

# **COS30045 Data Visualisation**

## **Sarawak Visitor Arrivals (2017-2021)**

Github Link:

[https://github.com/ChongKaiBo/COS30045\\_Data\\_Visualisation](https://github.com/ChongKaiBo/COS30045_Data_Visualisation)

Group 1

Kai Bo CHONG  
(102774161)

Chun Wei PANG  
(102773142)

Year 2 Semester 2

# **Table of Contents**

<b>1. Introduction</b>	<b>3</b>
1.1 Background and Motivation	3
1.2 Visualisation Purpose	3
1.3 Project Schedule	3
<b>2. Data</b>	<b>4</b>
2.1 Data Source	4
2.2 Data Processing	4
<b>3. Requirements</b>	<b>4</b>
3.1 Must Have Features	4
3.2 Optional Features	4
<b>4. Visualisation Design</b>	<b>4</b>
<b>5. Validation</b>	<b>7</b>
<b>6. Conclusion</b>	<b>7</b>
<b>Reference</b>	<b>7</b>

# 1. Introduction

## 1.1 Background and Motivation

First of all, this visualisation will be very beneficial for tourism officials and planners as they would be highly interested in the Sarawak visitors' related data. This is because these data can directly impact their decisions and strategies. They might want to analyse visitor demographics, track seasonal variations, identify popular tourist destinations, and assess the economic impact of tourism on the region.

Secondly, tourism businesses for example hotels, restaurants and tour operators as they would want to understand visitor patterns and preferences to tailor their services. They may want to analyse the geographic origin of visitors, peak booking times, and types of activities that tourists engage in to optimise their offerings and generate the most incomes.

Lastly, Sarawak residents will be interested too as they might want to stay informed about the number of tourists in their area and the potential impact on their community. They are also interested in understanding how tourism affects their daily lives, such as traffic patterns and demand for local services. Thus it is vital as this will eventually increase the quality and benefit of their living.

## 1.2 Visualisation Purpose

The purpose of this visualisation is to provide a more informative and effective viewing experience to the users by not only showing them all kinds of numbers. The data is then processed into various types of visualisation for instance bar chart, scatter plot, line chart and so on hence the users can understand more about the summarised data rather than just some messy tables. Besides, users may observe the trend based on the data visualisation provided and use this information to gain benefits for their living.

## 1.3 Project Schedule

A meeting will be held every week to keep track of the members' progress and also the tasks that need to be completed on that particular week.

Week 3 - Completed the Part 1 of the project (Introduction)

Week 4: Done Part 2 of the report (Data) and search for reference.

Week 5: Start Part 3 (Requirement).

Week 6-7: Working on Part 3 (Requirement) and Part 4 (Visualization Design)

Week 8: Done Part 3 (Requirement) and Part 4 (Visualization Design)

Week 9-10: Start coding.

Week 11-12:

## 2. Data

### 2.1 Data Source

The data we are harvesting is from a portal of Sarawak (Link: <https://data.sarawak.gov.my/home/data/search/?group=travel-tourism>). The data inside recorded the previous tourist arrival to Sarawak. The data from this website is presented in tabular form. The data type of this table is quantitative. This dataset consists of the following attributes:

Attributes	Description	Type
Citizenship	This attribute contains the name of the countries of the visitors who have visited Sarawak	Categorical (String)
Months	This attribute records the number of visitors of a country in that particular month.	Ordinal (Integer)
Grand Total (Country)	This attribute calculates the grand total of visitors from a specific country in that whole year.	Ordinal (Integer)
Grand Total (Monthly)	This attribute displays the total number of the visitors who have visited Sarawak in a specific month	Ordinal (Integer)
Total Foreigner	This attribute records the total number of visitors from foreign countries.	Ordinal (Integer)
Total Domestic	This attribute records the total number of visitors from domestic countries.	Ordinal (Integer)

### 2.2 Data Processing

Since the data sets used are simplified enough and no missing value is found too. Hence we decided not to do the substantial data cleanup. The quantities that we plan to derive from the data will be the Grand Total and the Months to calculate the mean of arrival. Data processing will be implemented by using RStudio to calculate mean, median, missing value, skewness, and the others. The variable that will be derived will be the mean of arrival. The calculation of the mean of arrival will be using the

Grand Total divided by the total number of months of a particular citizenship.

## 3. Requirements

### 3.1 Must Have Features

The features below are the some of the features in this project must have:

- Users can filter the data based on years.
- Users can interact with the visualisation which when they hover their cursor towards any part of the map, the name of the country and the number of visitors will be shown on the top of the country location.
- Users are allowed to view the number of grand total of visitor arrivals.
- Users are able to view the visitors from a country which will be presented in the form of a line chart.

### 3.2 Optional Features

The following features are the optional features that will be implemented to make the visualisation more useful:

- A button that allows the user to toggle the data of comparison between the foreign and domestic visitors.
- Users are able to use a feature that allows them to compare the number of visitors from two or more countries.

## 4. Visualisation Design

### World Map (Index page)



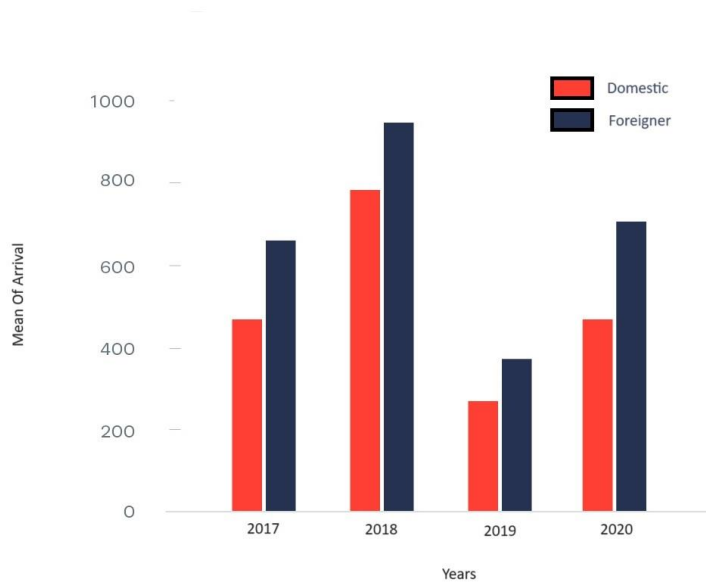
For this visualisation, it will show the grand total number of visitors arriving at Sarawak in the year 2021 by default. The users can change the year they want to view manually. By presenting the number of visitors on the world map can let the user easily visualise the visitors from whichever country the user hover it. However, this might look a bit messy when the number of visitors is shown in the smaller countries.

### Line chart



A line graph will be shown on the screen when the users click on the particular countries. The line graph will be displaying the trend of number of visitors based on the filter they chose either yearly or monthly. With these line graphs, it allows the users to better visualise the trend of visitors for the past few months or years.

### Bar Chart



The bar chart will be showing the comparison between the yearly grand total of domestic and foreign visitors to Sarawak. However, this is an optional feature that we might have in our project.

### Final Design

2021



This final design is which that when the user hovers their mouse to a country, it will shows the country name and its grand total of visitors in that particular year. Besides that, there will be golden curve lines connecting from the hovered country to Sarawak for specifying which country the visitors are from. Other then, if the user wants to view the trend of visitors from this country from month to month or year to year, they can click the country. From this visualisation, the user not only can view the grand total of visitors, but also the monthly visitor from a particular country.

5.Validation

6.Conclusion

Reference