Development of a COVID-19 Data Tracking Platform for Public Health Safety

Yulin Wang

School of Computing Science

Sir Alwyn Williams Building

University of Glasgow

G12 8RZ

A dissertation presented in part fulfilment of the requirements of the Degree of Master of Science at the University of Glasgow

17 Dec 2021

Abstract

COVID-19 is a contagious disease caused by corona virus. The first known case was identified in December, 2019. The disease has since spread worldwide, leading to an ongoing pandemic, and it is one of the deadliest pandemics in history. Data analysis is a very useful way to find out the implied relationship in data, and give a solution or suggestion to a certain problem. Therefore, it is sensible to apply data analysis and set up a data platform to find out the trend of cases, to trigger the alarm and remind the relevant officers in the healthcare sector and the general public. In the process of development, many methodologies are involved. Statistics and machine learning are the key to prediction, Python programming is applied in web development, and data visualisation is also important to intuitively show the results.

The aim of the project is to build up a platform to analyse the COVID-19 data in different countries or regions, and to predict the possible cases in the next period of time. It reasonably analyses and organises the data, and finally shows the visualised results on the website. The COVID-19 data platform will be useful in our daily life and act an essential role in public health safety and precautions.

Education Use Consent

I hereby give my permission for this project to be shown to other University of Glasgow students and to be distributed in an electronic form.

Name: Yulin Wang Signature: 

Acknowledgements

During the one-year study of MSc Software Development courses and the whole process of this MSc Project, I have received a lot of invaluable helps from many people. Their kind recommendations and comments contributed extremely to the accomplishment of this dissertation as well as my master degree.

First of all, my deepest appreciation flow to my supervisor, Dr Festus Oderanti, for his helpful guidance and valuable suggestions from beginning to end. His profound insight and accurateness about my dissertation taught me so much that they are engraved on my heart. He offered me precious comments during the whole process of coding and writing, without which the paper could not be what it is now.

Also, I would like to express my sincere gratitude to all the professors and lecturers who have taught me in the past year that greatly enriched my knowledge in my study.

Due to the COVID-19 epidemic, the classmates couldn't take classes on campus, but we still keep in touch with each other online, which makes me feel warm and not lonely in Glasgow through such a tough year.

Finally, I would like to thank my family and my friends. They always support me unconditionally, inspire and encourage me to work hard and to complete my postgraduate studies.

Contents

Chapter 1 Introduction

This chapter is a general introduction to the project, with some background of the related concepts. And it claims the aim and objectives of this project, as well as the structure of the dissertation.

1.1 Introduction to the Project

Since December 2019, COVID-19 has profoundly affected human's life. The increasing number of infected people and deaths obviously shows the severity of the epidemic [1]. The statistical work of epidemic data is usually complicated and time-consuming. In order to make data statistics more efficient, it is extremely important to develop a data analysis platform, to integrate multiple data-related functions. In such a data platform, users can get information as they want, including source data, results of data analysis, and data prediction.

The topic of the project is The Development of a COVID-19 Data Tracking Platform for Public Health Safety. It will analyse the data of COVID-19 cases and show the charts and forms to visualise the information. Besides, it will use machine learning algorithms to find out the relationship between data, and predict the new cases in the next period of time.

All the data, analysis results, predictions, and other outputs will be shown on a website. The website is developed with Django framework in Python. And it should be available for 1000 simultaneous accesses.

1.2 Aim and Objectives

The overall aim of the project is to build up a platform to analyse the number of cases in different countries or regions, and to predict the possible cases in the next period of time.

To achieve the aim, the following objectives would be addressed in the project.

- To get real data about COVID-19 cases from medical and healthcare websites and governments official websites, to ensure the reliability of the result of analysis.
- To analyse the data and tell the implied relationship in data and summarise the feature of data.
- To develop and perfect a model to predict the possible cases in the future based on machine learning algorithms.
- To visualise all the analysis and prediction results, and to make them easy to read and understand by showing charts and forms.
- To develop a user-friendly interactive website, and deploy all the works onto the website that would allow users to filter data about what they would like to focus on.

1.3 Structure of the Dissertation

This paper describes the development of the project. It is structured as follows.

Section 2.1 demonstrates the background of the COVID-19 pandemic, and why this topic was chosen. Then, it discusses how to develop a website properly to help analyse the data. Section 2.2 reviews the current literature and existing websites which have the similar functions. Section 2.3 analyses the target of the project and how to achieve the goal. Section 2.4 lists the functional and non-functional requirements in the process of development. Section 2.5 lists the software and technologies which are used in development.

Section 3.1 demonstrates the architecture of the whole system in this project. Section 3.2 lists the detailed structures as well as the structure diagrams. Section 3.3 gives details of the user interface and user experience design.

Chapter 4 demonstrates the implementation of the project, including the key points in the whole system, from data auto-update and pre-process, to data analysis and visualisation, data prediction, and building up the website.

Chapter 5 includes the system testing and product evaluation. Both the web development and the functions implemented by Python programming successfully pass the tests. Meanwhile, the evaluation from voluntary users gives many recommendations with high quality.

Chapter 6 summarises the whole process of this project, and proposes the future plan of further development.

Chapter 2 Analysis and Requirements

This chapter gives details of the background of this project, analyses the current COVID-19 data tracking platforms, and explains the reason why this new data tracking platform should be developed. The requirements and how to achieve the aim are also discussed, to ensure the feasibility of the project. Finally, the technologies used in this project are listed at last.

2.1 Background

COVID-19 is a highly contagious disease caused by the SARS-CoV-2 virus. In December 2019, the first known case was discovered. Since then, the disease has spread globally, resulting in an ongoing pandemic [2]. As of 4th October 2021, more than 234 million cases and 4.8 million deaths have been confirmed, making it one of the deadliest pandemics in history [3].

Data analysis is a very useful way to find out the implied relationship in data, and give a solution or suggestion to a certain problem. Here the data analysis can tell the changing trend of the number of cases. And it will also predict the possible trend in the future to trigger the alarm and remind the relevant staff in the healthcare sector.

2.2 Current Literature and Existing Platforms

There are plenty of websites which focus on the data of COVID-19. Many countries' government release the latest data every day on their official websites. There are always detailed information and reasonable analysis along with the original data.

The governments' official websites show the data of new cases, new deaths, vaccination, etc. There is also reasonable analysis on the websites. These data have the highest reliability while a website only has the data of a single country.

In addition, there are some organisations' websites which collect and summarise the global data. After browsing these websites, it was found that few of them could provide a future data prediction function. In fact, the prediction of future cases in the next period of time could be beneficial and imperative. It can guide the government to make reasonable and wise decisions in a certain condition. And it can also remind people to be prepared when the epidemic gets worse.

Fortunately, a lab of Johns Hopkins University does the global statistics work every day, which includes the data from all over the world. And the datasets are released by them on their website daily. It is convenient and reliable to use these datasets as the source data for this project.

2.3 Target Analysis and Contribution of this New Platform

The target of the project is to build up a data analysis platform to integrate global COVID-19 data together, analyse the data in different countries or regions, and to predict the possible cases in the future.

This platform should get real data about COVID-19 cases from reliable sources and ensure the data are up to date. And it should reasonably analyse and organise the data, and finally show the visualised results to make it easier for users to understand the data. Besides, the platform should also be able to predict the possible cases in the next period of time. And all of these functions should be established on a website which is accessible and user-friendly.

