Recommending the Bicycle lane for Reduction of Bicycle accident

Minjun Lee

May 18, 2020

1. Introduction

1.1 Background

Nowadays there are many people who ride bicycle in Seoul (Korea). Actually, releasing 'Ttareungi'service from Seoul administration, the people who ride bicycle is increasing. And, also the bicycle accidents tent to increase. even if the rider would be careful, the unsuspected accidents will be happened. To solve this problem, It needs to facility for safety of the rider. So, Seoul administration will install the Bicycle Lane for the safety of the rider. But it's hard to decide the place where should be installed, where should be begun and where should be finished. To decide the place where is put Bicycle Lane, we will use the Data that include the number of accidents, location, the number of victims of bicycle accident from 2012 ~ 2018 years.

1.2 Problem

There are many cause of Bicycle accidents. therefore it's hard to decide the place where should be installed (actually, it's huge construction and takes huge money).and it's important to be efficient. The data includes the Location of accidents, the number of accidents and victims. we will use this data and suggest the place for installation.

1.3 Interest

Actually, administration will be interested in this project, for saving money and reducing accident. And also other city administrations may also be interested

2. Data Acquisition and Cleaning

2.1 Data Sources

the Data of bicycle accidents can be found in "TAAS Caraccidents System" that is open source data center. if you want to reach this data, follow this link "http://taas.koroad.or.kr/". this data is quite preprocessed, so we can save the time to preprocessing data.

2.2 Data cleaning

After downloading data, we will make table. but there are a lot of unnecessary columns in this data. we will drop the ID columns, code columns, polygon code column because it is unnecessary. and we will change the column names from korean to english for you guys who cannot read korean. the last cleaning is the dropping rows that have other citys' accidents

2.3 Feature selection

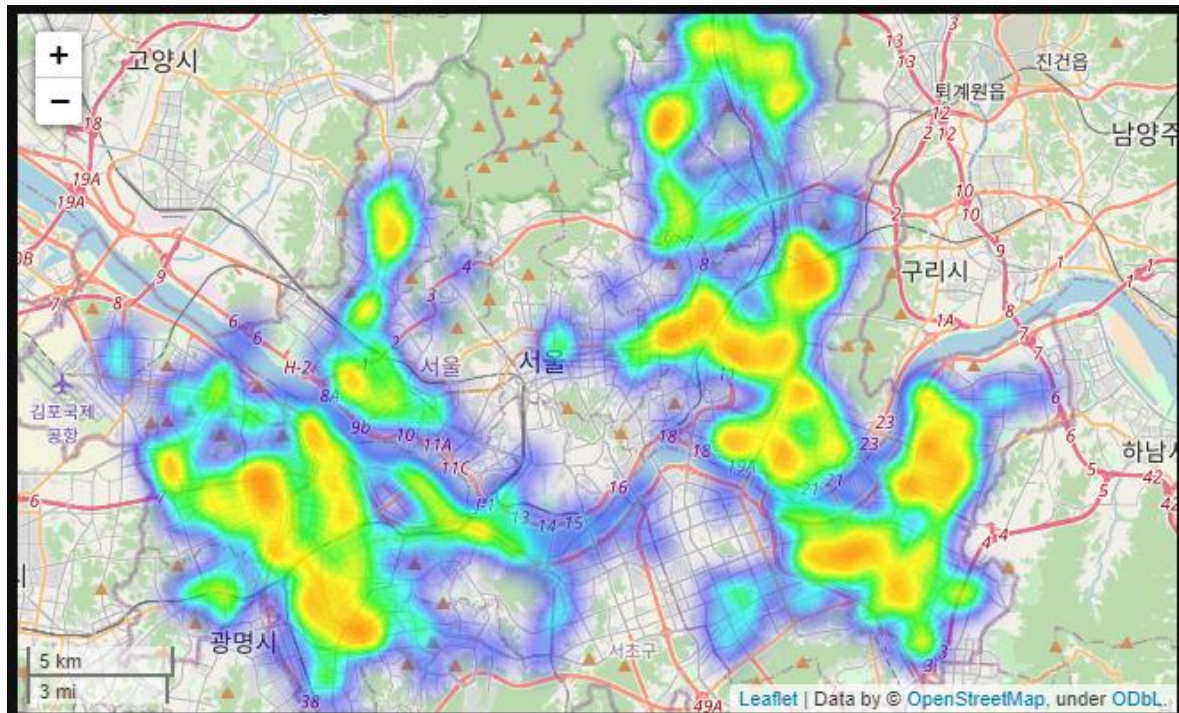If i finished the data cleaning there were 2516 rows and 10 columns in the data.

