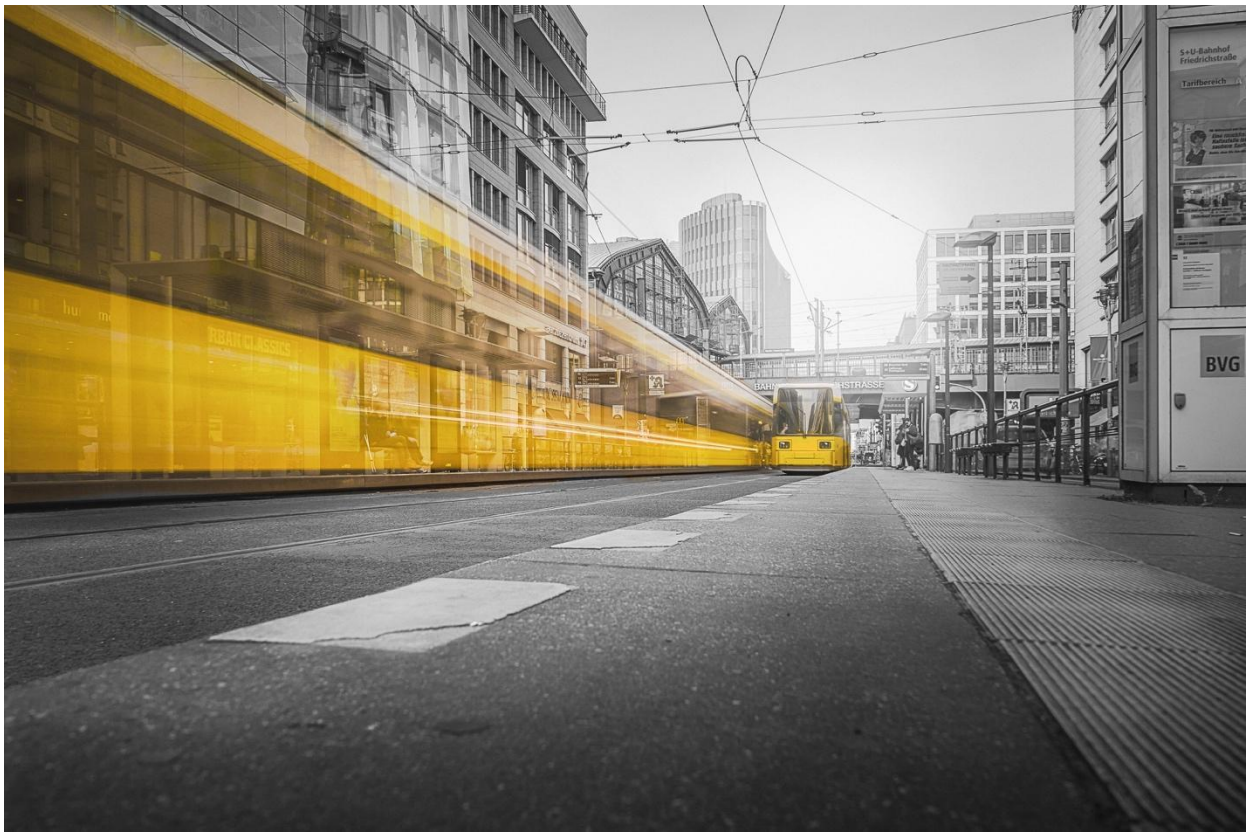

(Pixabay, 2017)

Data Analysis and Decision Support

: How to improve tram passenger experience in Melbourne CBD

—

Miji Kim

Executive Summary

Operating the largest tram network in the world, Yarra Trams has been striving to provide a better experience to passengers travelling on a tram. Despite its effort, overcrowding is still a major problem with Yarra Trams operating exceeding their capacity in the CBD. It is considered a severe factor as it can not only be fatal to passenger safety but also contribute to Yarra Trams failing to meet its reliability and punctuality. All of these factors are imposing Yarra Trams a significant amount of financial burden as compensation can be made by eligible passengers accordingly. In this context, this report would seek a possible solution for Yarra Trams to provide a better passenger experience and to minimize the overcrowding issue in the Melbourne CBD by changing tram classes for certain routes.

A description of the problem statement and dataset used for analysis is provided at the beginning of the report. Then, quantitative and qualitative analyses regarding the problem along with the methodology applied are shown throughout the report.

Table of Contents

1. Introduction

In April 2019, it was reported that Yarra Trams has recorded the worst performance since the new contract was initiated; as a result, passengers who used the Myki pass were able to apply for compensation, which meant free fares (Rooney, 2019). While the target punctuality was 82%, Yarra Trams was only able to achieve around 78%. Being the major culprit of such failure, overcrowding on trams continues to be the problem Yarra Trams has been facing. Overcrowding is an imposing great burden to Yarra Trams for several reasons. First, overcrowding contributes to passengers slipping, tripping, or falling when they are on the tram or when they try to get on/get off the tram. Such would increase the financial burden for Yarra Trams, as Yarra Trams are responsible for compensating them. In fact, according to research, the hospital-related cost arising from tram-related injuries in Victoria from 2011 to 2013 was around $900,000 (Victorian Injury Surveillance Unit (VISU), 2017). Another reason is that when the trams are overcrowded, they get delayed and this delay causes other trams being run late on the same route.

When looking into overcrowding in the evening peak, it was revealed that trams travelling along Collins Street in the Melbourne CBD were "the only services found to breach load standards", averaging around 105% (Carey and Butt, 2013). In fact, the intersection of Collins St and Swanston St is considered to be one of the busiest tram intersections in the world (Laskle, 2016). Moreover, one of the most crowded tram routes, 48, is also on Collins St (Carey and Butt, 2013).

Upon analysing pedestrian traffic around the Melbourne CBD, the top five sensor locations having the highest average traffic and the highest traffic flow in total from 2009 to 2014 (excluding 2011) are Town Hall West, Melbourne Central, Bourke Street Mall (South), Bourke Street Mall (North), and Flinders St Station Underpass (See Appendix A Figure A4-1).

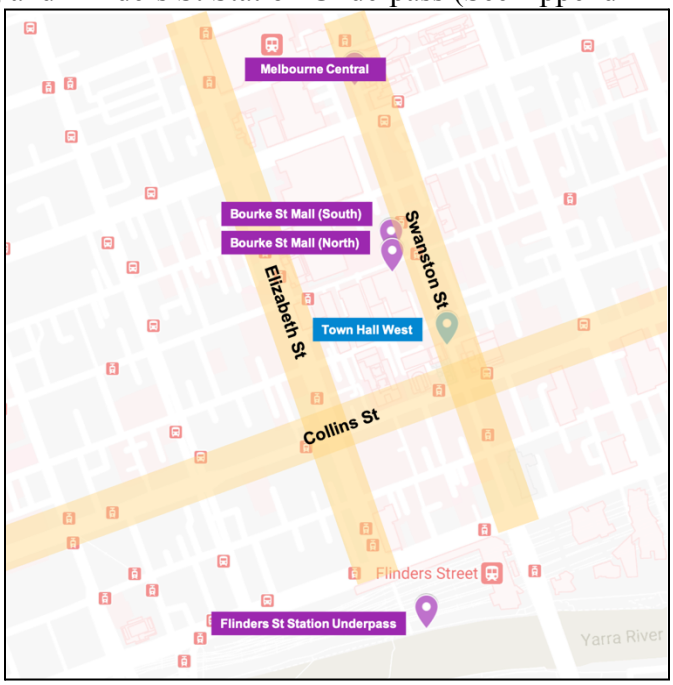

Figure 2. Sensor locations of the top five average traffic (Google, 2019) (For a broader map, see Appendix A Figure A1.)

