Final Report

CITY OF KELOWNA: BIKE NETWORK ANALYSIS

Marzieh Rafieenia, Liza Wood, Mohsen Zardadi

June 21, 2019

City of Kelowna:
Bike Network Analysis Final Report

Table of Contents

Executive Summary

In 2018, the City of Kelowna entered into a license agreement with Dropbike to operate a dockless bikeshare pilot in and around the downtown core. The bikes were tracked by the user's cell phone GPS through the Dropbike app. The City's Active Transportation team recognized that this GPS data could help the City understand the routes used by cyclists which would then inform decision-making for infrastructure improvements. Using OSMnx and NetworkX in Python we converted the map of Kelowna into a graph network to map inaccurate, infrequent GPS points to the nearest street intersection, calculate the potential paths taken by cyclists and count the number of trips by street segment. Combined with the data from four counters around downtown, we used a mixed effects statistical model and a least squares optimization to estimate a relationship between the very different traffic patterns of the bikeshare and counter data. That relationship was used to estimate and visualize the Annual Daily Bicycle volume in downtown Kelowna. This visualization helped us understand how the bike network was being used, how cyclists crossed Highway 97 and how laneways were used by cyclists.

Background

According to the *Kelowna 2030 Official Community Plan*, the community envisions "urban communities that are compact and walkable" with "walking paths and bicycle routes… to key destinations". (City of Kelowna, 2011)

In response to this vision, the City of Kelowna developed *Kelowna On the Move: Pedestrian and Bicycle Master Plan* to identify "infrastructure, planning and policy requirements to promote and facilitate walking and cycling throughout the community". The report also shares the Active Transportation Vision, which "seeks to improve safety, connectivity, and accessibility by:

- Improving the quality and attractiveness of pedestrian and cycling facilities by establishing a low-stress Primary Network for users of all ages and abilities;
- Reducing conflicts due to truck, transit, and bicycle network overlaps;
- Enhancing route connectivity and continuity with new routes through gap areas; and
- Adding connectivity through high speed, high vehicle traffic volume areas with new connections and direct routes, including grade-separated crossings where appropriate."

One of the goals of the Pedestrian and Bicycle Master Plan is to "increase year-round walking and cycling so that within 20 years, 25 per cent of all trips less than five kms in length are made by walking and cycling." (City of Kelowna, 2016)

Kelowna currently has over 70 km of off-road pathways and 280 km of bike lanes available. (City of Kelowna, 2019) According to Pedestrian and Bicycle Master Plan, future infrastructure projects are currently prioritized based on the following criteria with respect to utility:

- Geographic area
- Closing gaps between existing infrastructure
- Connectivity to transit
- Primary Network route
- Connectivity to schools

Bicycle and Pedestrian Counters

To help the City understand pedestrian and bicycle activity around downtown Kelowna, there are eight counters in key locations. Five of those counters are able to distinguish bicycle traffic from pedestrian traffic. They can also distinguish the direction of traffic. These counters collect real-time data and upload the data to the City via cell uplink once per day. Since they have started collecting data at their respective locations, these five counters have counted 1,831,042 cyclists over a period of two to five years. (City of Kelowna, 2019)

2018 Dropbike Bikeshare Pilot

In February 2018, the City of Kelowna entered into a license agreement with Dropbike to operate a bikeshare pilot in and around the downtown core. Dropbike is a dockless system, so users were not confined to starting and stopping their trips at specific locations. The bikes were tracked by the user's cell phone GPS through the Dropbike app. Because of this, the City recognized that the pilot would provide data to help the City understand the routes used by cyclists which would then inform decision-making for infrastructure improvements. (City of Kelowna, 2018)

The pilot ran from 11 June to 11 November 2018. During the first three months of the pilot, there were "more than 33,000 (trips) from 9,000 unique users on 331 bikes… roughly a third of the total trips were made by 600 frequent users…" For a community the size of Kelowna, these were strong indicators of the potential for a bikeshare system going forward. (Worona, 2019)

The City has the Dropbike GPS data for each uniquely identifiable trip and has done some visualizations similar to Figure 1 and Figure 2 shown below.

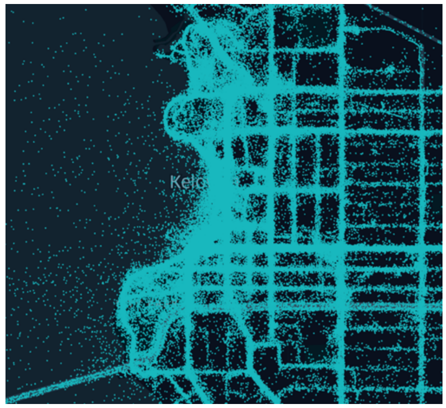

Figure 1: Bikeshare GPS points in and around downtown Kelowna, 11 June – 9 Sept 2018

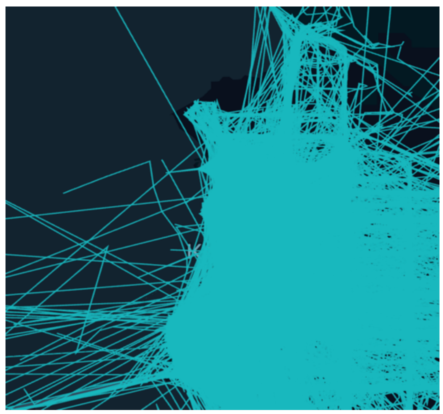

Figure 2: Lines connecting bikeshare GPS points per trip, 11 June - 9 Sept 2018

Visualizations like Figure 1 give a sense of where bikeshare riders have cycled, along with relative density of bikeshare traffic. When the points are connected to visualize the individual routes, as shown in Figure 2, the routes are indistinguishable. The City needs a more quantified analysis to inform decision-making for infrastructure improvements. In addition, bikeshare data only represents a portion of all the bicycle traffic in Kelowna.

Using Bikeshare GPS and Counter Data to Analyze Kelowna's Bike Network

Similar studies using a combination of bike share GPS data, GPS data from a city app and/or counter data have been done in San Francisco (Proulx & Pozdnukhov, 2017) , Montreal (Strauss, 2015) and Seattle (Strava Metro, 2019). In these studies, the ratio of total bicycle traffic to bikeshare traffic at the counter locations was found. The Average Annual Daily Bicycle traffic (AADB) of each street segment was calculated from the GPS data, scaled up by that ratio, then visualized as the example in Figure 3 from the Montreal study:

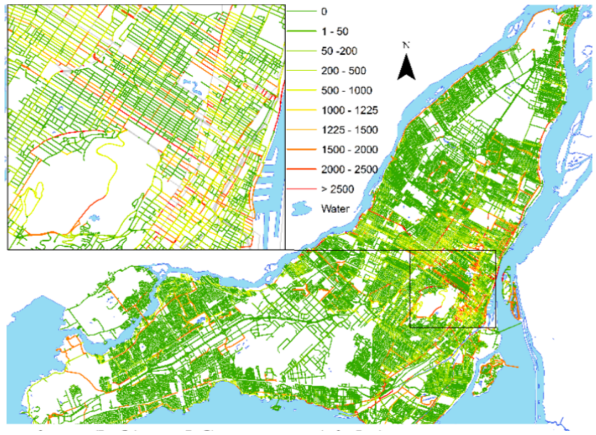

Figure 3: AADB by segment produced from combining GPS and counter data, Montreal

For this analysis, we combined the bikeshare GPS data with the counter data to estimate and visualize the Average Daily Bicycle (ADB) volumes in downtown Kelowna from 11 June to 9 September 2018. The first three months of the bikeshare pilot had the most bikeshare rentals and the largest, most consistent fleet of bikes available. The specific area is bounded by the following four counters:

- City Park Multi (south-west corner)
- Waterfront Walkway (north-west corner)
- Cawston Street Corridor (north-east corner)
- Ethel Street ATC (south-east corner)

The Sunset Drive North of Brandts Creek Crossway counter is not included since it was out of commission during most of the bikeshare period.

This analysis will be valuable for city planning and prioritization of future infrastructure projects such as new bike lanes. Routes with high bicycle traffic where there is no current infrastructure for cyclists are gaps and opportunities for improvement in the existing network. The City also wants to understand which laneways are being used and how much they are used compared to streets and active transportation routes. Knowing how and where cyclists cross Highway 97 (Harvey Ave) will help the City prioritize improvement projects with the Province of British Columbia, who are responsible for the highway.