		Reason for columns
Dropped Columns	Accident ID	unnecessary
	Accident code	
	Polygon code	
Kept Columns	location name	describe the location in english
	num_accidents	describe the number of accidents
	victims	describe the number of victims
	location	using the location data

# 3. Exploratory Data Analysis

there are various analysis at this case. when we see the Scatter plot first, we can find it simply scattered evenly. becaues it is the marker that the accident are happened, not the number of the accident.
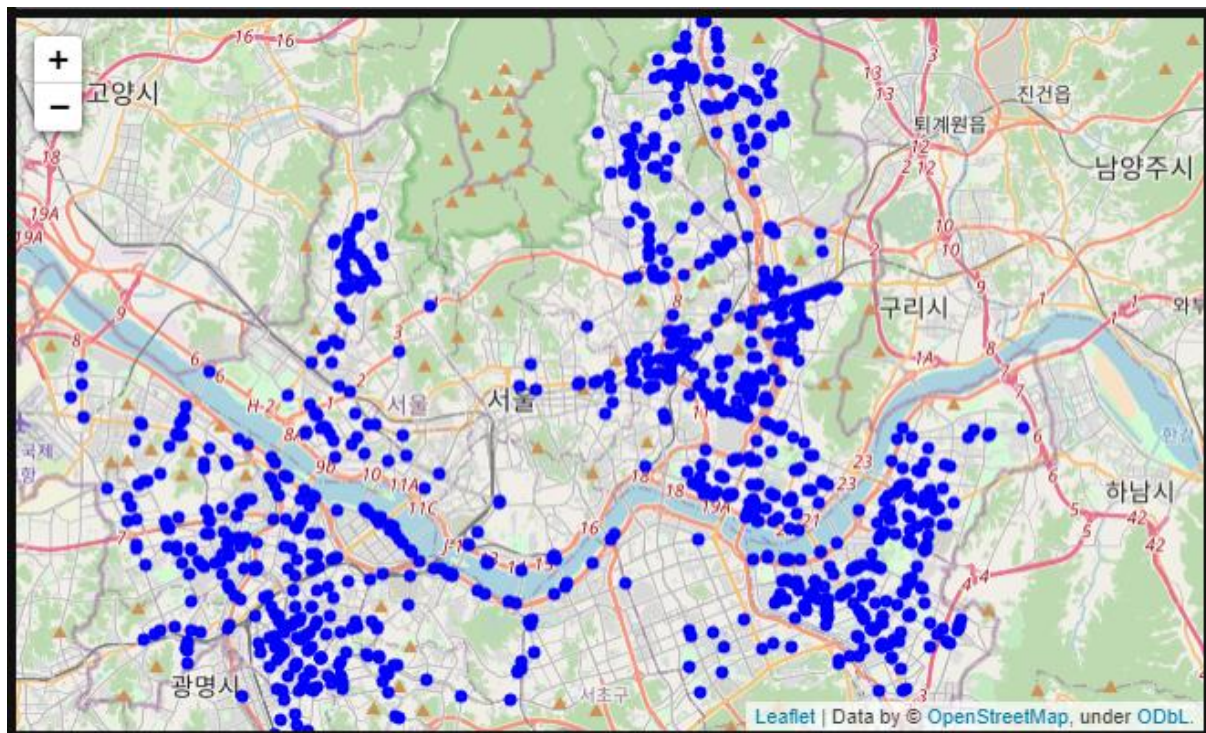

So, i also plotted Geo-HeatMap. then we could find the region Songpa-gu, Gwangjin-gu, Yeongdeungpo-gu, Dongdaemun-gu, Jungnang-gu thoes are concentrated on the Accident. Additionally, Extracting the data that is the location happend more than 9 accidents. And it was added in Geo-HeatMap. then we had also found the place thoes concentrated on accident. Songpa-gu, Gwangjin-gu, Yeongdeungpo-gu, Dongdaemun-gu, Jungnang-gu.

