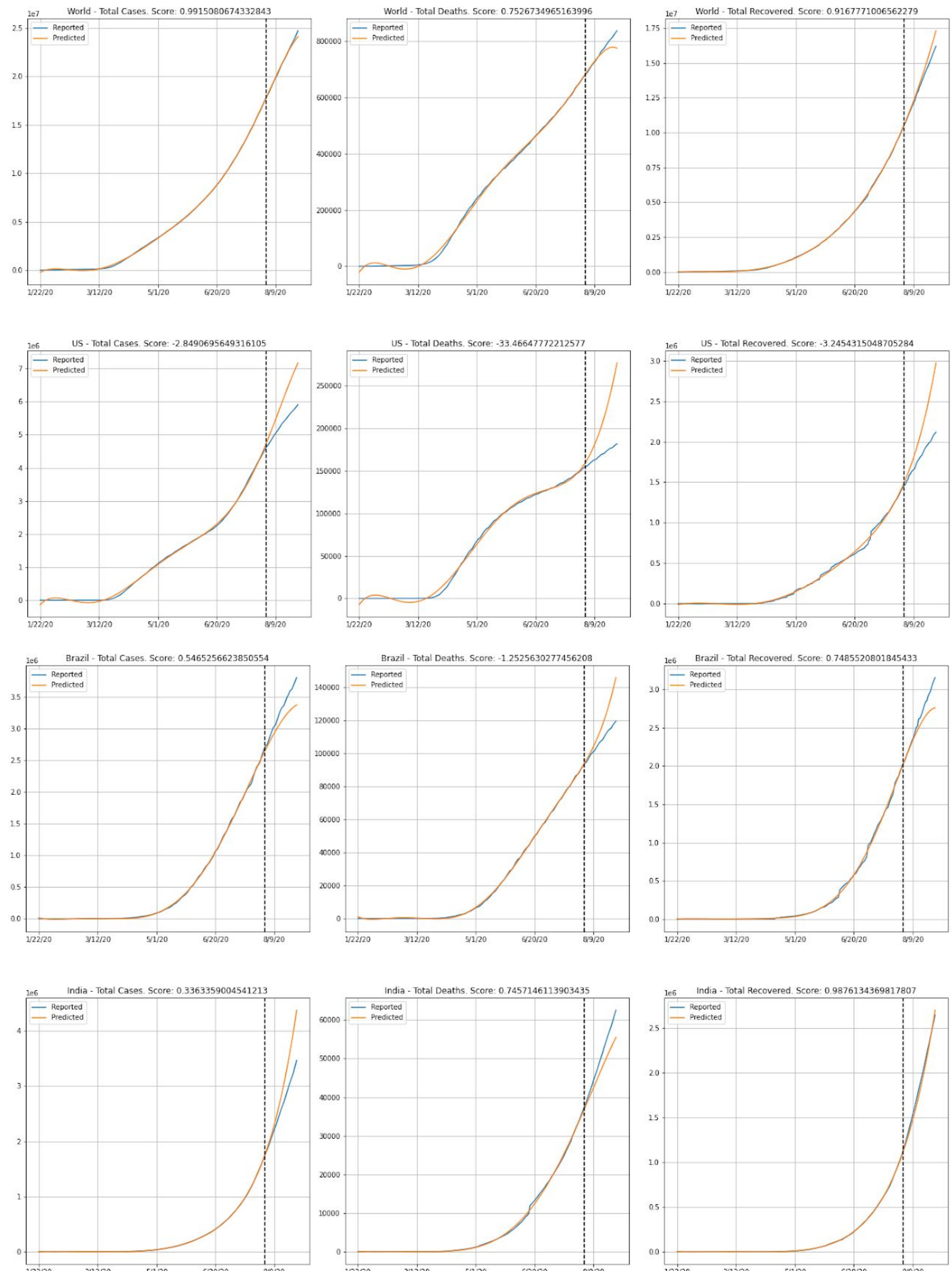
where n is called the degree. After selecting a degree and generating features x^2, x^3, …, x^n, a linear regression model can be used to fit the polynomial function. We temporarily name this as a polynomial linear model.

3.2. Applying to world and top 3 countries' data

For each country, one model was used to train on each of the 3 time series data tables: cumulative confirmed cases, cumulative deaths, and cumulative recoveries. However, these 3 models have these in common:

- Date to split train and test data. Train data is all data from the first day of data collection (which is Jan 22, 2020) until split date; and test data is all data from split date until current day, or up to a number of days.
- Number of degrees - the only hyperparameter.

