

**VAMAP**

OSS 프로젝트 프로젝트 제안서

오픈소스 프로젝트

김효수 교수님

이상민 | 20164204

주석천 | 20151363

이현재 | 20133096

민성재 | 20145223

정유석 | 20141359

# 프로젝트 주제 소개

국가적 정책의 일환으로 수 많은 공공데이터가 생겨나고 있다. 또한 지리적 위치와 연결 되어있는 데이터는 누구나 가지고 있고 누구나 만들 수 있다. 하지만 이런 데이터가 많은 것과는 별개로 분석하는것과 시각화 하는 것은 별개의 문제이다.

기존의 프로그램들은 지도상에서 단순 특정데이터의 밀집도 혹은 분포도만 보여주는 경우가 많다. 하지만, 정작 밀집도를 통해 어떠한 분석결과를 한눈에 알아보기는 쉽지 않다. 따라서, 우리의 개발목표는 사용자가 가지고 있는 위치기반 (위도 및 경도) 데이터를 입력을 하면 도로, 건물, 지하철의 데이터와 함께 분석하여 각 위치기반 데이터 객체마다 가치를 계산하고 시각화 해주는 것을 목표로 한다. 이때 서울에 국한하여 하는 것을 목표로 한다.

# 주제의 필요성 / 중요성

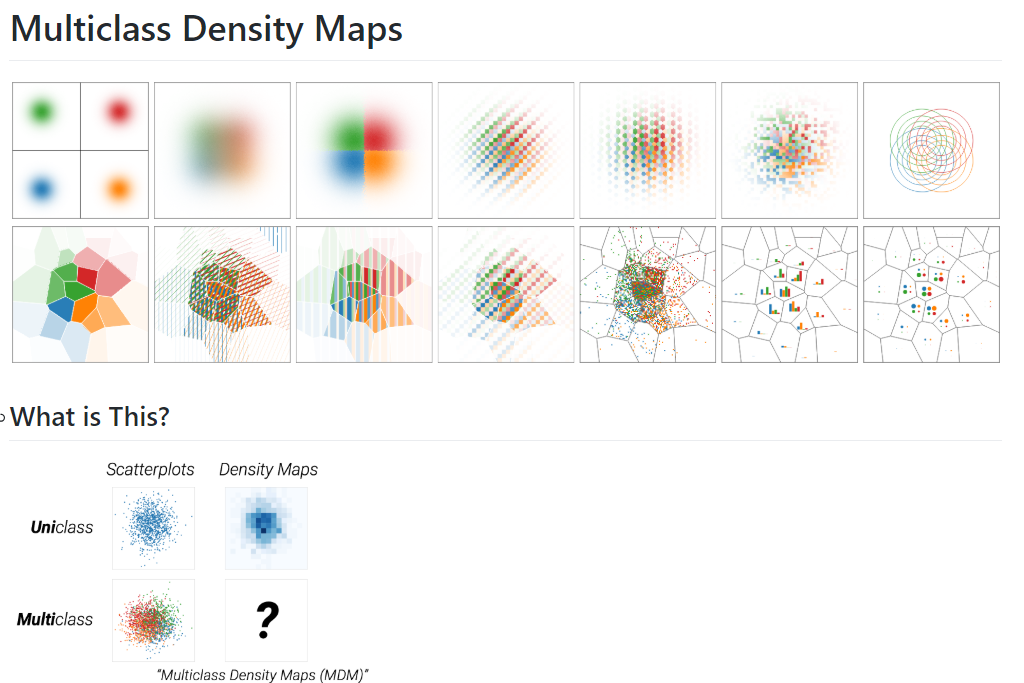
많은 지역 기반 데이터가 있지만 입지분석을 위해서는 사실상 사람의 손, 분석으로 이루어지는 경우가 많다. 지도상으로만 보고 어떠한 건물의 가치를 파악하는것은 쉬운일이아니다.  그렇기 때문에 도로, 건물의 분포와 그 밀집도를 사용하여 평균적인 입지, 분포를 분석하여 유저가 응용할 수 있도록 시각화, 정량화를 해주는 것으로 목표로 하고있다. 이때 건물이나 도로정보는 서울시 공공데이터를 가공하여서 사용할 예정이다.

