City Bike Network Analysis

Federico Ceccato

Università Cà Foscari Venezia

Agenda

1. Introduction
2. EDA
3. Analysis
4. Coclusions

Introduction

In this project I try to analyse the network of the bike-sharing service in Helsinki. Within the context of transportation, the term network refers to the framework of routes within a system of locations, identified as nodes. A route is a single link between two nodes that are part of a larger network that can refer to tangible routes such as roads and rails, or less tangible routes such as air and sea corridors.

Purpose

The purpose of this project is to analyse the network to find some improvements for the business model.
In fact, citizens purchase access for a day, week or the entire cycling season that lasts from April to November. All passes include an unlimited number of 30-minute bike rides. For an extra fee of 1€/hour, you can use the bike for longer. Bikes are picked up and returned to stations that are located all around Helsinki and Espoo.

Dataset

The dataset informations are:

- departure and arrival position
- avg. speed
- duration
- longitude and latitude of departures and arrivals
- distance
- station name
- temperature

Results

Interesting facts:

- The mean of the duration is 13 minutes and the median is 10
- The mean of the distance is 2,2km and the median is 1,8km
- The usage is concentrated during the work week, in the early morning and between 16.00 and 18.00

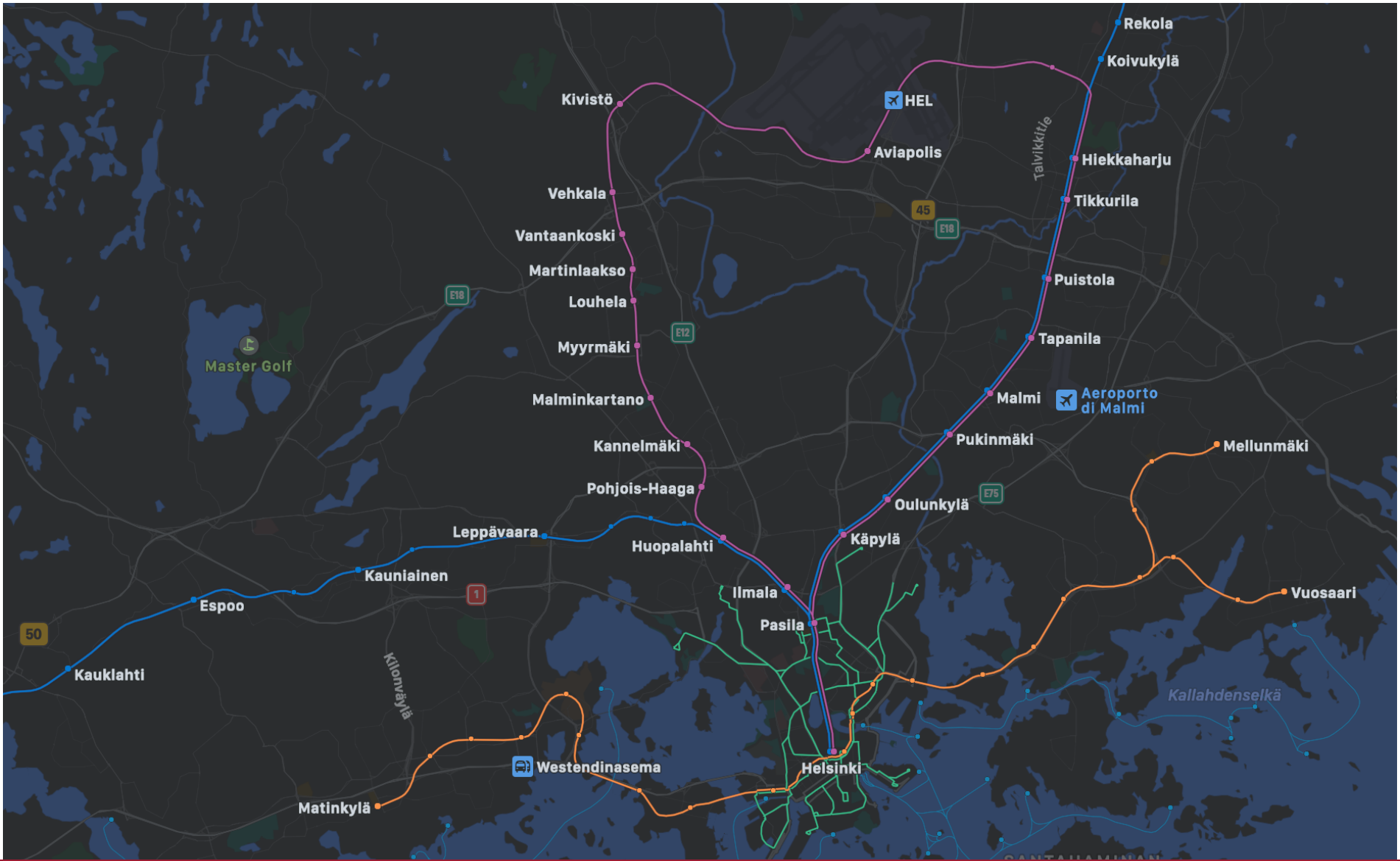
Analysis

The next pictures represents our network structure, based on the numbers of trips that arrive in each station. To understand the network I applied different centrality measures.
The basic stats of the network are:

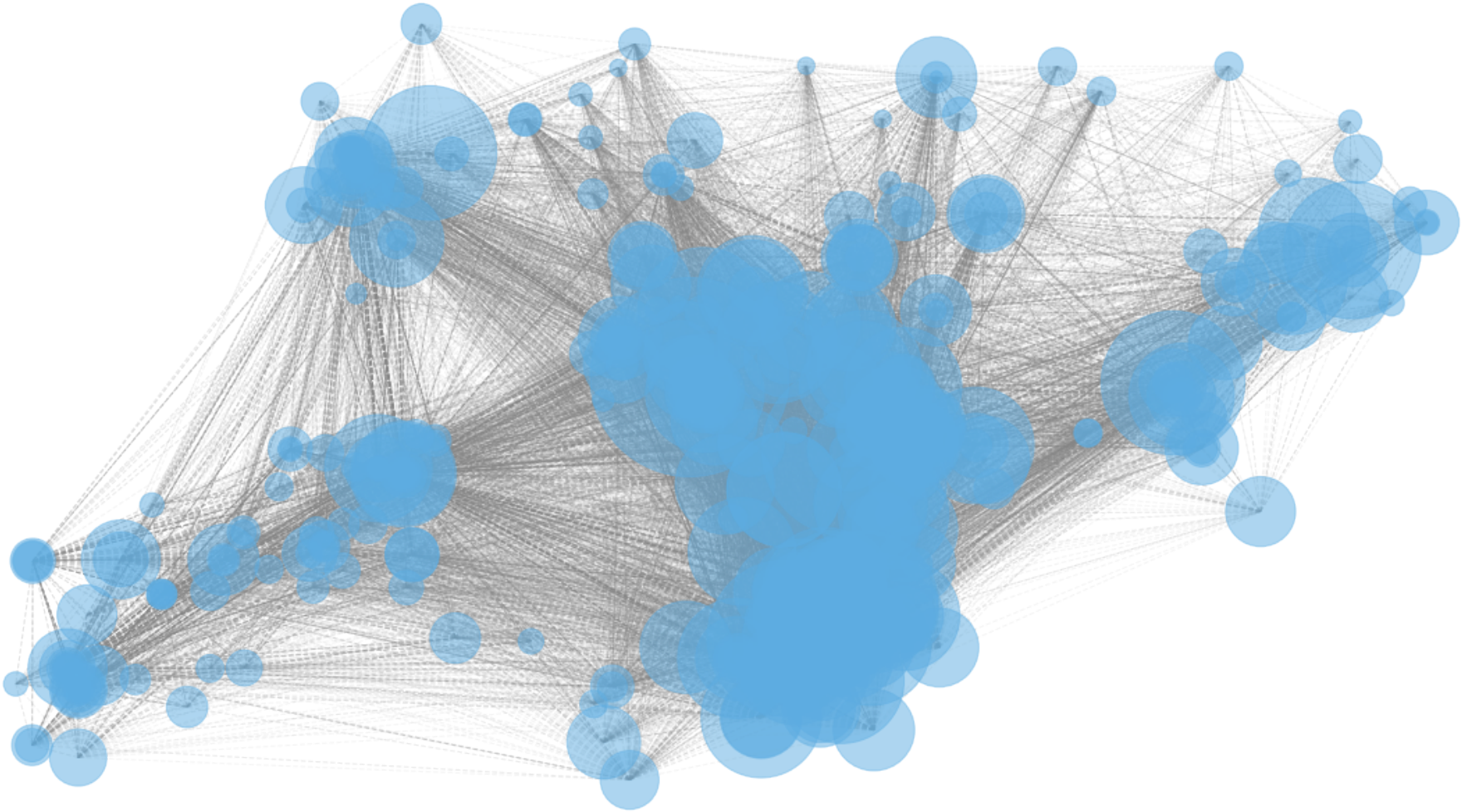
- Number of nodes: 347
- Number of edges: 29604
- Average degree: 170.6282
- Network density: 0.493

