Pedestrian activity in the City of Melbourne before and during COVID-19

Members: Huma Chaudhry, Prabhdeep Sandhu, James Akerman, (Mona Mansouri)

Presentation Outline

- Introduction
    - Background
    - Motivation
    - Research questions
- Dataset
- Data exploration and Cleanup
- Analysis Process
- Summary
- Findings and Limitations

Introduction:

Background

- The City of Melbourne has been recording hourly pedestrian activity within its region since at least 2009 via its pedestrian counting system.
- The system's stated aims include to
  - monitor retail activity in the city
  - inform decisions about urban planning and management
  - identify opportunities to improve city walkability and transport

Motivation

- We suspected pedestrian activity patterns in the City of Melbourne had changed since the arrival of COVID-19.

Our Research and Scope

- **Hypothesis One:** The overall pedestrian activity in the City of Melbourne had <u>decreased</u> since March 2020 when the first lockdown happened in Melbourne.

- **Hypothesis Two:** Pedestrian activity had <u>increased</u> after the first lockdown, but not to pre-pandemic levels (average of August 2019 to February 2020).

- **Hypothesis Three:** Pedestrian activity would <u>increase</u> as JobKeeper claims increased.

- **Hypothesis Four:** Pedestrian activity would <u>decrease</u> after JobKeeper was tapered.

- **Hypothesis Five:** Pedestrian activity <u>increases</u> as the number of Victorian full-time workers increases.

Datasets

Dataset Name	Source	Dataset description
Pedestrian Counting System - Monthly (counts per hour)	https://data.melbourne.vic.gov.au/Transport/Pedestrian-Counting-System-Monthly-counts-per-hour/b2ak-trbp	Each row is a Total hourly sensor count of pedestrians.
Pedestrian Counting System - Sensor Locations	https://data.melbourne.vic.gov.au/Transport/Pedestrian-Counting-System-Sensor-Locations/h57g-5234	This dataset contains status, location and directional information for each pedestrian sensor device installed throughout the city.
Labour force status by Sex, Victoria - Trend, Seasonally adjusted and Original	https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-australia/aug-2021/6202005.xls	Victorian employment data for each month from February 1978 to August 2021.
Jobkeeper Postcode data	https://treasury.gov.au/coronavirus/jobkeeper/data	Business JobKeeper per postcode per month for the duration of JobKeeper.
Google Maps API	https://developers.google.com/maps	Postcode and address data based on sensor location directional information.

Data Exportation and Clean Up

Common clean up strategies across all hypothesis		1. Get data from APIs 2. Select common sensors for year 2019, 2020, 2021 3. Remove Duplicate rows 4. Correct data type for columns
Hypothesis	**Datasets**	**Hypothesis Specific Clean Up Steps**
Hypothesis 1	Pedestrian count (per hour data) - csv	1. Removed inactive values 2. Used sum() to get monthly data per location
Hypothesis 2	Pedestrian count (per hour data) - csv Pedestrian Counting System - Sensor Locations	1. Convert Month column to categorical type and sort 2. Merge and aggregate average monthly pedestrian count across all sensors

Data Exportation and Clean Up (continued)

Hypothesis	Datasets	Clean Up Steps
Hypotheses 3 and 4	Pedestrian Counting System - Monthly (counts per hour)	- Filtered by year and date (Aug-19 to Aug-21) - Removed all null values - Aggregated hourly values by month-year
	Pedestrian Counting System - Sensor Locations	- Filtered by year and date (Aug-19 to Aug-21) - Removed all null values
	Google Maps API	Filtered locations by latitude and longitude coordinates found in the sensor location data.
	JobKeeper data	Filtered by sensor location postcodes found from the Google Maps API
Hypothesis 5	Labour Force, Australia (Full time employment)	- Filtered data for full-time employment for both genders across year 2019, 2020, 2021 - Save dataframe as a separate .csv file

Analysis Process

| Hypothesis	Analysis Process	
	Descriptive Analysis	*Statistical Analysis*
Hypothesis 1	● Line graphs ● Google map locations ● Table format of final result	● Independent T-test ● Logical analysis
Hypothesis 2	● Bar chart	● Independent T-test
Hypothesis 3	● Created line graphs of summed pedestrian count data for relevant postcodes for the months during the time JobKeeper was implemented.	● Conducted correlation tests between pedestrian count and JobKeeper business application counts for each relevant postcode.
Hypothesis 4	● Created line graphs of summed pedestrian count data for relevant postcodes for the months after JobKeeper was implemented.	● Conducted independent t-tests to verify apparent downward trend in pedestrian activity after the end of JobKeeper.
Hypothesis 5	● Scatter plot	● Linear Regression ● T-test

Findings

Hypothesis 1

the overall pedestrian activity in the City of Melbourne had decreased since March 2020 when the first lockdown happened in Melbourne.