2.4 Requirements

The platform needs to be developed after a comprehensive knowledge of the background and requirements. This section is a general introduction to the requirements of the project, including functional and non-functional requirements.

2.4.1 Functional Requirements

A functional requirement in computer science and software engineering defines a function of a system or the component, where the function is defined as the specification of behaviour between inputs and outputs [4].

In this project, the main functional requirements are as follows.

i. Showing the latest data of COVID-19, and make sure the data are true and reliable. The data source should be put on the website.

   a. The website should show the new cases and new deaths of COVID-19.

   b. The website should show the total cases and total deaths of COVID-19.

   c. The website should provide the latest dataset which is available to be downloaded by the users.

ii. To ensure the data is real-time, the website should show the latest data from the data source. So, the database should update data frequently from the source. In this project, the system should update data every 6 hours automatically.

iii. Visualising the data in a proper way (charts and forms). The data is usually less intuitive than the charts. Visualisation could be quite important in the process of displaying the data and the results of analysis. In this project, the results of analysis and prediction should be both visualised properly by different types of charts and graphs. The detailed requirements are:

   a. The website should show the form of new cases, total cases, new deaths, and total deaths in every country.

   b. The website should show the column charts of new cases in certain countries by date.

   c. The website should show the cases' distribution by pie chart.

   d. The website should show the number of cases intuitively on a world map.

iv. Giving the prediction of the number of cases in the future. The prediction of the future cases could be fairly important to the pandemic. In this project, the system uses the prediction algorithm of Tableau, which is effective. The requirements are:

   a. The system should choose reliable algorithm and parameters as the basis of prediction.

b. The system should make reasonable predictions of new cases and new deaths of certain countries in the next period of time.

c. The website should display the visualised prediction results by line charts to show the potential trend.

2.4.2 Non-functional Requirements

A non-functional requirement in computer science and software engineering is a requirement that specifies criteria that can be used to judge the performance of a system rather than specific behaviours [5].

In this project, the main non-functional requirements are as follows.

i. The website should be available for 1 000 simultaneous accesses.

The website should be usable for a certain scale of simultaneous accesses.

ii. The website should run 24*7.

The website should be usable whenever users want to access.

iii. The website should be accessed via Chrome/Edge/Firefox/Safari.

The website should be compatible to the mainstream and common browsers.

iv. The development of the website should be at most 12 weeks.

The development should not be too long and it must be finished in one semester.

2.5 Software and Technologies Used

This section gives details of the software and technologies used in the process of this project.

i. Programming languages

Python is chosen as the main language of the whole project. Python is an interpreted high-level programming language which was first released in 1991. Its language structure and object-oriented (OO) approach are intended to assist programmers in writing clear, logical codes for almost all kinds of projects [6].

Python libraries allow the effective use of Python in many domains including scientific computing [7], which could be very useful in the process of data analysis [8].

Besides, a little JS might be involved in this project because the final product would be a website.

ii. Statistics

Statistics focuses on data collection, organisation, analysis, interpretation, and presentation [9]. Some fundamental statistics knowledge was used to deal with and analyse the data in this project.

The rapid and ongoing gains in processing power have had a significant impact on the practise of statistical research. The traditional statistical models were nearly always linear in early years, but powerful computers combined with appropriate numerical techniques sparked interest in non-linear models as

well as the development of new types, such as generalised linear models and multi-level models [10]. In this project I will try different statistical models to optimise the results.

iii. Machine Learning

The study of computer algorithms that can improve themselves automatically through experience and the utilisation of data is known as machine learning. [11]. It is commonly regarded as a component of artificial intelligence (AI). Machine learning algorithms employ training data to create a model that can anticipate or make decisions without being explicitly taught to do so [12].

In this project, logistic regression algorithm might be used to make predictions. And linear regression algorithm is also useful in the supervised learning during the process of data analysis.

iv. Web Development

Web development is the process of developing a website. Since the website of this project is not very complicated, Django would be used as the framework of the website.

Django is a Python-based web framework, and its main aim is to ease the creation of complicated websites [13]. It is very suitable for this project.

v. SDLC stages

The systems development life cycle (SDLC), is a procedure for planning, developing, testing, and deploying a software or information system [14]. Here are the key stages of this project's SDLC.

a. Planning and requirements analysis: To do some investigations about the users' requirements.

b. Defining Requirements: To summarise the key requirements and write them down, to form a specification document for software development requirements.

c. Designing the Product Architecture: To follow the requirements specification, use multiple design approaches for the architecture.

d. Building or Developing the Product: To analyse the data, train the model, visualise the results, and develop the website.

e. Testing the Product: To test all the functions of the website.

f. Deployment: To release the product when everything is properly done.

Chapter 3 Design

This chapter demonstrates the architecture of the whole system in this project, lists the detailed structures as well as the structure diagrams, and gives details of the user interface and user experience design.

3.1 System Architecture

In software engineering, the Unified Modelling Language (UML) is a developmental and general-purpose modelling language that is intended to provide a standardised method for visualising system architectures [15]. UML has many types of diagrams, which are divided into two categories [16]. Some types of UML diagrams represent structural information, while others represent general behaviours, including different approaches of interactions.

The project will be shown as a website. The system architecture and the structure of the website should be like the following UML diagram.

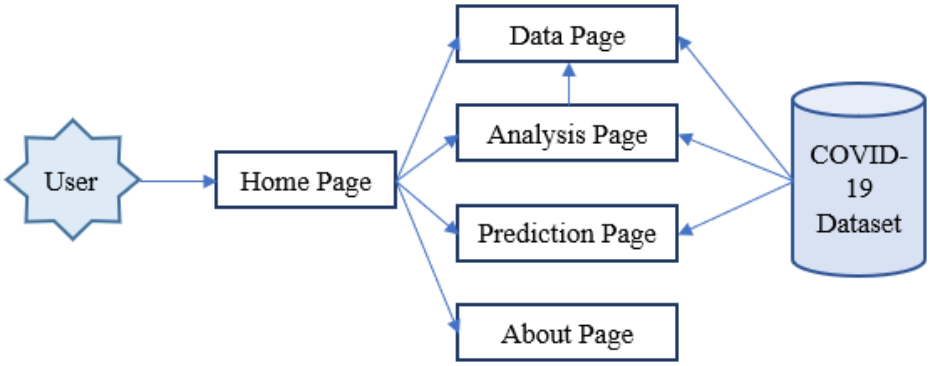

Figure 3-1. System Architecture UML Diagram. The figure shows the main structure of the website, including a Home page and four other pages.

The users will have access to the system via the homepage. In the homepage, users can browse key information and download data. Besides, they can also jump to other pages by clicking on the navigation bar. In the analysis page, users can see the data (including total cases, total deaths, new cases, and new deaths) as well as the visualised results of data analysis. The results are shown with reasonable charts and forms. In the prediction page, users can see the results of future prediction. The instruction of how the prediction algorithm works is also shown in this page. In the about page, there are details about the developer, sources of data, tools and technologies used in the process of development, and other related information in the project.

3.2 Detailed Structure

This section gives details of the structure of every page in the website as well as the structure diagrams.

3.2.1 Home Page

The Home page is the first page when a user enters the system. It has the most links to other pages in the system. Meanwhile, there are plenty of other functions in the homepage.

The entrances to Analysis page, Prediction page and About page, are clearly shown on the top sidebar of the Home page, and the users can also have access to those pages by clicking on the buttons of the dynamic slides.