Geographic Area

Below the geographic area of Helsinki with the most popular transportation routes.



Network Graph

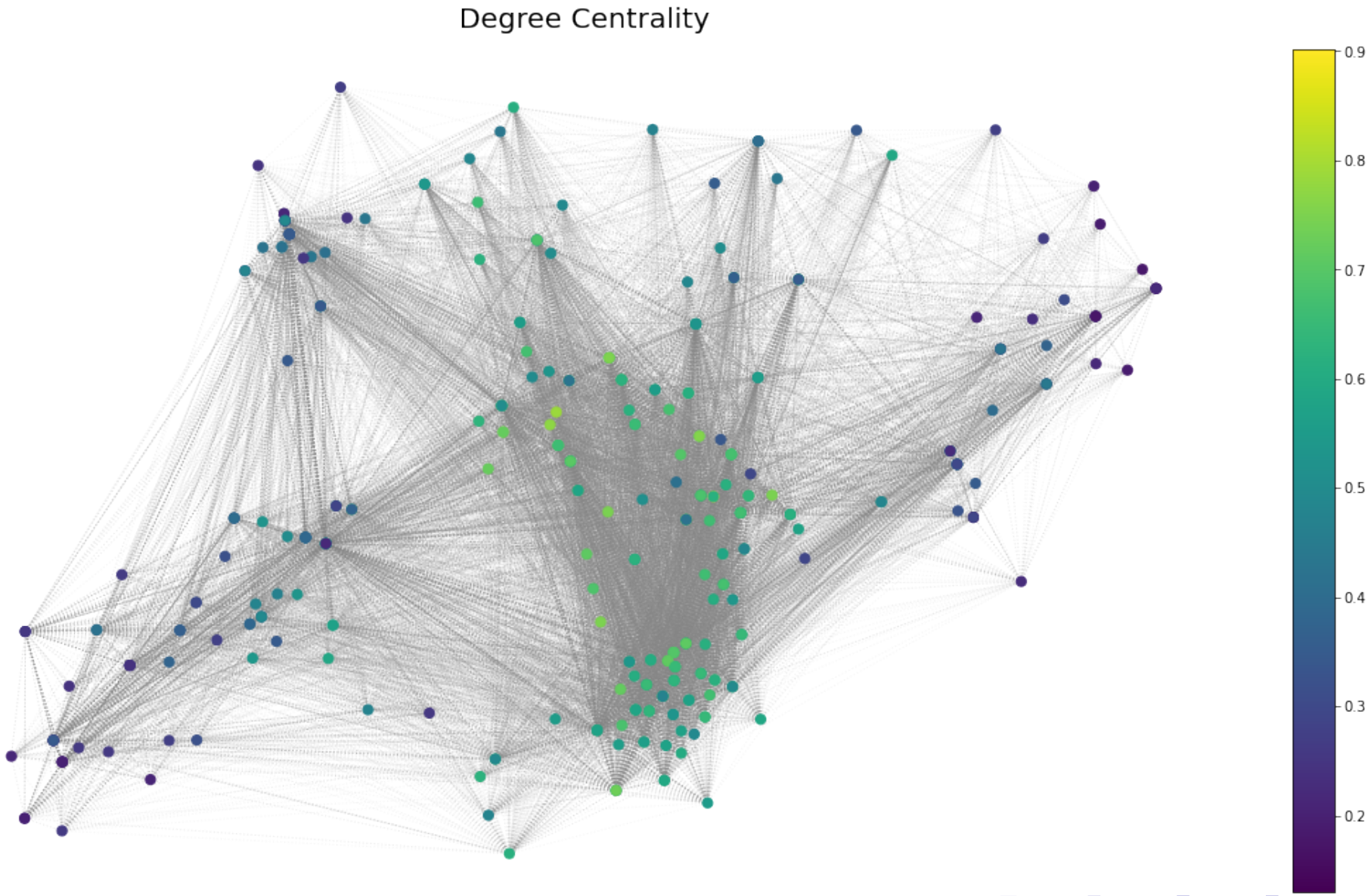


Degree Centrality

The degree centrality is the simplest measure, the degree corresponds to the number of links that a node has. In