SAFEPASS:
AN IMPLEMENTATION OF AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) FOR CROWD FORECASTING APPLIED IN QUARANTINE PASS

A Capstone Project Presented to the Graduate Program
College of Engineering and Technology
Pamantasan ng Lungsod ng Maynila

In Partial Fulfillment of the Requirements for the Degree
Master's in Information Technology

By
Joane Marie F. Llamera

Dr. Khatalyn E. Mata
Thesis Adviser

August 2021

APPROVAL SHEET

The capstone project hereto titled

SAFEPASS:
AN IMPLEMENTATION OF AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) FOR CROWD FORECASTING APPLIED IN QUARANTINE PASS

prepared and submitted by Joane Marie F. Llamera in partial fulfilment of the requirements for the degree of Master's in Information Technology has been examined and is recommended for acceptance and approval for **ORAL EXAMINATION**.

____
DR. KHATALYN E. MATA
Adviser

PANEL OF EXAMINERS

Approved by the Committee on Oral Examination
with a grade of __________ on ________.

PROF. MANUEL L. OCAMPO
Panel Chair
Chairman

______________________________	______________________________
Panel Member	Panel Member
Member	Member

Accepted and approved in partial fulfilment of the requirements for the degree of Master's in Information Technology.

______________________________	______________________________
Dr. Denvert C. Pangayao	Dr. Clydelle M. Rondaris
Director	Dean
Graduate Program	College of Engineering and Technology

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter One

INTRODUCTION

1.1 Background of the Study

Although the novel coronavirus illness (COVID-19) outbreak in Wuhan, China, was originally reported more than a year ago, it continues to be a global concern. Despite the fact that numerous vaccines are already available to assist prevent individuals from becoming extremely ill or dying as a result of the virus, most countries are still fighting the epidemic as new virus variants emerge.

In the study entitled "Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020" by Cai et al. (2020), COVID-19 instances related with a shopping centre in Wenzhou were studied by the researchers. Indirect transmission of the Covid-19 virus was shown to occur as a result of contamination of common objects, virus aerosolization in confined places, or virus propagation from an asymptomatic infected individual, according to the findings.

In another study entitled "Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study" by Badr et al. (2020), the findings revealed a robust and statistically significant link between social distancing in the USA, as measured by movement patterns, and COVID-19 case growth reduction. The researchers have stressed the importance of social distancing as an effective method of mitigating COVID-19 transmission, and it should continue to be a part of individual and institutional responses to the epidemic.

Forecasting is critical for effective governmental decision-making, supply chain resource management, and understanding extremely tough political decisions such as imposing a lockdown or curfews during a pandemic.

The Philippine government's instrument for assessing, monitoring, controlling, and preventing the spread and local transmission of COVID-19 is the Inter-Agency Task Force for the Management of Emerging Infectious Diseases (IATF – EID). They classify Provinces, Highly Urbanized Cities (HUCs), and Independent Component Cities (ICCs) according to their risk levels, and then conduct targeted lockdowns for designated vital regions.

The most stringent quarantine grade is Enhanced Community Quarantine, which requires tight home quarantine in all families. People's mobility will be restricted to vital goods and services, as well as employment in offices or industries that are permitted to function, such as public and private hospitals, health, emergency and frontline services, essential goods producers, and so on. A quarantine pass enabling one person per home to purchase necessary products or services is issued to communities under an Enhanced Community Quarantine (ECQ).

The present quarantine pass implementation is done manually, which has resulted in a number of difficulties. On March 20, 2020, DILG Region V issued an advisory advising Provincial Directors to instruct Local Officials that the distribution of quarantine pass slips in their areas of responsibility must be done on a house-to-house basis, as releasing Barangay passes in Barangay Halls defeats the intent of the Enhance Community Quarantine and Social Distancing Act. A barangay captain in Lanao del Sur was detained for "selling" passes that would allow individuals to leave their houses during the quarantine period, according to a news report published on Inquirer.net on March 23, 2020.

These challenges, as well as the human work involved in generating and validating quarantine passes, can be addressed by developing a mobile responsive application. The suggested application, when combined with a decision support system, will allow the government and communities to make more informed decisions in order to limit COVID-19 virus transmission.

1.2 Statement of the Problem

The existing procedure of generating, issuing, and validating quarantine passes primarily entails face-to-face interaction, which defeats the Enhanced Community Quarantine's goal. It not only takes a lot of time and effort, but it can also be used by individuals who wish to get over the quarantine rules.

Authorities are unable to capture any offenders since the manual method lacks adequate means to validate and authenticate the quarantine pass. Although just one person per home is permitted to leave the house to obtain needed commodities, the likelihood of a crowded institution is also a source of concern. These gaps may eventually result in a rise in the number of unauthorized people outside of the home, rendering the quarantine restrictions ineffective.

If the policies developed by the IATF during Enhanced Community Quarantine are not implemented with proper tools and risks are not controlled, they can be regarded ineffective.

1.3 Objectives of the Study

The main objective of this capstone project is to develop a mobile responsive web application that would automate the process of generating and validating quarantine pass and show statistical data for a decision-making system.

Specifically, this capstone project seeks to achieve the following objectives:

1. To provide digital quarantine passes for residents that will be validated at essential establishments by using Quick Response Code scheme.
2. To provide information to residents about the current crowd count of essential establishments.
3. To provide users with crowd forecast on essential establishments for upcoming days by using Autoregressive Integrated Moving Average (ARIMA) model.

1.4 Scope and Limitations

To define the scope of this capstone project, the proponent would create a mobile responsive online application that would automate the laborious process of generating and validating quarantine passes while also providing users with a crowd forecasting feature.

This capstone project will use dummy data for information of residents, households per barangay, barangay officials and essential establishments to give emphasize on how the automated system can provide usable crowd forecasts during the implementation of Enhanced Community Quarantine for better decision making and even risk management.

To limit the scope, the system to be developed will not cover other types of quarantine passes such as Special Quarantine Pass, Self-Employed Pass, 1 Week Transit Pass, 1 Day Transit Pass and the likes but only on Quarantine Pass intended for access to essential establishments.

1.5 Significance of the Study

Results obtained from this capstone project will benefit the following stakeholders:

Local Government Officials. During the implementation of Enhance Community Quarantine, this tool will eliminate the manual process of generating, issuing, and validating quarantine passes, allowing them to focus on other vital tasks. By offering crowd projections for decision assistance, the application can operate as an extra support system in ensuring that the Enhanced Community Quarantine requirements are followed accordingly.

Residents. Individuals can simply obtain a quarantine pass by using a mobile application that does not require face-to-face interaction. They are provided with informative data via the application within their grasp, allowing them to select when is the optimum time to purchase vital commodities while minimizing their risk of exposure to congested places.

Future Proponents. This capstone project could be used as a resource for students who are interested in conducting similar research.

1.6 Definition of Terms

COVID-19

Refers to the Coronavirus Disease 2019 which is caused by the virus known as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

Asymptomatic

Asymptomatic means there are no symptoms.

Aerosolization

The production of an aerosol -- a fine mist or spray containing minute particles.

Non-pharmaceutical Interventions

Non-pharmaceutical Interventions (NPIs) are measure that people and communities can take in addition to getting vaccinated and taking medicine to help reduce the spread of illnesses like pandemic influenza (flu). Community mitigation techniques are another name for NPIs. Pandemic flu occurs when a novel flu virus spreads around the world, producing widespread disease. The human population has little or no immunity to a pandemic flu virus since it is new. As a result, the infection can easily spread from person to person all over the world. When vaccines are not yet available, NPIs are one of the best approaches to fight pandemic flu.

Community Quarantine

Refers to the restriction of movement within, into, or out of the area of quarantine of individuals, large groups of people, or communities, designed to reduce the likelihood of transmission of an infectious disease among persons in and to persons outside the affected area.

Enhanced Community Quarantine (ECQ)

The enhanced community quarantine in Luzon is a series of stay-at-home orders and cordon sanitaire measures implemented by the Inter-Agency Task Force for the Management of Emerging Infectious Diseases (IATF-EID) on the island of Luzon and its associated islands. It is part of the COVID-19 community quarantines in the Philippines, a larger scale of COVID-19 containment measures with varying degrees of strictness. The "enhanced community quarantine" (ECQ) is the strictest of these measures and is effectively a total lockdown.

Quarantine Pass

A quarantine pass designates a household member as the representative of his or her household in procuring necessities. This disallows other family members to buy goods and conduct transactions outside their homes.

Essential goods and services

Covers health and social services to secure the safety and well-being of persons, such as but not limited to, food, water, medicine, medical devices, public utilities, energy, and others as may be determined by the IATF.

Forecasting

Forecasting is about predicting the future as accurately as possible, given all the information available, including historical data and knowledge of any future events that might impact the forecasts. Forecasting should be an integral part of the decision-making activities of management, as it can play an important role in many areas of a company. Modern organizations require short-term, medium-term, and long-term forecasts, depending on the specific application.

Time Series

A time series is a sequence of data points that occur in successive order over some period of time. This can be contrasted with cross-sectional data, which captures a point-in-time.

Time Series Forecasting

Time series forecasting uses information regarding historical values and associated patterns to predict future activity. Most often, this relates to trend analysis, cyclical fluctuation analysis, and issues of seasonality. As with all forecasting methods, success is not guaranteed.

Mean

The arithmetic mean, more commonly known as "the average," is the sum of a list of numbers divided by the number of items on the list. The mean is useful in determining the overall trend of a data set or providing a rapid snapshot of your data. Another advantage of the mean is that it's very easy and quick to calculate.

Regression

Regression models the relationships between dependent and explanatory variables, which are usually charted on a scatterplot. The regression line also designates whether those relationships are strong or weak. Regression is commonly taught in high school or college statistics courses with applications for science or business in determining trends over time.

Chapter Two

REVIEW OF RELATED LITERATURE

This chapter presents the different research and other literatures form both foreign and local researchers, which have significant bearings on the variables included in the research. It focuses on several aspects that will help in the development of this study. The literatures of this study come from books, journals, articles, electronic materials such as PDF or E-Book, and other existing thesis and dissertations, foreign and local which are believed to be useful in the advancement of awareness concerning the study.

2.1 Related Literature and Related Studies

COVID 19 Overview and its Transmission Mode

A cluster of individuals with pneumonia of unknown origin was linked to a seafood wholesale market in Wuhan, China, in December 2019. A study by Huang et al., (2020), had used an unbiased sequencing in samples from pneumonia patients on which betacoronavirus, the 2019 novel coronavirus (2019-nCoV), had been found. 2019-nCoV, unlike MERS-CoV and SARS-CoV, is the seventh member of the coronavirus family that infects humans (Huang et al., 2020).

