**빅데이터(8752) 텀프로젝트 보고서**



과목 : 빅데이터기초(8752)

담당 교수 : 신문선 교수님

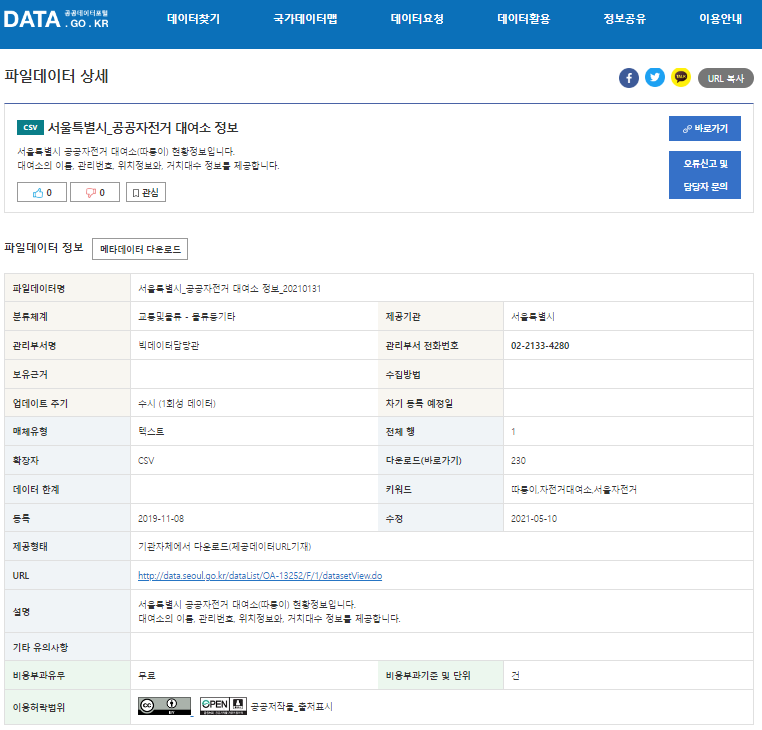
학과 : 소프트웨어전공

학번 : 202020899

이름 : 박보겸

**텀프로젝트 선정 주제 : 서울시 공공자전거 대여소 정보**

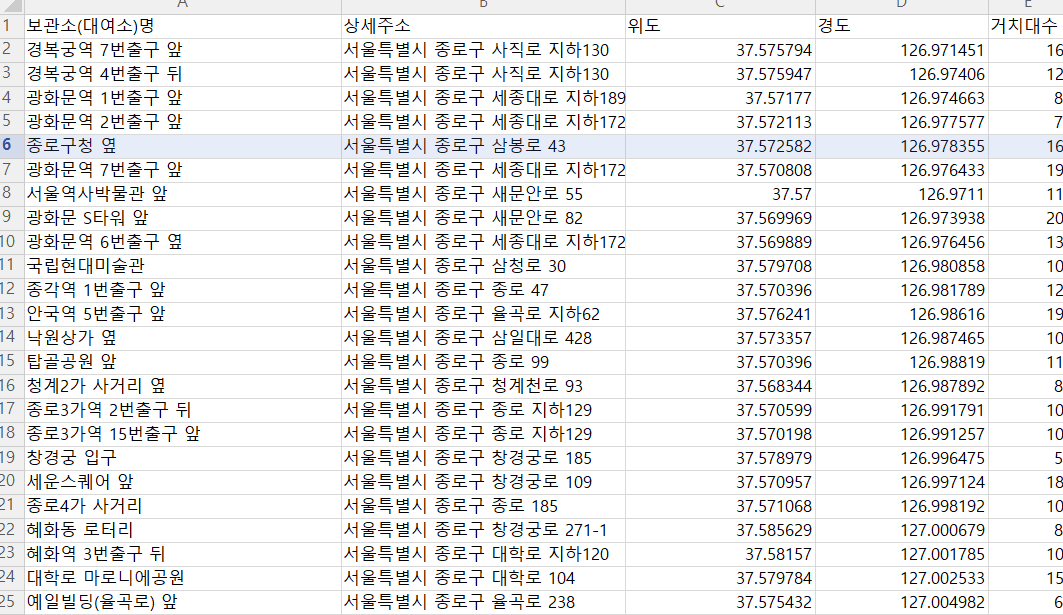
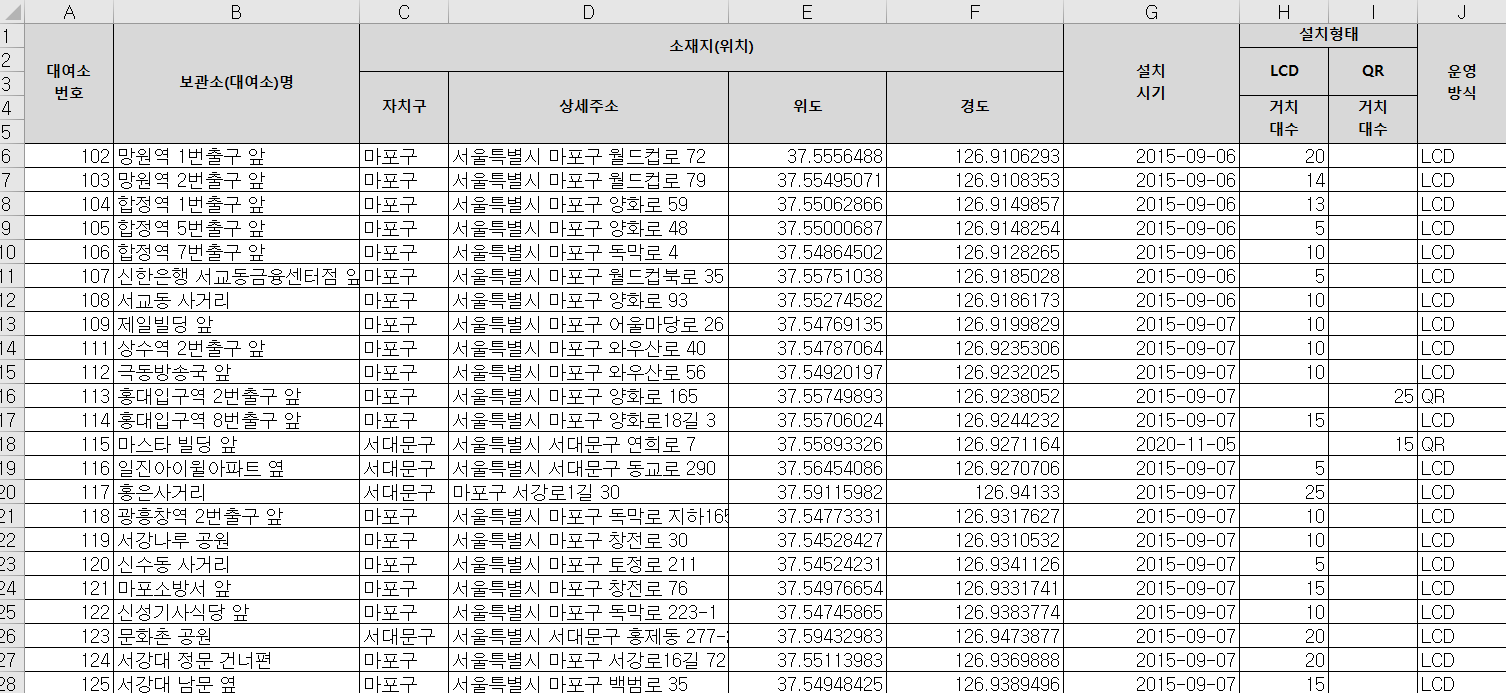
자동차 배출가스는 지구온난화를 가속화하고, 도시의 스모그를 발생시키는 등 현대 사회 대기오염의 주범으로 알려져 있다. 또한, 자동차의 매연에서는 일산화탄소, 탄화수소, 질소산화물, 오존, 미세먼지 등의 환경오염물질이 배출되는데, 특히 자동차에서 배출된 미세먼지는 입자가 매우 작아서 폐를 거쳐 혈관이나 혈액까지 침투하기도 한다. 이렇게 자동차의 이용은 심각한 환경 오염을 일으키는 만큼, 자동차 대신 자전거를 이용하는 것은 계속해서 환경문제 감소를 위한 대안방안으로 제시되고 있다. 그러나, 방치 자전거가 늘어남에 따라 방치 자전거 수거량도 늘고 있다. 오랜 시간 방치된 자전거는 녹이 슬게 된다. 이렇게 방치된 자전거의 녹은 비와 눈에 의해 물과 함께 땅으로 흘러 토양과 수질을 오염시킨다. 또한 녹슬고 고장이 난 방치 자전거는 고철 처리와 재활용 비용이 크게 발생하기 때문에 재활용하기도 어렵다. 개인 자전거를 사용하다 보면 이용하지 않게 되어 방치되는 자전거가 발생하기 마련이다. 따라서 개인 자전거 보다는 모두가 공공으로 대여하여 사용할 수 있는 공공 자전거를 이용하면 자동차로 인한 환경오염과, 개인 자전거 방치로 인한 환경 오염을 방지할 수 있으므로 공공 자전거를 더 적극적으로 이용할 수 있도록 공공자전거 대여소 정보를 수집하여 분석 및 시각화하기로 하였다.

