Travel Advisories of Ten Popular Countries during COVID-19

Shuyao Wang[1*], Zifan Lin[2]

[1]*Eberly College of Science, Pennsylvania State University, State College, 16803, United State*
[2]*Faculty of finance, City University Of Macau, Macau, 999078, China*
Corresponding author Email: ginawsywork@gmail.com

ABSTRACT

With the outbreak of the coronavirus in 2019, almost all travelers had to cancel their travel plans over the world and follow most countries' "Stay at Home" policy. But fortunately, with the situation improved this year, 2021, an increasing number of travelers have a chance to go out. This research will help travelers analyze the situation of ten popular countries and give them advice before actually going there. The graphs which help illustrate the points include bar charts that show total conformed and fatal cases, pie charts that show the proportion of confirmed cases to fatal cases, and scatterplots of daily and weekly confirmed and fatal cases separately. Finally, the work also made the prediction of confirmed and fatal for the next two weeks. In the end, the work will put all the data together and analyze them to advise the countries that the summary recommended. However, even though some countries show a decreasing trend for confirmed cases, people should also think twice before making decisions to go there with other considerations the graphs show.

Keywords: *COVID-19, Ten popular countries for traveling, Analysis of confirmed cases, Analysis of fatal cases, Rstudio*

1. INTRODUCTION

COVID-19 is declared by WHO on March 11, 2020. The disease which first broke out in Wuhan, China become a worldwide epidemic later that year. Until now (August 6, 2021), nearly two hundred million cases are confirmed and about four million of them died[1]. This worldwide epidemic has greatly changed the world economy, including tourism. In most countries, people have been asked to follow the policy of "stay at home".

About one and a half years later, human beings are still fighting the virus. Fortunately, vaccines are invented to help protect people from the disease. More than four billion doses of vaccine have been administered[2]. As a result, some countries loosened the lock-down policy [3]. People started to choose a place and take a trip after long restrictions. However, different countries act differently toward the virus. People should have a way to know the situation of the places they want to go and decide where they should go.

Our aim of the report is to give some travel suggestions for those people. The report will analyze the status of COVID-19 in ten popular countries for tourists and predict the development of the epidemic in those countries. By doing this, searches are made for the data about popular countries, people would like to go to before COVID-19 on the official website of the World Tourism Organization and get the names of the ten popular countries. Then the COVID-19 data of those countries were got on Johns Hopkins Coronavirus from June 1, 2021, to July 26, 2021. With the help of R Studio[4], the paper compares the total confirmed cases and total fatal cases within those countries for this period and calculates the trend of cases for each country at this time. After that, the prediction of the progress of the epidemic and scientific advice for tourists who would like to choose a good place to go will be made according to those data. In section 2, the work will describe the data and how those data to process and merge. In section3, the work will present our exploration. This part includes the comparisons between countries in two different ways, the process of getting the cases' trends, and the method of getting further predictions in each country. In section 4, the results of all the countries' data will be given, and analysis will also be addressed here. In section 5, the work would like to make a conclusion about all the data

they work previously gathered and give suggestions.

2. ABOUT THE DATA

2.1 Sources

The research would like to give suggestions to the travelers about the countries they would like to visit for the next two weeks. An initial search on information about popular countries around the world before the COVID-19 is gathered through the Official website of the World Tourism Organization[5], and decide to focus on ten popular countries which are France, China, Germany, Italy, Mexico, Spain, Thailand, Turkey, the UK, and the US.

Then through the website(https://covid.observer)[6], the data for each country that the research needed was found. The website is reliable as the data is based on the collection of Johns Hopkins University Center for Systems Science[7]. To use the data easily, these ten datasets were combined into one dataset.

2.2 Choosing variables

There are 13 columns in the combined dataset. The first column tells the date of the data collected which starts from 2021/1/1 to 2021/7/26 for each country. The columns on the right also include the cumulative data of confirmed cases, fatal cases, and active and recover cases. There are also daily growth, recovery rate, mortality rate, and the affected population which records the proportion of the data. Moreover, the columns also include the data of confirmed and fatal cases per 1000 which excludes the effect of differences in population. The columns named "one confirmed per every" or "one fatal per every" mean that among how many people there have one confirmed or fatal case which indicates that the larger data, the better situation.

However, after carefully reviewing the data, it is surprised to find that there are possible flaws in this dataset. First of all, there are many NA values of the recovered cases and the recovered rate of the UK and US. Moreover, there are many NA values of Daily growth in Spain. It is reasonable to have a value of zero in daily growth. To deal with these problems, this paper decides to not include the variables of recovered cases and recovered rate in the analysis.

Considering the aspects the work wants to analyze, the work both analyzes the data horizontally and

vertically by comparison of the conditions between ten countries during the last two months and seeing the condition of each country, specifically which includes the trend of cases during the last two months. After that, the paper would like to predict each country's future condition for the next two weeks.

The work decides the variables it includes in the following report by taking these factors together. Firstly, for comparison between ten countries, the variables of 1 confirmed or fatal per every will not include in the report. Moreover, the variables which represent the rate will not be used because the work would like to predict the number of cases in the future. The columns for confirmed and fatal cases per 1000 exclude the effect of the difference in population.

Overall, the variables the work may use include: Date, Confirmed cases per 1000, Fatal cases per 1000, Total confirmed cases, Total fatal cases, Affected population, and Country.