Coronavirus is an enveloped, positive single-strand RNA virus. It belongs to the Orthocoronavirinae subfamily, as the name, with the characteristic "crown-like" spikes on their surfaces. The exact origin, location, and natural reservoir of the 2019- nCoV remain unclear, although it is believed that the virus is zoonotic, and bats may be the culprits because of sequence identity to the bat-CoV. (Perlman, 2020).

An article entitled “The outbreak of COVID-19: An overview” by Wu et al. (2020), discuss that nonspecific syndromes, such as fever, dry cough, and weariness, generally precede the onset of symptoms. Respiratory (cough, shortness of breath, sore throat, rhinorrhea, hemoptysis, and chest discomfort), gastrointestinal (diarrhea, nausea, and vomiting), musculoskeletal (muscle ache), and neurologic systems may all be affected (headache or confusion). After onset of illness, the symptoms are somehow mild and the median time to first hospital admission is 7.0 days (4.0–8.0). But the disease progresses to short of breath (~8 days), acute respiratory distress syndrome (ARDS) (~9 days), and to mechanical ventilation (~10.5 days) in about 39% patients (Huang et al., 2020). Patients with fatal disease develop ARDS and worsened in a short period of time and died of multiple organ failure. The mortality rate in the early series of hospitalized patients was 11%–15%, but the later statistics was 2%–3% (Huang et al., 2020).

There is currently no treatment for COVID-19 that has been proven to be effective. Symptomatic and supportive therapy, such as monitoring vital signs, maintaining oxygen saturation and blood pressure, and treating consequences such secondary infections or organ failure, are the key techniques (Wu et al., 2020).

The causal agent of coronavirus disease (COVID-19), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is thought to spread predominantly through respiratory droplets and close contact. In July 2020, 239 scientists signed an open letter, “appealing to the medical community and relevant national and international bodies to recognise the potential for airborne spread of covid-19 (Morawska and Milton, 2020)”. While a World Health Organization article admitted that "airborne transmission cannot be ruled out," the response was cautious and possibly incorrect in continuing to suggest that airborne and droplet transmission are distinct categories and that airborne transmission occurs only during medical "aerosol generating procedures”.

Contact, droplet, aerosol, fomite, fecal-oral, bloodborne, mother-to-child, and animal-to-human transmission are all probable routes of SARS-CoV-2 infection, according to the World Health Organization. WHO defines droplets as ≥5-10 μm diameter and aerosols as <5 μm. However, both can be generated as a continuum of particle sizes during numerous respiratory activities and their behaviours are not distinct. This has important practical implications for infection control, the prevention of outbreaks and superspreading events, and for the new social behaviours that are being implemented in an effort to control the pandemic. They defined that COVID-19 virus spreads mainly via droplets of saliva or nasal discharge ('mucosalivary droplets') emitted when an infected person coughs, sneezes, talks or breathes. Evidence from chamber studies and models of computational fluid dynamics provide perspectives on that process. During normal breathing and speech, particles are emitted by a mechanism involving 'fluid-film burst' in the small airways, which leads to the emission of particles ≤ 1 μm in diameter (Asadi et al., 2019). By contrast, the more forceful 'explosive' exhalations associated with sneezing, coughing, shouting and loud singing result in greater numbers of much larger particles (Zy et al., 2013) Thus, a social distance of 1–2 meters may guard against viral spread caused by civil speaking, but not by yelling, singing, coughing, or sneezing.

The primary goal of the WHO's COVID-19 Strategic Preparedness and Response Plan is to control COVID-19 by decreasing virus transmission and preventing related sickness and death. The virus is primarily communicated by touch and respiratory droplets, to the best of their knowledge. It's possible that aerial transmission will occur in some cases (such as when aerosol generating procedures are conducted in health care settings or potentially, in indoor crowded poorly ventilated settings elsewhere).

"Inhalational risk may be reduced by social distancing, limiting interaction indoors, avoiding air recirculation, improved natural and artificial ventilation, and innovative engineering solutions which collect and neutralise aerosols to provide clean air in personal and community spaces (Morawska and Milton, 2020)".

Non-Pharmaceutical Interventions for COVID-19

To combat the coronavirus disease (COVID-19) pandemic, a number of non-pharmaceutical interventions (NPIs) have been launched around the world. School closures, remote employment, and quarantine have all been used as social distancing (SD) measures in the past.

The 1918-19 H1N1 influenza pandemic was the last time the world responded to a global emergent disease epidemic on the size of the present COVID-19 pandemic with no access to vaccines. During that epidemic, certain communities, particularly in the United States (US), used a number of nonpharmaceutical interventions (NPIs) - methods aimed at decreasing transmission by lowering general population contact rates (Bootsma and Ferguson, 2007). Closing schools, churches, pubs, and other social facilities were among the tactics implemented at this time. Cities that implemented these treatments early in the pandemic were successful in reducing case numbers while the interventions were in place, and had lower total mortality (Bootsma and Ferguson, 2007).

While our understanding of infectious diseases and their control has advanced significantly since 1918, most countries throughout the world are currently confronted with COVID-19, a virus with mortality equal to H1N1 influenza in 1918. There are two basic tactics that can be used: Suppression and Mitigation (Anderson et al., 2020).

Suppression. The goal is to lower the reproduction number (the average number of secondary cases each case generates), R, to 1 and thereby reduce case numbers to low levels or eradicate human-to-human transmission (as in SARS or Ebola). The fundamental drawback of this strategy is that NPIs (and medications, if available) must be kept in place – at least intermittently – for as long as the virus is circulating in the human population or until a vaccine is developed.

Mitigation. The goal here is to employ NPIs (and, if available, vaccines or medications) to decrease the health impact of an epidemic, similar to the technique used by several US cities in 1918 and the rest of the world in the 1957, 1968, and 2009 influenza pandemics. Early vaccine supplies were focused at people with pre-existing medical disorders who were at risk of a more severe sickness during the 2009 pandemic, for example (World Health organization, 2009). In this scenario, population immunity develops over the course of the epidemic, resulting in a rapid drop in case numbers and low transmission levels.

A study entitled “Adoption and impact of non-pharmaceutical interventions for COVID-19” by Imai et al. (2020), summarizes the social distancing measures that were considered or implemented in response to the pandemic.

Table 2.1 Summary of social distancing measures considered and/or implemented in response to the COVID-19 epidemic

Measure	Description
Contact tracing	Identifying individuals who might have been in contact with a confirmed case.
Isolation	Separation of ill persons with contagious diseases from susceptible persons.
Quarantine	Restriction of persons who are presumed to have been exposed to a contagious disease but are not ill, either because they did not become infected or because they are still in the incubation period or because they did not become infected
School closures	Closure of schools nationally or across a region. This is distinct from reactive closures of schools in response to identified cases.
Workplace closure and measures	Closure of workplaces and advisories to work remotely.
Crowding	Advisories to avoid crowded places such as concerts. This includes mandatory cancellations of mass gatherings such as conferences, weddings, and funerals.
University closure	Regional or nationwide closure of universities.

Countries and regions afflicted by the COVID-19 epidemic have employed social distancing measures to varying degrees. The most severe interventions took place in Hubei Province (China), when 40–60 million people were subjected to severe travel restrictions

(Wu and McGoogan, 2020). Outside of Hubei province, China, where long-term Social Distancing layered with strict movement restrictions in Wuhan City and Hubei have reduced the reproduction number R 0, which was estimated to be greater than 2 during the early stages of the outbreak, it is likely too early to evaluate or quantify the true effectiveness of specific SD interventions on the outbreak (Zhang et al., 2020).

In fact, because most countries have introduced a variety of non-pharmaceutical methods to encourage behavior change, such as travel restrictions, health screenings, and guidance on hand and cough cleanliness, it's impossible to estimate the effectiveness of Social Distancing in the absence of additional control measures. However, early research has revealed that case isolation and contact tracing are more successful than travel limitations or contact reduction. They also discovered that if these combined NPIs had been implemented one, two, or three weeks sooner, case numbers might have been decreased by 66 percent, 86 percent, and 95 percent, respectively, up to three months after their introduction (Lai et al., 2020).

The timing and duration of Social Distancing interventions have also been demonstrated to affect their effectiveness in pandemic influenza studies (Imai et al., 2020). Ferguson et al. observed that the effectiveness of social separation, rapid case ascertainment, and targeted prophylaxis were similar when evaluating targeted layered containment techniques, with school closures playing a key part in each scenario, especially if R0 values were less than 2. Various levels of evidence for avoiding crowding, workplace measures, and case isolation in the community were found in a comprehensive evaluation of the effectiveness of SD strategies for pandemic influenza (Lai et al., 2020).

The majority of COVID-19 isolation has occurred in a hospital setting. As the number of cases reported in the community grows, case isolation techniques may shift to voluntary home isolation or household quarantine. Household influenza quarantine was found to have an overall effect, although it may raise the risk of infection among quarantined persons inside an infected household. When combined with other interventions like quarantine and isolation, other resource-intensive methods like contact tracking have been demonstrated to be successful in preventing influenza spread (Imai et al., 2020).

Quick Response Code

The acronym QR Code refers to a machine-readable optical label that contains information about the related item or product. When compared to 1-D Codes, 2-D Codes can store more data in a less amount of area. Unlike in Bar Codes on which information is only coded in one direction, QR code information is encoded in two directions: horizontally and vertically

Table 2.2 Comparison of QR Code and Bar Code

Features	QR Code	Bar Code
High Capacity	Upto 7089 numeric digits	10-20 digits
Durability against damage	Reading is possible (upto 30% damaged)	Reading is impossible
Reduced Space	40 digits numeric	10 digits numeric
360◦ degree	Supports 360◦	Horizontal reading
Language Supported	Numeric, Alphanumeric, Kanji, Kana etc.	Numeric, Alphanumeric

Denso Wave, a Toyota subsidiary firm in Japan, was the first to deploy QR codes in 1994. Other industries attempted to embrace this technology after its successful adoption at Denso Wave. The QR code is patented by Denso Wave, although it can be used elsewhere in the world. In 2011, the telecoms industry was the first to commercialize the QR code.

The popularity of these codes is primarily due to the following features, according to an article titled "Two-Level QR Code for Private Message Sharing and Document Authentication" (Tkachenko et al, 2016): they are robust to the copying process, easy to read by any device and any user, they have a high encoding capacity enhanced by error correction facilities, they have a small size and a low cost.

Those evident advantages, however, are not without their drawbacks:

1) Even though the information encoded in a QR code is ciphered and hence only legible to authorized users (the distinction between "see" and "understand"), it is always accessible to everyone.

2) Due to its insensitivity to the Print-and-Scan (P&S) procedure, it is impossible to identify an original printed QR code from a copy.

This machine-readable matrix code is made up of black and white squares. It can hold URL (Uniform Resource Locator) information, contact information, links to videos or photographs, plain text, and other sorts of material (ISO/IEC 18004, 2000).

