FORECASTING FUTURE FERTILITY RATES IN MALAYSIA THROUGH
REGRESSION MODEL

WAN NUR SYAZANA BINTI WAN RUSLAN

UNIVERSITI TEKNOLOGI MALAYSIA

UTM
UNIVERSITI TEKNOLOGI MALAYSIA

UNIVERSITI TEKNOLOGI MALAYSIA
DECLARATION OF Choose an item.

Author's full name	:	WAN NUR SYAZANA BINTI WAN RUSLAN
Student's Matric No.	:	MCS241016 Academic : Session
Date of Birth	:	UTM Email :
Choose an item. Title	:	FORECASTING FUTURE FERTILITY RATES IN MALAYSIA THROUGH REGRESSION MODEL

I declare that this thesis is classified as:

☒ **OPEN ACCESS**   I agree that my report to be published as a hard copy or made available through online open access.

☐ **RESTRICTED**   Contains restricted information as specified by the organization/institution where research was done. *(The library will block access for up to three (3) years)*

☐ **CONFIDENTIAL**   Contains confidential information as specified in the Official Secret Act 1972)

(If none of the options are selected, the first option will be chosen by default)

I acknowledged the intellectual property in the Choose an item. belongs to Universiti Teknologi Malaysia, and I agree to allow this to be placed in the library under the following terms :
1. This is the property of Universiti Teknologi Malaysia
2. The Library of Universiti Teknologi Malaysia has the right to make copies for the purpose of only.
3. The Library of Universiti Teknologi Malaysia is allowed to make copies of this Choose an item. for academic exchange.

Signature of Student:
Signature :

Full Name
Date :

Approved by Supervisor(s)

Signature of Supervisor I: Signature of Supervisor II

Full Name of Supervisor I   Full Name of Supervisor II
NOOR HAZARINA HASHIM   MOHD ZULI JAAFAR

Date :   Date :

NOTES : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction

This letter should be written by a supervisor and addressed to Perpustakaan UTM. A copy of this letter should be attached to the thesis.

Date:

Librarian

Jabatan Perpustakaan UTM,

Universiti Teknologi Malaysia,

Johor Bahru, Johor

Sir,

CLASSIFICATION OF THESIS AS RESTRICTED/CONFIDENTIAL

TITLE: FORECASTING FUTURE FERTILITY RATES IN MALAYSIA THROUGH REGRESSION MODEL

AUTHOR'S FULL NAME: WAN NUR SYAZANA BINTI WAN RUSLAN

Please be informed that the above-mentioned thesis titled _____ should be classified as RESTRICTED/CONFIDENTIAL for a period of three (3) years from the date of this letter. The reasons for this classification are

(i)

(ii)

(iii)

Thank you.

Yours sincerely,

SIGNATURE:

NAME:

ADDRESS OF SUPERVISOR:

"I hereby declare that I have read this proposal and in my opinion this proposal is sufficient in term of scope and quality for the award of the degree of Master in Data Science"

Signature   :   _____

Name of Supervisor I   :   ASSOCIATE PROF. DR. MOHD SHAHIZAN BIN OTHMAN

Date   :   17 JANUARY 2025

Signature   :   _____

Name of Supervisor II :

Date   :

Signature   :   _____

Name of Supervisor III :

Date   :

Declaration of Cooperation

This is to confirm that this research has been conducted through a collaboration WAN NUR SYAZANA BINTI WAN RUSLAN and ASSOCIATE PROF. DR. MOHD SHAHIZAN BIN OTHMAN

Certified by:

Signature   :

Name   :

Position   :

Official Stamp

Date

* This section is to be filled up for theses with industrial collaboration

Pengesahan Peperiksaan

Tesis ini telah diperiksa dan diakui oleh:

Nama dan Alamat Pemeriksa Luar   **:**

Nama dan Alamat Pemeriksa Dalam   **:**

Nama Penyelia Lain (jika ada)   **:**

Disahkan oleh Timbalan Pendaftar di Fakulti:

Tandatangan   :

Nama   :

Tarikh   :

FORECASTING FUTURE FERTILITY TRENDS IN MALAYSIA THROUGH
REGRESSION MODEL

WAN NUR SYAZANA BINTI WAN RUSLAN

A proposal submitted in fulfilment of the
requirements for the award of the degree of
Master in Data Science

School of Computing
Faculty of Computing
Universiti Teknologi Malaysia

JANUARY 2025

DECLARATION

I declare that this proposal entitled *"Forecasting Future Fertility Trends in Malaysia Through Regression Model"* is the result of my own research except as cited in the references. The proposal has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature   : ....

Name   :

Date   : 17 JANUARY 2025

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious, the Most Merciful. Peace be upon the Prophet Muhammad SAW.

I would like to express my utmost gratitude and deepest appreciation to which has given guidance and consultation in coordinating this research project. A clear explanation did help me a lot along the journey especially in writing this paper.

I would also like to acknowledge with much gratefulness to my beloved friends who provided their best moral support, comfort and inspiration to complete this challenging journey. Without them, I will not be able to complete this report.

Many thanks to my coursemates in giving me the much-needed information and always offering their cooperation and full support. May Allah bless those who were involved in this journey with me.

Finally, I am most thankful and lucky enough to receive constant encouragement from my beloved family to continue my education. All of my prayers and love goes to them.

ABSTRACT

The problem of an aging population is increasing worldwide, and Malaysia is also dealing with this challenge. Previous research has acknowledged that the decreasing fertility rate is a key factor contributing to this issue. Nonetheless, in addition to the decreasing fertility rate in Malaysia over time, there are also alterations in social and economic factors, including a rise in female labor participation, greater household income and consumption spending, escalating inflation, and shifts in the population's demographics. The research has raised worries as the elements might influence the decreasing fertility rate. Hence, this study is conducted (1) to analyse the current population trends in Malaysia from year 1970 to 2023; (2) To identify the direction of the causal relationship declining fertility rate (TFR) to economic performance of a country; (3) to forecast future fertility trends in Malaysia through regression model approach. The first objective of this research will be achieved through Exploratory Data Analysis (EDA). Granger Causality test will be conducted to achieve the second research objectives. However, Autoregressive Distributed Lad (ARDL) model will be used to achieve the third objectives which to predict the future trends of the fertility rates in Malaysia.

ABSTRAK

Masalah penduduk yang semakin tua semakin meningkat di seluruh dunia, dan Malaysia juga menghadapi cabaran ini. Penyelidikan terdahulu telah mengakui bahawa penurunan kadar kesuburan adalah faktor utama yang menyumbang kepada isu ini. Namun begitu, sebagai tambahan kepada penurunan kadar kesuburan di Malaysia dari semasa ke semasa, terdapat juga perubahan dalam faktor sosial dan ekonomi, termasuk peningkatan dalam penyertaan buruh wanita, pendapatan isi rumah yang lebih besar dan perbelanjaan penggunaan, peningkatan inflasi, dan peralihan dalam demografi penduduk. Penyelidikan telah menimbulkan kebimbangan kerana unsur-unsur mungkin mempengaruhi penurunan kadar kesuburan. Oleh itu, kajian ini dijalankan (1) untuk menganalisis trend populasi semasa di Malaysia dari tahun 1970 hingga 2023; (2) Untuk mengenal pasti arah hubungan sebab akibat penurunan kadar kesuburan (TFR) kepada prestasi ekonomi sesebuah negara; (3) untuk meramalkan trend kesuburan masa depan di Malaysia melalui pendekatan model regresi. Objektif pertama penyelidikan ini akan dicapai melalui Analisis Data Penerokaan (EDA). Ujian Granger Causality akan dijalankan untuk mencapai objektif kajian kedua. Walau bagaimanapun, model Autoregressive Distributed Lad (ARDL) akan digunakan untuk mencapai objektif ketiga untuk meramalkan trend masa depan kadar kesuburan di Malaysia.