this case we can see that in the city centre the nodes have a higher degree than the others outside.
Top 5 nodes by degree:

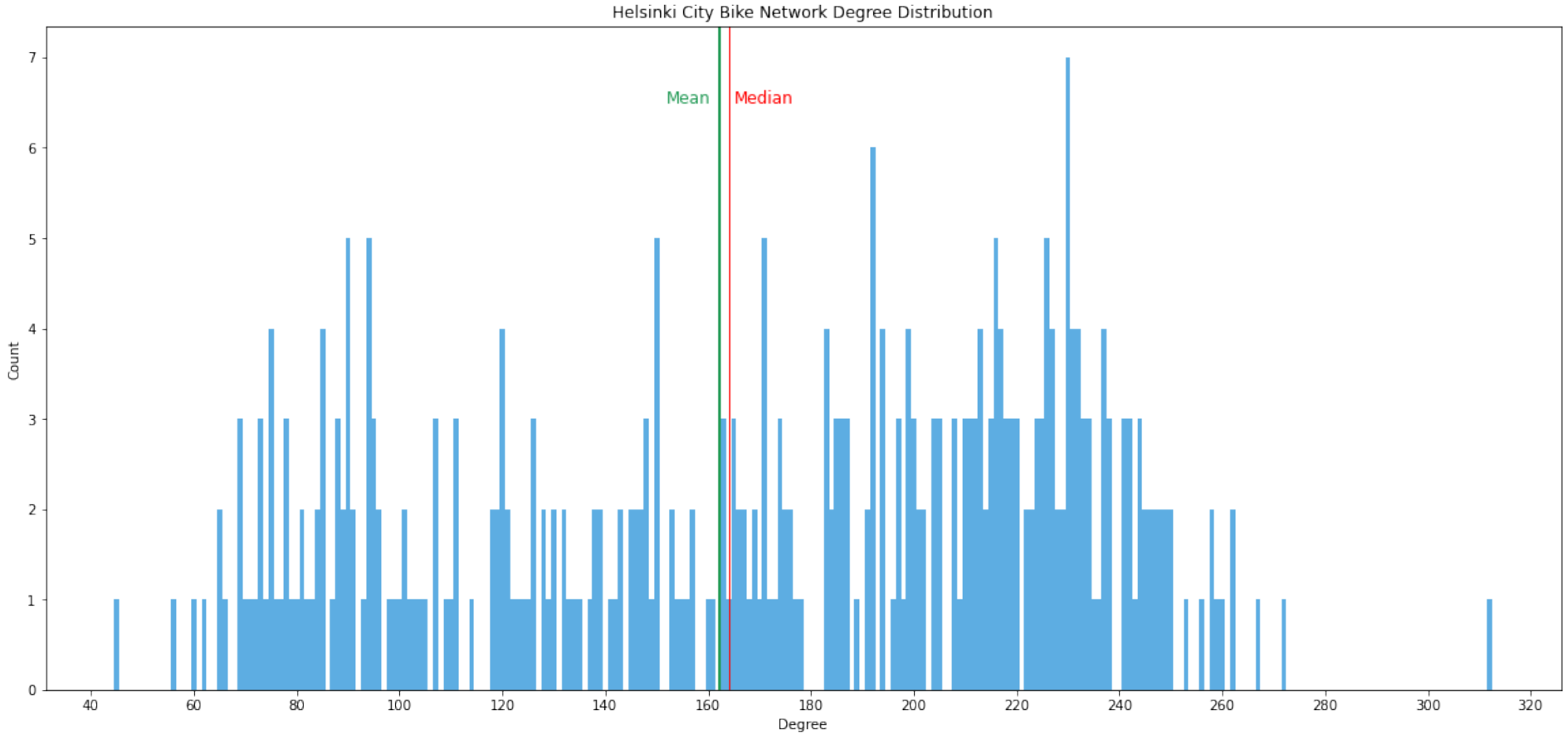
- 'Haukilahdenkatu', 312
- 'Paciuksenkaari', 272
- 'Huopalahdentie', 267
- 'Laajalahden aukio', 262
- 'Munkkiniemen aukio', 262