2.3 Further process before dealing with the variables

As the prediction for the next two weeks needs the trend, the range of data that could help for prediction narrows down. The paper then creates a new data set that only contains the data of the latest two months from 2021/6/1 to 2021/7/21 which is data for eight weeks.

Moreover, to make the calculation easy and see the number of days the paper includes, the paper decides to turn the variable of Date into a real date in real life so that the trend in the plot between time and number included could be seen.

As the conflict between the cumulative data of confirmed and fatal cases the dataset given and the daily confirmed and fatal cases the research really need exists, new columns were created to show the data of yesterday. Then daily cases could be found by subtracting the cumulative data today and yesterday.

3. EXPLORE THE DATA

3.1 Comparison between ten countries.

As shown in figure1 and figure 2 above, the first two bar plots illustrate the total confirmed cases per 1000 and the total fatal cases per 1000 in ten different countries from June 1st to July 26th, 2021.
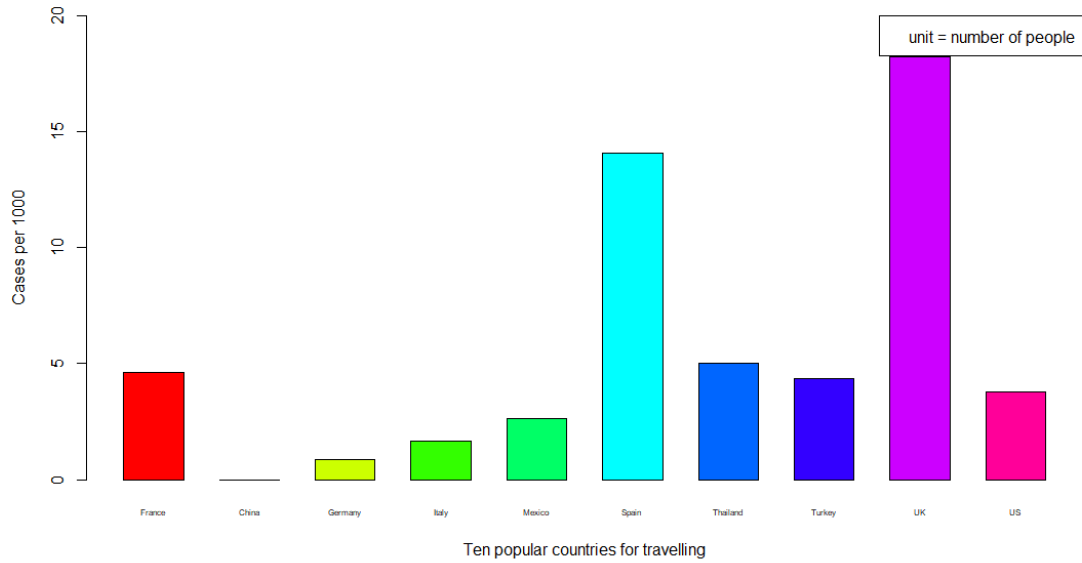
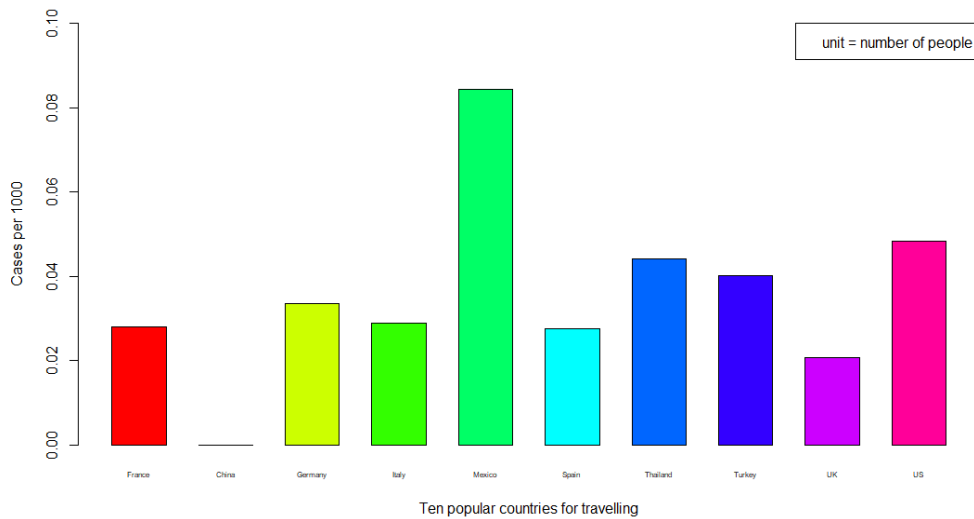**Figure 1** Total confirmed cases per 1000 from 2021.6.1 to 2021.7.26

Figure 2 Total fatal cases per 1000 from 2021.6.1 to 2021.7.26

On the graph, Mexico, Germany, and Italy have a similar situation. The confirmed cases per 1000 are relatively low and the fatal cases per 1000 are relatively high. This may show that the coronavirus is still not very constant in these countries. Out of concern for the safety of the travelers, they are not highly recommended to be visited. In contrast, Spain and the UK, both countries have a high rank in total confirmed cases but a low rank in total fatal cases, which means that these countries are more likely to be recovered. Visitors need to pay attention to following the suggestion of the hospital to protect themselves if they decide to go there.

