BIKE THEFT ANALYSIS

OTTAWA & VANCOUVER DURING COVID-19

ANDRIY KHALIK | PRIYA S | INSIYA KANJEE

PROJECT OUTLINE

Covid19 has disrupted several industries, and the biking industry is no exception. With several activities not allowed, individuals have resorted to outdoors activities and biking has been a favourite. However, with increased demand many stores have reported low inventory. Simultaneously, there have been **increased reports of bike thefts in many cities.**

In Seattle, Wash., there was a 54% increase in reported bicycle thefts. Denver, Colo., has seen around 26% increase in bike theft between 2019-2020, according to the Denver Police Department. As for Boston, Mass., it has climbed from 790 in 2019 to 1015 in the last year, a 28% increase. New York City has similarly seen a nearly 30% spike from March to Sept. 21 of last year when compared to the same period in 2019, according to The New York Times. (link)

In this project we will explore if similar bike theft trends are observed in Ottawa & Vancouver in Canada.

RESEARCH QUESTIONS

1. EXPLORE IF MAJOR CITIES (OTTAWA & VANCOUVER) HAVE SEEN SIMILAR TRENDS IN BICYCLE THEFTS DUE TO COVID

2. IDENTIFY THE RELATIONSHIP BETWEEN BIKE THEFT RATES AND TOTAL BIKE VOLUME

3. IDENTIFY THE RELATIONSHIP BETWEEN THE BIKE THEFT RATES AND COVID CASES

AGENDA

1. Methodology
2. Executive Summary
3. Bike Theft Situation Analysis
4. Bike Theft Relationship Analysis
5. Limitations and Further Research
6. Appendices

1. METHODOLOGY

APPROACH

To answer the research questions we:

Examined open theft data available through the the official city website or police department domain to identify and analyse daily, monthly and yearly trends in bike theft incidents in Ottawa and Vancouver during 2018-2020

Used location data and geo API to plot a heatmap in order to understand which parts of city had most bike theft incidents

Analysed open bike traffic/usage data to reveal monthly and yearly trends and how they changed through the years 2018 to 2020

Performed statistical analysis of the collected data by looking at key tests, examining a correlation between bike thefts and usage, and running a regression analysis

Collected COVID cases data for two cities and studied correlation between trends in COVID cases, bike thefts, and bike usage to establish whether COVID had an impact on biking and thefts

DATA ACQUISITION

City	Data Set	Format	Source
VANCOUVER	BIKE THEFT: Crime Data, Vancouver Police Dep	csv	https://geodash.vpd.ca/opendata/
	BIKE VOLUME: City of Vancouver ● Bike volume data ● Historic data from bef Aug 2018	csv pdf	https://vancouver.ca/streets-transportation/how-we-collect-bike-volumes.aspx
	COVID CASES:	csv	http://www.bccdc.ca/health-info/diseases-conditions/covid-19/data
OTTAWA	BIKE THEFT: Bicycle Thefts, City of Ottawa	csv	https://open.ottawa.ca/datasets/ottawa::bicycle-theft-1/about
	BIKE VOLUME: Bicycle Trip Counters,City of Ottawa	csv	https://open.ottawa.ca/documents/f218592c7fe74788906cc6a0eb190af9/about
	COVID CASES:	csv	https://open.ottawa.ca/datasets/ottawa::covid-19-cases-and-deaths-in-ottawa/about
APIs	GEO locations Maps JavaScript API Census Canada		

DATA CLEANING

- Used split function to delimit dates in various formats
- Converted date to datetime type
- Replaced empty/NaN data points with 0 where applicable (for eg. no bikers on a particular day on a trail)
- Summarized daily numbers to weekly and monthly for bike theft and volume data
- Converted X,Y UTM coordinates to Lat and Lon
- Removed irrelevant columns in Covid cases and bike theft data sets
- Filtered only the rows corresponding to Vancouver from BC covid dataset
- Exported cleaned up datasets into csv files for further use

2. EXECUTIVE SUMMARY

EXECUTIVE SUMMARY

The bike theft trends in US have not translated to Ottawa and Vancouver

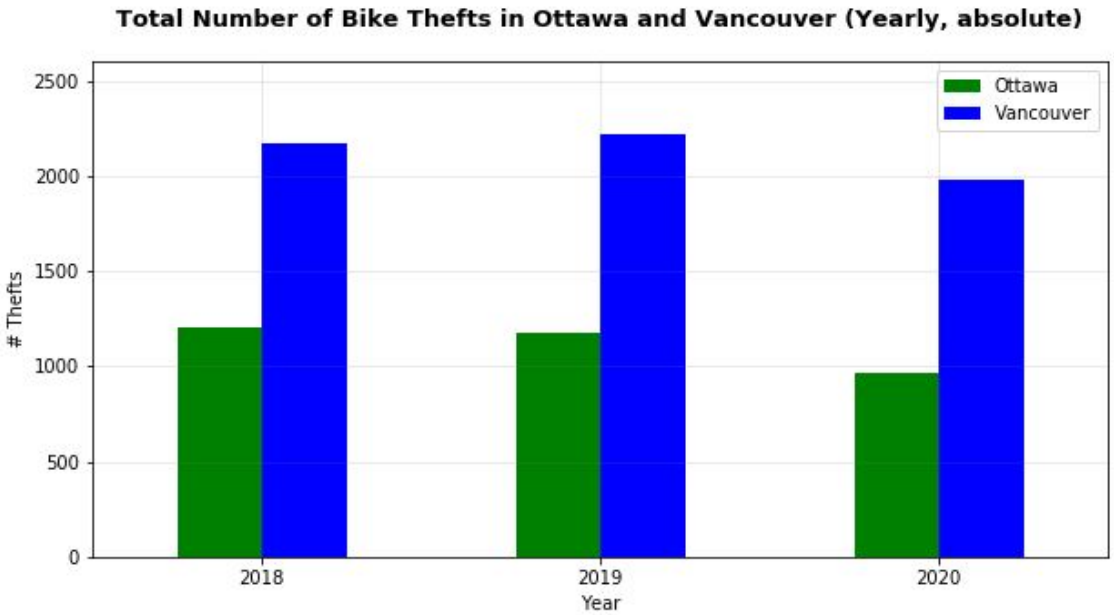
- **Bike thefts decreased** during covid in both cities
- **Bike volume decreased** in Ottawa during covid and stayed the same in Vancouver
- Bike Volume and Bike Thefts trends **move together** and have a significant direct relationship for both cities. However bike volume is not the core driver of bike theft
- While **no significant correlation** was found **between COVID cases and biking volume/thefts** for the entire 2020, there was a significant correlation between COVID cases and bike volumes in May-Oct 2020
- There is a **partial relationship** observed between number of bike thefts and covid cases in Vancouver

3. BIKE THEFTS SITUATION ANALYSIS

BIKE THEFTS YEARLY TREND

Key Takeaways

- Vancouver has more bike thefts in total and per capita compared to Ottawa
- Total number of thefts in both cities decreased in 2020



Ottawa Bike Theft Volume

YEAR	THEFTS			sum per capita
	sum	mean	median	
2018	1204	100.33	70.0	12.89
2019	1175	97.92	80.5	12.58
2020	962	80.17	72.5	10.30