< 데이터 수집 >

‘서울 열린데이터 광장’이라는 페이지에서 서울시 공공자전거 대여소 정보데이터를 다운받아 csv파일을 저장

**1. 지오맵에 대여소 나타내어 시각화**

< 데이터 준비 >



필요한 정보만 남긴 뒤 추가로 정보가 없는 위도, 경도를 입력하여 데이터를 저장

<분석 모델 구축 및 시각화>

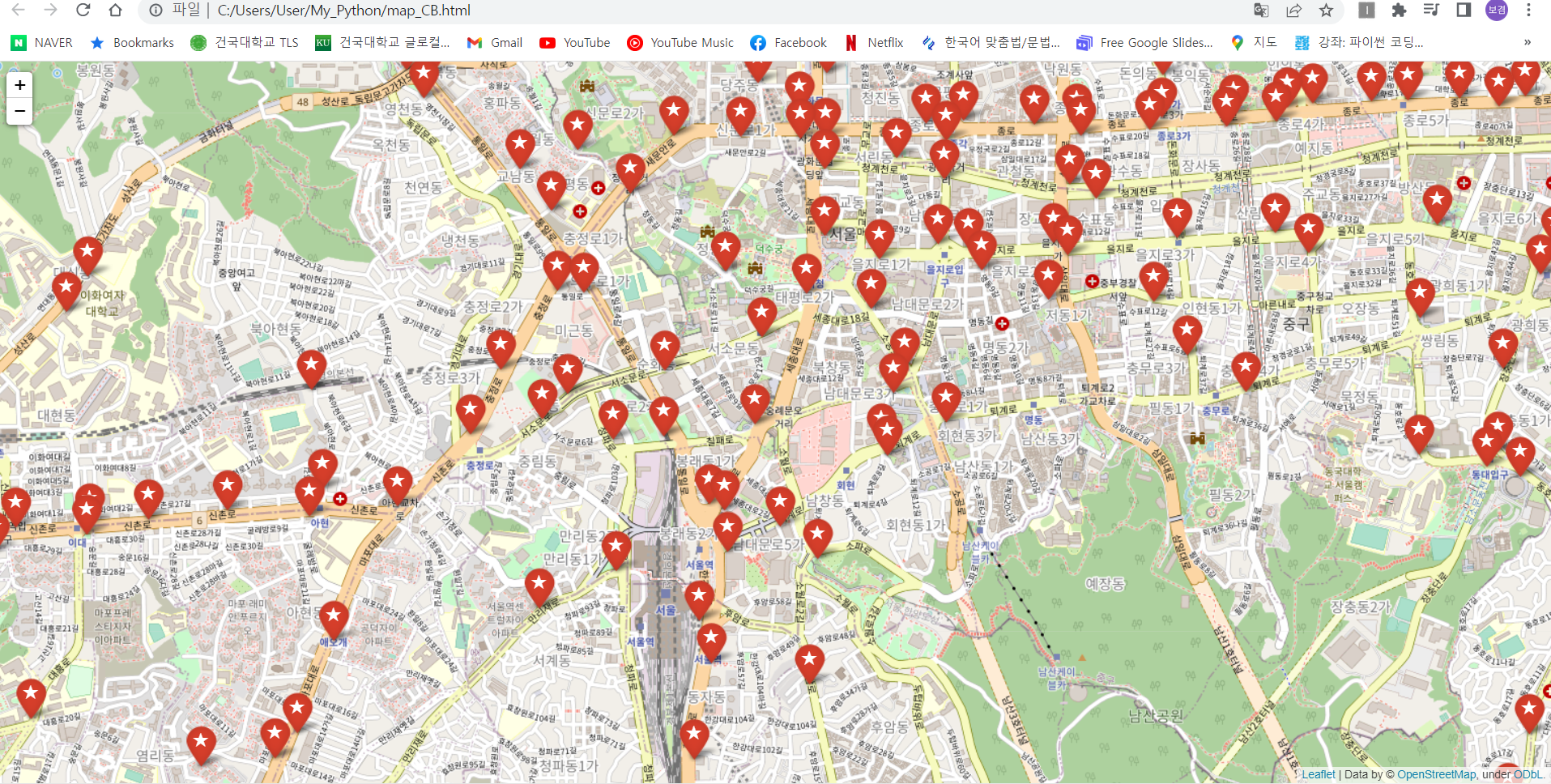


정리한 csv파일을 CB\_geoData 객체로 로드한 후 지도 객체인 map\_CB를 생성

For 반복문을 이용하여 CB\_geoData 객체에 있는 자전거 보관소 위치 정보를 하나씩 읽음

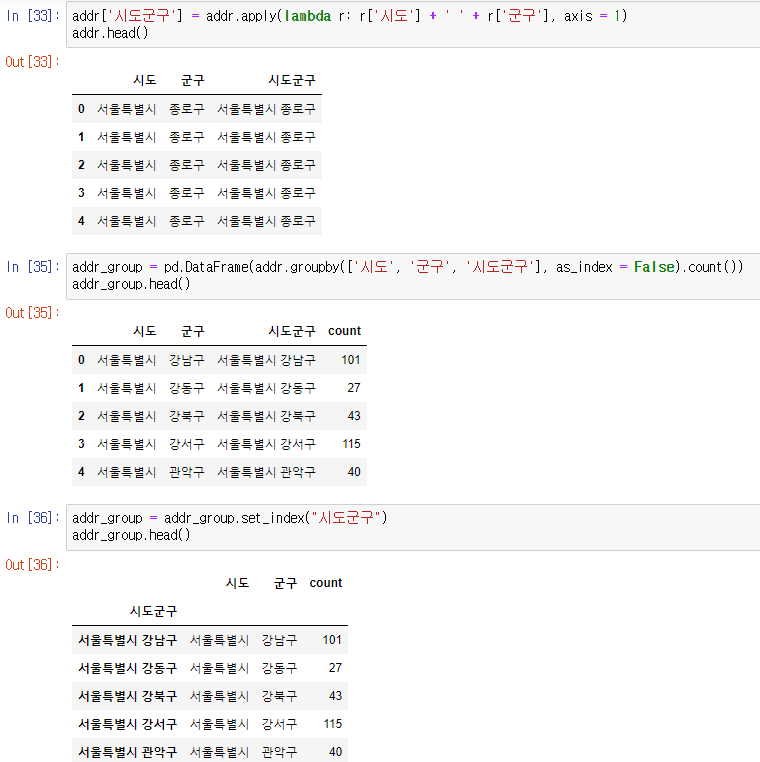