The QR code architecture is discussed in a study titled "QR Code Analysis" (Singh, 2016). Each QR code symbol has a square pattern to it. There are two regions in this square pattern: the encoding region and the function patterns. The location where the encoding region indicates the data encoding is the focus of the function patterns. Finder patterns, timing patterns, and alignment patterns are all part of the function pattern. Finder patterns are three frequent structures found on the three corners of a QR code symbol. The Finder pattern is used to determine the symbol's proper orientation. The decoder software uses timing patterns to determine which side of the pattern to use. In the case of image distortion, alignment patterns are utilized to ensure that decoder software accurately decodes the symbol. Other than the function pattern, the rest of the region is the encoded region, which stores data code words and error correcting code words.

Error correction level and mask pattern are also included in the format information regions. The version information regions store the code version and error correcting bits. Information encoding utilizing the Reed-Solomon error correction code, information division on codewords, application of the mask pattern, and placement of codewords and function patterns into the QR code are all part of the QR code generating algorithm. The scanning process, picture binarization, geometrical correction, and decoding algorithm are all part of the QR code recognition method (Tkachenko et. al, 2016).

The following are the main characteristics of QR codes (according to Singh, 2016):

1) High Storage Capacity

When compared to a 1-D barcode, a QR code symbol can store up to 7,089 characters of data, which is a tremendous number.

2) Encodable Character Set

Numeric data (Digits 0-9)

Alphanumeric data (upper case letters A-Z; Digits 0 - 9; nine other characters: space, : % * + - / _ $)

Kanji characters

3) Small Printout Size

The information in a QR code is stored in two directions: horizontally and vertically. Because of this feature, QR codes take up a fourth of the area that a 1-D barcode takes up for the same amount of data.

4) 360 Degree Reading

The QR code may be read from any angle. The finder patterns found in three corners of the symbol provides this feature. The finder pattern aids in the detection of the QR code.

5) Capability of Restoring and Error Correction

Data can be recovered if a component of the code symbol is destroyed or unclean. The error-detection procedure can concentrate on the proper information region. L, M, Q, and H are the four layers of error correction for QR codes. Level L has the least error correction capabilities, whereas level H has the strongest.

QR codes are employed in the majority of commercial market items in several nations. In essence, QR codes are a practical way to combine the virtual and physical worlds to give useful information at a moment's notice. QR codes are a low-cost technology that is simple to use and implement. QR codes have a wide range of uses in a variety of disciplines. The speed and convenience with which information is given is the true value of QR code access.

In a study conducted by Uzun (2016), a healthcare application was created consisting of QR-codes and QR-code reader applications, installed on smartphones or tablets, were placed in various locations of the hospital, which scans the QR codes to obtain vast amount of information. In the same study, they implemented a QR Code Identity Tag which allows for quick access to important medical information. The general information about the patient, including name, address, and emergency point of contact, will be displayed after the QR code is scanned by the reader, such as a smartphone or any other electronic device capable of scanning.

An article entitled “The Usefulness of the QR Code in Orthotic Applications after Orthopedic Surgery” (Cho et al., 2021), they pointed out the significance of QR code, in the orthotic field due to its advantages of repeatability and convenience. The usage of QR code prevented the dissemination of false orthosis information to the patients, because it provided audiovisual information in an easy way via a smart phone. They emphasized that QR code can be used more effectively by repeated patient education and accessibility than conventional oral training in an outpatient clinic.

During the current COVID-19 pandemic and previous pandemics, a variety of digital health initiatives were used to control disease spread. These control strategies have been shown to be effective in decreasing COVID-19's effect in a number of nations.

A study conducted by Nakamoto et al. (2020), proposed a pandemic management paradigm based on symptom-based quick reaction (QR) codes to contain the spread of

COVID-19 QR codes were designated as electronic certificates of an individual's health state under this framework, and they can be utilized for contact tracking, exposure risk self-triage, self-update of health status, health care appointments, and contact-free psychiatric consultations. Since the first case was officially reported in January 2020, the framework outlined in this paper has effectively restricted the spread of COVID-19 in

Fujian Province, which has a population of about 40 million people. As of July 12, 2020, 361 (99.4%) of the 363 reported cumulative cases had been recovered, 1 patient had died (mortality rate 0.3%), and one patient had remained positive for COVID-19 (0.3%). Due to the approach's early deployment and strict application, it was able to successfully transition the GDP from –6.8% to a positive figure by July. Firms have been able to progressively resume normal operations while maintaining efficient containment. The technology has aided government agencies in achieving successful containment and travel control.

In an article entitled “Usefulness of an Online Preliminary Questionnaire under the COVID-19 Pandemic” by Hur and Chang (2020), it has been discussed how a mobile self-report questionnaire with quick response (QR) codes can help prevent hospital-acquired secondary infections by reducing contact between patients with COVID-19 and hospital personnel or other patients inside the hospital. The authors have emphasized how the method will save time and human resources needed for investigation, as well as minimize any potential turmoil in the hospital as a result of such screening activity.

Time Series Forecasting

Forecasting future values of an observed time series is useful in almost every aspect of research and engineering, including economics, finance, business intelligence, meteorology, and telecommunications (Palit and Popovic, 2005).

A forecasting method is a method for calculating forecasts based on current and historical data. As a result, it might just be an algorithmic rule with no need for an underlying probability model. Alternatively, it could result from establishing a specific model for the provided data and determining the best forecasts based on that model. As a result, the terms "method" and "model" should be maintained separate. Unfortunately, the word "forecasting model" is thrown around in the literature with a lot of looseness, and it's occasionally used incorrectly to define a forecasting strategy (Chatfield, 2000).

In his book "Time-series Forecasting," Chatfield (2000) divided forecasting methodologies into three categories:

(a) Predictions based on subjective judgment, intuition, "inside" commercial knowledge, and any other relevant data.

(b) Univariate approaches, in which forecasts are based solely on the current and previous values of a single series, potentially supplemented by a time function such as a linear trend.

(c) Multivariate approaches, in which the forecasts of one or more extra time series variables, referred to as predictor or explanatory variables, are based, at least in part, on the values of one or more additional time series variables.

Authors reviewed how different intervals were employed in diverse literature in an article titled "Financial time series forecasting with machine learning techniques: A survey" (Krollner et al, 2010). The varied predicting intervals utilized in the literature are depicted in Figure 2.1. The prediction periods are divided into three categories: one day, one week, and one month ahead. 'Several / Others' lists publications that use multiple or distinct time frames. The majority of periodicals make one-day forecasts, such as projecting the next day's closing price. However, being able to anticipate the stock index one day ahead of time does not always imply that an investor may profit from this information in terms of trading, especially because the index is not traded.

Time-frame	Number	Publications
Day	31	[1], [2], [3], [4], [6], [7], [8], [9], [10], [13], [14], [17], [19], [20], [21], [22], [24], [27], [28], [31], [32], [33], [34], [35], [36], [37], [40], [41], [42], [44], [45]
Week	3	[18], [23], [43]
Month	3	[26], [38], [39]
Multiple / Other	9	[5], [11], [12], [15], [16], [25], [29], [30], [46]

Figure 2.1 Sample Time Series Data

A time series is denoted by Y = y1, y2...,yn . The technique of estimating future values of Y, $yn + h$, where h specifies the forecasting horizon, is known as forecasting. There are two types of quantitative approaches to time series forecasting: univariate and multivariate. Univariate methods are procedures that use previous observations to model future observations of a time series. Multivariate techniques build on univariate approaches by taking into account additional time series as explanatory variables. In this paper, we'll concentrate on univariate techniques. When dealing with time series prediction difficulties, another factor to consider is the forecasting horizon. One step ahead forecasting, or the prediction of the next value of a time series (yn+1), is usually the focus of forecasting methodologies. Occasionally, one is intrigued by the prospect of foreseeing many steps into the future. Multi-step forecasting is a term used to describe these jobs (Taieb et al., 2012).

The authors reviewed typical time series models in a study titled "Machine Learning vs Statistical Methods for Time Series Forecasting: Size Matters" (Cerqueira et al, 2019).

1. Naive method, also known as the random walk forecast, predicts the future values of the time series according to the last known observation:

$$y^{\wedge} n + \hbar = y_n$$

2. The seasonal naïve model functions in the same way as the naive technique. The seasonal naïve technique differs in that it employs a previously known value from the same season as the intended forecast:

$$y^{\wedge}_{\ n} + \hbar = y_n + \hbar - m$$

   where *m* denotes the seasonal period.

3. One of the most often used approaches for modeling univariate time series is ARMA (Auto-Regressive Moving Average). ARMA(p, q) combines two components: AR(p), and MA(q). According to the AR(p) model, the value of a

given time series, y_n, can be estimated using a linear combination of the *p* past observations, together with an error term $\in_n$ and a constant term *c*.

$$y_n = c + \sum_{i=1}^{p} \emptyset_i \, y_{n-1} + \in_n$$

4. Exponential smoothing model is similar to the AR(p) model in the sense that it models the future values of time series using a linear combination of its past observations. Exponential smoothing methods produce weighted averages of the past values, where the weight decays exponentially as the observations are older.

$$y_{n+1} = \; y_n\beta 0 + y_n - 1\beta 1 + y_n - 2\beta 2 +. \; . \; .$$

Recursive, Direct, and DirRec are the three basic techniques for Multi-step Time Series Forecasting. In the Recursive strategy (Weigend and Gershenfeld, 1994), a one-step model is first trained. *F*

$$y_{t+1} = f(y_t, \ldots, y_{t-n+1}) + w_t + 1,$$

with $t \in \{n, \ldots, N - 1\}$ and then uses it recursively for returning a multistep prediction. The recursive method's sensitivity to estimation error is well-known, as estimated values, rather than actual ones, are increasingly used as we move further into the future.

The Direct strategy (Cheng et.al, 2006) learns independently *H* models f_h

$$y_{t+h} = f_h(y_t, \ldots, y_{t-n+1}) + w_{t+h},$$

with $t \in \{n, \ldots, N - H\}$ and $h \in \{1,\ldots,H\}$ and returns a multi-step forecast by concatenating the H predictions. The Direct method is immune to accumulation of errors because it does not employ any approximated data to produce forecasts. Despite this, there are certain flaws. First, since the *H* models are learned independently no statistical dependencies between the predictions $\hat{}y_{N+h}$ is considered. To model the stochastic dependency between two series values at two distant instants, second direct approaches sometimes require more functional complexity than iterated methods. Finally, because the number of models to learn is equal to the size of the horizon, this technique necessitates a lengthy processing time.

