

---

## KULLIYYAH OF INFORMATION & COMMUNICATION TECHNOLOGY

---

**INFO 4314 BUSINESS DATA ANALYTICS**

**SEMESTER 2, 2024/2025**

**SECTION 2**

**A REPORT OF ANALYSIS ON NATIONAL MORTALITY FOR SDG 3: GOOD  
HEALTH AND WELL-BEING**

**PREPARED BY: DND**

NAME	MATRIC NO.
Muhammad Hafiz Faruqi bin Md Saifuddin	2217241
Ahmad Adlan bin Abdul Halim	2215775

---

**SUPERVISOR: ASST. PROF. DR. ABDUL RAFIEZ BIN ABDUL RAZIFF**

---

# 1.0 INTRODUCTION

## 1.1 Background

The Sustainable Development Goals (SDGs) are a universal blueprint adopted by the United Nations to promote prosperity while protecting the planet. Among these, SDG 3: Good Health and Well-Being aims to “ensure healthy lives and promote well-being for all at all ages.” A critical aspect of achieving this goal is reducing mortality rates and addressing disparities in health outcomes across different populations and regions. Therefore, this project focuses on developing a data-driven analytics application using Microsoft Excel to analyse mortality data by state, sex, and ethnicity. The purpose is to uncover potential disparities in death counts and identify trends that may indicate underlying health inequalities.

## 1.2 Objective

1. To analyse the number of deaths recorded across different states in Malaysia.
2. To explore how mortality varies based on sex and ethnicity.
3. To identify regions or demographic groups that may experience higher mortality, and by extension, potential health disparities.

## 1.3 Data Overview

The dataset used in this project contains 7,854 records and includes the following columns:

- **state** – The Malaysian state where the deaths were recorded.
- **date** – The date associated with the recorded deaths.
- **sex** – male, female, or both.
- **ethnicity** – Ethnic classification of the deceased.
- **abs** – The absolute number of deaths.

This dataset was obtained from OpenDOSM ([open.dosm.gov.my](http://open.dosm.gov.my)), which is the official government open data portal. It has been preprocessed to ensure it is clean, consistent, and suitable for analysis

## 2.0 DATA PREPROCESSING

### 2.1 Initial Inspection

The raw dataset consists of **7,854 records** and includes **five key columns**:

- **state (string)**: Represents the Malaysian state.
- **date (Date)**: The date of the record, initially in date format.
- **sex (string)**: Categorical variable representing male, female, or both.
- **ethnicity (string)**: Categorical variable denoting ethnic classification.
- **abs (numerical)**: Number of deaths recorded (absolute count).

This inspection confirmed the overall structure and identified key preprocessing steps.

### 2.2 Handling Missing Values

A total of 64 missing values were identified in the abs (death count) column, accounting for less than 1% of the dataset. These were treated as follows:

- Missing values were replaced with the mean death count (~1785).
- This method preserves the dataset's overall distribution while maintaining data consistency.

### 2.3 Consistency Checks

- Values in the sex and ethnicity columns were reviewed for formatting inconsistencies (e.g., different spellings or case sensitivity).
- All categorical values were standardised to lowercase and validated for correctness.
- No incorrect or malformed entries were found after this cleanup.

## 3.0 DESCRIPTIVE ANALYSIS

A foundational descriptive analysis was conducted to summarise and understand the distribution of deaths across different demographics and regions in Malaysia. This stage provided quantitative insights that support later trend analysis and policy recommendations.

### 3.1 Analysis of Overall Summary

Quartiles				
25%	50%	75%	OVERALL SUMMARY	
Mean	Median	Standard Deviation	Minimum	Maximum
1785	242	3181.458319	0	42051

Statistic (Dataset!E:E)	Value
Mean =AVERAGE(Dataset!E:E)	1,785
Median =MEDIAN(Dataset!E:E)	242
Standard Deviation =STDEV.P(Dataset!E:E)	3181.458319
Minimum =MIN(Dataset!E:E)	0
Maximum =MAX(Dataset!E:E)	42,051
25th Percentile =QUARTILE.INC(Dataset!E:E, 1)	34
75th Percentile =QUARTILE.INC(Dataset!E:E, 3)	2,239

These values indicate high variability in death counts, with a right-skewed distribution where most values are low but a few are extremely high.