Degree Centrality



Degree Centrality Distribution

In this picture we see the degree distribution. On average each station is connected to other 162. We can explain the most connected station to the proximity of other crucial points like bus station, train stations, schools and so on.



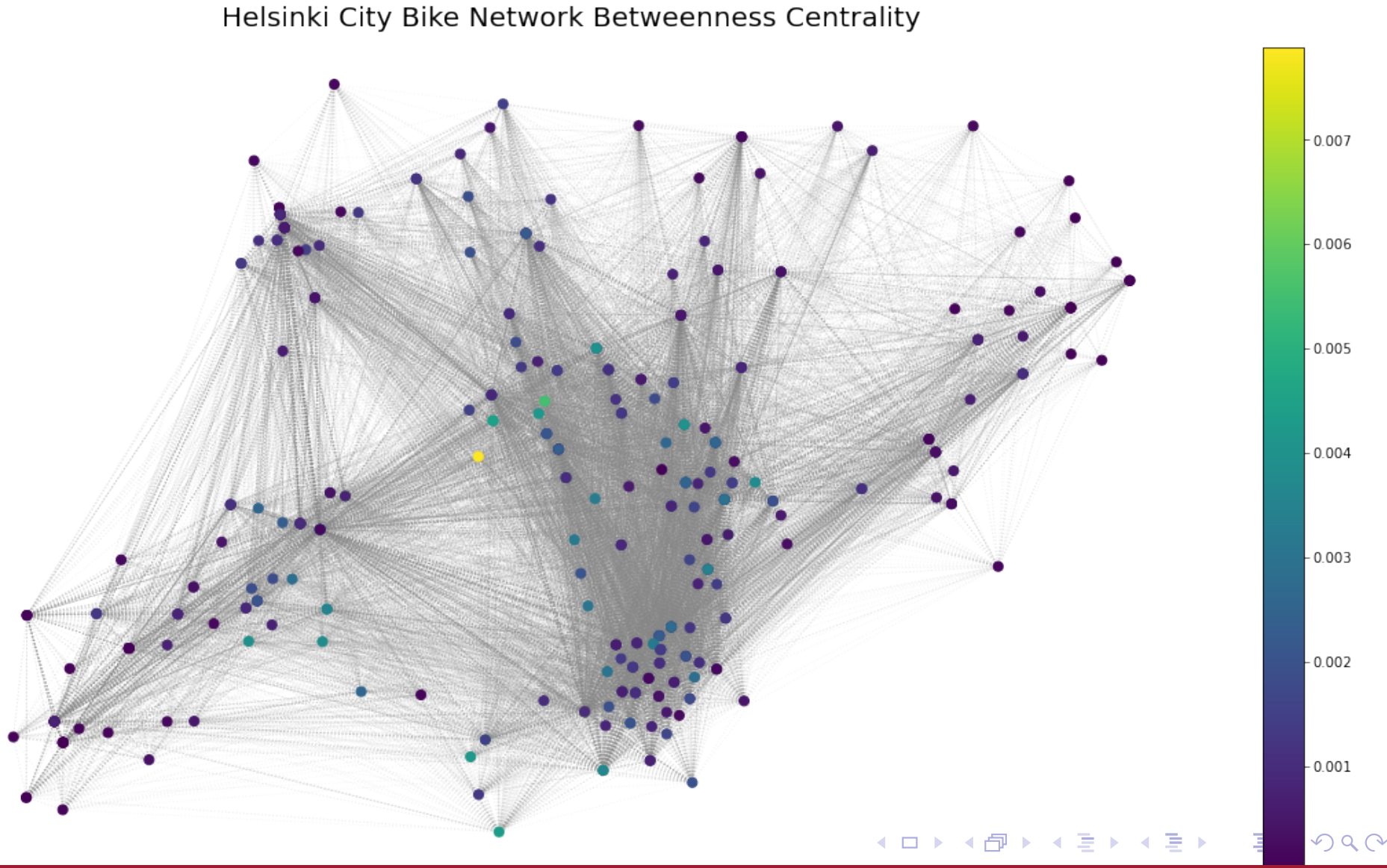
Betweenness Centrality

The betweenness centrality calculates the shortest paths between all the nodes of the network and assigns each node a measure based on the number of shortest paths going through the target node.
We'll see that some nodes are very crucial due to geographical constraints. In fact Helsinki is an archipelago so some bridges are the only passage to the city centre and their betweennes centrality is very high.
Top 5 nodes by betweenness:

- 'Haukilahdenkatu', 0.0493
- 'Lehtisaarentie', 0.0071
- 'Lauttasaaren ostoskeskus', 0.0061
- 'Aalto-yliopisto (M), Korkeakouluaukio', 0.0057
- 'Paciuksenkaari', 0.0056

Betweenness Centrality

The graph below illustrates the city bike stations by betweenness centrality.