In the Home page, users can see some key data at first glance, including the key data of some selected countries and the daily global data of new cases, new deaths, total cases, and total deaths.

It is also available for users to download the source data. Users can check the latest data form first, and download the up-to-date datasets to the local storage as they like.

Besides, the Home page also gives some tips to help people get rid of the COVID-19 virus and provides links to some authoritative organisations' websites.

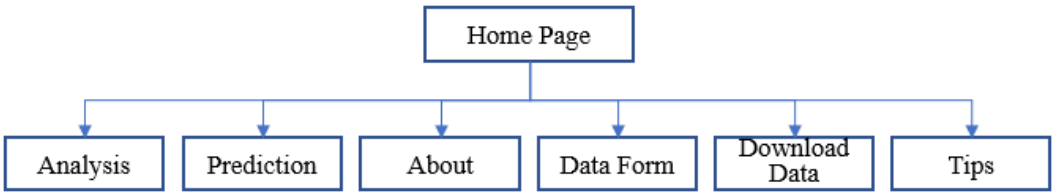

Figure 3-2. Structure Diagram of Home Page. The figure shows that the Home page contains 6 modules.

3.2.2 Analysis Page

The Analysis page has two main functions. On the one hand, users can check the data form for detailed data of every country. On the other hand, some charts are displayed in this page, to show the visualised results of data analysis. These charts are beneficial for the users to understand the trend and other features of the data intuitively.

If the user wants to check the data form, the website will jump to the Data page. All the detailed data of every country are shown by the order of alphabet, which is convenient for searching.

The main part of the Analysis page includes several well-visualised charts and graphs, showing the trend and other features of data. These charts are all interactive, so users can get detailed data by clicking on the charts. These charts are linked to the latest dataset, so they are also up-to-date, along with the data form in Data page.

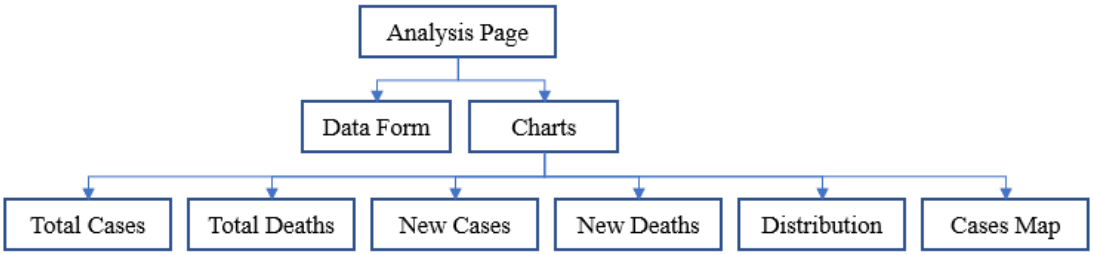

Figure 3-3. Structure Diagram of Analysis Page. The figure shows that the Analysis page contains the link to data form and 6 charts which show the results of data analysis.

3.2.3 Prediction Page

The Prediction page includes the line charts of new cases prediction in five selected countries. In this page, users can intuitively know the most possible trend of the new cases in the future.

Five line charts are displayed in this page, and users can get the number of new cases on any single day in these charts.

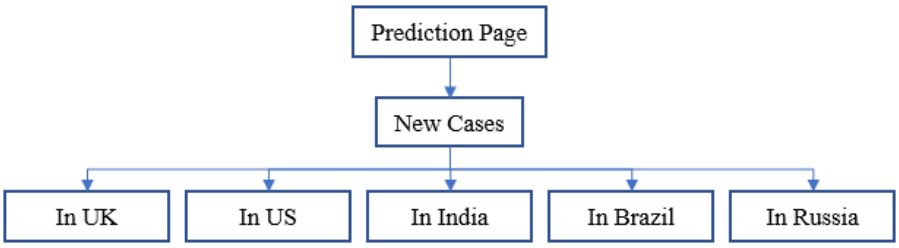

Figure 3-4. Structure Diagram of Prediction Page. The figure shows that the Prediction page contains the new cases prediction of 5 selected countries.

3.2.4 About Page

The About simply includes some necessary technical information of the platform, including a basic instruction to guide the users while using this platform.

In addition, the related information is also published here, including data sources, software versions, author info, etc.

Figure 3-5. Structure Diagram of About Page. The figure shows that the About page contains the related information of the platform.

3.2.5 Global Structure

Overall, the structure of the whole system is well-designed. And the levels of contents are organised properly, so that users can easily have access to the target page or target part in the platform.

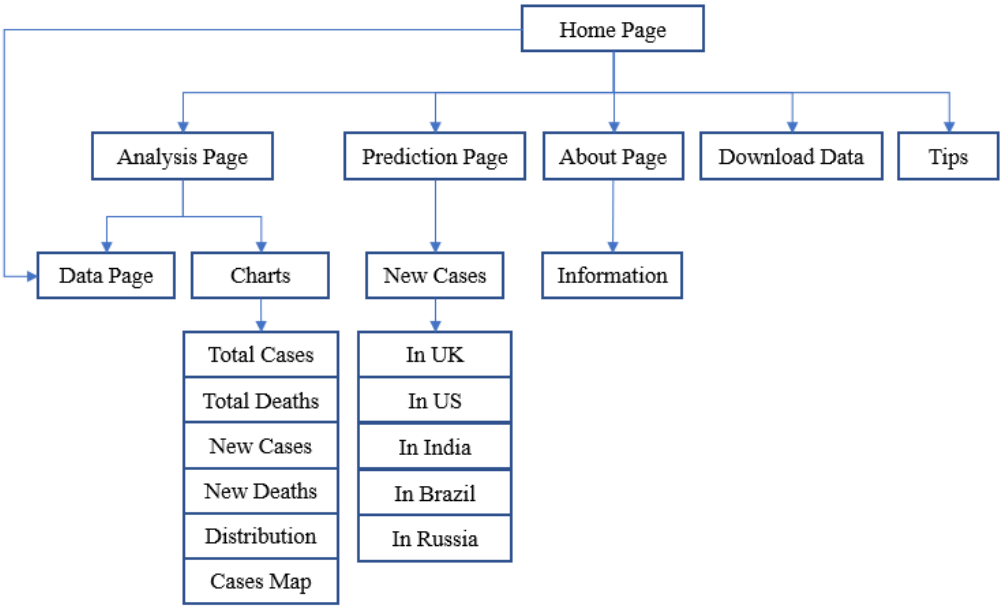

Figure 3-6. Structure Diagram of the Entire System. The figure shows the relationship of pages in the whole project.

3.3 User Interface and User Experience

A user interface (UI) is the space where humans and machines interact [17]. The goal of this interaction is for the human to be able to effectively manage and control the machine, while the machine simultaneously provides information that aids the operators' decision-making process. User interface design issues are related to or involve disciplines like psychology and art design.

The user experience (UX) is how a user perceives and interacts with a system, software or service [18]. It encompasses a person's ideas of utility, ease of use, and efficiency. When developing and refining goods, most organisations, designers, and producers prioritise improving user experience because poor user experience can reduce product usage. User experience is subjective, while the attributes that comprise the user experience are objective.

In general, the purpose of user interface design is to create a user interface that makes it simple, efficient, and pleasurable to operate a machine in order to get the desired result (i.e., maximum usability). This often means that the operator must supply minimal input in order to produce the intended output, and that the machine reduces unwanted outputs to the user [19].