Dataset used for analysis

Month	Bourke Street Mall (North)	Bourke Street Mall (South)	Melbourne Central	Town Hall (West)	Princes Bridge	Flinders Street Station Underpass	Birrarung Marr	Webb Bridge	Southern Cross Station	Victoria Point	...	QVM-Queen St (East)	QVM-Therry St (South)	Faraday St-Lygon St (West)	QVM-Franklin St (North)	Elizabeth St-Lonsdale St (South)
Aug-19	906524	641786	981893	1064689	963378	1051024	324456	144995	621153	155030	...	206593	137175	193302	99286	325035
Sep-19	943758	643780	869674	971314	925437	1008301	234247	140767	561236	131512	...	179612	93315	184283	87344	313040
Oct-19	1028680	495941	919196	1047185	1028103	1151297	187355	58490	638247	148089	...	194461	94980	205178	101168	332225
Nov-19	1037267	744945	901523	1074747	673975	1105507	320914	137894	536947	134803	...	205257	135467	205253	94935	328209
Dec-19	1192463	927155	911297	1244392	1068849	1187116	209137	146964	442756	118303	...	207353	132729	223789	97110	310402

Pedestrian count vs lockdown stages
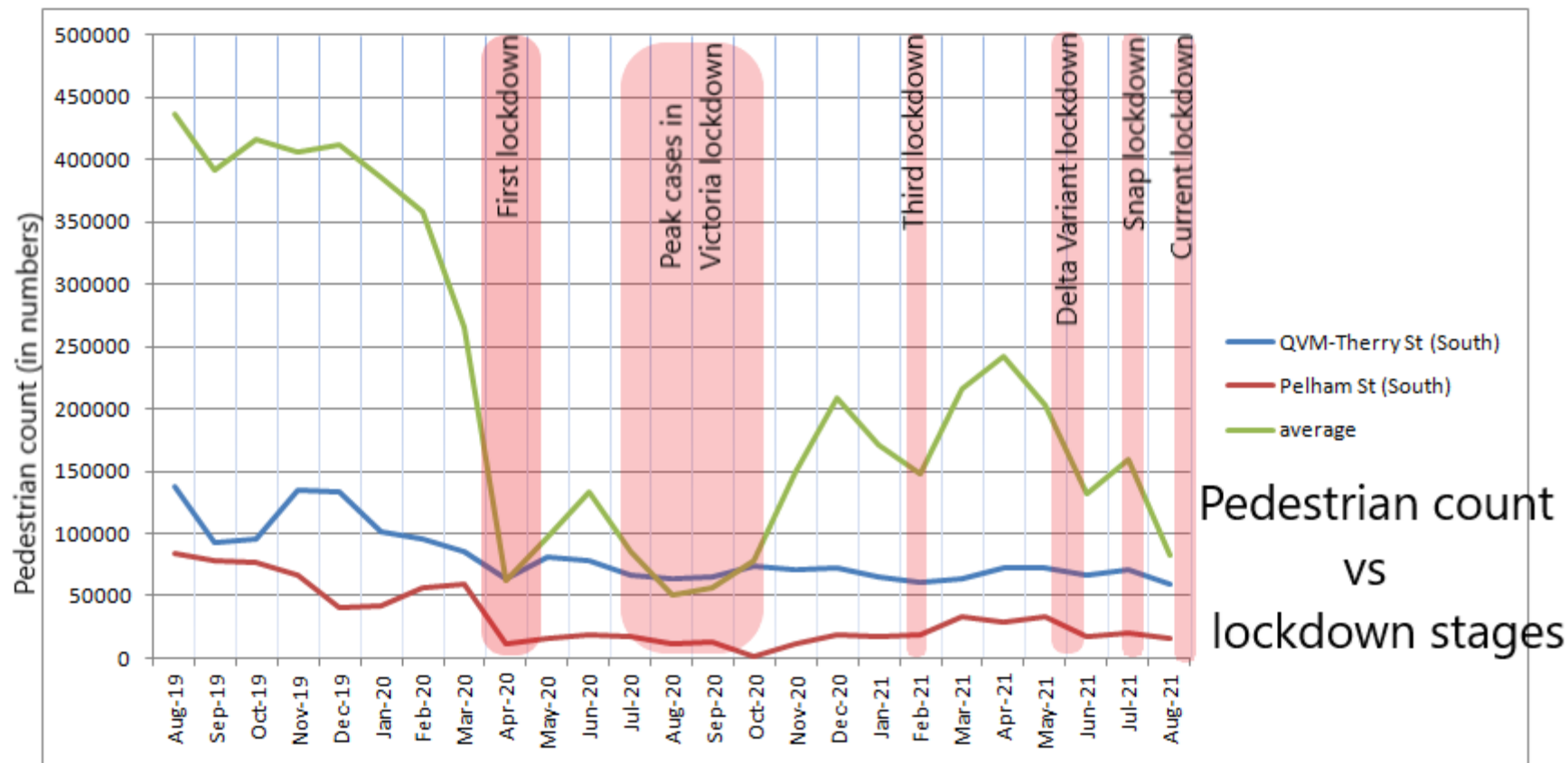
Largest drop of pedestrian count

Location	Max month	Max value	Min month	Min value	Difference	Difference percentage
Pelham St (South)	Aug 2019	83553	Oct 2020	1790	81763	97.86

Smallest drop of pedestrian count

Location	Max month	Max value	Min month	Min value	Difference	Difference percentage
QVM-Therry St (South)	Aug 2019	137175	Aug 2021	59596	77579	56.55

