Dataviz2

# 1 Overview

# 2 Critique

## 2.1 Clarity

1. Titles: The title of the dashboard is lacking information for the reader to comprehend the chart sufficiently.

Even though it can be inferred that this is an origin-destination analysis, it is unclear to the audience what is the mode of transportation used for measuring the trips, what is the time period of the data used, and at what level of region/zone is the data being compared at. The information should be mentioned clearly in the title to allow the viewer to understand what data exactly he is looking at. In this case, it would be the inter and intra zonal bus trips in Singapore, compared at subzone level, using data in January 2022.

1. Abbreviations used in Adjacency Matrix Axes, Tooltips and Origin/Destination Filters

The ‘sz’ abbreviation is used in the adjacency matrix and filter to represent the subzone of the location labelled or chosen. However as ‘sz’ was not explained or mentioned in the dashboard, it will be unclear to the reader on what is the exact meaning of ‘sz’, especially someone who does not have access or knowledge of the data used.

1. Y-axis on the bar chart

As the x and y axes of the bar charts are not labelled, it is not clear that the bars are measuring the total trips for each hour across the day until the user hovers over the bars and read the tooltip. It may be intuitive from the title ‘Trips generated…’ or ‘Trips attracted…’ but it is still not as obvious as labelling on the axis.

1. Weekend adjacency matrix inaccurate label

Even though the holidays in January are on a weekend, there is no data impact to the visualisation. However, just to be accurate, 'Weekend/Holidays adjacency matrix' should be used instead. In the event a holiday falls during a weekday, the viewer may think that the data is stored under the 'Weekday adajacency matrix', which will also inaccurately skew the data as holidays are non-working days and the behaviour of passengers should follow that of weekend instead.

1. Null values in the filters

The origin and destination still consist of null values, showing that the data may not have been cleaned up before plotting the visualisation. This means that there may be data that is invalid or unaccounted for in the visualisation, and hence, may not be an accurate representation.

1. Lack of Source and Caption

There is no recognition of the data source that enable the data to be verified or any insight to highlight significant findings for the visualisation.

## 2.2 Aesthetic

1. Adjacency Matrix for weekday and weekends/holidays height not the same

The size of the weekday adjacency matrix is shorter than the weekday adjacency matrix even though they should have the same number of rows and columns. Reader may misunderstand and think the weekend has a smaller number of origins. The charts should be resized to have equal heights.

1. Bar Charts Y-axis header truncated

The y-axis headers of the bar charts are truncated for “weekends/holidays”. The audience may not be able to understand the full picture if the header is truncated. Furthermore, a straight header occupies a lot of real estate space in the dashboard. To prevent truncation and optimise the space, the header can be rotated 90° instead.

1. Radio buttons for origin and destination selection

As the origin and destination consists of over 300 choices, by using radio buttons, the length of the filter pane will not be sufficient to contain all the choices of the origin and destination. Furthermore, it makes the pane looks cluttered and messy. This can be improved by using a dropdown list instead.

1. Tooltips and Legend naming

For adjacency matrix tooltips “% of Total Total Trips along Table (Across)”. The title of the tooltip can be improved to “% of Total Trips” instead as the rest may be redundant. For adjacency matrix legend “% of Total Total Trips”. The title of the legend can be improved to “% of Total Trips” instead as the additional “Total” is redundant.

# 3 Proposed Design

## 3.1 Clarity

1. Present complete data by accounting for null values in the filters.

The data cleaning step will enable the null values to be accounted for via their bus stop code, allowing completeness and more accuracy in the data presented in the visualisation.

1. Title, caption and source.

The title of the dashboard will incorporate what the charts are about incluging the mode of transportation used for measuring the trips, what is the time period of the data used, and at what level of region/zone is the data being compared at. This will allow the viewer to understand what data exactly he is looking at. Source will also be included to recognise the data source and enable the data to be verified.

1. Using geospatial to draw clarity to the locations of the subzone and its profile.

With the chloropleth map, the audience can better appreciate the location of the subzones and the summary of the percentage of trips originating or going to a certain subzone.

1. Tooltips and labelling. With cluttered graphs like the adjacency matrix, it will be difficult to read the data off the axis. Hence, the viewer will have to depend on tooltips to derive the information required. The interactive information of subzone origin, destination, number of trips, percentage of total trips will be clearly provided in the tooltips and labels, and unclear abbreviations such as ‘sz’ will be avoided.

## 3.2 Aesthetic

1. Consistent size of same charts.

The sizing of the same type of charts will be adjusted to ensure it is consistent throughout the dashboard.

1. Rotate axis to view headers clearly

The axes and headers will be rotated 90° to ensure that the words will not be truncated, and the reader can gain the full understanding of the measurement used.