Sorjamaa and Lendasse (2006) proposed the DirRec approach, which combines the designs and ideas of the Direct and Recursive techniques. DirRec computes forecasts using several models for each horizon (similar to the Direct technique) and expands the set of inputs at each time step by adding variables according to the previous step's forecasts (like the Recursive strategy). However, unlike the previous two solutions, the embedding size n is not uniform across all horizons. To put it another way, the DirRec strategy learns H models f_h from the time series $[y1, \ldots, yN]$ where

$$y_{t+h} = fh(y_{t+h-1}, \ldots, y_{t-n+1}) + w_{t+h} \,,$$

with $t \in \{n, \ldots, N-H\}$ and $h \in \{1,\ldots,H\}$.

Despite the fact that several nations have already begun COVID-19 mass vaccination campaigns in order to quickly manage the illness outbreak, a more contagious and lethal version of coronavirus is causing extraordinary spikes in new COVID-19 cases in many countries throughout the world. As the number of new cases rises, government officials face additional problems in combating the pandemic, including pandemic fatigue and public apathy toward various intervention measures. As a result, it is critical for government authorities to have a thorough understanding of COVID-19's future dynamics in order to establish strategic readiness and resilient response plans.

Mathematical models have been crucial in the ongoing crisis; they have been used to inform governmental policies and have been helpful in many of the social distancing measures that have been implemented around the world.

A study conducted by Li et al. (2021), DELPHI, a unique epidemiological model, was created to model the impact of under-detection and government action on COVID 19 transmission. If the limits were applied one week sooner, the DELPHI model estimates a 75 percent drop in both cases and deaths for several countries around the world. If limits had been enforced sooner, Western European countries like as Switzerland, Spain, and Italy would have benefited the most. This supports the theory that these areas saw some of the greatest outbreaks outside of Asia, and hence did not have as much time to respond as nations like Romania and Iceland, which experienced an outbreak later. DELPHI estimates that if every country in the world enforced its limits one week sooner, nearly 280,000 deaths, or 68 percent of the world's total mortality count, could have been prevented by May 17th of 2021. This reveals that extreme actions taken by governments and communities around the world preserved a considerable section of the global population.

The authors of the study entitled "Mathematical Models for COVID-19 Pandemic: A Comparative Analysis" by Adiga et al. (2020), discusses a few key computational models for COVID-19 pandemic preparation and response created by academics in the United States, the United Kingdom, and Sweden. Policymakers and public health officials in each country have utilized the models to assess the pandemic's evolution, design and analyze control measures, and investigate potential what-if scenarios. As previously stated, all models were challenged by a lack of data, a quickly expanding epidemic, and extraordinary control efforts. Despite these obstacles, they believed that mathematical models can offer policymakers with important and timely information.

In past studies, many researchers used various time series models to forecast pandemic occurrence. Exponential smoothing (Tseng & Shih, 2019) generalized regression (Imai et al., 2015), multilevel time series models (Spaeder & Fackler, 2012), and autoregressive integrated moving average (ARIMA) models are used in several of these methods (Li et al., 2012).

ARIMA

A family of stochastic processes known as auto regressive integrated moving average (ARIMA) processes is used to evaluate time series (Box and Jenkins, 1994). The following is the general plan:

Step 0) Certain hypotheses are assumed in the formulation of a class of models.

Step 1) For the observed data, a model is found.

Step 2) The parameters of the model are estimated.

Step 3) If the model's hypotheses are confirmed, proceed to Step 4, otherwise, return to Step 1 to revise the model.

Step 4) Forecasting is now possible with the model.

A straightforward type of an AR model of order p, i.e., *AR(p)*, can be written as a linear process given by:

$$x_t = c + \sum_{i=1}^{p} \emptyset_i x$$

Where x_t is the stationary variable, *c* is constant, the terms in φ_i are autocorrelation coefficients at lags 1, 2, *p* and *t*, the residuals, are the Gaussian white noise series with mean zero and variance σ_2. An MA model of order q, i.e., *MA(q)*, can be written in the form:

$$x_t = \mu + \sum_{i=0}^{q} \theta_i \, \epsilon_{t-i}$$

Where μ is the expectation of x_t (usually assumed equal to zero), the θ_i terms are the weights applied to the current and prior values of a stochastic term in the time series, and $\theta_0 = 1$. We assume that *t* is a Gaussian white noise series with mean zero and variance

σ^2. We can combine these two models by adding them together and form an ARIMA model of order (*p*, *q*):

$$x_t = c + \sum_{i=1}^{p} \emptyset_i \, x_{t-i} + \epsilon_t + \sum_{i=0}^{q} \theta_i \, \epsilon_{t-i}$$

Where $\varphi_i = 0$, $\theta_i = 0$, and $\sigma^2 > 0$. The parameters p and q are called the AR and MA orders, respectively. Because of its "integrate" stage, ARIMA forecasting, also known as Box and Jenkins forecasting, can deal with non-stationary time series data. In fact, the "integrate" component entails differencing the time series in order to convert a non-stationary one into a stationary one. ARIMA(p, d, q) is the general form of an ARIMA model (Siami-Namini et.al, 2018).

ARIMA modeling is one of the best modeling strategies for forecasting a time series. In terms of natural catastrophe prediction, this model outperforms other models such as the wavelet neural network (WNN) and the support vector machine (SVM) (Zhang et al., 2019). A study entitled "Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future" by Tandon et al., suggested that ARIMA model showed the most accurate forecast compared to Linear Trend, Quadratic Linear, S-Curve Trend, Moving Average, Single Exponential as well as Double Exponential models in predicting future COVID-19 instances in India.

Chapter Three

THEORETICAL FRAMEWORK

Quarantine Pass Manual Process

Upon issuance of IATF resolution declaring high-risk Provinces, Highly Urbanized Cities (HUCs), and Independent Component Cities (ICCs) under localized Enhanced Community Quarantine, the Local Government Units (LGUs) are advised to implement lockdown procedures limiting people mobility for essential activities only such as buying of food, groceries, medicines etc.

To ensure the public's access to essential goods quarantine passes are issued by respective LGUs to residents within their jurisdiction. Only one quarantine pass will be given per household.

The generation and issuance of quarantine pass is purely manual process done by LGU officials. Quarantine pass templates are printed out on a piece of paper and distributed either in the Barangay halls or delivered on house-to-house basis in their areas of responsibility.

At checkpoints/essential establishment, the quarantine pass will need to be presented to the checkpoint/barangay officials for manual verification.

Quick Response Code

QR Code is a two-dimensional symbol. Denso, a significant Toyota group company, designed it in 1994, and it was accepted as an ISO international standard (ISO/IEC18004) in June 2000. This two-dimensional sign was created with the intention of being used in the production control of automotive parts, but it has since extended to other industries.

They encode either an e-mail address or a specific Web site URL, allowing smartphone users to directly access Web sites encoded by QR codes without having to type, copy, or memorize the Web site address.

QR codes can hold a variety of information in addition to being used to transmit links:

- A business card can have a QR code that contains an electronic version of the contact information. When you scan the code, the reader app adds the contact to your address book.
- Event details can be stored as a QR code. Scanning the code on a concert poster prompts the software to add the event's name, date, and location to your smartphone or computer's calendar.
- An SMS with a phone number and text can be stored in a QR code. Scan the code, and the scanning software will immediately enter you into a contest where you can win fantastic prizes.
- An e-mail message with a topic and message text can be stored in a QR code. That message could be a request for information, and you might receive a response email with additional information and files attached.
- A geographical location can be encoded in a QR code. Scanning the code on a restaurant poster makes the location of the restaurant available to your navigation program, which will tell you how to get there.
- WIFI configuration data can be stored in a QR code. After scanning the code, your Android device will immediately set itself to use the hotel's wireless network.

Description of Quick Response bar codes

In 1994, Denso-Wave, a Japanese company, invented the QR matrix code. It is an open standard that does not need the payment of a license fee. Various standards bodies, such as JIS and ISO (for example, the ISO/IEC 18004:2006 standard), are currently in charge of the physical encoding of QR codes. NTT DoCoMo, a Japanese telecom firm, devised the standard for encoding URLs.

Both the horizontal and vertical axes of QR codes carry information. When compared to ‘traditional' barcodes, this allows for significantly more raw data to be inserted. These can be numeric, alphanumeric, or binary data, with a maximum storage size of 2953 bytes. Actual data, including mistake correction information, is only found in a portion of each QR bar code. The URL data from the above QR code has been removed in the image below. As you can see, a significant portion of the bar code is dedicated to describing the data format and version, as well as positioning, alignment, and timing.

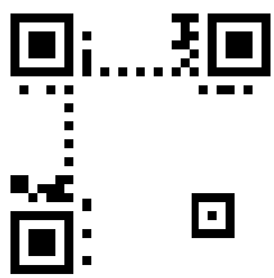

Figure 3.1 Positioning, alignment, data in a QR code

The larger the barcode, the more data that must be encoded. This page's QR code may be found below. Because the URL is longer than the home page, the bar code has grown as well. The following barcode does not include a URL, but rather the first five sentences of this page.

Figure 3.2 QR barcode pointing to a web page

Figure 3.3 QR barcode with five lines of text in it

A module is the smallest square dot or pixel element in a QR code. An empty region around the graphic, as with other forms of bar codes, is encouraged to make it easier for devices to read the bar code. This peaceful region should be at least four modules wide.

The resolving power of the cameras used to scan the code determines the minimum dimensions of a QR code. For QR codes that contain a URL, a minimum size of 32 32 mm or 1.25 1.25 inches, excluding a silent zone, is recommended, according to a Kaywa white paper. This ensures that all camera phones on the market can read the bar code correctly. Changing the size to 26 mm in width and height (approximately 1 square inch) still covers 90% of the phones on the market. However, newer camera models with increased macro capabilities can already handle QR codes that are less than 10 mm (0.4″) broad and high.

The preceding rule applies to codes that are perfectly printed and to which the user has direct access. When QR codes are used on a poster or billboard, things change. The common belief is that the physical dimensions of a QR code and its scanning distance are directly proportional. That ratio is about 1/10, thus if the reader is 50 centimeters away from the code, the QR code should be at least 5 centimeters in size. The height of the code should be at least 1 meter for a billboard that can be seen from a distance of 10 meters.

A good contrast between the background and the bar color itself is critical for reading accuracy. A dark color should be used on a bright background for the bar code. It is typically printed black on white background. If a colored background is required, it should be a solid color rather than a screening tint. Cyan or magenta should be avoided, but a solid yellow backdrop should suffice. If the contrast with the bar code is high enough, very light Pantone colors would also work.

Autoregressive Integrated Moving Average (ARIMA) Model

To understand the data and make future forecasts, the Autoregressive Integrated Moving Average (ARIMA) model employs time-series data and statistical analysis. The ARIMA model seeks to explain data by using time series data on previous values and making predictions using linear regression.

Each of the components of an ARIMA model can be understood by outlining them as follows:

- Autoregression (AR): a model in which a changing variable regresses on its own lagged (or prior) values.
- Integrated (I): denotes the differencing of raw observations to allow the time series to stabilize (i.e., data values are replaced by the difference between the data values and the previous values).
- Moving average (MA): A moving average model applied to lagged observations incorporates the dependency between an observation and a residual error.