Challenges with the Bikeshare GPS Data

Nearly 9000 trips of data were gathered over the first 13 weeks of the pilot. This is an average of 99 cyclists per day, a small portion of which pass one of the four counters. During this same period, the four counters counted 232,835 cyclists, an average of 640 cyclists per counter per day. Therefore, we have relatively few bikeshare trips to determine a statistically significant relationship between the bikeshare and counter data.

In ideal conditions, cell phone GPS has a location accuracy of 4.9m, but that accuracy worsens near buildings, bridges and trees. (GPS.gov, 2017) As shown in Figure 4, we have many GPS points that are far away from the nearest path, including in the middle of Okanagan Lake. As a result, we needed to find a way to assign the GPS points to the path and find a route between each of the points per trip.

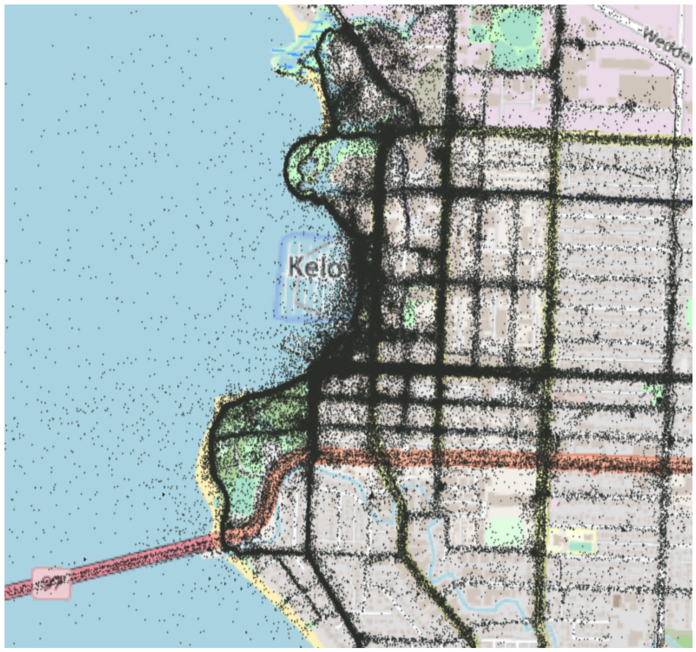

Figure 4: Map of bikeshare GPS points, downtown Kelowna

In addition, the frequency of the GPS updates for most of the trips is low, as shown in Figure 5. Ideally, the update frequency would only be a few seconds.

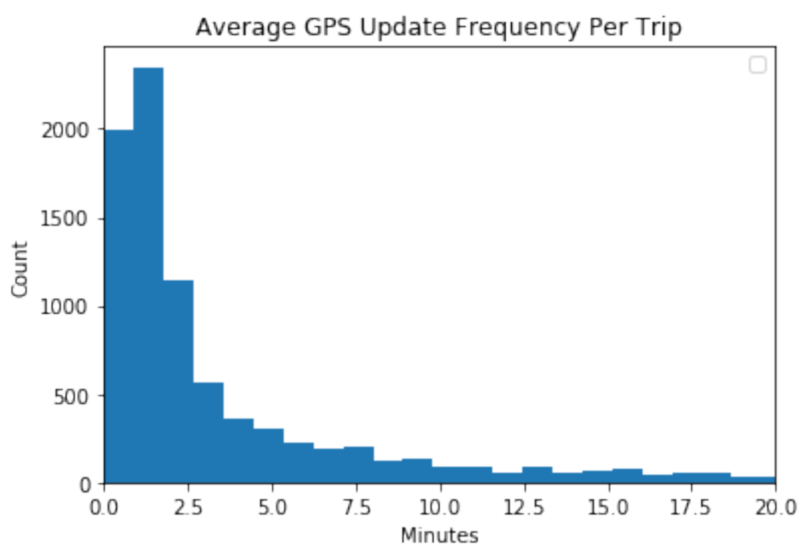

Figure 5: Distribution of Average GPS Update Frequency

The median update frequency is 2 minutes 18 seconds, during which an average cyclist could travel up to 400 meters. With that much distance between GPS updates, there is more than one possible route the cyclist may have taken to travel between GPS update points.

Finally, we only have one season of data. We will not be able to validate the accuracy of the ADB estimates year-over-year.

Applying Graph Theory and Network Analysis

Graphs are used to visualize and study the connections and relationships between things. These things are represented by ***vertices*** or ***nodes*** and they are connected by ***edges***. Given a set of nodes and edges, graphs can represent anything from social networks to electrical circuits to transportation networks to street maps. (Najera, 2018) In the case of street maps, the nodes represent intersections while the edges represent streets, lanes, or bike paths. Figure 6 below shows the map of downtown Kelowna represented as a graph.

Figure 6: Graph representation of downtown Kelowna

By representing our map as a graph, we are able to use helpful tools to map GPS points to nodes or edges, find and calculate the shortest path between those points, and understand how the nodes and edges relate to each other.

OSMnx and NetworkX are two Python packages provide those tools to create and analyze downtown Kelowna as a graph. OSMnx downloads street networks from OpenStreetMap, which is "built by a community of mappers that contribute and maintain data about roads, trails, cafés, railway stations… all over the world." (OpenStreetMap, 2019) OSMnx was created by an urban planning professor to "construct, project, visualize, and analyze complex street networks in Python with NetworkX." (Boeing, 2017) Figure 6 was created with OSMnx. NetworkX is a

“Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks.” (NetworkX Developers, 2019)

Data Cleaning

In the process of doing their own analysis of the bikeshare data, the City of Kelowna had identified which bikeshare trips and GPS points needed to be cleaned out of the data set. The City used the following criteria:

- Bounded to Kelowna
- minimum number of trip points = 3
- minimum segment length (metres) = 20
- minimum trip length (metres) = 150
- maximum trip length (metres) = 20000
- minimum trip duration (sec) = 120
- maximum trip duration (sec) = 10800 (3 hrs)
- maximum average trip speed (m/s) = 20/3.6 (20 km/h converted to m/s)
- maximum trip segment speed (m/s) = 30/3.6 (30 km/h converted to m/s)
- maximum fast segments = 0.1
    - i.e. 10% of segments can exceed maximum trip segment speed

While the trips and points violating these criteria were identified, they were still in the data set, which represented 22153 trips. We filtered out the identified points in Python, leaving us with 8853 trips and 97110 GPS points to work with.

Constructing Paths from GPS Points

There are algorithms in NetworkX to find the path between any two nodes in a graph. Therefore, the first step in our analysis used the `get_nearest_node` algorithm in OSMnx to calculate the nearest node to each GPS point and the corresponding distance. The amount of calculations needed to find the nearest node for each GPS point would take several hours on a single computer. We used Google’s Colaboratory to reduce computation time with cloud computing. With nearest node, each GPS point was represented by its nearest intersection on the map.

Many of the GPS points were in Okanagan Lake or travelled outside of Kelowna. We needed a way to exclude them from the analysis. We used the distance to the nearest node to filter out GPS points that were greater than 150m from the calculated nearest node. With those points removed, we then removed any trips with less than three points remaining. This left us with 8815 trips and 95905 GPS points.

We then used the `shortest_path` algorithm in NetworkX to calculate the path between those nodes. There were 72 trips for which the algorithm could not find a path. 59 of those trips passed through Highway 97 and Abbott intersection. This was mainly due to GPS inaccuracy or how cyclists are crossing that intersection. 59

of approximately 400 total trips crossing this intersection had this problem. To fix it, we mapped the problem nodes to the nodes of the path the cyclists are expected to take to cross that intersection. Figure 7 below shows how those nodes were remapped.

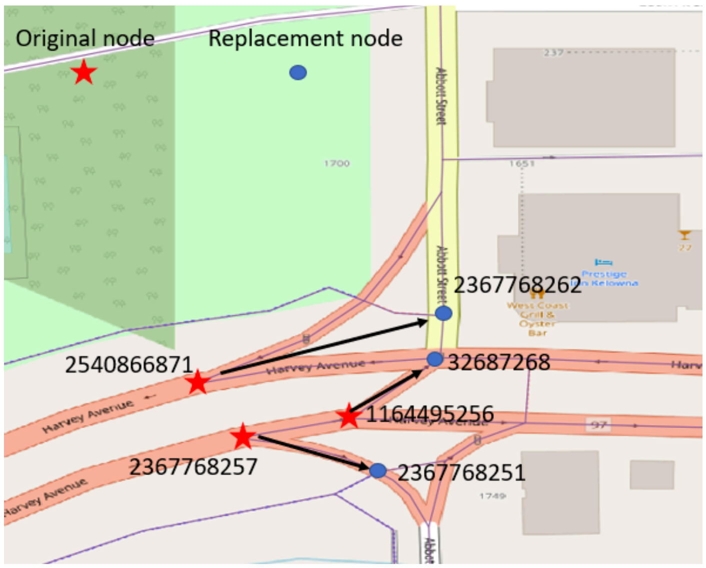

Figure 7: Remapping of problem nodes at Highway 97 and Abbott St.