Furthermore, in Thailand, Turkey, the US, and France, these four countries have about 4 people confirmed in 1000 and the total fatal cases per 1000 are 0.044, 0.040,0.048, and 0.028 respectively. So, these countries are not very safe. Visitors should visit these places after vaccination if they really want to go there. Moreover, there is a very different situation for countries

like China. Both confirmed cases and fatal cases per 1000 are very low since it only has a small number of active cases. Therefore, China is more suitable for travelers to visit according to these data, but they still need to take care of it.

3.2 Conditions of each country

3.2.1 The trend of each country's cases for the last two months

Next, the trend of these two months can be seen through the scatterplot of daily cases. To get the graph, the numbers of days that are considered in the following process through the scatterplot of daily cases are calculated. It turns out to be fifty-six days (or eight weeks) of daily cases will be used. In this part, the paper investigates daily confirmed cases and daily fatal cases separately. In each country, R Studio is used to plot a graph that shows the relationship between the

development of daily cases and the number of days and make a regression line. Figure 3 and Figure 4 show the scatterplot and the regression line of daily confirmed and fatal cases in France as an example.
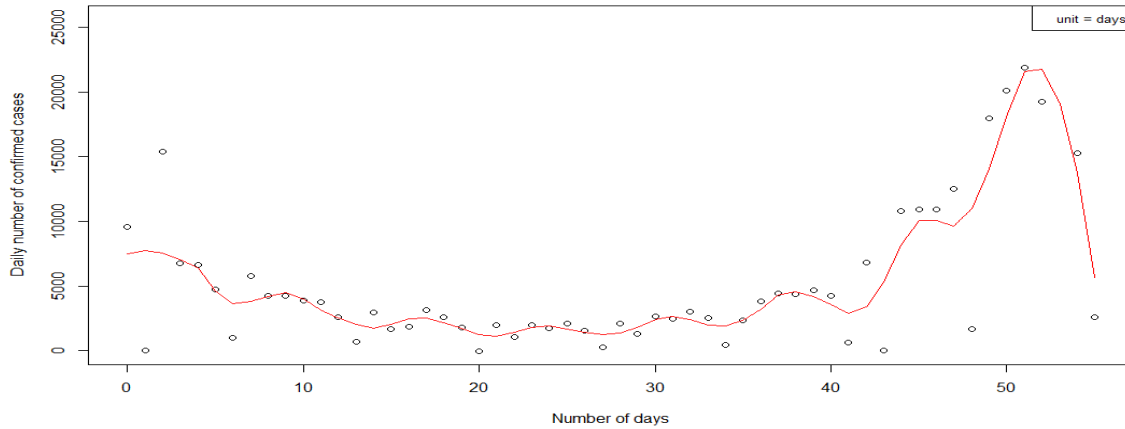

Figure 3 Daily confirmed cases in France from 2021.6.1 to 2021.7.26
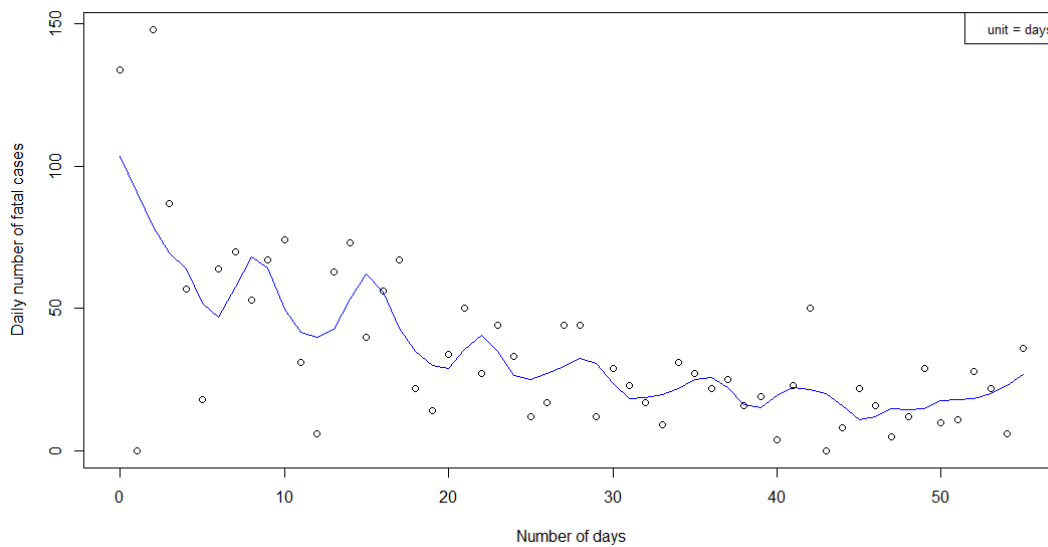

Figure 4 Daily fatal cases in France from 2021.6.1 to 2021.7.26

It is surprising to find that two days at the end of one week will often have relatively fewer cases. This weird pattern may disrupt the result of the prediction like the graph here for the confirmed cases for France. The reason why this pattern exists is that hospitals in Europe and America often close on weekends so that a much lower number of cases are counted on those days. It means that further consideration should be made to the data.

No matter what, the graphs with the loess model with regression lines are fit. The development for daily confirmed cases and daily fatal cases are shown in all of the graphs. Some rough predictions on the development of daily confirmed cases and daily fatal cases in a few

weeks through the graphs can be made. However, to solve the problems above, it is important to find a better way to treat the data instead of using the number of daily cases.