Not surprisingly, the top four locations for pedestrian traffic are surrounded by Elizabeth St, Collins St, and Swanston St, and the fifth location is still located near Collins St, where Elizabeth St and Swanston St end (See Figure 2). The location which had the highest average and total traffic flow was Town Hall West and it is important to note that this is the closest sensor location to Collins St. In this context, the pedestrian traffic flow captured at Town Hall West is analysed and forecasted throughout the report in order to provide a solution to overcrowding issue.

2. Description of the Dataset

2.1. Data Format Issues

Any white spaces and/or parenthesis within each column name in the original dataset were removed to facilitate the analysis. For eighteen locations, data types of some columns were converted from character to integers, so as to get the mean value. Likewise, to get a mean value, data type for some columns were converted from character to integer. Moreover, date format was converted for each dataset to have the same format.

2.2. Missing Data and Removal of Fields

Looking at the dataset of eighteen locations, it is noted that data for some locations that were to be analyzed were missing. Initially, it was important to identify the average and total pedestrian traffic among eighteen locations. As the probability of missing variable was the same for all observations, list wise deletion was used; this means the whole row of observations where the data is missing (Mekala, 2018). Therefore, if data were missing for any of eighteen locations, such period was not taken into consideration for initial analysis. (see Appendix A Figure A2 for the list of periods and related location missing data). In addition, a number of location fields were removed from certain data frames as the data for only eighteen locations were needed for the initial analysis.

When it comes to the data set of Town Hall West, the data was missing for seventeen days, 6th of October, from the 13th to 14th of December in 2013, from the 24th of July to the 31st of July and from the 1st of August to 6th of August in 2014 (see Appendix A Figure A2). Among these days, only the 6th of October was NA value; others were 0. Therefore, first, the NA value was interpolated (See Appendix B Figures B1 and B2. Then, other missing values such as data from the 13th to 14th of December in 2013 were dealt with mean imputation, similar case imputation in particular; this means that missing values (0s) were substituted with estimated ones based on the month (December = 1738, July = 1407, August = 1317) (See Appendix B Figure B2).

3. The Quantitative and Qualitative Analysis

3.1. Analysis Overview

Pedestrian traffic captured at Town Hall West was analyzed and forecasting for five and ten years was based on such data. After analyzing the past data and forecasted data, a certain trend was found; forecasted data for ten years suggests that there will be constant pedestrian traffic flowing around Town Hall West (See Figure 14). The forecasted trend is closely reflecting

the current data as figure 10 suggest; December will still be the month having the greatest traffic among the twelve months. Figure 14 also implies a similar trend between months. This implies that as Town Hall West is just around Collins St and as it will still have a great traffic volume in the future; a solution to the overcrowding issue of trams should be first focused on routes passing through Collins St.

According to the walking plan proposed by City of Melbourne (2014), more than 50% of tram stops in CBD which were forecasted to be over capacity by 2030 are located on Collins St. Currently, according to the Melbourne tram network, there are four routes on Collins St: 11, 12, 48, 109 (Public Transport Victoria, 2017). Among those routes, 11 were partly using E class with the capacity of 64 seated and 146 standing passengers (Johnstone, 2019). However, 12, 48, 109 are still using A class trams which is the smallest tram having the capacity of 40 seated and 65 standing passengers. If routes 12, 48, 109 were also using E class trams, there could have been less overcrowding around tram stops and on trams as they could carry more passengers at once.

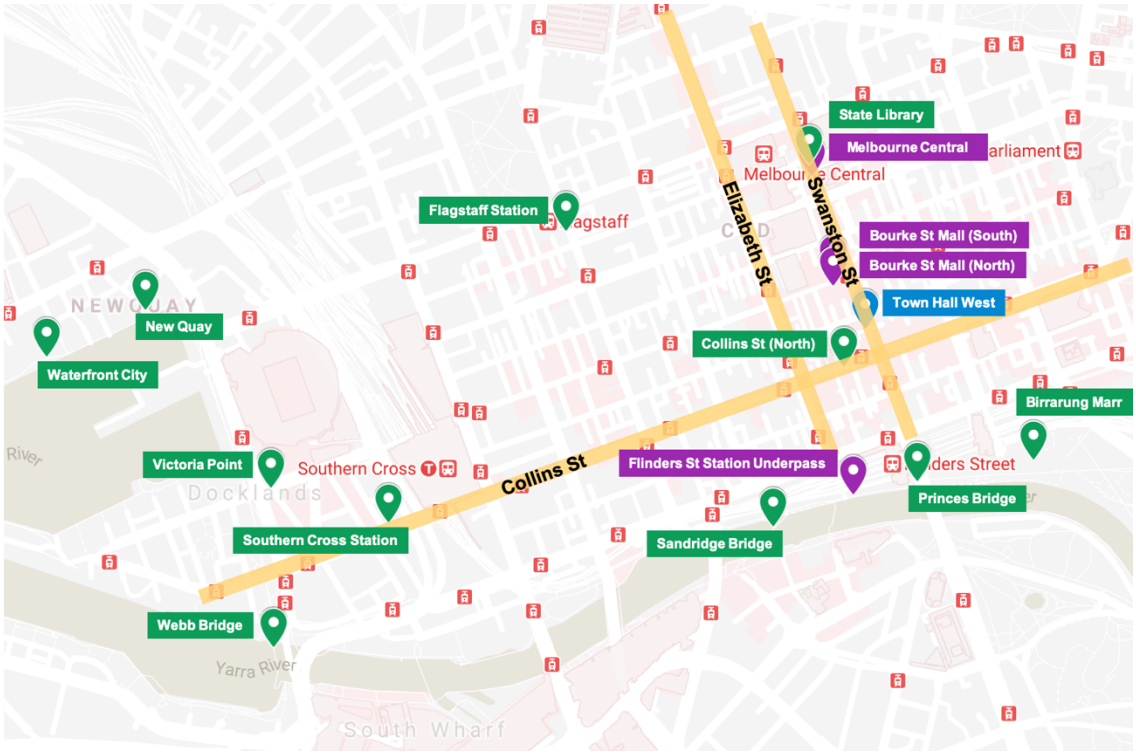

Figure 3. Locations in green refer to those not in the top five list based on traffic

3.2. Summary Statistics

Past data for pedestrian traffic at Town Hall West is analyzed; some instances are shown as the following figures (See Appendix B for more instances). Throughout the years, it was noted that 2012 was irregularly high at Town Hall West (See Figure 4) and when it comes to months, it is highlighted that December seems to be the month having the highest traffic throughout five years (See Figure 5 and Figure 6). In regard to the day, as figure 7 suggests, Friday is the day where the most pedestrian traffic occurred. Moreover, the peak hour at this location seems to be 1 pm, which is illustrated in figure 8 and figure 9.