Of the remaining problem trips, 7 were in the Strathcona Beach Parking lot and 9 were at the entrance to the library parkade. In both cases, there were nodes that did not connect back to a road or cycleway. These were fixed in the same way as the Highway 97 and Abbott intersection.

Graph edges can be given any numerical value, called ***weight***, to represent the distance or cost between nodes. The `shortest_path` algorithm calculates the path between two nodes with the minimum weight. Since cyclists do not always travel the shortest distance, we calculated our potential paths with the following weights:

- **Length** in metres of the street/path segments
- **Path Preference** – a numerical value given to each street/path category based on research done in Vancouver on cyclist path preferences (Winters & Teschke, 2010)
- **Simplified Path Preference** – numerical preference given to cycleways only
- **Weighted Length** – length of the street multiplied by the simplified preference value for cycleways
- **Closeness Centrality** – streets that have a higher degree of closeness centrality are preferred
- **Unbiased** – all streets were set to a weight of one.
- **Corner-weighted** – weights adjusted for Ethel, Cawston and the lakeside cycleways to match the split of bikeshare volumes at counter locations with the same split for the counter data
- **Corner-weighted Length** – the length of the street multiplied by the weights in the corner-weighted value for those edges.

Table 1 below shows how the values from the Winters & Teschke article were mapped to weights in our **Path Preference** and **Simplified Path Preference** path-finding models.

Table 1: Mapping of cyclist path preference to weight

Winter-Teschke Article		**Weights for Path-finding Calculation**		
Path Type	**Preference Factor (higher is better)**	**Path Type**	**Weight (lower is better)**	**Comments**
Paved off-street cyclepath - shared use	0.5	Cyclepath + Ethel	0.5	Ethel is classified as a tertiary street in OpenStreetMap.
Unpaved off-street cyclepath - shared use	0.4	N/A downtown		There is a short path behind the buildings on Sunset. Not worth breaking out.
Residential streets	0.1	Residential	0.9	
		Lanes, All Links, Footpaths, Unclassified	1	This is considered neutral.
Major streets with bike lanes - parked cars	-0.1	Not used		
Major streets with bike symbols - parked cars	-0.2	Secondary/ Tertiary	1.2	Unable to distinguish streets with bike lanes vs symbols in OpenStreetMap.
Major streets with parked cars	-0.5	Not used		
		Trunk (Hwy 97)	3	Rated very high since no one should be cycling directly on the highway. Any trips along here are assumed to be on the sidewalk.

Closeness centrality is a graph theory concept that measures the average shortest path between each node in the network and every other node. The more central a node, the closer it is to all other nodes. (Boeing G. , 2018) Edges bounded by nodes with high centrality are also considered to have high centrality. Figure 8 below shows the relative closeness centrality for downtown Kelowna. Since edges with higher closeness centrality have a higher numerical value, we assigned one minus the centrality as the weight in that graph. Edges with higher centrality would then be calculated as a shorter path than edges with lower centrality.

Figure 8: Closeness centrality for downtown Kelowna. Lighter streets have more centrality.

The **corner-weighted** and **corner-weighted length** path-finding models were created to get the same split of bikeshare counts at the counter locations as we have with the counter data. The counters are roughly at the four corners of our downtown map, as shown in Figure 9.

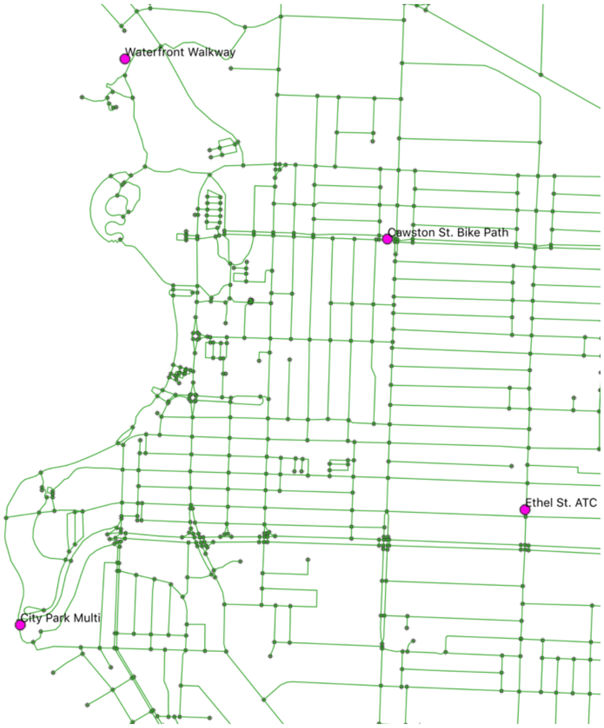

Figure 9: Downtown Kelowna counter locations (pink dots)

All four counters are on paved, off-street, shared-use paths: Ethel, Cawston and the lakeside path. Using the same logic as Simplified Path Preference, the weightings of these paths were iteratively adjusted until the split of the bikeshare counts at the counter locations were close to the same as those for the counter data. We started with weightings that represented the split of traffic between the counters. The weightings were increased or decreased until the counts from `shortest_path` algorithm resulted in the bikeshare counts at the counter locations having a similar split as the counter data.

Counting Bikeshare Trips

Once the paths for each trip were calculated in each of the path-finding models, we then needed to count how many bikeshare trips travelled along each street segment. To do this, the path for each trip was broken down into pairs of nodes. Each pair of nodes represents the intersections that bound each street or path segment. Each pair of nodes was summed to get the total number of trips for each segment.

The same was done for the counter locations. Three of the four counter locations are between two unique nodes. The exception is Cawston counter. With the poor accuracy of GPS points, the volume of bikeshare trips was split between the cycleway and the street. As a result, we had to add the total number of trips on both the street and the cycleway. Ethel did not have this problem since the Ethel St. cycleway is not a separate path on OpenStreetMap.

Visualization

The graphs with the bikeshare counts and ADB were exported from Python as shapefiles, a standard file format for mapping software. The counts and ADB for each path-finding model were then visualized in QGIS 3.6.

Calculating Average Daily Bicycle Volume

Finding the relationship between the counter data and the bikeshare data was the final step to scale up the bikeshare counts to the Average Daily Bicycle volume representing all cyclists. This was done in R, a programming language and environment used for a wide variety of statistical computing and graphics techniques. (The R Foundation, 2019) The relationship was found with a mixed effects model using daily bikeshare and counter counts at each counter location. There are more details in the next section.

Deliverables

Maps of each of the path-finding models and the final ADB calculations were provided as shapefiles to the City along with all of the Python and R code to replicate the work and do further analysis.

Analysis and Results

Each of the path-finding models was evaluated as follows:

- Visualization of counts.
- Percentage of segments within realistic speed bounds.
- Split of total bikeshare counts at each of the counter locations compared with the split of counter counts.
- Linear regression of bikeshare counts versus the counter data.

The path-finding model that performed the best in these evaluations was used to determine the ADB.

Visualization

Visualizing the counts in QGIS was our first verification of the path-finding model. Areas with no counts, wide variations in continuous paths, and a mismatch in relative size of the counts compared with the GPS point density indicated a problem with the model.

While Shortest Length performed the best in our other tests, the two areas circled in red in Figure 10 below were a concern in visualization. In the top left, more cyclists were counted along the sidewalk by Island Stage in Waterfront Park than along the boardwalk. In the bottom left, there is a noticeably smaller count in the lakeside segment after the City Park counter than in the rest of the segments along that path. The density of GPS points was relatively uniform throughout that area.