### 3.2 Analysis of Deaths by Gender

The total number of deaths was analysed across three gender categories: male, female, and both (aggregated totals). The results are summarised below:

DEATHS BY GENDER		
Sex	▼	Total of Deaths
both		6952257
female		3011605
male		4054892
<b>Grand Total</b>		<b>14018754</b>

Sex	Total Deaths
Male	4,054,892
Female	3,011,605
Both	6,952,257

This comparison shows that male deaths exceed female deaths, possibly due to gender-specific health factors, occupations, or behaviours. The "both" category reflects overall aggregated data across the sexes.

### 3.3 Analysis of Deaths by Ethnicity

The ethnicity breakdown reveals noticeable differences in total deaths among groups:

DEATHS BASED ON ETHNICITY	
Ethnicity	Total of Deaths
bumi_malay	3646322
bumi_other	667946
chinese	1837672
indian	580030
other_citizen	118118
other_noncitizen	210720
<b>Grand Total</b>	<b>7060808</b>

Ethnicity	Total Deaths
bumi_malay	3,646,322
chinese	1,837,672
bumi_other	667,946
indian	580,030
other_citizen	210,720
other_noncitizen	695,794
other	118,118

### 3.4 Analysis of Death by States

An important dimension of understanding health outcomes across Malaysia is the geographic distribution of mortality. This subsection presents a summary of total recorded deaths across all 16 Malaysian states and federal territories, offering insights into regional health disparities and potential infrastructure strain.

#### Total Deaths by States (2000–2023)

DEATHS BY STATES	
States	Total of Deaths
Johor	1713278
Kedah	1212132
Kelantan	981458
Melaka	470508
Negeri Sembilan	606060
Pahang	766448
Perak	1601962
Perlis	172328
Pulau Pinang	922644
Sabah	979848
Sarawak	1135804
Selangor	2081926
Terengganu	611400
W.P. Kuala Lumpur	712990
W.P. Labuan	31894
W.P. Putrajaya	18074
<b>Grand Total</b>	<b>14018754</b>

States	Total Deaths
Selangor	2,081,926
Johor	1,713,278
Perak	1,601,962
Kedah	1,212,132
Sarawak	1,135,804
Sabah	979,848
Kelantan	981,458
Pulau Pinang	922,644
Pahang	766,448
W.P. Kuala Lumpur	712,990
Terengganu	611,400
Negeri Sembilan	606,060
Melaka	470,508
Perlis	172,328
W.P. Labuan	31,894
W.P. Putrajaya	18,074

## **Key Observations:**

### **1. Urban States Lead in Mortality Totals:**

Selangor, Johor, and Perak recorded the highest number of deaths. These states are also among the most populous and urbanised, suggesting a correlation between population size, urban health challenges, and higher reporting coverage.

### **2. Smaller States Show Lower Totals:**

Territories like Putrajaya, Labuan, and Perlis report the lowest number of deaths. While this may reflect smaller populations, it could also indicate data reporting limitations or reduced access to healthcare services.

### **3. Implications for Health Planning:**

States with high death counts may require expanded healthcare capacity, especially in intensive care units, emergency response, and pandemic preparedness. Meanwhile, lower-death states should not be overlooked, as their lower figures might mask underlying healthcare access gaps.

## **3.5 Insights Summary**

1. Death counts vary significantly across gender, ethnicity, and states.
2. Males consistently show higher mortality than females across all groups.
3. The Bumi\_Malay community records the highest deaths, followed by the Chinese and Indian groups.
4. Selangor, Johor, and Perak have the highest state-level deaths, likely due to population size and urban density.
5. Putrajaya, Labuan, and Perlis show the lowest totals, possibly reflecting smaller populations or limited reporting.