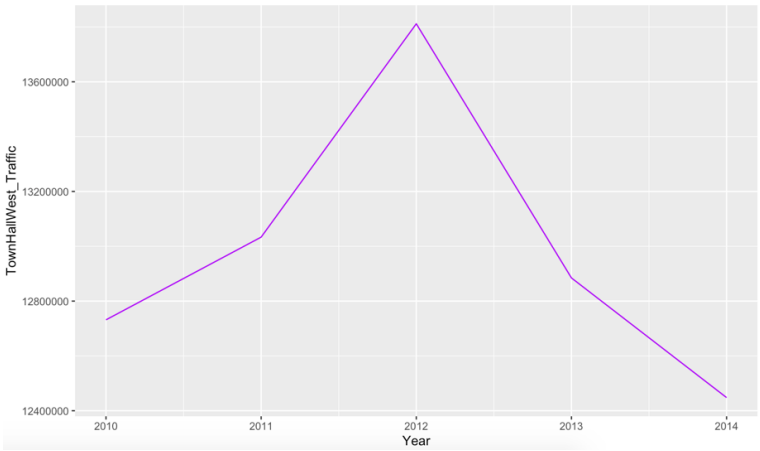

Figure 4. Pedestrian traffic over five years captured at Town Hall West

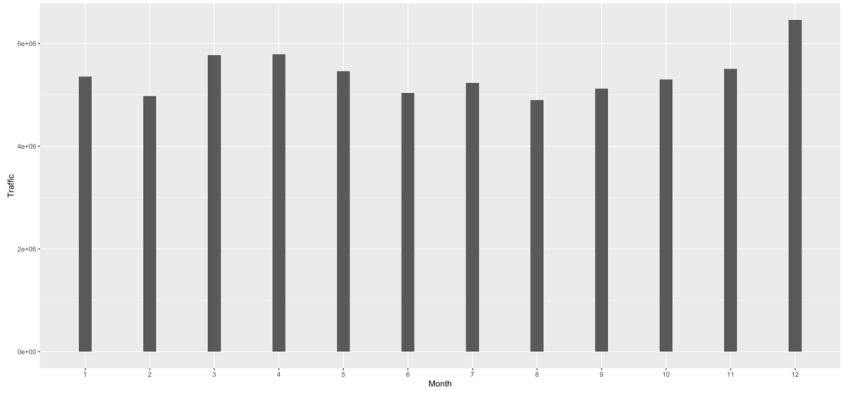

Figure 5. Pedestrian traffic over five years is shown per month at Town Hall West