TABLE OF CONTENTS

LIST OF TABLES

x

LIST OF FIGURES

LIST OF ABBREVIATIONS

xii

ARDL   -   Autoregressive Distributed Lag

FTR   -   Total Fertility Rate

GDP   -   Gross Domestic Product

CHAPTER 1

INTRODUCTION

1.1   Introduction

Ageing population is not something strange to any country in the world. It is part of a process in which every country will experience it. However, in Malaysia recently, this issue has been highlighted and becoming a concern among citizen's as well as the government's representatives. As Rafizi Ramli, a Minister of Economics of Malaysia said where Malaysia is expected to reach the status of an old country in 2044. More worryingly, other countries that go through this process is in a stronger economic and household financial position than Malaysia.

"Child Free" slogans is not ashamed to resound within social media platform of a Malaysian's newlywed couples recently. They express a statement that they are not wished to have an offsprings in their married life by giving their reasons that it is sort of economic burdensome to raise a child from infant stage until their young adult's life. This sentiment slowly arises and create anxiety to young adults to raise a child if the household income become the considerations. This case will contribute to the declination of birth rates among Malaysian population if this mindset were continuing to happen.

To enhance a rapid economy growth of a country, a large population were needed where bigger pool of labor, can increase productivity and drive industrial economic activities. Due to drops of birth rates in the long run, professionals need to increase their work-life spans due to shortage of local labor. To conclude, this issue needs to be emphasized, and a study needs to be done to address the declination of fertility rates that happen year by year to facilitates the government to take various measures and actions in order to sustain the future economic growth.

1.2   Problem Background

In recent development of advanced digital era, young generation especially Generation Z who born in the year 1997 to 2012 seems faced difficulties in developing their social skills in the real world as they have spent most of their time by surfing the internet such as mindlessly scrolling social media, addiction to online games and staying indoors resulting they were less exposed in real-world problems making them socially incompetent and developed social anxiety within themselves. They are unable to hold long, deep, meaningful and eye-to-eye conversations anymore.

This situation affects their relationship with other people, and they will miss many opportunities in their life if they keep maintained that kind of attitude and habits. Or even worse, they will face an obstacle in finding a life partner due to lack of communication skills or not exposing themselves to positive communities or gatherings. This is a serious problem when young adults do not even get married even after they have graduated from college or university due to the quarter life crisis. If this situation continues to happen in the long run, our world will experience a seriously ageing population in 2040.

1.3   Problem Statement

According to UNDP Malaysia Accelerator Lab authored by Rachel Yue, Malaysia is reaching a turning point in its demographic transition towards an increasingly ageing population. This is supported by the Department of Statistics Malaysia (DOSM), the proportion of the population aged 65 and older is expected to increase significantly, from 8.1% in 2024 to 14.5% by 2040. This proves that Malaysia will experience ageing nation with low productivity in workforce resulting to slow economic growth and might turn down the import and export activities in our country. Special measures and intervention from the authorities need to be taken as soon as possible to combat this issue for future long-term benefit of a country.

1.4   Aim of Research

The aim of this research is to study the drops of fertility rates in Malaysia and how it influences the economy of the country.

1.5   Research Objectives

The research conducted aims to achieve the following objectives:

(a)   To analyses the current population trends in Malaysia from year 1970 to 2023.

(b)   To determine the direction of the causal relationship declining fertility rate (TFR) to economic performance of a country.

(c)   To forecast future fertility trends in Malaysia through regression model approach.

1.6   Scope of Research

To conduct an impactful research analysis, time is an important element that needs to be considered by a researcher. Due to time constrain, the researcher aimed to perform several capable methods within a certain period of time such as systematic literature review to achieve the first objective of this research which is to identify the underlying factors that contribute to the drops of birth rates in Malaysia.

As this topic is not something 'rare' within the population of the world, a few research papers can be obtained. This benefits the researcher to not spend too much time to make the analysis in order to achieve the research objectives. The dataset obtained from the website of Malaysia's Open Data Portal (data.gov.my) with collaboration of Department of Statistics Malaysia (DOSM) is enough to represent the current population of Malaysia. The website provides easy-discoverable information and the relatable dataset categories can be attained such as demographic, annual live births, household income, labour rate market, national gross domestic product (GDP), education and healthcare.

CHAPTER 2

LITERATURE REVIEW

2.1   Introduction

This chapter discusses on the global overview on birth rates declination and factors that contributes to the declination. Besides, the content also consists of the impacts of fertility rate changes specifically on economic perspectives and past studies on the relationship between aging population and economic growth. In this literature review, it also highlights the forecasting model used to predict the future trends of birth rates of a population to help the researcher determine which model is the most and are capable to use to produce accurately and in short period of time.

2.2   Overview of Global Population Trends

Worldwide population undergoes birth rates declination since 1950 and continues to happen. According to (Bhattacharjee et al., 2024) in comprehensive demographic analysis report titled the Global Burden of Disease Study 2021 discovers that rate of fertility in all countries and territories declined since 1950 with Total Fertility Rate (TFR) remaining above 2.1 as figure 2.1 below. The figure 2.1 represents TFR, worldwide and by GBD super-region, 1950–2100. The dotted horizontal line represents replacement TFR (2.1), while the dotted vertical line marks the year 2022 (the initial forecast year).
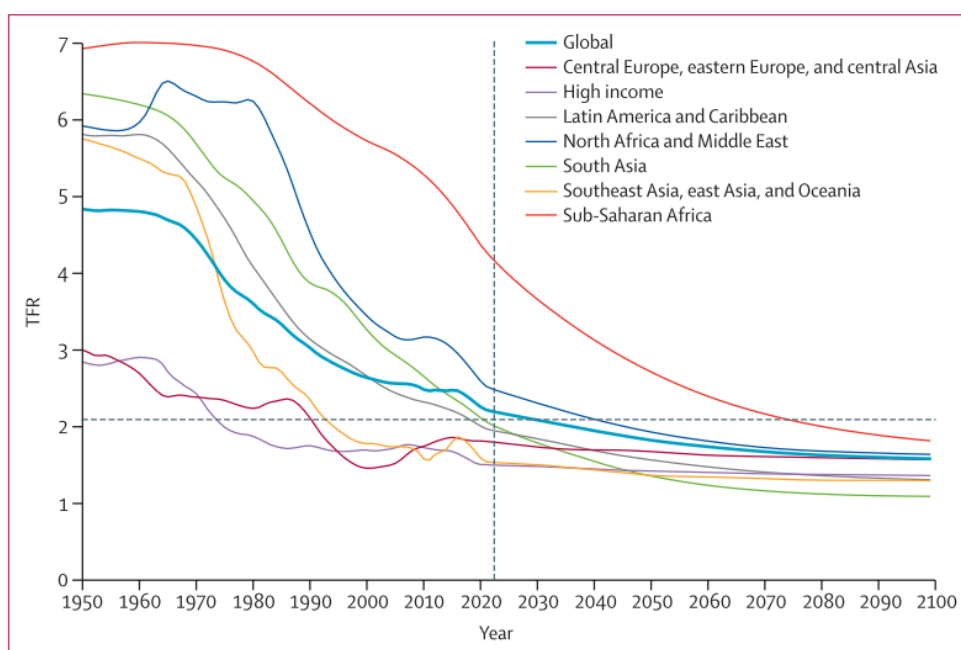
Figure 2.1 TFR, globally and by GBD super-region, 1950 - 2100

In a report that discuss of fertility rates issues cover OECD countries also highlights (OECD, 2009) that those countries such as Australia, Canada, Korea and Japan also experiencing a decline in the total fertility rate (TFR) for a longer time. In details, the decline was temporary stalled during 2000s but resumed its downward trend after 2007-08 Great Financial Crisis. The TFR of children per women fell to 1.5 in 2022 where average across the OECD where it is below "replacement level" of 2.1 children per woman.