User-centred interaction design should aim for maximum usability. In this project, the main UI design is based on the website. The homepage of the website shows some basic data and related information. And it also has the links to other pages, which are set in the top navigation bar, and it can be very quick to switch different pages by clicking on the links. Besides, the homepage shares the design of top navigation bar and bottom columns with all the other pages in this platform. Unified webpage design can help users utilise the platform easily and efficiently, and this ensures the user-friendly experience of the platform.

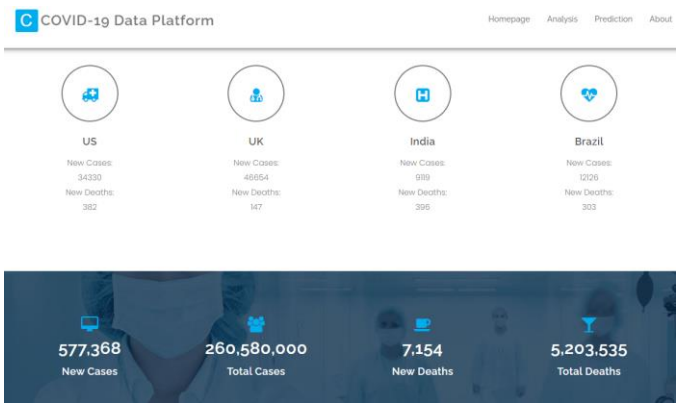

Figure 3-7. Screenshot of a Part of the Home Page. The figure shows that the navigation bar is always on the top.

The slides effect is widely used in the websites of many famous companies. It can offer several categories' introductions and links which are displayed cyclically. So, the slides effect is applied in the Home page of the platform, to provide related information and links to Data page, Analysis page, and Prediction page.

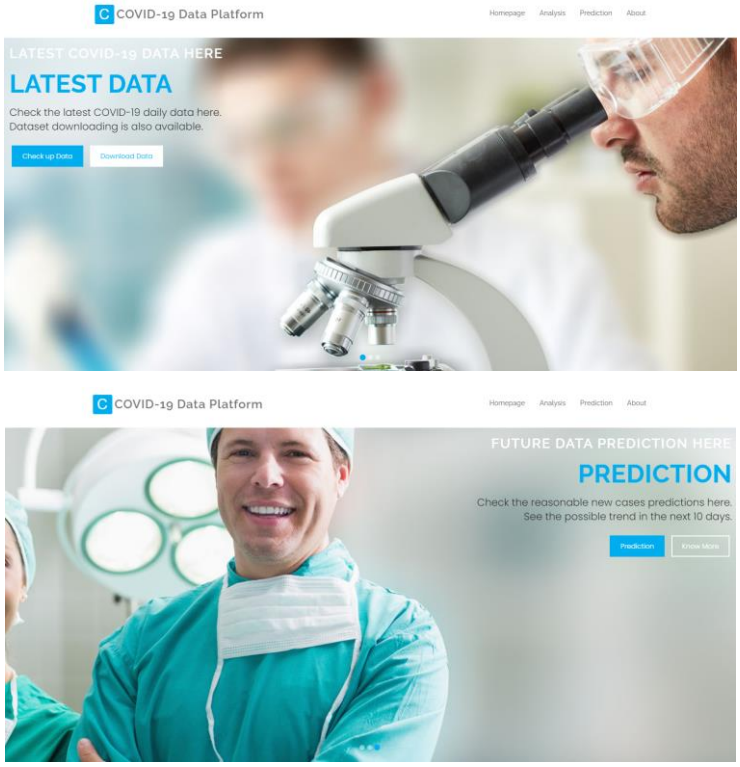

Figure 3-8. Slides Effect in the Home Page. The figure shows that the slides in Home page can intuitively show the core functions of the platform.

Chapter 4 Implementation

This chapter demonstrates the implementation of the project, including the key points in the process of development, from data auto-update and pre-process, to data analysis and visualisation, data prediction, and building up the website.

4.1 Data Auto-update and Pre-process

In this project, data are analysed in different dimensions. First, the dataset are downloaded from the data source automatically by the function *auto_download()* in the programme *auto_download.py*. The zip folder then should be unzipped by the function *un_zip()* in the programme *unzip.py*.

```python
import requests

def auto_download():
    url = "https://www.kaggle.com/antgoldbloom/covid19-data-from-john-hopkin
    cookie = {'Cookie':'_ga=GA1.2.199613981.1631564924; ka_sessionid=467eeee
    html_str = requests.get(url,cookies=cookie).content
    with open('D:/GLA Lessons/MSc Project/Data/test/data.zip', 'wb') as f:
        f.write(html_str)

if __name__ == '__main__':
    auto_download()
```

```python
    def un_zip(file_name):
        zip_file = zipfile.ZipFile(file_name)
        if os.path.isdir("data"):
            pass
        else:
            os.mkdir("data")
        for names in zip_file.namelist():
            zip_file.extract(names,"data")
        zip_file.close()

    if __name__ == '__main__':
        file = 'D:/GLA Lessons/MSc Project/Data/test/data.zip'
        un_zip(file)
```

Figure 4-1. Key Codes of auto_download.py and unzip.py Python Programs. The figure shows how the programme downloads the latest dataset automatically from Internet and unzips the zip folder to get the csv dataset.

Second, source data are pre-processed by the function *data_preprocess()* in the programme *data_preprocess.py*.

Third, the data of overseas territories or provinces are merged altogether by the function *merge_country()* in the programme *data_preprocess.py*.

```python
import pandas as pd
import numpy as np
from pandas.core.reshape.concat import concat
import os

def data_preprocess():
    # New Cases
    file_path1 = "D:/GLA Lessons/MSc Project/Data/data/CONVE

    df1 = pd.read_csv(file_path1).T
    df1 = df1.rename_axis('Country/Region').reset_index()
    df1 = df1.drop(df1.index[0:1, ]).reset_index(drop=True)
    df1.iloc[:, [-1]] = df1.iloc[:, [-1]].astype(np.float64)
    df1.iloc[:, [-1]] = df1.iloc[:, [-1]].astype(np.int64)
```

```
    # Concatenation
    df_processed = pd.concat([df1_processed, df2_processed['Total Cases'], df3_
    print(df_processed)
    os.getcwd()
    df_processed.to_csv('PROCESSED_DATA.csv')

if __name__ == '__main__':
    data_preprocess()
```

```
def merge_country():
    # Total Cases
    file_path2 = "D:/GLA Lessons/MSc Project/Data/data/RAW_global_confirmed_cases.csv"
    df2 = pd.read_csv(file_path2)

    df2 = df2.drop(df2.columns[1:4], axis=1)
    df2_1 = df2.iloc[0:8, :]
    df2_merged = df2.iloc[8:16, :]
    sum = df2_merged.iloc[:, 1:].sum(axis=0)
    df2_merged.loc[df2_merged.index.max()+1] = sum
    df2_merged.loc[df2_merged.index.max(),'Country/Region'] = 'Australia'
    df2_merged = pd.concat([df2_1, df2_merged]).reset_index(drop=True)
    df2_2 = df2_merged.drop(df2_merged.index[8:16, ]).reset_index(drop=True)
```

```
    df4_selected = pd.concat([df4_us, df4_india, df4_uk, df4_brazil, df4_russia, df4_world]).reset_index(drop=True)

    os.getcwd()
    df4_selected.to_csv('SELECTED_total_deaths.csv')
    df4_merged.to_csv('WORLD_total_deaths.csv')

if __name__ == '__main__':
    merge_country()
```

Figure 4-2. Key Codes of data_preprocess.py and merge_coutry.py Python Programs. The figure generally shows how the programme pre-processes the raw data to make it more suitable for further analysis.