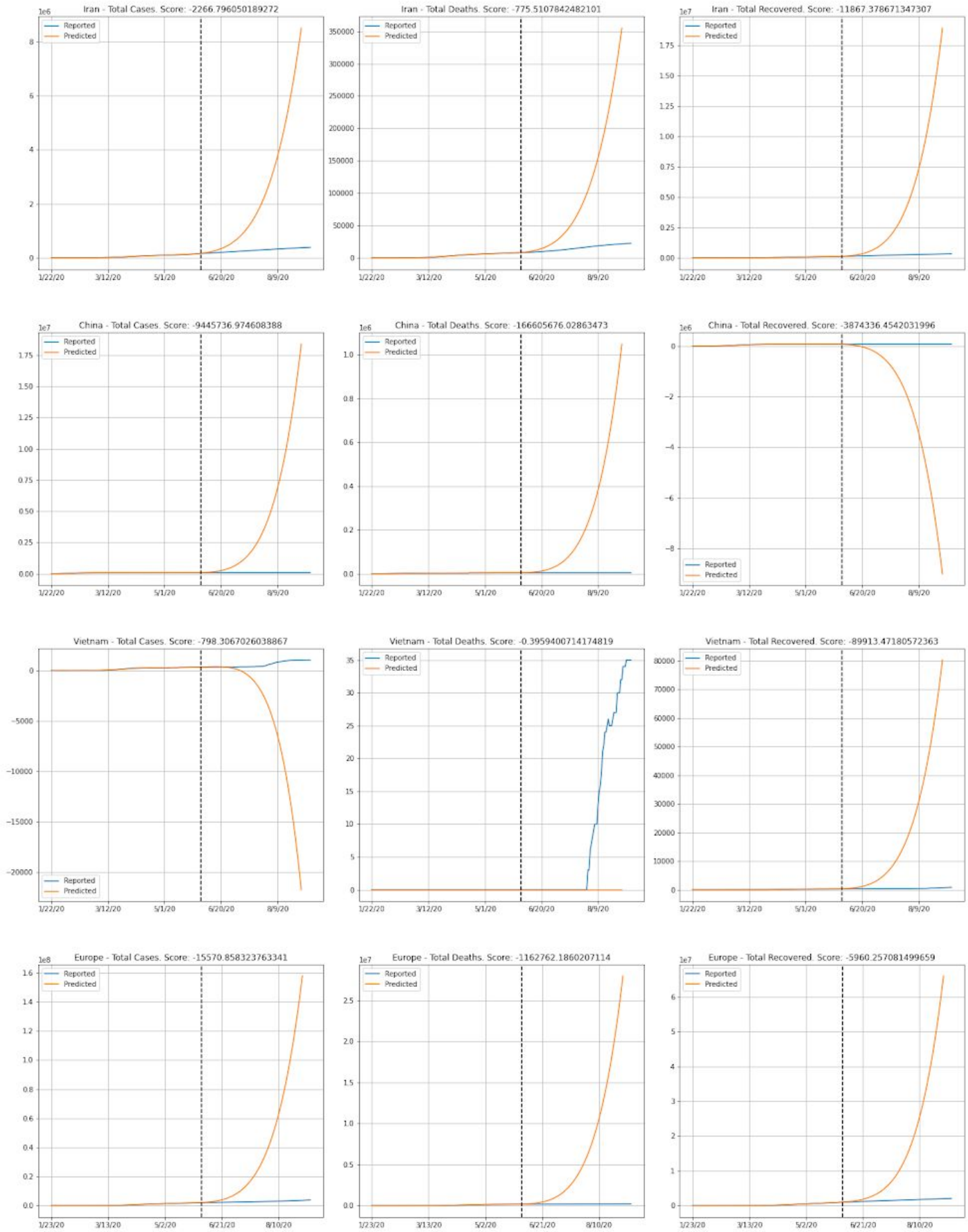
The polynomial model with degree of 6 proved to be accurate in predicting the data of August 2020 for the whole world and top 3 countries with most cases: US, Brazil and India.



Prediction of 6-degree polynomial linear models on worldwide, US, Brazil and India data. The dotted line is the split date (8/1/2020). Note that the score above each graph was calculated on only the prediction vs actual data after the split date, not for the whole period.