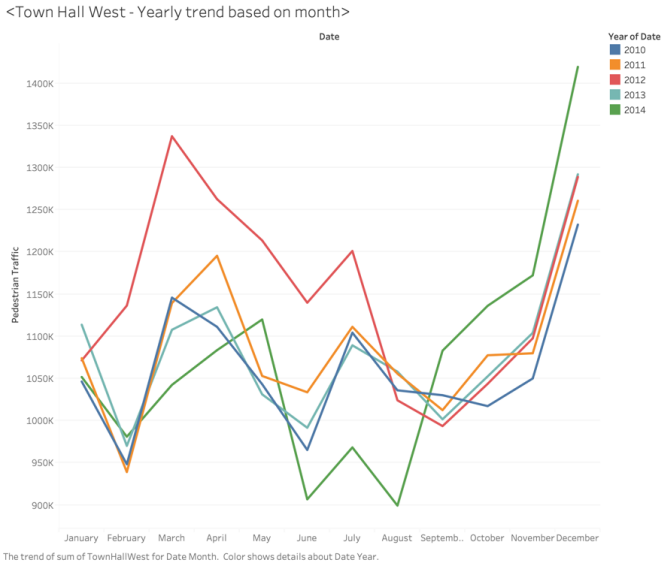

Figure 6. Pedestrian traffic over five years based on month at Town Hall West

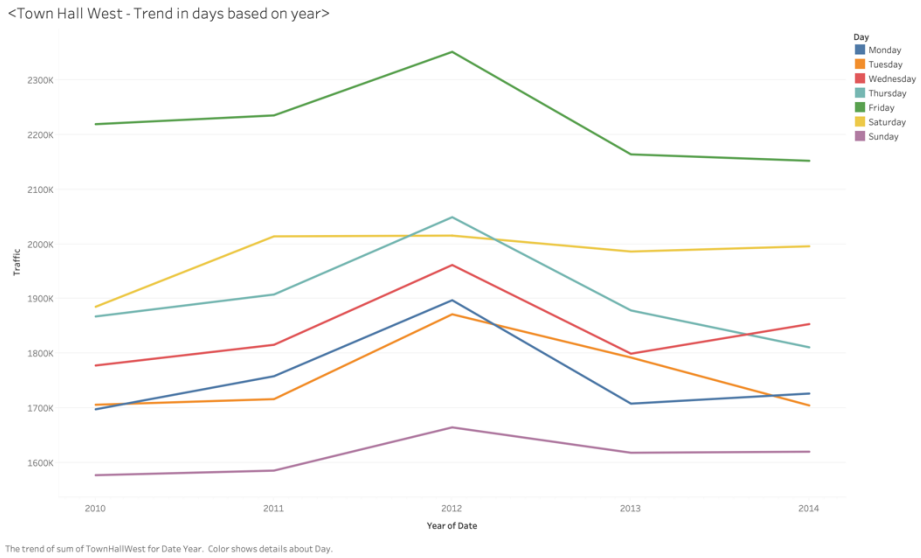

Figure 7. Pedestrian traffic over five years based on day

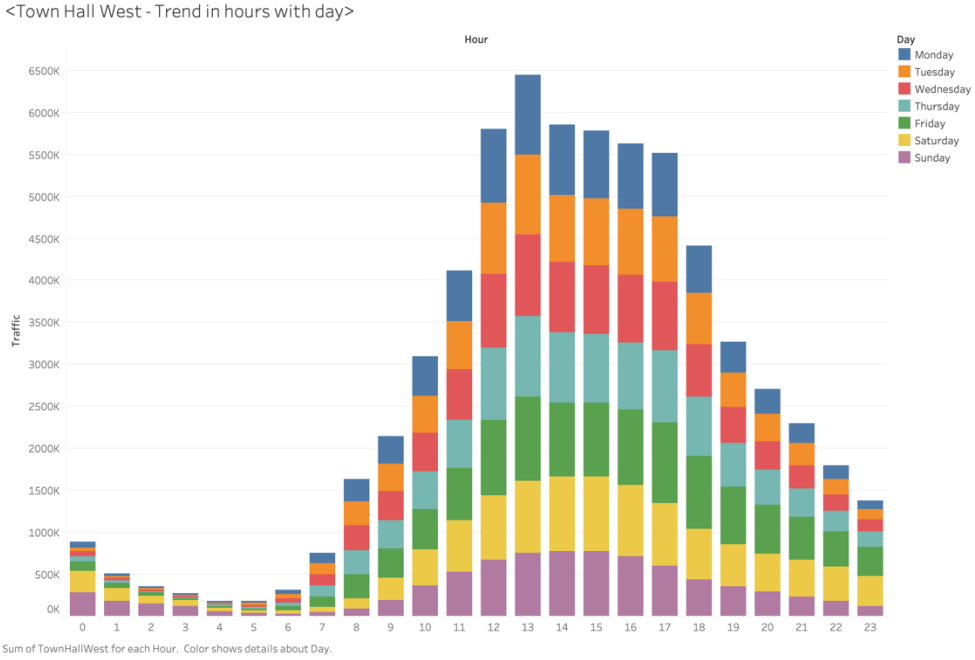

Figure 8. Pedestrian traffic per hour with day

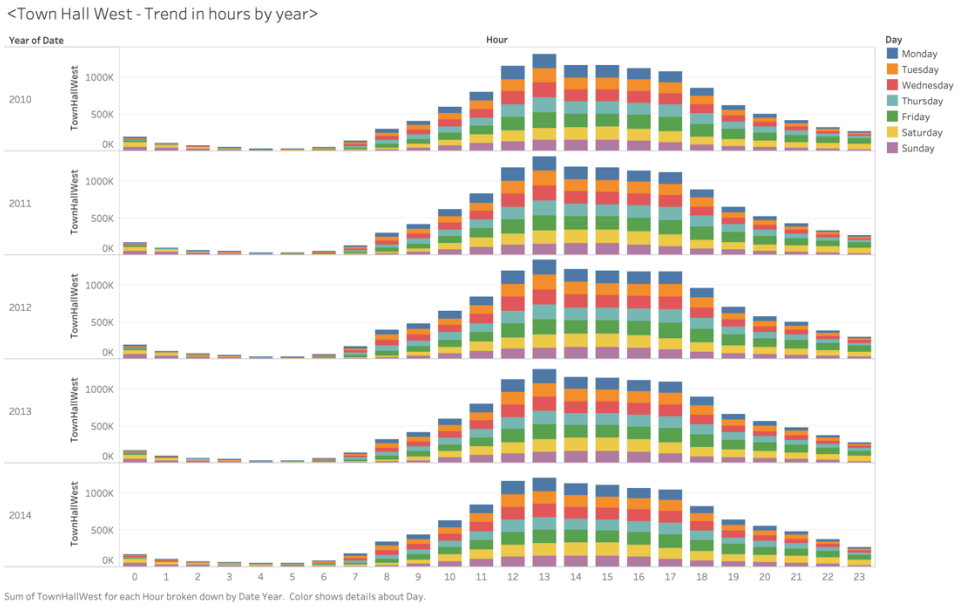

Figure 9. Pedestrian traffic per hour with a day based on year

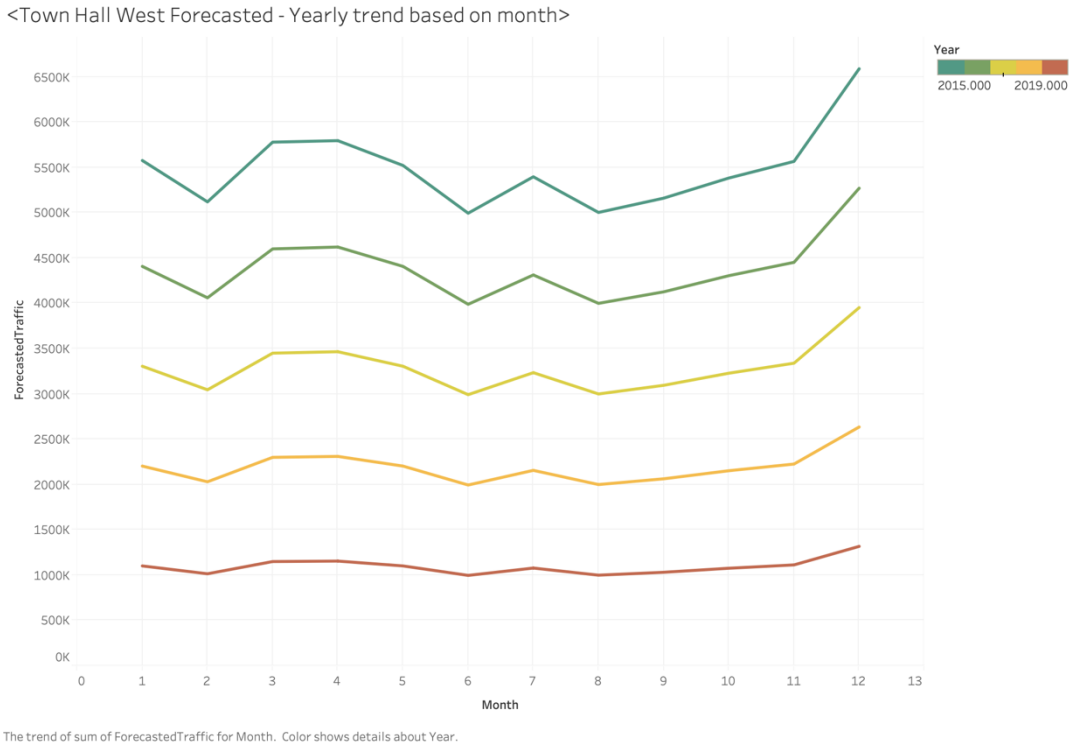

Figure 10. Forecasted traffic per year based on month

4. Methodology

The total number of eighteen sensor locations was initially analyzed as the list of locations can be found in Figure 4. This is because only eighteen locations (See Figure 4) listed below were consistently captured during each period described in Figure 3.

No.	Period	The Number of Locations
1	2009.09 – 2013.08	18
2	2013.09 – 2014.04	28
3	2014.05 – 2014.09	31
4	2014.10 – 2015.01	33
5	2015.02	37

Figure 11. The number of locations captured during each period

#	Location
1	State Library
2	Collins Place (South)
3	Collins Place (North)
4	Flagstaff Station
5	Melbourne Central
6	Town Hall (West)
7	Bourke Street Mall (North)
8	Bourke Street Mall (South)
9	Australia on Collins
10	Southern Cross Station
11	Victoria Point
12	New Quay
13	Waterfront City
14	Webb Bridge
15	Princes Bridge
16	Flinders St Station Underpass
17	Sandridge Bridge
18	Birranrung Marr

Figure 12. The list of eighteen locations

In order to identify pedestrian density for eighteen locations, average traffic per year in addition to total traffic was calculated. In the case of Town Hall West, it was having the highest density among eighteen locations, so additional analysis followed such as pedestrian traffic per hour or per day. Such analysis was done using both R and Tableau and analyzed based on various aspects such as day, hour, month, and year to identify the pattern or trend within the dataset.

The variable that the dataset offers is the number of pedestrians counted by sensors in various locations; the second variable is the time. With such variables, a time series analysis was appropriate for the following reasons (Edureka, 2017):

1) Values are not constant
2) Values cannot be represented using functions or formulas such as cos x

In this context, a time series algorithm was used to create a model, then the model was used to predict the future pedestrian traffic flow at Town Hall West based on the first variable, the number of pedestrians counted over five years.

In regard to forecasting, the following period was considered when making a time series (ts) object: 1st of January 2010 to 31st of December 2014. A time series object for Town Hall West was created before forecasting as figure 13 suggests. A time series object called "myts" was created based on "thwyearmonth" data frame; then when it was plotted, the NA values made a broken line. It is also noted that missing data in the period 13th to 14th of December in 2013, 24th of July to 31st of July in 2014, and 1st of August to 6th of August in 2014 was not dealt with yet (See Appendix B Figure B1). After adding interpolation using na.interp function, NA values were replaced with estimates (See Appendix B Figure B2).

