**ASSESSMENT COVER SHEET**

|  |  |  |  |
| --- | --- | --- | --- |
| Unit Name and Code: | **FIT5141 Advance Topic of Information Technology** | | |
| Campus: | **Caulfield** | | |
| Assignment Title: | **Data analysis and Decision Support** | | |
| Name of Lecturer: | **Ariel Liebman** | | |
| Name of Tutor: | **Muhammad Osama** | | |
| Tutorial Day and Time: | **Tuesday 10:00am - 12:00am** | | |
| Phone Number: | **0490778801** | | |
| Email Address: | **sfen0004@student.monash.edu** | | |
| Has any part of this assignment been previously submitted as part of another unit/course?  Yes  No | | | |
| Due Date: | **Nov 1st 2017** | Date Submitted: | Nov 1st 2017 |
| All work must be submitted by the due date. If an extension of work is granted this must be specified with the signature of the lecturer/tutor.  Extension granted until (date)\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Signature of lecturer/tutor \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Please note that it is your responsibility to retain copies of your assessments. | | | |
| ***Intentional plagiarism or collusion amounts to cheating under Part 7 of the Monash University (Council) Regulations***  **Plagiarism**: Plagiarism means taking and using another person’s ideas or manner of expressing them and passing them off as one’s own. For example, by failing to give appropriate acknowledgement. The material used can be from any source (staff, students or the internet, published and unpublished works).  **Collusion**: Collusion means unauthorised collaboration with another person on assessable written, oral or practical work and includes paying another person to complete all or part of the work.  Where there are reasonable grounds for believing that intentional plagiarism or collusion has occurred, this will be reported to the Associate Dean (Education) or delegate, who may disallow the work concerned by prohibiting assessment or refer the matter to the Faculty Discipline Panel for a hearing. | | | |
| **Student Statement:**   * I have read the university’s Student Academic Integrity [Policy](http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-academic-integrity-policy.html) and [Procedures](http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-academic-integrity-managing-plagiarism-collusion-procedures.html). * I understand the consequences of engaging in plagiarism and collusion as described in Part 7 of the Monash University (Council) Regulations http://adm.monash.edu/legal/legislation/statutes * have taken proper care to safeguard this work and made all reasonable efforts to ensure it could not be copied. * No part of this assignment has been previously submitted as part of another unit/course. * I acknowledge and agree that the assessor of this assignment may for the purposes of assessment, reproduce the assignment and:  1. provide to another member of faculty and any external marker; and/or 2. submit it to a text matching software; and/or 3. submit it to a text matching software which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking.  * I certify that I have not plagiarised the work of others or participated in unauthorised collaboration when preparing this assignment.   Signature ..........Siyang Feng..................................... Date……… Nov 1st 2017……………………  \* delete (iii) if not applicable | | | |

|  |  |
| --- | --- |
| **Student ID number** | 28246993 |
| Given Name | Siyang |
| **Family name** | Feng |

*The information on this form is collected for the primary purpose of assessing your assignment* and ensuring the academic integrity requirements of the University are met*. Other purposes of collection include recording your plagiarism and collusion declaration, attending to course and administrative matters and statistical analyses. If you choose not to complete all the questions on this form it may not be possible for Monash University to assess your assignment. You have a right to access personal information that Monash University holds about you, subject to any exceptions in relevant legislation. If you wish to seek access to your personal information or inquire about the handling of your personal information, please contact the University Privacy Officer:* [*privacyofficer@adm.monash.edu.au*](mailto:privacyofficer@adm.monash.edu.au)

FIT5141 Advanced Topics of IT

Assignment 3

Siyang Feng

28246993

Catalogue

[Introduction 3](#_Toc497331819)

[Description of the Dataset 3](#_Toc497331820)

[Data format issues 3](#_Toc497331821)

[Missing Data and addition or removal fields 4](#_Toc497331822)

[The Analysis and Why it was Done – business question 4](#_Toc497331823)

[Business Question 4](#_Toc497331824)

[Summary Statistics 4](#_Toc497331825)

[Methodology 4](#_Toc497331826)