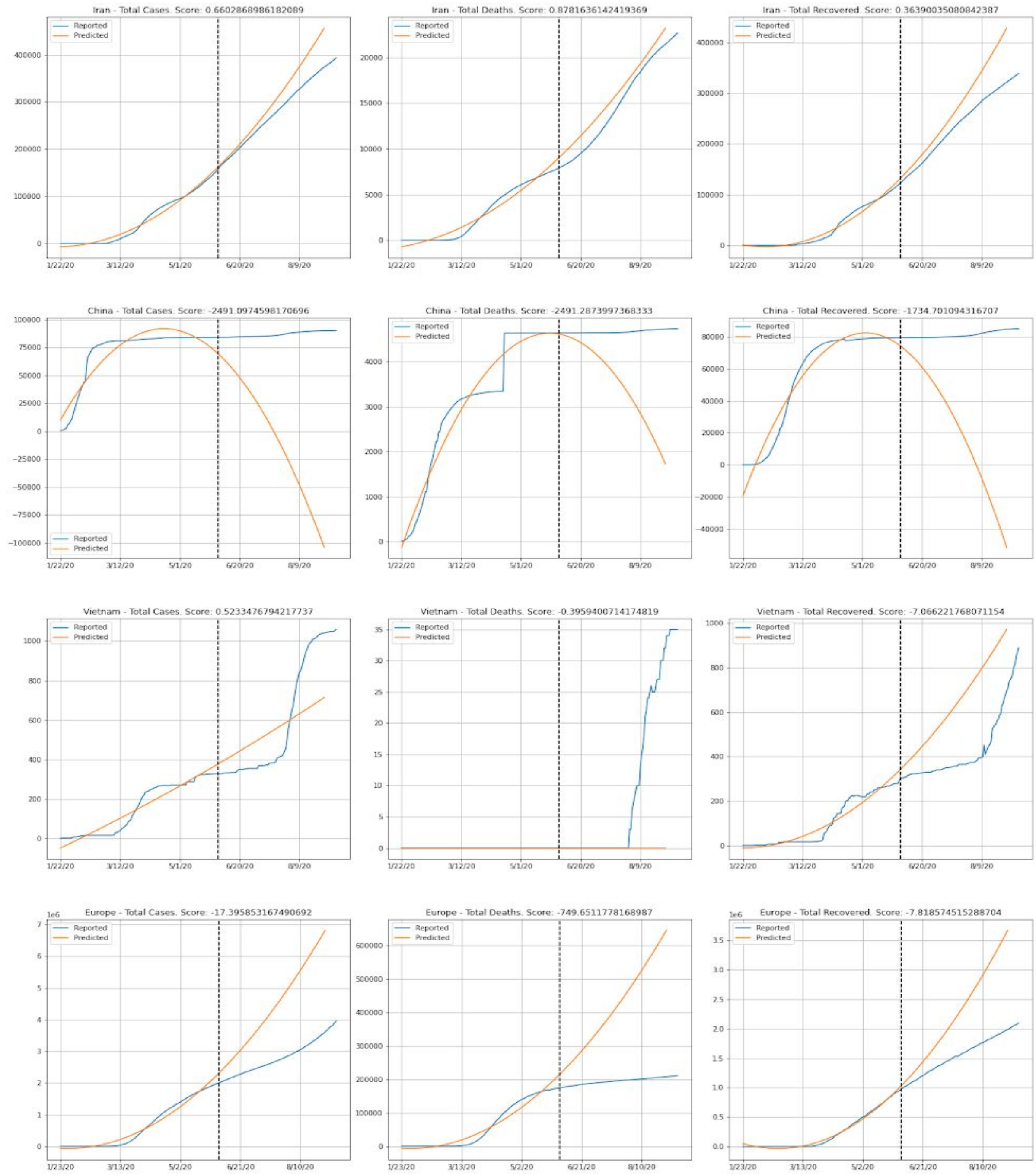
3.3. Comparison to SIR model

The comparison is made by predicting the number of cases in the similar period analyzed by the SIR model, for Iran, China, Vietnam and Europe.



Prediction of 6-degree polynomial linear models on Iran, China, Vietnam and Europe data.

The same model of degree 6 caused the polynomial linear model to overfit on these countries and Europe’s data, resulting in unrealistic predictions. The most plausible explanation is that the number of degrees is too high. Surprisingly, the polynomial linear model with degree of only 2 performed significantly better. However, it still cannot fit the actual data as close as the SIR model.



Prediction of 2-degree polynomial linear models on Iran, China, Vietnam and Europe data.

3.4. Conclusion

The model works well on data of countries with a high number of daily cases because the increase in the cumulative number of cases is relatively steady.

For countries with very few cases or irregular patterns in daily data, causing cumulative data to have a jagged shape, choosing a polynomial function with low degree to describe it is hardly possible. On the other hand, choosing a too large degree will lead to overfit, decreasing the model's generalizability. This is the model's major weakness.
Another drawback is tuning the number of degrees; however, this is not the major problem, as there it is the only hyperparameter, and tuning can be performed automatically, as long as we select the most appropriate loss function.

4. COVID-19 Impact to Economy

4.1. Overview

Since its emergence in Asia in late 2019, the coronavirus COVID-19 pandemic has been devastating. The virus spread to most countries, causing severe respiratory infections and many human casualties. The virus also put half of the world's population in lockdown which resulted in a slowdown of the world economy and a fall in stock prices.

The goal of this report is to introduce the steps for collecting and analyzing stock data in the context of the coronavirus pandemic.

4.1.1. The case study:

In this report, we will focus on the S&P 500 companies. We will start by collecting the following data :

- Stock prices in 3 different dates (January 1st, April 1st and August 1st)
- Number of outstanding shares for each company
- Industry/Sector where the companies operate (following the GICS classification)

4.1.2. Collecting and Storing Stocks Data:

We will be using the following services and libraries to collect and analyze the data:

- Wikipedia: We will use this page to get the list of S&P 500 companies.
- Google Finance: Google Finance is a website focusing on business news and financial information hosted by Google. Google Finance doesn't have an API that we can use directly in Python, but it can be accessed from Google Sheets using a formula called GOOGLE FINANCE. We will use Python to write down the GOOGLEFINANCE formulas.

4.1.3. Analyzing the stock data:

We got 505 companies in our list, and not 500... This is because some companies have a dual-class stock structure and are listed more than once in the list.

	Symbol	Security	SEC filings	GICS Sector	GICS Sub Industry	Headquarters Location	Date first added	CIK	Founded
0	MMM	3M Company	reports	Industrials	Industrial Conglomerates	St. Paul, Minnesota	1976-08-09	66740	1902
1	ABT	Abbott Laboratories	reports	Health Care	Health Care Equipment	North Chicago, Illinois	1964-03-31	1800	1888
2	ABBV	AbbVie Inc.	reports	Health Care	Pharmaceuticals	North Chicago, Illinois	2012-12-31	1551152	2013 (1888)
3	ABMD	ABIOMED Inc	reports	Health Care	Health Care Equipment	Danvers, Massachusetts	2018-05-31	815094	1981
4	ACN	Accenture plc	reports	Information Technology	IT Consulting & Other Services	Dublin, Ireland	2011-07-06	1467373	1989

The most important data that we need is:

- Symbol: Stock Symbol
- Security: Name of the company
- GICS Sector: Sector where the company operates following the Global Industry Classification Standard (GICS).
- GICS Sub Industry: Sub industry where the company operates following the Global Industry Classification Standard (GICS).

We can also check the number of companies by sector:

```
Industrials               71
Information Technology    71
Financials                66
Consumer Discretionary    64
Health Care               60
Consumer Staples          33
Real Estate               31
Utilities                 28
Materials                 28
Energy                    27
Communication Services    26
Name: GICS Sector, dtype: int64
```

And sub industry:

```
Health Care Equipment       19
Electric Utilities          13
Semiconductors              13
Industrial Machinery        13
Packaged Foods & Meats      12
                            ..
Real Estate Services         1
Motorcycle Manufacturers     1
Multi-Sector Holdings        1
Hotel & Resort REITs         1
Drug Retail                  1
Name: GICS Sub Industry, Length: 128, dtype: int64
```

We start by adding stock prices on 3 different dates: January 1st, April 1st and August 1st.

- January 1st, 2020 is the first date of the year. We want to have this price in order to calculate the price drop since the beginning of 2020.
- April 1st, 2020 is the date when the S&P 500 reached the bottom in 2020.
- August 1st, 2020 is the last date when the stock market was open.

Change in the total market cap of the S&P 500: The S&P 500 lost **6,6 trillion USD** from January 1st to April 1st, but it got back **7.1 trillion** from April 1st to August 1st.

