

TECHNICAL REPORT

An Analysis of Melbourne Parking



CONTENTS

1. Introduction
2. Hypothesis
3. Methodology
4. Background
 - 2.1 VicRoads Timeline
 - 2.2 Key Stakeholders
 - 2.3 Data Quality Assessment
 - 2.4 Core Systems
 - 2.5 Key Term Definitions
5. Data Pre Processing
 - 3.1 Data Pre Processing Methodology
 - 3.2 Pre Processing Justifications
6. Data Exploration
 - 4.1 Data Exploration Methodology
 - 4.2 Key Findings
7. Machine Learning
 - 5.1 Linear Regression
 - 5.2 Logistic Regression
 - 5.2.1 Transforming Data for Logistic Regression
 - 5.2.2 Test Batch
8. Findings
9. Conclusion
10. Bibliography
11. Appendix

1.0 Introduction

The purpose of this technical report is to present our data analysis findings and documentation on a Metropolitan Melbourne Parking dataset in a project agreed upon by VicRoads and our company Data Design Solutions (DDS). Data analysis is the process of collecting, cleansing, transforming and modelling data to provide useful insights into making important decisions. There are many methodologies to perform data analytics depending on the industry and requirements. DDS aims to implement machine learning algorithms by focusing on exploratory data analysis to bring advisory and technical solutions for the issue of car park funding allocation by VicRoads. This report contains detailed information on how our team approached the steps of pre-processing, exploring and analysing parking data along with the technical explanation and the advisory deliverables for VicRoads to consult on the findings and conclusions.

2.0 Hypothesis

The hypothesis is that streets or areas with more shops and work buildings, but limited car parking spaces will have the highest number of violations.

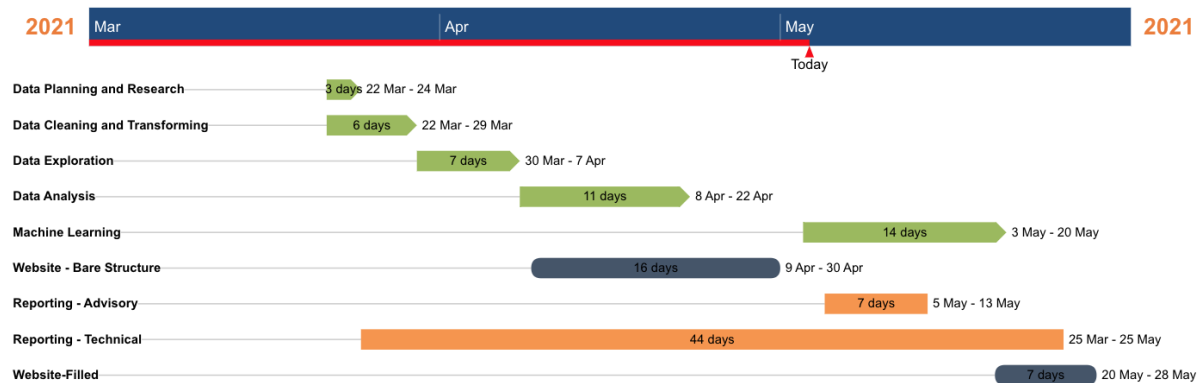
3.0 Methodology

A data set with records of parking events in metropolitan Melbourne was used in the analysis. The data set has several key attributes, including area, street, side of street, arrival/ departure time, parking duration, street sign and whether a violation occurred. Relations between attributes were analysed and visualised with various plots. A machine learning model was trained with the data set, which could make predictions on whether a parking event would result in a violation.

4.0 Background

2.1 VicRoads Timeline

The Gantt Chart below shows the main tasks DDS aims to undertake throughout the three months allocated for this project from 22nd Mar- 28th May 2021. There are 9 main tasks to complete which are allocated a reasonable amount of time for the team to work on data analytics and front-end website development along with every important documentation such as the advisory and technical reports.



2.2 Key Stakeholders

The success of the project will occur when the objectives are met and the expectations of the stakeholders are put into consideration. Stakeholders are people who care or have invested in the project and hope to gain from it. These individuals are constantly involved in the happenings of the project or have a stake in its success.

The Stakeholders for this project are as follows:

Internal Stakeholders

- Product Owner: Andong Wang
- Scrum Master: Alana Tobgui
- Development Team:
 - Yasiru Kodikara
 - Klarke Douvin Higuit Rueda
 - Alana Tobgui
 - Andong Wang

External Stakeholders

- Government: Provides regulations on how to undertake the project
- Customers: Individuals who will benefit on this project.
- VicRoads - Who will be using the product

2.3 Data Quality Assessment

Data Design Solutions holds data quality to the highest standard, and aims at improving the analysis of data through the initial cleaning of datasets provided. An important metric in the analysis of data quality includes the number of null values and the percentage of missing data based on columns and rows. These metrics can skew relationship data so modifications to the set must be made to minimise the risk of omitted information breaking or causing pseudo-relationships.

Quality metrics include:

- Duplicate names
- Incorrect inputs
 - For instance, extra digits in numeric values or the same categorical value capitalised in one entry and not capitalised in another.

It's important to ensure that a high standard is maintained in the analysis stage, so that recommendations to stakeholders are accurate and fall in line with the produced insights.

The 6-step Data Quality Framework [0]

1. Definition

Identify the business goals to improve the quality of data, the data owners / stakeholders, the impacted business processes, and the data rules. In this case, VicRoads will be able to use the analysed parking dataset to get useful insights and find solutions for future problems.

2. Assessment

Assess the existing data in relation to the criteria outlined in the Definition Stage. Evaluate data on a variety of dimensions, such as key attribute accuracy, completeness of all required attributes, attribute consistency across multiple data sets, and data timeliness.

3. Analysis

Analyse the assessment results on multiple fronts between current data and current goals. DDS will be using tools and methodologies defined in topic 2.4 core systems for data analysis.

4. Improvement

Based on prior analysis, create and implement improvement plans. The plans should include timeframes, resources, and costs. A clear example of an improvement stage is the data cleaning of the website's database, which includes detecting incorrect records from a record set, table, or database.

5. Implementation

Implement the solutions identified during the Improve stage. Understand any technical as well as business process changes.

6. Control

Business objectives and data rules must be clear and precise according to the Definition Step. Data quality is a continuous process and requires the entire organization to be data-driven and data focused.

2.4 Core Systems

The Core Systems are computers with these following programs installed and configured to help perform data analysis methodologies. The main programming language we used was Python.



Jupyter Notebook is an open-source web application that allows users to maintain live codes, equations, visualisations and narrative text. This application can be used for data cleaning and transformation, numerical simulation, statistical modeling, data visualisation and machine learning[1].



Pandas Library

Pandas is an open-source Python Library that helps users to perform data analysis and manipulation. It also offers data structures and operations for manipulating numerical data and time series. The team decided to use Pandas as the main library to perform data analytics throughout the project since it is fast , flexible and powerful[2].



Scikit-learn

Scikit-learn is a useful open source data analysis tool that allows users to use NumPy, SciPy and matplotlib to perform methodologies such as Classification, Regression , Clustering, Dimensionality reduction, model selection and preprocessing. Implementing machine learning techniques such as Linear Regression and Logistic Regression is made possible by the use of this application[3].



NumPy

NumPy is an open-source project aiming to enable numerical computing with Python. This library offers numerical computing tools to perform mathematical functions, number generators, linear algebra routines and Fourier transformations[4].



Matplotlib

This library is useful for creating static, animated and interactive visualisations in Python. Generating charts and graphs of the parking dataset during the exploratory data analysis stage, matplotlib was very useful for the team to get accurate results and visualisations for instances related to the parking dataset[5].



Seaborn

Like Matplotlib, Seaborn also is a Python based statistical data visualisation tool that allows users to draw attractive and informative graphs. This tool was used for visualisation of presenting graphs along with matplotlib[6].

2.5 Key Term Definitions

- Device ID: number plate of the car
- Street ID: unique street identifier
- Arrival Time: the time arrived
- Departure Time: the time departed
- Side of Street: parking on streets
- Street Marker: unique marker to identify street
- Street: street name
- Duration: time taken during an event
- Count: amount of events occurred
- Area: the location of a specific entity
- Violation: parking violation
- Charts -> Violin Plot -> consist of **white** dot which represents median value, **thick** black bar represents interquartile range and **thin** line represents the rest of distribution
- Violin plot-> useful in observing and comparing the distribution of numerical data across multiple groups
- Charts -> Scatter plot -> important in statistics to show correlation between values

5.0 Data Preprocessing

3.1 Data Pre-Processing Methodology

The team at DDS applied several cleaning techniques, including removing entries with null values, identifying entry errors, renaming and removing irrelevant ID columns and other data. Then attributes were transformed based on the assumed regression task.

3.2 Data Pre-Processing Justifications

3.2.1 Column Renaming

The data presented to the team required pre-processing before any relevant exploratory data analysis could be performed. This is vital for DDS to accurately analyse the data to deliver the best results for VicRoads.

The first instance of pre-processing focused on the columns of the dataset. In this instance, the following column names were changed:

- Area Name → Area
- Street Name → Street
- Duration of Parking Event (in seconds) → Parking Durations (s)
- Between Street 1 → Street Limit 1
- Between Street 2 → Street Limit 2
- In Violation? → Violation

All of the following column names remained the same:

- Side of Street
- Street Marker
- Arrival Time
- Departure Time
- Street ID
- Device ID

The next stage of pre-processing consisted of removing the columns which were unnecessary to the analysis, these specifically included identification or key columns, where each row consisted of a new input not relevant to the analysis. It was initially presumed that 'Street' and 'Street ID' held all of the same information. That is, each street name held the same, consistent ID number in the 'Street ID' Column. This would indicate that the table consisted of redundant values. The two columns remained in the dataset initially with the intention to use the ID column as a pseudo 'transformed' set of the 'Street' column.

3.2.2 Column Removal

Data cleaning is a good practise for getting rid of and changing out of date data values. This is because it serves no purpose and is not inferred by categorical variables that appear only once, the analysis ID should be removed. Similarly, columns with more than 50% of the data must be removed to avoid data skewing because there is insufficient information to infer relationships between variables. In addition, redundant columns must be removed as well, since they share the same information and most likely it will be simplified as one count.

3.2.3 Unique Identifier Check

This process will check for unique values in columns of interest, find outliers and possible mistakes relating to the dataset. This is vital towards getting successful results in the extrapolation methodology.

<i>Unique Entries in 'Area'</i>	<i>Unique Entries in 'Street'</i>
Banks Chinatown Courtney Princes Theatre Hyatt County RACV Spencer City Square The Mac Titles Magistrates Rialto Queensberry Victoria Market Supreme Hardware Regency Docklands Tavistock Southbank West Melbourne Jolimont	MARKET STREET RUSSELL STREET ELIZABETH STREET EXHIBITION STREET LONSDALE STREET Lt COLLINS STREET BOURKE STREET FLINDERS LANE SPRING STREET COLLINS STREET Lt LONSDALE STREET A'BECKETT STREET Lt BOURKE STREET KING STREET ERROL STREET FRANKLIN STREET Lt DRYBURGH STREET SOUTH WILLIAM STREET WILLS STREET SPENCER STREET FLINDERS STREET TERRY STREET QUEEN STREET LEVESON STREET O'CONNELL STREET CHETWYND STREET WALSH STREET ANDERSON STREET ROSSLYN STREET LA TROBE STREET BOND STREET

	CAPEL STREET COBDEN STREET FRANCIS STREET LA TROBE STREET QUEENSBERRY STREET EADES PLACE DRYBURGH STREET CHURCH STREET SWANSTON STREET PEEL STREET DUDLEY STREET CURZON STREET VICTORIA STREET HOWARD STREET ANTHONY STREET RODEN STREET MACKENZIE STREET ABBOTSFORD STREET PRINCESS STREET UNION STREET DODDS STREET GRANT STREET COVENTRY STREET BALSTON STREET KAVANAGH STREET MILES STREET DORCAS STREET SOUTHBANK BOULEVARD STURT STREET WELLS STREET FAWKNER STREET BATMAN STREET LANSDOWNE STREET ALBERT STREET CATHEDRAL PLACE GISBORNE STREET CLARENDON STREET PARLIAMENT PLACE ST ANDREWS PLACE CITY ROAD ST KILDA ROAD WELLINGTON PARADE NICHOLSON STREET JEFFCOTT STREET
--	---

Unique Entries in 'Side of Street' **[5 2 4 3 1]**

Unique Entries in 'Device ID' **7113**

Unique Entries in 'Sign' **329**

The values for Device ID indicate that the dataset has multiple events for the same car, and so using this to determine rates with repeating visits will be useful in the analysis

3.2.4 Null Rows Check

The team found no row with less than 50% of columns filled in the dataset, subsequently no row was removed.

6.0 Data Exploration

4.1 Data Exploration Methodology

Data Exploration focused on finding relationships between the different categorical and numerical attributes of the dataset. The initial goal was to explore relationships within specific areas, such as whether certain streets, signage and areas of the street have higher rates of infringements, and whether these patterns existed in other areas of the database.