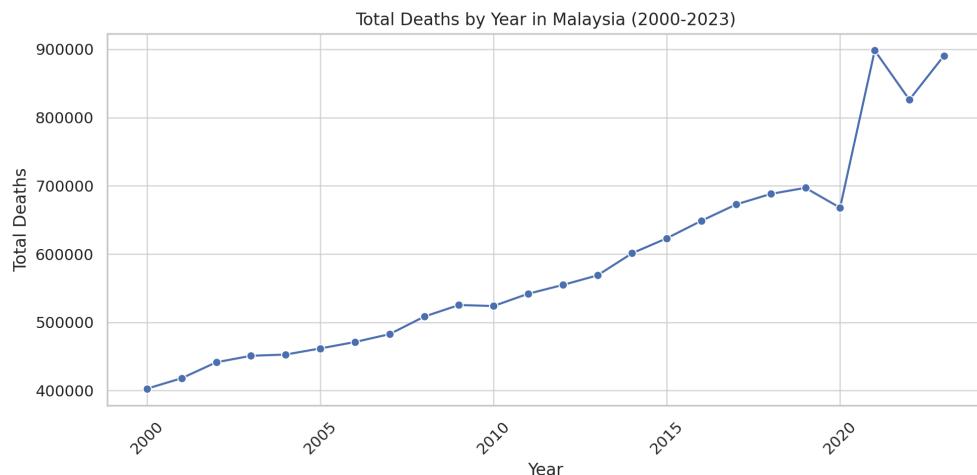
## 4.0 TREND ANALYSIS

To understand the progression of mortality in Malaysia over time, we conducted a year-by-year analysis of total death counts from 2000 to 2023.

### 4.1 Line Chart: Key Observations

- Gradual Increase (2000–2019): Annual deaths rose steadily from approximately 400,000 in 2000 to around 697,000 in 2019. This reflects population growth, urbanisation, and a potentially ageing population.
- COVID-19 Impact (2021): A dramatic spike occurred in 2021, where deaths surged to 898,000, marking the highest single-year total in the dataset. This likely corresponds to the COVID-19 pandemic.
- Post-pandemic stabilisation: Death rates remained elevated post-2021, suggesting ongoing effects of the pandemic, long-term complications, or increased healthcare strain.

A line chart visualising this trend was created using Excel, clearly illustrating the sharp spike in 2021 and gradual growth in earlier years.



The dramatic rise in 2021 underlines the need for emergency preparedness and pandemic response capacity. The overall trend shows a consistent rise in mortality, suggesting an ageing population or widening health gap.

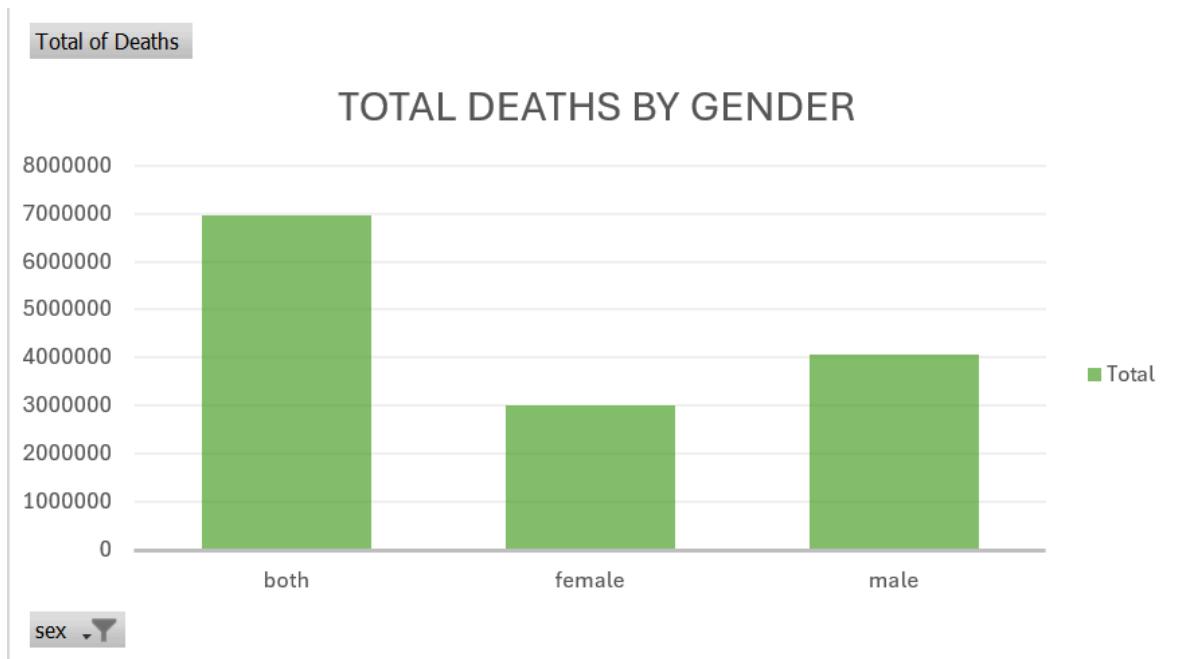
## 5.0 DATA VISUALIZATION

To support the findings, several visualisations were developed using Microsoft Excel to display death data by gender, ethnicity, and their intersections.