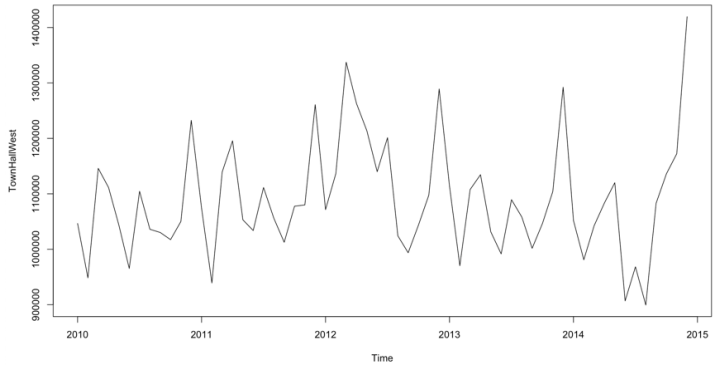

Figure 13. Plotting a time series object representing pedestrian traffic at Town Hall West

Using auto.arima function, the best ARIMA model was identified which was ARIMA(1,0,0)(0,1,1). This ARIMA model is then applied to the series, giving the forecasted values over the next 10 years with an accuracy level of 95% (See Figure 14).

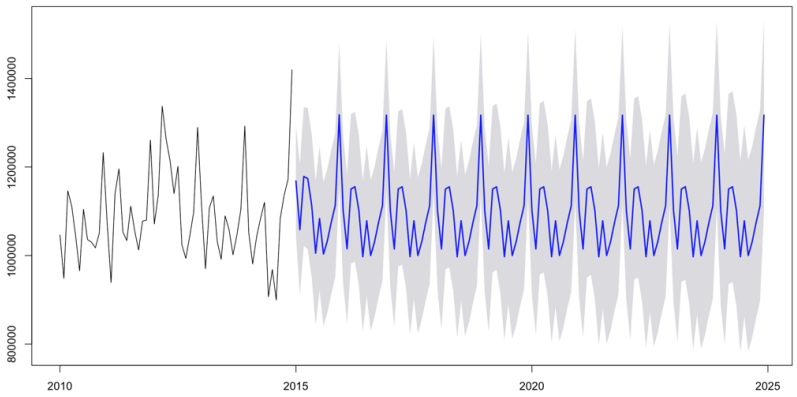

Figure 14. Plotting the forecasted values for 10 years

When testing this model, the values are similar to the actual value. For instance, the forecasted value for January 2015 is 1168346.6 whereas the actual value for January 2015 is 1154347 (See Appendix B Figure B5).

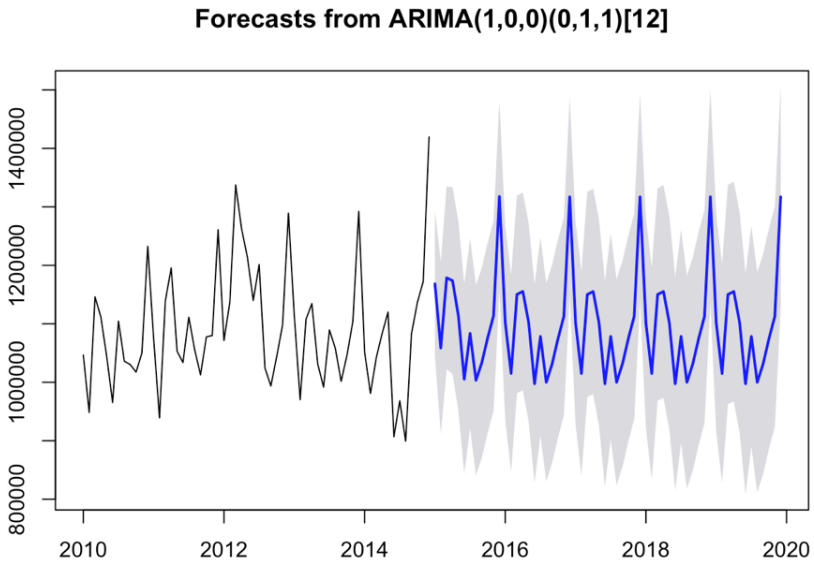

Figure 15. Plotting the forecasted values for 5 years

5. Conclusion

Based on this analysis and forecasting, it is noted that there will be consistent traffic around Town Hall West which is adjacent to Collins St which is known to have Melbourne's busiest tram routes. Taking this into consideration, it is recommended that Yarra Trams use E class trams along with current A class trams for routes 12, 48, and 109. Especially, 48 should be its top priority of concern as it is Melbourne's busiest route, having overcrowding issues. In this way, as E class trams carry much more passengers than A class trams do, this solution will be able to contribute to minimising the financial burden for Yarra Tram, reducing the overcrowding problem.

References

Carey, A. & Butt, C. (2013, June 19). Tram squeeze eases but some still suffer. https://www.theage.com.au/national/victoria/tram-squeeze-eases-but-some-still-suffer-20130619-2oj1i.html

City of Melbourne. (2014). Walking Plan. Retrieved from https://s3.ap-southeast-2.amazonaws.com/hdp.au.prod.app.com-participate.files/4114/3890/9931/Walking_Plan_full_version.pdf (2)

Edureka. [When Not To Use Time Series Analysis?]. (2017, July 27*). Time Series In R | Time Series Forecasting | Time Series Analysis | Data Science Training | Edureka* [Video file]. Retrieved from https://www.youtube.com/watch?v=wNB8AgZPFLU

Google. (2019). *Melbourne CBD*. Retrieved from https://www.google.com/maps/d/u/0/edit?mid=1BrWdvMHSu4RY2rg00dc_PhQcuXFlZeea&hl=en&ll=-37.814840569889554%2C144.9603517672448&z=15

Johnstone, R. (2019, October 16). Melbourne's trams ranked from best to worst. Retrieved from https://www.timeout.com/melbourne/blog/melbournes-trams-ranked-from-best-to-worst-033116

Laskle, A. (2016, November 26). Delays for Melbourne's busiest tram routes. Retrieved from https://www.theage.com.au/national/victoria/delays-for-melbournes-busiest-tram-routes-20161126-gsy2ek.html

Pixabay. (2017). Selective Color Photography of Yellow Train Beside Building [photograph]. Retrieved from https://www.pexels.com/photo/architecture-asphalt-blur-cables-417023/

Public Transport Victoria. (2017). Melbourne tram network. Retrieved from

https://www.ptv.vic.gov.au/assets/PDFs/Maps/Network-maps/417bb4810a/PTV_Tram-Network-Map_2017.pdf

Rooney, K. (2019, April 9). Compensation available after Metro Trains and Yarra Trams fail tough new targets. Retrieved from https://www.news.com.au/national/victoria/compensation-available-after-metro-trains-and-yarra-trams-fail-tough-new-targets/news-story/6b75c43f05d8fcefc7a54396850f086c

Mekala, H. (2018, June 30). Dealing with Missing Data using R [Web log post]. Retrieved from https://medium.com/coinmonks/dealing-with-missing-data-using-r-3ae428da2d17

Transport Safety Victoria. (2019). *2018 Annual incident statistics*. Retrieved from https://transportsafety.vic.gov.au/rail-safety/regulation-and-investigation/incident-statistics

Victorian Injury Surveillance Unit (VISU). (2017). Tram related injury statistics Victoria 2005/06 to 2014/15. Retrieved from https://transportsafety.vic.gov.au/__data/assets/pdf_file/0008/136619/Hopitalisation-data-analysis-Tram-2005-15.pdf