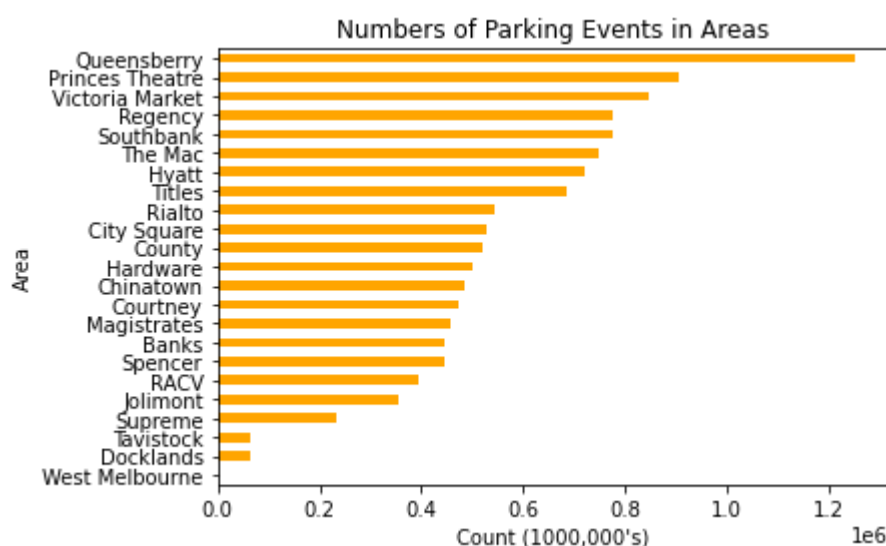
4.2 Key Findings

The key findings of the exploration can be linked to the following topics:

- Exploring Areas
- Street Exploration
- Parking duration exploration
- Device ID exploration
- Multivariate exploration
- Composition of All Violations by Violations per Device
- Individual Area Analysis

4.2.1 Exploring Areas

Figure 1 - Parking events



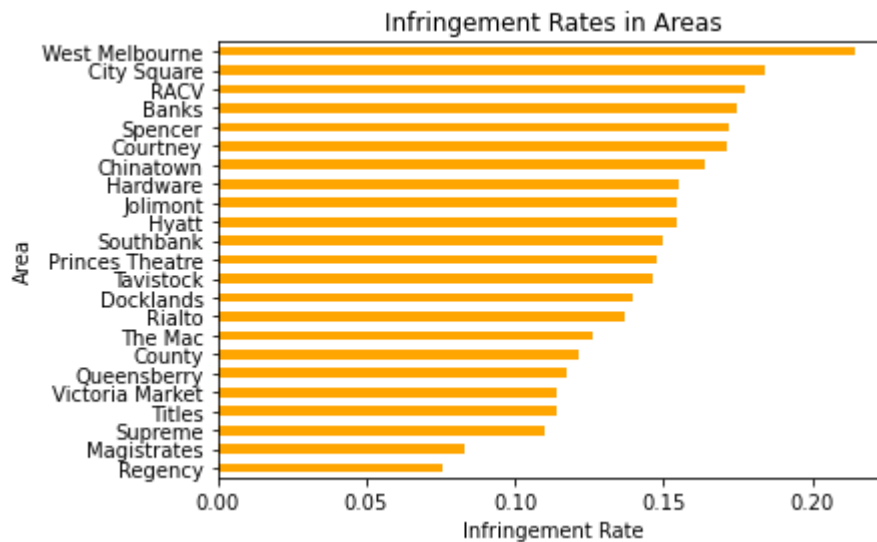
As shown in Figure 1, the graph above represents the number of parking events that occurred in the Metropolitan Melbourne area. This exploration was performed using the matplotlib method in Jupyter Notebook and the y-axis displays the names of the parking areas and the x-axis displays the numbers of parking events occurred in each area. The Queensberry area is reported to be the busiest area for parking events which has a value of slightly more than 1,200,000 since many people occasionally visit busy places like the Princes Theatre, Victoria Market, Regency, Southbank, The Mac, Hyatt and many other places surrounding the area. Therefore, it is common for many owners to park their vehicles around Queensberry.

The areas from Titles to Spencer recorded a value of more than 400,00 in parking events also making them areas with a lot of parking activity. The areas from RACV to West Melbourne are slowly decreasing in value since parking spaces tend to be lower in those areas.

Solutions to this would include:

The Queensberry area is reported to be the busiest area, the parking events happened when people parked their cars on this area which makes the area busy because of there so many markets, restaurant, people who are working, visitors, students who are driving to university and people who went to the bar and forgets their parked car therefore, prohibiting on-street parking on certain routes at certain times.

Figure 2 - Infringement Rates



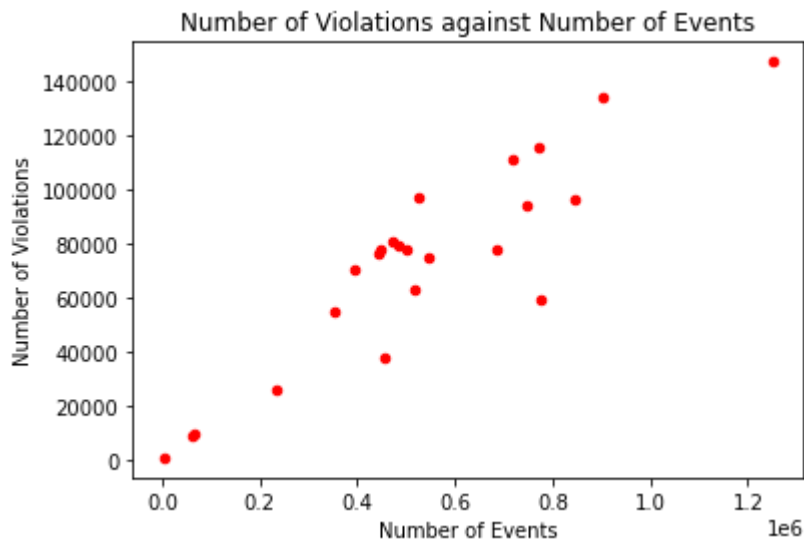
As shown in Figure 2, The graph represents the Infringement Rates found in the above areas. The y axis labels the areas where the infringement occurred and the x axis represents the infringement rates calculated in a median value. West Melbourne has a high rate of parking violations (>0.20) compared to all the other areas. Areas from City Square to Tavistock have rates between 0.15 and 0.20 making them locations where high infringement rates have been recorded.

The infringement rate in West Melbourne area are extremely high, the infringement rates happened when people went to shops and restaurants and forgets their parking duration and what time they been parked on the parking space

Solutions to this would include:

Implement parking meters in every area and locate parking meters and payment is only made in the valid bay number where the car is parked and increase parking spots

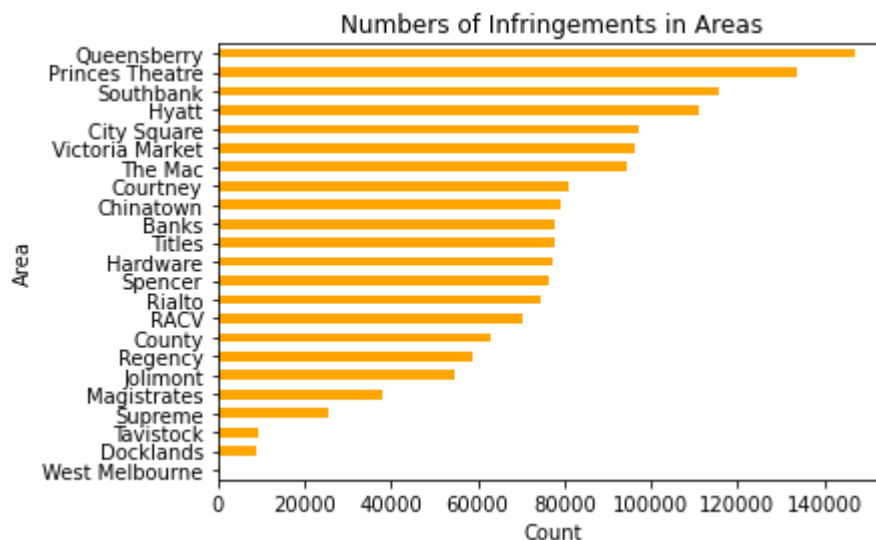
Figure 3 - Number of violations against number of events




As shown in Figure 3, we can see the y-axis represents the number of violations ranging from 0-140000, while the x-axis represents the number of events ranging from 0.0 - 1.2 in median values. There seems to be a positive linear correlation between the number of parking events and the number of violations in the same area. Therefore, linear regression will be attempted first before more categorical values.

This graph illustrates events and violations depending on the circumstances. We can tell that in one day whether these people go to the shop, buy food, go to work, visit their loved ones, in every event which happened in one day there are violations.

Figure 4 - Infringements





As shown in Figure 4, the graph above illustrates the numbers of infringements taking place in the Metropolitan Melbourne areas. The x-axis represents the total count of infringements in specific areas. The Queensberry area shows the highest number of parking infringements, with over 140000 recorded. Many people who parked in the West Melbourne area, did not receive any infringements. Queensberry appears to be the busiest area in Melbourne with a high number of infringements. Princes Theatre appears to be the next common area in the Metropolitan Melbourne area with a count of more than 12000 infringements.

The Queensberry area is reported to be the busiest area with the highest number of infringements. The infringements occurred when people parked their cars in this area, making the area busy because there are so many markets, restaurants, people working, visitors, students driving to university, and people who went to the bar and forgot their parked car.

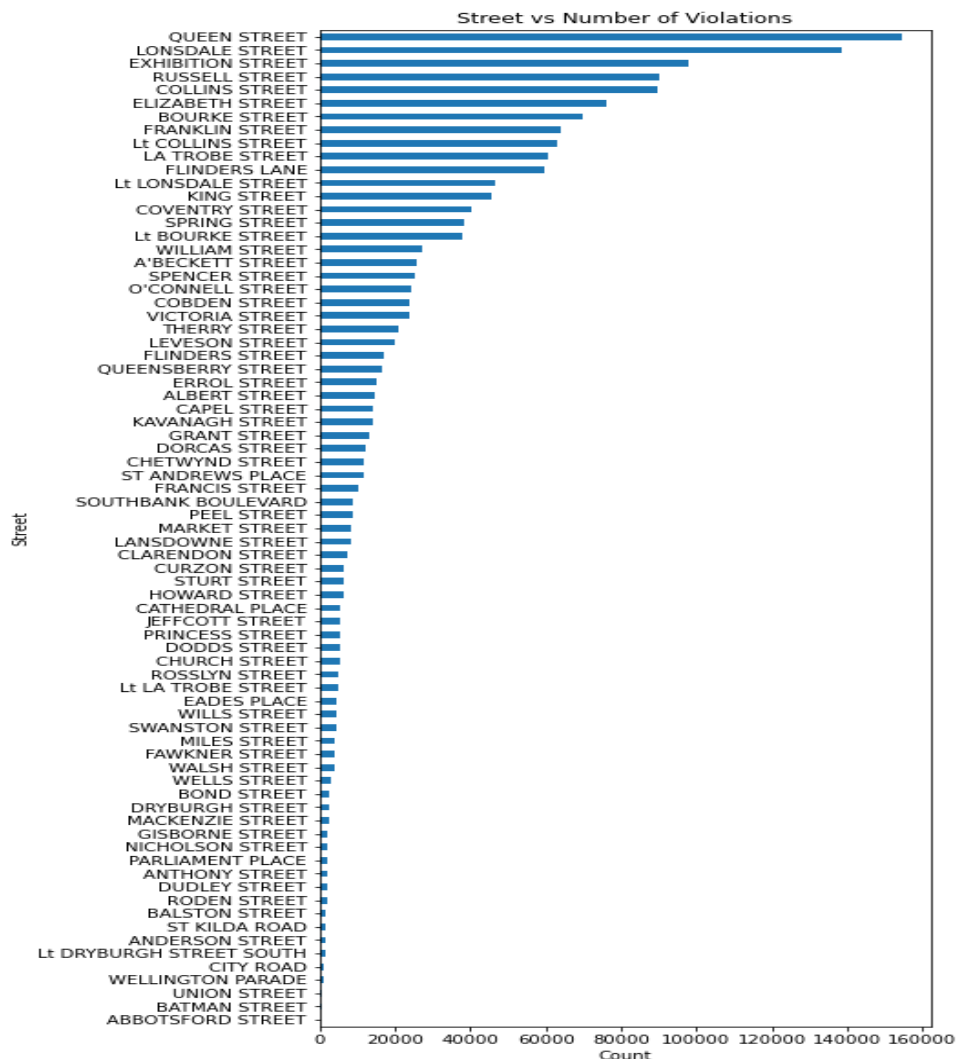
Solutions to this would include:

People can avoid infringements if they can purchase more parking time than permitted on the sign if they have commitments.

Parking restrictions will be beneficial to maximise parking spaces. For example, ¼ P Mon - Fri 8:30 AM to 3:30 PM.

4.2.2 Street Exploration

Figure 5 - Number of Violations per street



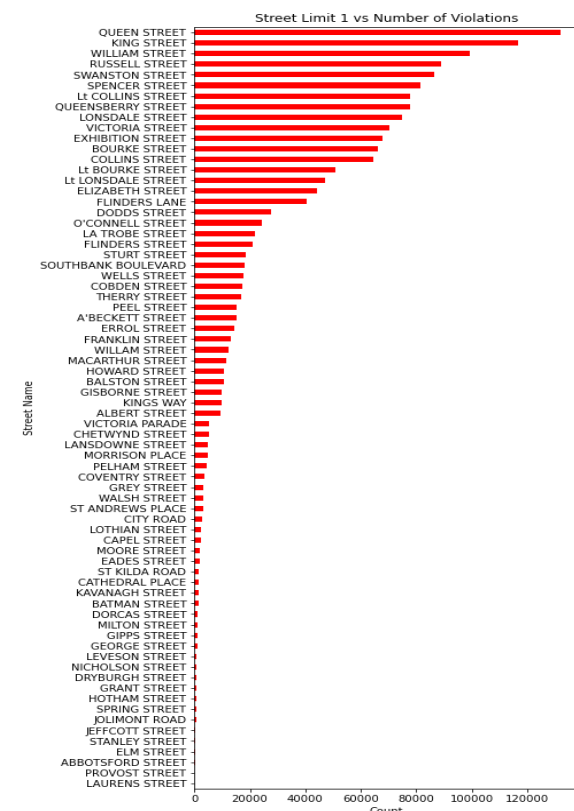
As shown in Figure 5, the graph above shows the numbers of parking violations in each street, the y-axis shows where the violations occurred. The x-axis represents the total count of violations that occurred in streets. Queen Street clearly reveals that it has the most street car park violations with a count of exactly 160000. The next busiest street car that has parking violation appears to be on Lonsdale Street, indicating that these streets do not have enough parking spaces. The areas from Exhibition Street to Abbotsford Street show significant changes in street car parking violations and appear to have more parking spaces than Queen Street and Lonsdale Street.