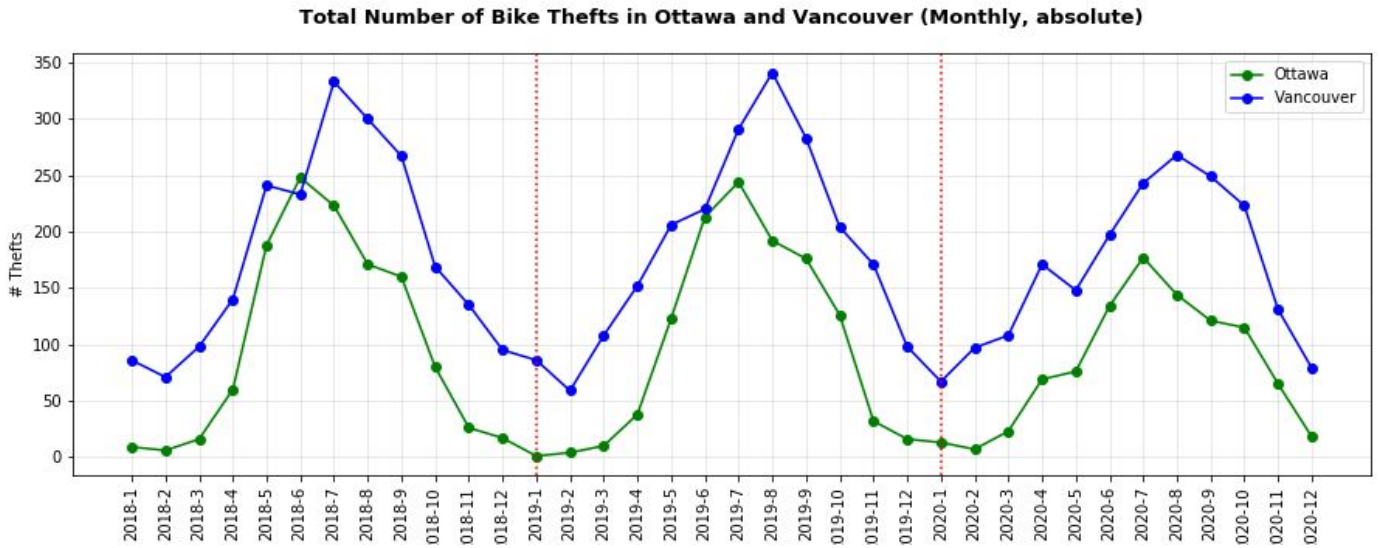
Vancouver Bike Theft Volume

YEAR	THEFTS			sum per capita
	sum	mean	median	
2018	2168	180.67	154.5	34.33
2019	2219	184.92	187.5	35.14
2020	1981	165.08	159.5	31.37

BIKE THEFTS MONTHLY TREND

Key Takeaways

- Bike thefts follow a cyclical trend, with most crime incidents happening in May through September
- There are less thefts in Ottawa compared to Vancouver in absolute numbers



Future opportunities: compare per capita

STATISTICAL ANALYSIS: BIKE THEFT SIGNIFICANCE

Key Takeaways

- Statistics confirms the difference in 2019 and 2020 theft data for both cities

Conducted T-Tests to check if bike volume from 2019 was significantly different from 2020

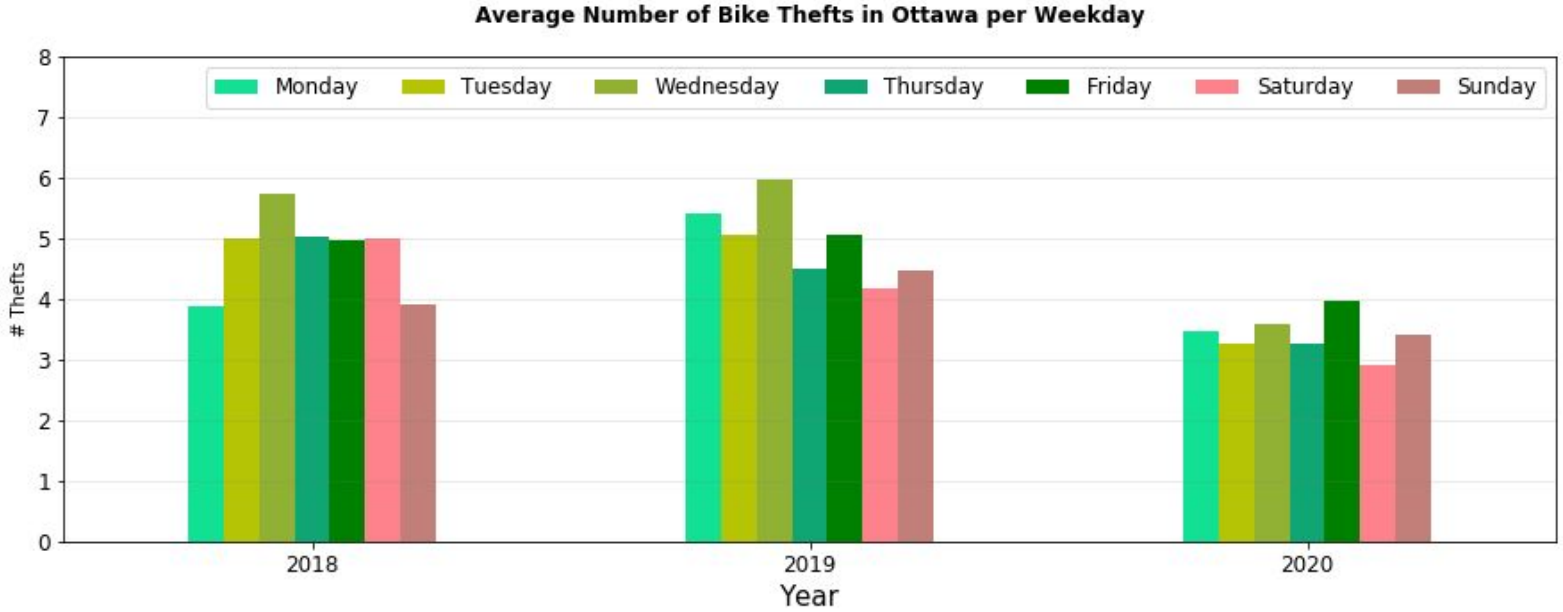
Vancouver Bike Theft	Ottawa Bike Theft
2018 vs 2019 **P Value: 0.61**	2018 vs 2019 **P Value: 0.69**
2019 vs 2020 **P Value: 0.02**	2019 vs 2020 **P Value: 0.002**

Used daily thefts to do t-test (more data points)