More worryingly, it is reported that the TFR of some countries falls below the replacement level where it is the indicator for the minimum rate that is necessary for generational replacement of the population (Bhattacharjee et al., 2024). Across the OECD, there is a general tendency toward a rise in childlessness, however the degree of this trend differs. The incidence of ongoing childlessness has at least increased twofold in Estonia, Italy, Japan, Lithuania, Poland, Portugal, and Spain, impacting almost one in four women in the 1975 cohort in Italy and Spain, based on a comparison of the cohort fertility of women born in 1935 and 1975. In Japan, it stands at 28%. (OECD, 2009).

2.2.1 Declining Rates

Fertility is declining globally, with rates in more than half of all countries and territories in 2021 below replacement level. This trend appears to be common in many countries, especially in Western nations, and is influenced by a variety of economic, cultural and social aspects. Since 1950, fertility rates have drastically decreased worldwide, and they will continue to do so in practically every nation and territory until the year 2100 (Bhattacharjee et al., 2024)

American families have changed in recent generations due to a variety of circumstances, including shifting social expectations. Birth rates fell into the early 1970s after the 1946–1964 baby boom, and then barely changed in the decades that followed. Fertility gradually increased in the years preceding 2008 before plummeting, and the Great Recession was another turning point. Since then, it has largely kept declining. According to the most recent data available, which covers 2020 and includes children born before the epidemic, 43 states saw their lowest general fertility rates in at least 30 years (State Fiscal Health, 2022). Figure 2.2 indicates the fertility rates have dropped considerably in recent decades on average across the OECD.
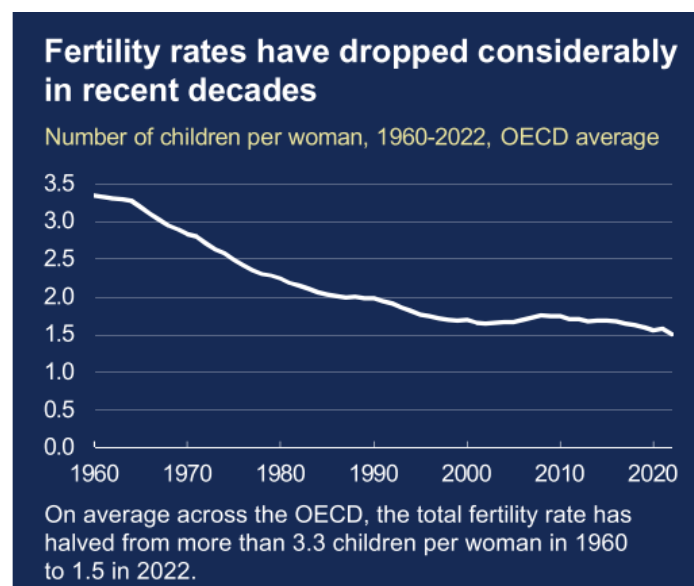

Figure 2.2 Number of children per woman, 1960 – 2022 (OECD, 2009)

However, it is reported that the future fertility rates were projected to continue to decline worldwide where it is reaching a global TFR of 1.83 (1.59 – 2.08) in 2050 and 1.59 (1.25-1.96) in 2100 (Bhattacharjee et al., 2024).

Past literature review also revealed that Malaysia has been confirmed Malaysia is moving towards ageing population. According to (Mohd, Ishak, & Selvaratnam, 2021) it is predicted around 7% (2.3 million people) of the population consists of age categories of 65 years old and above in year 2020. It is considered as aging population if the 7% of the population is 65 years and above. This statements can be proven by (Mohd et al., 2021) in their study revealed the increase trends of the age dependency ratio for older people. Conversely, the fertility rate in Malaysia indicates a decreasing trend as per Table 2.1 below.

Table 2.1 The age dependency ratio, fertility rate and government health expenditure from 2000 to 2018 (Mohd et al., 2021)

Year	Age dependency ratio	Fertility rate	Government health expenditure (% of GDP)
2000	6.241	2.78	1.17
2002	6.527	2.54	1.37
2004	6.695	2.36	1.46
2006	6.967	2.25	1.67
2008	7.238	2.19	1.60
2010	7.359	2.15	1.68
2012	7.812	2.11	1.86
2014	8.373	2.07	2.03
2016	8.980	2.04	1.89
2018	9.623	2.00	1.92

Source: World Bank (3–5).

2.3 Factor Influencing Fertility Declination

The drop in fertility rates involves many factors that contribute to it. This section will explain the factors that contribute to fertility rates decline in terms of socioeconomic development, access to education and healthcare, and cultural and social norms.
.

2.3.1 Socioeconomic Development

When discussing the factors of socioeconomic development, it correlates with the urbanization of the population. Logically, urbanization growth often leads to smaller family sizes due to limited space, higher living costs and different social norms. This statement, supported by (Jaffe, 1942) a number of well-known demographers conducted analyses on the topic and their conclusions all pointed to urbanization as a major contributing cause to the decline in fertility in developed nations.

Apart from that, economics conditions also influence the sustained low fertility rates of a population where the increasing costs of child nurturing and uncertainty in the economy inhibit the growth of bigger families. This evidence supported by (Muadz Bin Zulqarnain & Md Yusuf, 2022) tells that in past researchers conducted a quasi-experimental data that income and direct costs of children are significant towards fertility behaviors.

Moreover, another author highlights that some workers nowadays limit their family size since they need to commit to their work and need to send their children to child daycare while at work thus resulting to high cost to raise a child (Muadz Bin Zulqarnain & Md Yusuf, 2022). All of this evidence implies that the more the urbanization of a population, the high cost needed to raise a child and smaller the family sizes.

2.3.2 Education

This section reveals that in term of education landscape especially women that holds higher education and climb up the career ladder contributes to the declination rates for Malaysia. Few authors have concluded these statements several times such as research done by (Tang and Tey, 2017) in Malaysia found that higher education, career development of women, led to a decline in the fertility rate for Malaysia.

(Ketaki Chandiok, Prakash RanjanMondal, Chakraverti Mahajan, 2016) investigated the biological and social determinants of fertility in India. The study highlighted that women's self-determination, often reflected in their educational attainment and occupational status, significantly influences family size. Women who are self-determined often marry later and are more mindful about family planning. For working women, pregnancy can limit opportunities for personal development, as it involves extended maternity leave and attention to their health and their baby's well-being. Consequently, women's employment patterns may experience interruptions due to childbearing and family responsibilities (Yusuf, 2012).

Women are now participating in the labour force in significantly greater numbers and are staying active throughout their childbearing and child-rearing years. They are no longer seen as a secondary or reserve workforce. Historically, especially in developed countries, a "double peak" trend was common—women would join the labor force in their twenties, leave temporarily to have and raise children, and return later after completing their childbearing responsibilities (Mujahid & Uz Zafar, 2012).

2.4 Impacts of Fertility Rates Changes

Changes in fertility rates can influence population growth, age structure, and workforce composition, thereby affecting economic development, resource allocation, and social policies. High fertility rates often strain resources and infrastructure in rapidly growing populations, while declining fertility rates can lead to aging populations and labor shortages. This section will emphasize the impact of this changes especially in term of economic impacts.