Appendix A: Analysis of Melbourne Pedestrian Traffic Dataset

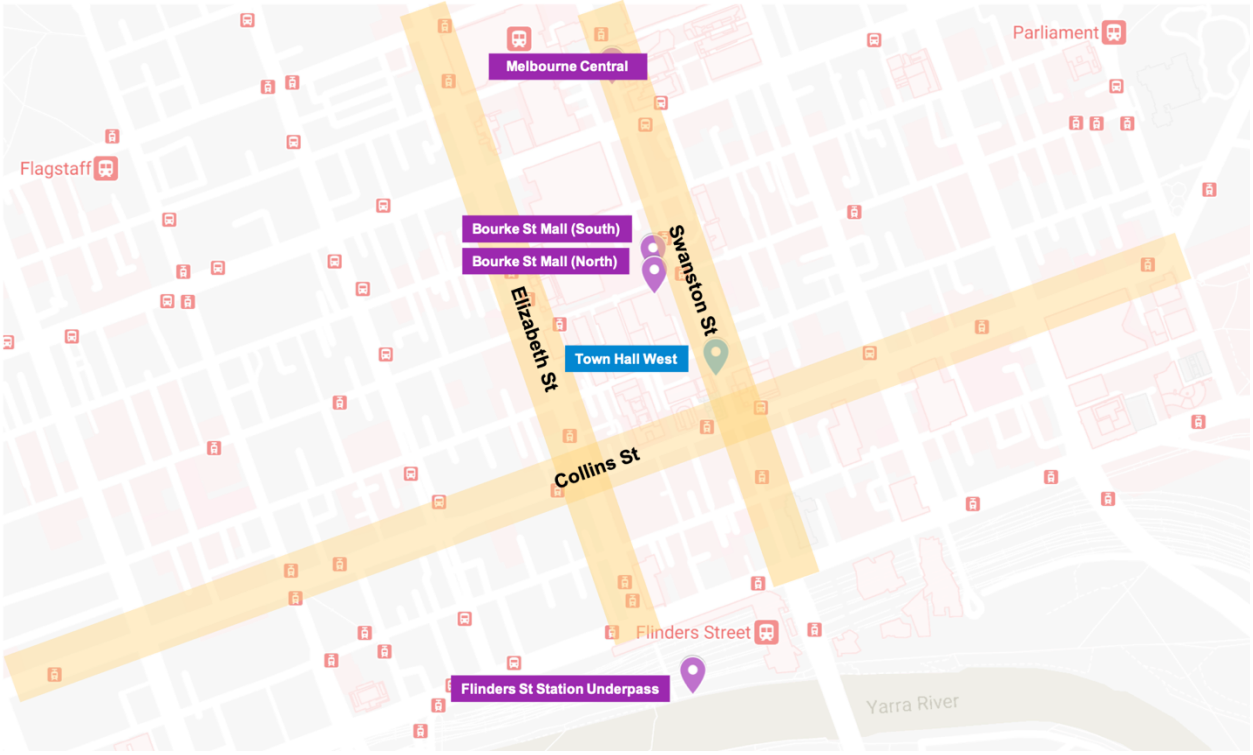

Figure A1. A broader map for sensor locations of the top five average traffic (Google, 2019)

Figure A2 illustrates the list of periods when missing data existed and it also indicates which location column was missing data.

No.	Period	Location with missing Data
1	2009.05.01 – 2009.05.18	Melbourne Central
2	2009.05.01 – 2009.05.20	Webb Bridge
3	2009.05.01 – 2009.09.24	Birrarung Marr
4	2010.08.06 – 2010.12.13	Birrarung Marr
5	2010.11.10 – 2010.12.31	Bourke Street Mall (North)
6	2011.01.01 – 2011.12.19	Bourke Street Mall (North)
7	2012.09.04 – 2012.09.19	State Library
8	2012.12.21 – 2012.12.31	State Library
9	2013.10.06	All eighteen locations
10	2013.01.01 – 2013.01.15	State Library
11	2013.06.01 – 2013.08.31	State Library
12	2013.08.01 – 2013.12.17	New Quay

13	2013.12.04 – 2013.12.20	Birrarung Marr
14	2013.12.13 – 2013.12.14	All eighteen locations
15	2014.03.19	Melbourne Central
16	2014.04.04	Southern Cross Station
17	2014.04.11	Bourke Street Mall (South)
18	2014.04.12	Flinders St Station Underpass
19	2014.06.07	Sandridge Bridge
20	2014.06.26 – 2014.06.30	Australia on Collins
21	2014.07.01 – 2014.07.31	Australia on Collins
22	2014.07.24 – 2014.07.31	Town Hall (West)
23	2014.08.01 – 2014.08.06	Town Hall (West)
24	2014.08.01 – 2014.08.31	Australia on Collins
25	2014.08.08 – 2014.08.31	Birrarung Marr
26	2014.09.01 – 2014.09.30	Australia on Collins
27	2014.09.16 – 2014.09.30	Birrarung Marr
28	2014.10.01 – 2014.10.31	Australia on Collins
29	2014.10.01 – 2014.10.31	Birrarung Marr
30	2014.11.01 – 2014.11.30	Australia on Collins
31	2014.11.01 – 2014.11.30	Birrarung Marr
32	2014.11.11 – 2014.11.30	Bourke Street Mall (North)
33	2014.12.01 – 2014.12.31	Bourke Street Mall (North)
34	2014.12.01 – 2014.12.31	Australia on Collins
35	2014.12.01 – 2014.12.31	Birrarung Marr
36	2015.01.01 – 2015.01.31	Bourke Street Mall (North)
37	2015.01.01 – 2015.01.31	Australia on Collins
38	2015.01.01 – 2015.01.31	Birrarung Marr
39	2015.02.01 – 2015.02.16	Bourke Street Mall (North)
40	2015.02.01 – 2015.02.28	Australia on Collins
41	2015.02.27 – 2015.02.28	Webb Bridge

Figure A2. The list of missing data according to the period

The average traffic for 2009, 2010, 2012, 2013, and 2014 are shown in Figure A3. Town Hall West had the highest average pedestrian traffic throughout the periods. From the second highest to the fifth highest traffic, the ranking varied among Melbourne Central, Bourke Street Mall South, Bourke Street Mall North, Flinders St Station Underpass, and Princes Bridge.

	Location	AverageTraffic
1	TownHallWest	1508
2	BourkeStreetMallSouth	1225
3	MelbourneCentral	1051
4	FlindersStStationUnderpass	1020
5	PrincesBridge	969
6	BourkeStreetMallNorth	944
7	AustraliaOnCollins	704
8	FlagstaffStation	688
9	StateLibrary	683
10	CollinsPlaceSouth	449
11	CollinsPlaceNorth	362
12	SandridgeBridge	335
13	SouthernCrossStation	268
14	BirrarungMarr	222
15	VictoriaPoint	198
16	NewQuay	198
17	WaterfrontCity	132
18	WebbBridge	73

2009

	Location	AverageTraffic
1	TownHallWest	1446
2	MelbourneCentral	1078
3	BourkeStreetMallSouth	1069
4	FlindersStStationUnderpass	981
5	PrincesBridge	962
6	BourkeStreetMallNorth	794
7	FlagstaffStation	704
8	StateLibrary	690
9	AustraliaOnCollins	690
10	CollinsPlaceSouth	449
11	CollinsPlaceNorth	369
12	SouthernCrossStation	348
13	SandridgeBridge	313
14	BirrarungMarr	257
15	VictoriaPoint	225
16	NewQuay	225
17	WaterfrontCity	132
18	WebbBridge	100

