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