Change in the total market cap by sector:

```
GICS Sector
Financials               -1299.816369
Information Technology   -1148.222922
Industrials               -781.125024
Consumer Discretionary    -656.135912
Health Care               -634.998289
Energy                    -628.265643
Communication Services    -617.054719
Consumer Staples          -298.397165
Materials                 -209.919810
Real Estate               -190.017675
Utilities                 -163.171664
dtype: float64
```

At the 2020 bottom of the S&P 500 (April 1st), the **Information Technology** and **Financials** sectors had the largest drop in total market cap (compared to January 1st) with 1.3 trillion and 1.1 trillion respectively.

```
GICS Sector
Financials                    -788.490849
Energy                        -475.908582
Industrials                   -300.712617
Real Estate                    -52.471633
Utilities                      -50.495636
Materials                       -9.537168
Consumer Staples                 7.746028
Communication Services         150.732835
Health Care                    199.600057
Consumer Discretionary         453.682302
Information Technology        1373.723367
dtype: float64
```

As of August 1st, we can see that the **Financials** and **Energy** sector had the largest drop in total market cap compared to January 1st.

Ranking of companies by percentage change of stock prices

	Security	PercentageChange_1_8_1_1
348	Norwegian Cruise Line Holdings	-77.800442
87	Carnival Corp.	-74.527383
130	Coty. Inc	-66.696833
406	Royal Caribbean Group	-64.805050
459	United Airlines Holdings	-64.486294

We can see from the table above that the companies that are hardest hit are the 2 major cruises companies: Norwegian Cruise Line Holdings, Carnival Corp., Royal Caribbean Cruises Ltd. Most of the companies operate in the tourism industry. These companies saw drops in their stock price of over 70%.

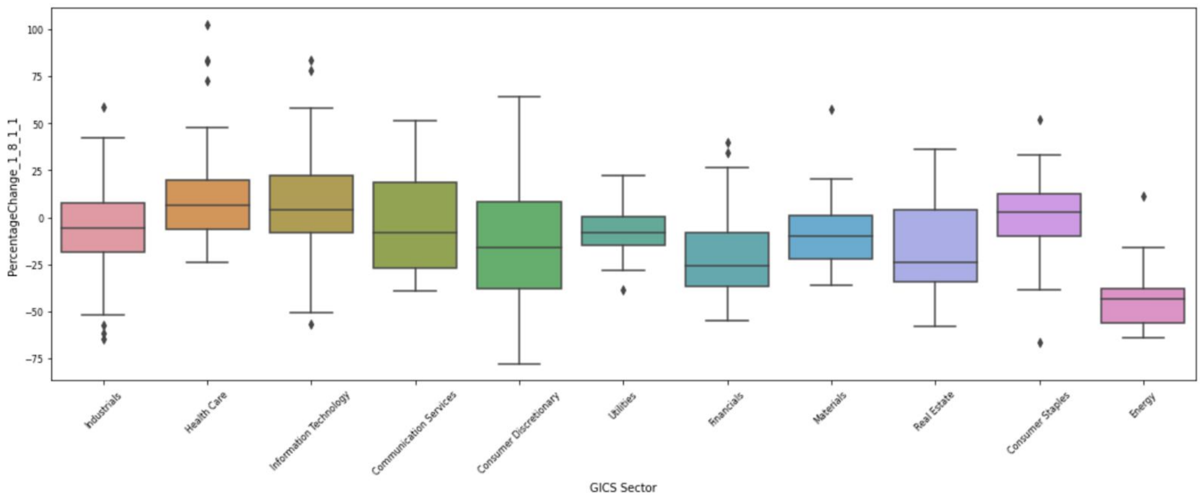
Only 311 stocks from the 505 saw positive growth from January 1st to August 1st.

Percentage Change of stock priced by sector

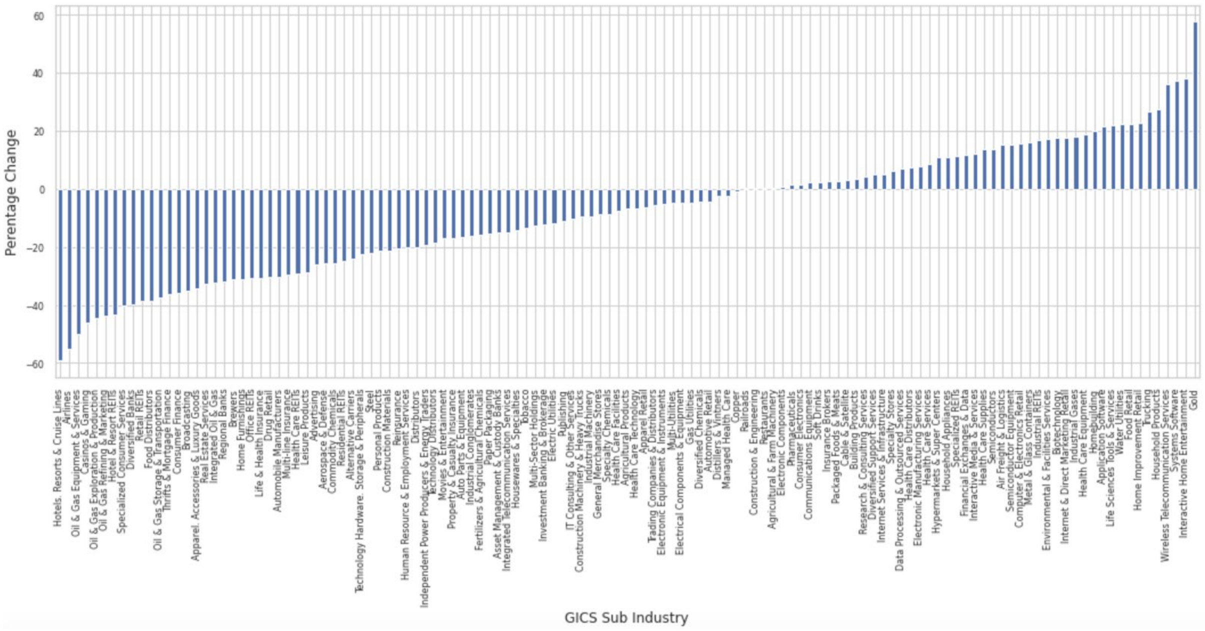
```
GICS Sector
Energy                    -43.276168
Financials                -19.505009
Real Estate               -15.835272
Consumer Discretionary    -13.106663
Industrials                -8.401737
Utilities                  -8.136877
Materials                  -7.499641
Communication Services     -3.480520
Consumer Staples           -0.232584
Information Technology      6.564060
Health Care                11.453025
Name: PercentageChange_1_8_1_1, dtype: float64
```