Fourth, the pre-processed data are displayed in a form embedded in a html file by the function *show_csv()* in the programme *csv_to_html.py*.

```
file_path = "D:/GLA Lessons/MSc Project/Data/PROCESSED_DATA.csv"
app = flask.Flask(__name__)

@app.route("/")
def show_csv():
    df = pd.read_csv(file_path)
    data_html = df.to_html('D:/GLA Lessons/MSc Project/Project/Cc

    return f"""
        <html>
            <body>
                <h3>Dataset of Covid-19</h3>
                </div>{data_html}</div>
            </body>
        </html>
    """

if __name__ == '__main__':
    app.run(host='0.0.0.0')
```

Figure 4-3. Key Codes of csv_to_html.py Python Programme. The figure shows how the programme displays csv data in a html website.

Finally, call these functions in order and add a scheduler in *daily_update.py*, to automatically update the data every 6 hours.

All the detailed codes and datasets in this project are presented in Appendix D.

```python
def func():
    auto_download.auto_download()
    file = 'D:/GLA Lessons/MSc Project/Data/test/data.zip'
    unzip.un_zip(file)
    data_preprocess.data_preprocess()
    merge_country.merge_country()
    csv_to_html.show_csv()

def update():
    # BlockingScheduler
    scheduler = BlockingScheduler()
    # update every 6 hours
    scheduler.add_job(func, 'interval', hours=6, id='test_job1')
    scheduler.start()

if __name__ == '__main__':
    update()
```

Figure 4-4. Key Codes of daily_update.py Python Programme. The figure shows that daily_update.py integrates all the programs together, to automatically update the latest data every 6 hours.

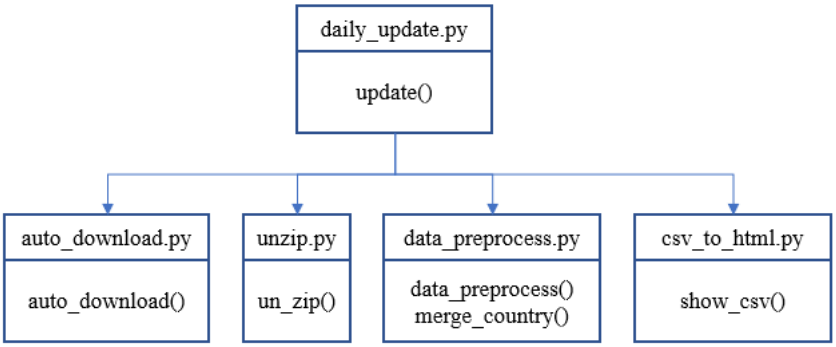

Figure 4-5. UML Diagram of Data Pre-process Python Programs. The figure shows that every small function was realised by a single programme.

4.2 Data Analysis and Visualisation

Data analysis is described as the process of cleansing, manipulating, and modelling data in order to extract valuable information for making reasonable decisions. The target of data analysis is to extract usable information from raw data and make decisions or analyse results based on that knowledge [20].

In this project, there are plenty of data in the dataset. So, it is a challenge to filter and select the data which are useful in the further application.

Data visualisation is highly widespread in everyday life, and it frequently appears in the form of charts and graphs. In other words, data presented graphically to make it easier for the human brain to absorb and process. Data visualisation is widely used to discover previously unknown facts and trends. You can discover useful knowledge by detecting relationships and comparing datasets.

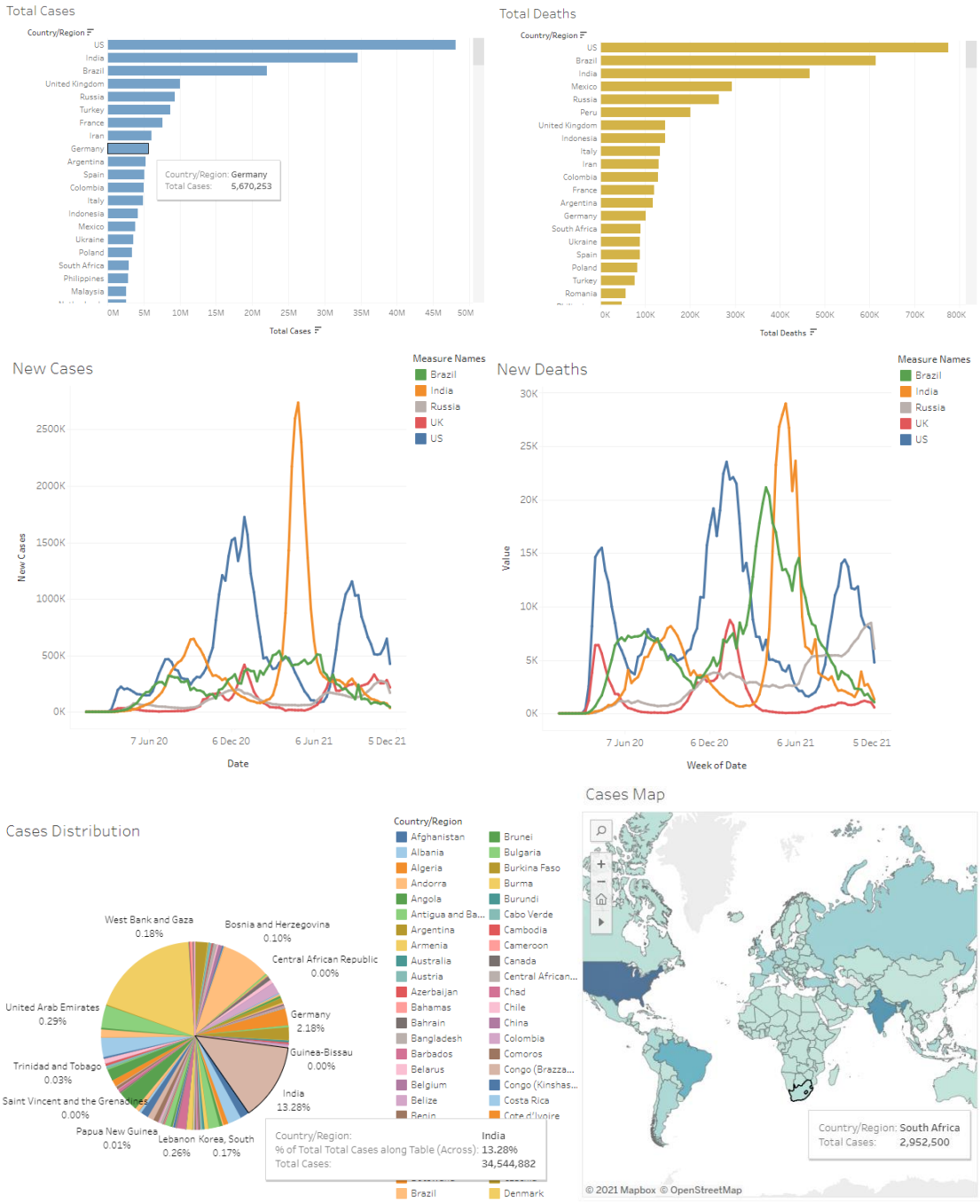

Figure 4-6. Visualised COVID-19 Charts in Analysis Page. The figure shows 6 charts in Analysis page, including the charts of total cases, total deaths, new cases, new deaths, cases distribution, and cases map.