3.2.2 Turn the data into weekly formal

In the previous section, a noticeable problem with the daily data is demonstrated. Since the fluctuations within a week might strongly affect the trend analysis, the paper then decided to combine the daily data into weekly ones so that the fluctuations and noises could be smoothed. Figure 5 is the scatterplots for confirmed cases in each weak after making this adjustment.
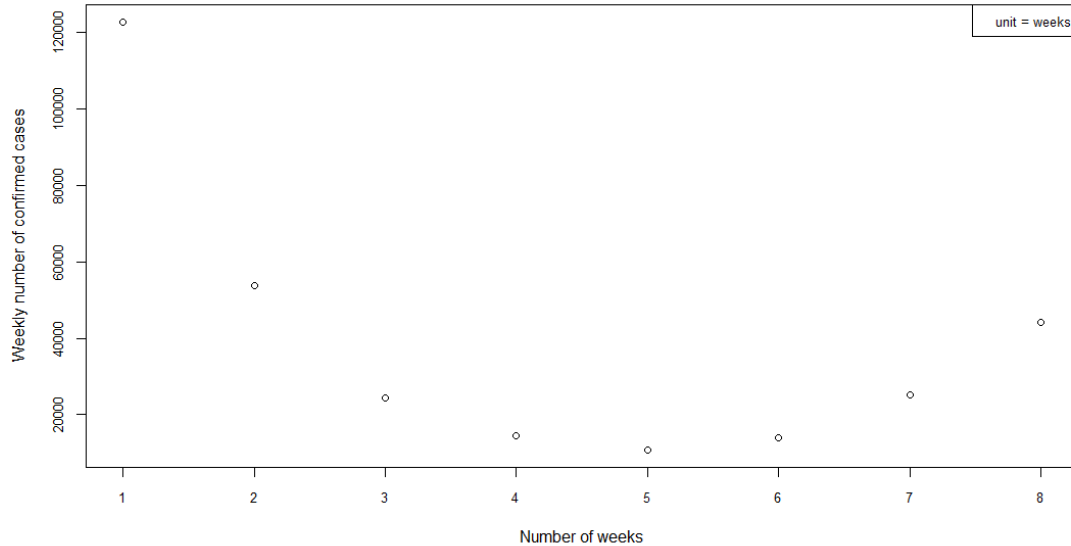
Figure 5 Weekly confirmed cases in France from 2021.6.1 to 2021.7.26

3.2.3 Prediction for the next two weeks

To predict the cases for the next two weeks, Table 1 shows two methods that could help with the prediction. One is the loess method which pays attention to the tail of the trend. The other is by calculating the linear regression line between the last two points and deducing the next two points by these lines. Table1 shows the confirmed and fatal regression for each country by using two ways of regression.

Table 1 Loess method and linear regression model for the prediction

	Confirmed Loess	Confirmed Regression	Fatal Loess	Fatal Regression
US	3008	80439	-21	-57
France	2136	39890	3	52
China	2	61	0	0
Italy	28	4638	2	33
Spain	-9026	-96599	4	37
UK	-13928	-162938	1	53
Mexico	-549	-45	39	673
Thailand	869	13016	-4	-23
Turkey	742	17672	1	60
Germany	-239	-1236	7	97

It is obvious that most of the countries' residuals obtained by using the loess model are closer to zero, apart from the confirmed cases in the UK and the fatal cases in China. Therefore the paper plans to predict the future by using the loess model instead of regression lines.