완성된 지오맵을 저장하고 파일을 웹 브라우저에서 열어 확인

<결과>



**2. 행정구역별 자전거 대여소 수 바 차트 이용해 시각화**

<데이터 준비>

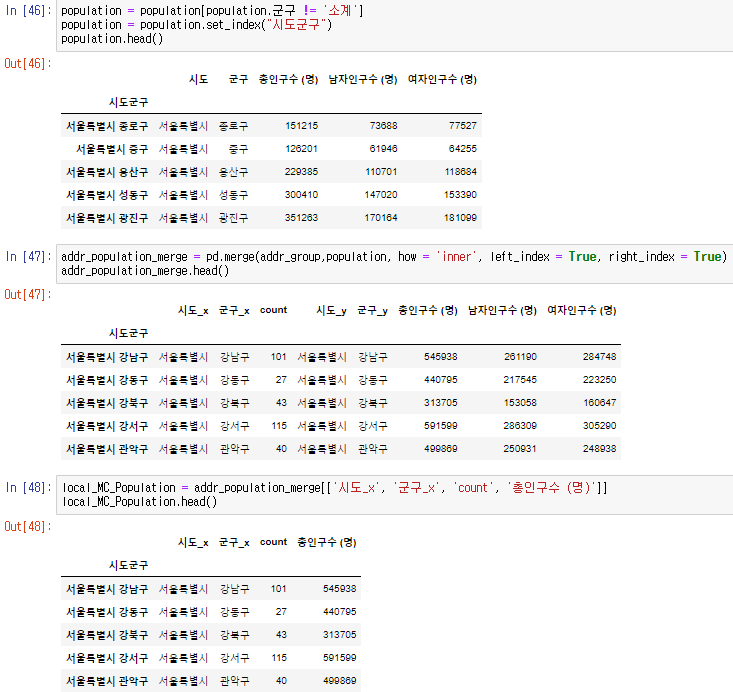
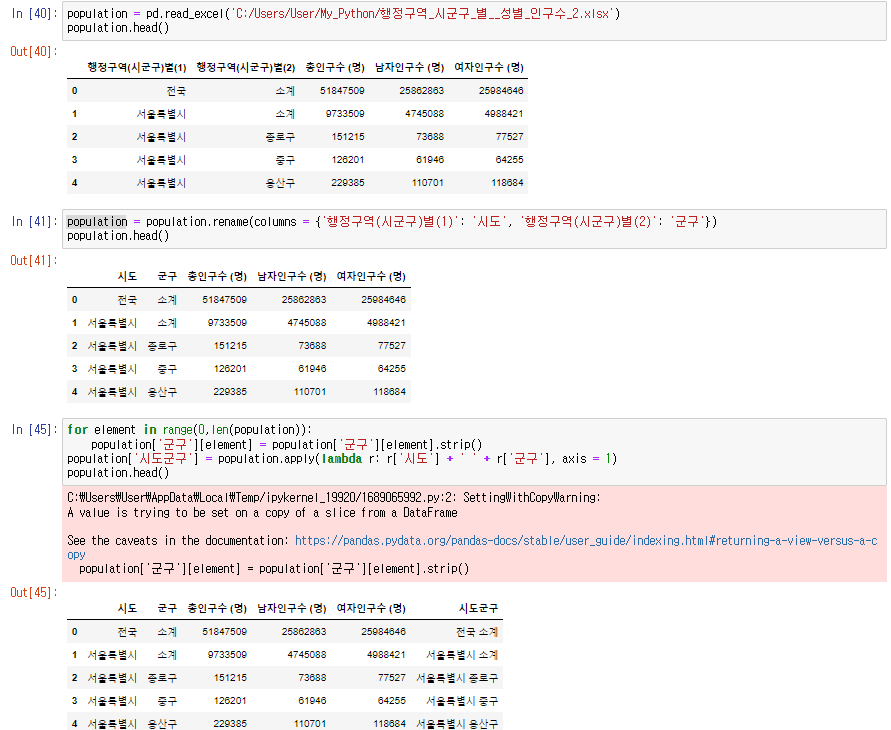


Bike\_rental.csv의 상세주소를 정리 후 bike\_rental2.csv파일로 저장하고, 파일을 data 객체로 로드

주소에서 시도, 군구 정보를 분리

행정구역별 공공자전거 대여소 수를 구하여 count 칼럼에 저장함

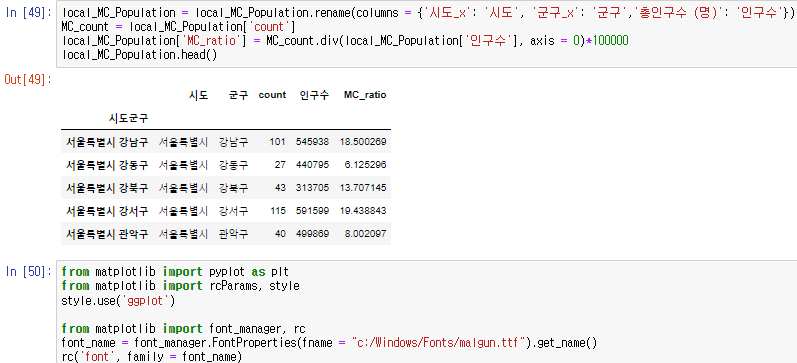
[‘시도군구’] 컬럼을 데이터프레임 병합에 사용할 인덱스로 설정



행정구역별 인구수 데이터 준비한 후 인구수 대비 공공자전거 대여수 비율을 구하여

local\_MC\_Population의 [‘MC\_ ratio’] 컬럼에 추가

<결과 \_ 바 차트>



첫 번째 바 차트 결과는 local\_MC\_Population 객체의 [‘count’] 컬럼 값을 오름차순으로 정렬하여, 행정 구역별 공공자전거 대여소 수에 대한 바 차트를 그린 것이고,

두 번쨰 바 차트 결과는 local\_MC\_Population 객체의 [‘MC\_ratio’] 컬럼 값을 오름차순으로 정렬하여, 행정구역별로 인구수 대비 공공자전거 대여소 비율에 대한 바 차트를 그린 것임