Solutions to this would include:

Queen Street is one of the busiest areas in the Melbourne CBD. The Queen street mall is visited by thousands of people every day and the surrounding

areas consist of many buildings where people come to work as well. The same scenario applies to Lonsdale St, where vehicle owners try to find a place to park their vehicles and sometimes forget to pay attention to the signs and parking meters (for example 1hr , 4hr). To solve this problem it is important to do a survey of the parking spaces around Queens St and Lonsdale St and try to remove time limits or increase them for car parks near work areas from Morning to Afternoon (average work hours). Another solution would be to add more parking signs to remind people of the parking information.

Figure 6 - Street Limit 1 vs Number of violations



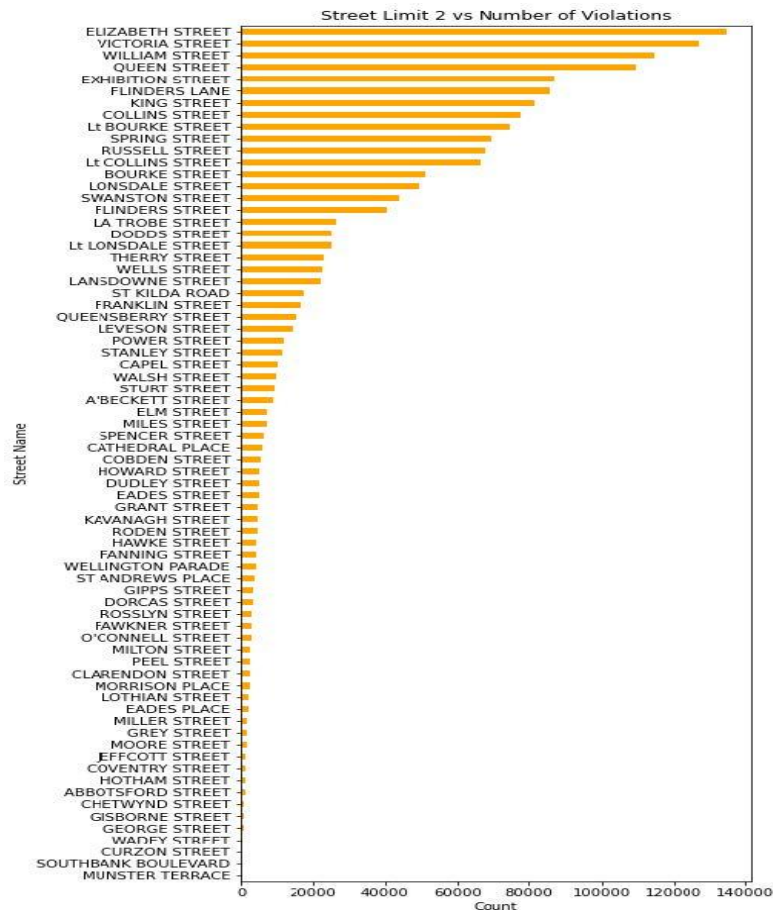
As shown in Figure 7, the graph shows the Street limit 1 in relation to the number of violations. The y-axis represents the street name, and the x-axis represents the number of violations. Queen Street, King Street, William Street, and Russell Street are the top four streets with the most violations.

Solutions to this would include:

One of the solutions is to do a survey of the parking spots surrounding Queen street, King street and William street and find out which ones are recording high amounts of violations. For example, the parking sign limits and the ticketing systems should be reviewed for the high priority parking areas.

Another solution is to remove or extend the parking time limit from parking signs near busy areas such as workplaces where people accidentally forget to change their parking spots to avoid getting fined.

Figure 7 - Street Limit 2 vs Number of violations



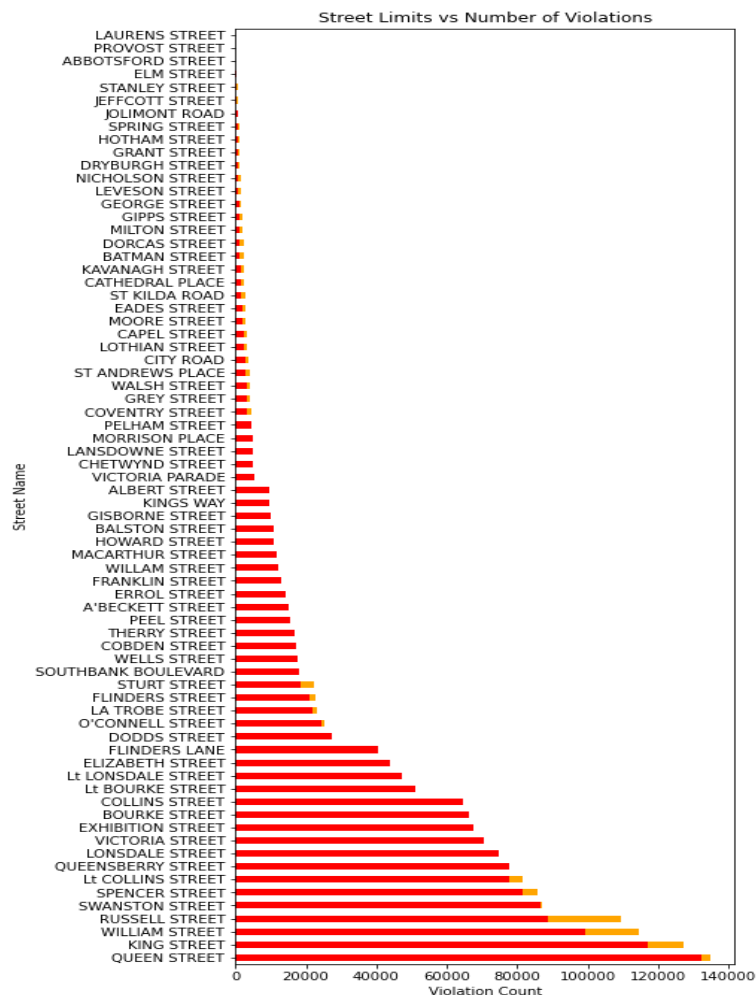
As shown in Figure 8, the graph shows the Street limit 2 in relation to the number of violations. The y-axis represents the street name, and the x-axis represents the number of violations. Lauren Street, Provost Street, Abbotsford Street, and Elm Street are the bottom four streets with the most violations.

Solutions to this would include:

One of the solutions is to do a survey of the parking spots surrounding Elizabeth street, Victoria street and William street and find out which ones are recording high amounts of violations. For example, the parking sign limits and the ticketing systems should be reviewed for the high priority parking areas.

Another solution is to remove or extend the parking time limit from parking signs near busy areas such as workplaces where people accidentally forget to change their parking spots to avoid getting fined.

Figure 8 - Street Limits vs Number of violations



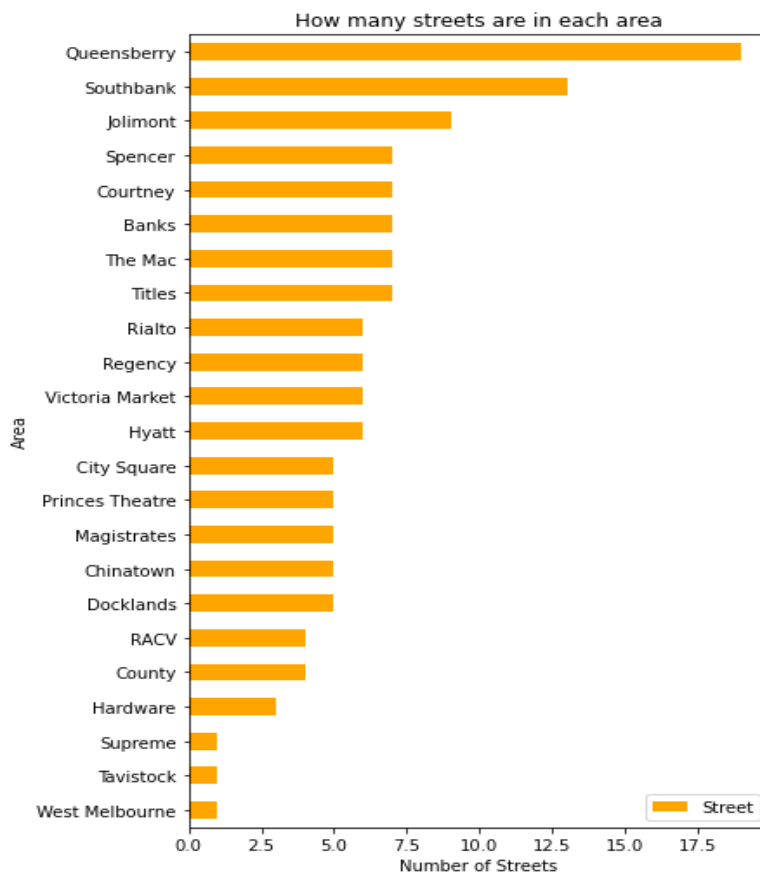
As shown in Figure 9, the graph indicates the number of parking violations in street limits 1 (red) and 2 (orange), with the y-axis indicating the streets in ascending order. The x-axis shows the total number of violations that occurred on the streets. From the figure we can see that the top 4 streets with the highest number of violations recorded more violations when being street limit 2 in comparison to when being street limit 1.

Solutions to this would include:

One of the solutions is to do a survey of the parking spots surrounding Queen street, King street and William street and find out which ones are recording high

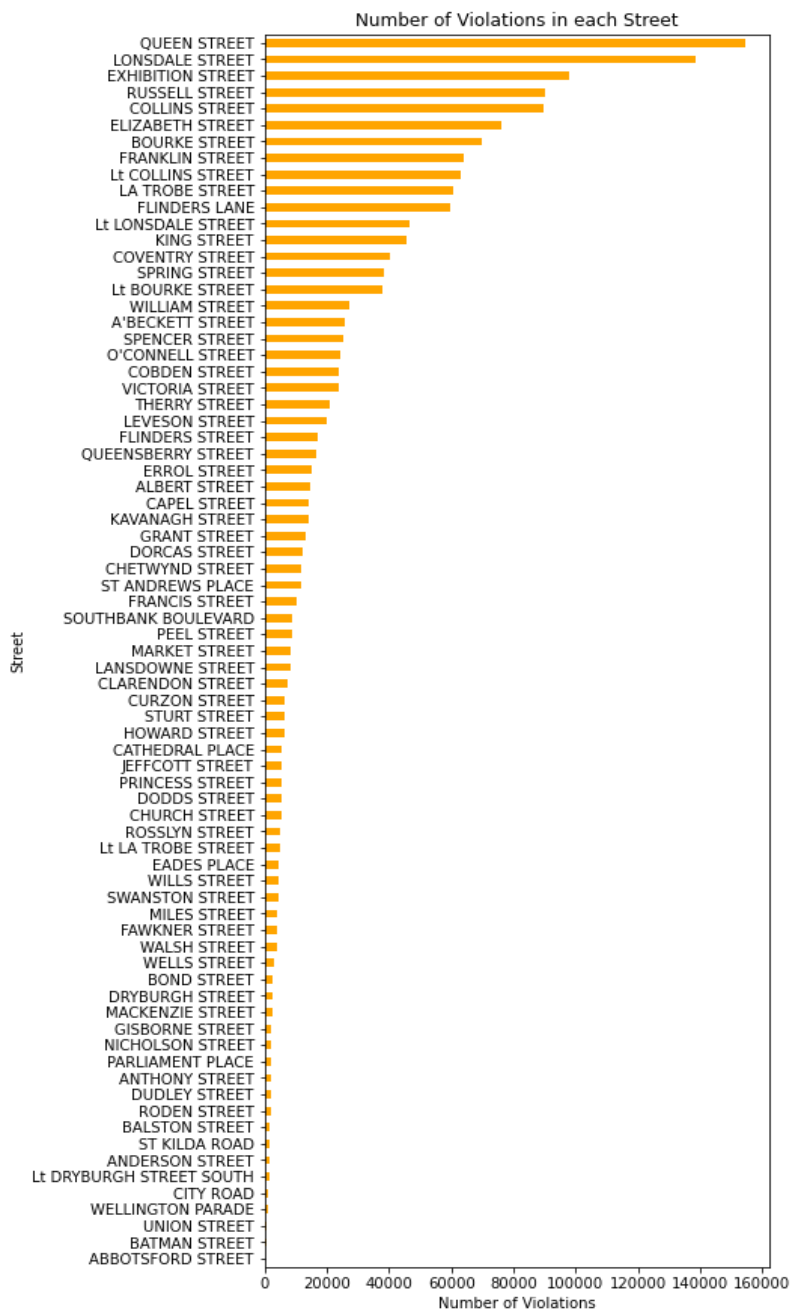
amounts of violations. For example, the parking sign limits and the ticketing systems should be reviewed for the high priority parking areas.

Figure 9 - How many streets in each area



As shown in Figure 10, the chart tells us how many streets are there in an area. The y- axis represents the name of an area, in conjunction to the x-axis which represents how many streets per area. Queensberry represents a high count since it is an area where there are many connecting streets, which consists of commercial buildings , houses and many side car parks with parking signs and meters displaying information.

Figure 10 - Number of violations in each street

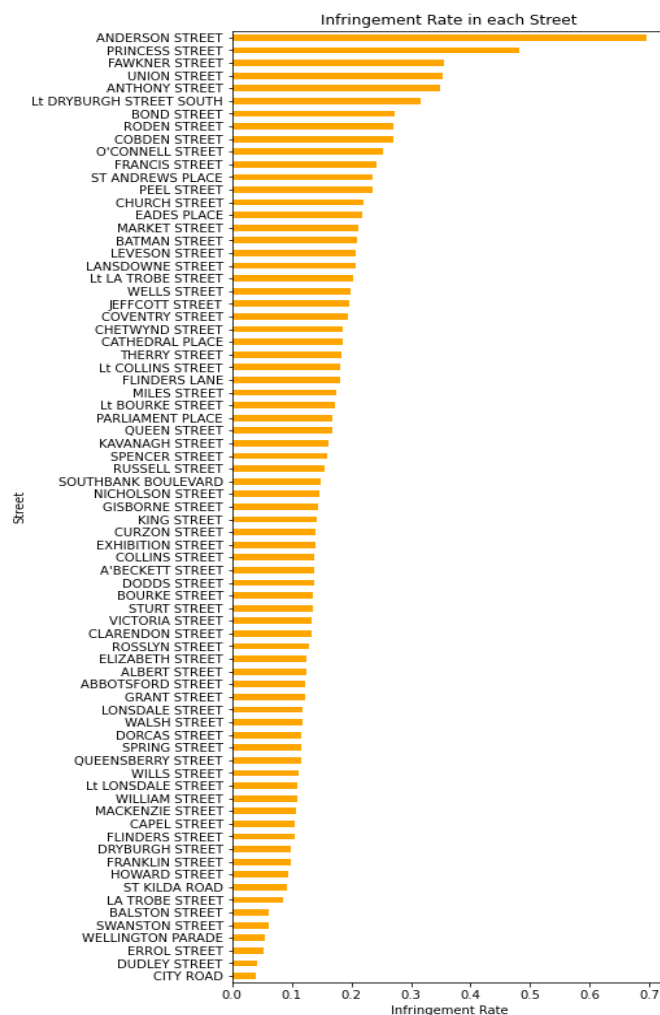


As shown in Figure 11, the graph above represents the number of violations that occurred in the streets that are shown in the y axis in descending order and the number of violations that are shown in the x axis. The Queen and Lonsdale streets recorded around 140000 - 160000 parking violations respectively making them the highest group in parking violations which indicates that there are no parking spaces on these streets, whilst the streets from Union to Abbotsford clearly shows that there are no parking violations recorded which shows that there is enough parking spaces.

Solutions to this would include:

One of the solutions is to do a survey of the parking spots surrounding Elizabeth street, Lonsdale street and Exhibition street and find out which ones are recording high amounts of violations. Never underestimate the time required displayed on the parking sign. For example, people are not allowed to park for more than one hour in a 1P signed area.

Figure 12 - Infringement rate per Street



As shown in Figure 12, the chart represents the Infringement rates in each street. The y-axis labels the street names and the x-axis labels the median value of the infringement rates. This exploration can determine that Anderson street has the highest rate of more than 0.6, while the majority of streets have less than 0.3 which indicates that Anderson Street is one of the busiest streets in Melbourne Metropolitan and does not have enough parking spaces for people that go to their work, university and other daily activities.

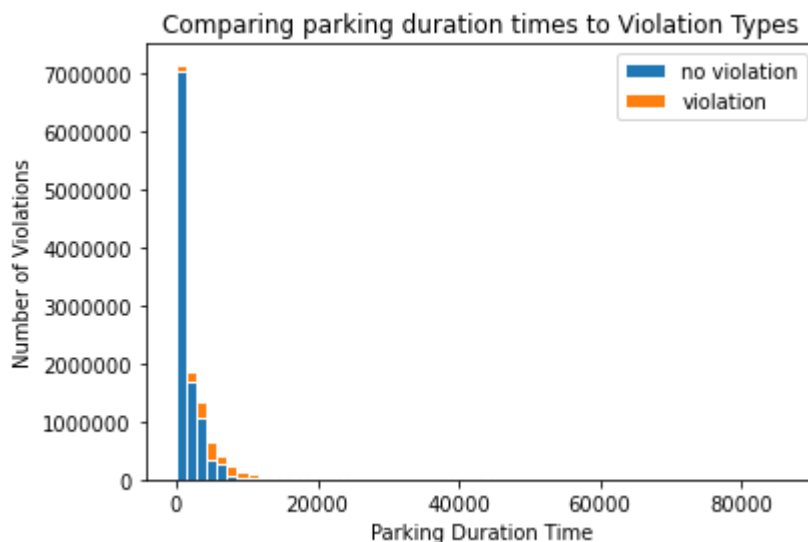
Solutions to this would include:

The main reason for a lot of violations are people driving to the city in their vehicles mainly for work and trying to find a proper parking space near their workplace. Some people get parking permits from their companies but some don't, therefore to avoid high parking fees they try to switch the parking spots before the time limit finishes so a parking fine won't be issued.

Therefore, some of the solutions are for VicRoads to make plans to extend/remove the time limits around the Anderson street area, while also making sure to reduce parking fees during busy times so normal working people can afford it.

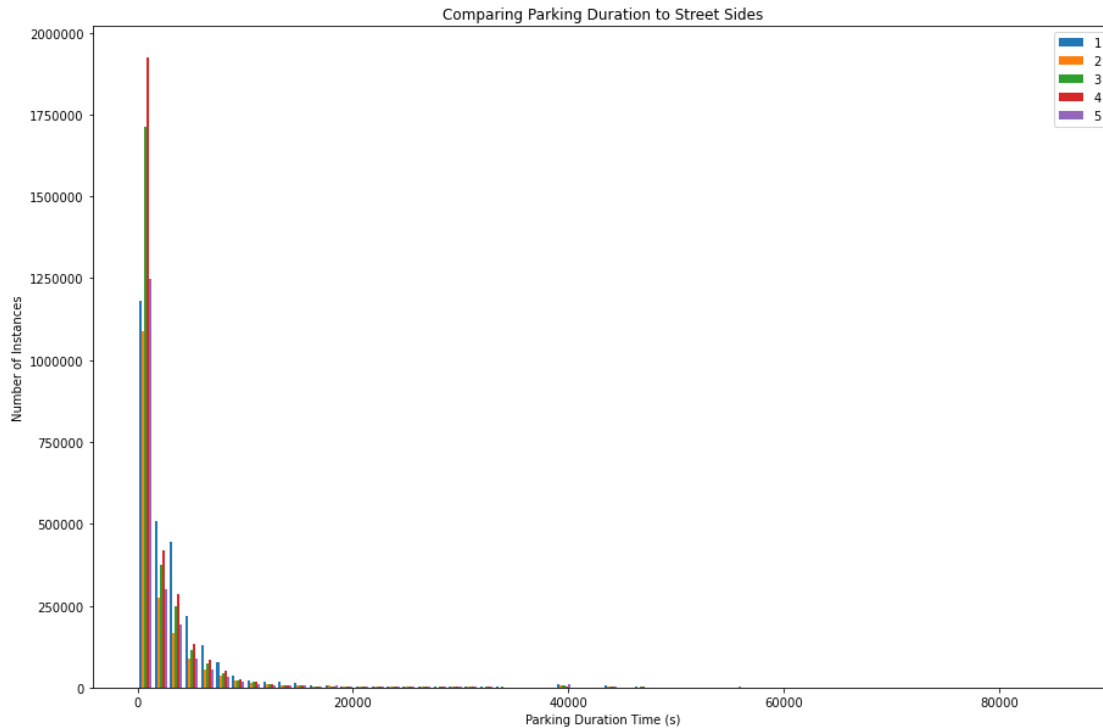
4.2.3 Parking duration exploration

Figure 13 - parking duration times vs violation types



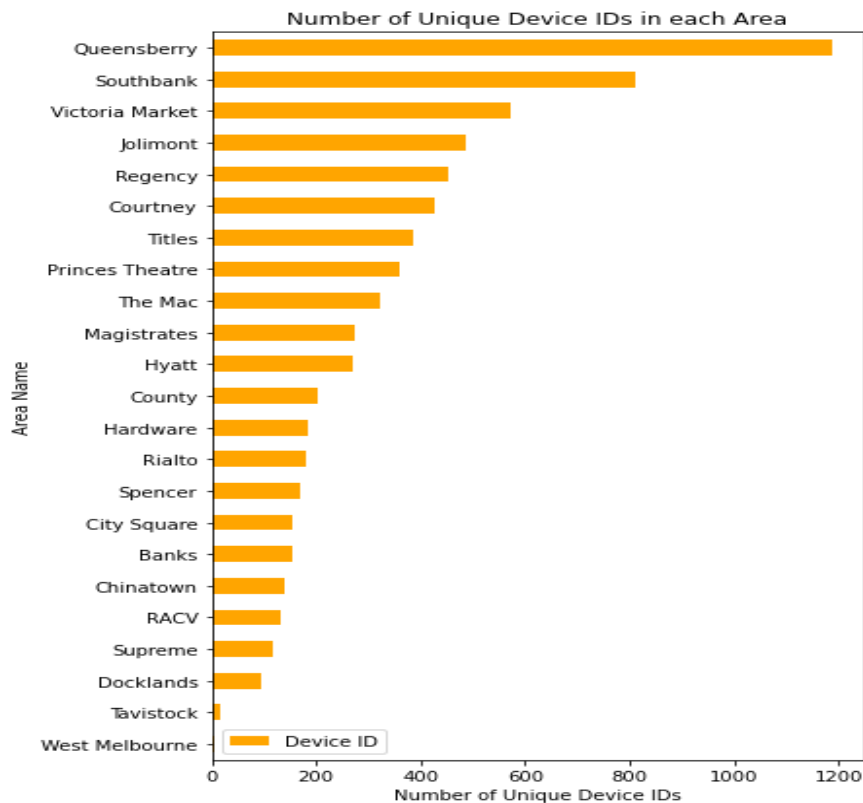
As shown in Figure 13, the y-axis labels the total number of violations that occurred, while the x-axis labels the total amount of time they were parked. The vast majority of parking events were less than 20000 seconds long. The majority of them lasted only a few seconds and recorded both violations and no violations. We see that the trend of the graph that occurred violation is relatively low. As a result, when planning a parking space, the council should prioritize short-term parking.

Figure 14 - parking duration vs street sides



From figure 14, we can see that the majority of parking events have extremely short durations, most of which occurred on sides 4 (~1900000), 3 (~1700000) and 5 (~1250000). While parking events with a longer duration were recorded more on side 1, as indicated by the blue lines. Based on that, the council should make more short-term parking space on sides 4, 3 and 5 while adding more parking space with longer duration on side 1.

Figure 15 - No. of unique device ID's in each area

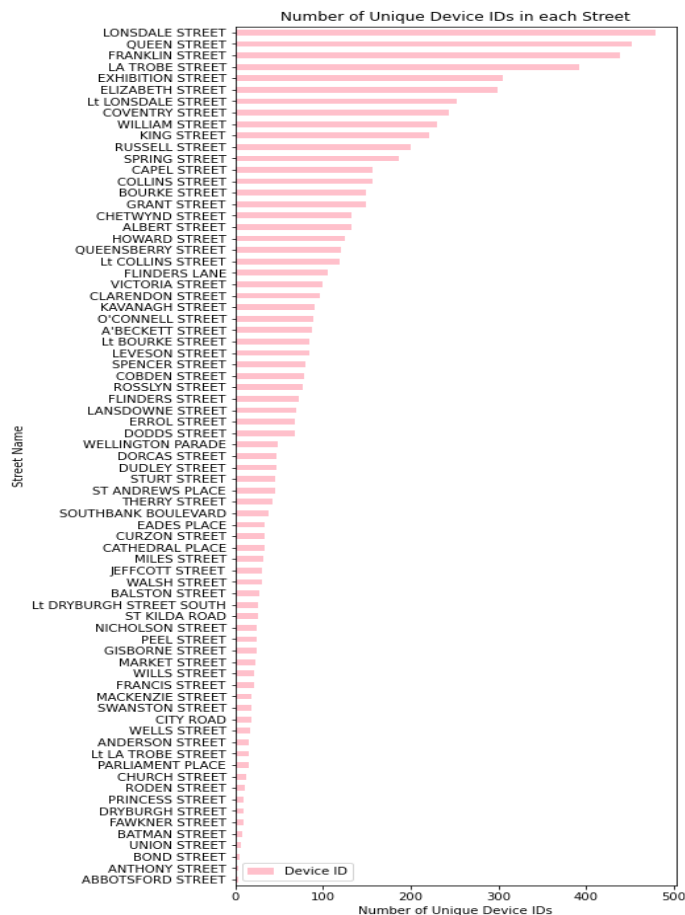


From figure 16, the graph above shows the total number of Device IDs recorded in each area. The y-axis displays the name of the area, and the x-axis displays the number of vehicle Device IDs. When compared to other parts of the city, the Queensberry area has nearly 1200 Device IDs, indicating a large number of vehicles parked in this area, while the rest of the streets from Southbank to Tavistock show a decreasing trend in terms of Device IDs, West Melbourne shows exactly 0 Device ID, indicating that these areas do not have parking events.

Solutions to this would include:

We can see that Lonsdale St. appears to be congested by vehicles who are working, buying food and visiting. As a result, some of the solutions include VicRoads making plans to increase the number of parking spaces in the Queensberry Street district. This proposal will help tremendously to avoid traffic congestion in one area and also helps to avoid parking infringement so that regular working people have a chance to switch parking spots and can avoid high parking fees before the time limit finishes so a parking fine won't be issued.

Figure 16 - No of unique device ID in each street



As shown in Figure 17, the chart above shows the total number of Device IDs in each area. The y-axis shows the street name where the trend of the graph tends to be decreasing, while the x-axis displays the number of Device IDs of vehicles parked on Melbourne streets. Lonsdale Street has approximately 500 Device IDs, which is a very high number of vehicles parked in comparison to the rest of the City Streets.