4.3 Data Prediction

Predictive Analysis shows what will probably happen based on previous data. And making reasonable and reliable predictions usually needs statistics knowledge. In this project, Tableau is used to make predictions. The prediction function of Tableau uses regression algorithms to find out the potential periodic changes or trends. After setting parameters well, it will release a possible prediction of the next 10 days.

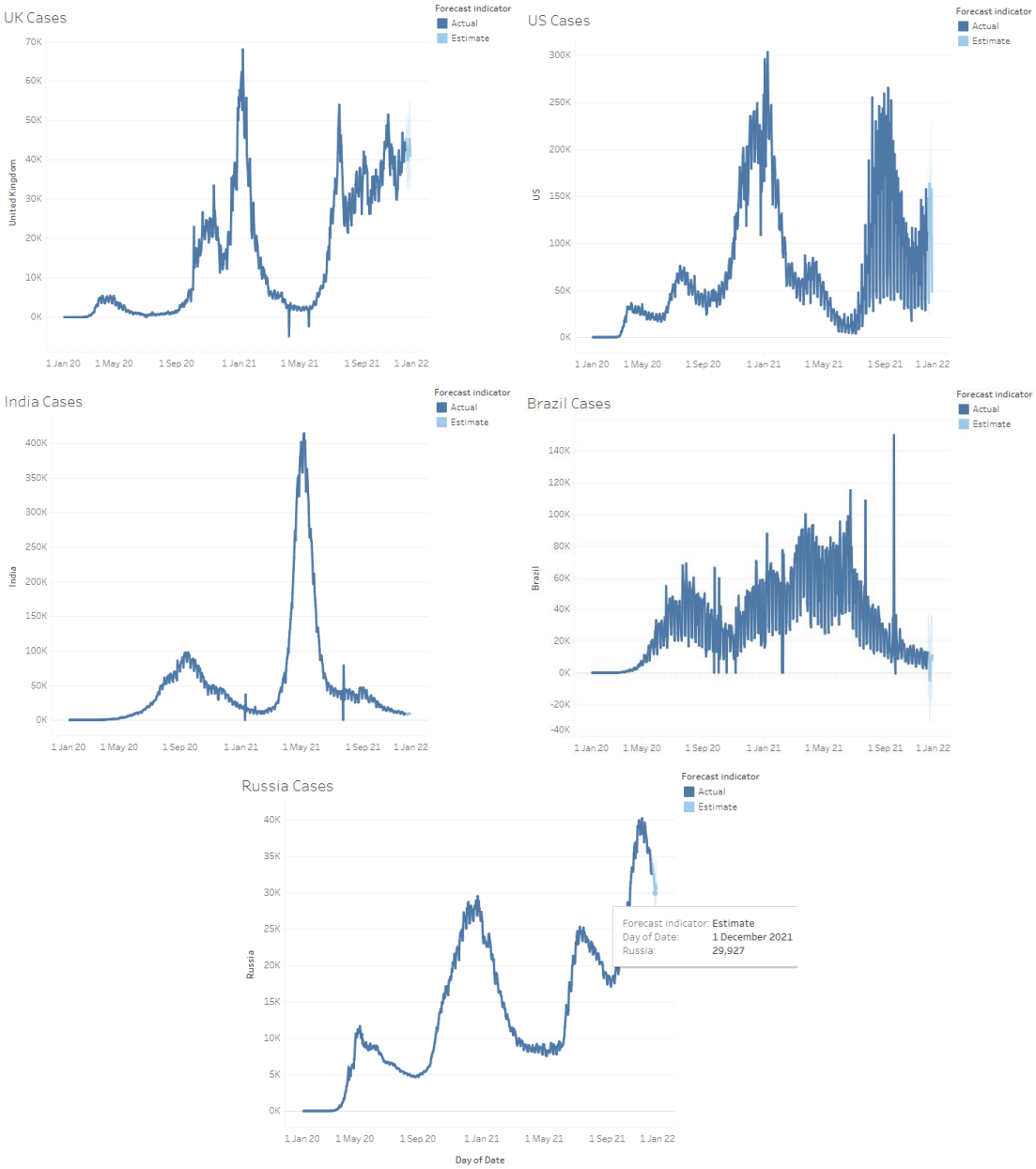

Figure 4-7. COVID-19 New Cases Predictions in Prediction Page. The figure shows the line charts of the new cases predictions of the next 10 days in 5 selected countries (UK, US, India, Brazil, and Russia).

4.4 Building up the Website

Building up the website is an important session to improve the human-system interaction. It can better deliver the contents to more people, and can more intuitively show the results of analysis.

The website is very flexible, the author can update the data as well as other contents in the website at any time. If the website is established on a server, it will be published to the Internet, which means people from all over the world can have access to the website.

In this project, Django framework is used and the Uniform Resource Locators (URL) are declared in the *urls.py* files.

```python
urlpatterns = [
    path('', views.index, name='index'),
    path('covid/', include('covid.urls')),
    path('admin/', admin.site.urls),
    path('about/', views.about, name='about'),
] + static(settings.STATIC_URL, document_root=settings.STATIC_ROOT)
```

```python
app_name = 'covid'

urlpatterns = [
    path('', views.index, name='index'),
    path('analysis/', views.analysis, name='analysis'),
    path('prediction/', views.prediction, name='prediction'),
    path('dataframe/', views.dataframe, name='dataframe'),
] + static(settings.STATIC_URL, document_root=settings.STATIC_ROOT)
```

Figure 4-8. Key Codes of urls.py Programme. The figure shows the URL settings of the website.

The index page is the homepage of the website, covid and about pages are under the index page. And there are three pages under covid page – analysis, prediction, and dataframe. The URL map is as follows.

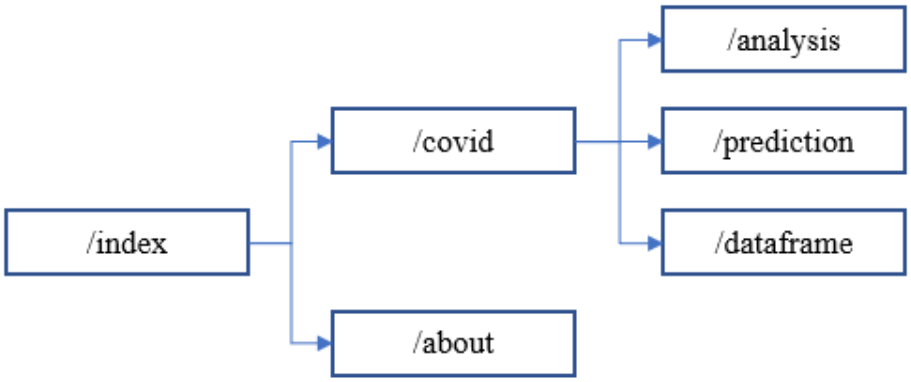

Figure 4-9. URL Map of the Website. The figure shows the URL structure of the website.

Chapter 5 Testing and Evaluation

This chapter includes the system testing and product evaluation. Both the web development and the functions implemented by Python programming successfully pass the tests. Meanwhile, the evaluation from voluntary users gave many recommendations with high quality.