1. Dropdown list for selection instead

As there are over 300 origin and destination to be selected, to prevent truncation from the radio button list, a dropdown list will be used instead. The user can also easily scroll through the dropdown list to select the subzone for viewing.

1. Remove redundant characters in tooltips and labelling. Repeated or unclear words in the labels/tooltips should be removed. For example, edit to “% of Total Trips” instead of “% of Total Total Trips”.

# 4 Visualisation

# 5 Step-by-Step Preparation

## 5.1 Data Preparation

|  |  |  |
| --- | --- | --- |
|  | Description | Photo |
| 1 | Load the origin-destination file into Tableau Prep Builder. Drag the worksheet into the main pane. | prep1 |
| 2 | Click on the ‘+’ button and select ‘Clean Step’ to clean the data. | Prep2 |
| 3 | Through exploring the data, you will realise some of the origin and destination are null, despite having a bus stop code. | Prep3 |
| 4 | To extract bus stop code of null subzones, PivotTable in Excel was used. | Prep4 |
| 5 | Using Pivot Table, filter the origin\_sz field for blank inputs, to obtain the origin bus stop codes with null values. | Prep5 |
| 6 | Using Pivot Table, filter the destination\_sz field for blank inputs, to obtain the destination bus stop codes with null values. | Prep6 |
| 7 | Access ura.gov.sg/maps and key in the bus stop code on the search bar. Scroll to the end of ‘Land Information’ tab and click on ‘See Other Services’. The planning subzone and area can be obtained there. | Prep7 |
| 8 | Repeat for all the bus stop codes with null values to obtain the subzone, planning area and planning region. | Prep8 |
| 9 | Create a custom calculation field to input the origin subzone names for origin bus stop codes. | Prep9 |
| 10 | Input the code to use existing ‘origin\_sz’ if not null. For null values, the subzone names will be input based on the origin bus stop code. | Prep10 |
| 11 | Repeat for destination subzones. Create a custom calculation field to input the destination subzone names for destination bus stop codes. | Prep11 |
| 12 | Input the code to use existing ‘destination\_sz’ if not null. For null values, the subzone names will be input based on the destination bus stop code. | Prep12 |
| 13 | The subzone names have been successfully input without null values for both origin and destinations. | Prep13 |
| 14 | Export the origin-destination data to tableau desktop. | Prep14 |
| 15 | Rename the flow from Tableau Prep Builder to ‘Origin-Destination Clean’. Load the ‘MP14\_SUBZONE\_WEB\_PL.shp’ file into Tableau Desktop. Join the shape file with the origin destination file where ‘DESTINATION’ = ‘Subzone N’. | Prep15 |
| 16 | Create a new path to join the shape file with the origin destination file where ‘ORIGIN’ = ‘SUBZONE N’. | Prep16 |

## 5.2 Visualisation Preparation

1. Chloropleth

|  |  |  |
| --- | --- | --- |
|  | Description | Photo |
| 1 | To create the map, drag ‘Geometry’ and ‘Subzone N’ to ‘Detail’. The map will be automatically created. Next drag ‘TOTAL\_TRIPS’ to ‘Color’ and the gradient of sum of trips to each location will the form in the chloropleth. | Chloro1 |
| 2 | Drag ‘ORIGIN’, ‘DAY\_TYPE’, ‘TIME\_PER\_HOUR’ to filter. | Chloro2 |
| 3 | For ‘ORIGIN’ and ‘DAY\_TYPE’ character data, select all inputs to be filtered. | Chloro3 |
| 4 | For ‘TIME\_PER\_HOUR’ numeric data, select ‘All values’ and ‘Range of values’ to be filtered. | Chloro4 |
| 5 | For ‘ORIGIN’ filter, select to a ‘Single Value (dropdown)’ instead as there are over 300 choices to choose from. | Chloro5 |
| 6 | For ‘DAY\_TYPE’ filter, select to a ‘Single Value (list)’ instead as there are only 3 choices. | Chloro6 |
| 7 | For SUM(TOTAL\_TRIPS) color, change the calculation type to ‘Percent of Total’ instead to show the percentage of trips to each location. | Chloro7 |
| 8 | Add ‘TOTAL\_TRIPS’ to ‘Tooltip’, and edit the Tooltip as follows. | Chloro8 |
| 9 | Rename the title of the chart using dynamic input from the ‘ORIGIN’ filter. | Chloro9 |
| 10 | Duplicate the sheet of the chloropleth with ‘ORIGIN’ filter, to create another chloropleth with ‘DESTINATION’ filter. This is to minimise effort on formatting of the chloropleth. | Chloro10 |
| 11 | Remove the ‘Geometry’, ‘Subzone N’ detail fields, and ‘ORIGIN’ filter fields by dragging it back into the Tables pane. | Chloro11 |
| 12 | Drag ‘GEOMETRY’ and ‘SUBZONE N’ from the 2nd shapefile to the ‘Detail’ section. | Chloro12 |
| 13 | Drag ‘DESTINATION’ field from the origin-destination data to ‘Filter’. Select all choices. | Chloro13 |
| 14 | For ‘DESTINATION’ filter, select to a ‘Single Value (dropdown)’ instead as there are over 300 choices to choose from. | Chloro14 |
| 15 | Edit the tooltip to show the subzone as origin instead. | Chloro15 |
| 16 | Rename the title of the chart using dynamic input from the ‘DESTINATION’ filter. | Chloro16 |