Parks and playgrounds

Business area (not essential)

School and colleges

University of Melbourne

University Square

Giblin Eunson Library

The Spot

Melbourne Business School

Church St

Hypetap

KFC Carlton
Fried Chicken · $

Subway
Sandwich · $

Tambo Ciek
Indonesian · $

The Park Hotel Melbourne
3.8 ★ (704)
4-star hotel

Swanston St

Little Pelham St

Lincoln Square N

Pelham St

Law School

Mama Tsai
Menulog Delivery

Lincoln Square/
Swanston Street

NORSIAHS KITCHEN

Lincoln Square/
Swanston St 3

Pelham St

MSS IMAGE

Power Tool Specialists
Tool store

The Creative Works

Barkly Pl

Canada Ln

Lincoln Square S

Student Living - 570
Swanston, Student...

Rooms International
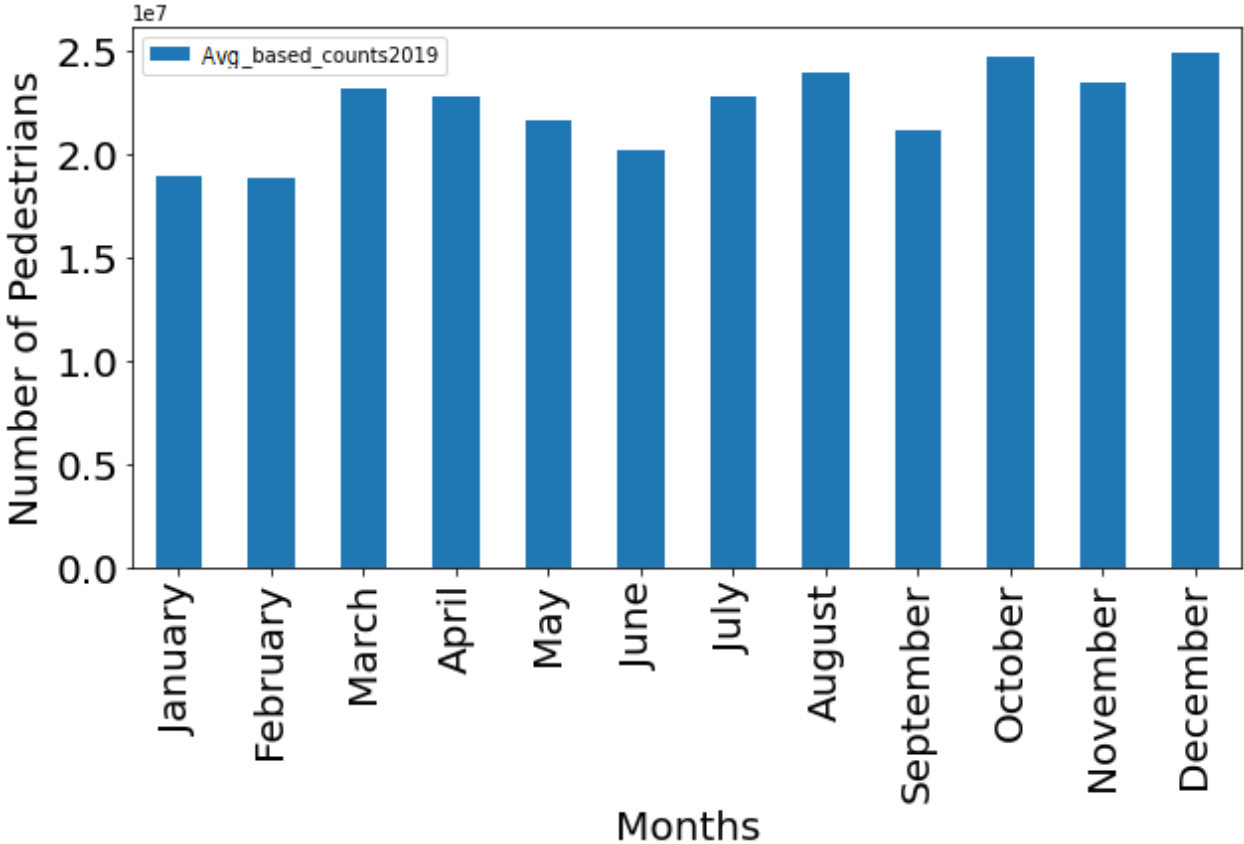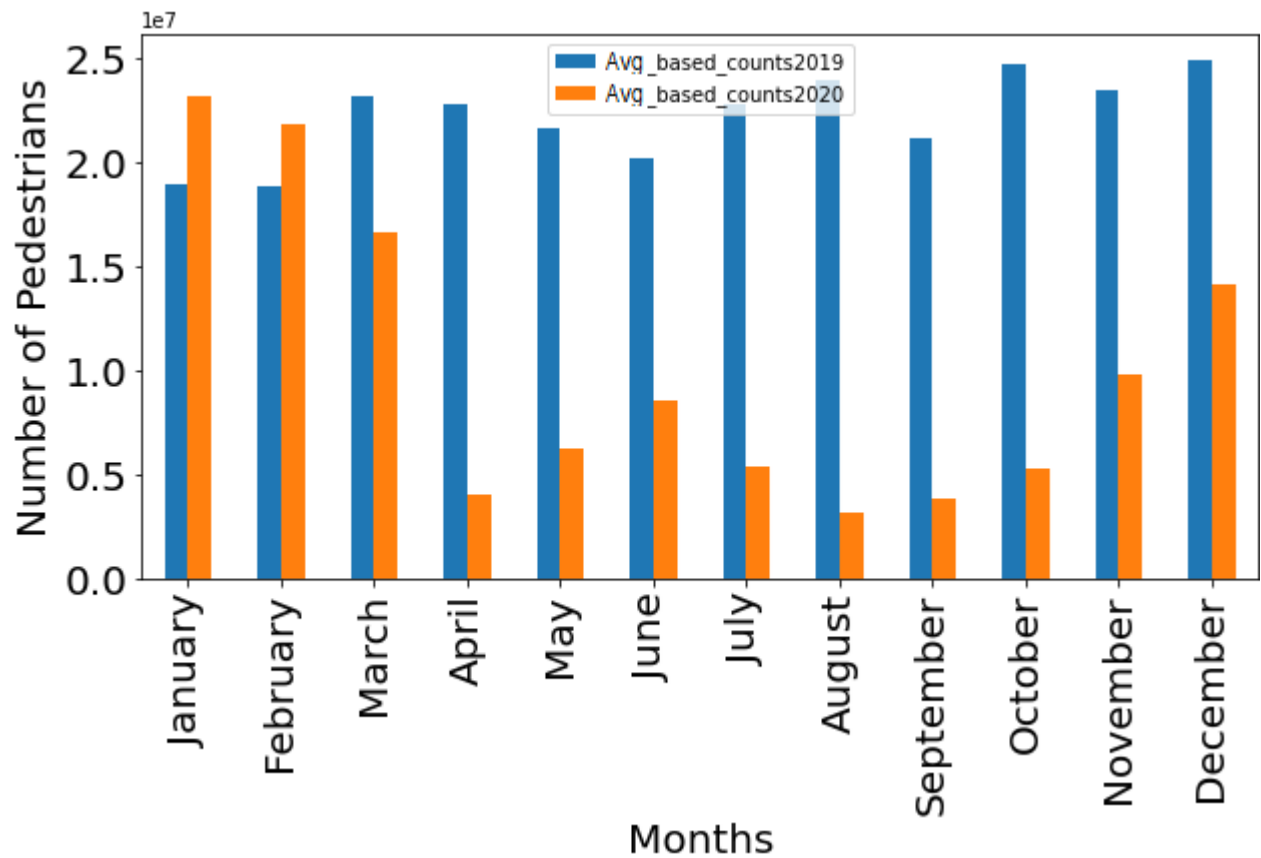
Ziman Pilates & Physio