### 5.1 Death by Gender

- **Male** deaths: 4,054,892
- **Female** deaths: 3,011,605
- **Both** (aggregated): 6,952,257

A column chart highlights the consistent pattern of higher mortality among males.

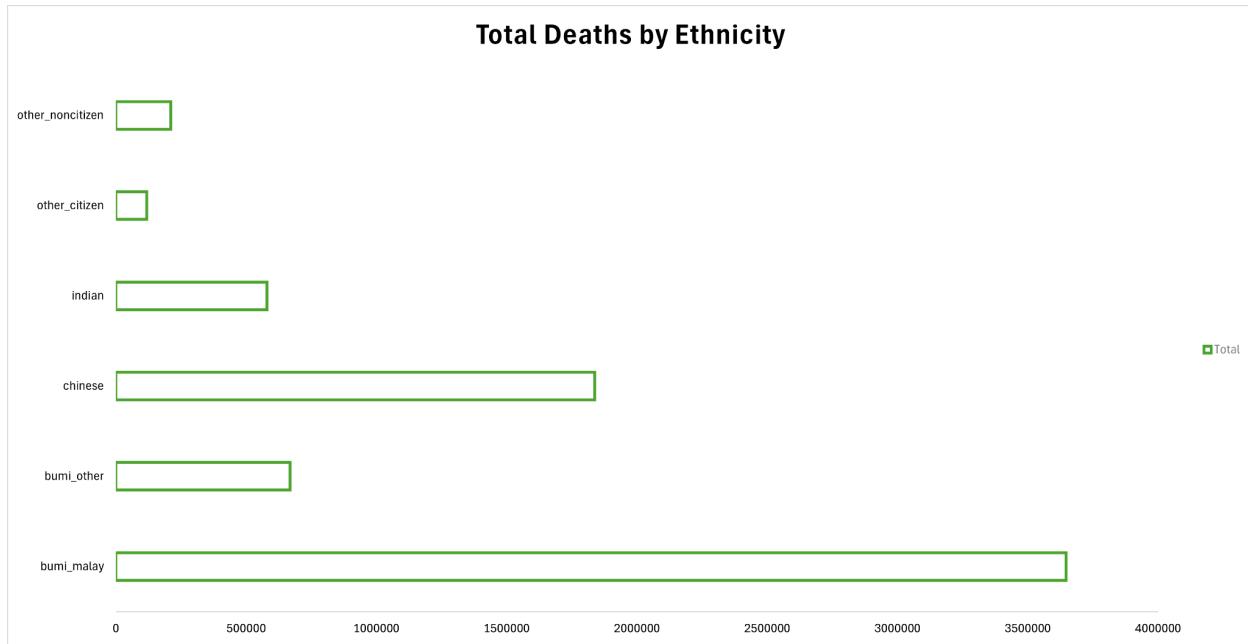


Men consistently showed higher mortality, supporting global findings related to higher vulnerability among males due to biological, occupational, and behavioural risk factors.

## 5.2 Death by Ethnicity

### Key Visual Insights

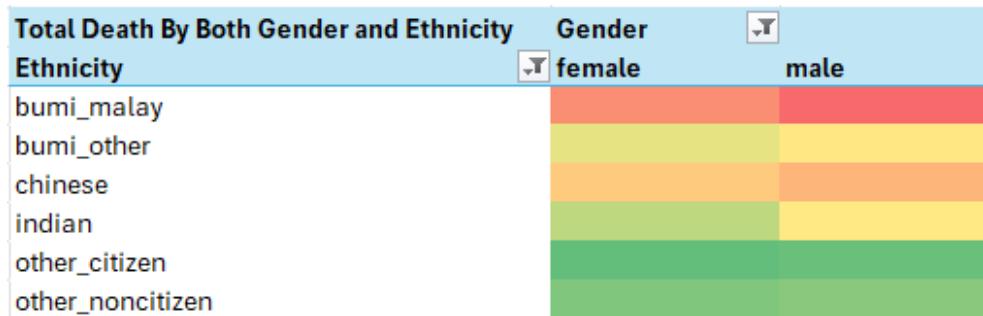
- Bumi\_malay accounted for the highest proportion of deaths, with 3,646,322 recorded.
- The Chinese population followed with 1,837,672 deaths.
- Other notable groups include Indian (580,030), Bumi\_other, and non-citizen categories.