2.4.1 Economic Impacts

When the size of the populations of a country shrinks over time, indirectly will give impact to its economy and unbearable to sustain. The implications of fertility declines include life expectancy rising where the more individuals over 65 years than young generations called aging population. The shortages of workers in a country will likely happen due to decreasing number of young generations to contribute to growth of economy. Lastly, the government needs to

2.4.1.1 Aging Population

Overall world population were expected to experience ageing nation due to sustained low fertility rates over the past few decades supported by (Mohd et al., 2021) claims that low fertility rates and increases in longevity can result in an increase in the aged population. Through researchers' findings there were several definitions regarding ageing population or nations and surprisingly there are differences between both "ageing society", "ageing nation" and "super-aged society".

Ageing society defined as where the population of a country consists of minimum 7 percent of the population who aged 65 and older according to (Sharayu Pillai, 2023). While the ageing nation has 14 percent or more in similar age groups. Moreover, the population who are aged 65 and older has more than 20 percent in the population called super-aged society.

Table 2.2 The definition of Aging Society, Ageing Nation and Super-aged society
(Sharayu Pillai, 2023)

Definitions		
Ageing Society	Ageing Nation	Super-aged society
"Population has at least **7 percent** of the population who are aged 65 and older."	"Population has **at least 14 percent** of the population who are aged 65 and older."	"Population has **at least 20 percent** of the population who are aged 65 and older."

Ageing population however will cause many consequences in country toward its population itself in terms of economic, health and social perspectives as claimed by (Bhattacharjee et al., 2024) where the authors stated that a sustained low fertility rate will result in ageing of the population will lead to challenges in the economics of the country and consequently creates pressure on healthcare systems, social security programmes as well as the labour workforce.

2.4.1.2 Workforce Shortages and Low Productivity

Following to the population of ageing nation slightly give an implication to reduction of labor force participants as higher proportion of older individuals with fewer children are being born leading to fewer young people entering the labor market in the future as portrays by (Mohd et al., 2021) the number of laborers may be impacted by an aging population since population growth is slowing down. Therefore, this situation could have many consequences to the job market and the economic growth of the country itself.

Since the number of labor participants has been reduced, the income of the country will decrease resulting to slow down the economy of the population as mentioned by (Otsu & Shibayama, 2016) the declination of the size of the workforce will decrease subside the income of the country as the number of workforce could participate to the industrial market is shrinking. Thus, the national income of a population can be challenging to sustain over the longer term, leading to various consequences if a country fails to implement effective measures to foster economic growth.

Consequences with the higher proportion of aged workers will decline in the productivity of the workers. The effect of this situation results in lower productivity outcome that can contribute to the decline growth of economy. (Börsch-Supan et al., 2021) who discovered that older people would produce less productive labor outputs after a certain age, can support this. The overall rate of technological advancement in society will decline due to the internal structure of the workforce is skewed toward "aging"(Mohd et al., 2021).

2.4.1.3 Economic Growth tied to Demographic Dividends

As we can review from the consequences of declination of fertility rates of a population where there is likely fewer new young generations and higher proportion of ageing people will lead to ageing society. The number of labor workforce participation will start to shrink, and the output of the workforce's productivity tends to be slower thus leading to a decrease in economic growth. This is proven by research done by (Maestas, Mullen, & Powell, 2016) in United States, the ageing population will lead to decrease in economic growth.

Above statement is further supported by another author (Mohd et al., 2021) shows that A 16.8% rise in the population aged 60 and above resulted in a 9.2% decline in GDP from 1980 to 2010. (Modigliani and Brumberg, 1954) via their life cycle theory, endorses the view that an ageing demographic can hinder economic progress. It is concluded by (Mohd et al., 2021) Malaysia might not achieve to become one of the developed countries of 2030 results of this aged population.

Due to slow down of economic growth, it is hard for a country to plan for their national savings due to the early stage of aging population. In a study done by (Modigliani and Brumberg, 1954), the life cycle hypothesis indicates that in the initial phase, the aging population will lead to an increase in national savings. Nevertheless, as the population keeps aging and certain fractions of the population attain their retirement age, this theory suggests a decline in overall savings resulting from a rise in the aging demographic.

In a theory by Solow's growth theory indicates that stable economic growth of country seems difficult to sustain in economy where the population is aging. This supported by (Gruescu. S, 2007) the ageing of a population has a negative impact on economic growth based on this theory. Only when the population's age distribution stays consistent can steady state growth or stable growth circumstances be achieved. However, the age distribution of an economy with an aging population is not stable. As a result, this is only feasible while the economy is moving toward stability.

Past studies by (Maestas et al., 2016) also discovered that the decreasing fertility rate and increasing aging indicators like the old age dependency ratio and the population aged 65+) are likely to hinder economic growth. (Teixeira, Renuga Nagarajan, & Silva, 2017) discover that the increasing age of the population adversely affects the expansion of developed nations, unlike that of less developed or developing nations.

However, it is essential to sustain the population size to promotes long-term economic development to have more national income growth as the income can be use by the government to provide more public goods and facilities to Malaysians (Mohd et al., 2021). This raises questions by (T. & James, 1992) to the recent argument regarding the impacts of an ageing population. He states that humanity is fated to exist in perpetual poverty since the increase in agricultural output will consistently lag the rise in population.

Another researcher gives their opinion on the reduction in size of populations impacts the evolution of long term economy through endogenous growth models such as (Romer, 1990) states the size of the population is important for the long term development of the economy through endogenous growth. Same goes to (Prettner, 2013) integrates endogenous growth models from (Romer, 1990) with semi endogenous growth models and discovered that an increase in longevity positively affects per capita economic growth. The decline in fertility rates creates adverse effects on economic growth.

Moreover, the beneficial impact of longevity prevails over the adverse effect of fertility. Population aging encourages sustained growth within the endogenous growth model. (Prettner, 2013) concludes that ongoing demographic shifts do not inherently obstruct technological advancement and are a factor in economic prosperity. He also contends that falling birth and death rates may result in a rise in the rate of economic growth.

2.4.1.4 Healthcare and Pension Systems

Conversely, general government expenditure on health is rising slightly (Mohd et al., 2021). In 2018, Japan allocated 10.9% of its GDP to health expenditure, reflecting its high proportion of elderly citizens (Mohd et al., 2021). Selvaratnam et al. (2021) note that Malaysia's increasing life expectancy has led to higher government spending on pension costs, the Employee Provident Fund (EPF), and health services for the elderly. This trend creates a financial burden on individuals, families, and policymakers, involving significant expenses related to housing and healthcare (Mohd et al., 2021).

Louria (2022) suggests that life expectancy increases could push average lifespans to 100-120 years, leading to various social challenges. These include rising health expenditures for those aged 65 and over, quality of life concerns, potential strains on social security and pension systems, and other related issues. Baharin and Saad (2023), using the autoregressive distributed lag (ARDL) method and vector error correction model, confirm that the growing elderly population has a significant impact on Malaysia's health expenditure.

Similarly, Ismail et al. (2015) argue that reducing the old dependency ratio could lessen the tax and social security burdens on employees, potentially easing the financial strain associated with retirement income and elderly healthcare. Despite, India faces similar challenges, with an aging population potentially reducing saving rates and increasing unemployment risks.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

In this chapter 3 will elaborate in detail the research method, design and approach used to attain and analyse all the data to accomplish the research objectives and answer the research questions. This research will follow the data science project lifecycle from the beginning until the end. The aspects that will be discussed and included in this chapter are research framework, problem formulation, data collection, data pre-processing and data analysis.