BIKE THEFTS DAILY: OTTAWA

Key Takeaways

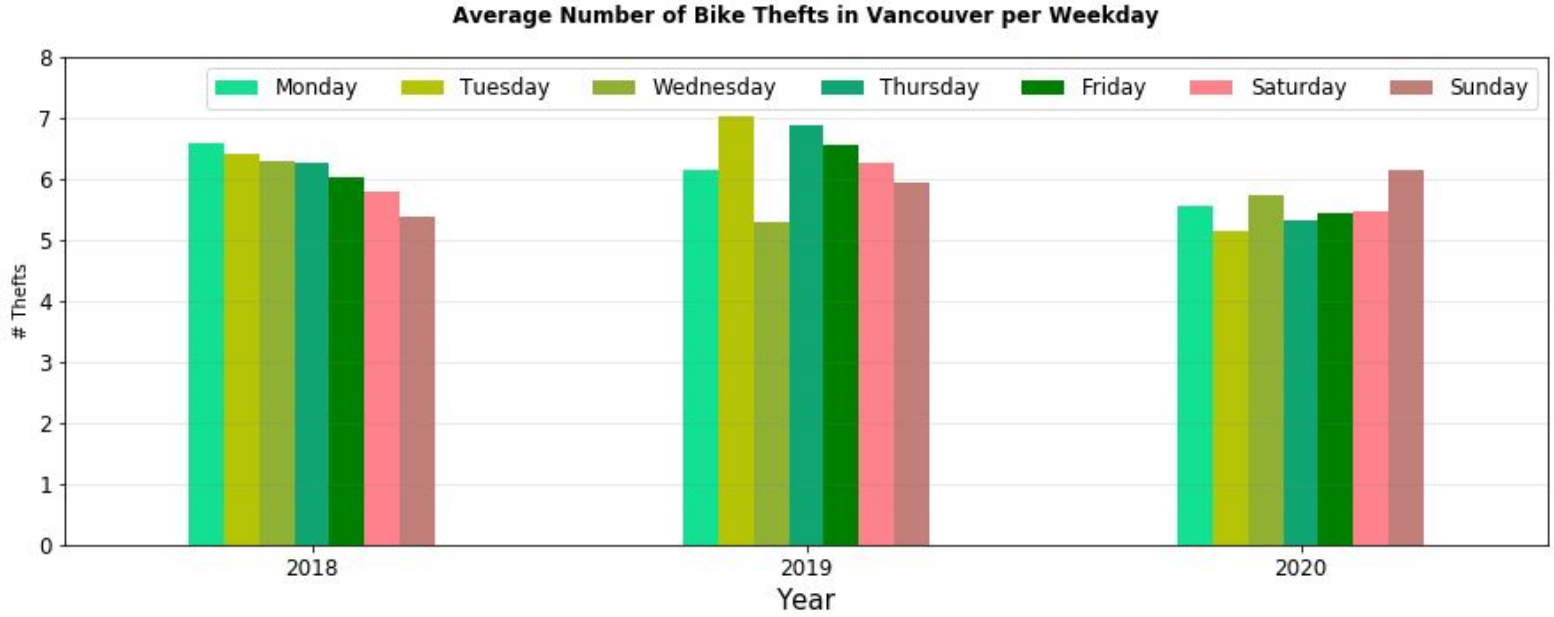
- On average the number of bike thefts in Ottawa was higher during weekdays in 2018-20
- Thefts were distributed more evenly across weekdays and weekends in 2020
- Daily average dropped across all days of the week in 2020



BIKE THEFTS DAILY: VANCOUVER

Key Takeaways

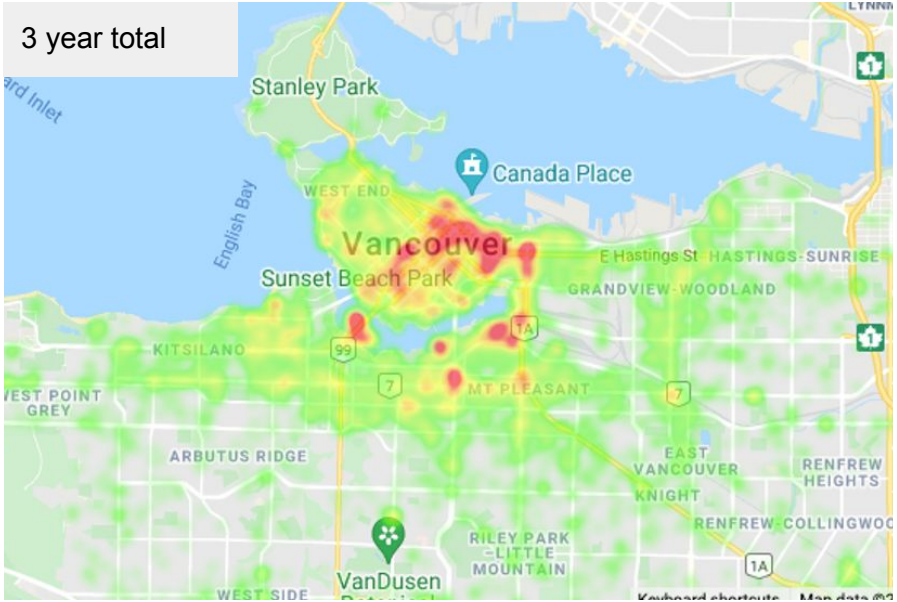
- While in 2018-19 the average number of bike thefts during weekends was lower, in 2020 the opposite trend is observed
- Overall, the daily average dropped (excl Sun) and the distribution of thefts was more even in



BIKE THEFT LOCATIONS

Key Takeaways

- Bike theft incidents take place predominantly around downtown / central area in both cities
- Ottawa theft incidents is more fragmented (could be due to data reporting issue)



BIKE THEFTS KEY OBSERVATIONS

Key Takeaways

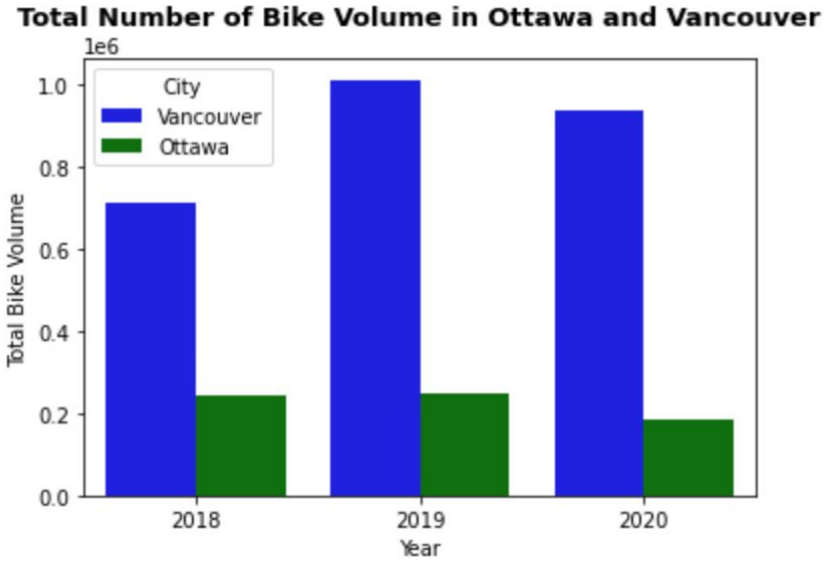
- Total number of **thefts** in both cities **decreased** in 2020
- **Vancouver has more bike thefts** in total and per capita compared to Ottawa
- Bike **thefts** f**ollow a cyclical trend**, with most crime incidents happening in May through September
- On average the number of bike thefts was **higher during weekdays**, except for Vancouver where there was a rise in theft incidents on Sundays