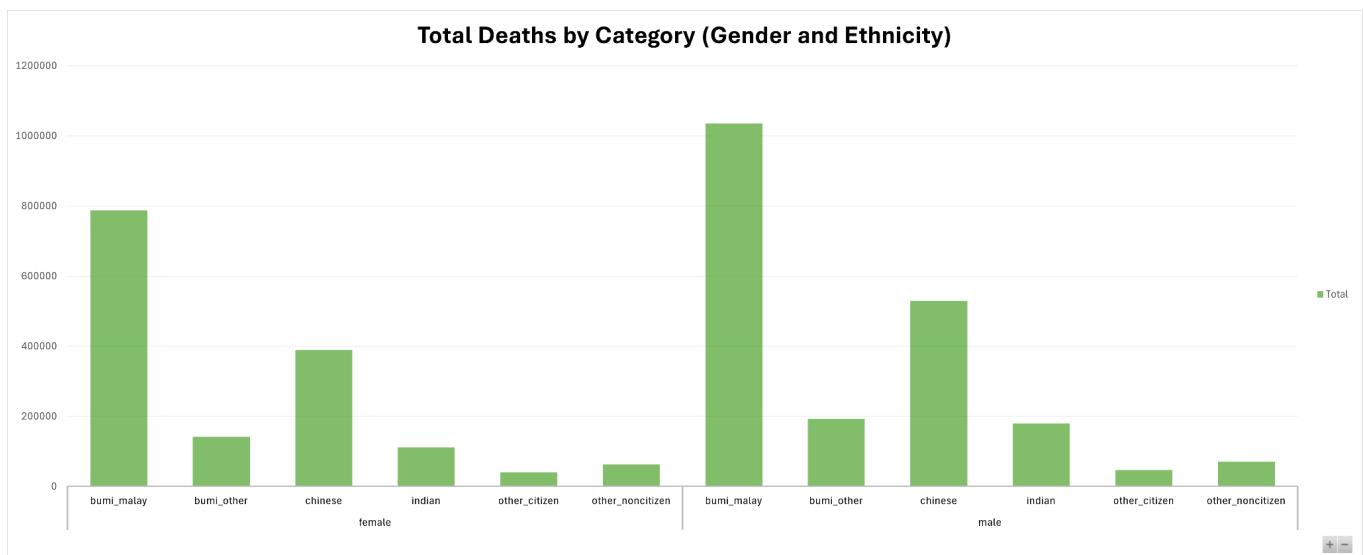
The dominance of Bumi\_malay and Chinese groups may reflect their population size, but could also signal disparities in healthcare access, especially for chronic conditions.

## 5.3 Death by Category (Gender and Ethnicity)

To better understand the intersectional impact of gender and ethnicity on mortality, a heatmap was constructed using a pivot table. It visually represents the total number of deaths from 2000 to 2023, categorised by ethnicity (rows) and gender (columns). A colour scale was applied, ranging from green (low) to red (high), to enhance visibility and emphasise high-risk groups.



A stacked bar chart is also used to represent the contribution of each sex-ethnicity pair to total deaths.