3.2 Research Framework

Research is an activity that involves observation, collecting more data and information while research methodology can show or describe how data is collected, analysed and data presented. As being mentioned, this research will follow the data science project life cycle phases which includes business problem understanding, data collection, data preparation, exploratory data analysis (EDA), data modelling, model evaluation and model deployment. All those steps must be followed and be done steps by steps to achieve meticulous results. Figure 3.1 below shows the typical data science project lifecycle process steps by steps.
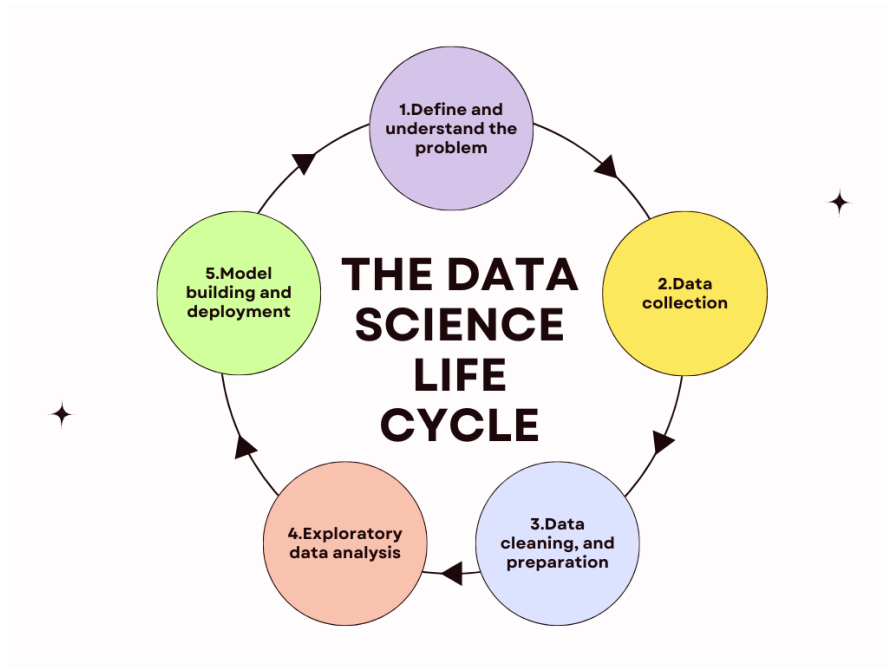
Figure 3.1 Data Science Project Life Cycles

A research framework is designed following the steps of typical data science life cycle starts from problem formulation, data collection method which involves on how the researcher find the sources of datasets that suitable for this research, data preparation, perform exploratory data analysis (EDA) acts as a preliminary analysis helps the researcher to get a comprehensive understanding of the dataset. This EDA process will achieve the first objective of this research. After exploratory data analysis (EDA) is performed, the suitable data modelling and deployment is identified to achieve the second and third of this research objectives which is to predict the future trends of fertility rates in Malaysia. The following figure 3.2 illustrates the process involved in the research frameworks.
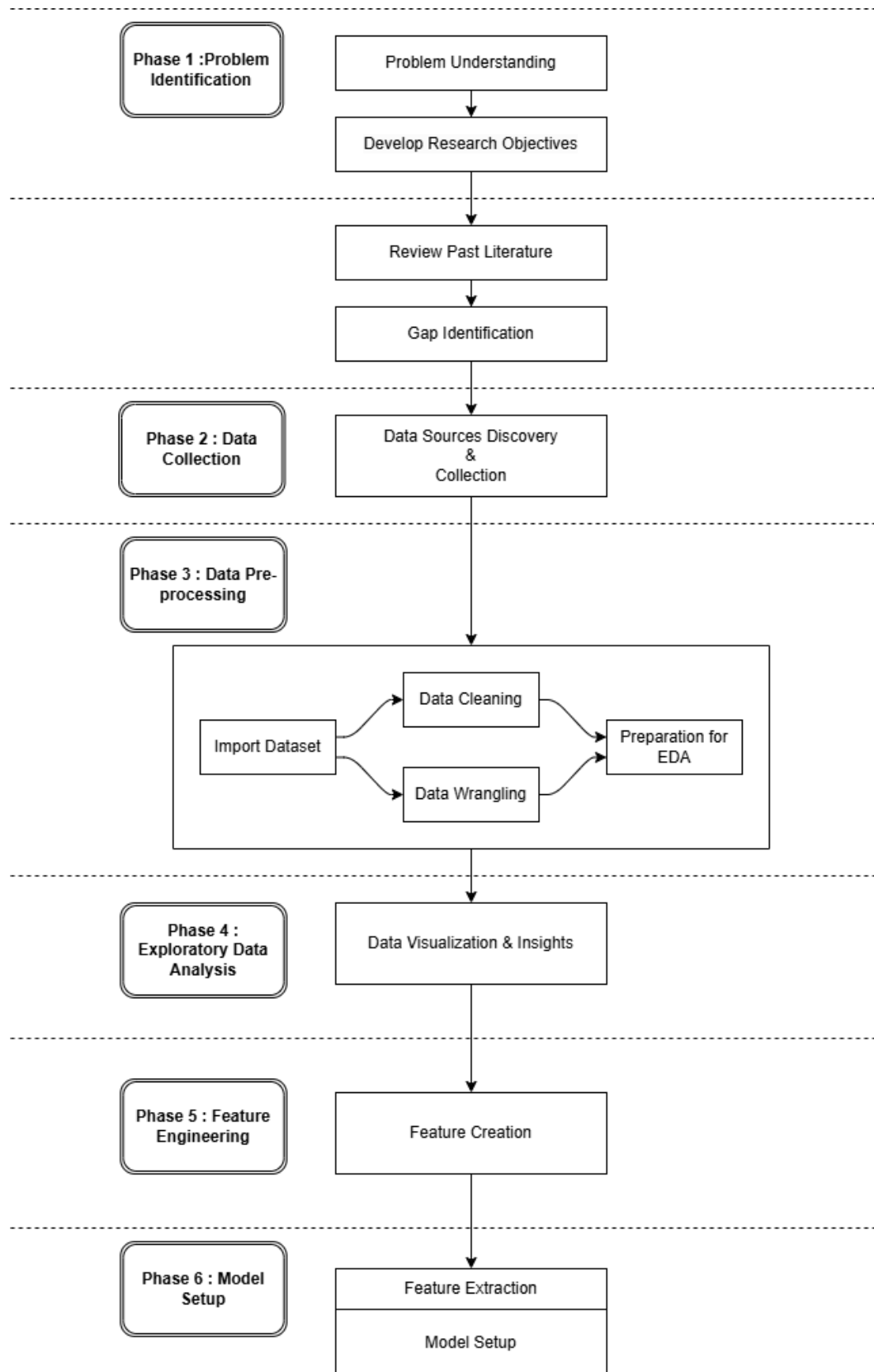
Figure 3.2 Research Framework

19

3.2.1 Problem Understanding

The initial of problem understanding and identification has been identified during literature review phase where the global trends of fertility has been declined, and Malaysian population has been labelled as ageing nation in 2022. So, the researchers are intended to find the root causes of declination trends and intended to predict the future trends in Malaysian populations to mitigate the prolonged effects cause by drop of fertility rates as well as to help the policy makers to refrain this problem from continues to happen.

3.2.2 Data Collection Method

Data collection is an activity of gathering data and information from all compatible sources to find an answer for the issue of research, test hypotheses, and analyze the results. Collecting data is an essential element for a research study. This section will explain how the data will be evaluated in this study. Basically, there are two types of approaches when collecting the data which are primary data and secondary data collection. The researcher decided to implement the primary data collection in this research to fulfill the research objectives.

a) Sources of Data

Table 3.1 Sources of Data

Research Objectives	Data Required	Sources of Data
i) To analyze the current population trends in Malaysia from year 1970 to 2023.	Historical data of Malaysian population from year 1970 to 2023.	• Department of Statistics Malaysia (DOSM) • Malaysia Official Open Data Portal (data.gov.my)
ii) To determine the direction of the causal relationship between declining fertility rate (TFR) to economic performance of a country.	• Gross Domestic Product (GDP) • Total Fertility Rate (TFR)	• World Bank Open Data • Department of Statistics Malaysia (DOSM)
iii) To forecast future fertility trends in Malaysia through regression model approach.	Historical data of Malaysian birth rates from year 1920 to 2023	• Malaysia Official Open Data Portal (data.gov.my)

The dataset obtained from the website of Malaysia's Open Data Portal (data.gov.my) with collaboration of Department of Statistics Malaysia is enough to represent the current population of Malaysia. The dataset that suitable for the analysis is births data from 1920 to 2023 and population by states from 1970 until 2023. These two datasets are considered suitable to forecast the future trends of birth rates in Malaysia.