[Conclusion 11](#_Toc497331827)

[Figure 1 Optimisation of mELbourne population 5](#_Toc497332276)

[Figure 2 6](#_Toc497332277)

[Figure 3 6](#_Toc497332278)

[Figure 4 7](#_Toc497332279)

[Figure 5 7](#_Toc497332280)

[Figure 6 7](#_Toc497332281)

[Figure 7 8](#_Toc497332282)

[Figure 8 Optimisation of population trend 9](#_Toc497332283)

[Figure 9 optimasition of population in every hour in a day 10](#_Toc497332284)

[Figure 10 population in each month 11](#_Toc497332285)

[Figure 11 population in weekdays 11](#_Toc497332286)

# Introduction

People prefer what kind of mall? What kind of factors influences the mall more succeed? All of these factors will influence development of a mall. The target of this project is to analysing the factors influencing the development of mall in Melbourne. Analysing result will be useful for a new mall orientation and old mall developing. In this analysis, pedestrians for working, shopping and eating are considered to analyse if the malls in city are still popular. The dataset, Melbourne Pedestrian Traffic, is selected to solve this problem.

# Description of the Dataset

## Data format issues

All the date in the datasets of Melbourne Pedestrian Traffic are recorded as character and the hour is recorded as integer. In date column, some date is recorded as “01/01/2017” and some are recorded as “01-Jan-2017”. The character date format like “01-Jan-2017” cannot be transferred into date format. Thus, I create a function to convert it into format like “01/01/2017“. Then, “as.Date” function is used to transfer character into date. Then, using function, times, translate integer hour into time hour. Finally, the record format of pedestrian traffic is integer. For the further analysing, I change it into numeric.

## Missing Data and addition or removal fields

The biggest problem in this dataset is the missing data. When combining the different collections into one, some places are recorded only in the collections of some year. Thus, set the NA into 0 which is easier for further analysing. Then, the useless fields are deleted and the useful fields are combined.

# The Analysis and Why it was Done – business question

## Business Question

This analysis is aim to find out the factors which influence the development of mall and how they influenced.

## Summary Statistics

This project analyses the percentage of the population in three different malls, Bourke Street mall, Melbourne central and Collins Place, in Melbourne city. It analyses the population of these malls comparing other places and the tendency of each mall for 2011 to 2016 and the population changes in months, weekdays and hours. To make the result more confidential, a new collection of Melbourne population is imported.