5.1 System Testing

The development of this platform could be basically divided into two parts. One is the development of the website, which carries miscellaneous data and information. And the other is the implementation of target functions, which enables the platform to be updated automatically.

Therefore, the system testing was done separately. The results are as follows:

Table 5-1. Tests of Website Development

Functions	Success or not
Having access to Home page	Success
Having access to Analysis page	Success
Having access to Prediction page	Success
Having access to About page	Success
Having access to Data page	Success

Table 5-2. Tests of Platform Functions

Functions	Success or not
Downloading the dataset	Success
Checking the data form	Success
Pre-processing data automatically	Success
Displaying key data in Home page	Success
Updating data every 6 hours	Success
Displaying the charts of analysis	Success

Displaying the charts of prediction	Success
Checking the details of tips	Success

5.2 Evaluation Design

Evaluations are essential to the iterative product development cycle. The recommendations and feedback which were captured from the evaluations have contributed extremely to the last period of development. Besides, they are beneficial to create a well-designed and user-oriented platform.

The target of the evaluation is to address the following questions:

- Would the users be happy to use this COVID-19 data tracking platform in their daily life or professional research?
- What function would the users like most?
- Which part of the platform could improve more in the future?
- How many marks would the users rate this platform?

All these 4 questions are included in the questionnaire. And the elected participants for this evaluation are in different age groups, and they are engaged in variety of professions.

The severe pandemic situation this year seriously limited the face-to-face participation for this evaluation. However, through emails, Telegram, and other communication tools, 30 participants filled in the form of system functional tests and evaluation questionnaires. Besides, they were also asked to give some recommendations to improve the platform.

5.3 Evaluation Results

The details of results are listed in Appendix B.

The results show that 21 / 30 users would be happy to use this platform.

The most popular function is the new cases prediction.

The participants think the data form should improve more in the future.

The average score of the platform is 8.47 / 10, which shows that most participants are satisfied with the platform generally, though it still can be improved in some parts in the future development.

Chapter 6 Conclusion and Future Work

This chapter includes the conclusion of this platform, the conclusion of the MSc project, and the future work.

6.1 Conclusion of the Project

This dissertation has introduced and recorded the current conditions of COVID-19 data platforms and the development of a new platform, from the investigation to its design, implementation, testing, and evaluation.

Through 11 weeks of intensive development, a completed COVID-19 data platform is delivered, with all the expected functional and non-functional requirements met. The target of building a user-oriented system and optimising the user experience has been met to the extent possible within the development time frame.

Thanks to the feedbacks and recommendations, several changes have been completed to improve the platform's functionality and usability. For example, the top sidebar is always displayed on the top now, to make users switch the web pages easily and quickly. The slides in the Home page now have a dynamic effect, to improve the user experience. The data form is listed by the alphabet order, to enable users find the data they want more efficiently.

6.2 Limitations and Future Work

Despite the system's overall effectiveness in reaching its aim and objectives, there are still certain flaws that can be addressed in future development. For example, the reliability of prediction is still not too high.

Due to the time constraints, the data platform is still not perfect, and there is still much to do in the future. The development plan in the future will mainly focus on improving human-system interaction and adding more functions to the platform, in order to maximise user experience. For example, more charts and graphs could be added to the Analysis page, to enrich the functions of the platform. In addition, the website could be established to an online server, so that users can have access to the website from all over the world.

Finally, the COVID-19 data platform is quite usable now, and the functions can meet the basic requirements of daily reminder. It is a useful tool for ordinary people as well as professional staff in public health, to get the latest pandemic data and some reasonable analysis and prediction results.

References

[1] Zimmer C. 2021. The Secret Life of a Coronavirus – An oily, 100-nanometer-wide bubble of genes has killed more than two million people and reshaped the world. Scientists don't quite know what to make of it. Retrieved 28 February 2021.

[2] Symptoms of Coronavirus. U.S. Centers for Disease Control and Prevention (CDC). 13 May 2020.

[3] Scientific Brief: SARS-CoV-2 and Potential Airborne Transmission. COVID-19 Published Science and Research. U.S. Centers for Disease Control and Prevention (CDC). 11 February 2020.

[4] Fulton R, Vandermolen R. 2017. Chapter 4: Requirements - Writing Requirements. Airborne Electronic Hardware Design Assurance: A Practitioner's Guide to RTCA/DO-254. pp. 89–93. ISBN 9781351831420.

[5] Chen, Lianping; Ali Babar, Muhammad; Nuseibeh, Bashar. 2013. "Characterizing Architecturally Significant Requirements". IEEE Software. 30 (2): 38–45. DOI:10.1109/MS.2012.174

[6] Kuhlman, Dave. 2012. A Python Book: Beginning Python, Advanced Python, and Python Exercises. Section1.1.

[7] Oliphant, Travis. 2007. Python for Scientific Computing. 9(3): 10-20. DOI:10.1109/MCSE.2007.58

[8] Millman, K. Jarrod; Aivazis, Michael. 2011. "Python for Scientists and Engineers". Computing in Science and Engineering. 13 (2): 9–12. DOI:10.1109/MCSE.2011.36.

[9] Romeijn, Jan-Willem. 2014. Philosophy of Statistics, The Stanford Encyclopedia of Philosophy.

[10] AISOMA. 2019. 10 Statistical Techniques Data Scientists Should Master. Retrieved 18 December 2021.

[11] Mitchell, Tom. 1997. Machine Learning. New York: McGraw Hill. ISBN 0-07-042807-7.

[12] Koza, John R.; Bennett, Forrest H.; Andre, David; Keane, Martin A. 1996. Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. Artificial Intelligence in Design '96. Springer, Dordrecht. pp. 151–170. DOI:10.1007/978-94-009-0279-4_9

[13] Django documentation of deployment. Archived from the original on 11 January 2012.

[14] SDLC – Overview. 2021. SDLC Tutorial in Tutorialspoint. Retrieved 18 December 2021.

[15] Addison-Wesley. 2005. Unified Modeling Language User Guide, The (2 ed.). p. 496. ISBN 0321267974.

[16] OMG Unified Modeling Language (OMG UML), Superstructure. Version 2.4.1. Object Management Group. Retrieved 9 April 2014.

[17] Adobe XD. UI Design. 2021. Adobe XD Ideas. Retrieved 18 December 2021.

[18] Don, Norman; Jakob, Nielsen. The Definition of User Experience (UX). Nielsen Norman Group. Retrieved 2 March 2021.

[19] Cohen, Philip R. 1992. The role of natural language in a multimodal interface. Proceedings of the 5th annual ACM symposium on User interface software and technology - UIST. pp. 143–149. DOI:10.1145/142621.142641. ISBN 0897915496.

[20] Transforming Unstructured Data into Useful Information. 2014. Big Data, Mining, and Analytics, Auerbach Publications. pp. 227–246. 2014-03-12. DOI:10.1201/b16666-14. ISBN 978-0-429-09529-0.

Appendix A Ethics Checklist

School of Computing Science Ethics Committee
University of Glasgow

Request for Ethical Approval

This form is to be used by 3rd year, 4th year, MSci, and taught MSc students in the School of Computing Science whose projects entail human participation and which do not conform to any one of the criteria on the project ethics checklist form (http://www.dcs.gla.ac.uk/~hcp/ethics/projects-form.pdf).