4. BIKE THEFT RELATIONSHIP ANALYSIS

BIKE VOLUME YEARLY TREND

Key Takeaways

- On Average, there are **4x more bikers in Vancouver** than Ottawa



Vancouver/Ottawa

2018	2019	2020
3.0	4.0	5.0

AVG Vancouver Bike Volume

2018	2019	2020
712.0	1013.0	939.0

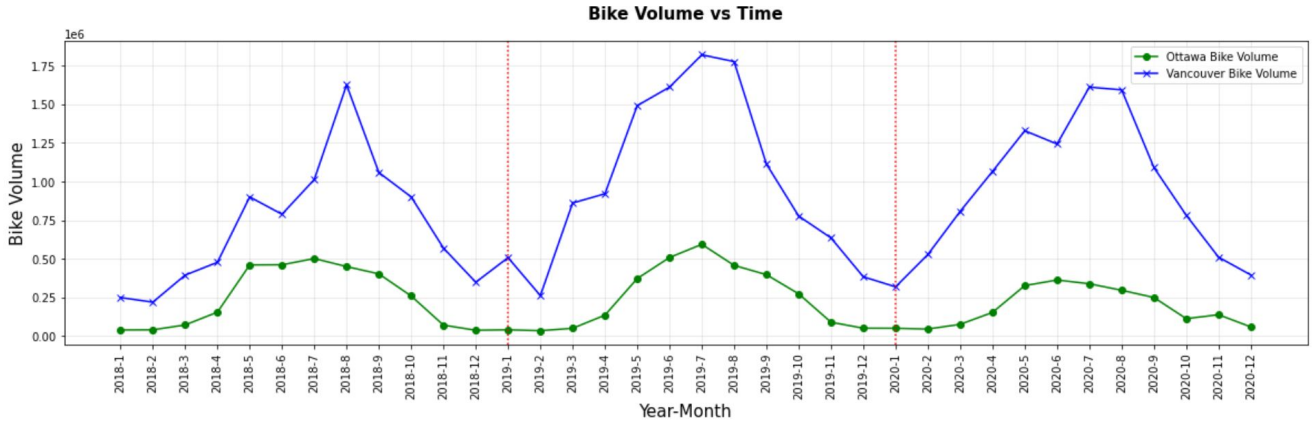
AVG Ottawa Bike Volume

2018	2019	2020
245.0	250.0	184.0

BIKE VOLUME MONTHLY TREND

Key Takeaways

- Both cities follow a **seasonal trend** with high number of bikers in summer



Vancouver Bike Volume

2018	2019	2020
8540372.0	12155570.0	11269963.0

Ottawa Bike Volume

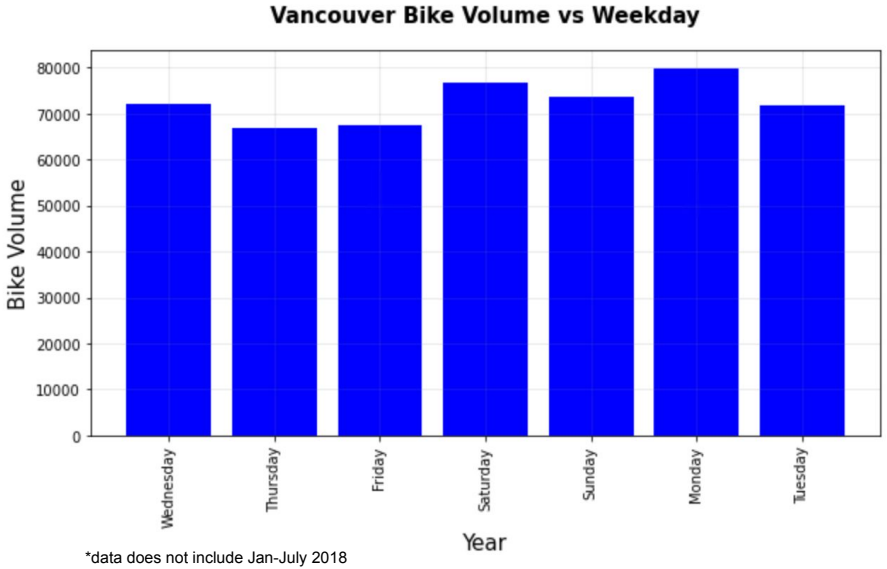
2018	2019	2020
2945947.0	2999935.0	2205156.0

Future opportunities: compare per capita

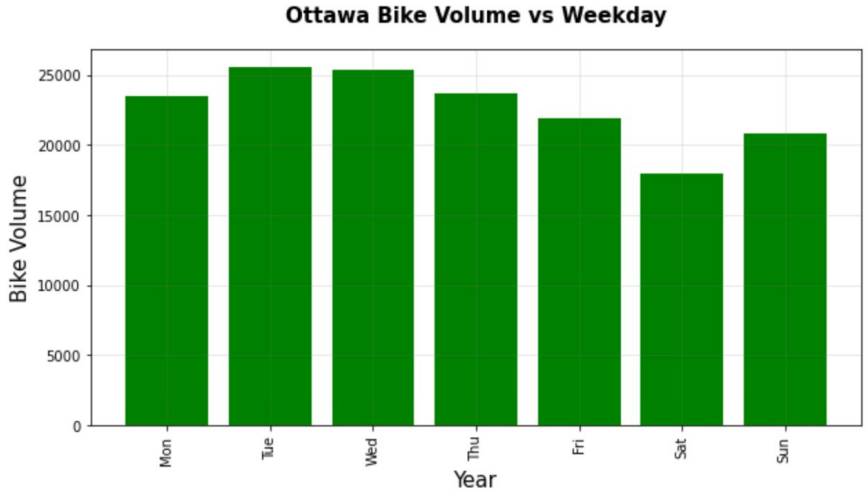
BIKE VOLUME WEEKLY TREND

Key Takeaways

- Ottawa has more bike volume during weekdays!



*data does not include Jan-July 2018



STATISTICAL ANALYSIS: BIKE VOLUME T-TEST

Key Takeaways

- Vancouver: Bike Volume **did not increase** during Covid
- Ottawa: Bike Volume **decreased** during Covid

Conducted T-Tests to check if bike volume from 2019 was significantly different from 2020

AVG Vancouver Bike Volume

2019	2020
1013.0	939.0

P Value: 0.07

AVG Ottawa Bike Volume

2019	2020
250.0	184.0

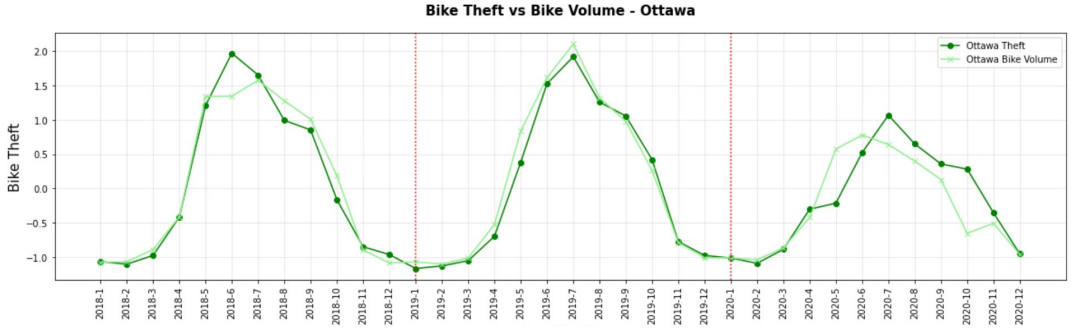
P Value: 1.16e-06

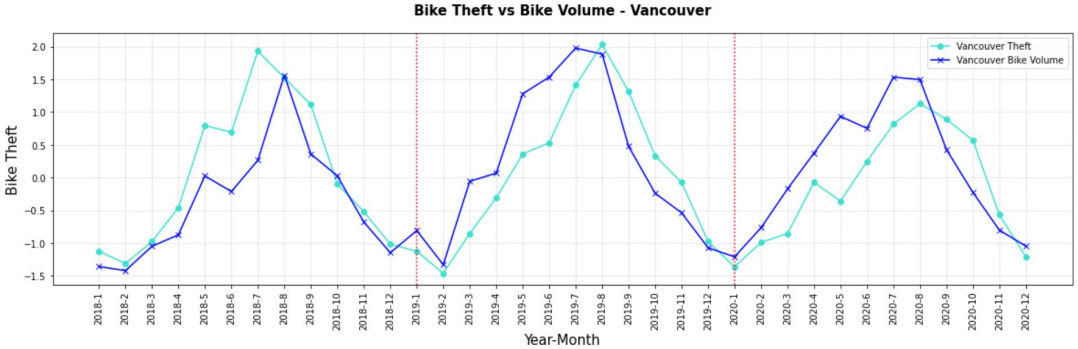
Used daily bike volume to do t-test (more data points)