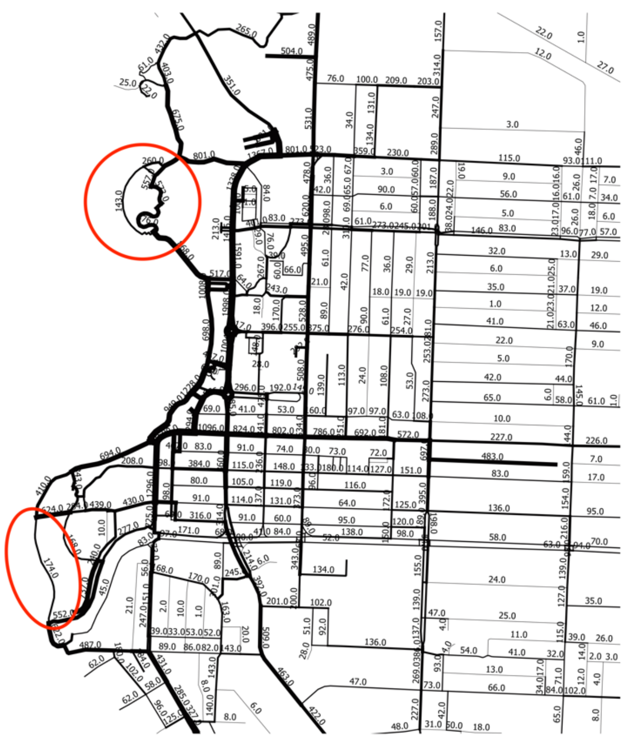

Figure 10: Bikeshare counts per path segment - shortest length model

Since St. Paul Street has the highest closeness centrality, that visualization has an unexpectedly high count along that street, circled in red in Figure 11 below. The Unbiased model also had unusually high counts along St. Paul Street and the surrounding area, circled in red in Figure 12 below. These relatively high counts were not reflected in the relative density of GPS points in those areas (see Figure 4).

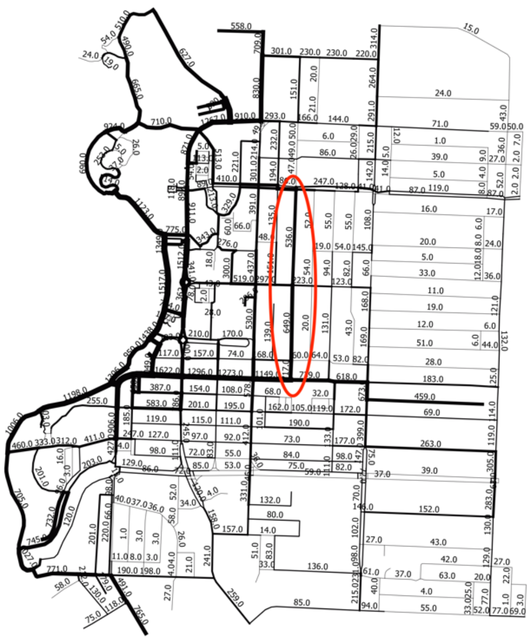

Figure 11: Bikeshare counts per segment – closeness centrality model

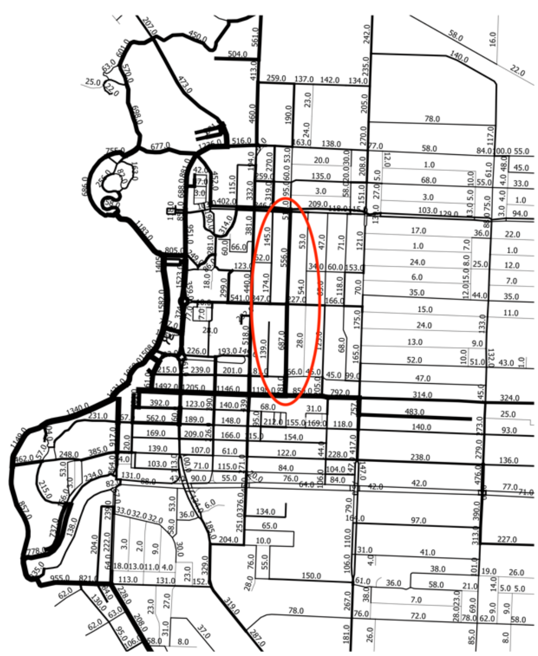

Figure 12: Bikeshare counts per street segment - unbiased model

All the other path-finding models did not have these issues and looked reasonably correct. After reviewing the Shortest Length, Path Preference, Closeness Centrality, Unbiased, and Weighted-Corner models with the City, the top three path-finding models based on the visualizations were Path Preference, Weighted-Corner and Shortest Length.

Percentage of Segments within Speed Bounds

The speed required to travel along the path between GPS points was way to gauge whether the calculated path was feasible. If the speed is too high, it is unlikely that path was travelled. Since cyclists need to stop at intersections, there is no minimum speed.

To determine the maximum speed, we calculated the speed between each pair of GPS points along the length of the calculated path in the Shortest Length model, using the timestamp for each GPS point. We then plotted a boxplot of the speeds, removing the outliers. The outliers are the infeasible speeds.

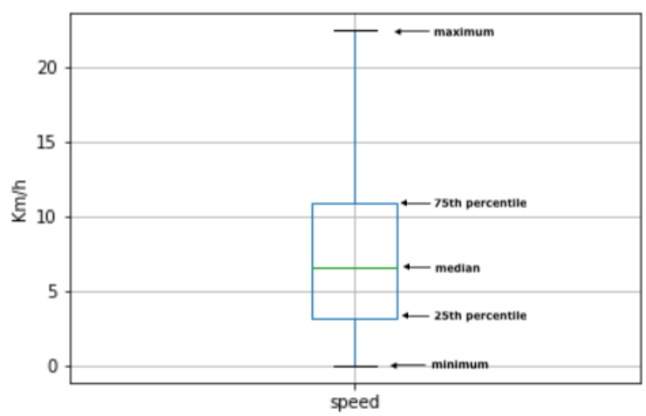

Figure 13: Boxplot of segment speeds - shortest length model

From the boxplot in Figure 13, the maximum speed is 22.5 km/h, which we set as the maximum for evaluating the other path-finding models. Table 2 below summarizes the percentage of segments under 22.5 km/h for each path-finding model. Shortest Length, Closeness Centrality and Corner-Weighted have the highest percentage of feasible segments based on speed.

Table 2: Percentage of segments under maximum speed for each path-finding model

Path-Finding Model	Segments Under 22.5 km/h
Shortest Length	88.10%
Path Preference	85.79%
Simplified Path Preference	86.19%
Weighted Length	85.32%
Closeness Centrality	87.80%
Unbiased	86.80%
Corner-Weighted	86.83%
Corner-Weighted Length	85.56%

Split of Bikeshare Counts by Counter Location

One of the base assumptions of this project was that the bikeshare data is representative of all bicycle traffic during the same period. Therefore, the best model would result in a similar split of bikeshare counts at the counter locations as we see in the counter data.

From the counter data, the split between counters is as follows:

- City Park: 40.7%
- Waterfront: 16.5%
- Cawston: 28.4%
- Ethel: 16.5%

For the bikeshare data, the split between counters for each of the path-finding models is summarized in Table 3.

Table 3: Bikeshare split between counter locations by path-finding models

Path-Finding Model	City Park	Waterfront	Cawston	Ethel
Shortest Length	41.3%	26.0%	22.9%	9.7%
Path Preference	45.5%	31.3%	14.2%	9.0%
Simplified Path Preference	47.0%	25.5%	15.4%	12.1%
Weighted Length	39.8%	23.5%	26.6%	10.1%
Closeness Centrality	53.4%	24.5%	15.8%	6.3%
Unbiased	51.7%	24.1%	12.8%	11.4%
Corner-Weighted	44.7%	17.1%	21.5%	16.6%
Corner-Weighted Length	35.4%	36.3%	22.9%	5.4%

After looking at the results for the first five models, we questioned the base assumption. *Is* the bikeshare data representative of all bikeshare traffic? The results showed only one or two counter locations would have similar splits, suggesting that the bikeshare traffic was different. To understand whether that was the case, we plotted the total bikeshare counts by hour for each counter location and compared the shape of those curves with the same data from the counters. This was for the Shortest Length model since that was statistically one of the better models.

As shown in Figure 14 below, the shape of the curves is most similar for City Park and somewhat similar for Waterfront. Both of those locations have more recreational traffic, which the bikeshare data seems to match. The shape of the curves for both Cawston and Ethel are quite different. As can be seen in the chart from the counter data, both those locations have peaks in the morning and the evening. This is a commuter traffic pattern. The bikeshare chart does not have those same peaks. Another reason these have a different shape is that Cawston and Ethel counters are in a residential area rather than the commercial/tourist area of Waterfront and City Park.