So, i also plotted Geo-HeatMap. then we could find the region Songpa-gu, Gwangjin-gu, Yeongdeungpo-gu, Dongdaemun-gu, Jungnang-gu thoes are concentrated on the Accident. Additionally, Extracting the data that is the location happend more than 9 accidents. And it was added in Geo-HeatMap. then we had also found the place thoes concentrated on accident. Songpa-gu, Gwangjin-gu, Yeongdeungpo-gu, Dongdaemun-gu, Jungnang-gu.
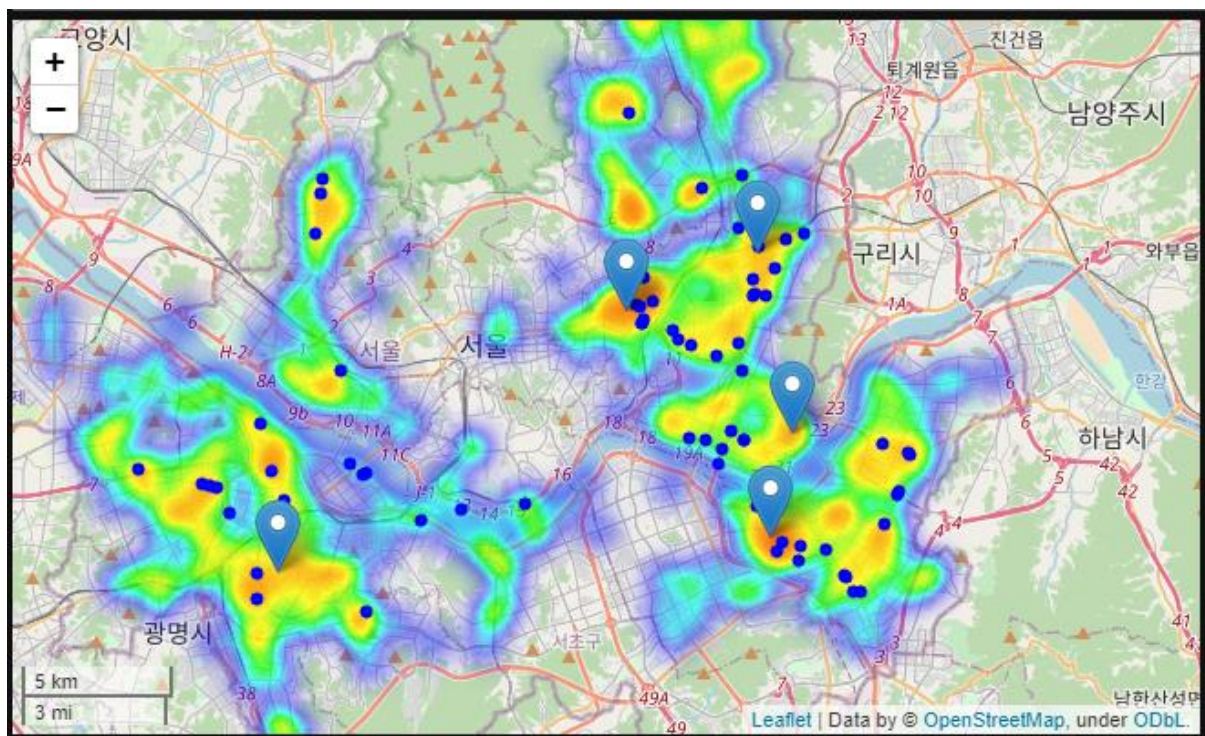
And again, I checked the Data that location happen the serious injuries more than 5 times. it had also found at same location. (Red circles are those Location)
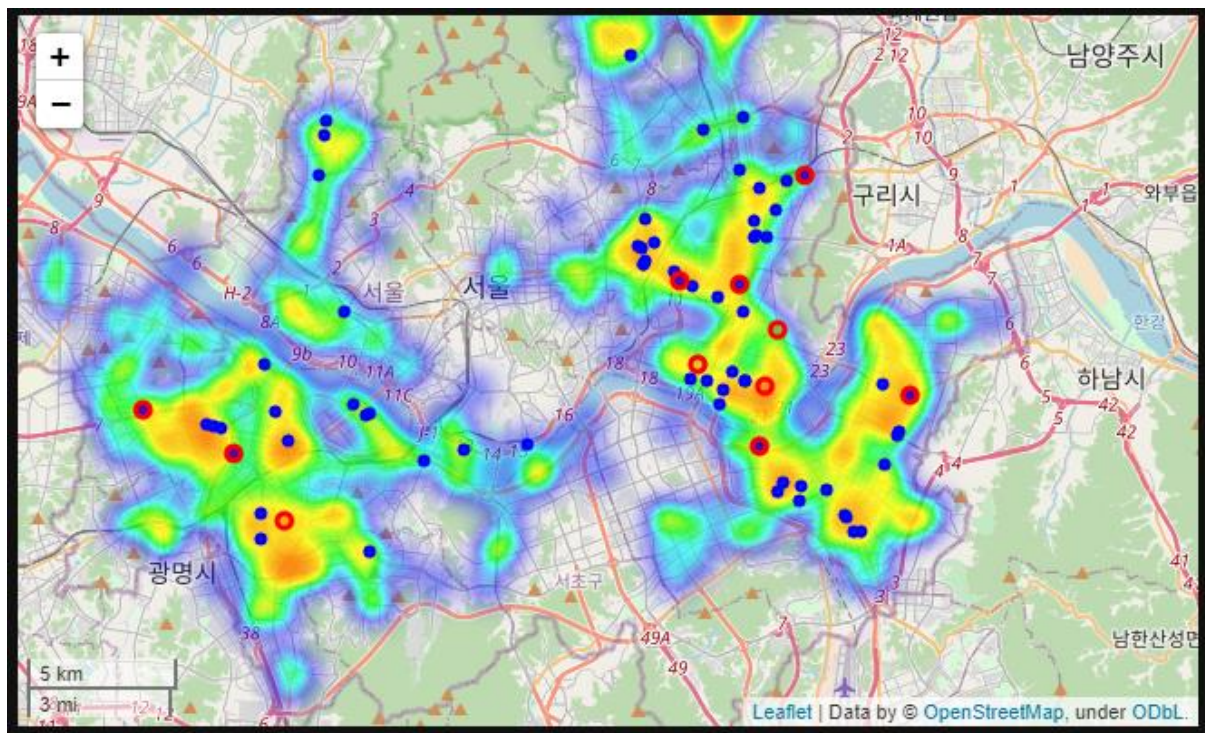
Finally, I plotted Choropleth Map to find the most common region at accident. then we could find the Songpa-gu is the most common region obviously.
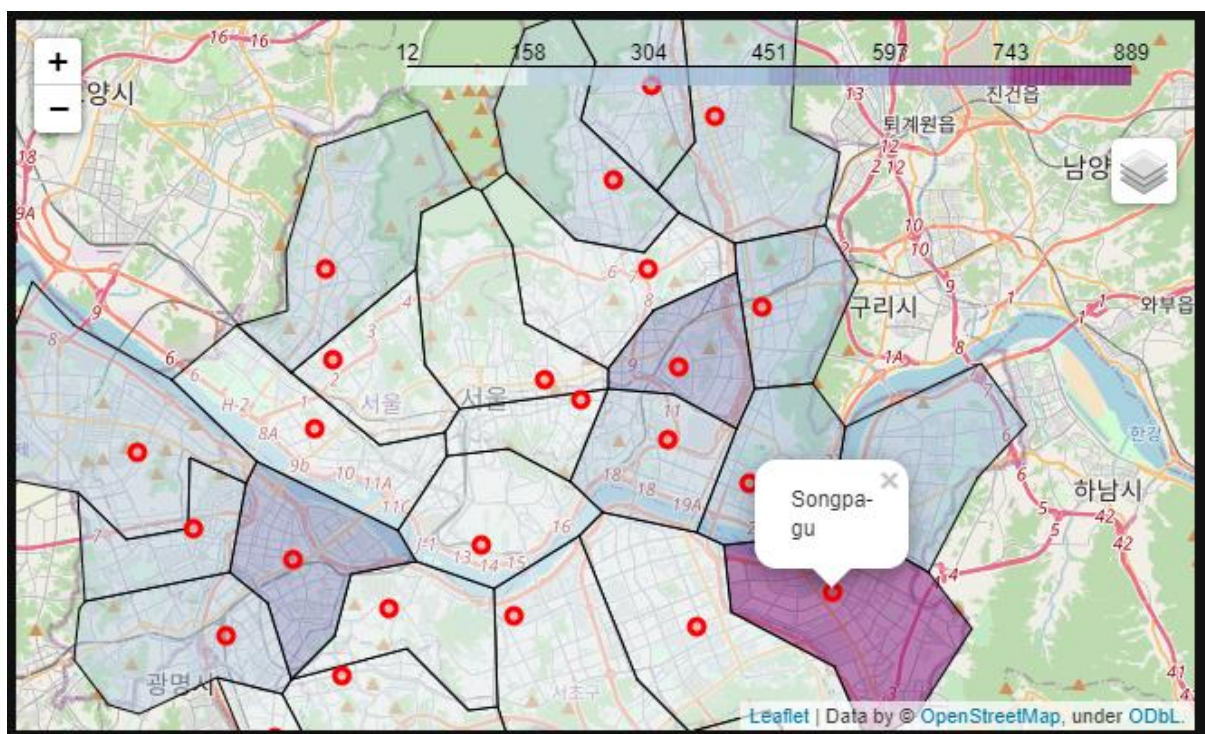
Collecting the data top 5 region of the number of accidents, the 1st is Songpa-gu, 2nd is Yeungdeungpo-gu 3rd is Dongdaemun-gu, 4th is Jungnang-gu. 5th is Gwangjin-gu.

	City_name	Num_acc	Num_total	Num_dead	Num_serious	Num_light
21	Songpa-gu	889	941	7	316	458
23	Yeongdeungpo-gu	555	612	5	157	329
1	Dongdaemun-gu	531	554	5	170	288
14	Jungnang-gu	393	417	2	118	188
11	Gwangjin-gu	318	333	3	135	151

4. Conclusion

Based on Data, I thought it is proper to connect the location where the accidents happened frequently for installing bicycle lane. So I selected the location where accidents happened more than 9 times and found the street where accident happened, There were some times accident in same street but these are different part of street.
Like below.

So, I recommend to install the bicycle lane along with that street. it is efficient and economical to install the bicycle lane. And definitely, it helps to reduce the ratio of bicycle accident.