# Methodology

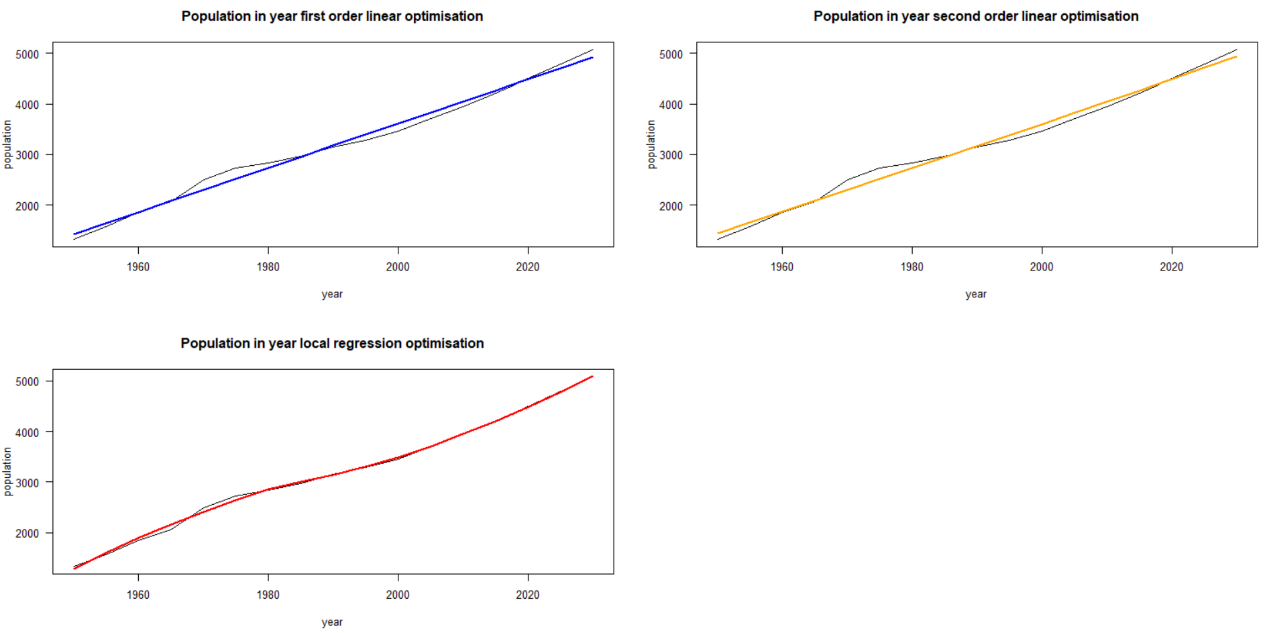


Figure 1 Optimisation of mELbourne population

With the development of population in Melbourne, the number of pedestrians in each place cannot represent the popularity of the mall. Thus, a new collection, population in Melbourne, is imported. The dataset of population is recorded every five years from 1950 to 2017 and predict to 2030. Figure 1 shows the regression lines based on the record population. As population of every year from 2011 to 2016 are needed. The Regression line is used to predict the population. From the figure 1, local regression has a better optimization in this model. Thus, parameters of local regression are chosen to predict.



Table 1 Population in year from 2009 to 2017

Table 1, the data frame, is the prediction result of figure 1. It records the population in each year from 2009 to 2017. The population is used to calculate the percentage of people in each place in each time.