We can see that **the energy sector** was the hardest hit with a 43% average drop in stock prices.

Below we can visualize a boxplot of the 11 sectors' percentage change in stock prices from January 1st to August 1st.



In sub industry:



If we look at the average percentage change in stock prices by Sub Industry; we can see that the **travel related industries**, **Oil & Gas** and **Airlines** were the hardest hit. 75 of the 125 Sub Industries had their average stock price declined from January 1st to August 1st.

4.1.4. Summary

We can see that the travel related industries, Oil & Gas and Department Stores were the hardest hit.

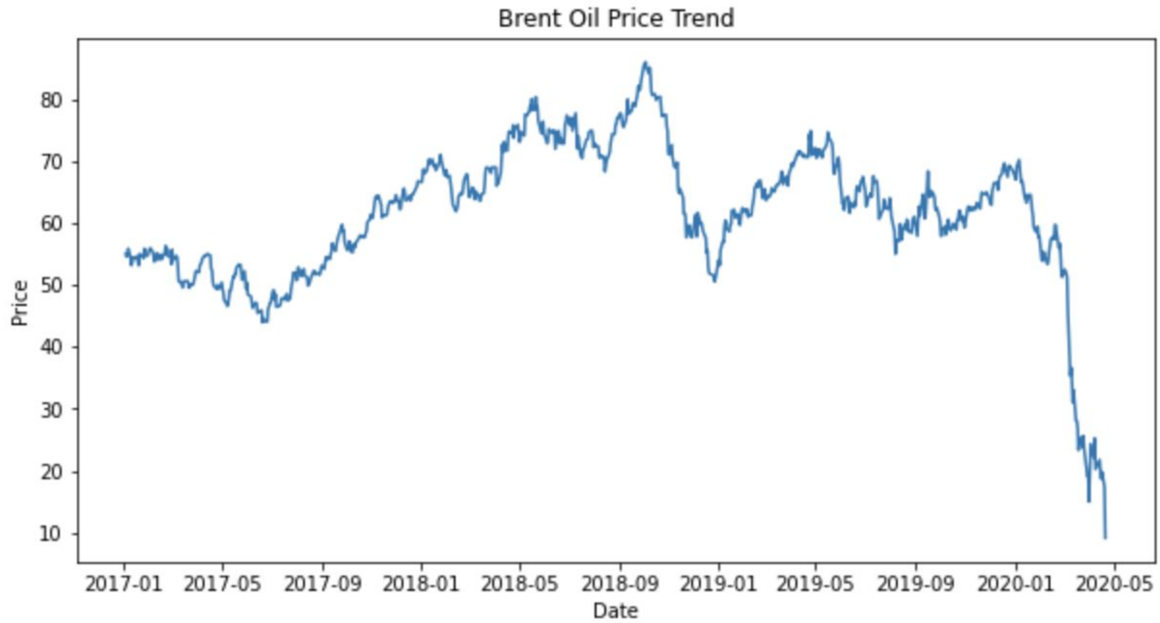
Only in Information Technology, Health Care is growing. Detail is: gold, interactive home entertainment, Telecommunication service, Food Retail, …

4.2. The hardest hit industry:

4.2.1. Impact of COVID-19 on oil price:

Global energy market demand, such as for oil, natural gas, and coal, is declining as the impacts of COVID-19 spread around the world. This decline in oil demand has been particularly significant as oil is mainly used in the transportation sector, and business closures, declines in domestic and international travel, and the lockdowns and quarantines in many countries have all shrunk the demand for oil. A sharp decline in domestic consumption and a possible decline in new investments, declines in tourism and business travel, spillovers of weaker demand to other sectors and economies through trade and production linkages, supply-side disruptions to production and trade, and shifts in health care expenditure are only some of the channels through which the pandemic is affecting the demand side in the oil market.

Global oil demand has contracted for the first time since the global recession of 2009.



On 5 March 2020, the Organization of the Petroleum Exporting Countries (OPEC) proposed a 1.5 mbd production cut for Q2 2020, of which 1 mbd would be by OPEC countries and 0.5 mbd from non-OPEC but aligned producers (most prominently, the Russian Federation). The following day, the Russian Federation rejected the proposal, and Saudi Arabia also announced unprecedented discounts of almost 20% in key markets. The result was a more than 30% plunge in prices to as low as $31.1 (WTI crude) per barrel on 9 March, and the crisis in the oil industry has continued to worsen. An intensifying recession due to COVID-19 fears drove global oil prices even further down to $11.57 per barrel (WTI crude) on 21 April.

When global oil (energy) prices go down sharply (when a negative price shock happens), like during the current oil price drop, a direct negative effect on the gross domestic product of a net oil exporter is expected due to a decrease in oil revenues.