3.2.3   Data Preparation

Data preparation phase considered as elementary phases where the dataset obtained need to go through the process of data cleaning, processing of raw data so that it is suitable for further processing and analysis. The researcher will import the dataset into Jupyter Notebook using Python to condition of the data and detect missing values. The researcher also intended to combine the two datasets to go through exploratory data analysis process (EDA) to easier the researcher to find the data information and conditions for further analysis.   The researcher intends to do the following data pre-processing as below:

Table 3.2 Data Wrangling and Its Purposes

Data Wrangling	Purposes
1) Convert "date" to a datetime format.	To ensure the "date" column is in the correct format for analysis.
2) Handle categorical variables	Convert categorical columns like "state", "age", "sex", and "ethnicity" into a format suitable for modelling.
3) Check for missing values	Handle missing values appropriately
4) Feature engineering	- Extract year or other useful features from the "date" column. - Group and aggregate data based on the analysis needed such as "total population per state"
5) Remove or transform outliers	Inspect the "population" column for outliers and handle them.
6) Normalize/scale numerical data	Scale the "population" values for use in machine learning models.

3.2.4 Preliminary Analysis

Preliminary analysis in data science lifecycles usually known as Exploratory Data Analysis (EDA). This phase needs to be performed to help the researcher to identify and understand the condition of the datasets and get the initial insights and findings through the datasets. The researcher intends to do the following preliminary analysis that is suitable.

Table 3.3 Data Analysis and Visualization

Analysis	Data Visualizations	Details
Population Distribution by Age Group	Bar Chart	Visualize the population across different age groups.
Gender-Based Population Analysis	Pie Chart	Display the proportion of each gender within a specific state or overall.
State-Wise Population Comparison	Bar Chart	Compare the total population across different states.
Time Series Analysis	Line Chart	Show how population changes over time for specific age groups, genders, or states.

3.2.5 Model Building & Deployment

To achieve the second and third objectives, two types of models have been identified by past literature review by (Mohd et al., 2021) which is using Autoregressive Distributed Lag (ARDL) and (Muadz Bin Zulqarnain & Md Yusuf, 2022) which is using Granger Causality test to identify the direction of causal relationship between socioeconomic factors and fertility rates.

Table 3.4 Types of Analysis and Model

Research Objectives		Types of Analysis	Methodology/ Modelling
i)	To analyze the current population trends in Malaysia from year 1970 to 2023.	Exploratory Data Analysis (EDA) • Descriptive Statistics & Analysis	
ii)	To identify the direction of the causal relationship between declining fertility rate (TFR) to economic performance of a country.	• Diagnostic Analysis	Granger Causality Test Model
iii)	To forecast future fertility trends in Malaysia through regression model approach.	• Predictive Analysis	Autoregressive Distributed Lag (ARDL) Model

CHAPTER 4

INITIAL RESULTS

4.1   Introduction

This chapter 4 will discuss regarding the presents the findings and analyzation of data obtained from the datasets. All the findings will be presented, tabulated, and summarized to obtain the results of the study. This chapter begins with the introduction of exploratory data analysis as well as data visualization and descriptive statistics. The contents also provide initial insights gained from the data visualization findings by the researcher. Lastly, the feature engineering that involves during the exploratory data analysis process will be emphasized in this chapter.

4.2   Exploratory Data Analysis (EDA)

In the data science life cycle, the preliminary analysis phase known as "Exploratory Data Analysis" should be made. The purpose of this phase that needs to be done is to help the researchers to gain insight into the conditions of the data itself. This phase involves gaining an understanding of the response and its influencing factors before building the actual model. The graphical examination of data distribution within unique variables of a character utilizes bar graphs, while correlations among various elements are illustrated through visual methods such as scatter plots and heat maps (GeeksforGeeks, 2023).

In this section will presents the initial findings that contain data analysis and suitable visualization to help the reader to understand and be able to interpret the findings of the data. The interpretation and insights will be elaborated after the visualization made to gain clear understanding of the correlation between the variables. Furthermore, the feature engineering will be present as the researcher made some feature engineering involved during data cleaning process.

4.2.1 Data Visualization & Descriptive Statistics

This section will present several initial analyses that can be done with dataset collected. The findings includes the distribution of population by age group, gender based population analysis and state-wise population comparison, as the dataset contains the following variables such as "population", "age" , "state", "ethnicity" and "date".
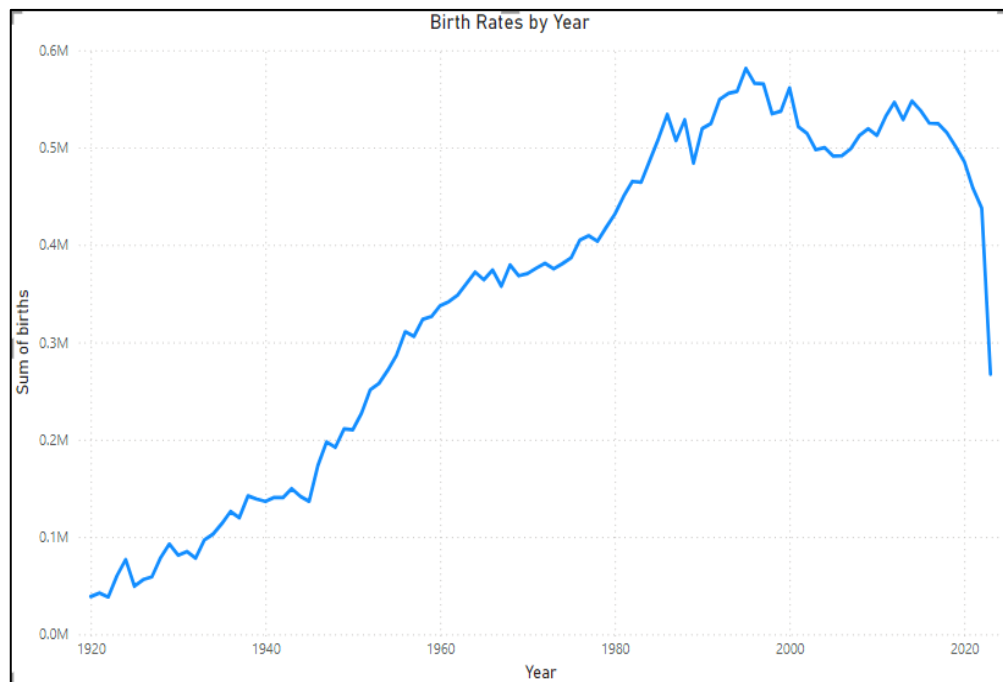

Figure 4.1 shows the line chart of birth trends from 1920 to 2023.