BIKE VOLUME VS BIKE THEFT MONTHLY TREND

Key Takeaways

For both cities, Bike Volume and Bike Thefts move closely together





Used Z-scores to standardise thefts and volume to plot on same scale

STATISTICAL ANALYSIS: BIKE VOLUME VS BIKE THEFTS

Key Takeaways

- Bike Volume and Bike Thefts have a significant relationship for both cities but volume is not the core driver of bike theft

Assess if there is a relationship between theft and biking volume:

1. Calculate average number of thefts each year
2. Conduct T-Test to see if the increase/decrease in thefts is significant
3. Regress bike volume against bike theft for each year (2018, 2019, 2020)

REGRESSION ANALYSIS: BIKE THEFT VS BIKE VOLUME

Key Takeaways

- All regressions were significant with R-squared close to 1 and p-value lesser than 0.05

Vancouver

2018:

- Equation: y = 0.0002x + 41.41
- R-squared: 0.79
- P-Value: 8.64e-05
- Corr: 0.89

2019:

- Equation: y = 0.0001x + 42.6
- R-squared: 0.74
- P-Value: 1.19e-07
- Corr: 0.86

2020:

- Equation: y = 0.0001x + 45.54
- R-squared: 0.67
- P-Value: 1.48e-06
- Corr: 0.82

Ottawa

2018:

- Equation: y = 0.0004x - 11.72
- R-squared: 0.95
- P-Value: 4.53e-08
- Corr : 0.98

2019:

- Equation: y = 0.0004x - 11.06
- R-squared: 0.98
- P-Value: 3.07e-10
- Corr: 0.99

2020:

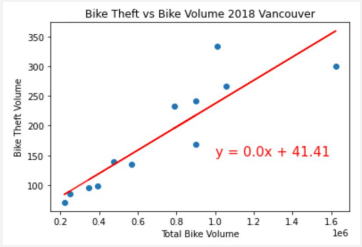
- Equation: y = 0.0004x + 7.51
- R-squared: 0.71
- P-Value: 0.006
- Corr: 0.84

If there is 10K in bike volume, there can be 11-12 bike thefts expected (0.0004*10K)+7.51

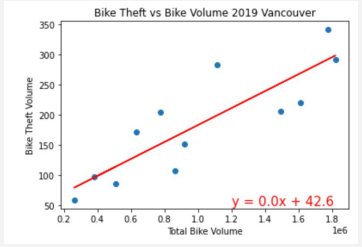
REGRESSION ANALYSIS: BIKE THEFT VS BIKE VOLUME

Vancouver

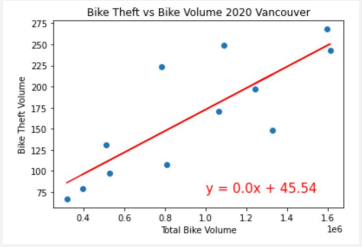
2018:



2019:

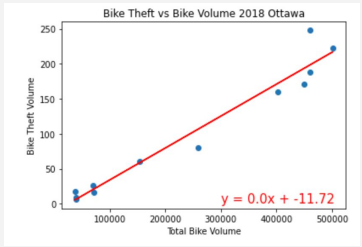


2020:

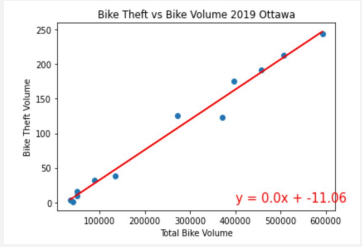


Ottawa

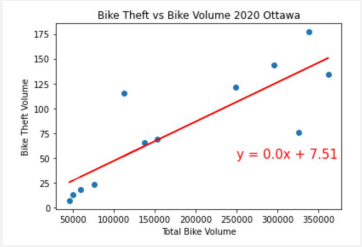
2018:



2019:



2020:



BIKE THEFTS VS BIKE VOLUME KEY OBSERVATIONS

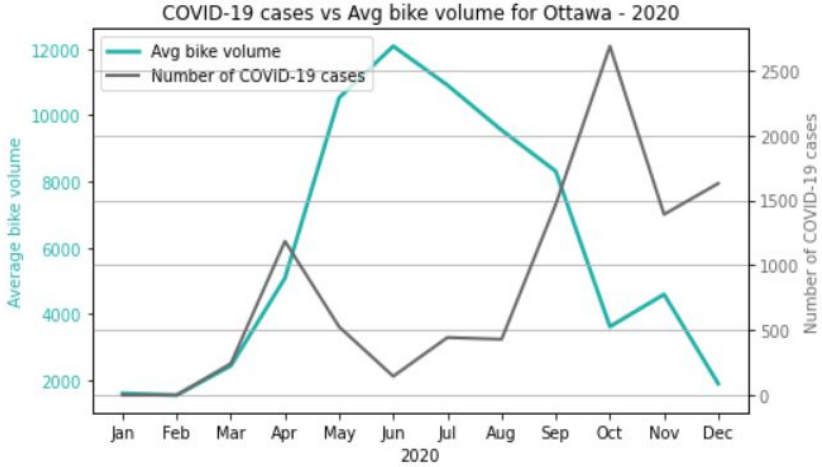
Key Takeaways

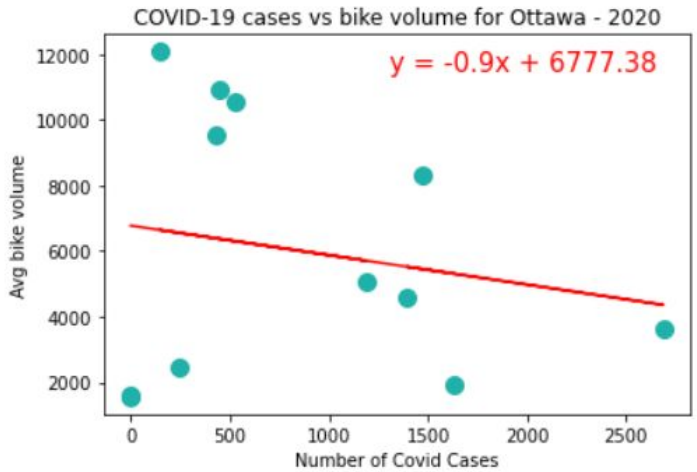
- Vancouver: Bike Volume **did not increase** during Covid
- Ottawa: Bike Volume **decreased** during Covid
- For both cities, Bike Volume and Bike Thefts **move closely together**
- Bike Volume and Bike Thefts have a significant relationship for both cities but **volume is not the core driver of bike theft**

COVID CASES VS NUMBER OF BIKERS (Ottawa) 2020

Key Takeaways

- No relationship observed between number of bikers and covid cases in Ottawa.
- Correlation Coefficient was -0.3 for the daily data.