2010

	Location	AverageTraffic
1	TownHallWest	1576
2	MelbourneCentral	1201
3	BourkeStreetMallSouth	1138
4	FlindersStStationUnderpass	1099
5	BourkeStreetMallNorth	1045
6	PrincesBridge	891
7	StateLibrary	778
8	FlagstaffStation	727
9	AustraliaOnCollins	713
10	CollinsPlaceSouth	463
11	SouthernCrossStation	446
12	CollinsPlaceNorth	346
13	SandridgeBridge	339
14	BirrarungMarr	333
15	VictoriaPoint	181
16	NewQuay	181
17	WebbBridge	143
18	WaterfrontCity	96

2012

	Location	AverageTraffic
1	TownHallWest	1524
2	MelbourneCentral	1287
3	BourkeStreetMallNorth	1166
4	BourkeStreetMallSouth	1124
5	PrincesBridge	1015
6	FlindersStStationUnderpass	873
7	StateLibrary	824
8	AustraliaOnCollins	721
9	FlagstaffStation	714
10	SouthernCrossStation	495
11	CollinsPlaceSouth	410
12	SandridgeBridge	364
13	BirrarungMarr	335
14	CollinsPlaceNorth	330
15	WebbBridge	164
16	VictoriaPoint	122
17	NewQuay	122
18	WaterfrontCity	119

2013

	Location	AverageTraffic
1	TownHallWest	1433
2	MelbourneCentral	1221
3	BourkeStreetMallNorth	1053
4	BourkeStreetMallSouth	992
5	PrincesBridge	961
6	StateLibrary	813
7	FlindersStStationUnderpass	784
8	FlagstaffStation	689
9	AustraliaOnCollins	670
10	SouthernCrossStation	473
11	CollinsPlaceSouth	422
12	SandridgeBridge	342
13	CollinsPlaceNorth	309
14	BirrarungMarr	235
15	WebbBridge	161
16	WaterfrontCity	109
17	VictoriaPoint	103
18	NewQuay	103

2014

Figure A3. The average pedestrian traffic in 2009, 2010, 2012, 2013, and 2014

After merging five years of data into one data frame, Figure A4 illustrates locations in order of the average traffic: Town Hall West, Melbourne Central, Bourke Street Mall South, Bourke Street Mall North, and Flinders St Station Underpass.

	Location	AverageTraffic
1	TownHallWest	1507
2	MelbourneCentral	1177
3	BourkeStreetMallSouth	1104
4	BourkeStreetMallNorth	1002
5	FlindersStStationUnderpass	973
6	PrincesBridge	947
7	StateLibrary	763
8	FlagstaffStation	710
9	AustraliaOnCollins	701
10	CollinsPlaceSouth	443
11	SouthernCrossStation	420
12	CollinsPlaceNorth	344
13	SandridgeBridge	338
14	BirrarungMarr	289
15	VictoriaPoint	169
16	NewQuay	169
17	WebbBridge	133
18	WaterfrontCity	114

	Location	TotalTraffic
1	TownHallWest	35839107
2	MelbourneCentral	27989209
3	BourkeStreetMallSouth	26257159
4	BourkeStreetMallNorth	23822439
5	FlindersStStationUnderpass	23138756
6	PrincesBridge	22528436
7	StateLibrary	18148237
8	FlagstaffStation	16877727
9	AustraliaOnCollins	16669522
10	CollinsPlaceSouth	10534108
11	SouthernCrossStation	9983011
12	CollinsPlaceNorth	8179532
13	SandridgeBridge	8040879
14	BirrarungMarr	6878826
15	VictoriaPoint	4027934
16	NewQuay	4027934
17	WebbBridge	3176748
18	WaterfrontCity	2709836

Figure A4-1. (Left) The average pedestrian traffic over five years, (Right) the total pedestrian traffic over five years having the same top five locations

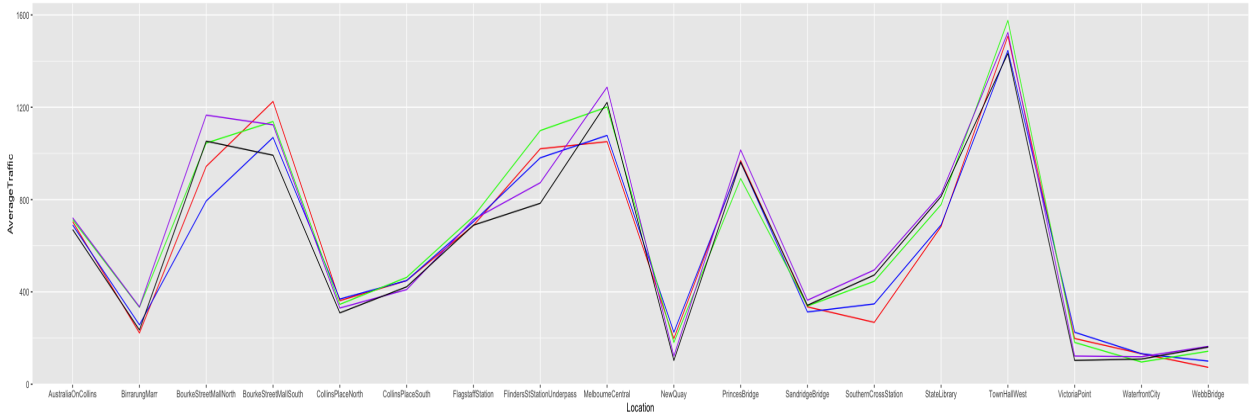

Figure A4-2. The average pedestrian traffic over five years per sensor location: red = 2009, blue = 2010, green = 2012, purple = 2013, black = 2014

Appendix B: Figures related to Forecasting

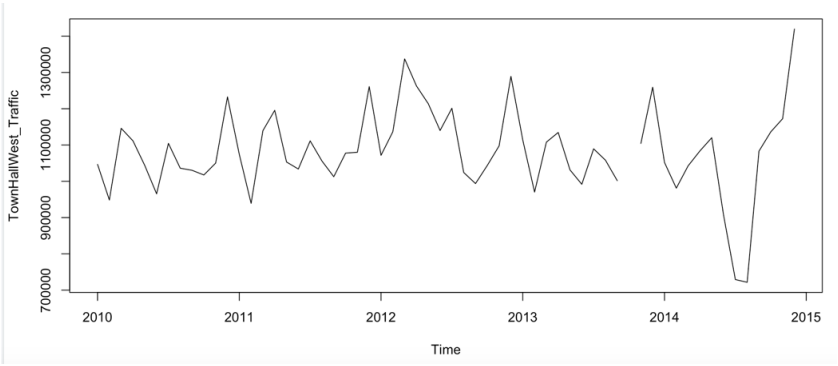

Figure B1. Plotting a time series object called “myts”