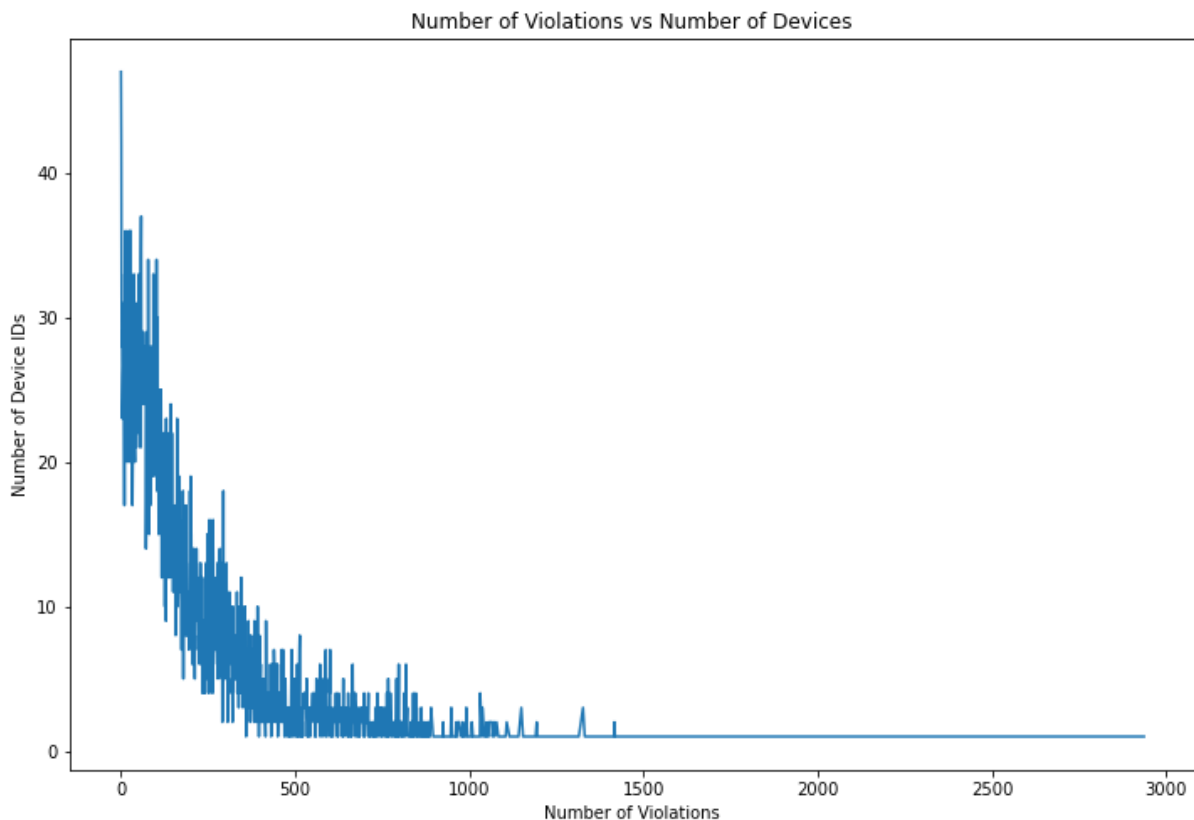
The Streets from Queen to Elizabeth recorded approximately between 300 to 450 Device ID which indicates the possibility of a large number of parking spaces. We can see that there is a decreasing trend in graph in Device Id after Elizabeth Street, which means the streets get less busy compared to the top 5 Streets which due to a lack of available parking events

Solutions to this would include:

We can see that Lonsdale Street, Queen Street, Franklin Street and La Trobe Street appear to be congested by vehicles who are working, buying food and going to the mall. As a result, some of the solutions include

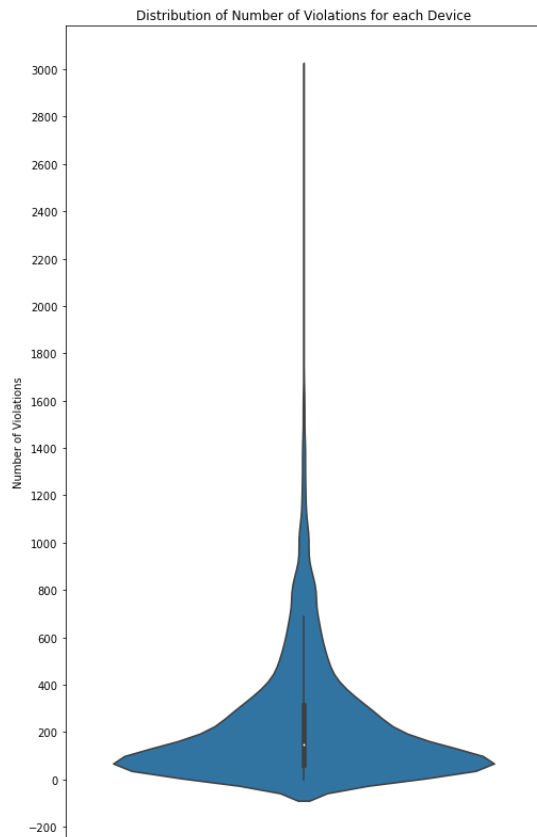
VicRoads making plans to increase parking spaces from Lonsdale Street to La Trobe Street. This also helps to avoid parking infringement so that regular working people have a chance to switch parking spots and can avoid high parking fees before the time limit finishes so a parking fine won't be issued and this solution will help tremendously in avoiding traffic congestion in one area.

Figure 17 - No of violations in each street



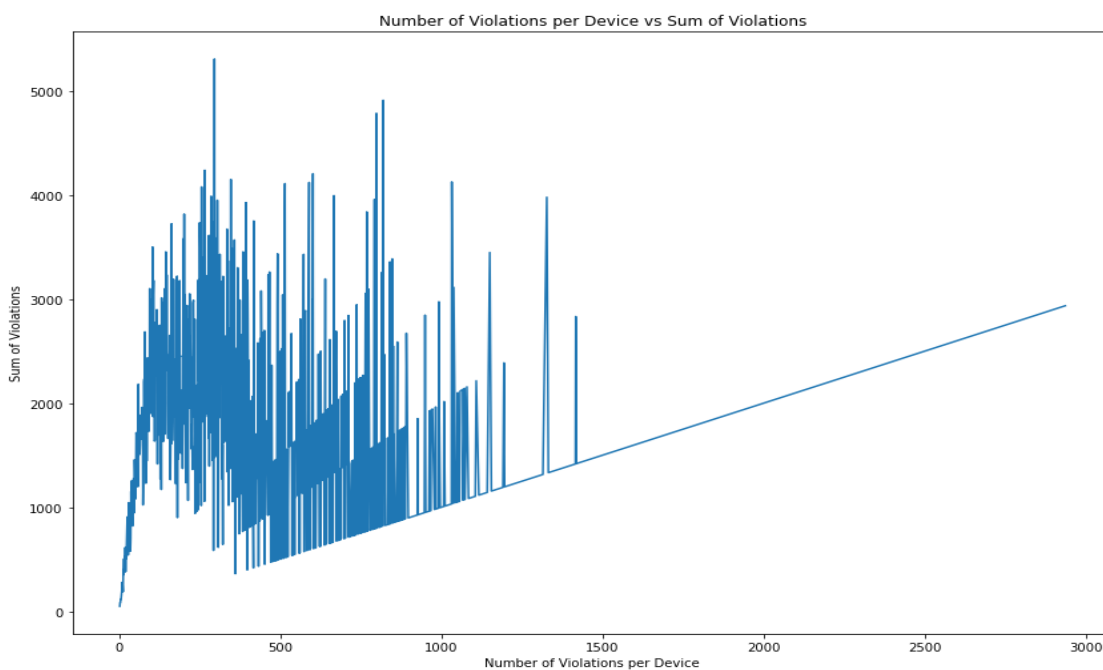
As shown in figure 18, the majority of devices which committed a violation committed less than 500 violations each, while there are a small number of devices that committed far more violations than the rest, with the highest number being close to 3000.

Figure 18 - Distributions of No of violations for each device



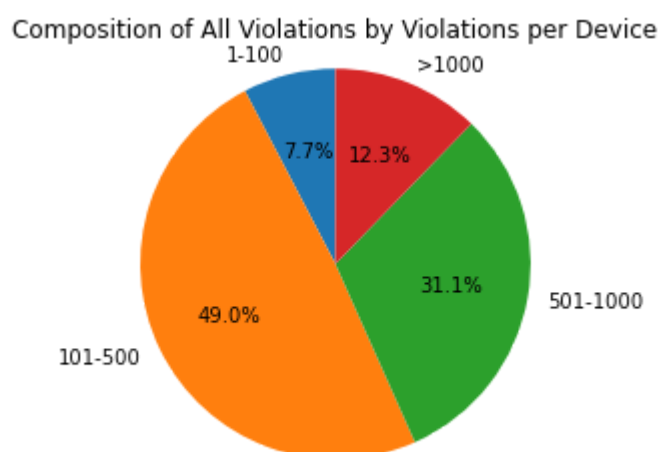
From figure 19, we can see that the majority of vehicles that are involved in a violation have less than 200 violations per vehicle, with a median value of around 150 and inter quartile range of ~50 to ~300.

Figure 19 - Number of Violations per Device vs Sum of Violations



According to Figure 20, we can see that the majority of infringements were committed by devices which committed 100-1000 violations each.

Figure 20 - Composition of all violations by violations per device



The pie chart in Figure 21 represents percentages of parking violations occurring per device ID's of vehicles. As we can see, vehicles which committed 101-500 violations each contributed the largest percentage in the total number of violations **49.0%(orange)**. Devices which committed 501-1000 violations each take up **31.1%(green)**, which is the second highest. Vehicles with more than 1000 violations each report **12.3%(Red)** and less than 100 report the lowest which is **7.7% (blue)**, which makes them the lowest entities in the chart for parking violations. A colourful pie chart like this is very insightful and saves time when trying to find a certain device ID violation statistic because it narrows down the search and makes it easier for users to search the dataset.

4.2.4 Individual Area Analysis

In this section we will be looking at all areas of Melbourne that have violations and determine what are the common factors in these violations.

4.2.4.1 Queensberry

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
45	ERROL	QUEENSBER	VICTORIA	2	01/10/2011	01/10/2011

Queensberry	STREET	RY STREET	STREET		12:01:29 AM	01:33:36 AM
661 Queensberry	LEVESON STREET	QUEENSBER RY STREET	VICTORIA STREET	1	01/10/2011 07:30:00 AM	01/10/2011 12:30:00 PM
672 Queensberry	CHETWYND STREET	VICTORIA STREET	STANLEY STREET	2	01/10/2011 07:30:07 AM	01/10/2011 12:30:00 PM
673 Queensberry	CHETWYND STREET	VICTORIA STREET	STANLEY STREET	2	01/10/2011 07:30:07 AM	01/10/2011 12:30:00 PM
674 Queensberry	CHETWYND STREET	VICTORIA STREET	STANLEY STREET	5	01/10/2011 07:30:07 AM	01/10/2011 12:30:00 PM

Queensberry is currently the place with the most violations, and the most parking, so it's important to have a look at the statistics of what is happening to determine the best course of action.

Finding the number of streets that recorded violations

Street Name	Violations
VICTORIA STREET	21761
LEVESON STREET	19703
ERROL STREET	15151
CAPEL STREET	14241
QUEENSBERRY STREET	13056
WILLIAM STREET	12591
CHETWYND STREET	11787
CURZON STREET	6308
HOWARD STREET	6254
ROSSLYN STREET	4911
KING STREET	4654
EADES PLACE	4546
WALSH STREET	3831

DRYBURGH STREET	2386
DUDLEY STREET	1885
RODEN STREET	1837
Lt DRYBURGH STREET SOUTH	1411
UNION STREET	769
ABBOTSFORD STREET	163

The street importance will be ranked as the following:

- High Priority
- Medium Priority

High priority streets have more than 10,000 violations, medium priority streets have 4,000-10,000 violations, and low priority streets have fewer than 5,000 violations. This means that funding for recommendations will be based on how high priority the streets in these specific areas are (also in comparison to the violation numbers in other areas). For now, let's only focus on high priority areas.

High priority streets that recorded violations in Queensberry

Side of Street	Violations	Percentage
1	31962	29.515191
2	30457	28.125404
5	17519	16.177856
3	15341	14.166590
4	13011	12.014960

From this we can therefore see that the violation types are mostly to do with side 1 and 2. Therefore if budget is an issue, recommendations to this budget will pertain to mostly side 1 and 2. If budget allows, side 5 will also be applicable.

We will now observe the parking signs to determine whether the signs should allow for more time, or if other parking solutions should be employed.

Testing worst ranked signs in Queensberry

Sign Num	Sign	Count	Parking Duration (s)		Sign Num	Sign	Count	Parking Duration (s)
One	1P TKT A-M-F 7:30-18:30	39532	998		Two	4P MTR M-F 7:30-18:30	39533	1798
		39533	560				39532	888
		39531	86				39531	107
		3965	18				39590	32
		3941	17				39589	29
		*****	*****				*****	*****
		23532	1				16484	1
		7224	1				18533	1
		17467	1				30827	1
		15420	1				16492	1
		25613	1				22525	1
	6307 rows x 1 cols					4936 rows x 1 cols		
	Max Length Stayed in hours	11				Max Length Stayed in Hours	11	
Three	1/4P M-F 7:30-18:30	39533	58		Four	1/4P 1P MTR RPA M-SUN 7:30-23:00	55792	963
		39532	39				55793	398
		1246	20				55794	8
		1321	19				6479	4
		1237	18				4081	4
		*****	*****				*****	*****
		10481	1				6386	1
		2285	1				19276	1
		6375	1				25421	1
		4326	1				25425	1
		5546	1				12286	1
	3409 rows x 1 cols					3969 x 1 cols		
	Max Length Stayed in hours	11				Max Length in hours	15.5	
Five	1P MTR M-SAT 7:30-18:30	4183	10		Six	1P MTR M-F 7:30-18:30	4100	10
		3953	9				39532	10
		3956	8				3985	8
		4009	8				4089	8
		4162	8				4095	8
		*****	*****				*****	*****
		9459	1				5758	1
		9477	1				10667	1
		5383	1				7807	1
		5387	1				5762	1
		4152	1				5350	1
	3024 rows x 1 cols					2714 rows x 1 cols		
	Max length in hours					Max length in hours	11	

A lot of the signs tend to have a really high margin for parking durations, where people stay for almost an entire day and subsequently get fined. This is irrelevant to the types of ticketing machine or charge used. These are considerable findings. This means that people, on average, during the working day do not stay for only 1 hour. Instead, they stay parked in these areas for more than 10 hours. This indicates that possible these individuals park their car for the work day (let's assume a 9-5 role). This takes away 8 hours already. Then let's assume a buffer time of 30 minutes to park and travel to work, and 30 minutes after work to get to the car. This adds an extra hour. So 9/11 hours are accounted for workers. Another 2 hours can include post-work outings, drinks or shopping.

Now this area is ticketed, which means they are paying for a spot, but are unable to continue paying every hour or move their car, furthering the point that they are probably working people. Another factor, as seen by graphs previous to this, is that the car IDs are constantly occurring in all areas. This indicates furthermore

that these drivers are regulars in the area, and that they park during the morning and afternoon of a work week.

Solutions to this would include:

Removing the 1 and 1/2 hour limits from Monday to Fridays, but keeping the ticketing system. This allows workers to use this space for their entire shift but allows for the owners of the parking spaces to still make money from the ticketing system. There can be instances where not all issues of this sign have to change, as keeping a few will allow others coming into the city to drop off packages or to complete other quick errands. This will also allow people to use more sustainable, eco-friendly options like cycling, taking public transport or using ride-share apps. Removing this limit also allows for people to park safely when they are attending a night out in the city, without worrying about parking fines. Keeping limits for Sundays and Saturdays, since 'day out' attendance can be supplemented by using public transport.

4.2.4.2 Princes Theatre

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
27 Princes Theatre	BOURKE STREET	RUSSELL STREET	EXHIBITION STREET	4	01/10/2011 12:01:24 AM	01/10/2011 03:36:42 AM
32 Princes Theatre	Lt BOURKE STREET	RUSSELL STREET	EXHIBITION STREET	4	01/10/2011 12:01:25 AM	01/10/2011 02:19:48 AM
40 Princes Theatre	Lt BOURKE STREET	RUSSELL STREET	EXHIBITION STREET	4	01/10/2011 12:01:28 AM	01/10/2011 06:21:31 AM
43 Princes Theatre	Lt BOURKE STREET	RUSSELL STREET	EXHIBITION STREET	4	01/10/2011 12:01:29 AM	01/10/2011 02:35:09 AM
60 Princes Theatre	BOURKE STREET	RUSSELL STREET	EXHIBITION STREET	4	01/10/2011 12:01:56 AM	01/10/2011 11:59:00 PM

Violations

Area	Violation	Street ID	Device ID
27 Princes Theatre	1	123	1401
32 Princes Theatre	1	907	1294
40 Princes Theatre	1	907	1305
43 Princes Theatre	1	907	1291
60 Princes Theatre	1	123	1392

Princess Theatre is the area with the second highest rate of infringements, closely behind Queensberry.

Side of Street	Violations	Percentage
1	62004	49.321476
4	32988	26.240514
3	19263	15.322876
5	6495	5.166489
2	4964	3.948645

Therefore, the biggest worry is side number 1, with an almost 50% share in violations. Side 4 is also considerable, covering 25% of the total violations. Both of these sides will be looked at, and sides 3,5 and 2 will not be considered.

Testing worst ranked signs in Princess Theatre

Sign Num	Sign	Count	Parking Duration (s)
Six	P10 M-SUN 0:00-23:59	921	14
		938	13
		926	11
		939	11
		942	10
	
		9469	1
		7418	1
		4083	1
		1269	1
		9600	1
	2691 rows x 1 columns		
	max length of stay	23.96666666666665	Hours

Another interesting observation has been made, specifically in the case of 'P10 M-SUN 0:00-23:59', where cases stayed for almost a full day. This could be indicative of those during the night life of the city, staying for longer periods of time, drinking and subsequently not being able to drive back home. This is one of multiple scenarios that could occur. But this is deduced given that this sign also applies to the Sunday, whereas most other signs do not.

One of the solutions can be removing the time limit from sign six on Sundays, therefore parking infringements can be reduced since accidental parking seems to be the common cause for many infringements in this case.

4.2.4.3 Southbank

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
453929 Southbank	DODDS STREET	SOUTHBANK BOULEVARD	GRANT STREET	5	14/10/2011 08:42:18 AM	14/10/2011 12:35:33 PM
455210 Southbank	DODDS STREET	SOUTHBANK BOULEVARD	GRANT STREET	2	14/10/2011 09:07:26 AM	14/10/2011 06:30:00 PM
456678 Southbank	DODDS STREET	SOUTHBANK BOULEVARD	GRANT STREET	2	14/10/2011 09:31:00 AM	14/10/2011 12:53:12 PM
457764 Southbank	DODDS STREET	SOUTHBANK BOULEVARD	GRANT STREET	2	14/10/2011 09:49:14 AM	14/10/2011 01:09:33 PM
458316 Southbank	DODDS STREET	SOUTHBANK BOULEVARD	GRANT STREET	2	14/10/2011 09:57:27 AM	14/10/2011 02:00:43 PM

Violations that occurred in the above areas in Southbank

Area	Violation	Street ID	Device ID
453929 Southbank	1	591	4465
455210 Southbank	1	591	4661
456678 Southbank	1	591	4639
457764 Southbank	1	591	4705
458316 Southbank	1	591	4623

Side of Street	Violations	Percentage
3	22214	56.260764
5	8060	20.413332
2	5991	15.173235
4	3219	8.152669

It appears as there are only 4 counts inside the street. The main areas of focus include sides 3 and 5, given side 3 counts for more than 50% of the violations. When incorporating side 5, we pass a violation rate of 75% and so this will also be included

Testing worst ranked sign in Southbank

Sign Num	Sign	Count	Parking Duration (s)
One	1P TKT A RPA M-F 7:30-18:30	39593	2369
		39592	1123
		39591	107
		39590	36
		39589	12
		*****	*****
		14401	1
		6213	1
		32838	1
		30793	1
		4094	1
	11020 rows x 1 columns		
	max length of stay	11.0 Hours	

This also confirms like in the other cases, where the violations are for a full work day give or take another couple of hours. This will only be tested on the above 10,000 violation figures as it is more of a confirmation about the pattern being found in many of these areas.

Southbank is a busy area both weekdays and weekends since it is home to the notable crown Melbourne Casino and Eureka Skydeck. Also other visitable places such as the National Gallery of Victoria, Melbourne Recital Centre, Arts Centre Melbourne etc.. This seems just like the Queensberry area where there is a lot of activity, therefore the parking sign 'one - 1p m-f 7:30am - 6:30pm records a high count of violations. This also might be people who come to the city to work. As a solution the 1-hour limit of parking signs around Southbank should be removed during 7:30 am to 6:30 pm, therefore this will lead to less parking violations.

4.2.4.4 Hyatt

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
19 Hyatt	SPRING STREET	FLINDERS LANE	FLINDERS STREET	2	01/10/2011 12:01:20 AM	01/10/2011 11:59:00 PM
35 Hyatt	Lt COLLINS STREET	EXHIBITION STREET	SPRING STREET	4	01/10/2011 12:01:25 AM	01/10/2011 05:43:12 AM
98 Hyatt	FLINDERS STREET	EXHIBITION STREET	SPRING STREET	3	01/10/2011 12:36:28 AM	01/10/2011 12:39:33 AM
106 Hyatt	FLINDERS STREET	EXHIBITION STREET	SPRING STREET	3	01/10/2011 12:44:02 AM	01/10/2011 12:50:16 AM
302 Hyatt	FLINDERS STREET	EXHIBITION STREET	SPRING STREET	3	01/10/2011 04:50:52 AM	01/10/2011 04:55:55 AM

Violations that occurred in High Priority areas in Hyatt

Street	Violation(s)
EXHIBITION STREET	47398
SPRING STREET	16587
COLLINS STREET	14656
Lt COLLINS STREET	14552
FLINDERS LANE	11729
FLINDERS STREET	5877

Given this, the high priority streets include:

- Exhibition
- Sprint
- Collins
- Lt Collins Street
- Flinders Lane

According to the statistics above it seems that Exhibition Street tops the number of violations with a count of 47398. Exhibition street is a big street in the Melbourne CBD that is surrounded by many notable theatres, bars, banks and many workplaces which requires many people to find a suitable parking spot for parking their vehicles. This factors into the high amount of parking violations.

Side of Street	Violations	Percentage
1	35676	34.002402
4	23874	22.754046
5	17914	17.073636
3	17063	16.262557
2	10395	9.907360

Testing worst ranked sign in Hyatt

Sign Num	Sign	Count	Parking Duration (s)
One	1P MTR M-SAT 7:30-19:30	43193	296
		43192	215
		3941	49
		4057	46
		3985	45
		*****	*****
		28418	1
		24292	1
		26339	1
		32480	1
		11653	1
8998 rows x 1 columns			
max length of stay		12.0 Hours	

From this we can see that the most problematic car parking spot is again a "weekday" slot of 1 hour, with a large number of parking durations of around 12 hours. The solution would be removing the 1-hour limit from Monday to Friday.

4.2.4.5 City Square

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
30 City Square	RUSSELL STREET	COLLINS STREET	COLLINS STREET	1	01/10/2011 12:01:24 AM	01/10/2011 12:03:07 AM
34 City Square	FLINDERS LANE	SWANSTON STREET	RUSSELL STREET	4	01/10/2011 12:01:25 AM	01/10/2011 01:10:28 AM
56 City Square	FLINDERS LANE	ELIZABETH STREET	SWANSTON STREET	4	01/10/2011 12:01:35 AM	01/10/2011 05:38:30 AM
61 City Square	RUSSELL STREET	COLLINS STREET	FLINDERS LANE	1	01/10/2011 12:02:02 AM	01/10/2011 12:04:28 AM
62 City Square	FLINDERS LANE	ELIZABETH STREET	SWANSTON STREET	4	01/10/2011 12:04:46 AM	01/10/2011 10:54:39 AM

Violations occurred in the above areas

Area	Violation	Street ID	Device ID
30 City Square	1	1221	479
34 City Square	1	669	1417
56 City Square	1	669	1366
61 City Square	1	1221	478
62 City Square	1	669	1419

--	--	--	--

Number of times street recorded violations

Street	Violation(s)
RUSSELL STREET	32595
COLLINS STREET	31249
ELIZABETH STREET	25455
FLINDERS LANE	4835
FLINDERS STREET	2867

As from the above table, Russell street tops the number of violations occurring at a value of 32595.

The percentage table below also shows a rate of 28.884982 % of violations for the third side of street, therefore we can determine that it is the busiest side of street with an amount of 25794 violations

Side of Street	Violations	Percentage
3	25794	28.884982
2	21640	24.233194
5	19614	21.964412
1	16796	18.808721
4	5455	6.108691

Testing the worst ranked sign in City Square

Sign Num	Sign	Count	Parking Duration (s)
One	1/2P MTR M-SAT 7:30-19:30	43193	232
		43192	94
		2121	32
		2169	31
		2123	31
	
		5875	1
		11647	1
		14032	1
		5859	1
		26429	1
	6697 rows x 1 columns		
	max length of stay	12.0 Hours	

The highest 2 value counts have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 1 and 2 hours limit from Monday to Friday.

4.2.4.6 Victoria Market

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
301 Victoria Market	QUEEN STREET	VICTORIA STREET	THERY STREET	2	01/10/2011 04:50:24 AM	01/10/2011 05:29:19 AM
330 Victoria Market	QUEEN STREET	VICTORIA STREET	THERY STREET	2	01/10/2011 05:30:27 AM	01/10/2011 05:34:16 AM
359 Victoria Market	QUEEN STREET	THERY STREET	FRANKLIN STREET	5	01/10/2011 06:00:07 AM	01/10/2011 07:42:19 AM
362 Victoria Market	QUEEN STREET	VICTORIA STREET	THERY STREET	1	01/10/2011 06:00:07 AM	01/10/2011 09:56:38 AM
365 Victoria Market	QUEEN STREET	VICTORIA STREET	THERY STREET	5	01/10/2011 06:00:07 AM	01/10/2011 09:21:28 AM

Finding the number of streets that recorded violations:

Street	Violation(s)
QUEEN STREET	36746
FRANKLIN STREET	31228
THERY STREET	20913
ELIZABETH STREET	3469
VICTORIA STREET	1949
PEEL STREET	1922

There are 3 high priority streets: Queen Street, Franklin Street and Therry Street.

Side of Street	Violations	Percentage
1	37130	41.772138
4	23369	26.290684
5	18531	20.847818
2	6215	6.992024
3	3642	4.097337

The street sides that also have the most violations are 1, 4 and 5. Which are

- Street side 1 - 37130
- Street side 4 - 23369
- Street side 5 - 18531

Testing the worst ranked signs in Victoria Market

Sign Num	Sign	Count	Parking Duration (s)
One	2P MTR M-SAT 7:30-20:30	46792	1494
		46793	515
		46791	132
		46790	52
		46789	34
		13762	1
		7623	1
		30402	1
		40401	1
		8196	1
	5701 rows x 1 columns max length of stay	13.0 Hours	
Sign Num	Sign	Count	Parking Duration (s)
two	SIGN TWO: 1/4P M-SAT 7:30-19:30	43193	34
		1326	24
		1246	23
		1319	23
		1240	23
		2812	1
		2804	1
		13043	1
		4839	1
		6586	1
	3395 rows x 1 columns max length of stay	12.0 Hours	
Sign Num	Sign	Count	Parking Duration (s)
three	1P MTR M-SAT 6:00-19:30	48593	123
		48592	102
		4042	12
		3951	11
		4372	11
		10854	1
		10846	1
		12885	1
		31316	1
		16384	1
	6848 rows x 1 columns max length of stay	13.5 Hours	

There were large numbers of violations with long parking durations (12 - 14 hours), which were likely contributed from those who came to the area for work. Thus, the solution would be removing 1, 2 and 4 hours limits from Monday to Friday.

4.2.4.7 The Mac

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
------	--------	----------------	----------------	----------------	--------------	----------------

21 The Mac	ELIZABETH STREET	LA TROBE STREET	Lt LONSDALE STREET	2	01/10/2011 12:01:21 AM	01/10/2011 12:27:44 AM
26 The Mac	A'BECKETT STREET	ELIZABETH STREET	SWANSTON STREET	4	01/10/2011 12:01:23 AM	01/10/2011 02:20:45 PM
91 The Mac	A'BECKETT STREET	ELIZABETH STREET	SWANSTON STREET	4	01/10/2011 12:31:45 AM	01/10/2011 12:51:04 AM
175 The Mac	ELIZABETH STREET	LA TROBE STREET	Lt LONSDALE STREET	2	01/10/2011 02:01:57 AM	01/10/2011 02:27:57 AM
222 The Mac	A'BECKETT STREET	ELIZABETH STREET	SWANSTON STREET	4	01/10/2011 02:56:20 AM	01/10/2011 03:47:13 AM

Violations record in the Mac area

Street	Violation(s)
FRANKLIN STREET	32577
LA TROBE STREET	15379
ELIZABETH STREET	15132
A'BECKETT STREET	11814
Lt LONSDALE STREET	10515
Lt LA TROBE STREET	4764
SWANSTON STREET	4222

There are 5 high priority streets: Franklin Street, La Trobe Street, Elizabeth Street, A'Beckett Street and Lt Lonsdale Street.

Testing worst ranked signs

Sign Num	Sign	Count	Parking Duration (s)
one	1P MTR M-SAT 7:30-19:30	43192	79
		43193	57
		3919	22
		4056	20
		3928	17
		8812	1
		29286	1
		8796	1
		6747	1
		14383	1
	6498 rows x 1 columns		
	max length of stay	12.0 Hour	
Sign Num	Sign	Count	Parking Duration (s)
two	2P MTR M-SAT 7:30-20:30	46792	130
		46793	41
		7511	15
		7572	14
		46791	14
		15988	1
		16209	1
		9839	1
		13933	1
		14572	1
	5850 rows x 1 columns		
	max length of stay	13.0 Hours	

The street sides that also have the most violations are 1, 4 and 5.

There were large numbers of violations with long parking durations (12 - 14 hours), which were likely contributed from those who came to the area for work. Thus, the solution would be removing 1 and 2 hours limits from Monday to Friday.

4.2.4.8 Courtney

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
2 Courtney	ELIZABETH STREET	PELHAM STREET	QUEENSBERY STREET	2	01/10/2011 12:01:11 AM	01/10/2011 11:00:19 AM
669 Courtney	ELIZABETH STREET	PELHAM STREET	QUEENSBERY STREET	2	01/10/2011 07:30:07 AM	01/10/2011 12:30:00 PM
670 Courtney	ELIZABETH STREET	PELHAM STREET	QUEENSBERY STREET	2	01/10/2011 07:30:07 AM	01/10/2011 12:18:15 PM
671 Courtney	O'CONNELL STREET	PEEL STREET	QUEENSBERY STREET	5	01/10/2011 07:30:07 AM	01/10/2011 08:54:20 AM
705 Courtney	ELIZABETH STREET	QUEENSBERY STREET	VICTORIA STREET	5	01/10/2011 07:30:07 AM	01/10/2011 12:30:00 PM

Violations occurred in Courtney Area

Street	Violation(s)
--------	--------------

O'CONNELL STREET	24507
COBDEN STREET	23904
ELIZABETH STREET	16590
PEEL STREET	6878
PRINCESS STREET	5418
QUEENSBERRY STREET	3551
CURZON STREET	32

There are 3 high priority streets: O'Connell Street, Cobden Street and Elizabeth Street.

Side of Street	Violations	Percentage
2	44206	68.008184
5	20795	31.991816

Testing the Worst Ranked Signs

SIGN ONE: 1/2P M-F 7:30-18:30

Value Counts for Parking Duration:

Parking Duration (s)

```

39593      59
39592      20
2185       17
2104       15
2178       14
...
4529       1
6576       1
4076       1
16795      1
6141       1

```

[3989 rows x 1 columns]

The time (hours) of the max length of stay is: 11.0 Hours

The highest 2 value counts have a duration of around 11 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 1/2 hours limit from Monday to Friday.

4.2.4.9 Chinatown

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
13 Chinatown	ELIZABETH STREET	Lt BOURKE STREET	BOURKE STREET	2	01/10/2011 12:01:18 AM	01/10/2011 06:13:22 AM

38 Chinatown	RUSSELL STREET	BOURKE STREET	Lt COLLINS STREET	5	01/10/2011 12:01:28 AM	01/10/2011 03:50:49 AM
39 Chinatown	RUSSELL STREET	BOURKE STREET	Lt COLLINS STREET	5	01/10/2011 12:01:28 AM	01/10/2011 12:58:51 AM
46 Chinatown	RUSSELL STREET	Lt BOURKE STREET	BOURKE STREET	2	01/10/2011 12:01:29 AM	01/10/2011 02:42:12 AM
55 Chinatown	ELIZABETH STREET	Lt BOURKE STREET	BOURKE STREET	2	01/10/2011 12:01:34 AM	01/10/2011 02:48:02 AM

Violations occurred in Chinatown

Street	Violation(s)
RUSSELL STREET	47229
ELIZABETH STREET	15316
LONSDALE STREET	10043
Lt COLLINS STREET	5501
BOURKE STREET	1021

There are 3 high priority streets: Russell Street, Elizabeth Street and Lonsdale Street.

Violations in side of street in Chinatown

Side of Street	Violations	Percentage
1	24450	33.683253
5	22874	31.512096
2	15221	20.969031
4	5298	7.298727
3	5298	6.536893

Around 33.68% of violations were recorded on side 1, followed by side 5 (31.51%) and 2 (20.97%).

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43200	89
43193	83
43192	46
4047	26
3930	24
...	...
12817	1
21005	1
24947	1
10754	1
9018	1

[5984 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

SIGN TWO: 1/2P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43193	281
43192	111
43200	72
2114	25
2177	25
...	...
7691	1
19969	1
17920	1
5626	1
4098	1

[5486 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

There were large numbers of violations with long parking durations (12 hours), which were likely contributed from those who came to the area for work. Thus, the solution would be removing 1 and 2 hours limits from Monday to Friday.

4.2.4.10 Banks

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
44 Banks	MARKET STREET	FLINDERS LANE	FLINDERS STREET	5	01/10/2011 12:01:29 AM	01/10/2011 02:15:40 AM
658 Banks	QUEEN STREET	FLINDERS STREET	FLINDERS LANE	1	01/10/2011 07:30:00 AM	01/10/2011 10:26:30 AM
693 Banks	MARKET STREET	COLLINS STREET	FLINDERS LANE	1	01/10/2011 07:30:07 AM	01/10/2011 10:15:58 AM
940 Banks	MARKET STREET	COLLINS STREET	FLINDERS LANE	1	01/10/2011 07:30:07 AM	01/10/2011 07:30:00 PM

1002 Banks	QUEEN STREET	COLLINS STREET	FLINDERS LANE	1	01/10/2011 07:30:08 AM	01/10/2011 10:03:07 AM
------------	--------------	----------------	---------------	---	------------------------	------------------------

77995 violations were recorded in the Banks area, which makes it the 10th highest among the 23 areas.

Street	Violation(s)
QUEEN STREET	28417
FLINDERS LANE	18759
COLLINS STREET	11985
MARKET STREET	8417
FLINDERS STREET	4074
WILLIAM STREET	3744
BOND STREET	2599

There are 3 high priority streets: Queen Street, Flinders Lane and Collins Street.

High priority streets that recorded violations in Banks:

Side of Street	Violations	Percentage
1	22547	50.694757
3	11990	26.958360
5	4705	10.578739
4	4069	9.148754
2	1165	2.619390

Side 1 recorded the highest number of violations, accounting for 50.69% of total violations, followed by side 3 (26.96%) and 5 (10.58%).

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43200	89
43193	83
43192	46
4047	26
3930	24
...	...
12817	1
21005	1
24947	1
10754	1
9018	1

[5984 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The majority of infringements have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution should be removing the 1 hour limit from Monday to Friday.

4.2.4.11 Titles

Area	Street	Street Limit 1	Street Limit 2	Side of street	Arrival time	Departure time
688 Titles	LA TROBE STREET	WILLIAM STREET	QUEEN STREET	3	01/10/2011 07:30:07 AM	01/10/2011 07:30:07
763 Titles	WILLS STREET	A'BECKETT STREET	LA TROBE STREET	2	01/10/2011 07:30:07 AM	01/10/2011 01:58:15 PM
764 Titles	A'BECKETT STREET	WILLIAM STREET	QUEEN STREET	4	01/10/2011 07:30:07 AM	01/10/2011 10:31:29 AM
853 Titles	A'BECKETT STREET	QUEEN STREET	ELIZABETH STREET	3	01/10/2011 07:30:07 AM	01/10/2011 07:30:08 AM
1069 Titles	LA TROBE STREET	QUEEN STREET	ELIZABETH STREET	3	01/10/2011 07:30:08	01/10/2011 09:44:01 AM

77814 violations were recorded in the Titles area, which makes it the 11th highest among the 23 areas.

Street	Violation(s)
LA TROBE STREET	25419
QUEEN STREET	15107
A'BECKETT STREET	14159
Lt LONSDALE STREET	12004
WILLIAM STREET	4907

WILLS STREET	4284
ANTHONY STREET	1934

There are 4 high priority streets: La Trobe Street, Queen Street, A'Beckett Street and Lt Lonsdale Street.

High priority streets that recorded violations in Titles:

Side of Street	Violations	Percentage
3	26550	39.811663
4	25032	37.535426
1	6248	9.368861
5	4986	7.476495
2	3873	5.807554

Side 3 recorded the highest number of violations, accounting for 39.81% of total violations, followed by side 4 with 37.53%.

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43193	63
43192	32
3987	15
4135	14
4203	12
...	...
12874	1
6727	1
10817	1
27979	1
7648	1

[4597 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The majority of infringements have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution should be removing the 1 hour limit from Monday to Friday.

4.2.4.12 Hardware

77459 violations were recorded in the Hardware area, which makes it the 12th highest among the 23 areas.

Finding the number of streets that recorded violations

Street Name	Violations
QUEEN STREET	49466
LONSDALE STREET	14717
Lt BOURKE STREET	13276

All of these streets have significant levels of violations, and are of high priority.

High priority streets that recorded violations in Hardware:

Side of Street	Violations	Percentage
1	50220	64.834299
4	15019	19.389613
5	7551	9.748383
2	3665	4.731535
3	1004	1.296170

Side 1 recorded the highest number of violations, accounting for 64.83% of total violations, followed by side 4 with 19.39%.

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43192	481
43193	321
43191	58
3901	53
4036	51
...	...
20666	1
8795	1
8404	1
16600	1
25091	1

[8589 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The majority of infringements have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus the solution should be removing the 1 hour limit.

4.2.4.13 Spencer

76344 violations were recorded in the Spencer area, which makes it the 13th highest among the 22 areas.

Finding the number of streets that recorded violations

Street Name	Violations
Lt COLLINS STREET	16222
FLINDERS LANE	13938
BOURKE STREET	12189
SPENCER STREET	11194
FRANCIS STREET	10477
COLLINS STREET	8092
FLINDERS STREET	4232

There are 5 high priority: Lt Collins street, Flinders Lane, Bourke Street, Spencer Street and Francis Street.

High priority streets that recorded violations in Spencer:

Side of Street	Violations	Percentage
4	28921	53.247781
3	14199	26.142431
5	5825	10.724675
2	5369	9.885112

Side 4 recorded the highest number of violations, accounting for 53.25% of total violations, followed by side 3 with 26.14%.

Testing the Worst Ranked Signs

SIGN ONE: 1P TKT A M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

3944	20
3942	15
3986	13
4016	12
4116	12
...	...
7368	1
9594	1
7344	1
9391	1
6081	1

[5120 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The durations of the highest value counts are slightly above the limit of 1 hour. The solution would be to increase the duration limit to 1.5 hours.

4.2.4.14 Rialto

74588 violations were recorded in the Rialto area, which makes it the 14th highest among the 23 areas.

Finding the number of streets that recorded violations

Street Name	Violations
-------------	------------

KING STREET	27627
COLLINS STREET	14094
FLINDERS LANE	10506
BOURKE STREET	9273
Lt COLLINS STREET	7783
CHURCH STREET	5305

There are 3 high priority streets: King Street, Collins Street and Flinders Lane.

High priority streets that recorded violations in Rialto:

Side of Street	Violations	Percentage
4	7700	54.633177
3	6394	45.366823

7700 violations were recorded on side 4, 6394 violations on side 3, which account for roughly 54% and 45% of total violations respectively.

Testing the Worst Ranked Signs

SIGN ONE: 1/2P M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43193	20
43192	20
2242	13
2108	12
2170	11
...	...
8650	1
6601	1
4548	1
6593	1
34821	1

[3774 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The majority of infringements have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 1 and 2 hours limit.

4.2.4.15 RACV

70159 violations were recorded in the RACV area, which makes it the 15th highest among the 23 areas.

Finding the number of streets that recorded violations

Street Name	Violations
QUEEN STREET	24561
BOURKE STREET	20782
Lt COLLINS STREET	19096
WILLIAM STREET	5720

There are 3 high priority streets: Queen Street, Bourke Street and Lt Collins Street.

High priority streets that recorded violations in RACV:

Side of Street	Violations	Percentage
3	24239	60.782888
4	15639	39.217112

About 60.78% violations were recorded on side 3, the rest on side 4.

Testing the Worst Ranked Signs

```

SIGN ONE: 1P MTR M-SAT 7:30-19:30
Value Counts for Parking Duration:      Parking Duration (s)
43192      97
43193      39
3951       27
3954       25
3932       23
...        ...
8786       1
35399      1
6725       1
8754       1
12350      1

```

```

[6549 rows x 1 columns]
The time (hours) of the max length of stay is: 12.0 Hours

```

The highest 2 value counts have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution should be removing the 1 hour limit.

4.2.4.16 County

62866 violations were recorded in the County area, which makes it the 16th highest among the 23 areas.

Finding the number of streets that recorded violations

Street Name	Violations
LONSDALE STREET	33846
Lt BOURKE STREET	11717
SPENCER STREET	11447
KING STREET	5856

There are 3 high priority streets: Lonsdale Street, Lt Bourke Street and Spencer Street.