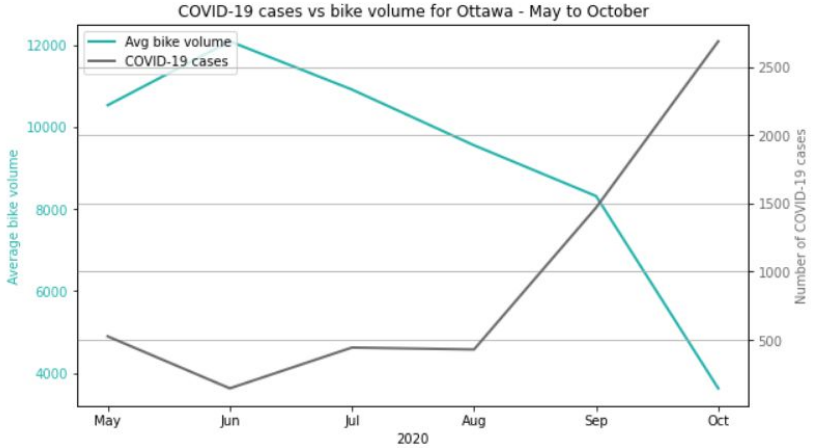


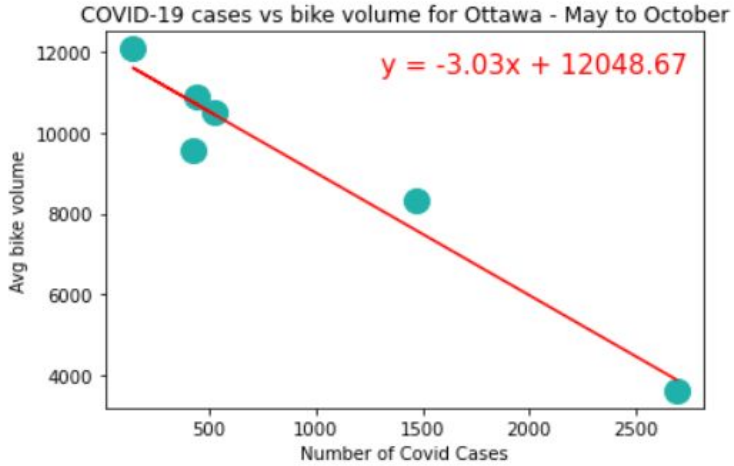
Correlation Coefficient is -0.19

COVID CASES VS NUMBER OF BIKERS (Ottawa) May - Oct

Key Takeaways

Increase in COVID cases shows a decrease in bike volume in Ottawa between May and October 2020



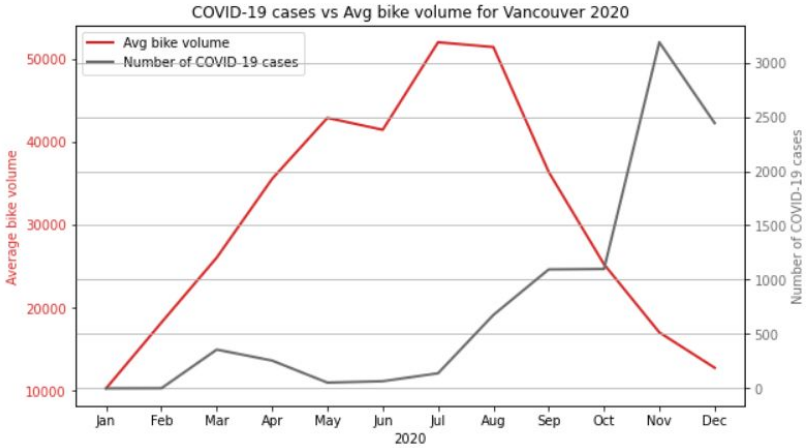


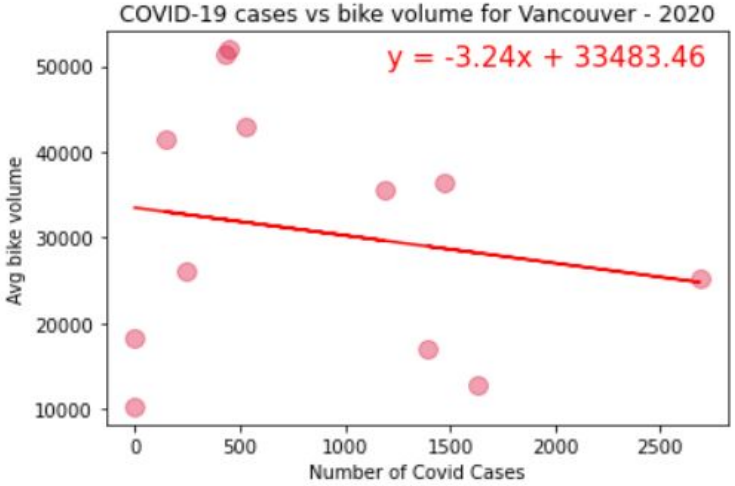
Correlation Coefficient is -0.97

COVID CASES VS NUMBER OF BIKERS (Vancouver) 2020

Key Takeaways

No relationship observed between number of bikers and covid cases in Vancouver





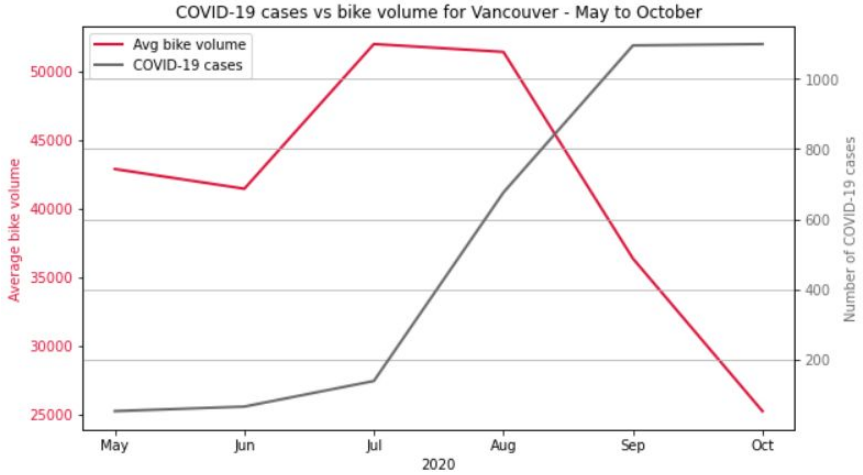
Correlation Coefficient is -0.18

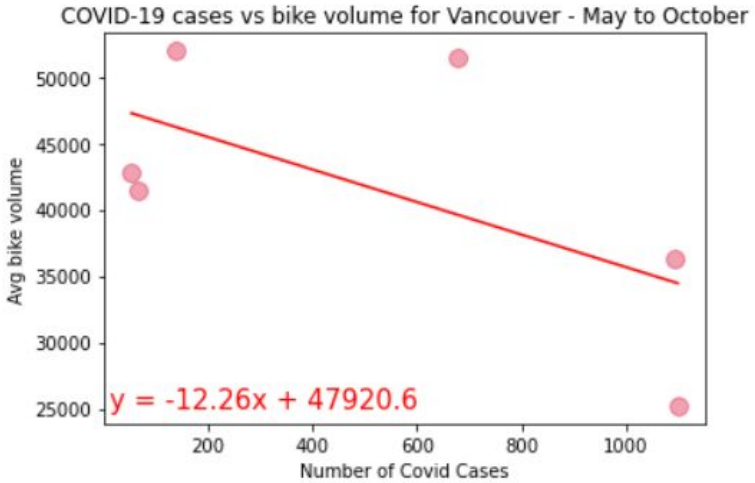
COVID CASES VS NUMBER OF BIKERS (Vancouver) May - Oct