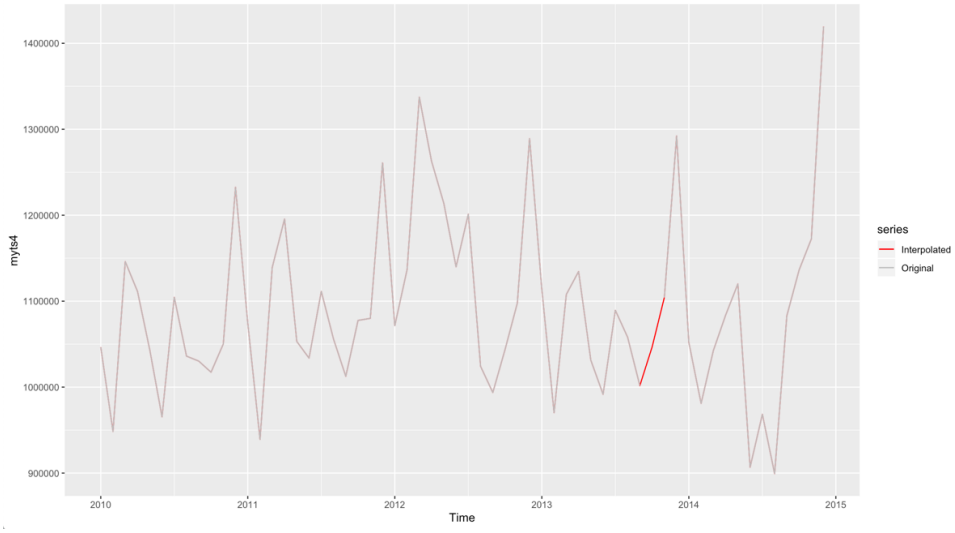

Figure B2. Auto-plot after adding interpolation for NA values

Figure B3 suggest plotting a time series object called “TownHallWest” after having dealt with NA values and missing data.

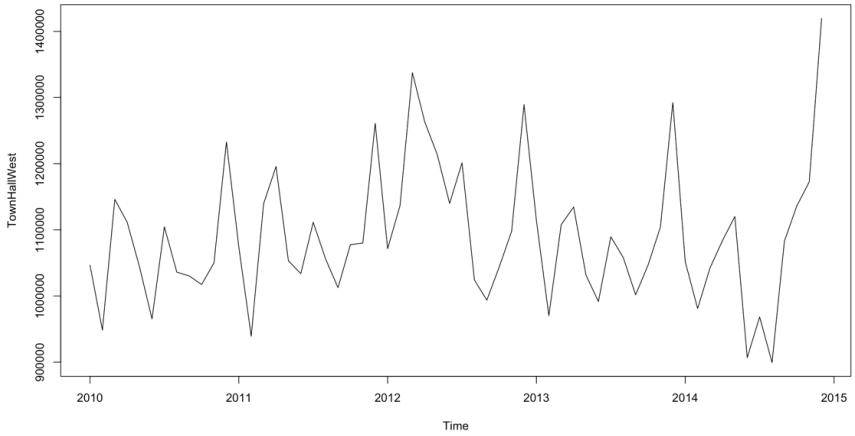

Figure B3. Plotting a time series object called “TownHallWest”

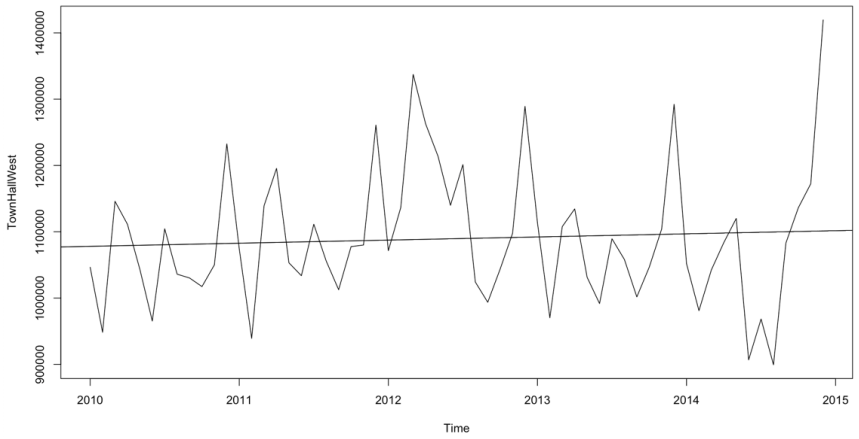

Figure B4. Line indicating mean value

```
> myforecast
         Point Forecast     Lo 95   Hi 95
Jan 2015      1168346.6 1044797.7 1291896
Feb 2015      1058567.2  911426.3 1205708
Mar 2015      1178578.2 1022475.9 1334680
Apr 2015      1173835.1 1014071.9 1333598
May 2015      1113715.0  952419.4 1275011
Jun 2015      1005283.8  843341.4 1167226
Jul 2015      1083544.8  921329.7 1245760
```

```
> jandata2015bymonth
  month(Date) TownHallWest_Traffic
1           1                1154347
```

Figure B5. Similar results were found in forecasted January 2015 and actual January 2015 values

	month(Date)	TownHallWest_Traffic
1	1	5357767
2	2	4975415
3	3	5773096
4	4	5787819
5	5	5461598
6	6	5037553
7	7	5235651
8	8	4895847
9	9	5121373
10	10	5297447
11	11	5504425
12	12	6461145

Figure B6. Town Hall West's pedestrian past pedestrian traffic by month

	Day	TownHallWest_Traffic
1	Friday	11122590
2	Saturday	9897346
3	Thursday	9514413
4	Wednesday	9208662
5	Tuesday	8791567
6	Monday	8788266
7	Sunday	8066450

Figure B7. Town Hall West's pedestrian past pedestrian traffic by day

Appendix C: External Sources

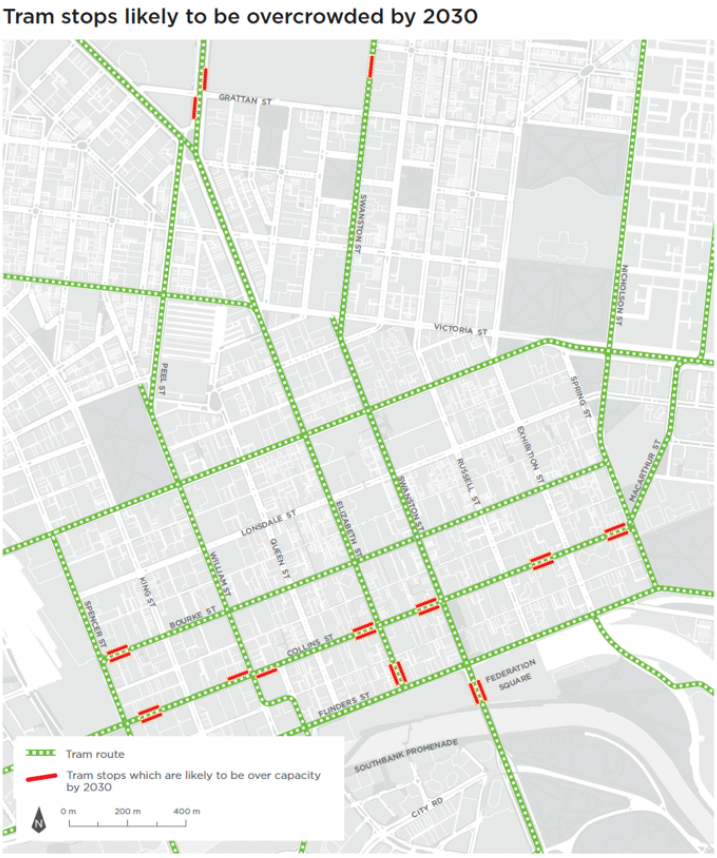

Figure 36: Tram stops likely to be at capacity by 2030

Figure C1. Expected overcrowded tram stops by 2030