4. RESULT

4.1 Trend analysis

To produce a generally illustrating trend for ten countries' epidemic trend, the paper produced some new graphs that are in the count of weeks instead of days in the shiny App[8] on the R Studio. The data the paper combined was still from June 1st to July 26th, 2021. The paper selected the data within a week during this period and yielded a composite number of death numbers and confirmed cases from our cases. dataset for further study purposes. Figure 6 and Figure 7 show the results we get for France of their confirmed cases and their fatal cases as an example. Other countries follow exactly the same process.
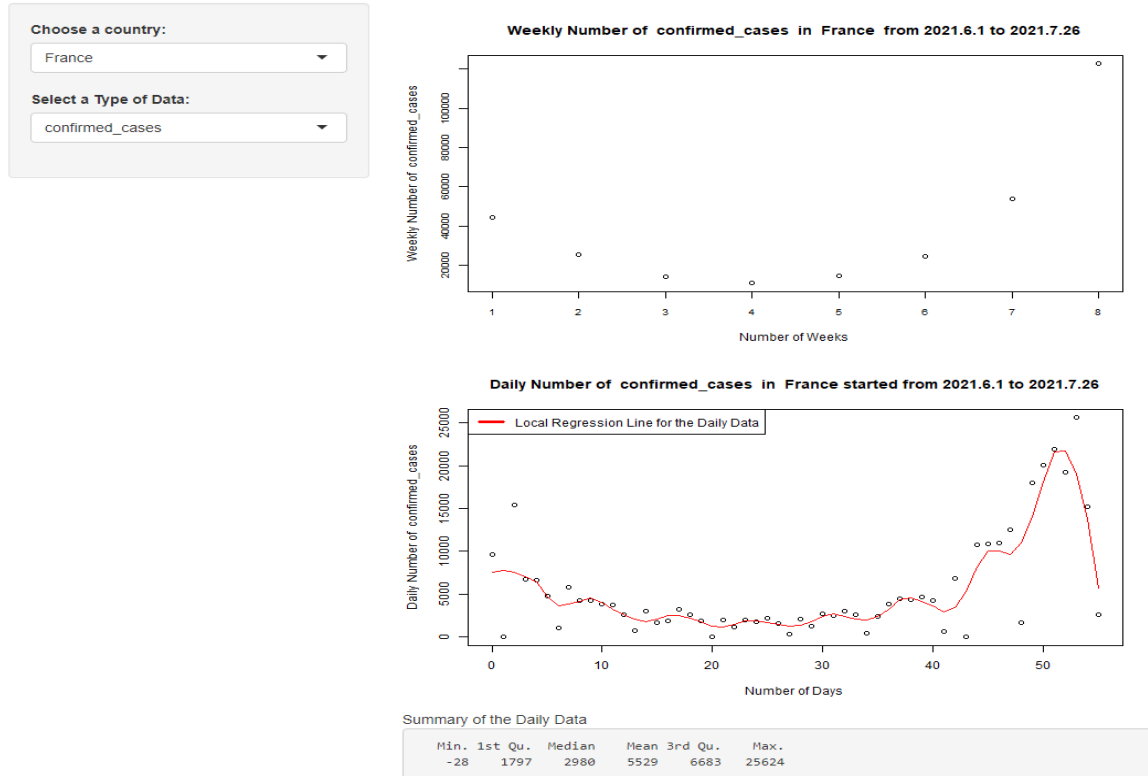
Figure 6 Weekly and daily confirmed cases in France from 2021.6.1 to 2021.7.26
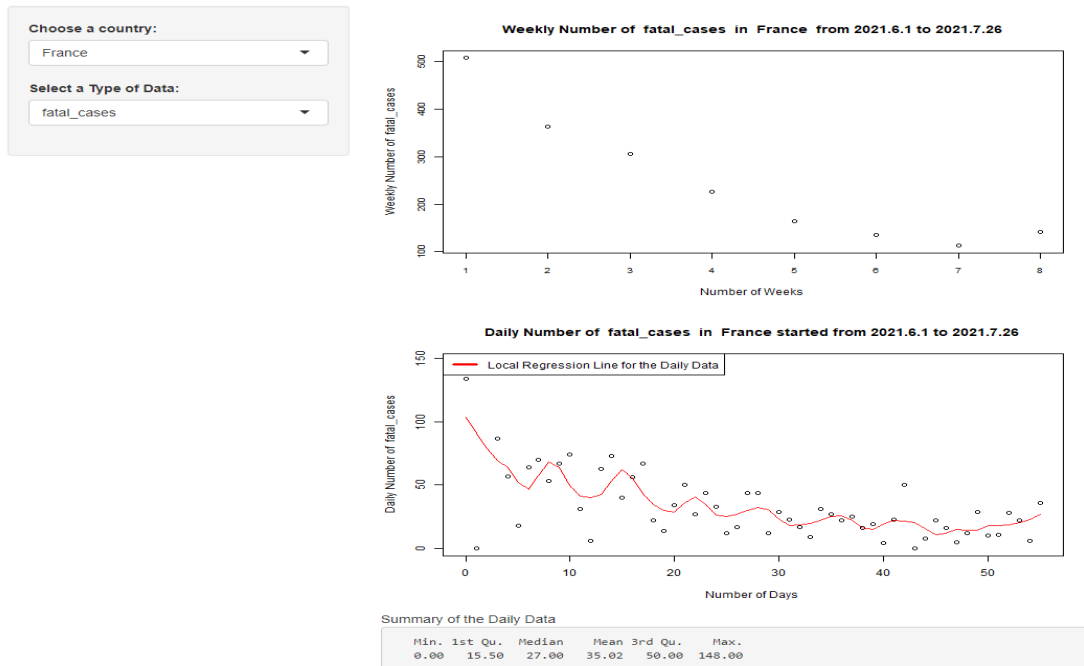

Figure 7 Weekly and daily fatal cases in France from 2021.6.1 to 2021.7.26

In the graphs of China, they show that China had good control in the first four weeks during the selected period since it has a stable count of around 150 to 200 cases. However, the weekly confirmed cases in the later four weeks rose to over 350 cases. Back to the fatal cases in China, the number was plain and seldom fluctuation except for the two deaths reported in the fourth week of the selected period.

In France, the number of new confirmed cases concerning its population showed that pandemic France was not promising, since the increase in confirmed cases was even steeper. Besides, the weekly new cases in France rose strikingly high to over 120,000 cases.

However, the death cases in France showed a good, decreasing scene which could partially be good news.

For Germany, the fluctuation appeared as well, which included a bottom confirmed cases in the fifth week. However, the bounce was still obvious. When it comes to the fatal cases in Germany, the number is decreasing at a relatively fixed rate and fell from near 800 to 200 deaths below.

Italy, as one of the most popular trip destinations in Europe, showed a similar trend as the former two Europe countries: a surge in confirmed cases during the fourth and fifth week then turned to ascend quickly again. The deaths in Italy descended steadily during the whole selected weeks which decreased from over four hundred to near a hundred.

In Mexico, the confirmed cases seemed to have increased drastically during our observed eight weeks. However, Mexico's death cases were noticeable compared with its active cases in the country. What's more, the fatal cases dropped to a bottom range which is between one to two thousand.

Spain, as a famous coastal country in Europe, showed a worrying trend with confirmed cases. The confirmed cases in Spain had climbed up quickly during the last five weeks: from 50,000 below to terribly over 150,000. However, the death cases in Spain were unstable since it has quite strong fluctuations between weeks, and their death cased decreased to near one hundred in the last two weeks. In this way, Spain showed a similar trend to Mexico in new cases, but still, not a safe place considering its not-as-large population.