High priority streets that recorded violations in County:

Side of Street	Violations	Percentage
----------------	------------	------------

1	18928	33.201193
3	14034	24.616734
4	12601	22.103140
5	8466	14.850026
2	2981	5.228907

Around 33.20% of violations were recorded on side 1, followed by side 3 (24.62%) and 4 (22.10%).

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43192	605
43193	465
4015	17
43191	16
4092	16
...	...
22653	1
15852	1
7672	1
10355	1
14049	1

[6507 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The highest 2 value counts have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus the solution would be removing the 1 hour limit from Monday to Friday.

4.2.4.17 Regency

58816 violations were recorded in the Regency area, making it the 17th highest among the 23 areas.

Finding the number of streets that recorded violations:

Street Name	Violations
-------------	------------

EXHIBITION STREET	17403
SPRING STREET	13573
RUSSELL STREET	10192
Lt LONSDALE STREET	9533
LA TROBE STREET	5806
MACKENZIE STREET	2309

There are 3 high priority streets: Exhibition Street, Spring Street and Russell Street.

High priority streets that recorded violations in Regency:

Side of Street	Violations	Percentage
1	23439	56.934998
2	9929	24.118247
5	7800	18.946755

Side 1 recorded the highest number of violations, accounting for over half of total violations (56.93%), followed by side 2 (24.12%) and 5 (18.95%).

Testing the Worst Ranked Signs

SIGN ONE: 2P MTR M-SAT 7:30-20:30
Value Counts for Parking Duration:

46792	274
46793	84
46791	40
7690	23
46789	21
...	...
14110	1
16159	1
20257	1
30502	1
32768	1

Parking Duration (s)

[8301 rows x 1 columns]

The time (hours) of the max length of stay is: 13.0 Hours

The highest 3 value counts have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 2 hours limit from Monday to Friday.

4.2.4.18 Jolimont

54712 violations were recorded in the Jolimont area, making it the 18th highest among the 23 areas.

Finding the number of streets that recorded violations

Street Name	Violations
ALBERT STREET	14491
ST ANDREWS PLACE	11510
LANSDOWNE STREET	8303
CLARENDON STREET	7412
CATHEDRAL PLACE	5548
GISBORNE STREET	2242
NICHOLSON STREET	2072
PARLIAMENT PLACE	2063
WELLINGTON PARADE	1071

There are 2 high priority streets: Albert Street and St Andrews Place.

High priority streets that recorded violations in Jolimont:

Side of Street	Violations	Percentage
4	16198	62.297604
3	9803	37.702396

About 62.30% violations were recorded on side 4, the rest on side 3.

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-18:30

Value Counts for Parking Duration:

Parking Duration (s)

39593	159
39592	23
4066	11
3941	10
4073	10
...	...
6228	1
6105	1
10199	1
8144	1
14732	1

[4345 rows x 1 columns]

The time (hours) of the max length of stay is: 11.0 Hours

The highest 2 value counts have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 1 hour limit from Monday to Friday.

4.2.4.19 Magistrates

37862 violations were recorded in the Magistrates area, making it the 18th highest among the 23 areas.

Finding the number of streets that recorded violations:

Street Name	Violations
Lt LONSDALE STREET	14314

LA TROBE STREET	14091
KING STREET	6990
SPENCER STREET	1317
LONSDALE STREET	1150

There are 2 high priority streets: Lt Lonsdale Street and La Trobe Street.

High priority streets that recorded violations in Magistrates:

Side of Street	Violations	Percentage
4	15777	55.543038
3	12628	44.456962

About 55.54% violations were recorded on side 4, the rest on side 3.

Testing the Worst Ranked Signs

SIGN ONE: 2P MTR M-SAT 7:30-20:30

Value Counts for Parking Duration:

Parking Duration (s)

```

46792      1439
46793      353
46791      148
46790       84
46789       54
...
13962        1
11913        1
9864         1
7815         1
12246        1

```

[5351 rows x 1 columns]

The time (hours) of the max length of stay is: 13.0 Hours

The majority of parking events have a duration of around 13 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 2 hours limit from Monday to Friday.

4.2.4.20 Supreme

25824 violations were recorded in the Supreme area, making it the 19th highest among the 23 areas.

Finding the number of streets that recorded violations:

Street Name	Violations
LONSDALE STREET	25824

High priority streets that recorded violations in Supreme:

Side of Street	Violations	Percentage
1	18547	71.820787
3	3859	14.943463
4	3418	13.235750

The vast majority of violations were recorded on side 1, accounting for 71.82% of total violations.

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

Parking Duration (s)

43192	525
43193	394
3946	20
4118	20
4005	17
...	...
22002	1
7675	1
30214	1
17936	1
18335	1

[7203 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The highest 2 value counts have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 1 hour limit from Monday to Friday.

4.2.4.21 Tavistock

1221 violations were recorded in the Tavistock area, making it the 22nd highest among the 23 areas.

Finding the number of streets that recorded violations:

Street Name	Violations
COLLINS STREET	9398

High priority streets that recorded violations in Tavistock:

Side of Street	Violations	Percentage
3	5248	55.841668
4	4150	44.158332

About 55.84% violations were recorded on side 3, the rest on side 4.

Testing the Worst Ranked Signs

SIGN ONE: 1P MTR M-SAT 7:30-19:30

Value Counts for Parking Duration:

43192	145
43193	51
3939	14
3962	14
3911	13
...	...
8328	1
8320	1
6271	1
12410	1
24496	1

Parking Duration (s)

[4241 rows x 1 columns]

The time (hours) of the max length of stay is: 12.0 Hours

The highest 2 value counts have a duration of around 12 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 1 hour limit from Monday to Friday.

4.2.4.22 Docklands

8770 violations were recorded in the Docklands area, making it the 21st highest among the 23 areas.

Finding the number of streets that recorded violations:

Street Name	Violations
JEFFCOTT STREET	5523
ANDERSON STREET	1509
SPENCER STREET	1358
KING STREET	321
BATMAN STREET	59

All streets will be included given that the number of violations is minimal.

High priority streets that recorded violations in Docklands:

Side of Street	Violations	Percentage
4	3525	40.193843
3	1998	22.782212
1	1568	17.879133
5	990	11.288483
2	689	7.856328

Side 4 recorded the highest number of violations (40.19%), followed by side 3 (22.78%) and 1 (17.88%).

Testing the Worst Ranked Signs

SIGN TWO: 1P RPA M-F 7:30-18:30

Value Counts for Parking Duration:

Parking Duration (s)

```
39593      156
39592      87
39591      11
39039       2
11084       2
...
32067       1
5440        1
30264       1
9534        1
30720       1
```

[1060 rows x 1 columns]

The time (hours) of the max length of stay is: 11.0 Hours

The highest 4 value counts have a duration of around 11 hours, which may suggest that those who parked here were for work purposes. Thus, the solution would be removing the 1 hour limit from Monday to Friday.

4.2.4.23 West Melbourne

668 violations were recorded in the West Melbourne area, making it the 23rd highest among the 23 areas.

Finding the number of streets that recorded violations:

Street Name	Violations
BATMAN STREET	668

High priority streets that recorded violations in Docklands:

Side of Street	Violations	Percentage
1	316	47.305389
3	255	38.173653
4	97	14.520958

Side 1 recorded the highest number of violations (47.31%), followed by side 3 (38.17%) and 4 (14.52%).

Testing the Worst Ranked Signs

SIGN ONE: 4P MTR M-F 7:30-18:30

Value Counts for Parking Duration:

Parking Duration (s)

39593	3
39592	3
19844	2
15018	1
32933	1
...	...
17013	1
22365	1
28065	1
14771	1
15907	1

[388 rows x 1 columns]

The time (hours) of the max length of stay is: 11.0 Hours

SIGN TWO: 2P M-SAT 7:30-18:30

Value Counts for Parking Duration:

Parking Duration (s)

8447	2
8156	2
7998	2
9700	2
8660	2
...	...
8012	1
15690	1
19273	1
15176	1
8192	1

[250 rows x 1 columns]

The time (hours) of the max length of stay is: 11.0 Hours

The highest 2 value counts of sign one has a duration of around 11 hours, which may suggest that those who parked here were for work purposes. Meanwhile, the violations recorded on sign 2 had significantly shorter parking durations, with most of them being just over 2 hours. Thus, the solutions would be removing the 4 hours limit of sign one from Monday to Friday, and increasing the limit of sign two to 3 hours.

7.0 Machine Learning

Machine learning (ML) is the process of studying computer algorithms that improve automatically through experience and by the use of data and artificial intelligence. ML algorithms can be developed for automation

tasks where computers discover how tasks can be performed without being explicitly programmed to do so. Some use cases surrounding Machine Learning are email filtering, medical applications, computer vision, data mining and many more applications that possess huge amounts of data. The team researched a Linear Regression model and a Logistic Regression model during the regression analysis stage and decided it is best to use Logistic Regression to implement machine learning to the metropolitan parking dataset. These methodologies are explained in more detail in the following subtopics.

5.1 Linear Regression

Linear Regression is the process of modelling the relationship between a scalar response and one or more explanatory variables. Linear Regression relies on the target variable to be a continuous Real value. Therefore the target variable that is being predicted is "Violation", which is a binary value, so linear regression cannot be used since when the values are in binary form and it relies on the target variable being a continuous real value, which means that the target variable "violation" is in binary form.

5.2 Logistic Regression

Logistic Regression is the process of modelling a binary dependent variable by using a logistic function. In regression analysis, logistic regression estimates the parameters of a logistic model therefore it relies on the target variable to be an integer or categorical value. Therefore we can use logistic regression instead of linear regression as the machine learning regression analysis since it relies on the target variable to be an integer or categorical value.

5.2.1 Data Transformation

A number of methods were considered in approaching the design of the model. It was known early on that the dataset contained mostly categorical variables, and so these attributes were required to be transformed into integer or float values. There were 3 main ways of approaching this task, with the third being the most appropriate for the task at hand.

The first included one-hot encoding these values, however given that each categorical attribute had at least 20 different inputs, this would mean a significant jump in the number of columns of the dataset (by the measure of hundreds). So therefore this approach was not used.

The second instance was to find a realistic way of structuring the variables so that they could fit in a 'cycle' to label them with an ordinal approach. For instance, one could label the seasons in an ordinal fashion to indicate what is closer to the other and how they relate to one another. However, there was no way of approaching this tactfully with the dataset available, so this approach was also not used.

The third suggestion was to incorporate a label encoder to assign integer values to all categorical attribute values. This was the simplest method, as this didn't involve relationships between the variables, nor did it require adding extra entries.

5.2.2 Hyperparameter Tuning

Machine learning algorithms have hyperparameters that allow tailoring the behaviour of the algorithm to tune the performance to a specific dataset, in this case the Metropolitan Parking Dataset. Hyperparameter tuning requires a lot of tuning and patience while the new model is being tested against the older models.

The default hyperparameters used gave an F1 score of 0.65. This wasn't too bad, but this can be drastically improved considering the size of the dataset should allow for an improved F1 score and accuracy measure. The default settings were a max iteration of 100 and a training and testing dataset split of 80%-20% respectfully.

Some of the modified parameters were:

- Max iterations: Typically increased so that the model could diverge closer to the global minimum (this is where we receive the least amount of loss for the model).
- Train-Test Split: The split was moved from 0.2-0.4 to try and reduce model overfitting to the training set.

5.2.3 Feature Selection

During the process of testing different models, it was decided to remove certain columns according to the confusion matrix. Variables that saw a minimal effect on the Violation attribute were removed. The main application of this was to remove Street Marker, since that showed a correlation almost at 0.0.

Street name was also removed, but this was because there was a redundant Street ID column present in the dataset.

The initial plans for arrival and departure times were to be divided into multiple columns: Day, Month, and Year. These are all integer values, and the time of day will be a sine-cos transition to demonstrate the continuity of time and how it circles back to midnight. However this was decided against, as the correlation matrix showed little correlation. And this information would be better used within a VGG neural network, which is out of scope for this project.

5.2.4 Final Model

The final model chosen presented with the following scores:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	2634498
1	1.00	1.00	1.00	417547
accuracy			1.00	3052045
macro avg	1.00	1.00	1.00	3052045
weighted avg	1.00	1.00	1.00	3052045

While this a highly accurate model, there is the risk that it has overfitted to the data, however with future partnerships with VicRoads, the team at DDS will be able to test this against new data. This model serves as a 'continuous' work with VicRoads, DDS to modify the parameters with new information, so as to keep up to date with new parking trends.

8.0 Conclusion

Given the thorough investigation, it was found that violations tended to be a result of workers who had parked for the entirety of their workday, with a buffer time of 1-2 hours. It was also found that very busy work districts of the city also saw a higher number of violations than others. Therefore

recommendations were based around these major findings.

Bibliography

- 1.) Jupyter.org. n.d. *Project Jupyter*. [online] Available at: <<https://jupyter.org>>.[1]
- 2.) Pandas.pydata.org. n.d. *pandas - Python Data Analysis Library*. [online] Available at: <<https://pandas.pydata.org/about/>>.[2]
- 3.) Scikit-learn.org. n.d. *scikit-learn: machine learning in Python — scikit-learn 0.24.2 documentation*. [online] Available at: <<https://scikit-learn.org/stable/>>.[3]
- 4.) Numpy.org. n.d. *NumPy*. [online] Available at: <<https://numpy.org/>>. [4]
- 5.) Matplotlib.org. n.d. *Matplotlib: Python plotting — Matplotlib 3.4.2 documentation*. [online] Available at: <<https://matplotlib.org/stable/index.html>>.[5]
- 6.) Seaborn.pydata.org. n.d. *seaborn: statistical data visualization — seaborn 0.11.1 documentation*. [online] Available at: <<https://seaborn.pydata.org/>>. [6]
- 7.) Dataversity.net n.d. *Data Quality* [Online] Available at: <<https://www.dataversity.net/data-quality-simple-6-step-process/>>. [0]