## Descriptive Title: Determining bikeability of census tract regions below 400% poverty line

## Introduction (2pts)

The premise of the project is based around Assembly Bill 117, which will allocate $10 million dollars in E-bike vouchers to households at 400% of the federal poverty level. That’s $51,000 for a single person and $106,000 for a family of four. Purchasable bikes include Class 1 and 2 e-bikes.

The project aims to evaluate the bikeability of census tracts that fall under this income level bracket and determine if the area around them are bike friendly as this would be a factor in determining how effective this bill would be.

We currently have census tract data from the US Census Bureau and bike street data from OSMNX. Census data is used to determine which census tracts the project will be focusing on. OSMNX will used to determine various factors about the bike network as explained in Data and Methods

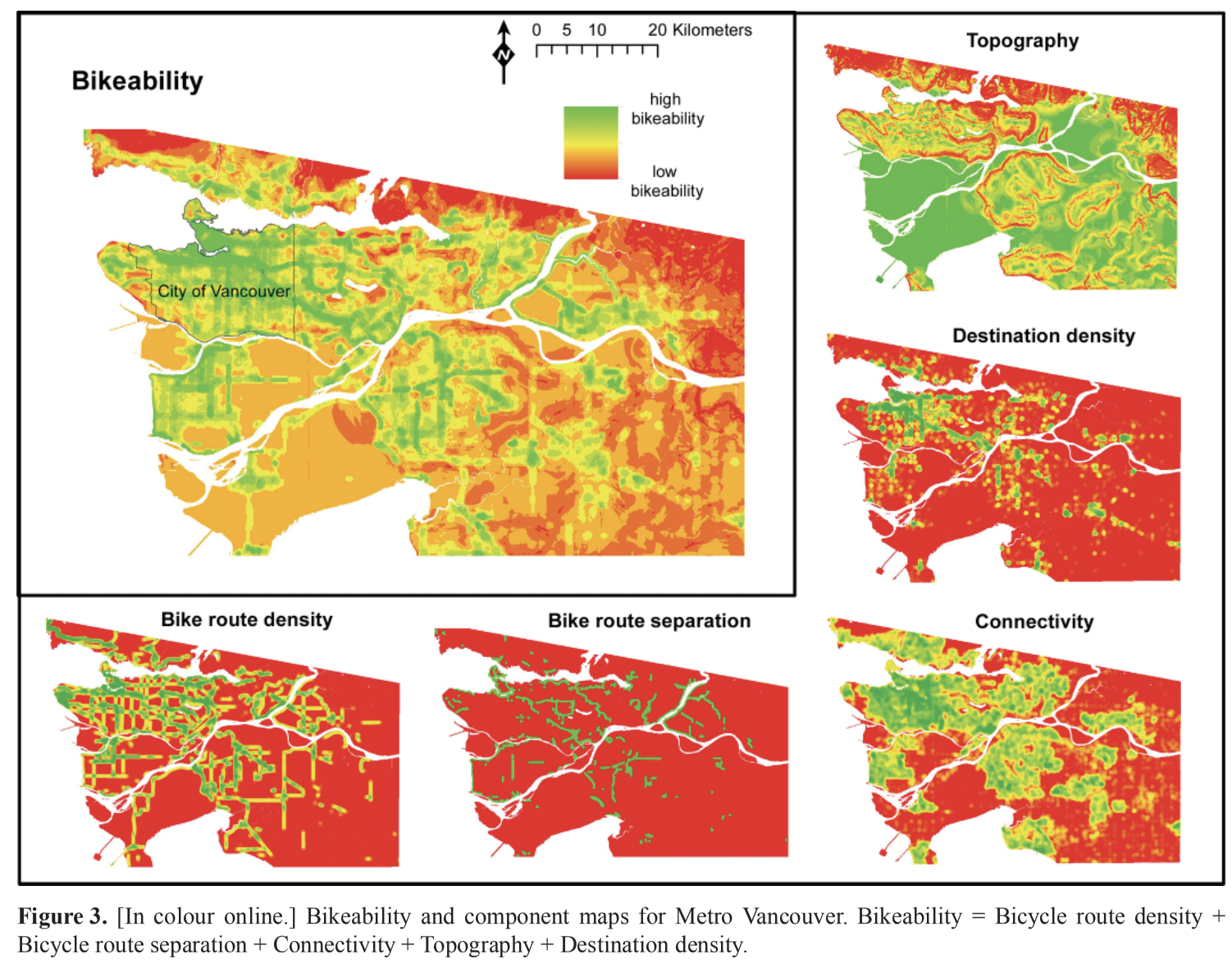
The Missing Link, Bicycle Infrastructure Networks and Ridership in 74 US Cities

This paper aims to develop a standard methodology for measuring bicycle facility network quality of 74 US cities.

Some things to note about the paper. Firstly, only cities that had a population greater than 100,000 were chosen, second, trips that would involve traveling within the city were selected.

We will use two indices that were used as metrics for evaluating bike network in the city, including

* β = edges / vertices, low β suggests an increased chance that any given route requires leaving dedicated bike infrastructure to ride with mixed traffic
* γ = e/ 3(v-2), edges vs theoretical edges. Higher gamma values indicate greater internal connectivity and increased redundancy.

Mapping bikeability: a spatial tool to support sustainable travel

In this paper, the authors survey Vancouver residents to see what factors are important in determining bikeability. They then collect the objective data behind these factors, and as seen in Figure 3, they weight each of these factors in accordance to its importance as determined by the survey, and assign a bikeability score to the whole city. We plan on using many of the same factors as Winters et al.; all of the factors shown, with the exception of bike route separation can be found via OSMNX. Additionally, the examples provided by Geoff Boeing give the code necessary to extract these statistics.

## Data and Methods (3pts)

### Census Data

Census data are collected using the cenpy package, which allows us to query the census api to find both ACS and decennial census data. We query the 2019 ACS at the census census tract level, fetching the variables corresponding to total population, population in poverty, and number of people without access to a vehicle. Additionally, cenpy gives geometric data, giving us the polygons that define the area covered by each tract. The resulting data structure from the query is a geopandas GeoDataFrame, and some statistics can be called on this geometric data, such as the area, which we use to then calculate the population density of each tract. We can also calculate the poverty rate by dividing the number of people in poverty by the tract population. We then select the top 25th percentile of the tracts by poverty rate, and filter out tracts with extremely low population density. This leaves us with poor urban tracts, which would receive the bulk of the money given by AB 117.

### Bike Street Network

Bike street network data is used for the project. The bike network data comes from OpenStreetMaps and the Python package OSMNX. The nodes are intersections while the edges are the bike paths. The graphs generated are undirected and weighted based on distance. Depending on the network location the graph may or may not have a small-world property.

### Basic Network Statistics

The project aims to use certain network statistics as a means of grading the bikeability of the various census tracts. These statistics are broken down into the following factors

* Size
  + Number of vertices
  + Number of edges
  + Number of subgraphs
  + Street length average
* Connectivity
  + Average degree
  + Beta Index
  + Gamma Index
  + Circuity average
  + Average betweenness
  + Highest betweenness
  + Standard deviation of betweenness
* Density
  + Population density
* Directedness
  + Average length of edges
  + Total length of edges

Using the data and statistics outlined above, we will adopt an approach similar to that described by Winters et al. We will also incorporate some further network analysis statistics developed from OSMNX. 

The first is the betweenness centrality of each node in the tract. We argue that tracts with a few nodes of high betweenness (like the one shown to the right) are less bikeable, as this implies a less permeable tract, giving longer travel times, which are especially impactful to cyclists, who tend to travel at lower speeds and over shorter distances than drivers.