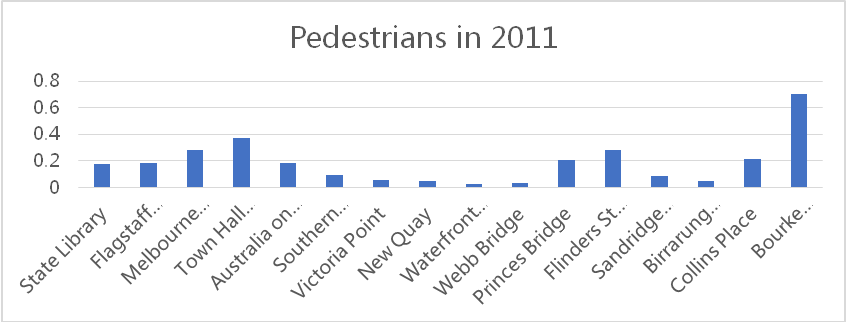


Figure 2

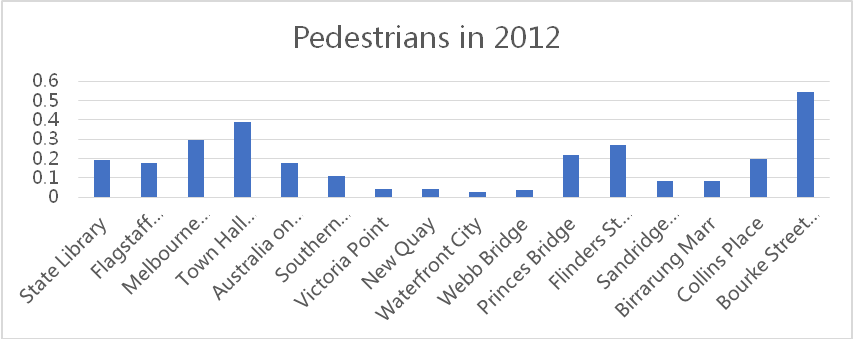


Figure 3

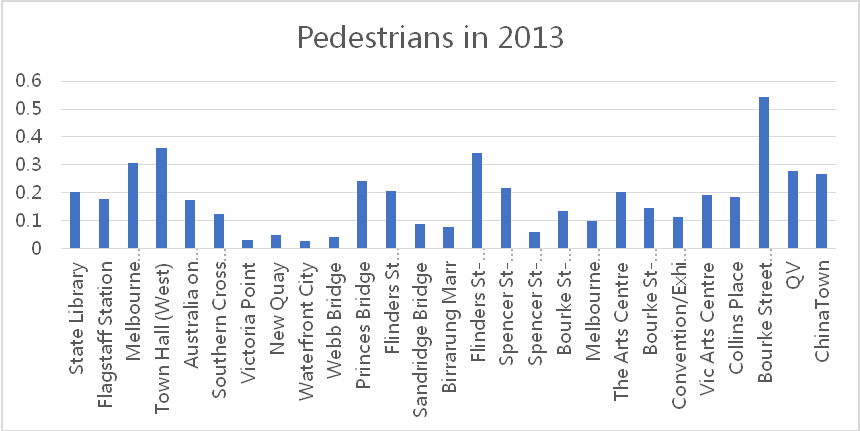


Figure 4

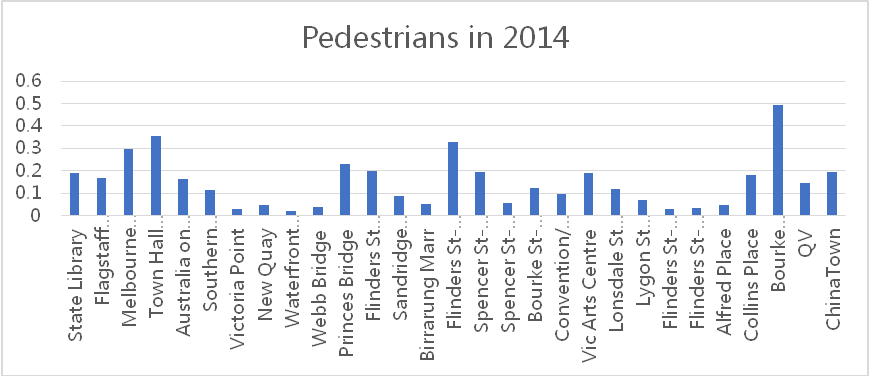


Figure 5

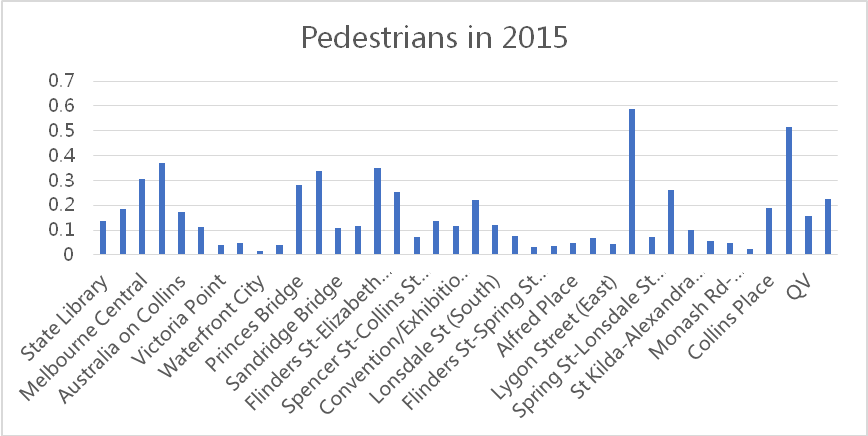


Figure 6

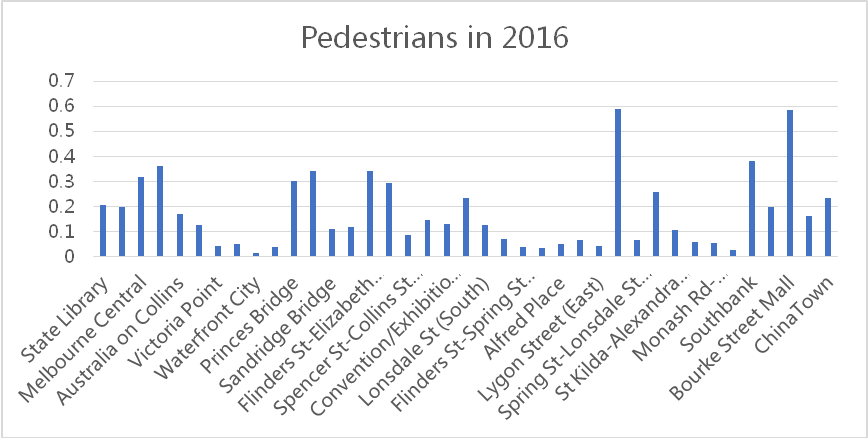


Figure 7

Bar charts from figure 2 to figure 7 shows the mean population for each year each place and indicates the population percentage from 2011 to 2016. It is easy to find out that Bourke street mall is almost the most popular place in every year. Melbourne central, with the station of the train and tram, looks not as attractive as considered. Collins Place which is far away from Melbourne central and Bourke street mall but near