For Thailand, the trends in the two kinds of data were all terrifying since the deaths and new cases increased at a strikingly high rate. Till the end of our observation, the weekly confirmed cases were near a hundred thousand.

In Turkey, the trend for new cases and deaths was similar to that of Italy. Turkey's confirmed cases were still high at the end and bounced after the fifth week, as the deaths continued to decrease during the end of the observation. Turkey's epidemic figures were not nice-looking, but the trend in new cases was worth noticing.

In the UK, both the fatality and new cases came to become larger during the whole observation. In the eighth month, the weekly confirmed cases were near 250,000 and the weekly death cases were over 400. These gave good news for the UK's pandemic situation, which could help travelers to be calm.

United States' figures were not optimistic in new cases. From the graph, the paper could heed that the cases were growing with an acceleration. Nevertheless, the US-controlled deaths are stable in the range from 1,500 to 2,000.

For countries in Europe, the similarity in the shape of their confirmed cases trend might be explained by the sharing of boundaries since they are quite close to each other in both geography and humanity.

4.2 Prediction analysis

As the trends mentioned in the perspective weeks above, the work then started to predict the ten countries' confirmed and fatal cases in the next week. To generate the expected values of confirmed cases and fatal cases in ten countries, the paper used the loess method in R studio. Then, two plots with ten countries' figures altogether are graphed to make a further comparison. Figure 8 puts the proportion of confirmed cases to the population of ten countries together. Figure 9 puts the proportion of fatal cases to confirmed cases for ten countries together.
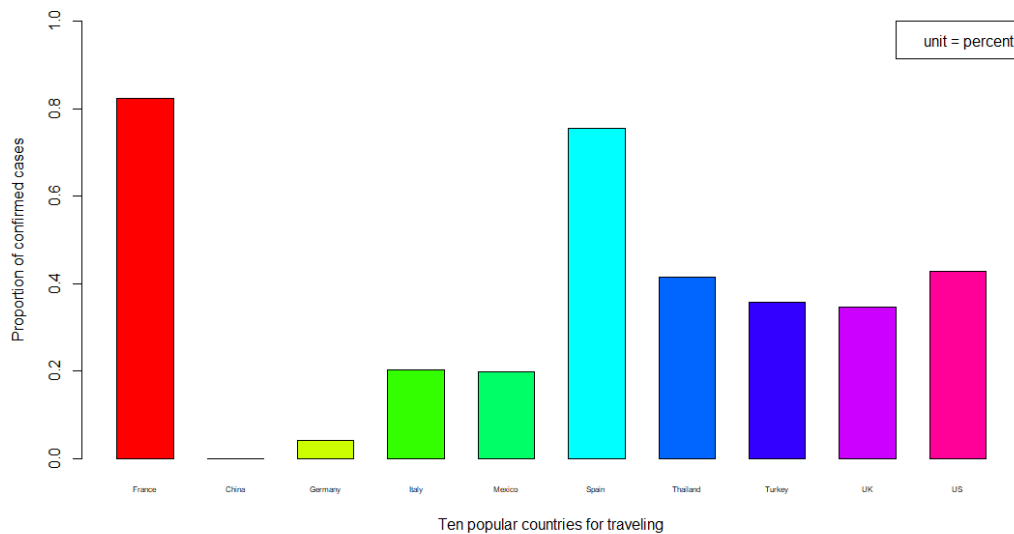

Figure 8 The proportion of confirmed cases in the population from 2021.7.27 to 2021.8
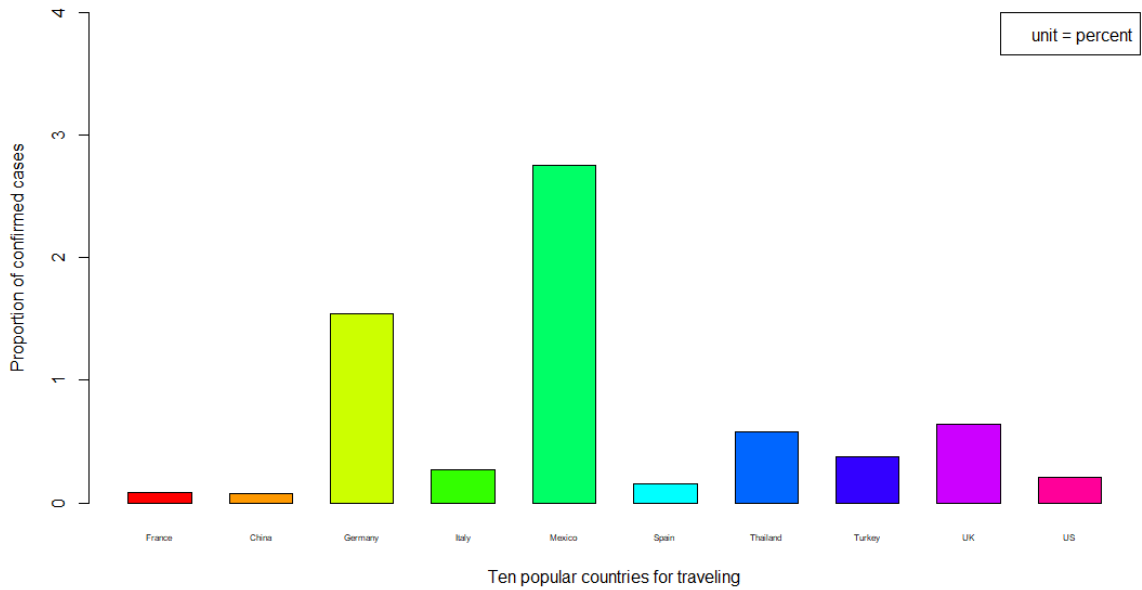
Figure 9 The proportion of fatal cases to the confirmed cases from 2021.7.27 to 2021.8.9