1. Bar Chart

|  |  |  |
| --- | --- | --- |
|  | Description | Photo |
| 1 | To create the bar chart, drag ‘DAY\_TYPE’ and ‘TIME\_PER\_HOUR’ to Columns, and drag ‘TOTAL\_TRIPS’ to row. | Bar1 |
| 2 | Add an ‘ORIGIN’ filter by dragging it to the filter pane and select all inputs. | Bar2 |
| 3 | Rename the title of the chart using dynamic input from the ‘ORIGIN filter’ | Bar3 |
| 4 | Drag ‘ORIGIN’ to the tooltip and click on the icon to edit the tooltip. | Bar4 |
| 5 | Click on the color icon to change the color of the bar to match the chloropleth. | Bar5 |
| 6 | Click on the size icon to fix the width of the bars and set alignment to center. | Bar6 |
| 7 | Double click on the y-axis to edit the y-axis name. | Bar7 |
| 8 | Double click on the x-axis to edit the x-axis scale and name. The scale of 4-24 was selected because the earliest bus rides start at about 5am daily. | Bar8 |
| 9 | Hide the labels for columns as it can be inferred, and to maximise graph space. | Bar9 |
| 10 | Duplicate the sheet of the bar chart with ‘ORIGIN’ filter, to create another bar chart with ‘DESTINATION’ filter. This is to minimise effort on formatting of charts. | Bar10 |
| 11 | Remove the ‘ORIGIN’ filter and tooltip by dragging it back to the table pane. | Bar11 |
| 12 | Drag ‘DESTINATION’ field from the origin-destination data to ‘Filter’. Select all choices. | Bar12 |
| 13 | Rename the title of the chart using dynamic input from the ‘DESTINATION’ filter. | Bar13 |
| 14 | Drag ‘DESTINATION’ field to the tooltip and edit tooltip to show the destination instead. | Bar14 |

1. Adjacency Matrix

|  |  |  |
| --- | --- | --- |
|  | Description | Photo |
| 1 | To create the adjacency matrix, drag ‘DESTINATION’ to columns and drag ‘DAY\_TYPE’ and ‘ORIGIN’ to Rows in the same order. Next, drag ‘TOTAL\_TRIPS’ to color icon. | Adj1 |
| 2 | To view the full matrix, select ‘Entire View’ on the top of the column pane. | Adj2 |
| 3 | For SUM(TOTAL\_TRIPS) color, change the calculation type to ‘Percent of Total’ instead to show the percentage of trips to each location. | Adj3 |
| 4 | Drag ‘DAY\_TYPE’ and ‘TIME\_PER\_HOUR’ to filter. | Adj4 |
| 5 | For ‘ORIGIN’ and ‘DAY\_TYPE’ character data, select all inputs to be filtered. | Adj5 |
| 6 | For ‘TIME\_PER\_HOUR’ numeric data, select ‘All values’ and ‘Range of values’ to be filtered. | Adj6 |
| 7 | For y-axis origin labels, right-click and select ‘rotate label’ to remove the names of the subzones. | Adj7 |
| 8 | Next, do the same to the day type label, by right-clicking and selecting ‘rotate label’ to optimise the space of the graph and view the day type text fully. | Adj8 |
| 9 | Rename the sheet to ‘Adjacency Matrix’ | Adj9 |
| 10 | Drag ‘TOTAL\_TRIPS’ to tooltip icon and edit the tooltip as follows. | Adj10 |