Queen Victoria Market (Essential)

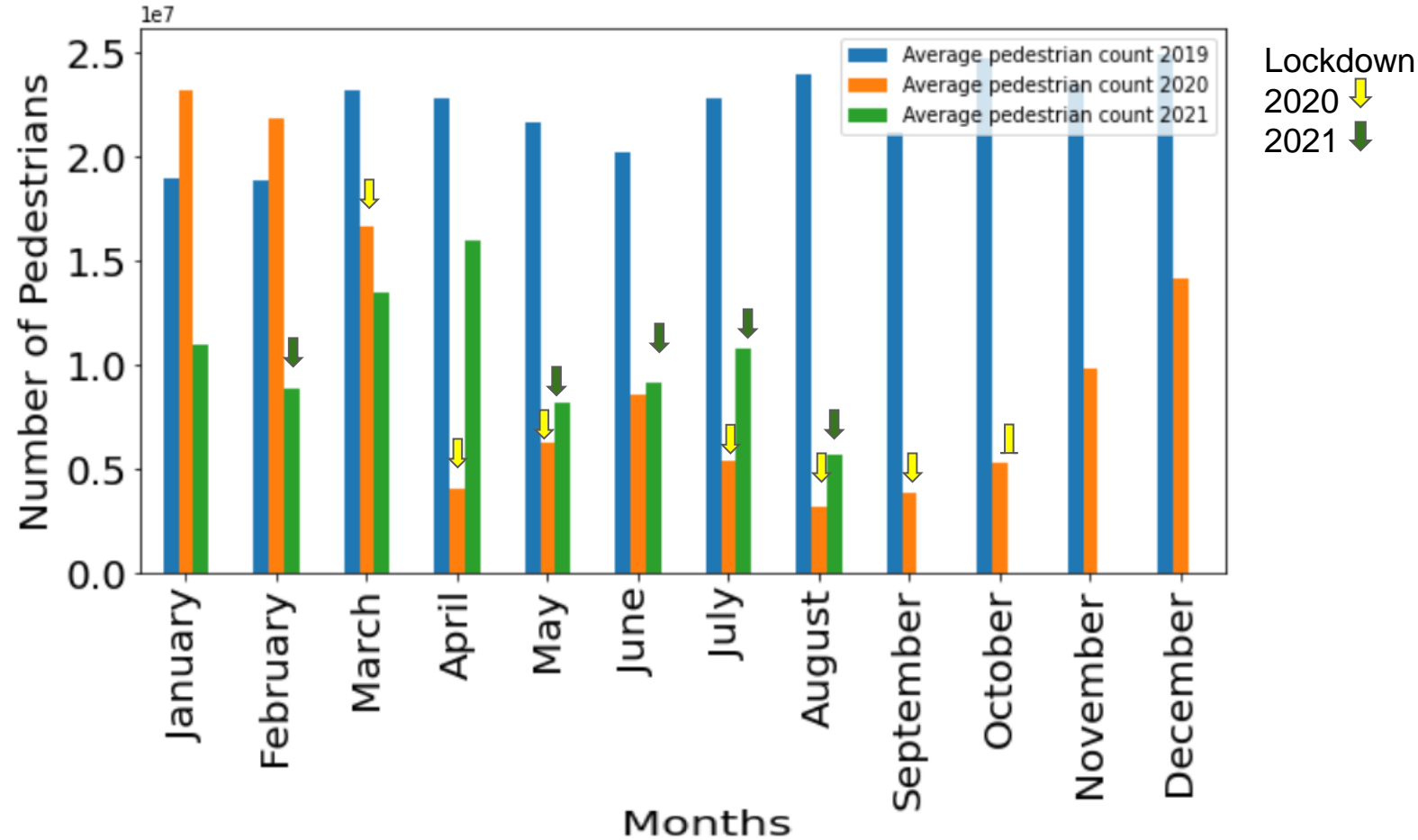NON-ESSENTIAL BUSINESS AREA

Hypothesis 2,5

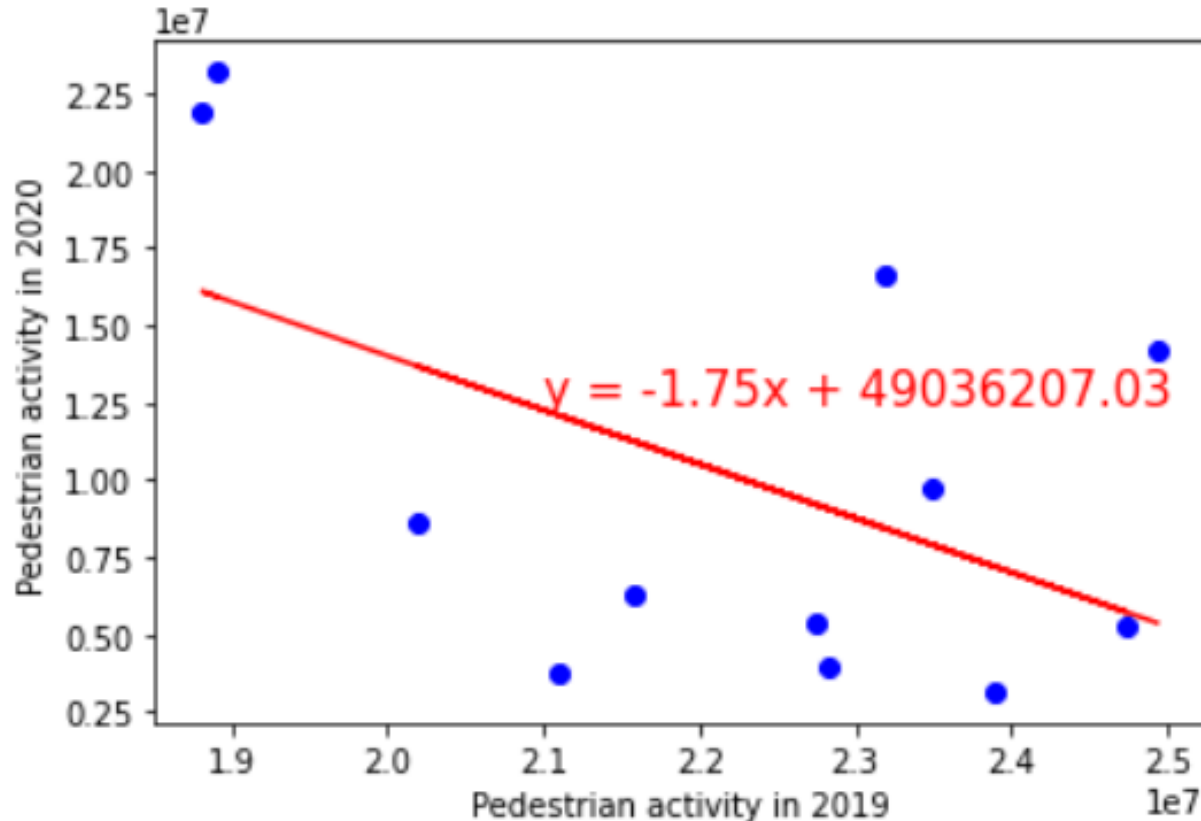**Average Pedestrian activity pre-pandemic year.**

Average Pedestrian activity in the first pandemic year 2020.

Hypothesis 2: Average Pedestrian activity increased after the first lockdown, but not to pre-pandemic levels .

Trend of pedestrian activity from PreLockdown (2019) to PostLockdown (2020)

Hypothesis 2: Average Pedestrian activity increased after the first lockdown, but not to pre-pandemic levels .

Scatter plot can illustrate a relationship been pedestrian count in 2019 and 2020.

It has a negative relation.

Hypothesis 2: Average Pedestrian activity increased after the first lockdown, but not to pre-pandemic levels .

Statistically prove significance of change using p-value

independent T-test on Hypothesis 2:

```python
stat,pvalEmployment=stats.ttest_ind(df["day_based_counts2020"], df["day_based_counts2019"], equal_var=False)
if pvalEmployment < 0.05:
    print("Reject the Null hypothesis- There is a significant change in both categories with p-value",\
          round(pvalEmployment,4))
else:
    print("Null Hypothesis: No change in pedestrian acitvity pre or post lockdown")
```

Reject the Null hypothesis- There is a significant change in both categories with p-value 0.0001

Summary Statistics across 2019, 2020, 2021

	count	mean	std	min	25%	50%	75%	max
day_based_counts2019	8.0	2.151833e+07	1.982256e+06	18816254.0	19871838.0	22167802.5	22908900.50	23896524.0
day_based_counts2020	8.0	1.112107e+07	8.177723e+06	3157701.0	5037630.0	7415702.5	17895621.75	23189121.0
day_based_counts2021	8.0	1.034907e+07	3.192701e+06	5678484.0	8628688.0	9957313.0	11531720.75	15917378.0

Hypothesis Five: Pedestrian activity increases as the number of Victorian full-time workers increases.

Trend of pedestrian activity vs. Fully time employment

$y = 0.0047x + 2.21e{+}06$
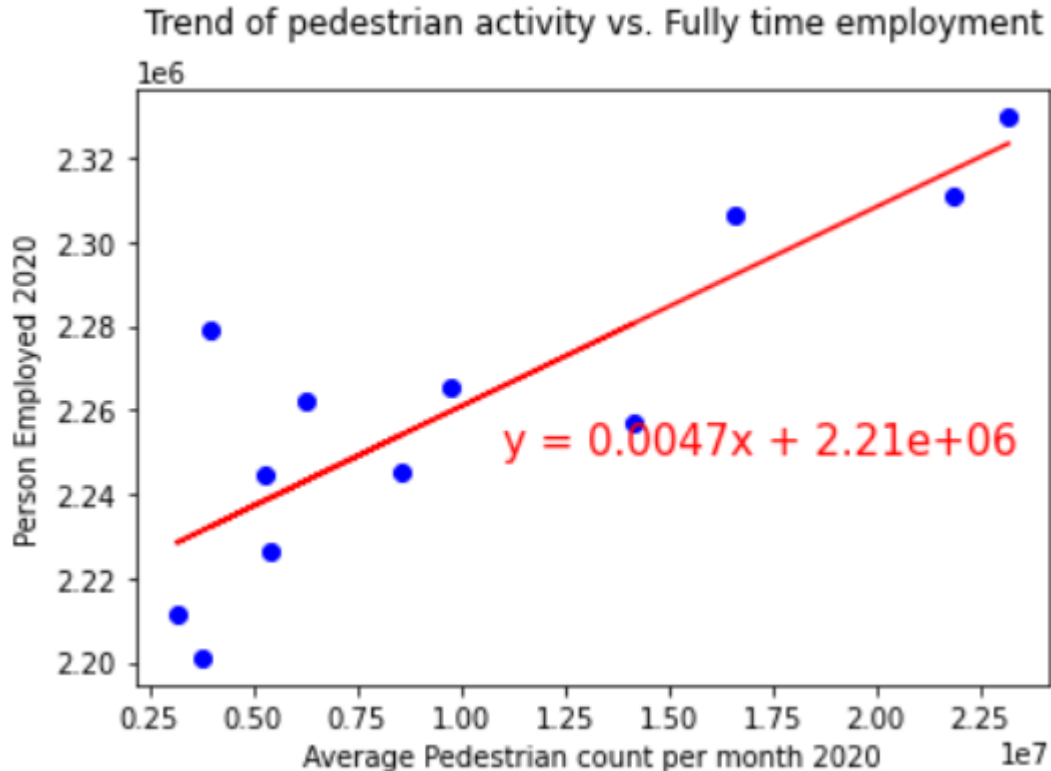
Scatter plot shows direct relationship between full time employed persons and pedestrian count

Hypothesis Five: Pedestrian activity increases as the number of Victorian full-time workers increases.