to southern cross, always as normal population percentage in each year. It looks like that the centre of Melbourne is still busy than other places.

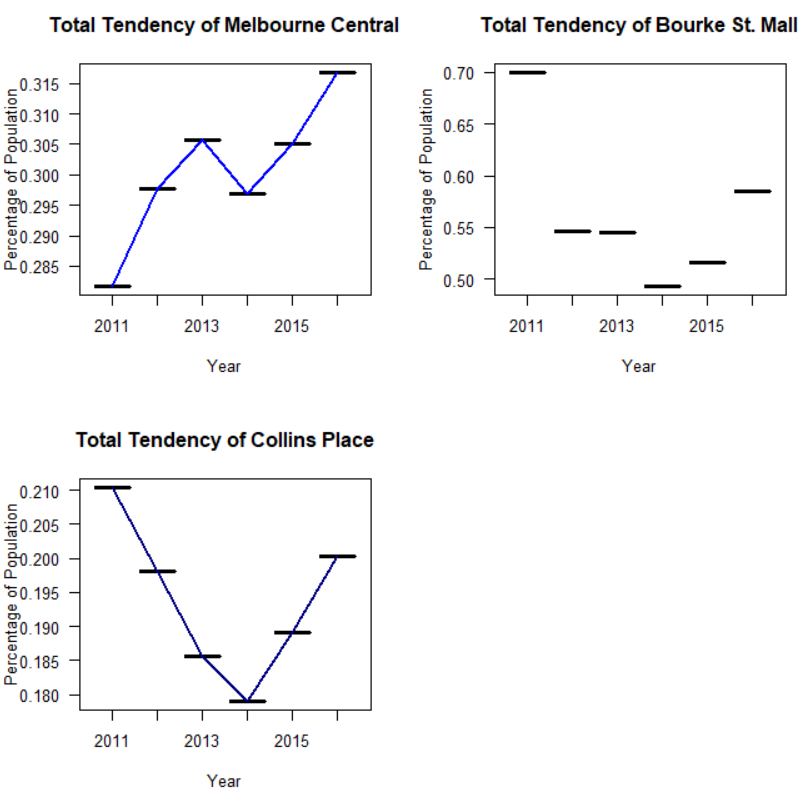


Figure 8 Optimisation of population trend

Figure 8 uses the line to show the tendency of these three malls for 2011 to 2016 using the mean value of each year. The population of these three malls is all decreased in 2014. It may be influenced by finance. However, in total, Melbourne central has the increasing tendency and other two malls decreased. Melbourne central is not as big as Bourke street mall but it lager than Collins Place. From this figure, it seems like that convenience traffic will be an important factor of the development of a mall.