ARIMA Parameters

ARIMA treats each component as a parameter with a consistent nomenclature. ARIMA with p, d, and q is a standard notation for ARIMA models, where integer values replace the parameters to denote the kind of ARIMA model utilized. The parameters are as follows:

- **p**: also called the lag order which refers to the number of lag observations in the model.
- **d**: The degree of differencing refers to the number of times the raw observations are differed.
- **q**: The order of the moving average is the size of the moving average window.

The parameters are integers that must be defined in order for the model to work. They can also be set to 0, indicating that they would be ignored in the model. The ARIMA model can then be transformed into:

- ARMA model (no stationary data, d = 0)
- AR model (no moving averages or stationary data, just an autoregression on past values, d = 0, q = 0)
- MA model (a moving average model with no autoregression or stationary data, p = 0, d = 0)

Therefore, ARIMA models may be defined as:

- ARIMA (1, 0, 0) – known as the first-order autoregressive model
- ARIMA (0, 1, 0) – known as the random walk model
- ARIMA (1, 1, 0) – known as the differenced first-order autoregressive model, and so on.

The ARIMA model seeks to estimate the coefficients α and θ, which is the outcome of using prior data points to anticipate values, once the parameters (p, d, q) have been set.

Limitations of the ARIMA Model

Although ARIMA models can be highly accurate and dependable under the right conditions and with enough data, one of the model's major drawbacks is that the parameters (p, d, and q) must be manually specified, making obtaining the best fit a lengthy trial-and-error process.

Similarly, the model is heavily reliant on the consistency and differencing of previous data. To ensure that the model offers reliable results and forecasts, it is critical to ensure that data was collected accurately and over a lengthy period of time.

Box- Jenkins Methodology

In 1970, statisticians George Box and Gwilym Jenkins published a book that described the Box-Jenkins technique to modeling ARIMA dynamics. An ARIMA process is a forecasting mathematical model. Identifying a suitable ARIMA process, fitting it to the data, and then utilizing the fitted model for forecasting are all part of Box-Jenkins modeling. One of the appealing aspects of the Box-Jenkins approach to forecasting is that ARIMA processes are a large class of models from which to choose, and it is usually possible to find one that adequately describes the data.

Model selection, parameter estimation, and model checking were all iterative three-stage processes in the original Box-Jenkins modeling approach. Makridakis, Wheelwright, and Hyndman (1998) added a preliminary stage of data preparation and a final stage of model application to their explanations of the process (or forecasting).

1. Data transformations and differencing are involved in data preparation. Data transformations (such as square roots or logarithms) can aid in the stabilization of variance in a series where variation varies with level. When it comes to business and economic data, this is a common occurrence. The data is then differentiated until no clear patterns, such as trend or seasonality, remain. The

term "differencing" refers to calculating the difference between two consecutive observations or two observations separated by a year. Modeling the differenced data is frequently easier than modeling the original data.

2. In the Box-Jenkins framework, model selection employs a variety of graphs based on converted and differenced data to try to discover probable ARIMA processes that might provide a good match to the data. Other model selection criteria, such as Akaike's Information Criterion, have emerged as a result of later advances.
3. Finding the values of the model coefficients that provide the best fit to the data is known as parameter estimation. This can be accomplished using powerful computational methods.
4. Model checking is putting the model's assumptions to the test in order to identify any flaws. If the model is judged to be insufficient, you must return to Step 2 and try to find a better model.
5. The entire technique is intended to accomplish forecasting. It is usually a simple operation to compute forecasts once the model has been chosen, estimated, and validated. Of course, this is accomplished through the use of a computer.

Although Box and Jenkins' fundamental technique was developed for modeling time series with ARIMA processes, it can be used to a wide range of statistical modeling applications. It gives a useful framework for an analyst to think about data and find a statistical model that can be utilized to answer relevant data questions.

3.1 Conceptual Framework

This section aims to demonstrate the overview of the final product of this capstone project. An I-P-O (Input-Process-Output) model will be used as the conceptual schema of the system It identifies relevant variables, inputs, mappings, and other components and how they will interact with each other. This includes all the underlying concepts and their associated mappings based on the system’s use.

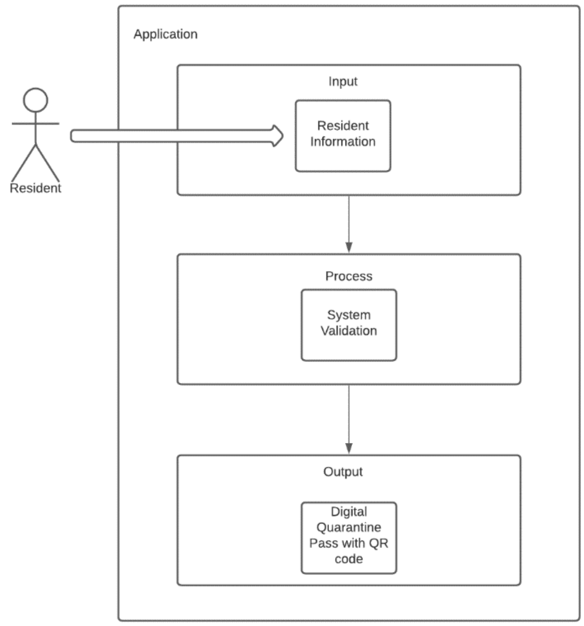

Figure 3.4 IPO Diagram of Quarantine Pass Generation

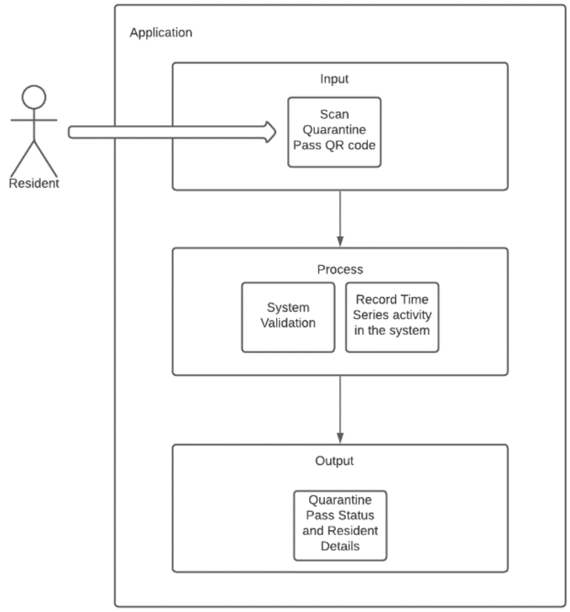

Figure 3.5 IPO Diagram of Quarantine Pass Validation

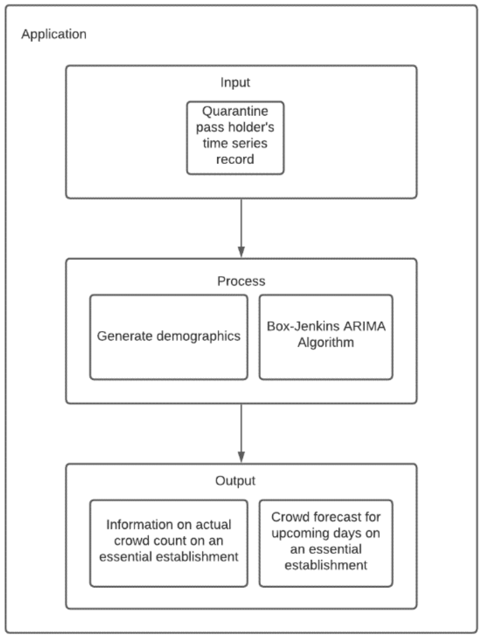

Figure 3.6 IPO Diagram for Decision Support System

The users of the proposed application are the residents who need to acquire a quarantine pass for access and procurement of essential goods or services identified by the IATF guidelines, local government officials implementing the quarantine policy, and essential establishments personnel allowed to operate during the quarantine period. They will have access to the same mobile responsive web application but will have role segregation to define their level of access.

The diagram on figure 3.4 and figure 3.5 illustrates the first component of the system where the system provides an automated process of generating and validating quarantine passes. Once a resident has been granted with quarantine pass, they will be able to download the generated QR code and present it to checkpoints/essential establishments whenever they go outside of residence. On the other end, when an essential establishment personnel scans the QR code, the system will display the information of the authorized quarantine pass holder along with the status of the quarantine pass. This action will then record the timestamp of scan in/scan out of the quarantine pass holder.

Diagram on figure 3.6 demonstrates how the data captured from quarantine pass holder’s mobility history can be utilized in generating insightful data regarding the current crowd count on a specific essential establishment and generate a one week forecast.

Chapter Four

METHODOLOGY

The proponent of this capstone project used prototype method in delivering the objectives of this project.

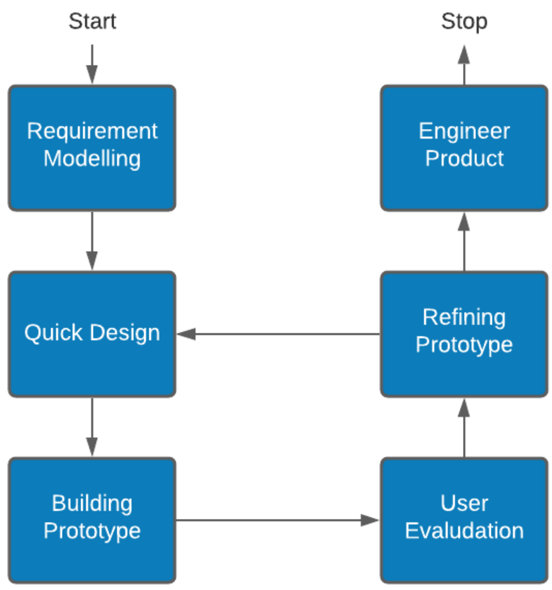

Figure 4.1 Prototype Model

Figure 4.1 shows the Prototype Model used by the proponent as a guide in developing the project entitled “SafePass: An Implementation of Autoregressive Integrated Moving Average (ARIMA) for Crowd Forecasting Applied in Quarantine Pass” which is a Systems Development Methodology (SDM) within which a paradigm output (or an early approximation of a final system or product) is constructed, tested, and then reworked.

Using this model, it will enhance the usability, design quality of the proposed application and it will also make the development process more cost-efficient since the development cycle becomes shorter.

The phases of the prototype model involve the following steps:

4.1 Requirements Modeling

This first step involves understanding the basics product requirements. The proponent of this capstone project possess firsthand experience on how the Quarantine Pass is obtained and utilized when going outside of residence. Apart from the proponent's personal knowledge, it is also in this stage that the proponent had requested permission from local government authorities and other parties to conduct the study and where all relevant data and information were examined.