Figure 4.1 shows the line chart illustrates the trend in birth rates over time, spanning from 1920 to 2020. Birth rates exhibit a consistent upward trend from 1920 to the 1960s. Birth rates reached a peak around the late 20th century (1980s-2000s), suggesting a period of demographic expansion and stability. From the mid-2000s onward, there is a noticeable decline in birth rates. The sharp drop near 2020 may be linked to extraordinary circumstances, such as the COVID-19 pandemic, which disrupted healthcare systems, economic stability, and family planning decisions.
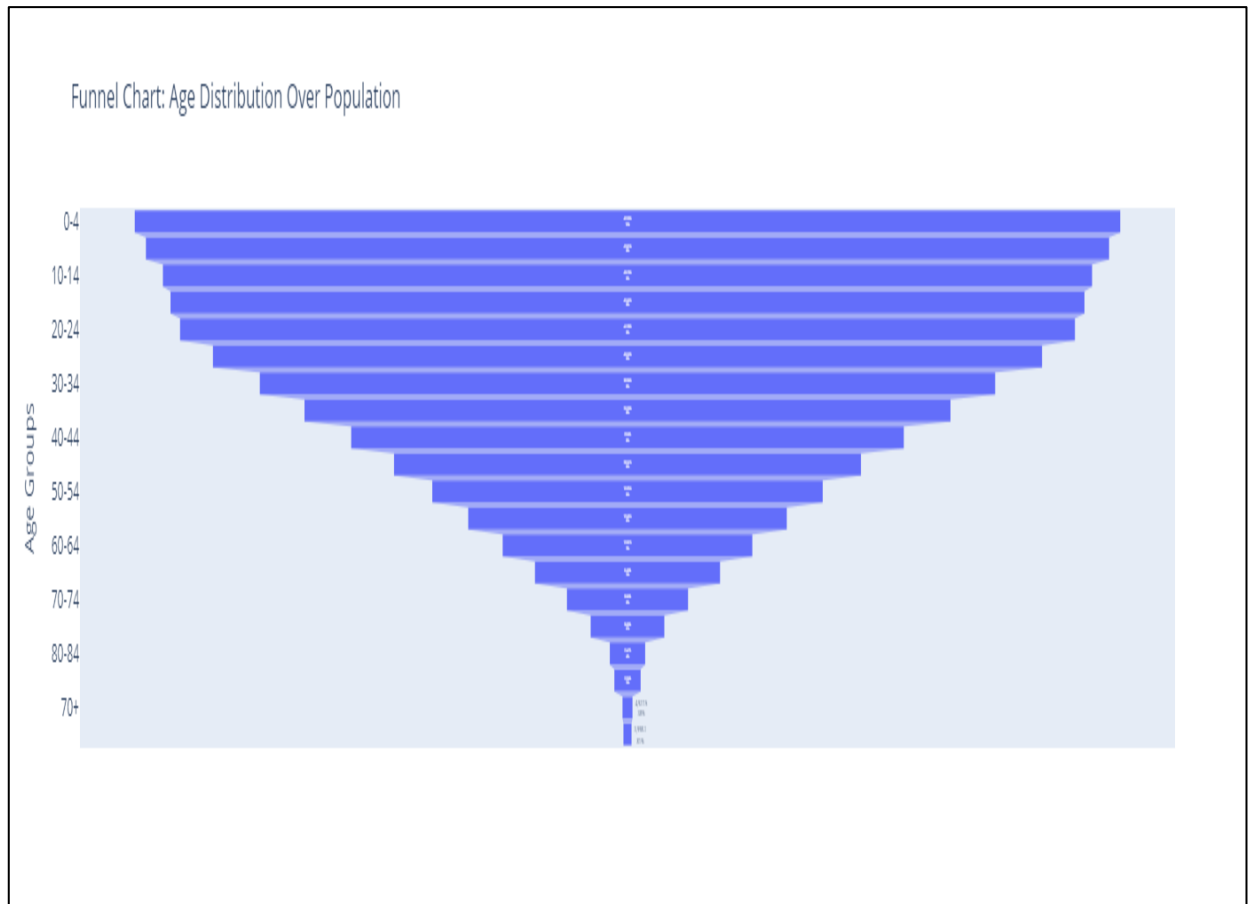
Figure 4.2 show funnel bar chart of age distribution of Malaysian population.

Figure 4.1 shows the funnel bar chart indicates the age distribution of Malaysian population. The bar chart is divided into age groups showing a breakdown of the population for each group. The funnel shape indicates a gradually declining population from younger to older groups. This is consistent with typical age distributions in many societies where younger generations are more numerous due to higher birth rates.
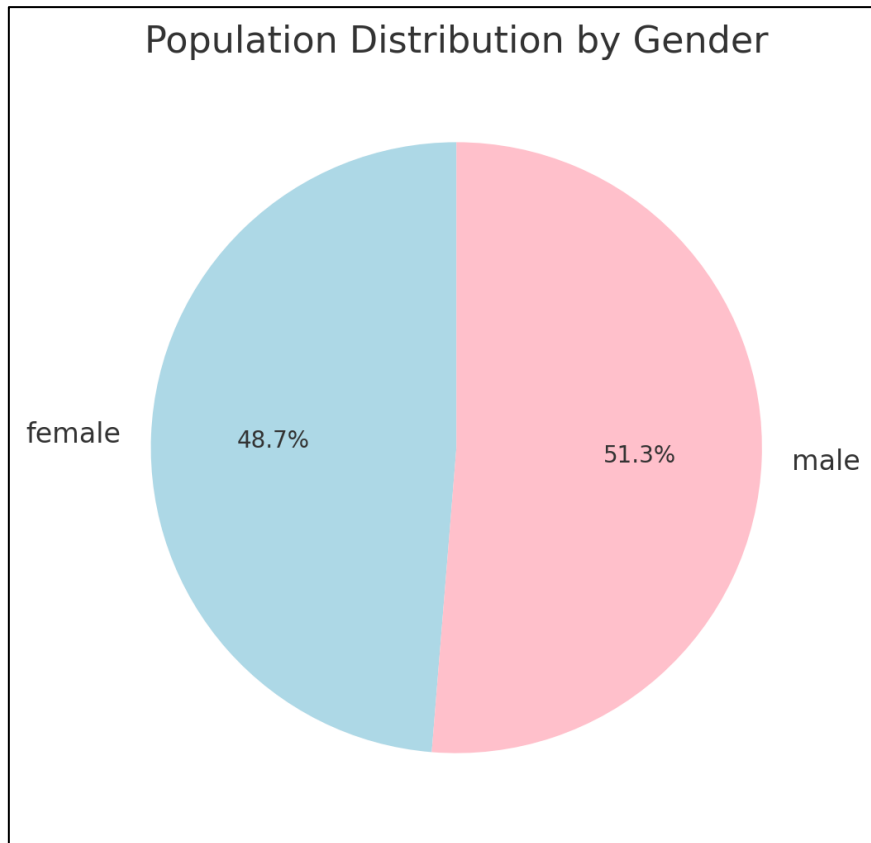
Figure 4.3 shows a pie chart of population distribution by gender.