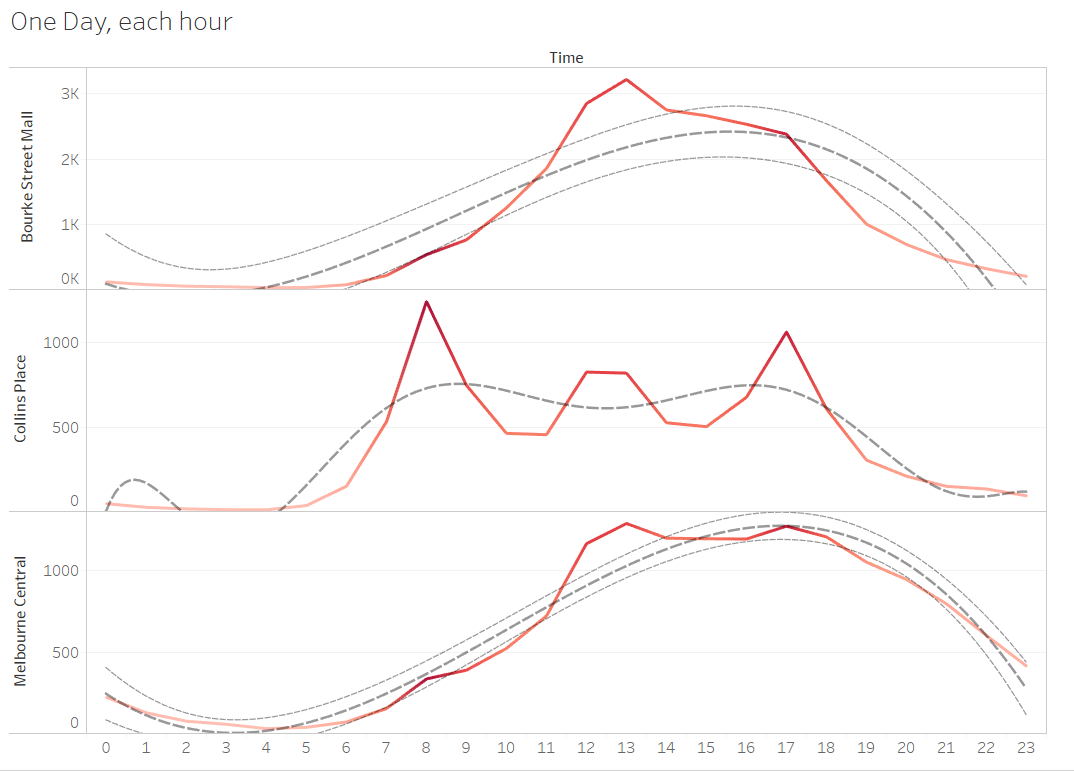


Figure 9 optimasition of population in every hour in a day

Line chart in figure 9 shows the percentage of population in each day using the sum population of all six years. Sum value makes figure more obviously. It records sum population of each hour from 2011 to 2016. It shows the tendency of a day. From the trend lines, Melbourne central and Bourke Street Mall are busy in the afternoon. However, in Bourke Street mall, most people go there between 10am and 7pm which is the exact the mall open time. Combining the figure 2 to figure 7, most of people will choose to go to this mall. Observing the line, the peak point is 12pm which is exact lunch time. It means that the mall also contains a lot of restaurants attracting people for lunch. Melbourne central has the similar trend line with Bourke Street mall. However, it contains two peaks, 12:00-13:00 and 17:00-18:00. The first one is lunch time and the second one is the time getting off work. Combining the trend after 18:00, because most shops have already closed, it could be most of the people will take the train back home or have other entertainment activities. And it can also explain the figure 8 what only Melbourne central has the positive tendency from 2011 to 2016. Combining the special point at 8:00, Melbourne central contains a lot of population but it seems not very succeed in shopping mall. In Collins Place, there are three peaks, 8:00 for working, 12:00 for lunch and 17:00-18:00. It totally satisfies the working schedule. It means that this mall has good traffic and has some restaurants but not popular on shopping. Even it is near to southern cross.

