UNIVERSITI KEBANGSAAN MALAYSIA

The National University of Malaysia

Semester 1, 2024/2025

TTTU6314 BUSINESS ANALYTICS

Analysis of Water Consumption and Crime Rates in Malaysia

Prepared by:

Name	Matric Number
MOHAMED MOUBARAK MOHAMED MISBAHOU MKOUBOI	P139575

Lecturer:

Dr. Ahmad Tarmizi Bin Abdul Ghani

Introduction

This study analyzes the connection between crime rates and water use in Malaysia's several states. To find possible relationships and patterns, we examine trends in both household and non-domestic water usage along with crime categories using data from data.gov.my.

Data Sources

1. Water Consumption Dataset

- Source: Data.gov.my - Water Consumption (Link)
- Description: Annual water consumption by state, divided into domestic and non-domestic sectors.
- Key Columns:
  - state: State of Malaysia
  - sector: Domestic or non-domestic
  - date: Yearly data
  - value: Water consumption in liters

2. Crime Rate Dataset

- Source: Data.gov.my - Water Consumption (Link)
- Description: Crime incidents reported by district and categorized into different types.
- Key Columns:
  - state: State of Malaysia
  - district: District of crime occurrence
  - category: Type of crime (assault, property crime, etc.)
  - type: Specific crime type
  - date: Yearly data
  - crimes: Number of reported crimes

Data Preprocessing and Merging

1. Challenges in Merging

One of the main obstacles to merging these datasets was the difference in detail:

- District-level data is provided by the crime dataset.

- State-level values are reported in the water usage dataset.

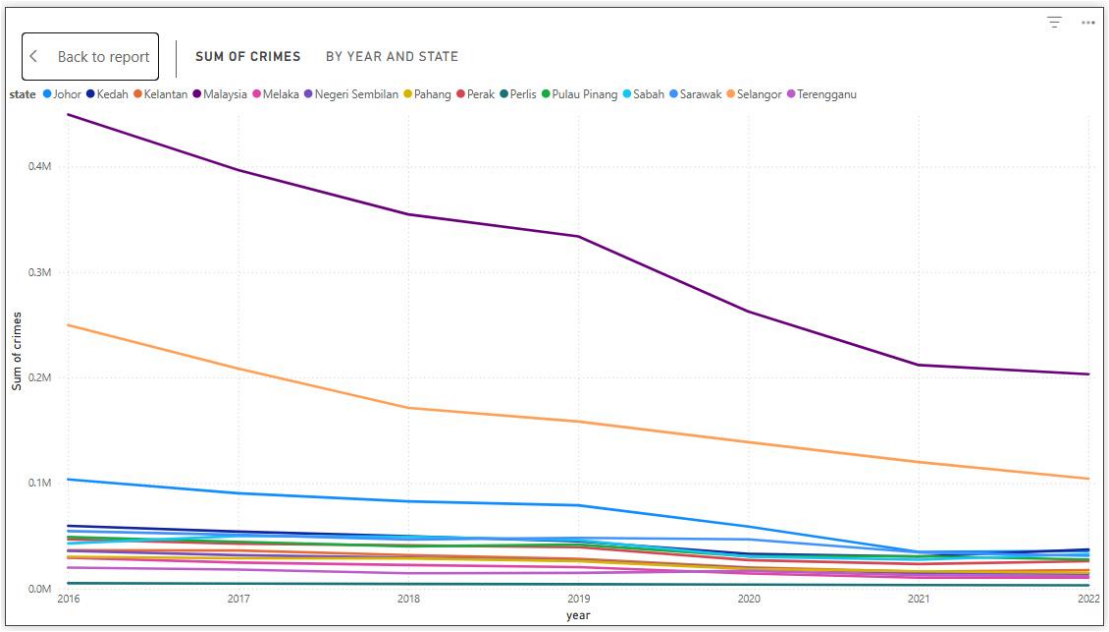
To address this problem, we counted the total number of crimes committed in each state annually in order to aggregate the crime data at the state level. As a result, we were able to align the two datasets using the same key (state and year).

2. Merging Process

1. Extracting Year: In order to guarantee a correct alignment during the merger, we extracted the year because date columns were included in both dataset.
2. Crime Data Aggregation: The total number of crimes was calculated by grouping the crime data by state, year, category, and type.
3. Merging on Common Fields: To make sure that only data points that were accessible in both datasets were kept, we carried out an inner join on state and year.
4. Managing Missing entries: A careful inspection was done to make sure the combined dataset had no missing data.

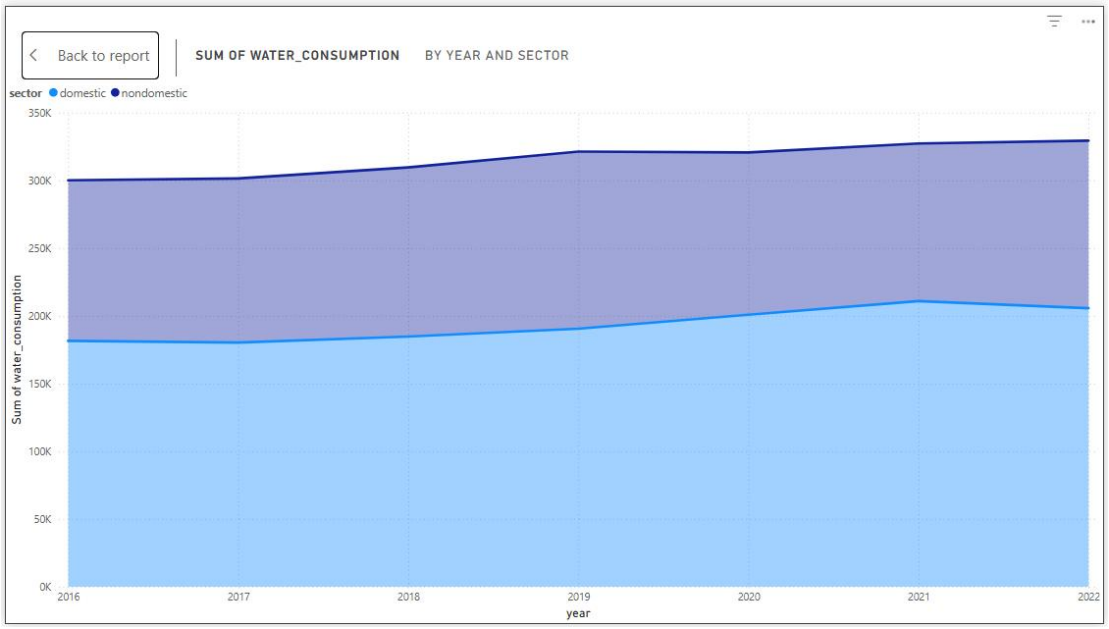
Graphs and Insights

1. Total Crimes by Year and State (2016-2022)



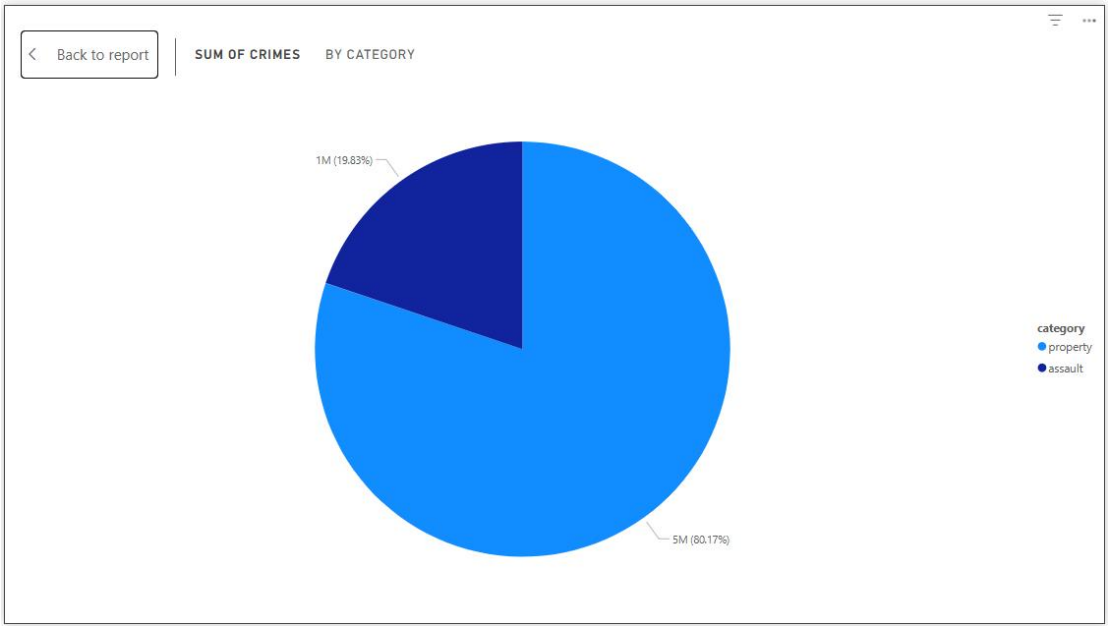
States have different tendencies in crime over time. There are variations, with certain states (such Johor and Selangor) continuously having greater crime rates. Finding the causes of crime rates in various areas might be improved by pattern analysis.

2. Total Water Consumption by Year and Sector (2016-2022)



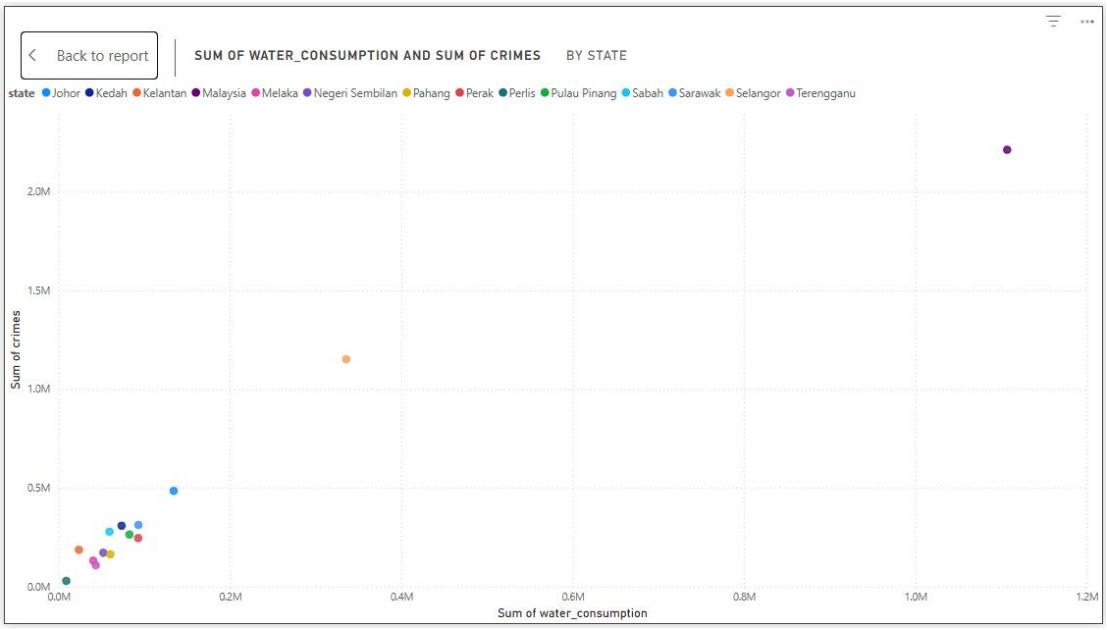
The two types of water usage are classified as domestic and non-domestic. Consumption at home seems to be steadily increasing. The patterns indicate whether consumption is rising or falling, which can help guide resource management policy decisions.

3. Distribution of Crimes by Category



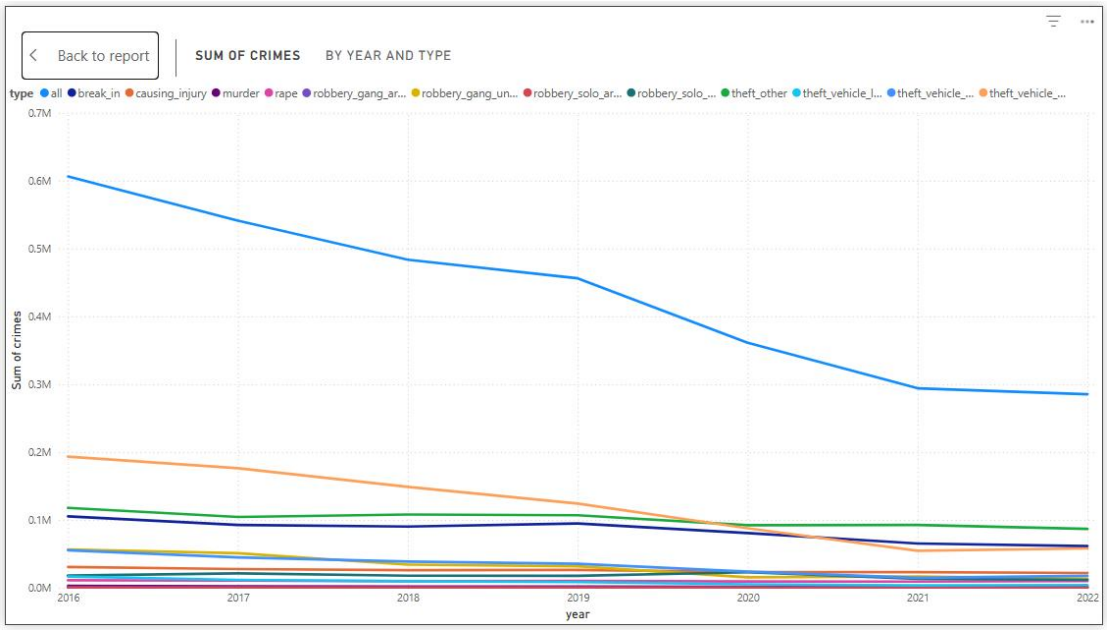
The majority of crimes (80.17%) are property crimes, with assault-related crimes making up the remaining 19.83%. This suggests that security precautions against theft and burglary are more important than ever.

4. Correlation Between Water Consumption and Crimes by State



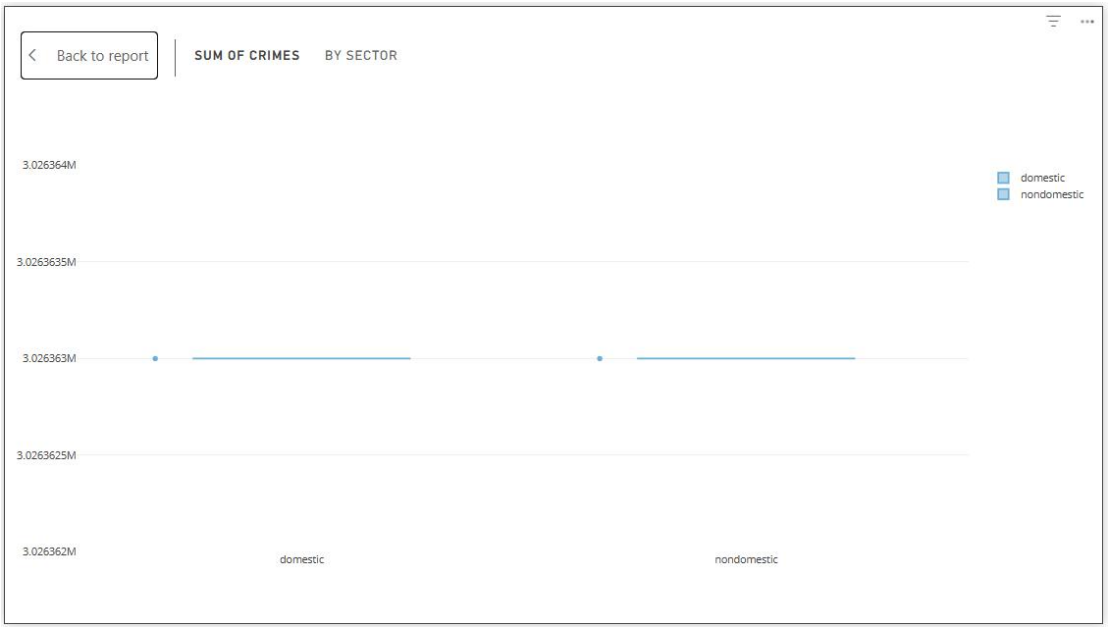
Certain states with greater crime rates also use a lot of water. The correlation between these variables, however, may reflect population density or urbanization levels rather than being a direct cause.

5. Total Crimes by Year and Type (2016-2022)



Trends in many sorts of crime have changed over time. While severe crimes like murder and rape have comparatively lower statistics, theft-related crimes seem to be the most common. Prioritizing crime prevention tactics can be aided by law enforcement's understanding of these tendencies.

6. Total Crimes by Sector

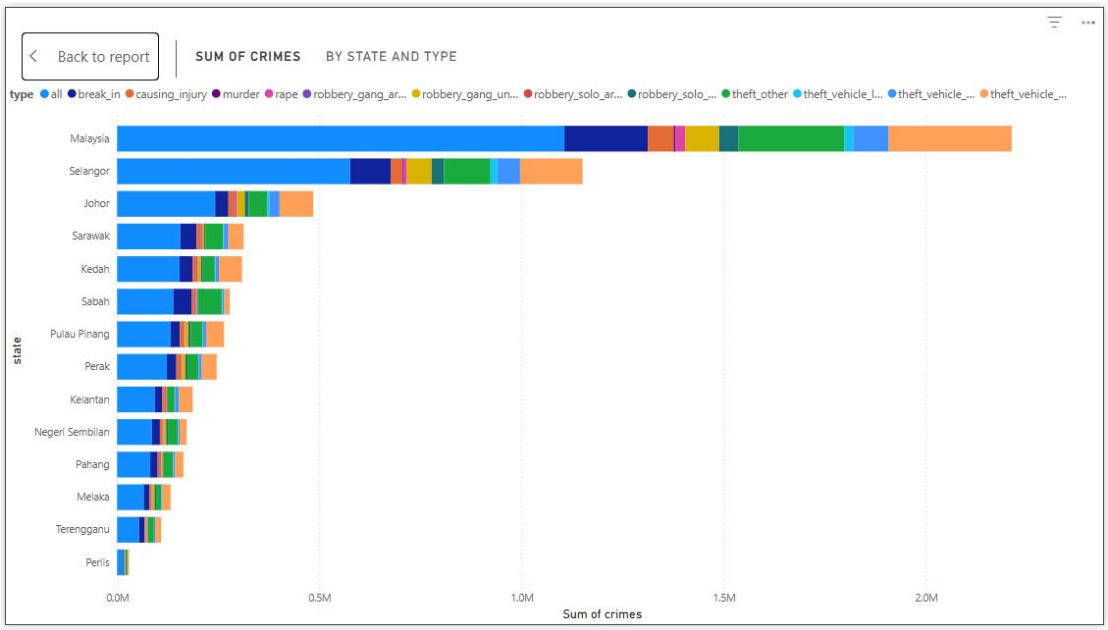


It is easier to determine where crimes are more common when offenses are broken down into domestic and non-domestic categories. While non-domestic crimes usually occur in public places or companies, domestic crimes may involve transgressions relating to the home.

7. Total Crimes by State

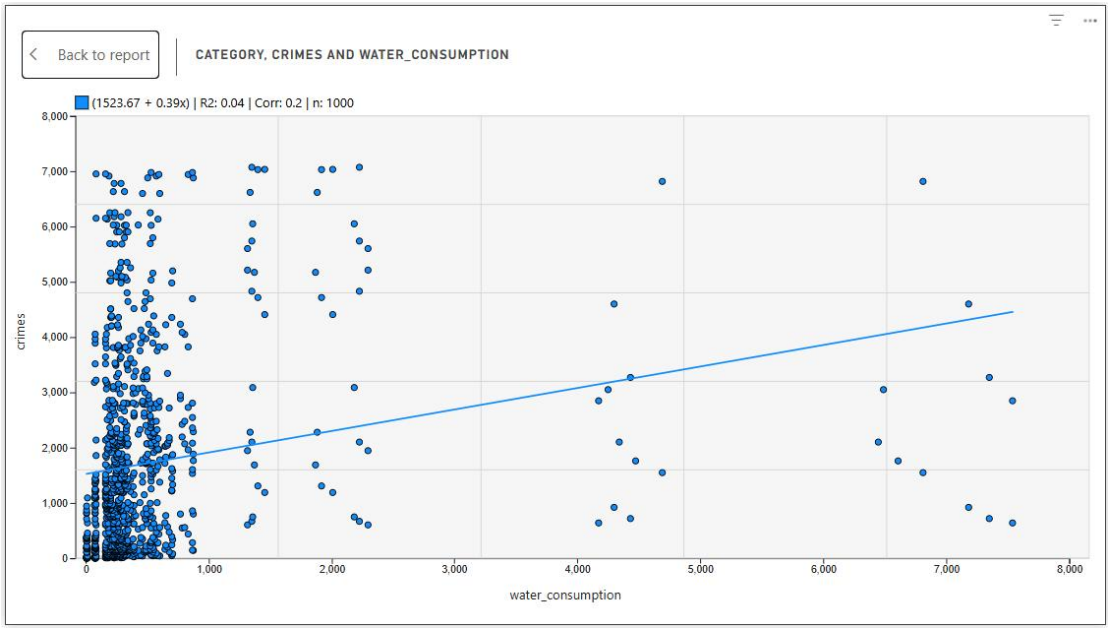
The greatest rates of crime in Malaysia are found in Selangor, Johor, and Sarawak. Perlis and Terengganu are among the states with the lowest rates of crime. This might be a result of variations in the population, the economy, or the efficiency of the police.

8. Total Crimes by State and Type



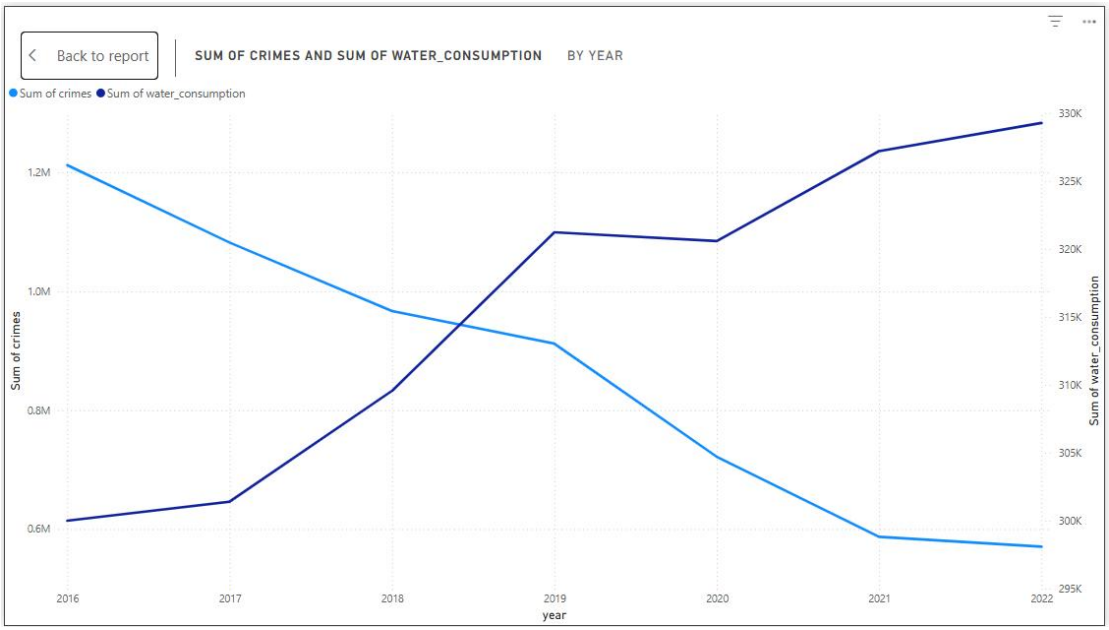
Certain types of crimes are more common in some states than others. For instance, robberies and thefts may be more common in urban states, but patterns may alter in rural locations. Customized law enforcement tactics require an understanding of state-specific crime distributions.

9. Relationship Between Crime Category, Water Consumption, and Crime Rate



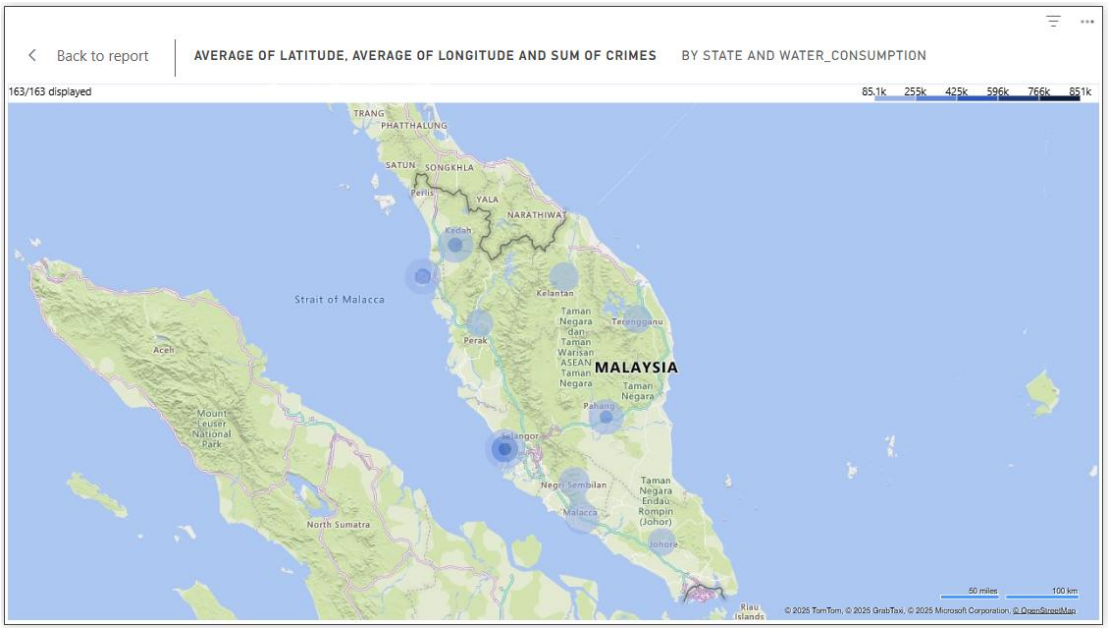
Crime and water use have a weak link ($R2 = 0.04$, correlation = 0.2), indicating that although urbanization may have an impact on both, the two variables are not directly related.

10. Comparison of Total Crimes and Water Consumption by Year



Trends in water use and crime rates over time reveal whether there is any apparent relationship. Although there are occasional variations in both parameters, there isn't much proof of a direct correlation.

11. Geospatial Distribution of Crimes and Water Consumption by State(Bonus)



Areas with high crime and use of resources can be visually represented by mapping crime and water consumption spatially. This can assist law enforcement and legislators in concentrating on hotspot areas.

Conclusion

Through an analysis of water consumption, this study suggests a potential connection between socioeconomic background and developments in crime. States with higher water usage typically report higher crime rates, most likely as a result of urbanization and population density, however a direct causal relationship has not been shown. For a more thorough examination, future studies could include other variables such as population growth and income levels.

References

- Data sources from data.gov.my
- Joining dataset using pandas library by Gözde Madendere from Medium.com
- How to plot graphs using Power Bi from GitHub