서울시 공공데이터는 대부분 xls파일의 형태이고 그 데이터 또한 정규적이지 못하다. 그렇기에 실 사용과는 상대적으로 거리가 있다. 그렇기에 우리 프로젝트는 이를 정돈하고 쉽게 개발자가 사용할수 있는 형태로 정규화 하는 것을 목표로 한다. 또한 사용자가 어떠한 건물의 분포적가치를 알고싶을때 유용하게 사용할수있는 툴을 제공할것이라고 생각한다.

# 관련 기술 소개

<https://github.com/e-/Multiclass-Density-Maps>

다중 클래스의 밀집도를 보여주는 맵

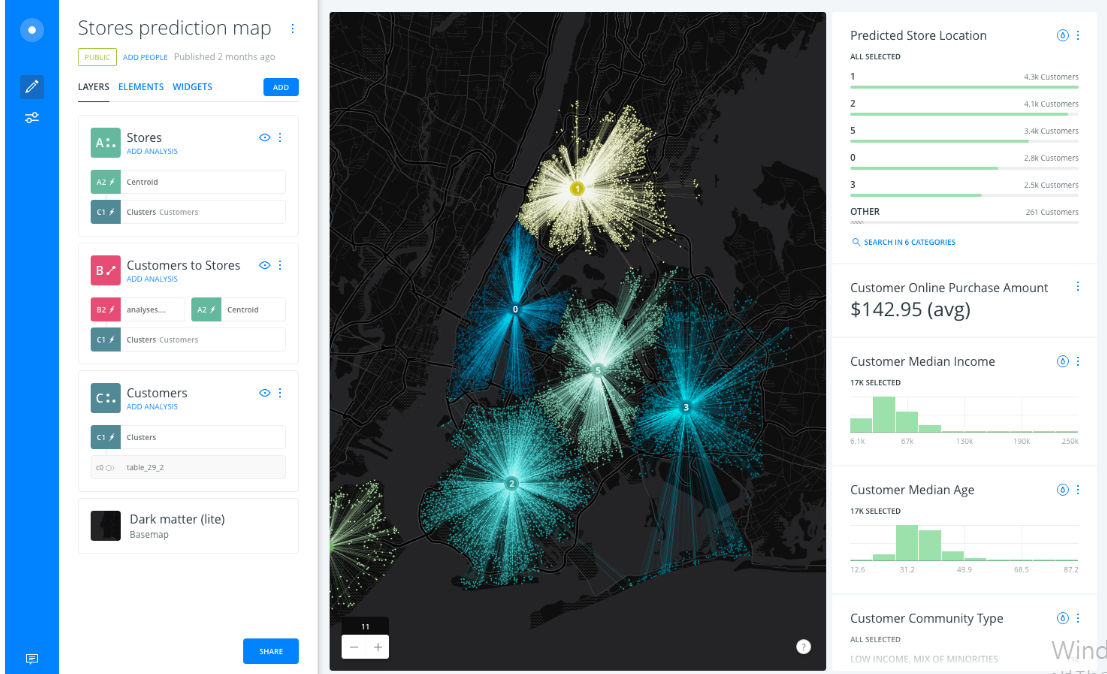


<https://github.com/greatsong/2016dataproject>

구 별 밀집도를 비교

<https://github.com/CartoDB/cartodb>

지역별 수에 따라서 색을 표현, 도로의 구조 등을 보여주며 유저가 modify할 수 있게 만든 플랫폼



<https://glasvase.tistory.com/14?category=672672>

서울 공공데이터 통계장인

<https://planet.openstreetmap.org/> 길 추출

<https://developers.google.com/maps/documentation/javascript/geocoding> 건물 request

<https://stackoverflow.com/questions/5401661/google-maps-how-to-get-buildings-polygon-coordinates-from-address>

google map에서 건물데이터 추출

<https://github.com/ndrberna/CarAccidents>뷰가 이뻐요

# 본인들만의 강점, 차별성

기반시설 혹은 편의시설에 대한 입지적 가치를 보여주는 지도라는 점에서 독자적인 정체성을 갖는다. 즉, 이 지도는 다양한 시설들 중에서 각 사용자들이 원하는 시설들에 대해 입지적 가치를 제시함을 통해 입점 등에 대한 판단의 기준을 제공한다.