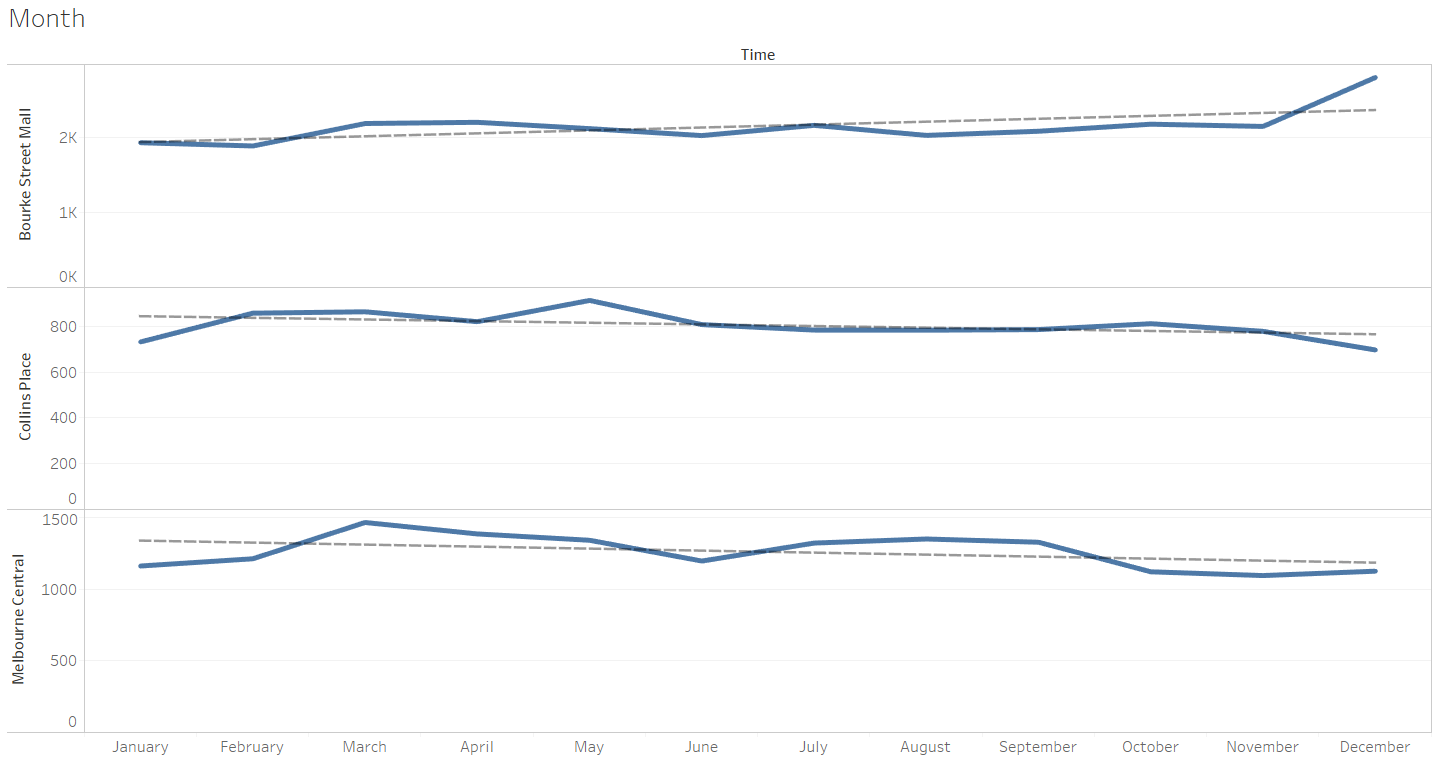


Figure 10 population in each month

Line chart in figure 10 shows the population in every month for past 6 years using the sum. The population in each month almost not changes a lot. However, in December, only Bourke Street mall increases a lot. It means that people are attracted by the activity of Christmas of Bourke Street mall.