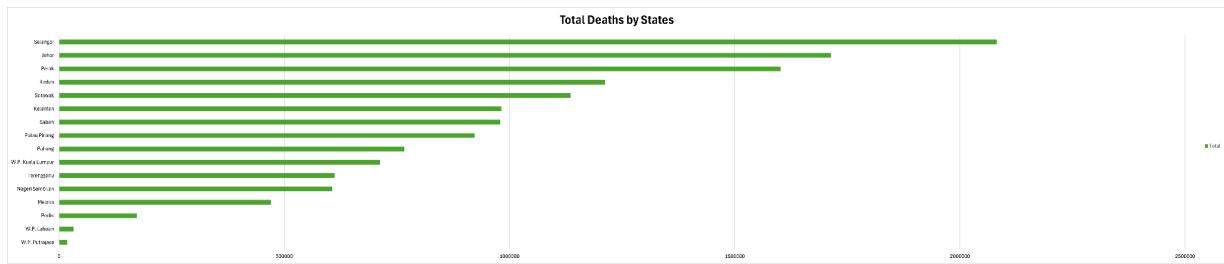
## Key Visual Insights

- **Highest Mortality Subgroup:**
  - The Bumi\_malay (Male) category stands out with the highest recorded death count (1,035,349), indicated by the deepest red tone.
  - This subgroup accounts for a significant portion of national mortality, pointing to a critical demographic for health policy targeting.
- **Gender Disparity Across All Ethnicities:**
  - Males have higher death counts than females in every ethnic group.

- **Relatively Low Mortality in Marginalised Groups:**
  - Ethnic groups such as Other\_citizen and Other\_noncitizen have noticeably lower death counts for both genders.
  - However, this may not directly indicate better health outcomes. Instead, it could reflect lower population sizes, underreporting, or barriers to healthcare access and documentation for these communities.
- **Moderate Risk Clusters:**
  - Chinese and Bumi\_other males form a moderate risk cluster, with mortality figures in the 500 – 200k range.
  - These groups warrant preventive intervention strategies, especially given their urban residency patterns and ageing population dynamics.
- **Indian Group Analysis:**
  - The Indian subgroup, while smaller in population, still reflects notable gender disparities, with Indian males accounting for 179,133 deaths vs 110,882 for females.
  - This reinforces the need for community-level engagement and screening programs.

## 5.4 Death by States

To enhance understanding of the regional distribution of mortality in Malaysia, a clustered bar chart was created to visually compare total death counts across all 16 states and federal territories from 2000 to 2023.



### Key Visual Insights:

- Selangor, Johor, and Perak are visually dominant on the chart, reinforcing their positions as the top three contributors to total mortality in Malaysia.
- States like W.P. Putrajaya, Labuan, and Perlis occupy minimal space on the chart, reflecting significantly lower death totals, possibly due to smaller populations or limited data granularity.
- The contrast between urban and rural states is stark, making it easier to identify regions requiring healthcare scaling versus those needing data collection or outreach improvements.

## **6.0 CONCLUSION**

Over the past 23 years, this study reveals that mortality in Malaysia is not evenly spread. One of the clearest patterns is that men consistently experience higher death rates than women. This reflects global trends and suggests that men may be more exposed to health risks through their jobs, lifestyle habits, or limited engagement with healthcare services. It's a reminder that health campaigns and services need to be more tailored to reach and support men effectively.

Ethnic differences in death counts also stand out. The Bumi\_malay community, which makes up the majority of Malaysia's population, recorded the highest number of deaths, followed by the Chinese and Indian groups. While this might partly reflect population size, it also points to possible gaps in healthcare access, education, or economic conditions that impact these communities differently. Notably, Bumi\_malay males appear to be the most affected group, showing the need for focused public health interventions.

At the other end of the spectrum, minority and non-citizen groups reported fewer deaths, but this doesn't necessarily mean better health. These numbers might be low due to underreporting or barriers to accessing the healthcare system. It's important not to overlook these groups when designing national health policies. The sharp increase in deaths during 2021, likely due to COVID-19, exposed vulnerabilities in the healthcare system. Even after the pandemic peak, mortality rates stayed high, suggesting long-term impacts that the country must be prepared for in future crises.

Ultimately, improving health outcomes in Malaysia requires more than just general policies. It calls for data-driven, targeted efforts that understand and respond to the different realities people face based on their gender, ethnicity, and social background. This is essential for making real progress toward SDG 3: Good Health and Well-Being for all.

### **6.1 Implications for SDG 3**

These findings demonstrate the importance of targeted, demographic-specific health policies. To achieve SDG 3: Good Health and Well-Being, Malaysia must prioritise interventions that address the health needs of specific gender-ethnic subgroups, particularly those identified as high-risk. Through data-driven planning and inclusive health initiatives, disparities in mortality can be reduced, leading to better overall public health outcomes.

## **7.0 RECOMMENDATIONS**

Based on the visual and statistical findings of this study, the following policy and action-oriented recommendations are proposed to strengthen Malaysia's public health response and ensure equitable progress toward SDG 3: Good Health and Well-Being.