Statistical Analysis with t-test

```
df3_employed["day_based_counts2020"].corr(df3_employed["Person Employed 2020"])
```

```
0.8471311321828738
```

```
corr2=st.pearsonr(x_val,y_val)
print(f'The correlation between full time employment and pedestrian count is {round(corr2[0],2)}
```

```
The correlation between full time employment and pedestrian count is 0.85 with p-value 0.0005
```

Because p-value<0.05, thus there is a significant relationship- it did not happen by chance

Hypothesis 3

Pedestrian activity would increase as JobKeeper claims increased

Statistically significant moderate to strong negative correlations were found between pedestrian activity and JobKeeper Business applications for 5 of the 7 Melbourne postcodes observed (3000, 3004, 3008, 3052, 3053)

Correlation Tests

The correlation between both factors for postcode 3000 is -0.76 with a p-value of 0.00436
The correlation between both factors for postcode 3004 is -0.87 with a p-value of 0.00027
The correlation between both factors for postcode 3006 is -0.26 with a p-value of 0.40575
The correlation between both factors for postcode 3008 is -0.69 with a p-value of 0.01349
The correlation between both factors for postcode 3010 is 0.15 with a p-value of 0.65185
The correlation between both factors for postcode 3052 is -0.59 with a p-value of 0.04526
The correlation between both factors for postcode 3053 is -0.72 with a p-value of 0.00776

The correlation between both factors is -0.76 with a p-value of 0.00436

Figure 6: Pedestrian Count vs JobKeeper Applications for Postcode 3000

$y = -0.54x + 5267.11$
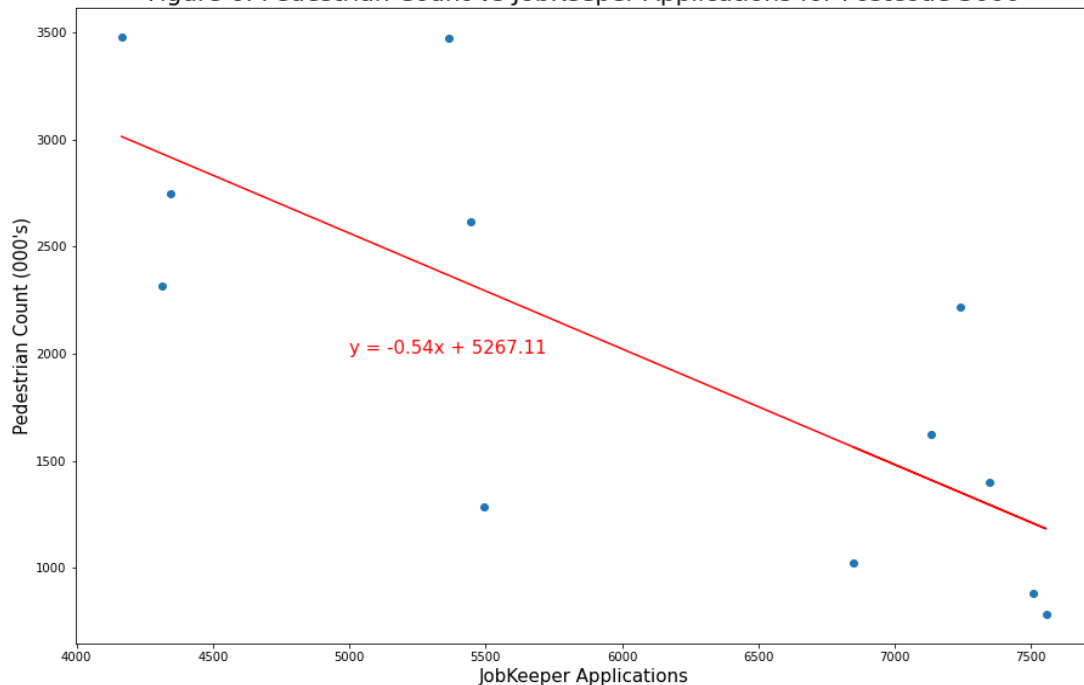
Hypothesis 4

Pedestrian activity would decrease after JobKeeper was tapered

Although there appeared to be a decline in pedestrian activity after the JobKeeper tapering, our analysis found that there was no statistically significant difference in pedestrian activity during these periods.

Figure 1: Pedestrian Traffic per postcode from August 2019 to August 2021 (000's)
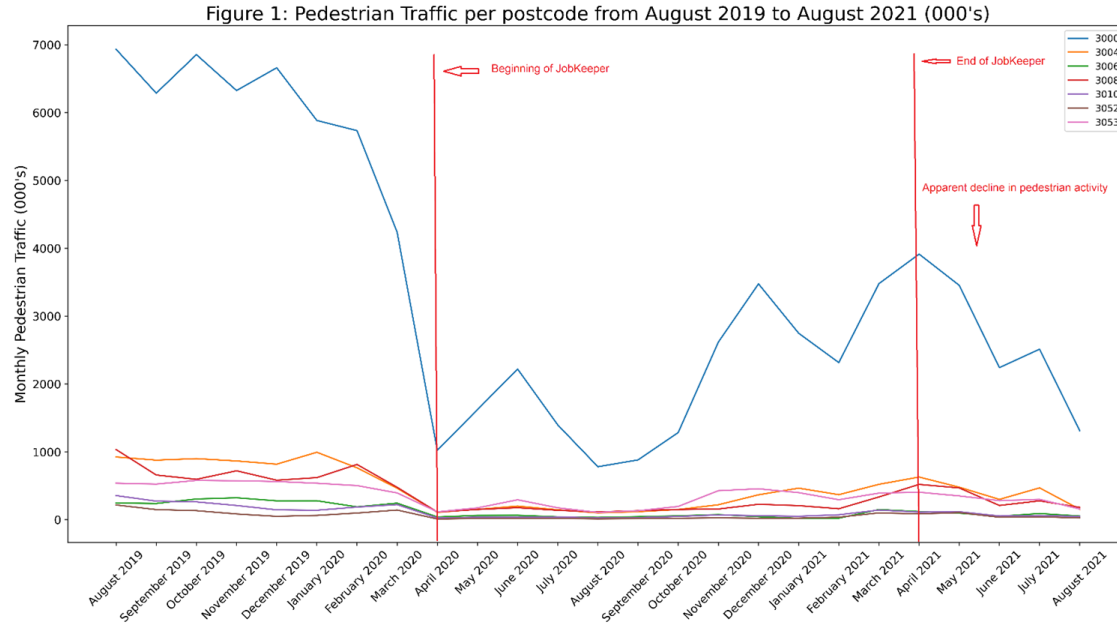
- Independent T-test Analysis
- the p-value of the Welch's t-test for postcode 3006 is 0.132
- the p-value of the Welch's t-test for postcode 3000 is 0.233
- the p-value of the Welch's t-test for postcode 3052 is 0.116
- the p-value of the Welch's t-test for postcode 3010 is 0.185
- the p-value of the Welch's t-test for postcode 3008 is 0.08
- the p-value of the Welch's t-test for postcode 3004 is 0.126