The demand of oil importers is shrinking as non-essential businesses and services have been shut down to halt the spread of the virus in affected countries across the globe and government officials have been urging their citizens to stay home. These measures have effectively stopped economic activity, shrinking the demand for oil, especially in the transportation sector such as: airlines, …

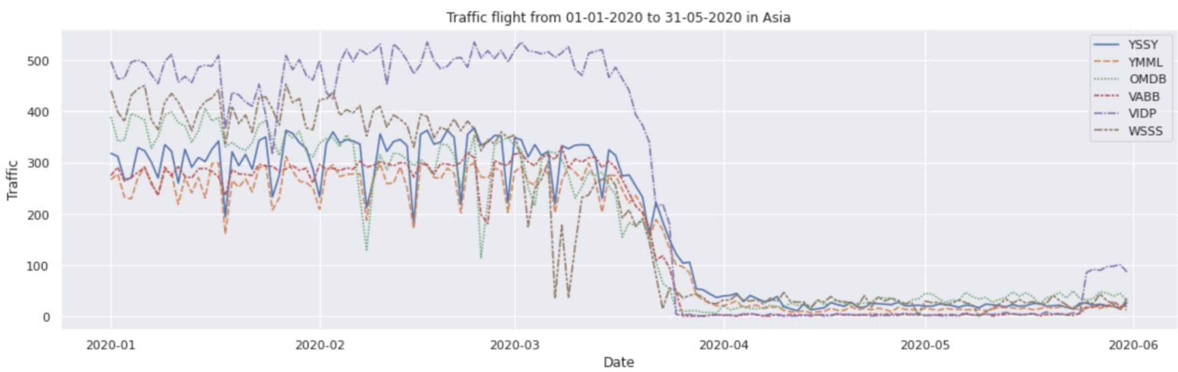
In Viet Nam:

Gasoline Prices in Vietnam remained unchanged at 0.47 USD/Liter in April from 0.8 USD/Liter in July of 2020. Because government apply work from home policy in March 2020

4.2.2. Impact of COVID-19 on worldwide aviation

The pandemic of coronavirus is having a serious impact on aviation around the world. The slowdown appears on data, with some regional peculiarities.

The following plot displays the current trend in the number of departing aircraft from airports in Asia (covered by The OpenSky Network).



The trend showed:

- A slow decrease from February in Asian airports (an early one in Hong-Kong);
- Airports plummeting since early day of March;
- India almost stopped all traffic (VABB, VIDP).

Traffic airlines decrease because some country in Asia lockdown their country:

- HongKong lockdown in 5/2/2020
- India lockdown 24/3/2020
- Australia lockdown 20/3/2020
- Viet Nam lockdown 1/4/2020

Significant reductions in passenger numbers have resulted in flights being cancelled or planes flying empty between airports, which in turn massively reduced revenues for airlines and forced many airlines to lay off employees or declare bankruptcy. Airliner manufacturers and airport operators have also laid off employees. According to some commentators,the ensuing crisis is the worst ever encountered in the history of the aviation industry.

4.2.3. Impact of COVID-19 on tourist:

The COVID-19 pandemic has caused a 22% fall in international tourist arrivals during the first quarter of 2020, the latest data from the World Tourism Organization (UNWTO) shows. According to the United Nations specialized agency, the crisis could lead to an annual decline of between 60% and 80% when compared with 2019 figures.

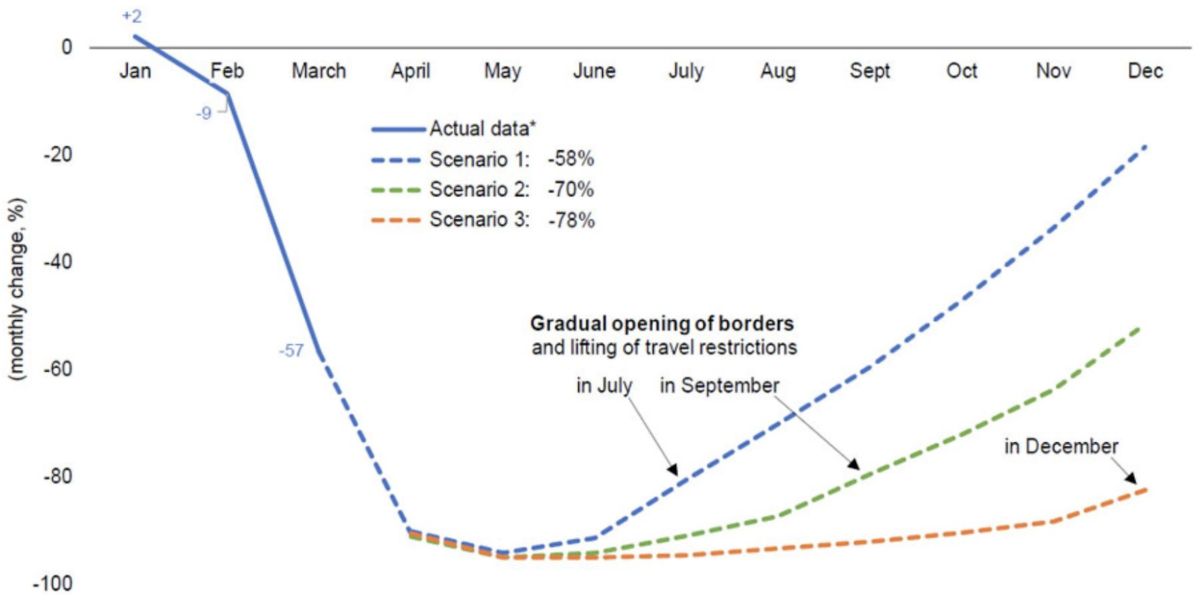
Arrivals in March dropped sharply by 57% following the start of a lockdown in many countries, as well as the widespread introduction of travel restrictions and the closure of airports and national borders. This translates into a loss of 67 million international arrivals and about US$80 billion in receipts (exports from tourism).