To build a logical model of the application, preliminary investigation was conducted via interviews with the involved parties. With these interviews, the researcher was able to identify the transactions involved and analyzed them against the proposed solution. This collected information had enabled the researcher to identify critical decisions geared toward implementing the proposed application.

In Molino IV, a barangay in the city of Bacoor, Cavite, the process of implementing home quarantine passes is done manually by designated LGU or barangay officials. Quarantine Pass templates are printed on a paper with the details of their barangay and some blank fields to be manually populated by the individual who will utilize it. The dissemination process of quarantine pass is for the barangay officials to deliver on house-to-house basis. In case the issued quarantine pass was lost, the resident should request for a replacement from their respective Barangay Hall.

In the 2015 census done by Philippine Statistics Authority, the estimated population in Barangay Molino IV, Bacoor Cavite is 51, 362 and the average number of people per household is 4. Given this data, there would be around 12,841 quarantine passes to be issued to each household on a manual basis requiring a lot of time and effort.

At each designated checkpoints or basic commodity establishments, residents are then required to present their quarantine pass allowing them access to the establishment. The probability of having crowd congestion in a specific establishment is high since the residents do not have visibility on the actual number of quarantine pass holders that are inside an essential establishment on a specific period. Additionally, instances of long queue in grocery stores are mostly observed as only 50% capacity is allowed inside the establishments.

It is for these reasons that the researcher opted to design and develop an application that would not only automate the manual and time-consuming processes but also to expand in utilizing the gathered data on quarantine pass holder's mobility activity to display useful information and statistics for decision making and risk management.

Figure 4.2 Sample Quarantine Pass

In compliance with the Data Privacy Act of 2021, the data to be used in the study such as the Resident Personal Information, Household Demographics, and Essential Establishment details will be dummy data.

Below are the functional requirements grouped by specific role:

- Resident
    - Register and Login – Register to gain access to the system.
    - Generate Quarantine Pass – Generate a digital quarantine pass by providing personal information to be validated by the system.
    - View Statistical Information – Access to the dashboard on actual crowd count and forecast per registered essential establishment.
- Essential Establishment Personnel
    - Register and Login – Register to gain access to the system.
    - Scan Quarantine Pass – Scan quarantine passes for individuals entering and leaving the essential establishment premises.
    - View Statistical Information - Access to the dashboard on the establishment's actual crowd count and forecast.

In generating the digital quarantine pass, QR code scheme will be implemented to help to speed up the flow of information. This would allow residents to download the QR code that will be scanned at essential establishment premises upon entry and exit.

Date time stamps of the entry and exit activity will be logged in the system and will be analyzed to generate crowd demographics. ARIMA model will be used to analyze the time series data on SafePass.

Forecasting using Auto Regressive Integrated Moving Average (ARIMA) Model

A time series is a collection of measurements taken at regular intervals. A time series can be yearly (for example, an annual budget), quarterly (for example, expenses), monthly (for example, air traffic), weekly (for example, sales quantity), daily (for example, weather), hourly (for example, stock price), minutes (for example, inbound calls in a call center), or even seconds in length, depending on the frequency (ex: web traffic).

The ARIMA model is a famous and commonly used statistical approach for time series forecasting.

$$Y_t = a + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_p Y_{t-p} \epsilon_t + \emptyset_1 \epsilon_{t-1} + \emptyset_2 \epsilon_{t-2} + .. + \emptyset_q \epsilon_{t-q}$$

ARIMA model in words:

Predicted Yt = Constant + Linear combination Lags of Y (upto p lags) + Linear Combination of Lagged forecast errors (upto q lags)

The first step is to look at the autocorrelations and partial autocorrelations shown on the Mobility Time Series data. An autocorrelation pattern like in figure 4.3 will appear in a series with a trend:

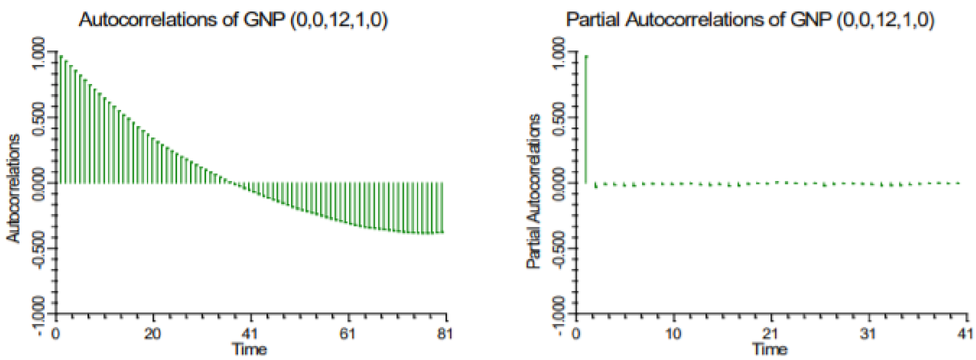

Figure 4.3 Sample of Non-Stationary Series Plot

If the data series in the system has substantial autocorrelations, the system will need to execute differencing, which involves transforming the original series, xt, to become stationary around its mean and variance. The autocorrelation plots could look like in figure 4.4:

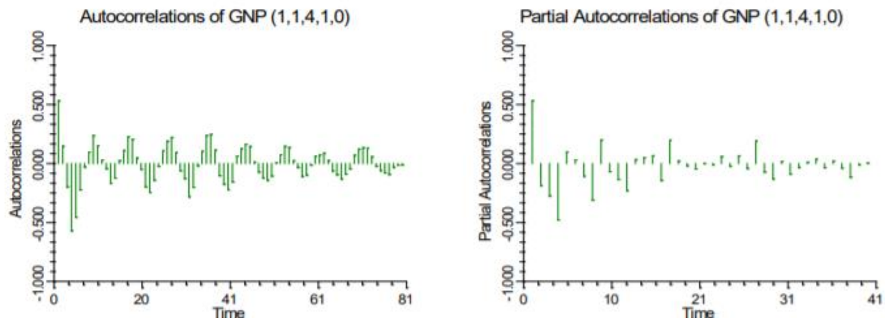

Figure 4.4 Sample Stationary Series Plot

Differencing usually significantly reduces the number of large autocorrelations. We'd have to difference the differenced series again if it didn't appear stationary.

The magnitude of a large autocorrelation and partial autocorrelation coefficient is frequently useful. To be statistically significant, an autocorrelation must be at least 2 / N in absolute value. The table below shows some frequent significant autocorrelation values for various sample sizes. Even though an autocorrelation is statistically significant, it may not be large enough to be a cause for concern.

Table 4.1 Sample series with large autocorrelation

N	Large Autocorrelation
25	0.40
50	0.28
75	0.23
100	0.23
200	0.14
500	0.09
1000	0.06

If the data series is steady in the first place, no differencing is required, and d=0.

In the ARIMA(p,d,q) model, the next step is to choose values for d, then p and q.

The partial autocorrelations of the correctly differenced series are used to calculate the value of p. The projected value of p would be the last lag with a big value if the partial autocorrelations broke off after a few lags. You have a moving average model (p=0) or an ARIMA model with positive p and q if the partial autocorrelations do not cut off.

The autocorrelations of the correctly differenced series are used to calculate the value of q. If the autocorrelations stop after a few lags, the predicted value of q would be the last lag with a big value. You either have an autoregressive model (i=0) or an ARIMA model with positive p and q if the autocorrelations do not cut off.

The diagnostic examination of a model is the final stage after it has been fitted. The checking is done by looking at the autocorrelation plots of the residuals to determine whether there is any additional structure (high correlation values). The model is deemed acceptable, and forecasts are created if all autocorrelations and partial autocorrelations are modest.

The values of p and/or q are modified, and the model is re-estimated, if some of the autocorrelations are high.

This procedure of reviewing the residuals and modifying the p and q values continues until the residuals have no more structure. The application can be used to generate forecasts and related probability limitations once an appropriate model has been chosen.

4.2 Quick Design

At this stage the initial prototype is developed, where the very basic requirements are showcased, and user interfaces are provided. This stage would provide a high-level view of the application to the client.

4.2.1 Context Diagram

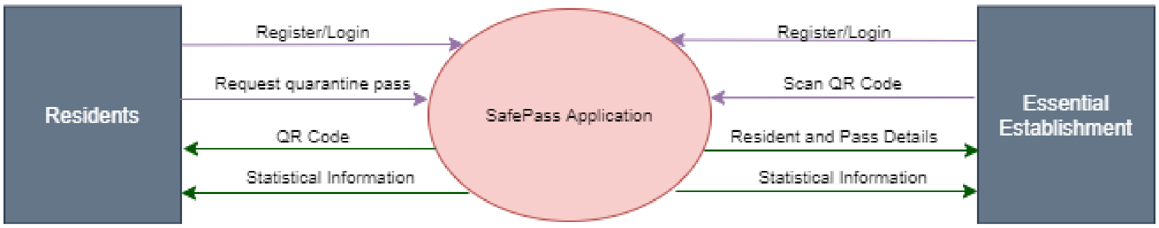

Figure 4.5 Context Diagram

The context diagram in Figure 4.5 illustrates the authorized users input to the system and the expected output information. The target users are the residents and essential establishment personnel. The expected output of the proposed application is: Digital Quarantine Pass (QR Code), Quarantine Holder Information and Quarantine Pass Status, Crowd Count per Establishment and Crowd Forecast per Establishment.