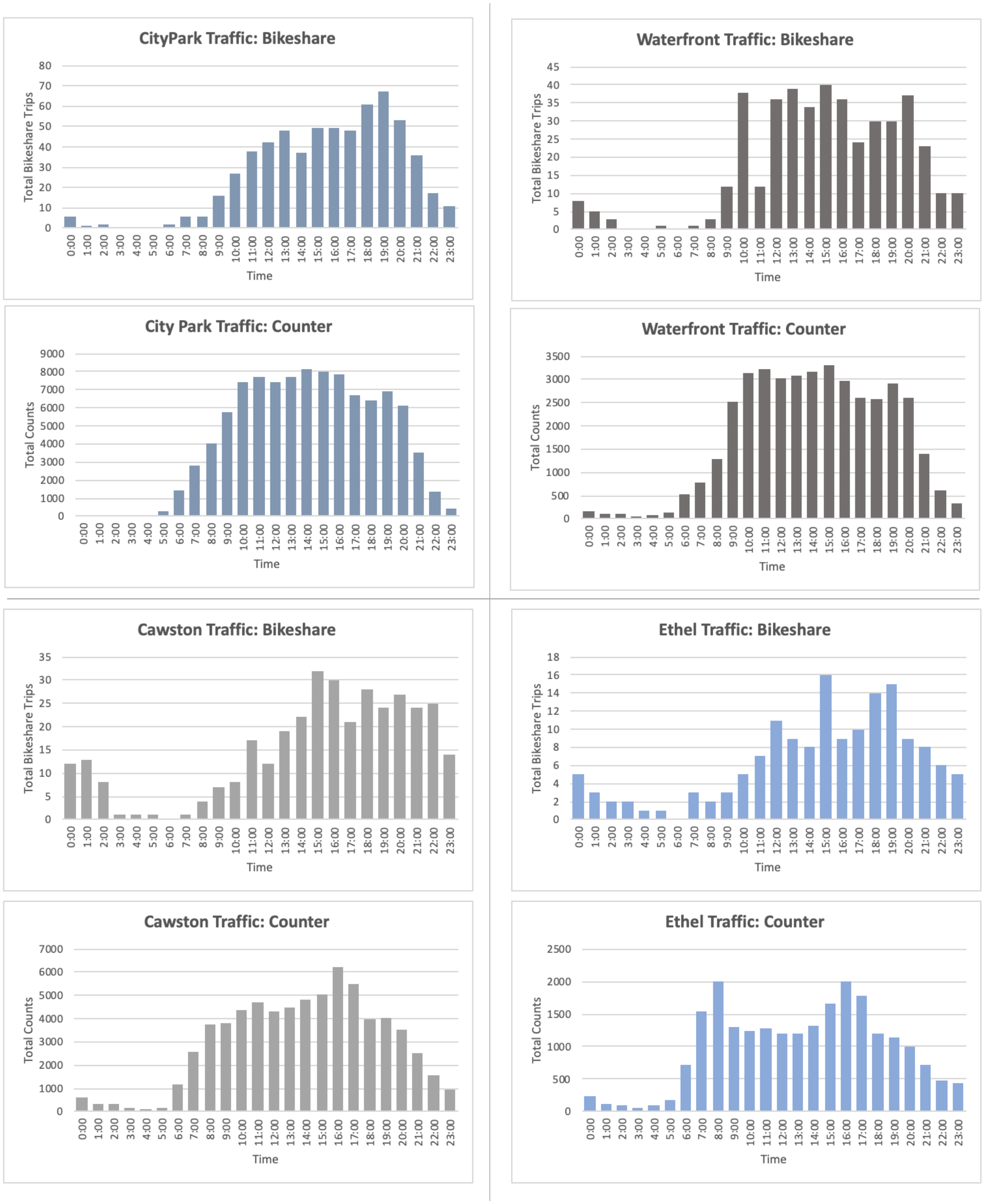

Figure 14: Comparison of bikeshare (top) vs. counter (bottom) traffic across all counter locations

We then wondered if we could determine weights for the Cawston, Ethel and lakeside cycleways that would result in a similar split as the counter locations and result in a good model of the relationship between the counter and bikeshare data.

We developed the Corner-Weighted model as previously described, with the weightings shown in Table 4 below.

Table 4: Weights for Corner-Weighted Model

Cycleway Segment	Weight
Cawston (Road + Cycleway, Gordon to Water St.)	0.005
Ethel (Clement to Sutherland)	0.95
City Park (South of Yacht Club to Bridge)	2.25
Waterfront (North of Yacht Club to Sunset Dr.)	2.75
All other roads, lanes and paths	1.0

While we were able to get a similar split of traffic as the counter data, the weights give an indication of how dissimilar the bikeshare traffic is. We had to make the lakeside cycleway 2-3x less favorable than any other road and Cawston 200x *more* favorable. As we'll discuss in the next section, this model still did not result in the strongest model of the relationship between bikeshare and counter data.

Since multiplying the Simplified Path Preference with Length for the Weighted Length model resulted in the best split between the counter locations, we created the Weighted Corner Length model to see if could improve on the results. Multiplying the length of the edges with the weights of the Weighted Corner model produced the one of the worst splits compared to the counter data, as shown in Table 3.

From this test, Weighted Corner, Shortest Length and Weighted Length performed the best.

Linear Regression of City Park Bikeshare Counts

Since the bikeshare traffic pattern at City Park is similar to the traffic pattern recorded by the counters, it was a good location to compare the relationship between counter data and bikeshare counts resulting from each of the path-finding models. We did a linear regression of weekly counter counts vs weekly bikeshare counts, an example of which is shown in Figure 15 below.

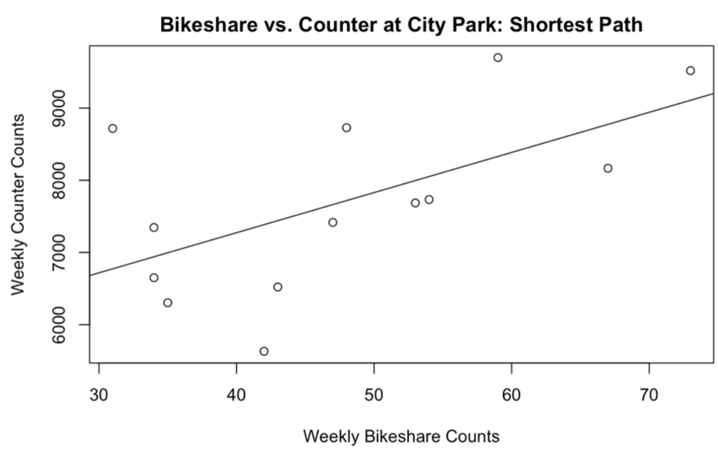

Figure 15: Linear regression of counter vs. bikeshare counts - Shortest Length model

We used R-squared value and p-value to compare the regression models. The smaller the R-squared value, the less we are explaining the variation in the data. Therefore, the closer the R-squared is to one, the better. The p-value is an indication of whether the multiplier for the bikeshare data is statistically significant. A p-value of less than 0.05 means that the multiplier is statistically significant.

Table 5 below compares the R-squared value and p-value for each of the path-finding models.

Table 5: Comparison of linear regression of weekly counter vs. bikeshare data at City Park

Path-Finding Model	R-Squared	P-Value
Shortest Length	0.3463	0.0343
Path Preference	0.1998	0.1260
Simplified Path Preference	0.2125	0.1130
Weighted Length	0.2341	0.0939
Closeness Centrality	0.3128	0.0469
Unbiased	0.2316	0.0959
Weighted Corner	0.2168	0.1089
Weighted Corner Length	0.2376	0.0911

For City Park and Weighted Corner models, we also did a linear regression for all the counter locations. The Weighted Corner model performed better than Shortest Length at Waterfront, but not for the other three locations. So, even intentionally weighting the model to try to match the counter data did not result in creating a similar traffic pattern at each counter location.

From this test, Shortest Length, Closeness Centrality and Weighted Corner Length performed the best.

Determining Best Path-Finding Model

Table 6 below summarizes the results from the evaluations of each path-finding model. The numbers indicate the order of the top 3 models for the particular evaluation.

Table 6: Comparison of path-finding models

Path-Finding Model	Visual	Speed	Split	Linear Regression
Shortest Length	**3**	**1**	**2**	**1**
Path Preference	1			
Simplified Path Preference				
Weighted Length			3	
Closeness Centrality		2		2
Unbiased				
Weighted Corner	2	3	1	
Weighted Corner Length				3

Although Shortest Length had a couple of issues in the visual evaluation, it was in the top two for all other evaluations. Because it is based on the length of the streets it was one of the more objective path-finding models. It was not weighted based on path preference or trying to fit the counter data. Therefore, this was the model chosen to calculate ADB volumes.