Predict International Tourism 2020 Scenarios:

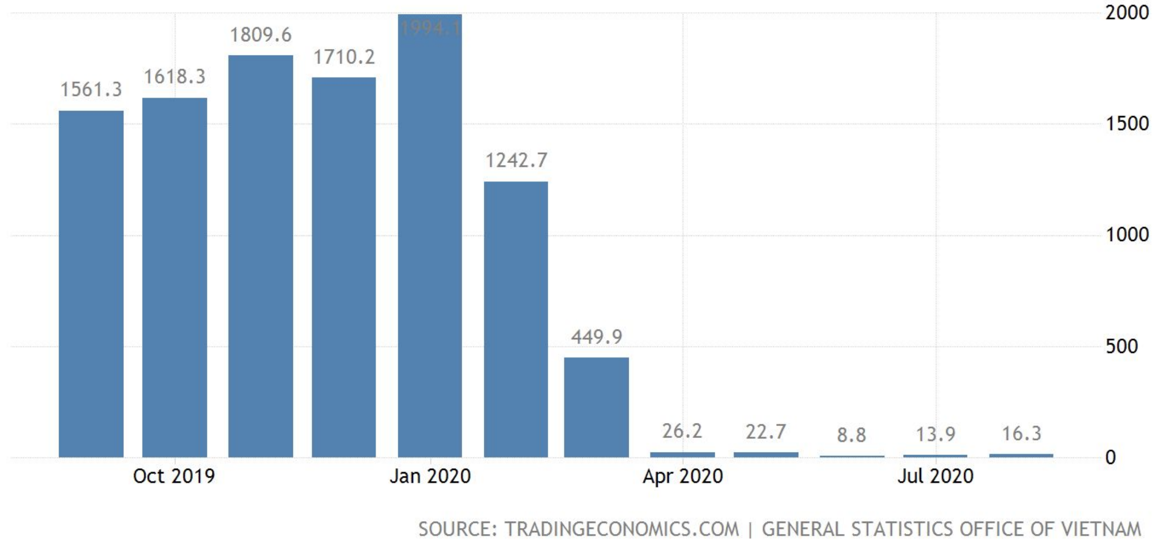
Current scenarios point to possible declines in arrivals of 58% to 78% for the year. These depend on the speed of containment and the duration of travel restrictions and shutdown of borders. The following scenarios for 2020 are based on three possible dates for the gradual opening up of international borders.

- **Scenario 1** (-58%) based on the gradual opening of international borders and easing of travel restrictions in early July
- **Scenario 2** (-70%) based on the gradual opening of international borders and easing of travel restrictions in early September
- **Scenario 3** (-78%) based on the gradual opening of international borders and easing of travel restrictions only in early December.

International tourist arrivals in 2020: three scenarios (YoY monthly change, %)



In Viet Nam:



International arrivals to Vietnam plunged 98.9 percent year-on-year to 16.3 thousand in August of 2020, the same as in the prior month, amid the prolonged impact of coronavirus cases, with the government still not opening up international tourism yet. Visitors fell from Asia (-98.8 percent), mostly China (-98.7 percent), South Korea (-99.2 percent), Japan (-99.8 percent); America (-99.6 percent), namely the US (-99.8 percent); Europe (-99.4 percent), of which Russia (-99.8 percent), the UK (-99.3 percent), France (-99.8 percent), and Germany (-99.2 percent); and Australia (-99.7 percent). Considering the first eight months of the year, tourist arrivals plunged 66.6 percent from the same period of 2019.

4.3. Some industry positive growth:

4.3.1. Top company Information Technology:

Apple (AAPL) saw its shares jump into record high territory after its June-quarter earnings report. The consumer electronics giant surged to a market capitalization of over $2 trillion.

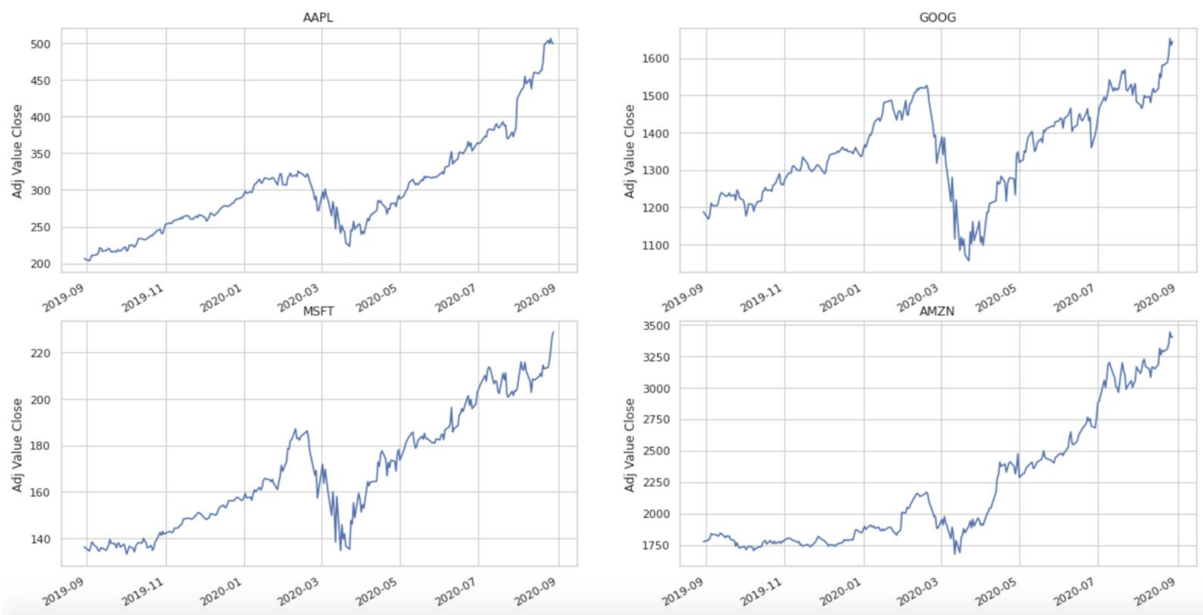
Lately, two businesses have given Apple's sales and profits a boost: services and wearables.