1. Coordinating Filters

|  |  |  |
| --- | --- | --- |
|  | Description | Photo |
| 1 | To align the ‘ORIGIN’ filter of the chloropleth and bar chart, click on the ‘ORIGIN’ filter and apply it to ‘selected worksheets’. | Fil1 |
| 2 | Select the corresponding bar chart worksheet to apply the ‘ORIGIN’ filter to. | Fil2 |
| 3 | To align the ‘DESTINATION’ filter of the chloropleth and bar chart, click on the ‘DESTINATION’ filter and apply it to ‘selected worksheets’. | Fil3 |
| 4 | Select the corresponding bar chart worksheet to apply the ‘DESTINATION’ filter to. | Fil4 |
| 5 | For the ‘DAY\_TYPE’ filter, click on the filter and apply to ‘All Using This Data Source’ | Fil5 |
| 6 | For the ‘TIME\_PER\_HOUR’ filter, click on the filter and apply to ‘All Using This Data Source’ | Fil6 |

1. Dashboard Building

|  |  |  |
| --- | --- | --- |
|  | Description | Photo |
| 1 | Open a new dashboard, and for the size of the dashboard select ‘Automatic’ so it will adjust based on the viewer’s screen size | db1 |
| 2 | Select charts to be displayed on the dashboard and drag the charts from the sheets pane to desired location on the dashboard. | Db2 |
| 3 | From the dashboard menu, select ‘Show Title’ to insert dashboard title. | Db3 |
| 4 | Edit the dashboard title to include a caption, description of the chart and the source. | Db4 |
| 5 | Edit the names of the filters by clicking on the filter and select ‘Edit Title’. | Db5 |
| 6 | After editing the filter names, the dashboard is ready to be exported to Tableau Public. | Db6 |

# 6 Observations

1. Intra-zonal and industrial areas public bus trips make up majority of passenger volume.

Intra-zonal travel by public bus seems to make up a higher percentage of total bus rides. This can be seen from the prominent darker colour diagonal line intersecting the adjacency matrix for both weekends and weekdays. This highlights the intra-zonal travel within the same subzone e.g. 51% of bus trips that originated in Taman Jurong ended in Taman Jurong, 44% of bus trips that were initiated in Tampines East ended in Tampines East and 32% of bus trips that started in Pasir Panjang 1 ended in Pasir Panjang 1. These trips are likely made via the feeder bus services which offer transfers from MRT stations and bus interchanges to surrounding housing estates and industrial areas. It also connects the passengers to various neighbourhoods within the same estate. This percentage is similar for both weekdays and weekends as passengers travel to work on weekday, and for recreational activities on weekends.

![]obs1

There are also other inter-zonal travel points on the matrix that had darker colour spots, especially industrial areas such as Joo Koon, Tuas, Peng Siang, Changi West e.g., Joo Koon-Gul Basin over 90% and Tuas-Tuas View Extension over 88%, Choa Chu Kang-Peng Siang over 50%, Tampines East-Changi West over 80%. These places are typically large industrial areas that do not have other transport infrastructure other than roads. Hence, the bus trips to these places are likely made with bus services that are designed to pick up passengers in a central location (i.e. MRT) and transfer than to their workplaces. The percentage is similar for both weekdays and weekends as workers in these industries may need to conduct 24-hour shift work, 7 days a week.

1. The bus traffic peak times and volumes differ on weekdays and weekends

From the distribution bar charts, we can identify the peak periods of bus trips on weekdays. For example in the chart below for Tampines East subzone, one of the most populous subzone in Singapore, we can identify the morning peak, 7-9am and evening peak, 5-7pm. This is consistent for both to and fro trips throughout the day as the Tampines Regional Hub (Tampines1, Tampines Mall, Century Square) are in this subzone. From the chart of distribution in Yunnan below, we can see the morning peak 7-9am for trips originating Yunnan and evening 5-7pm for trips heading to Yunnan. Both morning and evening peaks are not present in both the to and fro graphs. The chart is different from the Tampines East one because Yunnan is a dominantly residential estate, with its closest town centre at Boon Lay Jurong Point. Comparing with weekday trips, weekend trips on the other hand are more evenly distributed throughout the day.

Obs2, obs3

1. Majority of passengers travel to or from subzones within close proximity to their origin or destination respectively.

From the chloropleth below, we can tell that the subzones closer to the origin or destination have a higher colour intensity, indicating that most passengers use public buses to travel within locations of closer proximity. This is because buses are able to penetrate through the road system within a neighbourhood that MRTs are unable to access. To travel between zones with larger proximity, passengers may prefer to use more efficient or convenient forms of transport like the MRT instead, which is not subjected to road traffic conditions.

One exception to the observation was spotted in Punggol Field subzone, where many passengers also cross region from North-East to East to travel to locations like Tampines and Changi. This may be due to the convenience of the Tampines Expressway (TPE) which conveniently connects Punggol to Tampines within 2 bus stops, and less than 10 mins drive away. Hence, even though the subzones are not in close proximity, but the transport network enables efficiency for travelling between these subzones.