From the graph of confirmed cases, the result shows that China and Germany have relatively smaller proportions of confirmed cases within their own population, and countries like France and Spain have larger proportions. However, in the graph of fatal cases, countries like Germany and Mexico have the largest ratio of fatal cases within the group of confirmed cases which is much bigger than the other countries. The previous paragraph had shown a good analysis of our expectation counts.

The previous paragraph had shown a good analysis of our expectation counts. Nevertheless, some further investigations on extreme values on those graphs such as the high fatality ratio in Mexico and Germany are done. In order to explain these high fatality ratios, the paper picked several countries and made pie charts to see if their figures' proportions could explain part of the story. From Figure 10 to Figure 13, four such plots have been made to investigate the proportion of confirmed cases and fatal cases in four different countries.
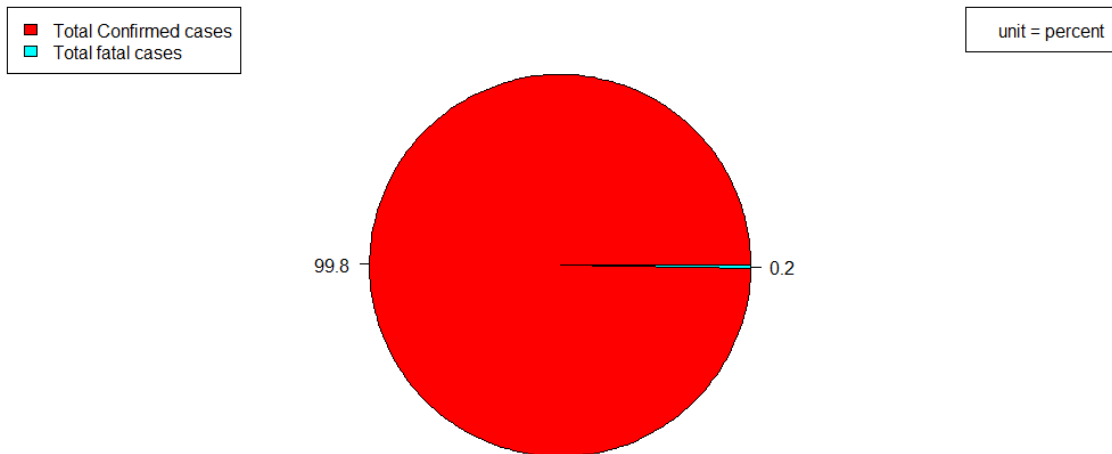

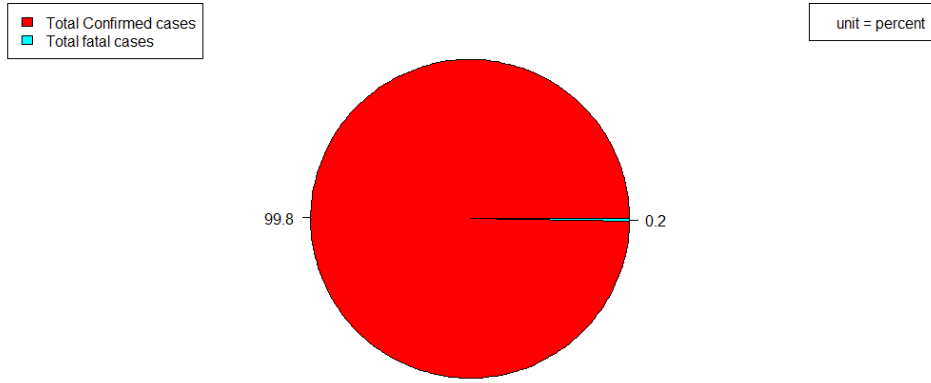**Figure 10** Confirmed cases and fatal cases in France from 2021.7.27 to 2021.8.9

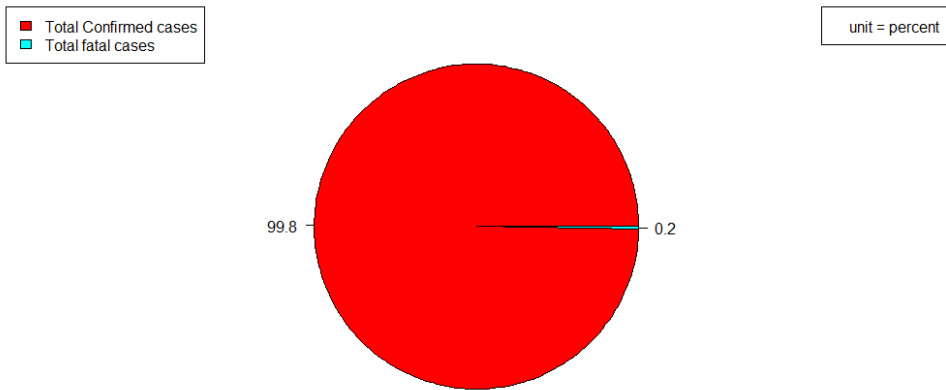**Figure 11** Confirmed cases and fatal cases in Spain from 2021.7.27 to 2021.8.9