Students enrolled for an MSc by Research or a PhD, and members of academic or research staff should submit their request for ethics approval to the Faculty Ethics Committee (see https://frontdoor.spa.gla.ac.uk/researchethics/)

The form should be completed and returned by email to Professor Matthew Chalmers (matthew.chalmers@glasgow.ac.uk) to whom all enquires or requests for advice should be directed.

All sections of this form must be completed.

Before completing this form, please read the British Psychological Society's Code of Conduct (available on http://www.dcs.gla.ac.uk/~hcp/ethics/). The relevant sections of the code are noted against questions in this form.

Copies of the participant information form and consent form should be submitted together with this form.

Project title: Development of a COVID-19 Data Tracking Platform for Public Health Safety

1. Describe the basic purposes of the proposed research:
   To track the latest COVID-19 data, analyse the data, and predict the future cases.

2. Describe the design of your experiment (e.g. conditions, number of participants, procedure, equipment):
   30 participants filled in the form to mark the platform and gave their recommendations and feedback.

3. Describe how the procedures affect the participants:
   Participants first tried the platform and then stated their feelings.

4. State what in your opinion are the ethical issues involved in the proposal:
   The participants' personal information should be maintained as secrecy.

5. Specify whether the research will involve children or those with mental disability or handicap:
   If so, explain the steps taken to obtain permission from LEAs, headteachers, parents etc:

No.

6. State if payment will be made to participants:
   No.

7. Describe procedures for advertising, for recruiting participants, and for obtaining consent from participants:
   Advertisement was sent on the social media and then 30 volunteers were selected for the product evaluation.

8. State whether the proposal is in accord with the BPS Code of Conduct or the ESRC Frame of Research Ethics.
   Yes.

9. Describe how the participants' anonymity and confidentiality will be maintained:
   All the questions in the questionnaire didn't mention private information and the survey was totally anonymous.

10. Date on which the project will begin and end: 11/11/2021 – 2/12/2021

11. Location at which the project will be carried out: *online.*

12. Describe how participants will be debriefed at the end of the experiment. This must include the opportunity to contact the experimenter (or supervisor) for feedback on the general outcome of the experiment.
    I will send the feedback and results to the participants.

Student's Name _____Yulin Wang_____

Year level (3rd, 4th, MSci, MSc) _____MSc_____

Student's Signature _____王雨琳_____

Date ____16/12/2021____

Supervisor's Name _____Festus Oderanti_____

Supervisor's Signature _____Festus Oderanti_____

Date ____16/12/2021____

Appendix B Results of Evaluation

Total participants amount: 30.

Results statistics of the questionnaire:

- *Would the users be happy to use this COVID-19 data tracking platform in their daily life or professional research?*
  Yes: 21
  Not sure: 3
  No: 6

- *What function would the users like most?*
  New cases prediction: 16
  Visualised data analysis: 8
  Check the latest dataset: 5
  Download raw data: 1

- *Which part of the platform could improve more in the future?*
  Data form: 17
  Data analysis charts: 6
  Data prediction functions: 5
  Website design: 2

- *How many marks would the users rate this platform?*
  10: 5
  9: 9
  8: 12
  7: 3
  6: 1

Appendix C Screenshots of the Website

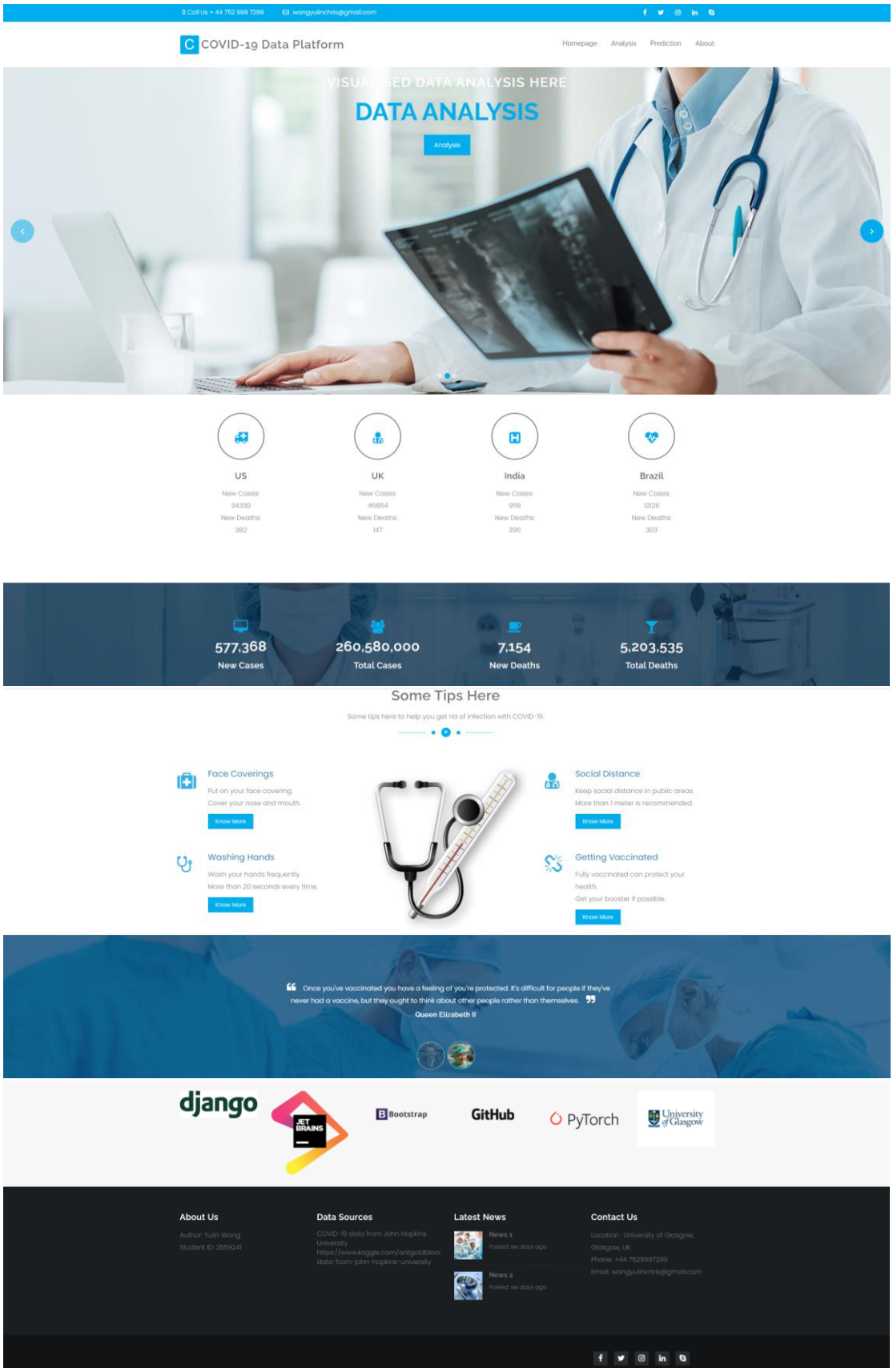

Figure C-1. Screenshots of Home Page. The figure shows the Home page of the platform.

Appendix D Codes and Datasets

All the codes of the project are released on GitHub, and they are stored separately by back-end or front-end in two repositories.

Back-end:

https://github.com/dlwyl1997/Covid-19_Data

Front-end:

https://github.com/dlwyl1997/Covid-19_Data_Analysis

Data sources:

https://www.kaggle.com/antgoldbloom/covid19-data-from-john-hopkins-university