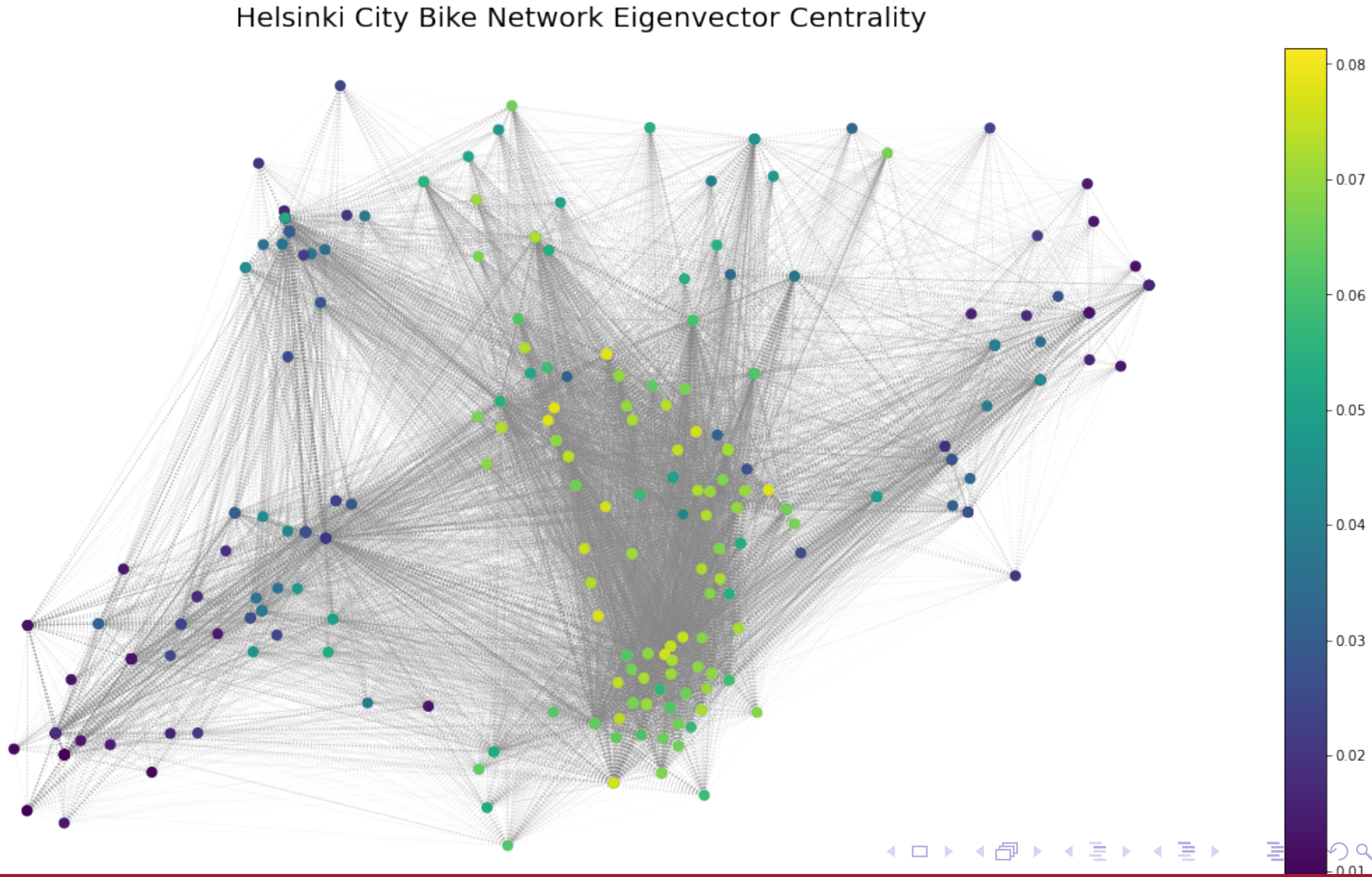
Eigenvector Centrality

In this type of context the Eigenvector centrality isn't indicate because we can forget a lot of factors that make a node important for the people.
Top 5 nodes by eigenvector centrality:

- 'Haukilahdenkatu', 0.0813
- 'Töölöntulli', 0.0781
- 'Paciuksenkaari', 0.0780
- 'Pasilan asema', 0.0779
- 'Huopalahdentie', 0.0777