Figure 12 Confirmed cases and fatal cases in the UK from 2021.7.27 to 2021.8.9
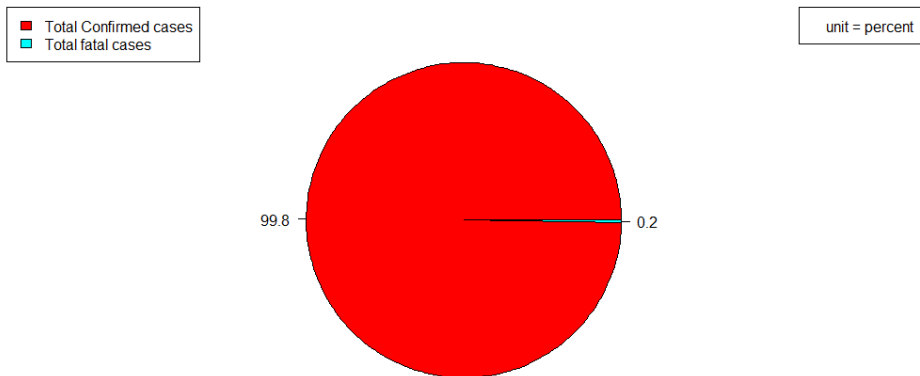

Figure 13 Confirmed cases and fatal cases in Mexico from 2021.7.27 to 2021.8.9

From the pie charts from the R Studio, they showed that the fatality ratios of our expected counts are generally small. Even the highest proportion in Mexico, which is 0.2%, is, in fact, a convincing figure.

In all, the paper has analyzed ten countries in two aspects: Confirmed cases and fatal cases. In these two sections, the paper discussed the trend that ten countries showed from June 1st to July 26th and finalized the predictions for all countries while picking some extreme values, and made the further investigation. The paper would then be able to safely rank the countries' predicted figures and give our end suggestion.

4.3 Summary

As the rank above, three safe countries are suggested for traveling. They are China, the UK, and the US. China stands out as the undeniable safe country due to its stable low counts in both confirmed cases and death cases. When it comes to the UK, though the figures still showed an unsatisfying picture of its current situation, the trend in the two kinds of data could tell a different story. Since the UK's counts decreased drastically in the last four weeks, it is reasonable to conclude that the epidemic circumstance may become easier in the following two weeks. The US, as a country that didn't do well in the very first period of the epidemic, showed an optimistic scene after having been through some pinpoint policies. Thus, these three countries came to the final recommendation list of traveling countries.

5. CONCLUSION

This research primarily wants to give people who would like to go traveling during the next two weeks some suggestions by analyzing the conditions of the coronavirus in ten popular countries. To make a prediction, our analysis mainly focuses on the last two week's data which started from June, 1st to July, 26th. To make a trend easily, the paper also creates a new variable that would record the daily cases for countries. In all, the paper analyzes the data horizontally by making a comparison between ten countries' confirmed and fatal cases and vertically by making scatterplots of cases throughout the timeline. Moreover, the paper turns the data into a weekly format to deal with unusual points that appeared on the weekend in Europe and America. Compared with the regression method, the loess method is used because it has a smaller residual value to predict the next two weeks' confirmed cases and fatal cases. Pie charts of these two variables in the countries that appear to have dangerous conditions are made to help illustrate the points.

Above all, this paper concludes that China, the UK, and the US are the best choices of destinations since they either have both small fatal cases and confirmed cases or display a large decreasing trend of cases during the last two months. In addition, they all present good prediction values by our model which show an optimal condition for the travel.

However, some limits exist in the dataset. The paper needs to pay attention to them and find the reasons as much as possible. For example, there is one negative value in the Chinese dataset. That may be due to the movement of cases from one city to another, or simply by the mistakes taken by the government. Moreover, there are also some possibilities that there are some cases that did not be recorded in the dataset, the traveler should think twice and check for the updated news of the countries they would like to go to before actually going there. Vaccines and masks should always be taken into consideration to protect one's own, too.

REFERENCES

[1] Johns Hopkins University & Medicine (2021): Deaths. Retrieved July 26, 2021. url:https://coronavirus.jhu.edu/map.html

[2] Johns Hopkins University & Medicine (2021). Vaccination Doses Administered. Retrieved July 26, 2021. url:https://coronavirus.jhu.edu/map.html

[3]Covid rules: What's changing in Wales, Scotland, England and Northern Ireland Retrieved July 26, 2021. url: https://www.bbc.com/news/explainers-52530518

[4]R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved July 26, 2021. url:https://www.Rproject.org/.

[5] UNWTO (2020): International Tourism Highlights, 2020 Edition,8. Retrieved July 25, 2021. url: https://www.e-unwto.org/doi/epdf/10.18111/9789284422456

[6] Andrew, S. COVID-19 World Statistics. Retrieved July 26, 2021. url:https://covid.observer/.

[7] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis. 20(5):533-534. Retrieved July 26, 2021. doi: 10.1016/S1473- 3099(20)30120-1.

[8] Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert and Barbara Borges (2021). shiny: Web Application Framework for R. R package version 1.6.0. Retrieved July 26, 2021. url:https://CRAN.R-project.org/package=shiny .