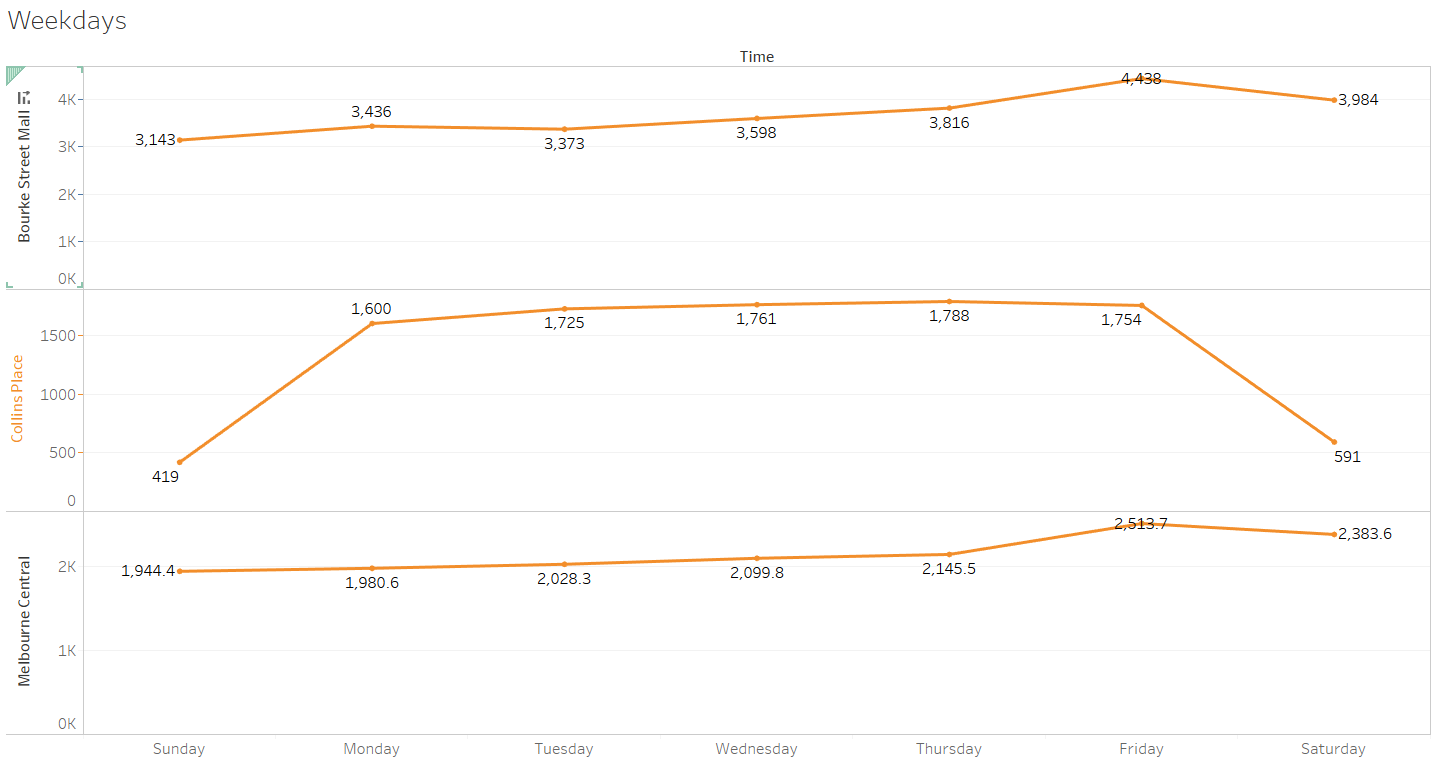


Figure 11 population in weekdays

Line chart in figure 11 shows the population from Sunday to Saturday in the three malls also using the sum of six years. Melbourne central and Bourke Street mall has a similar tendency. The population increases on Friday and Saturday which is the time out of work. It means that people spend their rest time on these two places. It might be shopping or other entertainment activities. However, Collins Place only contains more population on the workday. It means that people cannot consider to go to this mall when they are in rest.

# Conclusion

From the analysing in this report, Bourke Street mall is more succeed than another two malls. It does not have the traffic as convenience as Melbourne central has. But it is between the Melbourne central station and flinders station. Thus, it is still convenience. It offers a lot of restaurants for worker and shopper as other two malls. However, it has more attractive activities in festival than other malls. From the figure 2 to 7, it is easy to find out that Bourke Street mall attracts more people than other two. Thus, the comfortable and convenient shopping environment will impress people and help mall develop better.

In Melbourne central, is there anything attracting? It has the similar feature of Bourke street mall but does not have the population as many as the Bourke mall. The reason cannot be figure out using this dataset. For the future research, the dataset of pedestrian of every level in every time can be analysed to find out the result.

This analysing report would be useful for the manager of mall to decide what factors would be improved for their mall to get more populations.