Key Takeaways

The correlation coefficient moved from -0.18 to -0.61 for May - October

As the covid cases increases, the bike volume decreases when the cases are high



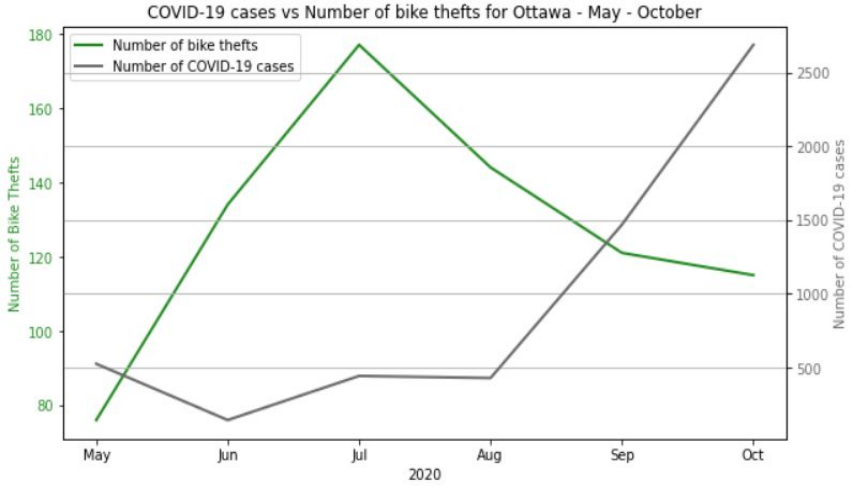


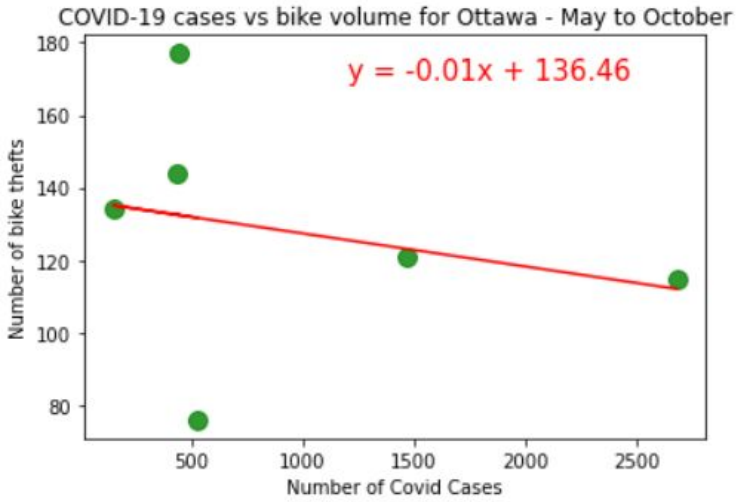
Correlation Coefficient is -0.61

COVID CASES VS BIKE THEFTS (Ottawa) May - October

Key Takeaways

No strong relationship observed between number of bike thefts and covid cases



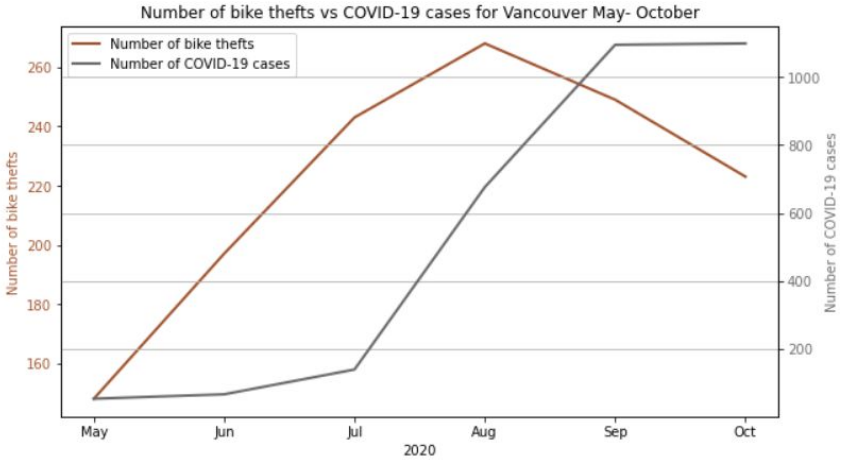


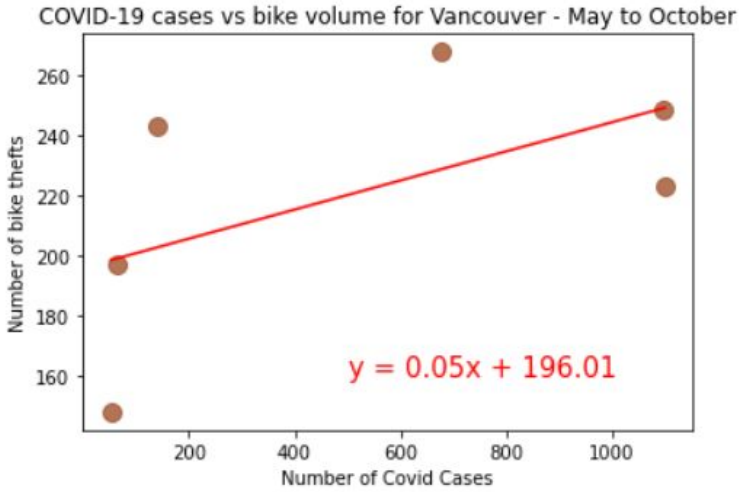
Correlation Coefficient is -0.26

COVID CASES VS BIKE THEFTS (Vancouver) May - October

Key Takeaways

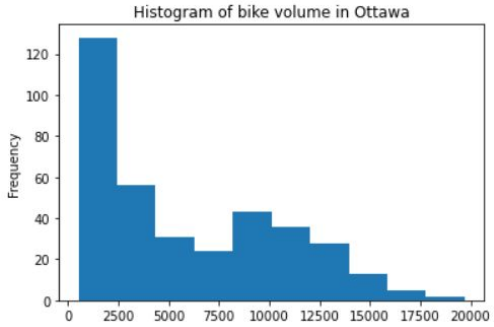
- There is a partial relationship observed between number of bike thefts and covid cases
- The increase in covid has increased the thefts

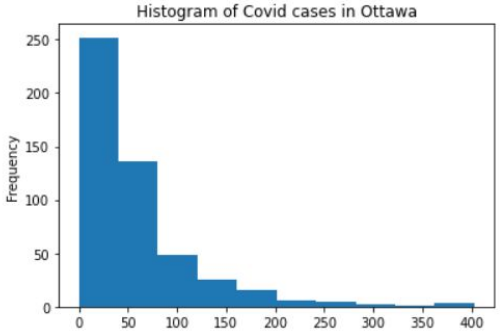


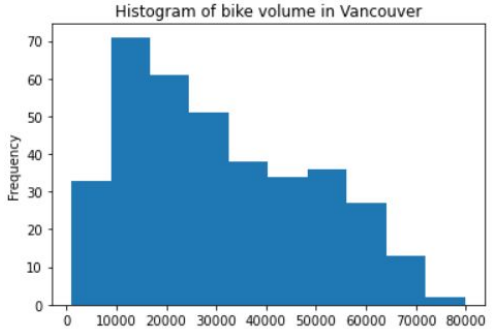


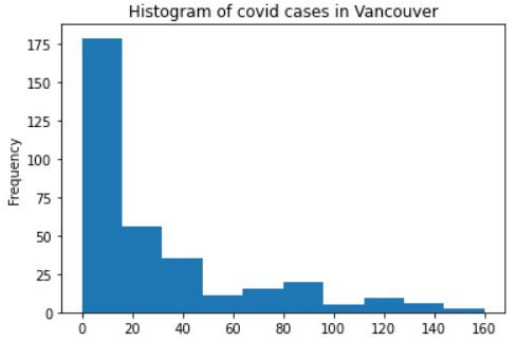
Correlation Coefficient is 0.56

STATISTICAL ANALYSIS: COVID CASES VS BIKE VOLUME









Summary of data grouped by months

	Thefts	Covid Cases	Bikes Volume
count	12.000000	12.000000	12.000000
mean	80.166667	845.916667	6017.075470
std	57.682254	827.295974	4004.088422
min	7.000000	0.000000	1558.620690
25%	21.750000	218.500000	2308.580645
50%	72.500000	485.000000	4842.716667
75%	124.250000	1411.250000	9793.580645
max	177.000000	2687.000000	12082.366667

CONCLUSION

During COVID the bike theft incidents decreased in both Vancouver and Ottawa. The drop in bike thefts was closely followed by a decrease in the bike traffic, which in turn could also be affected by an increase in COVID cases, especially during May-Oct period. The strong correlation between the variables suggested that changes in bike ridership do have an impact on number of thefts, however bike volume is not the core driver of bike theft.

Research Questions

- Explore if Ottawa & Vancouver have seen similar US trends in bicycle thefts due to covid
- Identify the relationship between bike theft rates and total bike volume
- Identify the relationship between the bike theft rates and covid cases

Findings

- The trends in US have not translated to Ottawa and Vancouver, as there was a decreased in bike theft incidents during covid in both cities
- Similar to bike thefts, bike volume also decreased, and it was found that Bike Volume and Bike Thefts trends move together and have a significant direct relationship for both cities.
- There was no significant correlation between Covid cases and biking volume/thefts for the entire 2020, however there was a significant correlation between COVID cases and bike volumes in May-Oct 2020. There was also a partial relationship observed between number of bike thefts and covid cases in Vancouver

5. LIMITATIONS AND FURTHER RESEARCH

LIMITATIONS

- Limited number of cities in the analysis due to lack of data for other cities in Canada
- Bike index API - The date of registration of bikes was not supported.
- Bike Volume Dataset
    - Inconsistent data format for Vancouver
    - Ottawa bikers not entire city
- Geographical locations - Ottawa theft locations coordinates were deliberately modified due to privacy, so the analysis was based on the Neighborhood locations, rather than an actual incident location
- 2021 data was not available in all cases to compare covid trends with 2020.

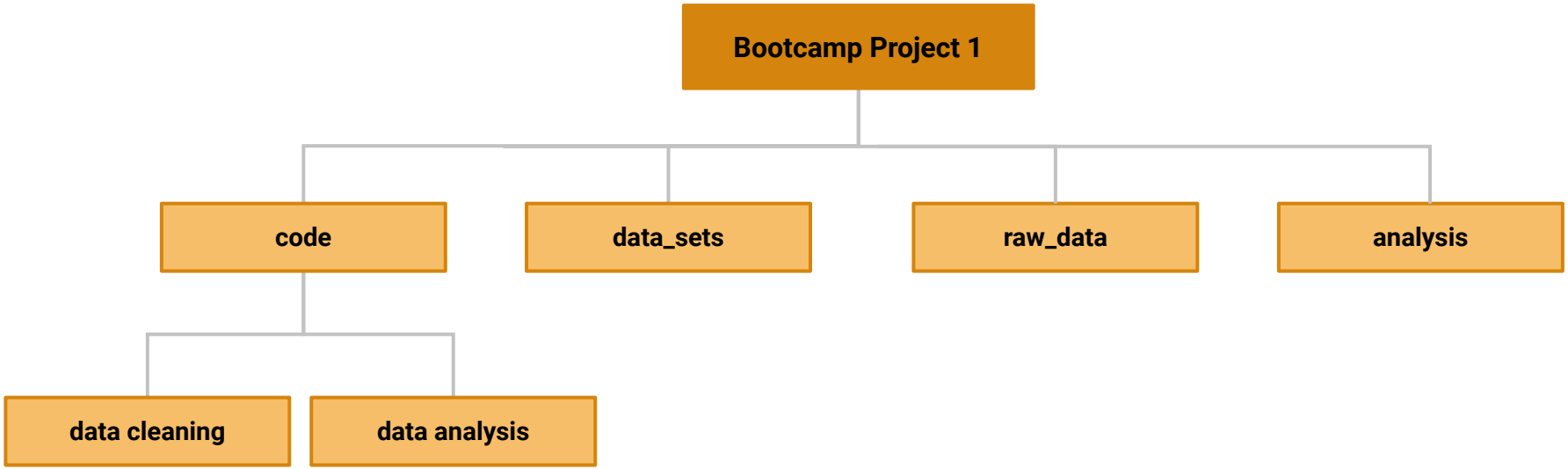
FURTHER RESEARCH

- Obtain data for 2021 and see if similar trends continue from 2020
- Analyse if US cities in question have also seen strong correlation with bike riderships
- Explore bike sales in the same cities - new vs used
- Most of the major cities police departments suggest bike registration with Garage 529/ Bike Index. Analyzing these areas could show more details on bike thefts and awareness.
- Explore few cities in Europe/Australia on same topics
- Explore weather trends and income based segmentation

QUESTIONS

6. APPENDICES

GITHUB WORKFLOW



Development Requirements

You will also be responsible for:

- Using Pandas to clean and format your dataset(s).
- Creating a Jupyter Notebook describing the **data exploration and cleanup** process.
- Creating a Jupyter Notebook illustrating the **final data analysis**.
- Using Matplotlib to create a total of 6–8 visualizations of your data (ideally, at least two per "question" you ask of your data).
- Saving PNG images of your visualizations to distribute to the class and instructional team, and for inclusion in your presentation.
- (Optional) Using at least one API if you can find an API with data pertinent to your primary research questions.
- Creating a write-up summarizing your major findings. This should include a heading for each "question" you asked of your data and a short description of your findings and any relevant plots.

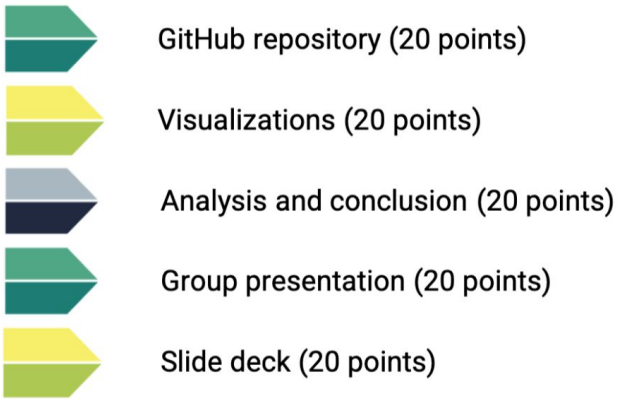
Presentation Requirements

You will also be responsible for preparing a formal, 10-minute presentation that covers:

- Questions you found interesting and what motivated you to answer them
- Where and how you found the data you used to answer these questions
- The data exploration and cleanup process (accompanied by your Jupyter Notebook)
- The analysis process (accompanied by your Jupyter Notebook)
- Your conclusions, which should include a numerical summary and visualizations of that summary
- The implications of your findings: What do your findings mean?

Rubric at a Glance

Categories for grading



GitHub repository (20 points)

Visualizations (20 points)

Analysis and conclusion (20 points)

Group presentation (20 points)

Slide deck (20 points)