Calculating Average Daily Bicycle (ADB) Volume

Calculating the ADB at each counter can be done by finding the relationship between counter and bikeshare data at each counter. However, to find the ADB for all downtown street segments, we needed to estimate the effect each counter has on the ADB. As we can see in Figure 16, the relationship between counter and bikeshare data is different for each counter.

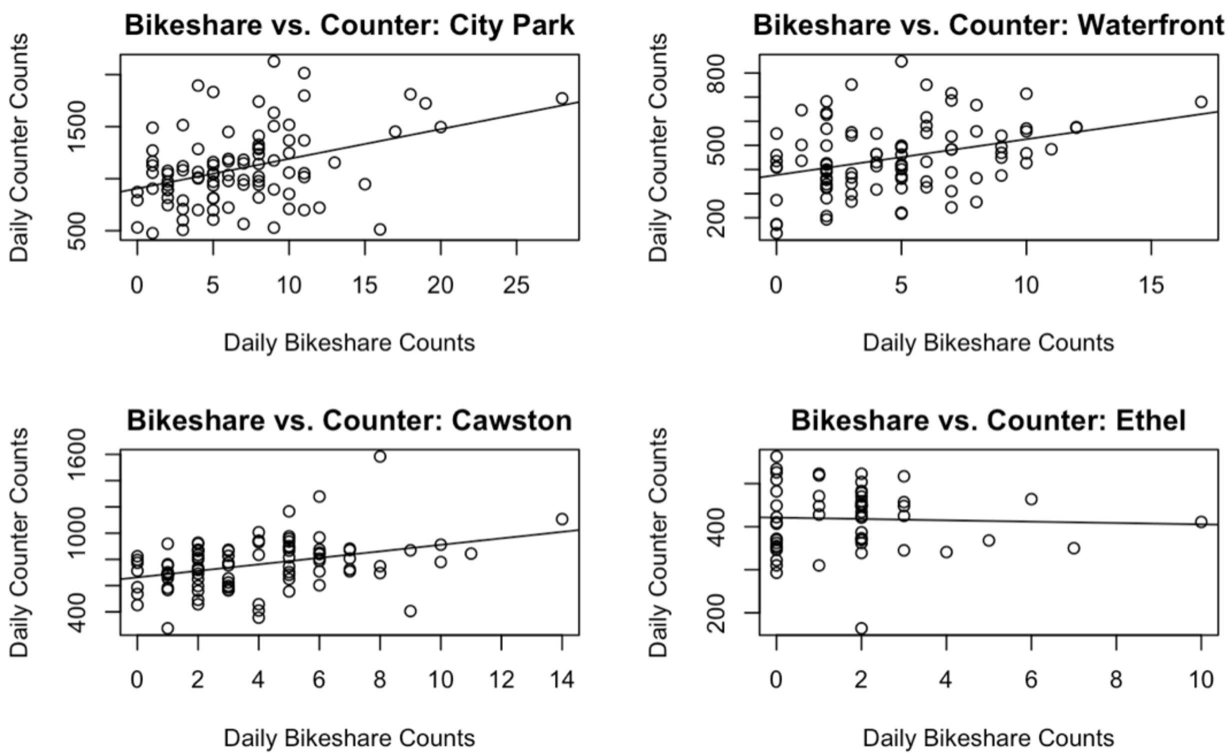

Figure 16: Relationship between daily bikeshare vs counter data at each counter location

A mixed effects model is a generalized regression model that takes into account the effect of each counter location as it calculates the overall relationship between the counter and bikeshare data. Since we have count data with many zeros and repeated values, the relationship was calculated using a negative binomial distribution.

From this model, the general relationship between counter and bikeshare data is shown in Figure 17 below:

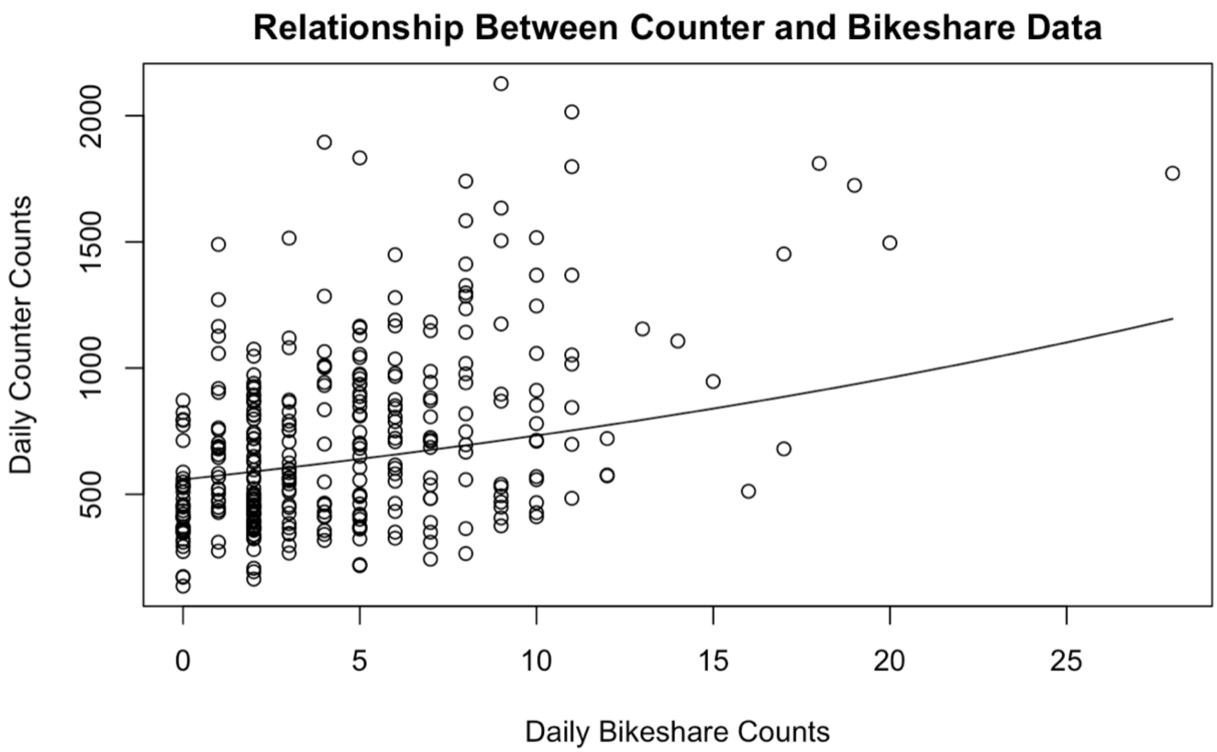

Figure 17:Mixed effects model of the relationship between counter and bikeshare data

The equation to calculate ADB from this model is:

$$\text{ADB}= e^{(0.02717094 * \text{average daily bikeshare} + 6.325313)}$$

where

$$average\ daily\ bikeshare\ = \frac{total\ bikeshare}{91\ days}$$

Using this equation, the ADB was calculated for each street segment and mapped as shown in Figure 18 below.

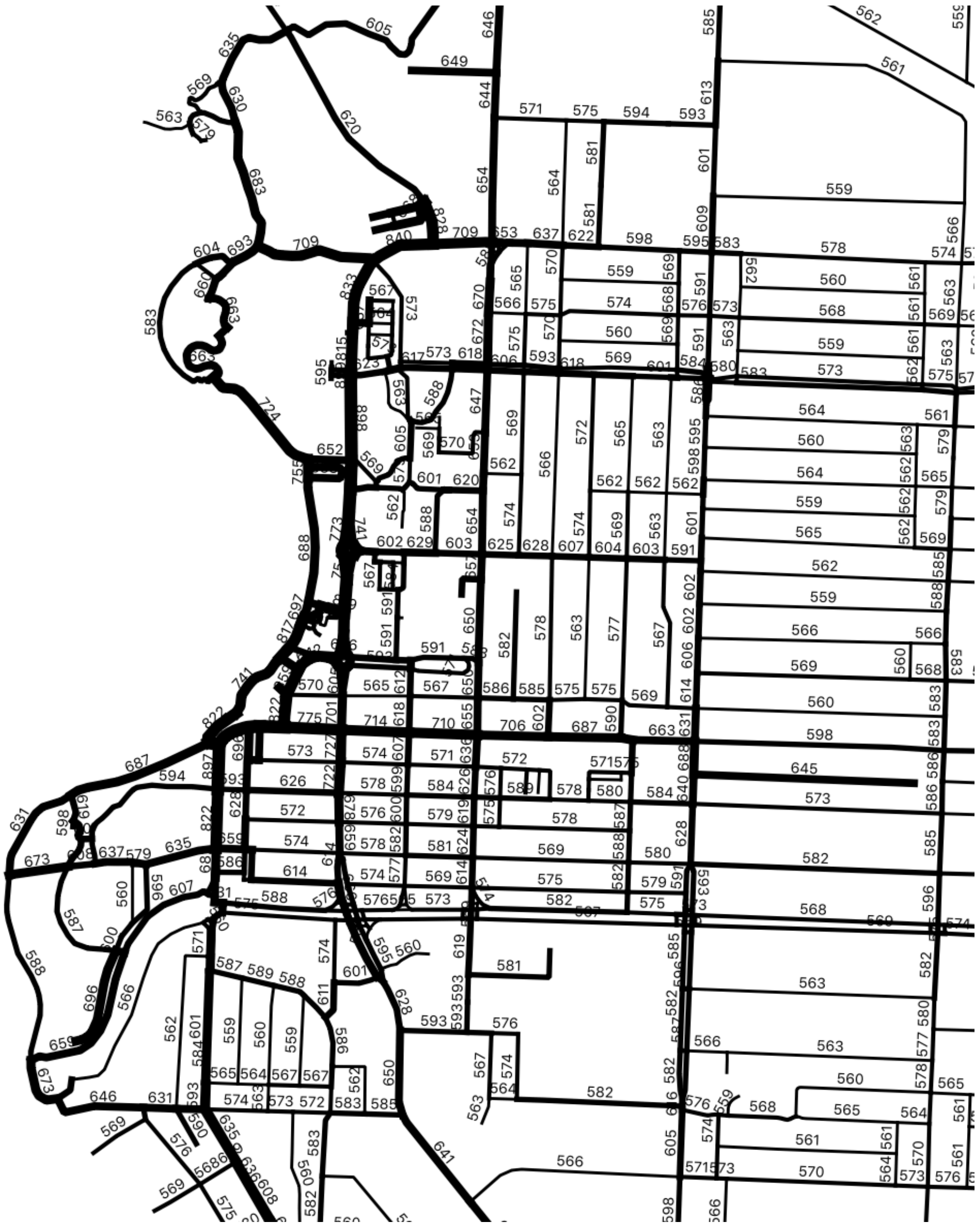

Figure 18: ADB calculated with mixed effects model approach

One of the drawbacks of finding one statistical model from the different counter models is that there is an “averaging out” effect. As a result, we don’t see the same level of variance in the map as we do for the total bikeshare counts.

To retain the variation, we needed to find a single multiplier for the average daily bikeshare data to approximately equal the ADB from the counter locations. One strategy to find this multiplier is a least squares optimization, which minimizes the following equation:

$$f(x) = \sum_{i=1}^{4} (ax_i - y_i)^2 \; w_i$$

where, for each counter location, *i:*

$a = the\ multiplier$
$x_i = bikeshare\ counts$
$y_i = counter\ counts$
$w_i = split\ of\ counts\ at\ each\ counter\ location$

After minimizing the above equation, the equation for ADB became:

$$ADB = 159 \left(\frac{total\ bikeshare\ count}{91} \right)$$

In Python, we calculated the ADB for each street segment and visualized the map as shown in Figure 19 below. This approach provides an estimate of ADB without losing the variation as we did with the mixed effects model.

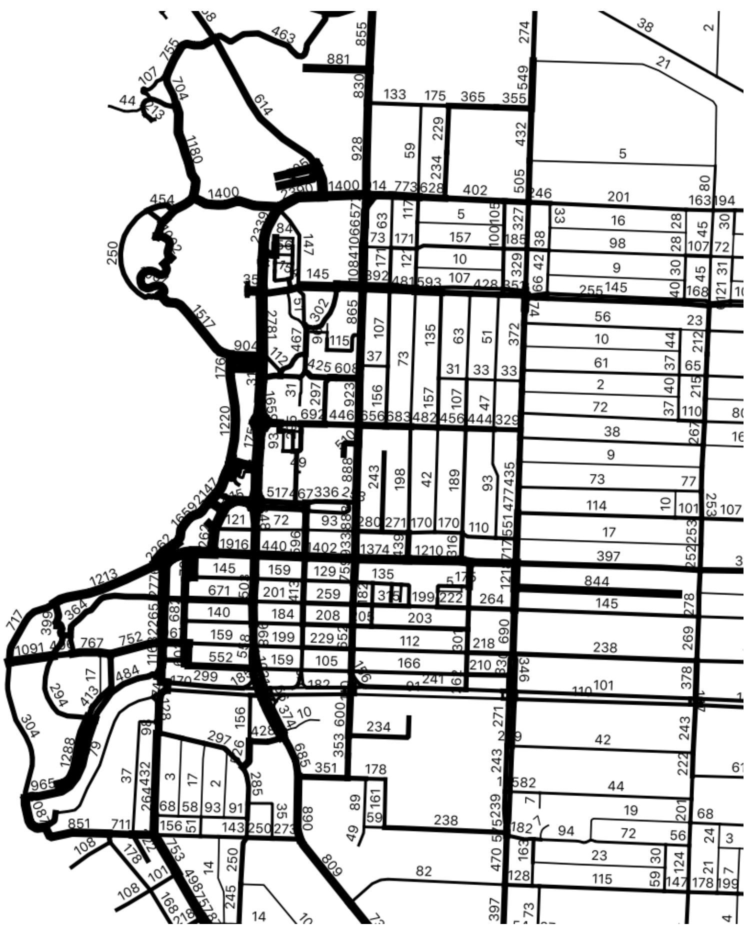

Figure 19: ADB calculated with least squares optimization approach

As we have shown throughout this analysis, the traffic at each counter location is different. In addition, the bikeshare traffic is different from the general traffic. While the mixed effects model and least squares optimization gave us estimates of ADB, they are only estimating how the bikeshare data scaled up to overall volumes. Since the bikeshare traffic patterns are so different from the general traffic, we can only conclude how the bike network was used by bikeshare riders. Therefore, we used the map of the total bikeshare counts shown in Figure 20 for our final conclusions and recommendations.

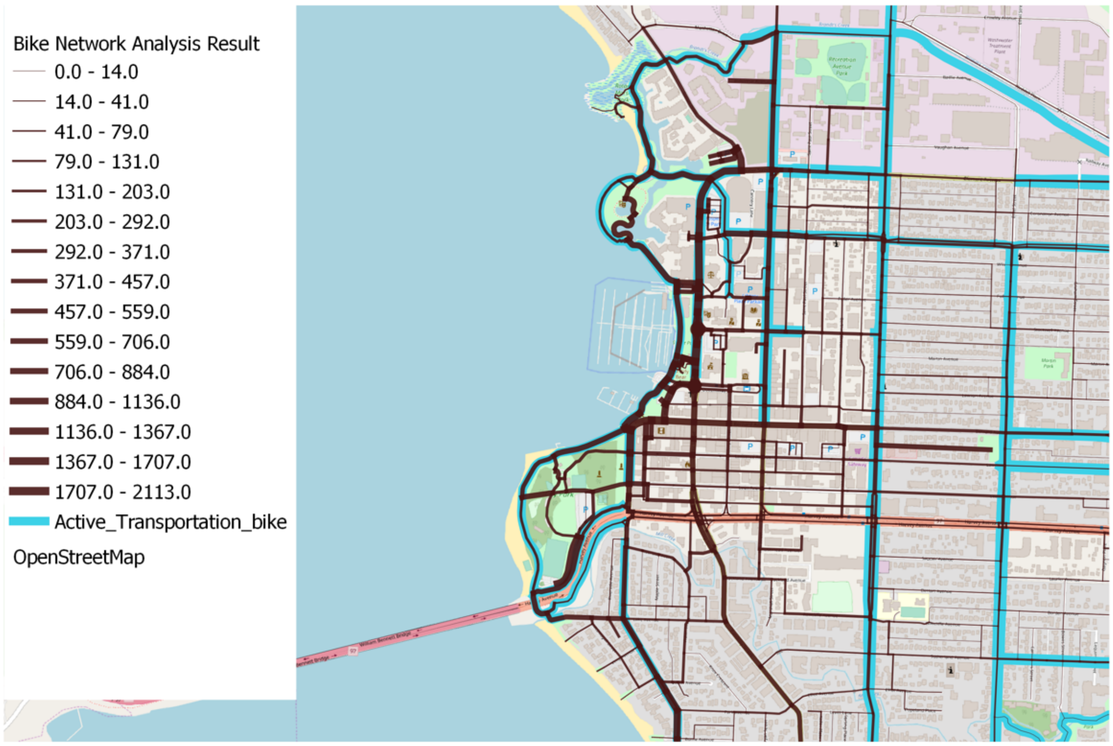

Figure 20: Total bikeshare counts with bicycle infrastructure (highlighted in blue)

Gaps in Infrastructure

The above map show that Bernard Ave. and Pandosy St./Water St. were highly travelled by bikeshare users, even though there are no bike lanes for cyclists. Bikeshare users either cycled with traffic or on the sidewalks. There are also gaps in infrastructure at Clement and Doyle, which were often cycled by bikeshare users.

Crossing Highway 97 (Harvey Avenue)

Table 7 summarizes the number of bikeshare trips that crossed Highway 97 at each intersection and the tunnel at City Park.

Table 7: Total number of bikeshare trips at each crossing of Highway 97

Highway Crossing	Number of Bikeshare Trips
Tunnel at City Park	622
Abbott	426
Pandosy/Water	504 (Southbound) 207 (Northbound)
Ellis	312
Richter	165 (Southbound) 216 (Northbound)
Ethel	164

As previously mentioned, the `shortest_path` algorithm had trouble finding paths at the Highway 97 and Abbott intersection. It may be due to GPS inaccuracy. From experience with that intersection, it may also be due to northbound traffic continuing along the sidewalk from Riverside Ave to the Highway and using the crosswalk at the highway to cross Abbott St.

Use of Laneways

One laneway stands out as being exceptionally popular: a residential lane circled in green in Figure 21 below.

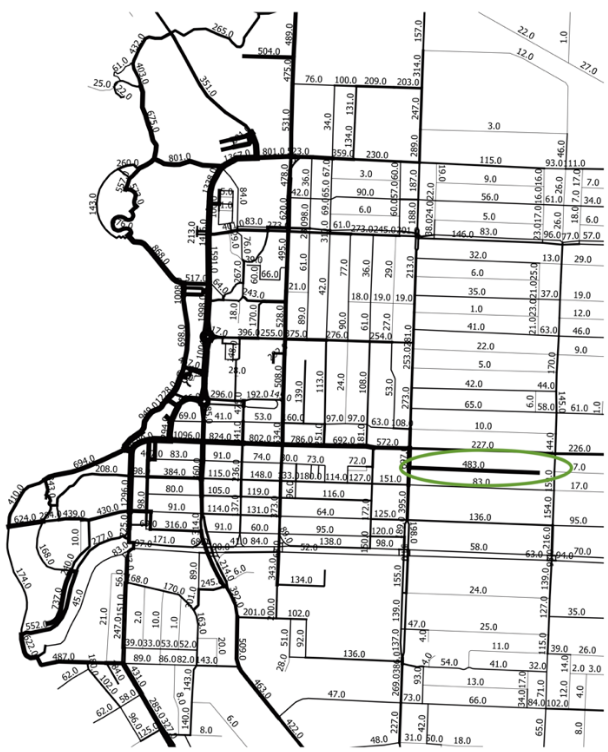

Figure 21: Laneway with unusually high bikeshare trips (circled)

The laneway is parallel to and directly south of Bernard Avenue, a major road to/from downtown. It indirectly connects Richter to Ethel through a park at the east end of the laneway. Richter has a bike lane and Ethel has a separated, paved path. Since it is a residential neighbourhood, it may mostly be local traffic, though the route through the park is clearly shown on Google Maps. There was a bikeshare haven at the southeast corner of the park, where bikeshare users could initiate or end a trip.

Outside of this one, the laneways in the residential blocks are used less than the residential streets. Many of the other residential laneways don't directly connect two streets or cyclists may be treating those laneways as private property. In the commercial blocks, the laneways are used as much as the tertiary downtown streets, since they connect to streets on both sides. Commercial laneways may also be viewed as public property.

Conclusions

From the analysis of Kelowna's bicycle network, using a combination of bikeshare and counter data, the key conclusions are as follows:

- Using OSMnx to apply graph theory gave us the tools to map inaccurate, infrequent GPS points to intersections, find paths to connect those GPS points for each trip and analyze the structure of the bike network.
- Based on our four evaluation criteria, the best path-finding model was shortest path, which weighted street segments based on their length.
- Bikeshare traffic patterns are more recreational and similar to bicycle traffic recorded by the counters at City Park and Waterfront. The bikeshare traffic pattern is different from the traffic recorded by the counters at Cawston and Ethel.
- Mixed effects model and least squares optimization gave us estimates of ADB. However, because the bikeshare and general traffic patterns were so different, we can only draw conclusions how bikeshare users travelled through the bike network. For that we used the visualization of the total count of bikeshare trips for each street segment.
- Bernard Ave. and Pandosy St./Water St. were highly travelled by bikeshare users, even though they don't have infrastructure for cyclists. There are also infrastructure gaps at Clement and Doyle, which were often used by bikeshare riders.
- After the tunnel, Highway 97/Abbott is the most popular highway crossing.
- The residential laneway shown in Figure 20 is really popular, likely with local traffic.

Recommendations

We recommend the following for further analysis of Kelowna's bike network in the future:

- Continue collecting more data. Another bikeshare pilot or an ongoing bikeshare service will be a good source of GPS data. Consider partnering with other BC cities to develop and deploy a free GPS app for active transportation, similar to what Montreal did. Deploying additional counters will collect more data on how all cyclists use the network.
- Consider partnering with UBCO to design statistical experiments to better understand the types of bicycle traffic in key areas of the network.
- Validate how cyclists are crossing the intersection at Highway 97 and Abbott Street.

Works Cited

Boeing, G. (2017). OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. *Computers, Environment and Urban Systems*, 65, 126-139.

Boeing, G. (2018, Jan 19). *Urban Street Network Centrality*. Retrieved from Geoff Boeing: https://geoffboeing.com/2018/01/urban-street-network-centrality/

City of Kelowna. (2011, June 2). *Kelowna 2030 - Official Community Plan.* Retrieved from City of Kelowna: https://www.kelowna.ca/city-hall/city-government/bylaws-policies/kelowna-2030-official-community-plan

City of Kelowna. (2016, April). *Kelowna On the Move: Pedestrian and Bicycle Master Plan.* Retrieved from City of Kelowna: https://www.kelowna.ca/sites/files/1/docs/related/pbmp_final_draft.pdf

City of Kelowna. (2018, February). *Bikeshare Pilot*. Retrieved from City of Kelowna: https://www.kelowna.ca/roads-transportation/active-transportation/cycling/bikeshare-pilot

City of Kelowna. (2019, April). *Benefits of Active Transportation.* Retrieved from City of Kelowan: https://www.kelowna.ca/roads-transportation/active-transportation/benefits-active-transportation

City of Kelowna. (2019, April). *City of Kelowna Counters*. Retrieved from Eco-public: http://www.eco-public.com/ParcPublic/?id=4198

GPS.gov. (2017, Dec 5). *GPS Accuracy.* Retrieved from GPS.gov: https://www.gps.gov/systems/gps/performance/accuracy/

Najera, J. (2018, Nov 26). *Graph Theory - History & Overview.* Retrieved from Towards Data Science: https://towardsdatascience.com/graph-theory-history-overview-f89a3efc0478

NetworkX Developers. (2019). *NetworkX: Software for Complex Networks*. Retrieved from NetworkX: https://networkx.github.io

OpenStreetMap. (2019). *About OpenStreetMap*. Retrieved from OpenStreetMap: https://www.openstreetmap.org/about

Proulx, F. R., & Pozdnukhov, A. (2017). Bicycle Traffic Volume Estimation using Geographical Weighted Data Fusion. *Journal of Transport Geography*.

Strauss, J. (2015). *New Methods for Modeling and Integrating Bicycle Activity and Injury Risk in an Urban Road Network* . Montreal: McGill University.

Strava Metro. (2019). *A Story of Synergy: Bike Counts and Strava Metro.* Retrieved from Strava Metro: http://cdn2.hubspot.net/hubfs/1979139/Bike%20Counter%20Correlation.pdf

The R Foundation. (2019). *What Is R?* Retrieved from The R Project For Statistical Computing: https://www.r-project.org/about.html

Winters, M., & Teschke, K. (2010). Route Preferences Among Adults in the Near Market for Bicycling: Findings of the Cycling in Cities Study. *The Science of Health Promotion*, 40-47.

Worona, M. (2019, March 4). *Report to Council: Bikeshare Pilot Update and Next Steps*. Retrieved from City of Kelowna: https://kelownapublishing.escribemeetings.com/filestream.ashx?DocumentId=21470