## 

## Results (8 pts)

To begin with, we look at various cities, capturing a point of interest with radius 2km. Cities with higher density as that has been shown to likely lead to a better bike network.

Table 1: Basic Statistics of densely populated census tract region

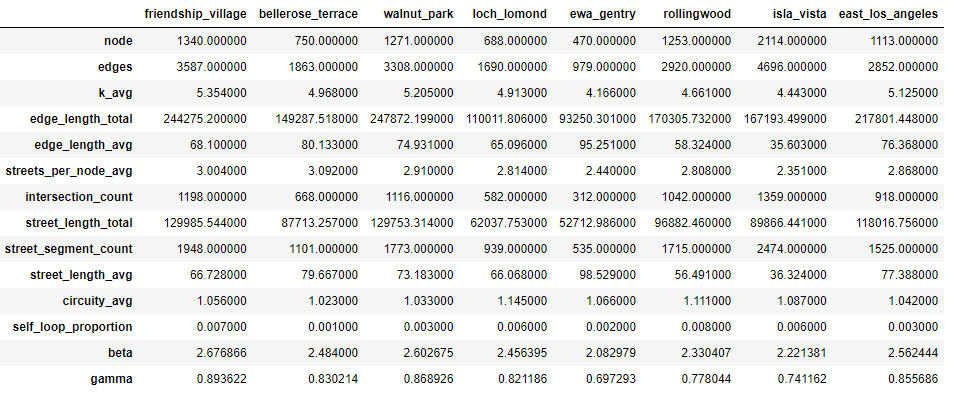
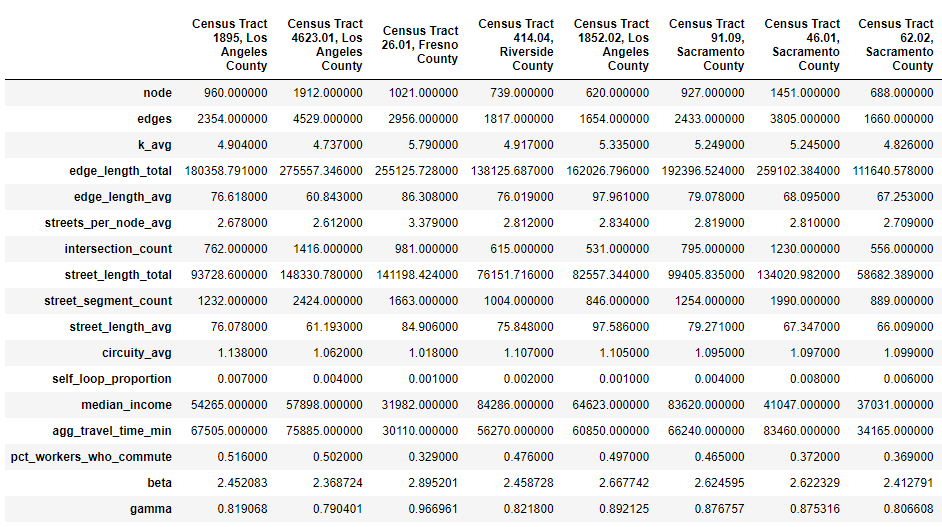


Table 1 shows basic statistics of various densely populated census tract regions. These statistics will serve as a benchmark when evaluating other census tract regions.

Filtering out the census tracts using the methodology in Census Tracts, we get around 2000-3000 unique census tracts. As a starting point, we decided to take a sample of 15 census tracts, graph out their network of 2km and collate the following statistics.

Table 2: Basic Statistics of densely populated low income census tracts (8 tracts shown)



As can be seen from Table 1 and Table 2. The values for most of the statistics are relatively similar, making it difficult to determine if the network is bike friendly or not.

The next section dives deeper into other methods through the use of networkx as well as using Places of Interest data to classify land use of the region. In addition, we will also begin to trim the networks in terms of only selecting roads that have bike lanes on them.

## Future Work and Conclusions (2 pts)