Eigenvector Centrality

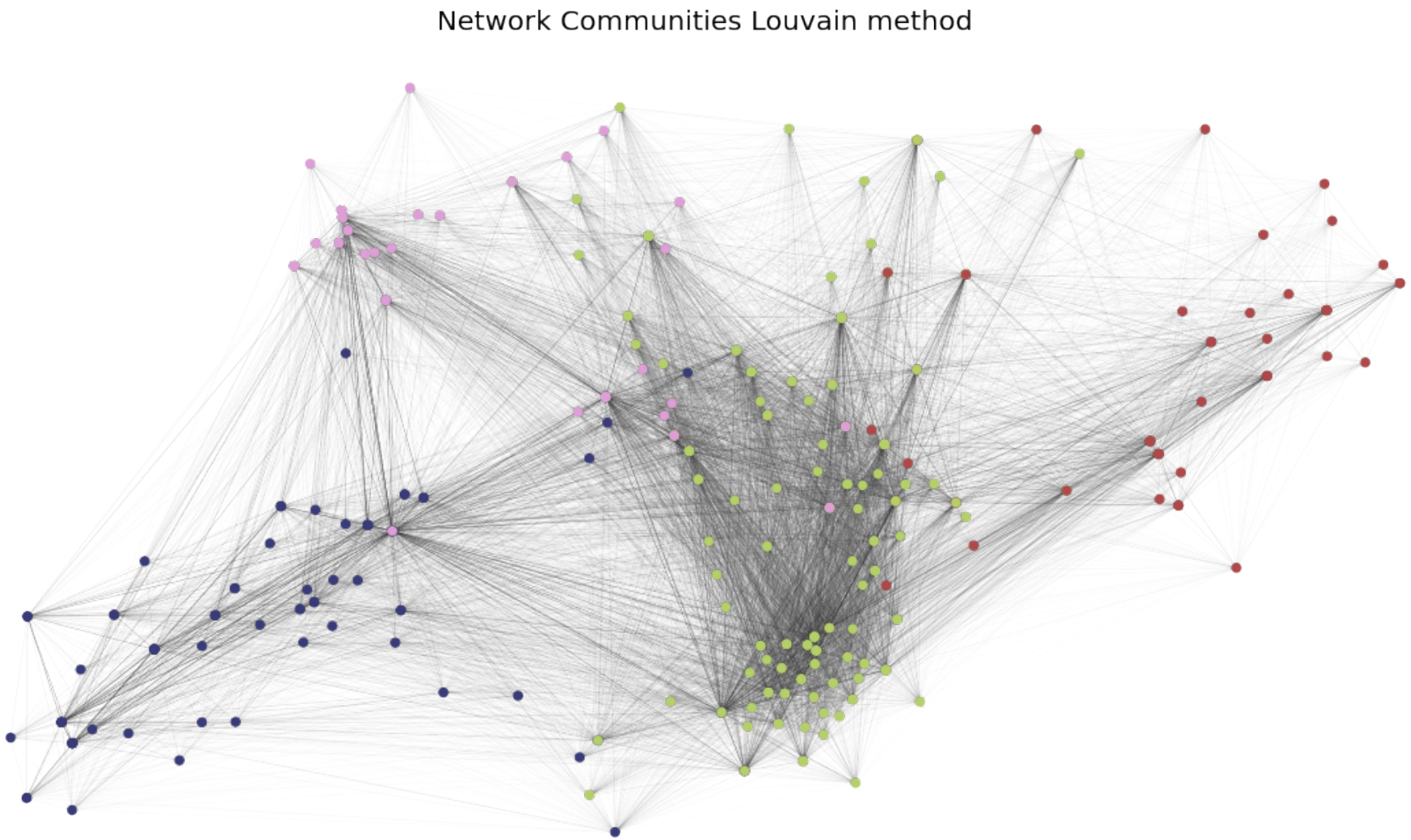
The graph below illustrates the city bike stations by Eigenvector centrality.



Louvain Method

The Louvain method reveals that bike usage inside these four communities is higher than the bike usage between these communities. Nodes with high betweenness centrality are usually the places where the communities might be divided into modules. This division doesn't come as a surprise taking into account the geographic character of the Helsinki archipelago.
It is important to highlight that all transportation networks are spatial networks and their structure and evolution are closely intertwined with physical constraints.

Louvain Method



Conclusions

In the end we can say that the part with the highest usage of this service is in the communities of the centre of Helsinki and Espoo.
The stations with the highest betweenness are those for which the passage is obliged to reach a lot of other stations.
Now, how can we improve this service?

Conclusions

Green solutions:

- more bicycles in the most popular zones
- maintain the same pricing (that is very low)
- maybe create a discount or a reward for students or after an amount of kilometers to boost the usage of bicycles
- organize some events to sponsor the service

Conclusions

Business solutions:

- more bicycles in the most popular zones
- propose new prices based on the zones. Higher prices in most used areas
- reduce the free ride to 30 from 15 minutes
- charge the prices