In the June quarter, Apple's services revenue rose 15% to $13.16 billion. Services include App Store, AppleCare, iCloud, Apple Pay, Apple Music, Apple TV+, Apple Arcade and other offerings.

Meanwhile, Apple's Wearables, Home and Accessories unit saw sales jump 17% to $6.45 billion in the June quarter. This unit includes wearables like the Apple Watch, AirPods wireless earbuds and Beats headphones. It also contains the Apple HomePod wireless speaker and other miscellaneous gadgets.

In the June quarter, iPhone sales climbed 2% to $26.42 billion. The company's new lower-cost iPhone SE gave sales a lift in the quarter.

Mac computer sales jumped 22% to $7.08 billion in the June quarter. IPad sales leapt 31% to $6.58 billion. The work-from-home and school-at-home trends driven by the Covid-19 pandemic fueled the gains in both product categories.



With work from home policy some company has positive grow up faster:

- Apple grow up to **2 trillion USD**
- Microsoft, Amazon, Alphabet(Google) grow up to **1.6 trillion USD**

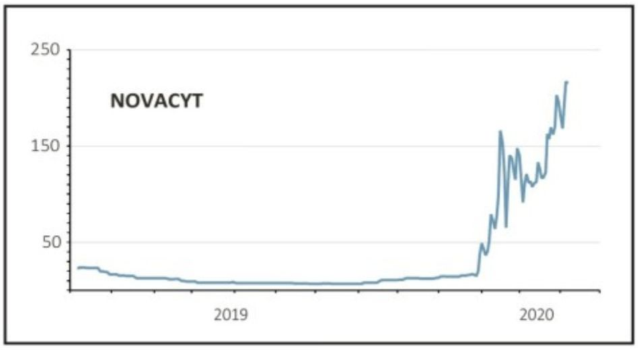
4.3.2. Top HealthCare companies:

Health care, one of the largest and most complex sectors, comprises a broad range of companies that sell medical products and services. The healthcare sector includes companies that sell drugs, medical devices, and insurance, as well as hospitals and health care providers. Some of the largest healthcare companies in the world include UnitedHealth Group Inc. (UNH), Pfizer Inc. (PFE), and Abbvie Inc. (ABBV). Several health care companies have received U.S. Emergency Use Authorization (EUA) to develop either tests or treatments for COVID-19, helping to fuel optimism about their shares.

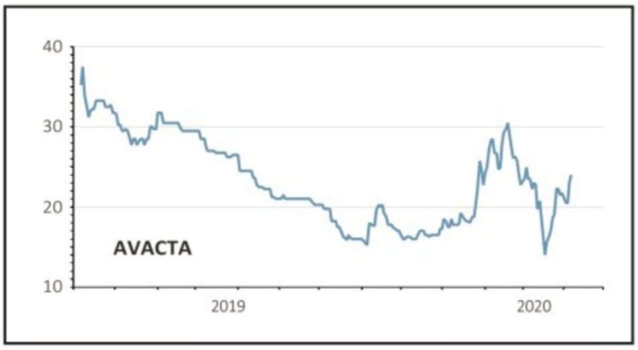
Healthcare companies are heroes, and invest in the companies trying to save the world from coronavirus.

The best performing shares over the last three months are healthcare companies. All of them are at the leading edge, developing testing kits, therapies and vaccines to combat the coronavirus pandemic.

Novacyt (NCYT:AIM) has developed a unique testing kit that gives results in under 30 minutes and is being used in hospitals and testing labs. The shares have risen 13-fold since 1 January.



Testing is fast becoming very competitive as a number of companies enter the space, such as UK biotherapeutics firm Avacta (AVCT:AIM) which on 8 April announced a collaboration with therapeutics company Cytiva to develop a rapid coronavirus test, pushing its shares up 69% on the day.



Fastest Growing Health Care Stocks

	Price ($)	Market Cap ($B)	EPS Growth (%)
Perrigo Co. PLC (PRGO)	53.09	7.2	528.6
Regeneron Pharmaceuticals Inc. (REGN)	616.89	65.6	353.0
Cardinal Health Inc. (CAH)	50.16	14.7	243.1

Perrigo Co. PLC: Perrigo is an Ireland-based health care company offering over-the-counter consumer goods and specialty pharmaceutical products. Its products include pharmaceuticals, infant formulas, nutritional products, active pharmaceutical ingredients, and pharmaceutical and medical diagnostic products. The company reported net income growth of 573.3% on net sales growth of 6.1% for Q2 2020, which ended June 27, 2020.5

Regeneron Pharmaceuticals Inc.: Regeneron Pharmaceuticals is a biopharmaceutical firm that discovers, develops, and markets treatments for a variety of serious medical conditions. Regeneron announced in early July that, under the U.S. government's Operation Warp Speed program, it had been awarded a $450 million contract by the Biomedical Advanced Research and Development Authority (BARDA) and the Department of Defense. Under the contract, Regeneron will manufacture and supply REGN-COV2, an antiviral antibody cocktail that could be used for the treatment of COVID-19 and potentially its prevention.6

Cardinal Health Inc.: Cardinal Health is a provider of healthcare services and products. Its services include pharmaceutical distribution, health care product manufacturing, distribution and consulting services, drug delivery systems development, pharmaceutical packaging, retail pharmacy franchising, and more.

5. Reference

[1] Yi-Cheng Chen. “A Time-dependent SIR model for COVID-19 with Undetectable Infected Persons”.
[2] Luca Magri and Nguyen Anh Khoa Doan. “First-principles Machine Learning for COVID-19 Modeling”.
[3] Adilmoujahid. “Analyzing the Impact of Coronavirus on the Stock Market using Python, Google Sheets and Google Finance”
[4] Eoin Kilbride. “COVID-19: Measuring Industry Impact With Media Signals”