4.2.2 Data Flow Diagram

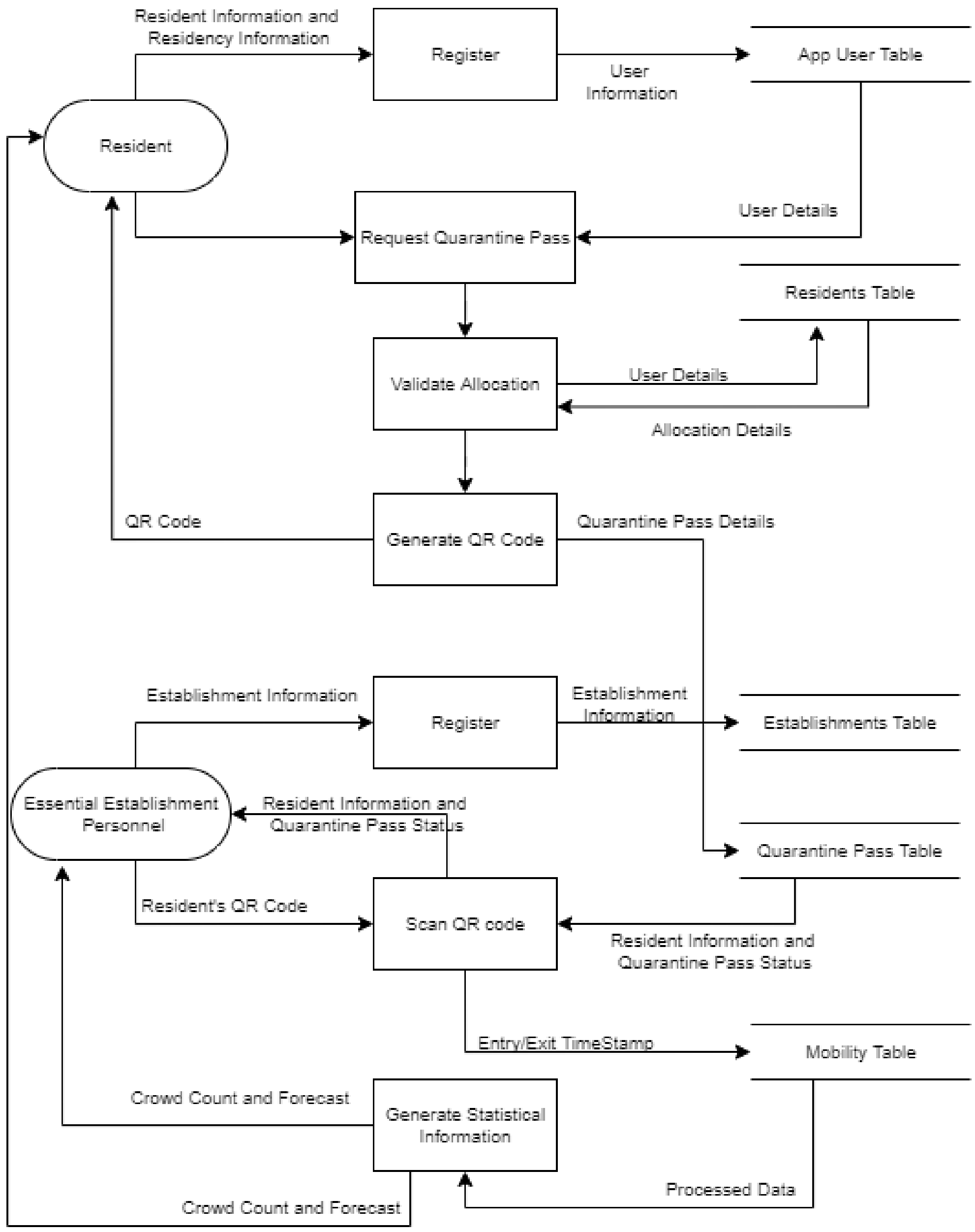

Figure 4.6 Data Flow Diagram

The researcher used the Data Flow Diagram, which is a dramatic representation of the information flow within a system that shows how information enters the system and leaves the system, what changes the information and where it is stored (Kendall, 2005).

4.2.3 Use Case Diagram

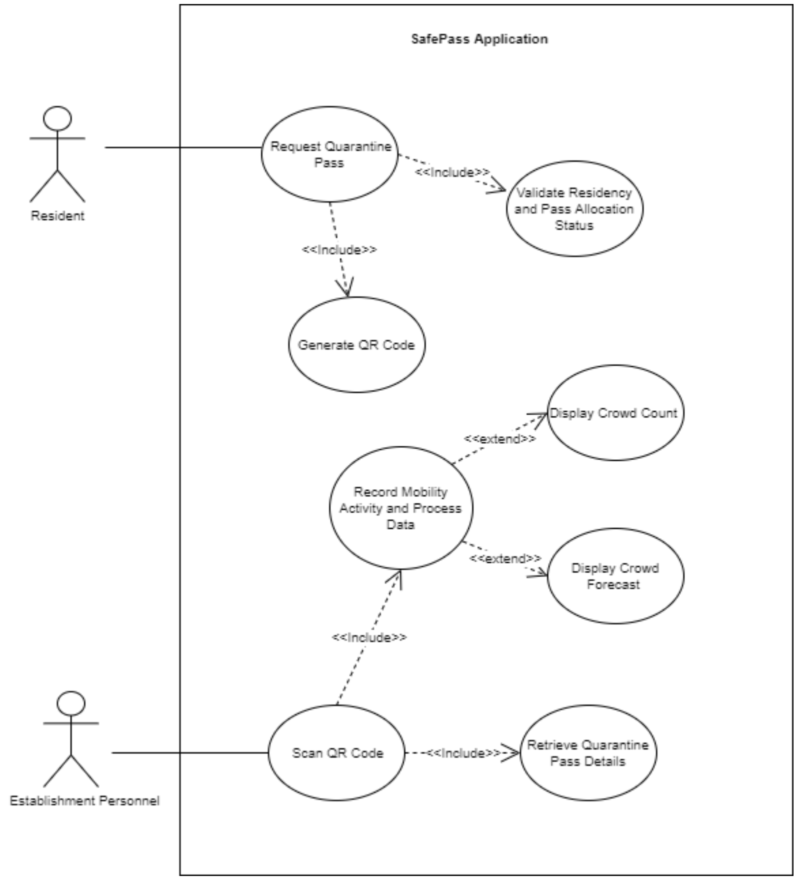

Figure 4.7 Use Case Diagram

The development of the proposed application does not solely depend on how the current system is manually implemented. It also contains a workflow process that was identified by the proponent as a necessary step to improve the system flow and accomplish the target results. The components of the proposed application “SafePass: An Implementation of Autoregressive Integrated Moving Average (ARIMA) for Crowd Forecasting Applied in Quarantine Pass”, is illustrated in Figure 4.7 with the use of Use Case Diagram. It describes its user, processes, and the relations between the system components that give the overall behavior of the application.

4.2.4 System Flowcharts

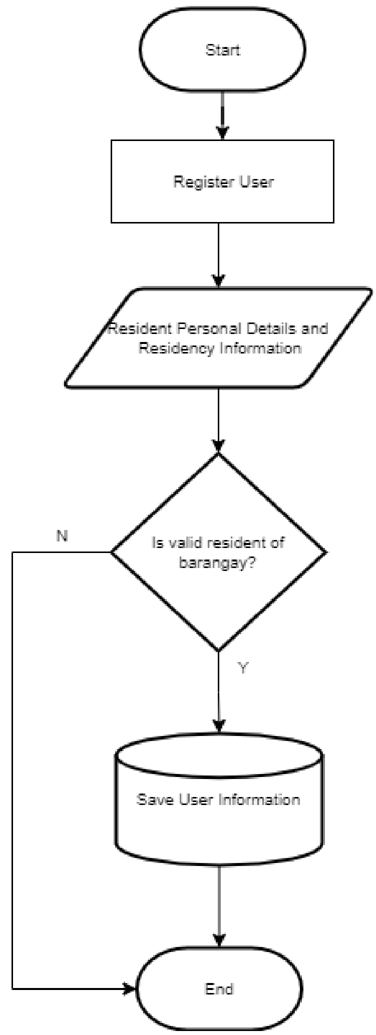

Figure 4.8 Resident System Registration Flow

Figure 4.8 shows the flow of resident registration wherein they need to provide their personal information, residency information and attach a valid ID for verification purposes. Personal information includes First Name, Middle Name, Last Name, Date of Birth and Contact Number. Residency information such as Barangay Name and Street/Subdivision/Purok/Zone are needed to validate if the user belongs to the correct

Barangay base on the resident's database in the system. Lastly, the user must provide an identification card and upload an image for further system validation.

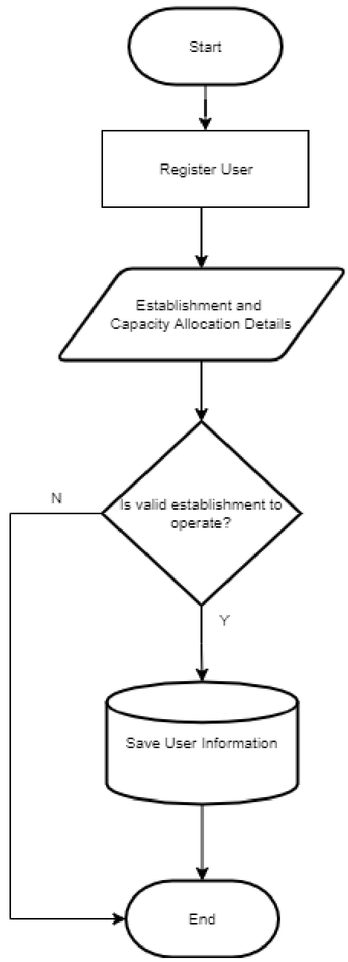

Figure 4.9 Essential Establishment System Registration Flow

The illustration on Figure 4.9 defines the process flow when an essential establishment registers in the application. Information such as Registered Business Name, Registered Business Owner Name, Registered Business Address/Barangay, number of people allowed in the premises shall be provided together with Business Registration Certificate and will be validated against the essential establishment's database which contains a list of permitted establishments to operate during Enhanced Community Quarantine.

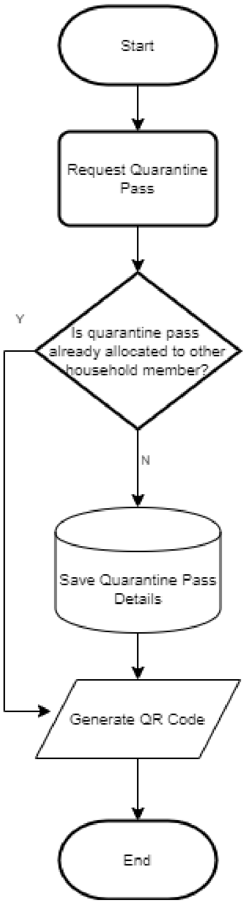

Figure 4.10 Generate Quarantine Pass Flow

The process flow for the generation of quarantine pass is illustrated in Figure 4.10 wherein the system will check on what household the user belongs to and validate if the quarantine pass allotted for their household was already granted. If no other member of the same household has not requested for a quarantine pass, then the system will generate a QR Code and update the database that a quarantine pass was already granted. The user will have the ability to download the QR Code or view it on the system every time they login.