### **7.1 Strengthen Urban Healthcare Systems**

High-density states such as Selangor and Kuala Lumpur experienced a disproportionately high number of deaths during the pandemic. These regions should be prioritised for healthcare capacity upgrades, including:

- Expanding ICU bed capacity and access to emergency medical equipment.
- Enhancing the supply chain for oxygen, medications, and critical care tools.
- Scaling up digital telehealth systems to relieve physical infrastructure and ensure continuous care during surges.

### **7.2 Implement Male-Focused Health Campaigns**

Throughout the dataset, male individuals, especially Bumi Malay and Indian men, consistently showed higher mortality rates. This group requires dedicated interventions such as:

- Awareness programs on cardiovascular health, smoking cessation, and occupational safety.
- Preventive screenings and education campaigns focused on men's health risks.
- Post-pandemic recovery support, including mental health services for male-dominant industries and communities.

### **7.3 Enhance Ethnic-Sensitive Health Outreach**

Health disparities across ethnic groups indicate the need for culturally competent strategies. Recommendations include:

- Developing multilingual health education materials in Malay, Mandarin, Tamil, and indigenous languages.

- Engaging local community leaders from Chinese, Indian, and East Malaysian backgrounds in awareness campaigns.
- Ensuring health messaging is culturally relevant, trusted, and accessible, especially during health emergencies.

#### **7.4 Build Predictive and Real-Time Health Monitoring Systems**

COVID-19 highlighted the need for real-time response and forecasting capabilities. To prevent future surges:

- Establish a centralised digital platform that tracks mortality and public health trends at the national and state levels.
- Integrate predictive analytics to detect early warning signs of health crises, enabling proactive decision-making.
- Use live data to mobilise medical resources where and when they are most needed.

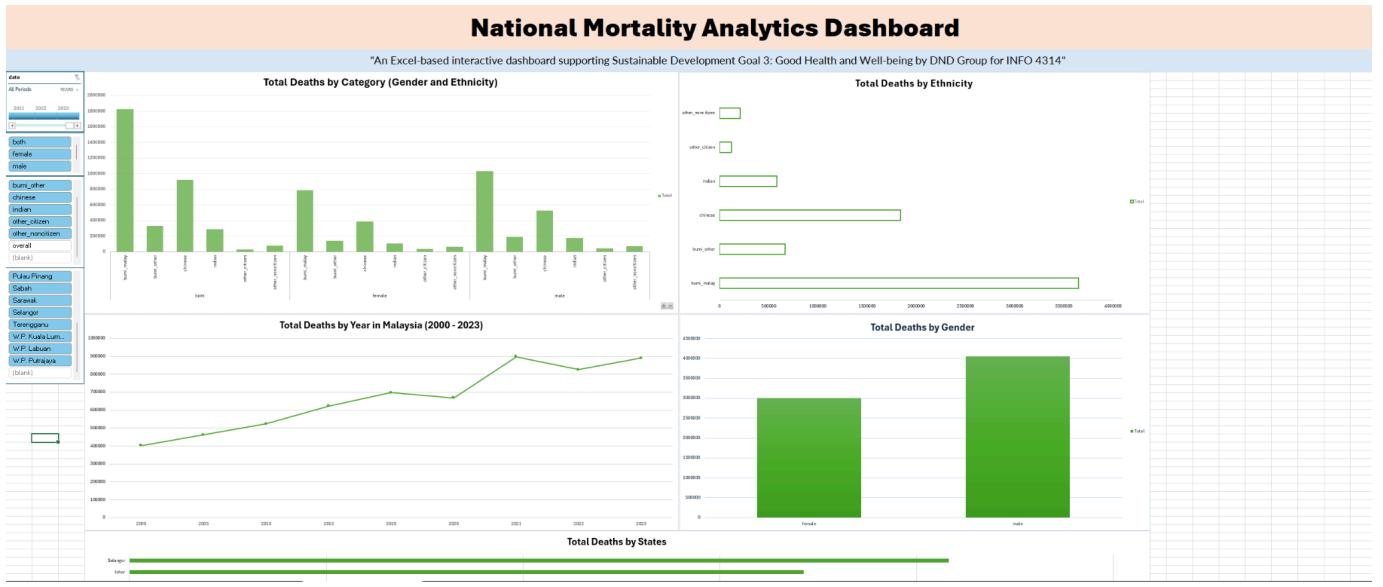
#### **7.5 Improve Equity in Underserved Territories**

Although Labuan, Putrajaya, and other small regions recorded fewer deaths, they remain vulnerable due to limited healthcare infrastructure. To ensure equitable healthcare delivery:

- Deploy mobile health units, especially during outbreaks or lockdown periods.
- Maintain strategic reserves of medical supplies and essential services in remote territories.
- Include these areas in national-level emergency planning, ensuring no region is left behind.