가치는 지도가 제시하는 기반시설, 편의시설간의 연관성 혹은 밀집도 기준으로 매겨지기 때문에 입점, 폐점 등을 원하는 각 사용자들에게 정확한 판단 기준을 제시한다.

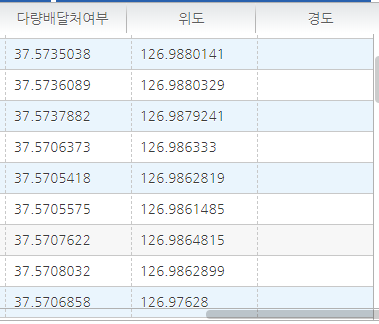
더불어 국내 지도 서비스는 google api를 통해 개발하기 때문에 국내 시설 파악에 취약하다. 이 지도는 국내 정책중 하나인 공공 api를 활용함으로써 정확한 정보를 기반으로 만들어진 지도 서비스이다.

# 개발 방법 및 계획 (구현 범위, 사용할 tool등)

<http://data.seoul.go.kr/dataList/datasetView.do?infId=OA-11679&srvType=S&serviceKind=1&currentPageNo=1&searchValue=&searchKey=null>

도로데이터

<http://data.seoul.go.kr/dataList/datasetView.do?infId=OA-13227&srvType=S&serviceKind=1&currentPageNo=1>건물데이터- 층수, 좌표, 건물이름



에서 도로 데이터와 건물데이터를 추출하여 실제 지도와 매핑하는 과정을 거친다. 그 후 사용자가 가지고있는 데이터와 지도 데이터를 연결하여 그 안에서 연관성 평균, 특이성 등 geological한 다양한 특징을 추출해낸다. 그후 그 특성에 비해 상대적으로 유저 데이터가 부족하거나 많은 장소를 추출하거나 히트맵의 형식으로 표현하여준다.

1.     xls형태로 되어있는 도로 데이터와 건물 데이터를 실제 지도와 매핑하는 과정->좌표 변환과 경로 표현을 위해 번거로운 작업이 필요

2.     사용자가 가지고있는 데이터와 우리의 지도와 매핑하는 과정-> google api에서 제공하는 geocoding과 유사한 형태로 진행- 길,건물,데이터(cu)가 지도위에 표시되어있는 상황

.

3.     그안에서 데이터를 추출하는 과정, 이때 pca와 같은 방식으로 깊지 않은, 3~4가지의 데이터들의 연관성만을 확인한다. 밀집도(시골에 길도 없고 건물도 없는데 cu도 없다, 그러면 괜찮 )라든지 뭐 그런거

4.     그 후 그 분석 데이터를 지도위에 매핑하고 높고, 낮은 데이터를 표현하여 준다.

openstreetmap api , naver open map api ( naver cloud api로 이전될 계획)

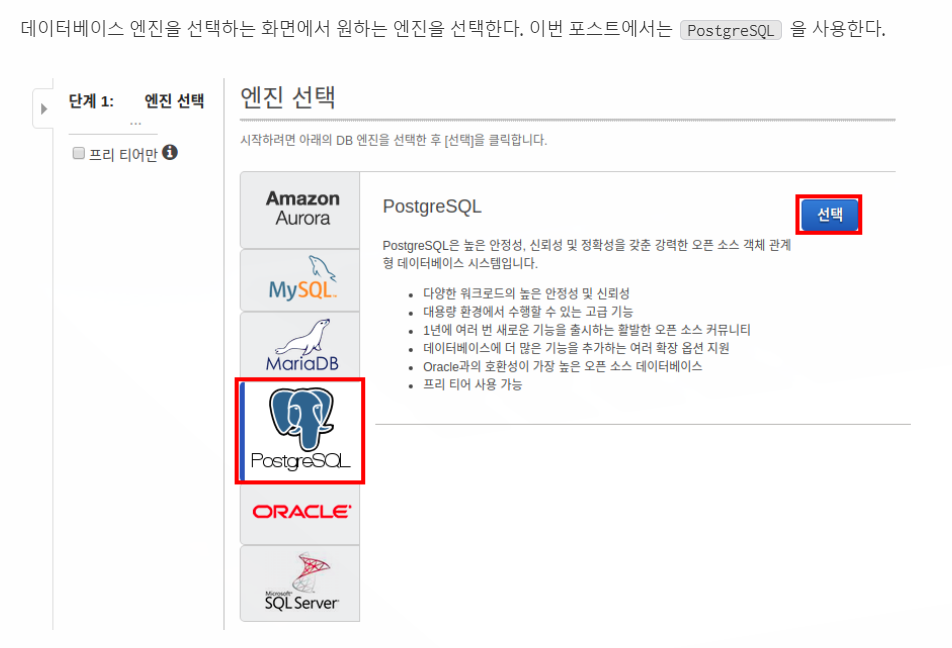
Openstreetmap api 와 서울시 공공데이터를 사용하여 길, 건물, 지하철 등 다양한 축을 기준으로 사용자가 가지고있는 데이터간의 상관관계를 분석하여 각 지점별 가치를 분석한다.  그후 유저는 지도를 확대, 축소를 하며 그에 맞는 데이터의 분석 결과를 보게 된다.

google에서 provisioning중인 GCP는 sql, 머신러닝, 서버, 데이터 분석등 많은 컴퓨터 자원이 필요한 여러 작업들을 하기에 최적화 되어있습니다. 서버 구축, 배포에 드는 시간을 줄이기 위해서 저희 팀은 구글 클라우드 플랫폼위에서 서비스를 진행하고자 합니다. 유저가 보낸 데이터셋을 gcp의 Dataproc이나 Bigquery에서 받아 저희가 가지고 있는 도로, 건물 과 같은 공공데이터셋과 상관분석을 통해 각 데이터의 상대적 가치, 지리적 중요도를 알아내고자 합니다.

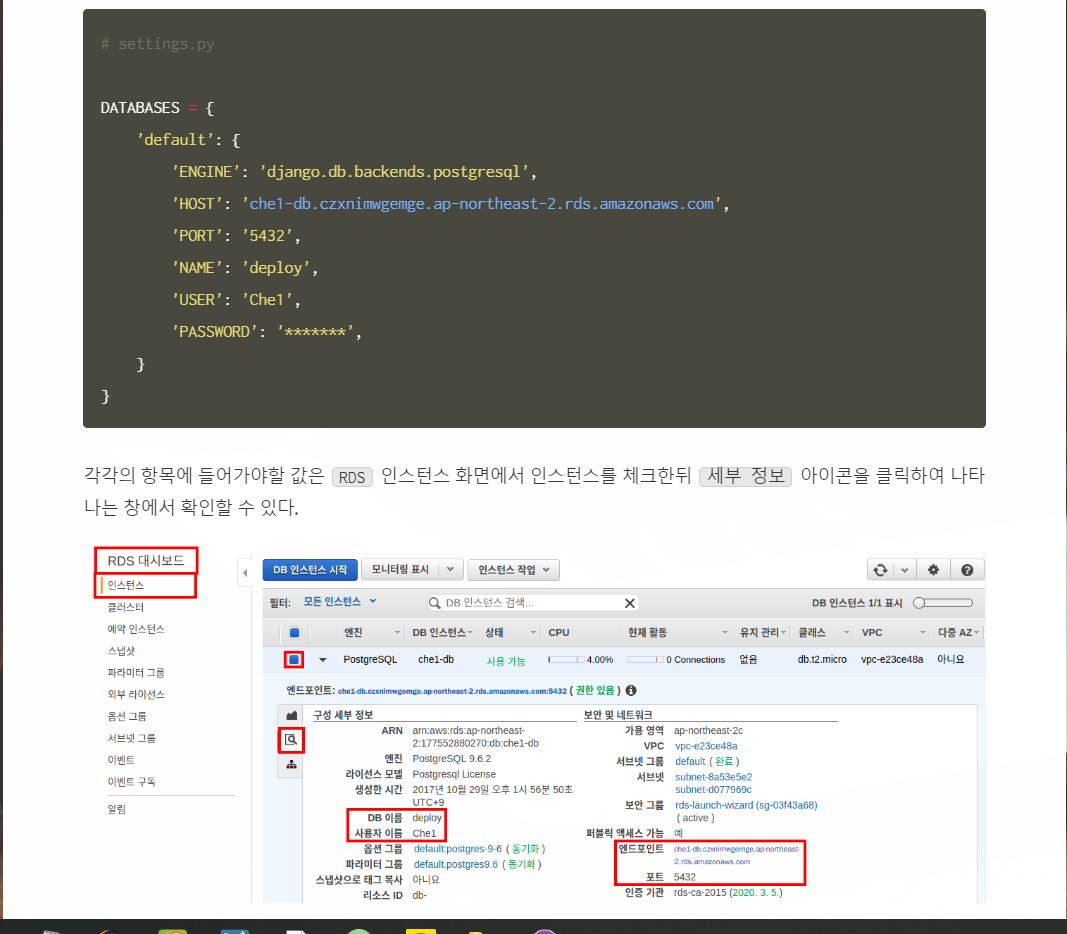
**- 사용할 tool**

**1. <**Amazone Web Service & Django 연동>

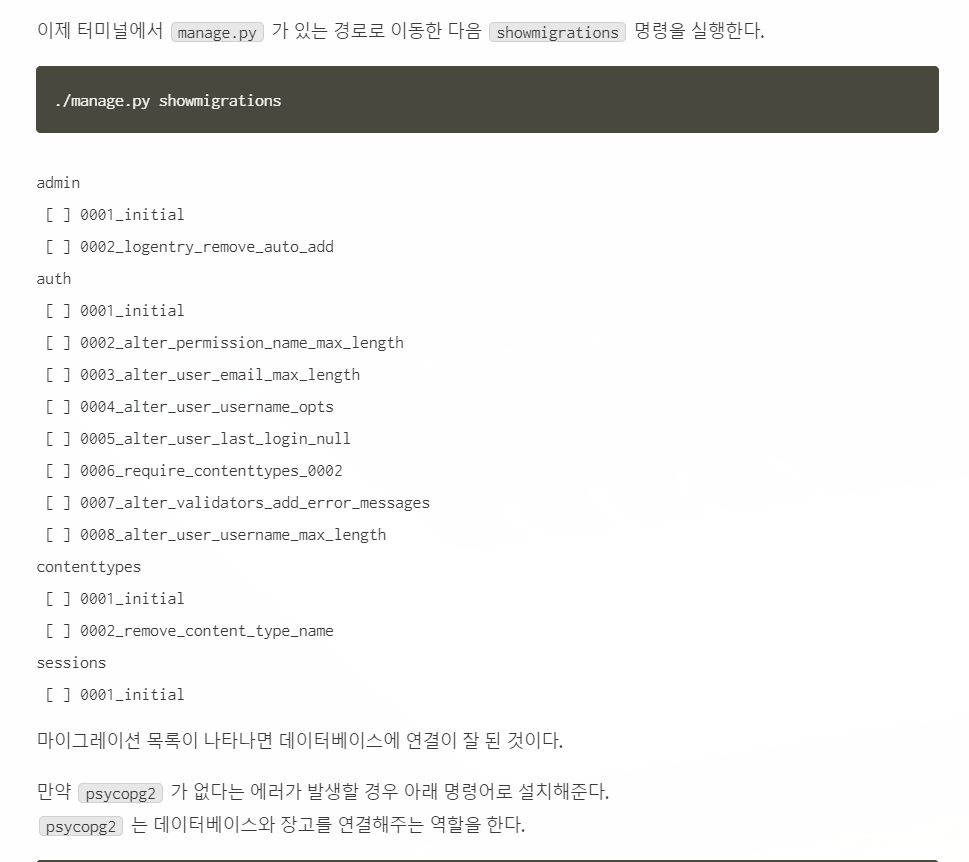
참고사이트 : <https://nachwon.github.io/django-deploy-5-rds/>



Django의 엔진 PostgreSQL 선택 후 AWS에서 RDS생성



RDS 인스턴스 내부의 세부 정보 설정



터미널 명령어를 통해서 AWS RDS와 Django를 연결