Findings Summary

Hypothesis One - the overall pedestrian activity in Melbourne had decreased since March 2020 when the first lockdown happened in Melbourne.
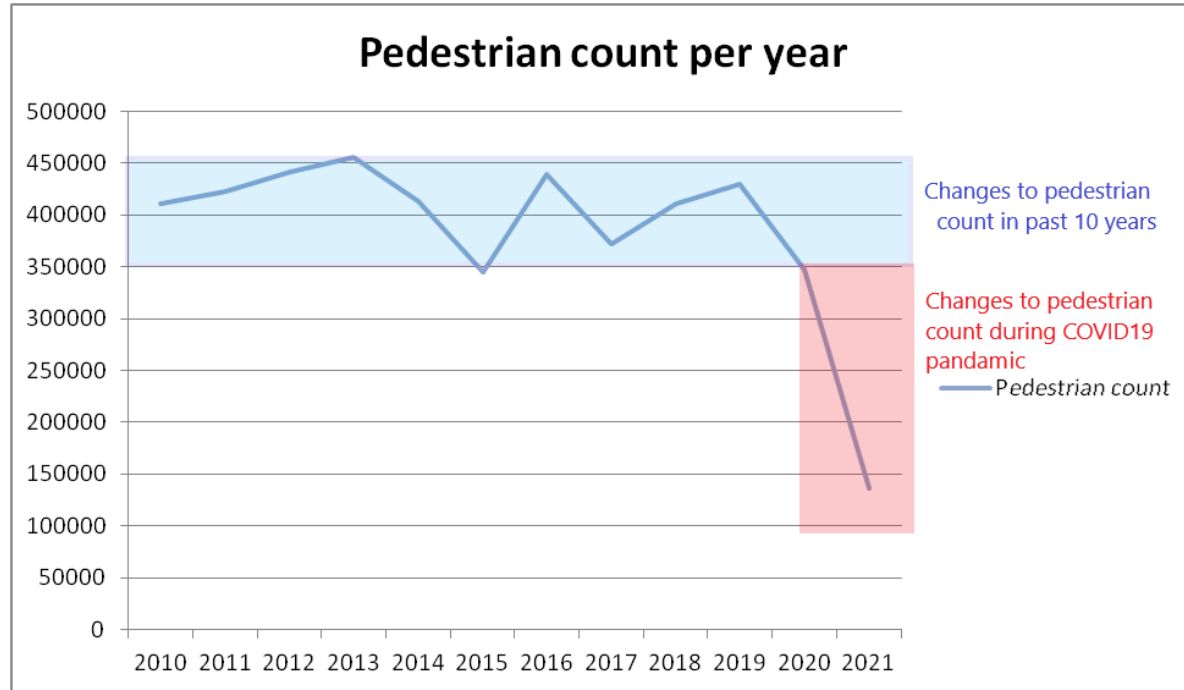
Conclusion

- There was an average drop of 87% in pedestrian activity at various locations in Melbourne during covid19.

- Since the first lockdown, the pedestrian activity has still not returned to pre-lockdown which might suggest a new trend.Only half of the pedestrian activity was returned back

  (50% instead of 87%)

- The negative correlation found in most postcodes may be because businesses were struggling to pay their workers as the number of pedestrians and hence customers decreased.

- No statistically conclusive evidence of pedestrian numbers decreasing after JobKeeper. Apparent decline maybe due to chance.

- Full time workers were still active during years affected by lockdown.

Research Limitations

- Data on the financial situation of businesses within each postcode was limited to only JobKeeper claims.

- Impossible to distinguish types of pedestrians (Tourists, event spectators, patrons etc..).
- Data does not represent entire population such as pedestrians who didn't walk past sensors (work from home, no purpose of travel to CBD).

- Dataset had few sensors inactive, because of which the accuracy might have been compromised.

Questions?