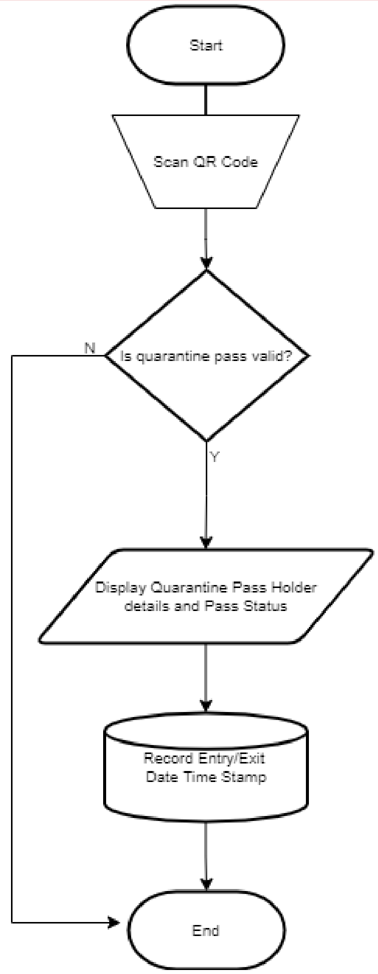

Figure 4.11 Scan Quarantine Pass Flow

Figure 4.11 displays the system flow when a QR Code is scanned in an establishment. Resident with valid QR Code should present the QR code to the essential establishment personnel using SafePass application upon entry/exit. The system will then validate the QR code validity and should display the quarantine pass holder's Full Name, Barangay, and the uploaded Image. This action will record the entry/exit date and time stamp of the holder and the establishment name in the system database.

Figure 4.12 Generate Statistical Information Flow

In the process flow presented in Figure 4.12, a user will have the ability to view the current crowd count in an essential establishment registered in SafePass application. The crowd count information will provide details if a certain establishment is already reaching their allowed capacity which will help the residents decide whether to proceed or not. By using the historical data in the system, a weeklong crowd forecast can be generated which will aid the residents in planning when is the best time and day of the week should they go to the establishment to avoid crowd exposure. The same statistical information can also be utilized by the establishment in risk management and planning.

4.3 Building Prototype

At this stage, system requirements and other components necessary to develop the proposed application will be identified.

4.4 User Evaluation

This is the stage where the application users would evaluate the application based on its required features. The capstone project will utilize a checklist type of survey questionnaire in which the respondents will be able to answer faster and easier at their convenience.

4.5 Refining Prototype

In this stage, any dissatisfaction with the prototype at this level will result to a revision based on the given requirements. The new prototype will be re-evaluated, and the process continued until such time that the requirements identified by the end-user were met. Revisions will be done based on the user's comments and suggestions during the evaluation of the developed application.

4.6 Engineer Product

The last stage of this approach will conclude with the confirmation and approval of the application by the end-users. This will also be referred to as the user acceptance phase. It is also in this phase that the proponent will be able to appraise the overall performance of the final system, using the predetermined indices or indicators such as functionality, efficiency, reliability, usability, and portability.

LIST OF REFERENCES

Cai, J., Sun, W., Huang, J., Gamber, M., Wu, J., & He, G. (2020). Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020. Emerging infectious diseases, 26(6), 1343–1345. https://doi.org/10.3201/eid2606.200412

Badr. H., Du, H., Marshall, M., Dong,E., Squire, M., Gardner, L. (2020). Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. The Lancet Infectious Diseases,Volume 20, Issue 11, 2,7. https://doi.org/10.1016/S1473-3099(20)30553-3.

COVID-19 Dashboard, Republic of the Philippines, accessed 7 August 2021, <https://www.covid19.gov.ph/security/ecq>

Department of the Interior and Local Government Bicol Region, Republic of the Philippines, accessed 7 August 2021, <http://region5.dilg.gov.ph/advisory-distribution-of-barangay-quarantine-id-pass/>

Inter-Agency Task Force for the Management of Emerging Infectious Diseases (IATF – EID), Republic of the Philippines, accessed 7 August 2021, <https://iatf.doh.gov.ph/?page_id=680>

Inquirer.net, Inquirer, accessed 7 August 2021, <https://newsinfo.inquirer.net/1247189/lanao-del-sur-village-chief-nabbed-for-selling-quarantine-pass-for-p20>

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet, Volume 395, Issue 10223, 1. https://doi.org/10.1016/S0140-6736(20)30183-5

Perlman, S. Another Decade, Another Coronavirus. N Engl J Med 2020. DOI: 10.1056/NEJMe2001126

Wu, Y. C., Chen, C. S., & Chan, Y. J. (2020). The outbreak of COVID-19: An overview. Journal of the Chinese Medical Association: JCMA, 83(3), 217–220. https://doi.org/10.1097/JCMA.0000000000000270

Morawska L, Milton DK. It is time to address airborne transmission of covid-19. Clin Infect Dis 2020;. doi: 10.1093/cid/ciaa939

World Health Organisation. Scientific brief. Transmission of SARS-CoV-2: implications for infection prevention precautions. 2020. <https://www.who.int/news-room/commentaries/detail/transmissionof-sars-cov-2-implications-for-infection-prevention-precautions?>

Asadi, S., Wexler, A.S., Cappa, C.D. et al. Aerosol emission and superemission during human speech increase with voice loudness. Sci Rep 9, 2348 (2019). https://doi.org/10.1038/s41598-019-38808-z

Han ZY, Weng WG, Huang QY. 2013 Characterizations of particle size distribution of the droplets exhaled by sneeze. J R Soc Interface 10: 20130560. http://dx.doi.org/10.1098/rsif.2013.0560

Bootsma MCJ, Ferguson NM. The effect of public health measures on the 1918 influenza pandemic in U.S. cities. Proc Natl Acad Sci U S A 2007;104(18):7588–93. https://doi.org/10.1073/pnas.0611071104

Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. Comment How will countrybased mitigation measures influence the course of the COVID-19 epidemic? 2020;2019(20):1– 4.

Imai, N., Gaythorpe, K., Abbott, S., Bhatia, S., van Elsland, S., Prem, K., Liu, Y., & Ferguson, N. M. (2020). Adoption and impact of non-pharmaceutical interventions for COVID-19. Wellcome open research, 5, 59. https://doi.org/10.12688/wellcomeopenres.15808.1

Wu, Z., McGoogan, JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. JAMA. 2020;323(13):1239–1242. doi:10.1001/jama.2020.2648

Zhang, Y., Jiang, B., Yuan, J., et al.: The impact of social distancing and epicenter lockdown on the covid-19 epidemic in mainland china: A data-driven seiqr model study. medRxiv. 2020. 10.1101/2020.03.04.20031187

Lai, S., Ruktanonchai, N.W., Zhou, L., et al.: Effect of non-pharmaceutical interventions for containing the covid-19 outbreak: An observational and modelling study. medRxiv. 2020. 10.1101/2020.03.03.20029843

Tkachenko, W. Puech, C. Destruel, O. Strauss, J. Gaudin and C. Guichard, "Two-Level QR Code for Private Message Sharing and Document Authentication," in IEEE Transactions on Information Forensics and Security, vol. 11, no. 3, pp. 571-583, March 2016, doi: 10.1109/TIFS.2015.2506546.

Singh, S., “QR Code Analysis,” International Journal of Advanced Research in Computer Science and Software Engineering, vol. 6, no. 5, 2016.

Uzun, V., "QR-Code Based Hospital Systems for Healthcare in Turkey," 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC), 2016, pp. 71-76, doi: 10.1109/COMPSAC.2016.173.

Cho, J.; Seo, G.-W.; Lee, J.S.; Cho, H.K.; Kang, E.M.; Kim, J.; Chun, D.-I.; Yi, Y.; Won, S.H. The Usefulness of the QR Code in Orthotic Applications after Orthopedic Surgery. Healthcare 2021, 9, 298. https://doi.org/10.3390/healthcare9030298

Nakamoto, I., Wang, S., Guo, Y., Zhuang, W. A QR Code–Based Contact Tracing Framework for Sustainable Containment of COVID-19: Evaluation of an Approach to Assist the Return to Normal Activity. JMIR Mhealth Uhealth 2020;8(9):e22321. https://mhealth.jmir.org/2020/9/e22321. DOI: 10.2196/22321

Hur, J., & Chang, M. C. (2020). Usefulness of an Online Preliminary Questionnaire under the COVID-19 Pandemic. Journal of Medical Systems, 44(7).

Palit, A.K.; Popovic, D. (2005). Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications

Chatfield, Chris (2000). Time-series Forecasting. CRC Press, 2000

Krollner, Bjoern; Vanstone, Bruce; Finnie, Gavin; (2010). Financial time series forecasting with machine learning techniques: A survey

Taieb, Souhaib Ben; Bontempi, Gianluca; Atiya, Amir F; Sorjamaa, Antti (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition.

Cerqueira, Vitor; Torgo, Luis; Soares, Carlos (2019). Machine Learning vs Statistical Methods for Time Series Forecasting: Size Matters Weigend, Andreas .S. and Gershenfeld, Neil A. (1994) Time Series Prediction: forecasting the future and understanding the past

Sorjamaa, Antti and Lendasse, Amaury (2006). Time series prediction using dirrec strategy

Li, M. L. et al. Forecasting COVID-19 and analyzing the effect of government interventions. (2020) v Preprint at medRxiv

Adiga, A., Dubhashi, D., Lewis, B. et al. Mathematical Models for COVID-19 Pandemic: A Comparative Analysis. J Indian Inst Sci 100, 793–807 (2020). https://doi.org/10.1007/s41745-020-00200-6

Tseng, Y., Shih, Y. (2020) Developing epidemic forecasting models to assist disease surveillance for influenza with electronic health records, International Journal of Computers and Applications, 42:6, 616-621, DOI: 10.1080/1206212X.2019.1633762

Imai, C., Armstrong, B., Chalabi, Z., Mangtani, P., Hashizume, M., Time series regression model for infectious disease and weather. Environmental Research, Volume 142, 2015. https://doi.org/10.1016/j.envres.2015.06.040.

PAEDER, M., STROUD, J., & SONG, X. (2012). Time-series model to predict impact of H1N1 influenza on a children's hospital. Epidemiology and Infection, 140(5), 798-802. doi:10.1017/S0950268811001749

Box, George E.P.; Jenkins, Gwilym; Reinsel, Gregory C. (1994). Time Series Analysis Forecasting and Control

Siami-Namini, Sima; Tavakoli, Neda; Siami Namin, Akbar (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series

Zhang, Y., Yang, H., Cui, H. et al. Comparison of the Ability of ARIMA, WNN and SVM Models for Drought Forecasting in the Sanjiang Plain, China. Nat Resour Res 29, 1447–1464 (2020). https://doi.org/10.1007/s11053-019-09512-6

Tandon, H., Ranjan, P., Chakraborty, T., Suhag, V., Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future,2020. arXiv2004.07859

Hyndman, R.J., & Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on Aug 7, 2020, https://otexts.com/fpp2/arima.html

PANKRATZ, A. (1983) Forecasting with univariate Box–Jenkins models: concepts and cases, New York: John Wiley & Sons. Accessed on Aug 7, 2020, https://robjhyndman.com/papers/BoxJenkins.pdf

PhilAtlas, accessed 7 August 2021. <https://www.philatlas.com/luzon/r04a/cavite/bacoor/molino-iv.html>

Philippine Statistics Authority, Republic of the Philippines, accessed 7 August 2021. <https://psa.gov.ph/content/highlights-household-population-number-households-and-average-household-size-philippines>

APPENDIX A: