



SCHOOL OF COMPUTING

Faculty of Engineering

Project Proposal Form MCSD 6215
Sem: Session:

SECTION A: Project Information

Program Name: **Masters of Science (Data Science)**

Subject Name: **Project 1 (MCSD 6215)**

Student Name: **Sun Qi**

Metric Number: **MCS231030**

Student Email & Phone: **sunqi@graduate.utm.my +60 17-208-2853**

Project Title: **Traffic flow analysis about the Malaysian city Johor Bahru**

Supervisor 1:

Supervisor 2 / Industry

Advisor(if any):

SECTION B: Project Proposal

Introduction:

Traffic problems have been a major issue related to people's livelihood since ancient times, and smooth and convenient transportation is conducive to people's travel. However, in the city of Johor Bahru in Malaysia, traffic problems affect people's travel every day, especially in several large arterial roads, Jalan Wong Ah Fook, Jalan Skudai, Jalan Abdullah Ibrahim, these arteries are important traffic arteries connecting the city center and surrounding areas because they are connected to the city center and intersect with many main roads, so they are more likely to become hot spots for traffic congestion during peak hours. Commuting to and from work in the morning and evening is a serious hindrance to people's travel. At the same time, due to the proximity of Johor Bahru to Singapore, there are a large number of migrant workers coming and going in Singapore every day, and these people also exacerbate the traffic problem in Johor Bahru, so accurate and timely traffic flow forecasting is essential to help people plan their trips properly and help the traffic management department to carry out effective traffic control.

Problem Background:

The problem of traffic congestion in Johor Bahru is a problem that plagues both local residents and Malaysian and Singaporean workers, and the problem is caused by a combination of factors

(1) Special geographical location

Johor Bahru is close to Singapore and is an important hub for the movement of people and goods between the two countries, with a large number of tourists and cross-border workers flowing around every day. Carrying nearly 300,000 people a day, Johor Cswy is one of the busiest land borders in the world.

(2) Rapid urban development and population growth

Johor Bahru is one of the fastest growing cities in Malaysia, with an impressive rate of population growth and urban expansion, and the location of many universities attracting many students from home and abroad, and the transportation infrastructure is far from keeping up with the rapid development of the city, which greatly increases the pressure on existing roads.

and (3) the limitations of public transportation

Public transport in Johor Bahru is not well developed, and people tend to travel independently, with private cars and motorbikes being the main means of transport, which also leads to an increase in vehicles on the road during rush hour, which congests traffic.

(4) Insufficient road planning

Some major roads such as Jalan Tebrau, alan Wong Ah Fook were not designed to accurately predict future traffic flows, resulting in insufficient design capacity of the roads. In addition, the layout and timing of traffic lights in Johor Bahru are not reasonable.

(5) The impact of tourism

Johor Bahru has several large shopping malls as well as some tourist attractions, which attract many tourists and locals on weekends and holidays, thus exacerbating the traffic problem during these times.

Aim of the Project:

The goal of traffic flow analysis is to analyze vehicle information on the road, as well as road information on congested road sections at different times, as well as other traffic conditions to ensure that traffic is safe, efficient, and sustainable. Analyzing traffic flow information is of great significance for urban planning, reducing the work of road management departments and improving people's travel efficiency.

Scopes of the Project:

Geographical Scope : Focused on the Malaysian city of Johor Bahru

Data Scope: Leverage a variety of data sources, including historical traffic flow data, real-time traffic flow data, public transit data, and more

Methodology: Data Analysis & Visualization, Python

Expected Contribution of the Project:

1. Improve the road safety index: Alleviate traffic pressure by optimizing and analyzing traffic flow data
2. Optimize the use of infrastructure: ensure that transportation hubs and various roads exert their capacity
3. Reduce the burden of government work

Expected Contribution of the Project:

Project Requirements:

Software: MySQL, Visual Studio Code, RStudio, scikit-learn, TensorFlow
Hardware: High-Performance Computers, Cloud storage options
Technology/Technique/
Methodology/Algorithm: Python, R, data visualization, machine learning, EDA,

Type of Project (Focusing on Data Science):

- [] Data Preparation and Modeling
[] Data Analysis and Visualization
[] Business Intelligence and Analytics
[] Machine Learning and Prediction
[] Data Science Application in Business Domain

Status of Project:

- [] New
[] Continued
If continued,
what is the
previous title?

SECTION C: Declaration

I declare that this project is proposed by:

- [] Myself
[] Supervisor/Industry Advisor ()

Student
Name:

Signature

Date

SECTION D: Supervisor Acknowledgement

The Supervisor(s) shall complete this section.

I/We agree to become the supervisor(s) for this student under aforesaid proposed title.

Name of Supervisor 1: _____

Signature _____ **Date** _____

Name of Supervisor 2
(if any): _____

Signature _____ **Date** _____

SECTION E: Evaluation Panel Approval

The Evaluator(s) shall complete this section.

Result:

[] FULL APPROVAL

[] CONDITIONAL APPROVAL

(Major)*

[] CONDITIONAL APPROVAL

[] FAIL*

(Minor)

* Student has to submit new proposal form considering the evaluators' comments.

Comments:

Name of Evaluator 1:

Signature

Date

Name of Evaluator 2:

Signature

.....

Date

CHAPTER 1

INTRODUCTION

1.1 Introduction

Traffic problems have been a major issue related to people's livelihood since ancient times, and smooth and convenient transportation is conducive to people's travel. However, in the city of Johor Bahru in Malaysia, traffic problems affect people's travel every day, especially in several large arterial roads, Jalan Wong Ah Fook, Jalan Skudai, Jalan Abdullah Ibrahim, these arteries are important traffic arteries connecting the city center and surrounding areas because they are connected to the city center and intersect with many main roads, so they are more likely to become hot spots for traffic congestion during peak hours. Commuting to and from work in the morning and evening is a serious hindrance to people's travel. At the same time, due to the proximity of Johor Bahru to Singapore, there are a large number of migrant workers coming and going in Singapore every day, and these people also exacerbate the traffic problem in Johor Bahru, so accurate and timely traffic flow forecasting is

essential to help people plan their trips properly and help the traffic management department to carry out effective traffic control.

This study aims to analyze the traffic in Johor Bahru through the data provided by the municipal department, drone monitoring data, data obtained from questionnaires, and research through tools such as python, MATLAB, VISSIM, etc. In addition, the objective of the study is to understand the characteristics of traffic flow and identify the temporal distribution of traffic congestion, evaluate the efficiency of the transportation system and improve road safety issues, and then conduct a reasonable analysis of the traffic construction in Johor Bahru, Malaysia based on the analysis of various data and provide suggestions and measures for improvement. This study can help contribute to transportation in the Johor Bahru urban area in Malaysia.

1.2 Problem Background and Problem Statement

The problem of traffic congestion in Johor Bahru is a problem that plagues both local residents and Malaysian and Singaporean workers, and the problem is caused by a combination of factors (1) Special geographical location Johor Bahru is close to Singapore and is an important hub for the movement of people and goods between the two countries, with a large number of tourists and cross-border workers flowing around every day. Carrying nearly 300,000 people a day, Johor Cswy is one of the busiest land borders in the world. (2) Rapid urban development and population growth Johor Bahru is one of the fastest growing cities in Malaysia, with an impressive rate of population growth and urban expansion, and the location of many universities attracting many

students from home and abroad, and the transportation infrastructure is far from keeping up with the rapid development of the city, which greatly increases the pressure on existing roads. and (3) the limitations of public transportation Public transport in Johor Bahru is not well developed, and people tend to travel independently, with private cars and motorbikes being the main means of transport, which also leads to an increase in vehicles on the road during rush hour, which congests traffic. (4) Insufficient road planning Some major roads such as Jalan Tebrau, alan Wong Ah Fook were not designed to accurately predict future traffic flows, resulting in insufficient design capacity of the roads. In addition, the layout and timing of traffic lights in Johor Bahru are not reasonable. (5) The impact of tourism Johor Bahru has several large shopping malls as well as some tourist attractions, which attract many tourists and locals on weekends and holidays, thus exacerbating the traffic problem during these times.

1.3 Research Questions

- (a) terms of traffic in Johor Bahru, how much is the difference between the traffic flow data during peak and off-peak hours?
- (b). What are the differences in people's travel patterns in different regions (city center vs. suburbs)
- (c). What are people's worries and confusions about road planning in Johor Bahru.
- (d). What is the coverage rate of public transportation.

(e). What are the obstacles to the travel mode of Johor Bahru residents due to traffic congestion.

(f) In terms of future planning, what measures can be taken in Johor Bahru to improve the road problem.

1.4 Objectives of the Research

In order to solve the above problems, this research following objectives:

The goal of traffic flow analysis is to analyze vehicle information on the road, as well as road information on congested road sections at different times, as well as other traffic conditions to ensure that traffic is safe, efficient, and sustainable. Analyzing traffic flow information is of great significance for urban Project1 Proposal Form MSc (Data Science) planning, reducing the work of road management departments and improving people's travel efficiency.

1.5 Scopes of the Project:

In the traffic flow analysis in Johor Bahru, this study will be scoped according to the objectives, area and time of the study to ensure the comprehensiveness of the study. Therefore, the study should include the city centre and suburbs, the bustling and backward areas, the commercial and residential areas of Johor Bahru, and collect data on traffic flow in recent years to capture long-term trends, seasonal trends and make data forecasts.

The data will be sourced from government agencies and officials as well as intelligent transport systems and monitoring devices in Johor Bahru including Johor Public Works Department (JKR), Johor Land Transport Department (JPJ), Johor Police Traffic Department (Royal Malaysia Police Traffic Department), Malaysian Highway Management Company (PLUS Malaysia Berhad), Johor Bahru Local Government (MBJB), Closed Circuit Television (CCTV), Sensors and Automatic Traffic Counters (ATC), drones and aerial photography equipment, vehicle-to-everything data (V2X).

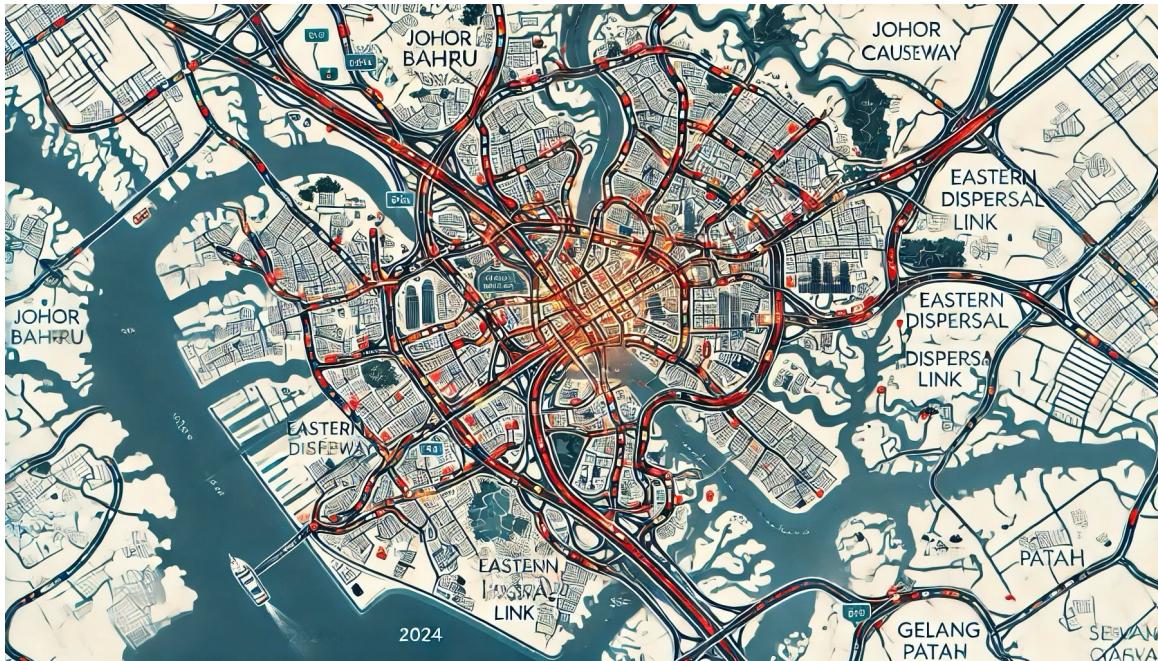
Methodologically, different data analysis and visualization tools, such as Python, will be applied in this study.

Chapter 2

Literature Review

2.1 Introduction

This chapter conducts a literature review of the traffic flow in Johor Bahru, Malaysia in recent years. By studying the flow of cars and pedestrians on the road, road use can be optimized, traffic congestion reduced and safety improved. This chapter will summarize and organize the direction of data-driven methods and future trends.



2.2 Visualization of road congestion in Johor Bahru

This is a visualization of traffic congestion on major roads in Johor Bahru in 2024. The red mark in the picture represents severe congestion, yellow represents moderate congestion, and green represents basically smooth traffic. The figure also shows several key congestion locations, such as the Singapore-Johor Bahru Bridge and the Eastern

Evacuation Link (EDL). The reason for the traffic congestion is that the Johor Bahru government is planning a number of road traffic upgrades and planned construction projects. The projects are divided into short, medium and long-term projects and cover the construction of a new road section from Gelang Patah CIQ interchange to Tanjung Pelepas four-junction, road expansion project from Bandar Seri Alam to Pasir Gudang, Senai-Desaru Expressway and from the east Evacuation Highway (EDL) viaduct project connecting the city center to Stulang Laut, etc. These projects are currently exacerbating the traffic pressure in Johor Bahru, which will be discussed in this chapter from two aspects.

2.3 Data-driven approach

In recent years, the advancement of data collection and processing technology has greatly promoted the development of road traffic flow analysis, which makes real-time observation possible.

- (a)Big Data and Internet of Things (IoT)
- (b)Machine learning model
- (c)Regression model for traffic analysis
- (d)Reinforcement learning for traffic signal control

2.3.1

The following are some key technologies and methods for combining the above technologies to analyze traffic flow

- (a)Data collection

IoT devices (sensors, radars, infrared cameras, GPS devices) collect traffic data in real time. ITS transmits data to headquarters via wireless networks.

Historical data collects traffic records, weather conditions, and time distribution to help analyze.

- (b)Data processing and storage

Big data platforms (Hadoop, Spark) process and store large amounts of traffic data.

Cloud computing High-speed and efficient data processing through the cloud.

(c)Real-time monitoring

The traffic management platform integrates and processes the data collected by IoT devices and servers and exports them into visual traffic flow data, helping traffic management personnel to monitor in real time and handle emergencies in a timely manner.

(d) IV.Data Visualization and Reporting

Visualization tools (3D models, dashboards, maps) display real-time traffic conditions.

Analysis reports provide managers with traffic flow reports and trend analysis reports.

2.4 Future Trends

Recently focused technologies can also provide support for traffic flow analysis

(a)Autonomous driving technology for vehicles

(b)Smart cities and traffic management systems

(c)AI-driven real-time traffic control systems

(d)Dynamic traffic assignment (DTA)

2.4.1 Vehicle Autonomous Driving Technology

Autonomous vehicles can provide accurate data by collecting and analyzing the speed, location and road conditions of vehicles in real time. Autonomous vehicles have built-in V2V and V2I systems that can communicate with each other, which can effectively improve the efficiency of traffic network coordination. Through these data, traffic flow can be accurately predicted and traffic congestion can be reduced.

2.4.2 Smart City and Traffic Management System

Smart cities collect traffic flow data, vehicle speed, road conditions and other data in real time by integrating sensors, cameras and IoT devices. Correspondingly, the traffic management system uses big data analysis and machine learning models to optimize traffic light control, route planning and emergency response based on the collected data. The two work together to effectively reduce congestion and improve road utilization.

2.4.3 AI-driven Real-time Traffic Control System

The AI-driven real-time traffic control system collects traffic data (such as vehicle speed, traffic accident information, road conditions) in real time and uses AI to perform in-depth calculations to organize and analyze traffic patterns. The system can adjust the duration of traffic lights based on real-time data to allocate lanes, thereby optimizing traffic flow and reducing congestion. AI can also predict potential traffic accident risks in advance through deep algorithms, improving overall traffic efficiency and safety.

2.4.4 Dynamic Traffic Assignment (DTA)

Dynamic traffic assignment uses real-time monitoring to flexibly allocate lanes and guide traffic signals to adapt to current traffic conditions.

Chapter 3

Methodology

3.1 Overview

This chapter describes in detail the research methods used in the Johor Bahru traffic flow analysis project. When analyzing the traffic flow of Johor Bahru, Malaysia, it is necessary to collect a variety of data from multiple channels based on the specific characteristics of the city, and conduct an overall analysis based on traffic patterns and different regions. In this process, several steps are required, and the methods used in each step are different. The following will briefly describe the methods involved in each step.

3.2 Methods for clarifying the range of selected data

Before collecting data, the selected scope needs to include important traffic arteries and important transportation hubs in Johor Bahru, such as bridges, viaducts, highways, intersections leading to large shopping malls, tourist attractions, municipal departments, etc. At the same time, it is necessary to determine the time range, such as the whole day or commuting time, the transportation modes involved, such as private cars, buses, motorcycles, bicycles, etc. Finally, weather conditions need to be included.

3.3 Data Collection Methods

Data analysis requires rigor, so the source of data is particularly important. Wrong data will have wrong consequences on the final result.

To collect traffic flow data in Johor Bahru, the following methods can be used

(a) Manual counting is performed through monitoring records on traffic lights. The advantage of this method is that the recorded data is more accurate and reliable, but it is time-consuming and labor-intensive, and it will waste human resources.

(b) Automatic counting is performed through cameras installed at key intersections. This method solves the problem of manual counting, which is labor-intensive and records quickly, but the corresponding problem is not accurate enough

(c) Counting traffic at intersections through drones or counting vehicles

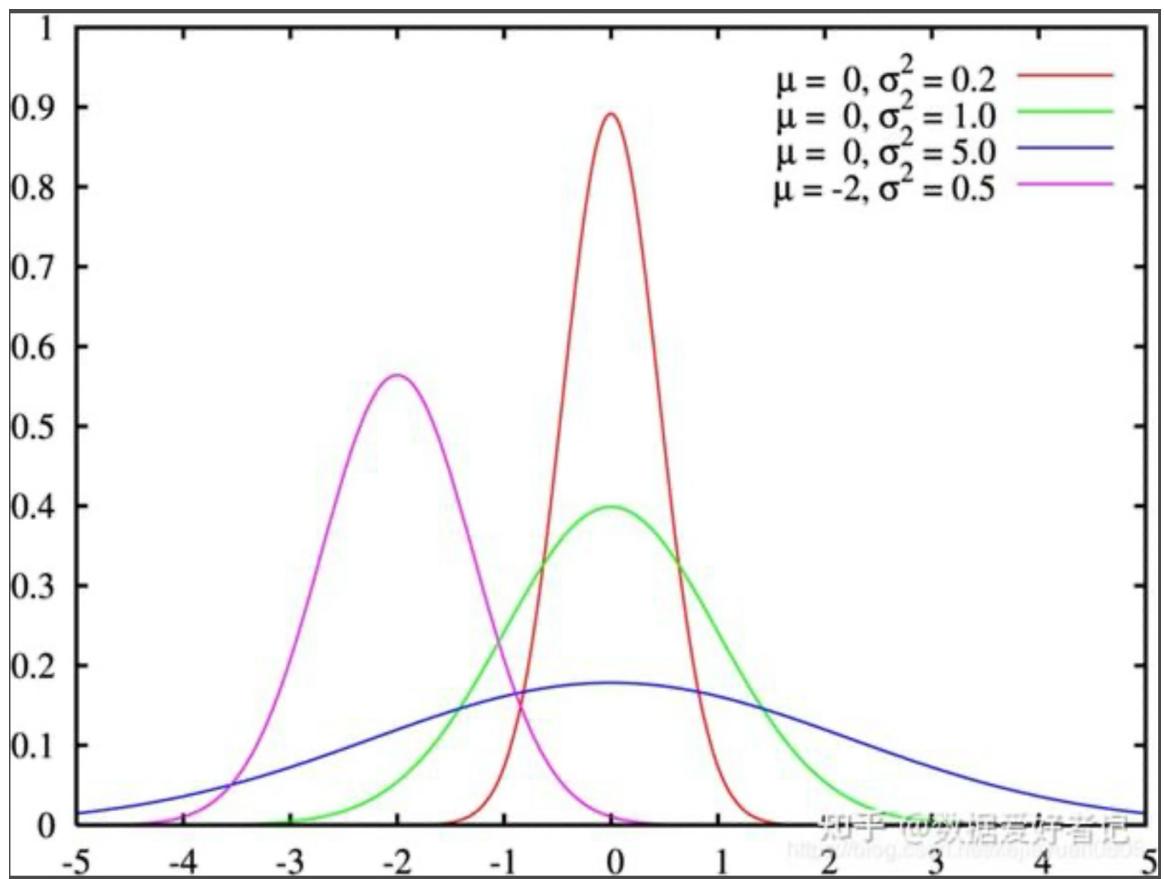
(d) Collecting data through third-party data platforms (such as Google Maps, Waze platforms)

(e) Directly collecting data from government agencies and traffic management departments

3.4 Data preprocessing methods

The collected complex information needs to be summarized and sorted first to extract useful information.

(a)Data anomalies



Outliers in data refer to values that are beyond or below the normal range. There are several common methods for dealing with outliers, using simple statistical methods, the

3σ principle, normal distribution, etc.

(b) Handling missing values in data

In the process of data collection, it is inevitable to encounter missing values. The common method for missing values is to delete them directly, use the mean, median, and median in statistics to supplement the missing values, or mark the missing values and find the corresponding values in other data to supplement them.

(c) Remove duplicate values

When collecting multiple data sets, there is a high probability of encountering duplicate data. To remove duplicate data, you can use Power BI to process the data to remove duplicate values and use Python scripts to process the data.

(d) Unified data format

In order to make the data beautiful and neat, the data format needs to be processed. Unifying the units, time format, and font size are all good methods.

3.5 Data analysis methods

3.5.1 Statistical analysis

The processed data can be analyzed using statistical methods, such as calculating the variance, mean value, and peak flow of the data. At the same time, the flow of different time periods can be analyzed for comparison.

3.5.2 Data visualization

Another common solution is data visualization, which integrates data into various pie charts, bar charts, and line charts for people to observe and record.

3.6 Model selection and construction method

For the collected traffic flow data, various models can be constructed and classified according to the available scenarios.

(a) Time series model

This model is suitable for analyzing the trend of traffic flow in the future. The

following methods can be used: ARIMA (autoregressive integrated moving average model), SARIMA (seasonal ARIMA), Exponential Smoothing (exponential smoothing method), LSTM (long short-term memory network), GRU (gated recurrent unit).

(b) Specific classification model

This model is suitable for judging whether the traffic is congested. The methods that can be used are: logistic regression, decision tree, support vector machine (SVM), gradient boosting model (XGBoost/LightGBM/CatBoost).

(c) Clustering classification model

This model is suitable for discovering traffic patterns. The methods that can be used are K-Means, DBSCAN (density clustering), and Gaussian mixture model (GMM).

3.7 Subsequent summary and collation methods

Problems encountered should be recorded to prevent the possibility of recurrence next time

CHAPTER 4

INITIAL RESULTS

4.1 Introduction

This chapter delves into the exploratory data analysis (EDA) of Johor Bahru traffic flow to understand the main data and obtain preliminary conclusions related to the research. Exploratory data analysis (EDA) is often used by data scientists to study and analyze data sets and summarize their main features based on the research process. The commonly used method is data visualization. Therefore, the data collected in this paper will be organized and analyzed using simple statistical data methods and visualization tools.

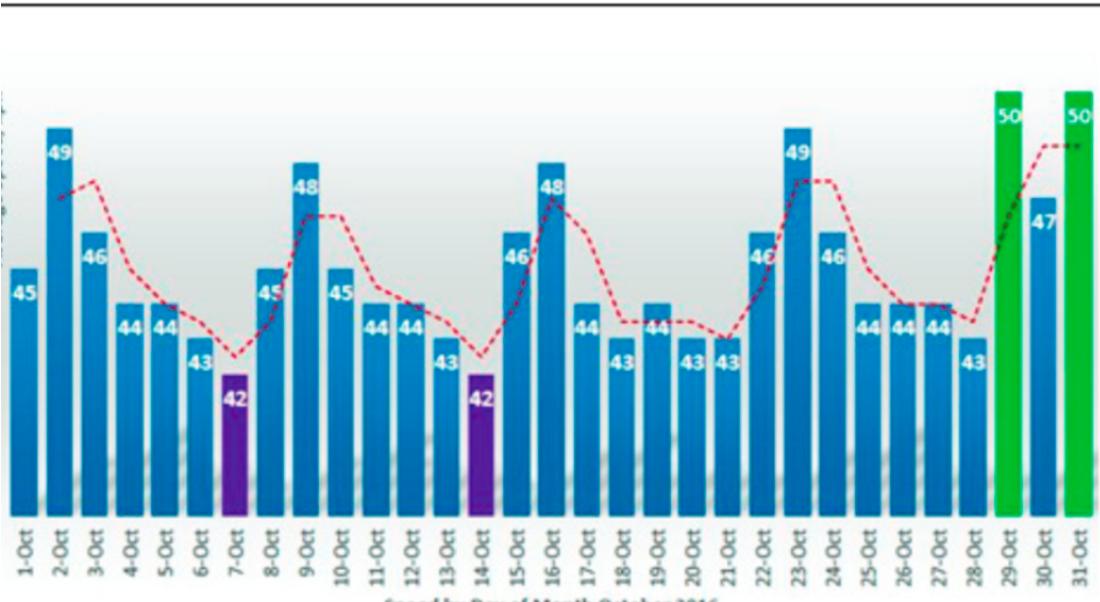
4.1 Previously used data sources and statistical methods

1. The data collected in this paper mainly come from the following sources: intelligent traffic information system, crowdsourcing-based traffic information system, social geography network, open transportation system and online car-hailing application network called Grab Taxi.

2. OSM, or Open Street Map, is an open source map application that allows third-party editing. The traffic flow data for Johor Bahru, Malaysia was exported from this website.

4.2 The following figure is a traffic analysis of the entire Johor Bahru area in September 2024

According to the data in the figure, the situation of the main roads in Johor Bahru in September is basically similar every week. The average speed of vehicles on Friday and Saturday is between 40 and 50 kilometers per hour because they are rest days, and slightly faster at other times. The lowest speed occurred on November 8, 15 and 22, 2024, which shows that the overall average speed on Fridays of each week is lower. This is because on the weekend rest day, students are off school, migrant workers return home and tourists gather, which leads to traffic congestion on Fridays, while the heavy traffic on Sundays is because it is the first day of the working day.



4.3 Traffic flow analysis of each road

The figure below shows the average speed of each street from November 24 to 30, 2024. It is observed that the traffic speed of Inner Ring Road and Jalan Bukit Cagar is slower compared to other roads.

Data	Jalan Stulang Laut	Inner Ring Road	Eastern Dispersal Link Expressway	Jalan Sultan Ibrahim	Jalan Bukit Cagar
24 NOV 2024	33	22		35	33
25 NOV 2024	33	20		30	35
26 NOV 2024	32	20		33	33
27 NOV 2024	33	17		32	35
28 NOV 2024	32	16		37	36
29 NOV 2024	30	20		35	36
30 NOV 2024	31	19		42	37

4.4 Conclusion

This work collected the average traffic speed of certain blocks in Johor Bahru in November 2024. The relevant data were collected through Grab Taxi and OSM software, sorted using excel tables, and summarized and analyzed through visualization tools.

CHAPTER 5

CONCLUSION

5.1 Summary

This study, through an in-depth analysis of traffic flow information in Johor Bahru, Malaysia, shows the causes of congestion during peak hours on major streets and the characteristics of different streets and areas. This study not only provides data support for Johor Bahru's traffic problems, but also provides a basis for urban planning.

5.2 Future Works

After completing the preliminary analysis of traffic flow in Johor Bahru, Malaysia, the following aspects can be explored in future work.

1. The data collected at this stage is limited to a few main streets and is not complete. In the future, data from more street areas can be collected to improve the reliability of the data.
2. The data processing tools currently used are not complex enough to show more in-depth information. Technical iterations can be performed later to use more advanced data analysis tools.
3. Conduct more diversified data analysis, such as adding climate factors, population factors, sound effects, etc.

5.3 References

- Papacostas, C. S., & Prevedouros, P. D. (2001). *Transportation Engineering and Planning*. Pearson Education.
- Ortúzar, J. de D., & Willumsen, L. G. (2011). *Modeling Transport*. John Wiley & Sons.
- Chin, H. C., & May, A. D. (1991). Managing urban traffic congestion. *Transportation Research Part A: Policy and Practice*, 25(5), 295-301.
- Barter, P. A. (1999). An international comparative perspective on urban transport and urban form in Pacific Asia: The challenge of rapid motorisation. *Habitat International*, 23(4), 447-457.
- Zhang, L., & Levinson, D. (2004). Road network evolution and intelligent transportation systems. *Transportation Research Part C: Emerging Technologies*, 12(2), 113-129.
- Fwa, T. F., & Goh, Y. M. (1995). Traffic flow analysis on Singapore-Malaysia link. *Journal of Transportation Engineering*, 121(2), 121-129.

Tey, L. S., & Abdul Samad, M. (2015). Challenges in managing cross-border traffic between Malaysia and Singapore. *Journal of the Eastern Asia Society for Transportation Studies*, 11, 1203-1215.

Abdullah, M., & Shariff, N. (2016). Public transport service quality in Johor Bahru, Malaysia. *International Journal of Economics and Management*, 10(Special Issue 1), 123-136.

Barter, P. A. (2011). Challenges and opportunities in Southeast Asian urban transport. *Transportation Research Board Annual Meeting Proceedings*.

Tsan-Ming, C., & Qiang, M. (2013). Data-driven approaches for traffic flow analysis. *Transportation Research Part C*, 24, 88-103.

Longley, P., Goodchild, M., Maguire, D., & Rhind, D. (2015). *Geographic Information Systems and Science*. Wiley.

Papacostas, C. S., & Prevedouros, P. D. (2001). *Transportation Engineering and Planning*. Pearson Education.

Ortúzar, J. de D., & Willumsen, L. G. (2011). *Modeling Transport*. John Wiley & Sons.

3 Chin, H. C., & May, A. D. (1991). Managing urban traffic congestion. *Transportation Research Part A: Policy and Practice*, 25(5), 295-301.

Barter, P. A. (1999). An international comparative perspective on urban transport and urban form in Pacific Asia: The challenge of rapid motorisation. *Habitat International*, 23(4), 447-457.

Zhang, L., & Levinson, D. (2004). Road network evolution and intelligent transportation systems. *Transportation Research Part C: Emerging Technologies*, 12(2), 113-129.

Fwa, T. F., & Goh, Y. M. (1995). Traffic flow analysis on Singapore-Malaysia link. *Journal of Transportation Engineering*, 121(2), 121-129.

Tey, L. S., & Abdul Samad, M. (2015). Challenges in managing cross-border traffic between Malaysia and Singapore. *Journal of the Eastern Asia Society for Transportation Studies*, 11, 1203-1215.

VIII. Abdullah, M., & Shariff, N. (2016). Public transport service quality in Johor Bahru, Malaysia. *International Journal of Economics and Management*, 10(Special Issue 1), 123-136.

Barter, P. A. (2011). Challenges and opportunities in Southeast Asian urban transport. *Transportation Research Board Annual Meeting Proceedings*.

Tsan-Ming, C., & Qiang, M. (2013). Data-driven approaches for traffic flow analysis. *Transportation Research Part C*, 24, 88-103.

Longley, P., Goodchild, M., Maguire, D., & Rhind, D. (2015). *Geographic Information Systems and Science*. Wiley.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). *Time Series Analysis: Forecasting and Control*. Wiley.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*. OTexts.

Wu, Y., & Tan, H. (2016). “Short-term traffic flow prediction with LSTM neural network.” 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC).

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Chen, T., & Guestrin, C. (2016). “XGBoost: A scalable tree boosting system.” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge

Discovery and Data Mining.

Zhang, Z., Wang, J., & Sun, H. (2020). "Traffic flow prediction using machine learning and comparative analysis." *Transportation Research Part C: Emerging Technologies*.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). "A density-based algorithm for discovering clusters in large spatial databases with noise." *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD)*.

Huang, L., & Zhang, Y. (2018). "Clustering traffic patterns for urban road networks." *Transportation Research Record: Journal of the Transportation Research Board*.

Jain, A. K. (2010). "Data clustering: 50 years beyond K-means." *Pattern Recognition Letters*.

Barceló, J. (2010). *Fundamentals of Traffic Simulation*. Springer.

Krajzewicz, D., Erdmann, J., Behrisch, M., & Bieker, L. (2012). "Recent development and applications of SUMO - Simulation of Urban Mobility." *International Journal On Advances in Systems and Measurements*.

Nagel, K., & Flötteröd, G. (2012). "Agent-based traffic assignment: Going from trips to behavioral travelers." *Proceedings of the 91st Annual Meeting of the Transportation Research Board*.