Figure 4.3 shows the bar chart that illustrates the percentage of the total population by state. The chart shows that 51.3% of the population is male, slightly more than half of the total population. Females make up 48.7% of the population, a slightly smaller proportion compared to males. The distribution suggests a near balance between the genders, with males being marginally more prevalent. Such small differences may reflect natural variations or cultural and social factors affecting the gender ratio. This level of gender parity is common in many populations and may suggest an equitable distribution in terms of natural birth rates and survival rates across genders.
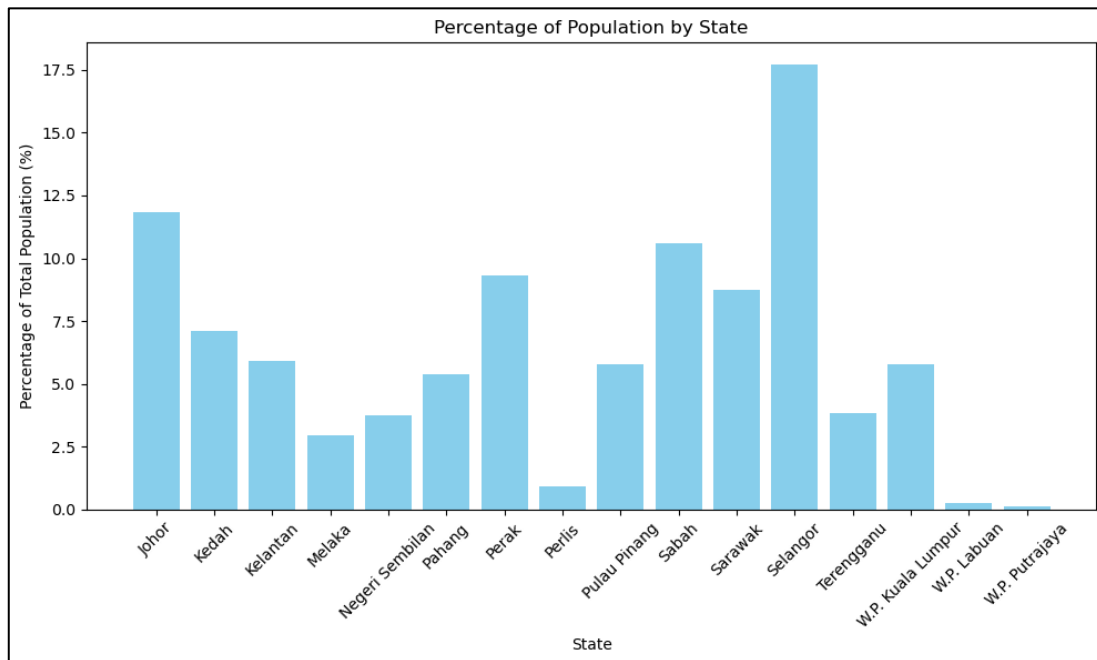
Figure 4.4 shows bar chart of the percentage of the total population by state in Malaysia.

Figure 4.4 shows the bar chart that illustrates the percentage of the total population by state. Selangor has the highest percentage of the population, exceeding 17%, making it the most populous state. This could be due to urbanization, industrial development, and better job opportunities. Johor has the second-highest population percentage, slightly above 12%. This reflects its strategic location and economic importance. States like Perlis**,** Wilayah Persekutuan Putrajaya, Labuan**,** and Wilayah Persekutuan Putrajaya have the smallest population percentages, likely due to their smaller geographical sizes or administrative roles.

CHAPTER 5

CONCLUSION AND FUTURE WORKS

5.1   Conclusion and Future Works

World population are moving towards aging population as the fertility rates slowly decline. Malaysia is not exceptional and expected to reach the status of an old country in 2044 claimed by Minister of Economics of Malaysia, Rafizi Ramli in a conference. Therefore, this study aims to analyse the current population of Malaysians citizen from year 1970 to 2023. In Chapter 4, during the exploratory data analysis (EDA) phases, a funnel bar chart has illustrated the age distribution of Malaysian population. The results show that the widest sections at the top suggest that younger age groups have a higher population. This may indicate a relatively high birth rate or a young population demographic. Conversely, fom the mid-2000s onward, there is a noticeable decline in birth rates.

For second objectives, the researcher intendeds to identify the direction of the causal relationship declining fertility rate (TFR) to economic performance of a country. Suitable independent and dependant variables need to be identified to deploy into the intended model which is Granger Causality test modelling. Before that, several tests need to be done to test whether the variables is stationary state.

Third objectives of this research are to achieve the aim of this research which is to predict the future birth rates. This aim is significance as to prepare the government as well as stakeholders to plan several measures to mitigates the issue of declination of economic growth in the future and factors that contribute to it such as shortages number of labor participants, healthcare spending and pension systems, national savings and so on as per stated in the literature review. Autoregressive distributed lag (ARDL) model will be used to predict the future trends of fertility rates in Malaysia based on past historical data as this model is somewhat capable to be done .

REFERENCES

Bhattacharjee, N. V., Schumacher, A. E., Aali, A., Abate, Y. H., Abbasgholizadeh, R., Abbasian, M., … Vollset, S. E. (2024). Global fertility in 204 countries and territories, 1950–2021, with forecasts to 2100: a comprehensive demographic analysis for the Global Burden of Disease Study 2021. *The Lancet*, *403*(10440), 2057–2099. https://doi.org/10.1016/S0140-6736(24)00550-6

Dr. Sandra Gruescu. (2007). Population Ageing and Economic Growth Contributions to Economics. *Contributions to Economics*.

GeeksforGeeks. (2023). Data Science Lifecycle - GeeksforGeeks.

Jaffe, A. J. (1942). Urbanization and Fertility. In *American Journal of Sociology* (Vol. 48). https://doi.org/10.1086/219078

Ketaki Chandiok, Prakash RanjanMondal, Chakraverti Mahajan, and K. N. S. (2016). *Journal of Anthropology - 2016 - Chandiok - Biological and Social Determinants of Fertility Behaviour among the Jat Women.pdf*.

Maestas, N., Mullen, K. J., & Powell, D. (2016). THE EFFECT OF POPULATION AGING ON ECONOMIC GROWTH, THE LABOR FORCE AND PRODUCTIVITY Nicole Maestas Kathleen J. Mullen David Powell Working Paper 22452 http://www.nber.org/papers/w22452. *NBER Working Papers*, 89. Retrieved from http://www.nber.org/papers/w16019

Mohd, S. N. A., Ishak, A. A., & Selvaratnam, D. P. (2021). Ageing Population's Impact on Economic Growth in Malaysia From 1981 to 2019: Evidence From an Autoregressive Distributed Lag Approach. *Frontiers in Public Health*, *9*(November), 1–9. https://doi.org/10.3389/fpubh.2021.731554

Muadz Bin Zulqarnain, A., & Md Yusuf, M. (2022). the Relationship Between Socioeconomic Factors With Fertility Rate in Malaysia. *Malaysian Journal of Computing*, *7*(2), 1222–1235. https://doi.org/10.24191/mjoc.v7i2.18841

Mujahid, N., & Uz Zafar, N. (2012). Economic growth-female labour force participation nexus: An empirical evidence for Pakistan. *Pakistan Development Review*, *4*(51), 565–585.

OECD. (2009). Society at a Glance 2009. In *Society at a Glance 2009*. https://doi.org/10.1787/9789264077904-ko

Otsu, K., & Shibayama, K. (2016). Population aging and potential growth in Asia. *Asian Development Review*, *33*(2Special Issue), 56–73. https://doi.org/10.1162/ADEV_a_00072

Prettner, K. (2013). Population aging and endogenous economic growth. *Journal of Population Economics*, *26*(2), 811–834. https://doi.org/10.1007/s00148-012-0441-9

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, *98*(5), S71–S102. Retrieved from http://www.jstor.org/stable/2937632

Sharayu Pillai. (2023). Malaysia Unprepared For Aged Nation Status: MOH Official. Retrieved from https://codeblue.galencentre.org/2023/10/malaysia-unprepared-for-aged-nation-status-moh-official/#:~:text=An ageing society is defined,an ageing nation last year.

State Fiscal Health. (2022). The Long-Term Decline in Fertility—and What It Means for State Budgets.

T., C., & James, P. (1992). T. R. Malthus. An Essay on the Principle of Population; Or a View of Its Past and Present Effects on Human Happiness; With an Inquiry into Our Prospects Respecting the Future Removal or Mitigation of the Evils Which It Occasions. *Population (French Edition)*, *47*(1), 245. https://doi.org/10.2307/1533650

Teixeira, A. A. C., Renuga Nagarajan, N., & Silva, S. T. (2017). The Impact of Ageing and the Speed of Ageing on the Economic Growth of Least Developed, Emerging and Developed Countries, 1990–2013. *Review of Development Economics*, *21*(3), 909–934. https://doi.org/10.1111/rode.12294