## 8.0 APPENDICES

8.1 Analytics Dashboard (Excel) ▶ DND Business Data Analytics



8.2 Statistics Table Analysis (Pivot Table) ▶ DND Business Data Analytics

Overall Summary										Quartiles					
Mean	Median	Standard Deviation	Minimum	Maximum	25%		50%		75%						
1785	242	3181.458319	0	42051	34		242		2239						
<b>DEATH BY ETHNICITY</b>															
Ethnicity	Total of Deaths	DEATHS BY STATES		Ethnicity		Gender	female	male	Grand Total						
bumi_malay	3646322	Johor	1713278	bumi_malay	787812	female	1035349	male	1823161						
bumi_other	667946	Kedah	1212132	bumi_other	141449	female	192524	male	333973						
chinese	1837672	Kelantan	981458	chinese	389132	female	529704	male	918836						
indian	580030	Melaka	470508	indian	110882	female	179133	male	290015						
other_citizen	118118	Negeri Sembilan	606060	other_citizen	40829	female	46790	male	87619						
other_noncitizen	210720	Pahang	766448	other_noncitizen	63114	female	70806	male	133920						
<b>Grand Total</b>	<b>7060808</b>	Perak	1601962	<b>Grand Total</b>	<b>1533218</b>	female	<b>2054306</b>	male	<b>3587524</b>						
<b>DEATHS BY GENDER</b>															
Sex	Total of Deaths	Deaths by Gender		Total Death By Both Gender and Ethnicity		Gender	female	male	Grand Total						
female	3011605	Perlis	172328	Female	1533218	female	2054306	male	3587524						
male	4054892	Pulau Pinang	922644	Male	2054306	female	3587524	male	1533218						
<b>Grand Total</b>	<b>7066497</b>	Sabah	979848	Female	1533218	female	2054306	male	3587524						
<b>DEATHS BY STATE</b>															
State	Total of Deaths	Deaths by State		Ethnicity		Gender	female	male	Grand Total						
Johor	1713278	Sabah	979848	bumi_malay	787812	female	1035349	male	1823161						
Kedah	1212132	Sarawak	1135804	bumi_other	141449	female	192524	male	333973						
Kelantan	981458	Selangor	2081926	chinese	389132	female	529704	male	918836						
Melaka	470508	Terengganu	611400	indian	110882	female	179133	male	290015						
Negeri Sembilan	606060	W.P. Kuala Lumpur	712990	other_citizen	40829	female	46790	male	87619						
Pahang	766448	W.P. Labuan	31894	other_noncitizen	63114	female	70806	male	133920						
Perak	1601962	W.P. Putrajaya	18074	<b>Grand Total</b>	<b>14018754</b>	female	<b>2054306</b>	male	<b>3587524</b>						
<b>ETHNICITY</b>															
Ethnicity	Total of Deaths	Ethnicity		Gender		Female	Male	Grand Total							
bumi_malay	3646322	bumi_malay	787812	Female	1035349	female	1823161	male	3587524						
bumi_other	667946	bumi_other	141449	Male	2054306	female	333973	male	1533218						
chinese	1837672	chinese	389132	Female	529704	female	918836	male	4054892						
indian	580030	indian	110882	Male	179133	female	290015	male	118118						
other_citizen	118118	other_citizen	40829	Female	87619	female	210720	male	2054306						
other_noncitizen	210720	other_noncitizen	63114	Male	133920	female	3587524	male	1533218						
<b>Grand Total</b>	<b>7066497</b>	<b>Grand Total</b>	<b>14018754</b>	<b>Female</b>	<b>2054306</b>	female	<b>3587524</b>	male	<b>1533218</b>						