Table of Contents

[1. A description of project 2](#_Toc10022379)

[2. Related technology, service, opensource software 2](#_Toc10022380)

[2.1 Weka 2](#_Toc10022381)

[2.1.1 set up django 2](#_Toc10022382)

[2.2 RapidMiner 2](#_Toc10022383)

[2.2.1. install Chrome Extension 2](#_Toc10022384)

[2.2 Orange data mining 2](#_Toc10022385)

[2.2.1. install Chrome Extension 2](#_Toc10022386)

[2.2 R 2](#_Toc10022387)

[2.2.1. install Chrome Extension 2](#_Toc10022388)

[2.2 KNIME 2](#_Toc10022389)

[2.2.1. install Chrome Extension 2](#_Toc10022390)

[3. Vamap 3](#_Toc10022391)

[1. Architecture 3](#_Toc10022392)

[2. Workflow 3](#_Toc10022393)

[1. github 3](#_Toc10022394)

[2. personal 3](#_Toc10022395)

[3. implementation 3](#_Toc10022396)

[1. Parser. 3](#_Toc10022397)

[2. dataproc. 3](#_Toc10022398)

[2. scoring. 3](#_Toc10022399)

[2. dataproc. 3](#_Toc10022400)

[4. performance 3](#_Toc10022401)

[Main result 3](#_Toc10022402)

[5. Github repository 4](#_Toc10022403)

[5.1. commit, contributor, issue/pull request 4](#_Toc10022404)

[5.2. usage 4](#_Toc10022405)

# A description of project

## 1.1 Purpose

2.1

## 1.2 need

2.1

# Related technology, service, opensource software

## 2.1 Weka

### 2.1.1 set up django

사용자가 파일을 선택하고 (.arff형식) 데이터 전처리, 시각화등을 제공해줌

데이터 전처리 -> 표본추출(sampling) – 적당한 표본 크기를 추출하여 정보 손실 최소화,속성선택(feature selection) – 무관하거나 중복된 속성들을 배제, 중요도에 따라 가중치를 부여함,차원축소(dimensionality reduction) – 주성분분석(PCA) , 부분최소제곱회귀(PLSR) 적용,산화(discretization), 이진화(binarization), 변수변환 등을 이용한 전처리

## 2.2 RapidMiner

### 2.2.1. install Chrome Extension

1. access [https://github.com/choiyooung/summaryhelp/](https://github.com/choiyooung/summaryhelp/tree/feature/chromeExtension)

## 2.2 Orange data mining

### 2.2.1. install Chrome Extension

1. access [https://github.com/choiyooung/summaryhelp/](https://github.com/choiyooung/summaryhelp/tree/feature/chromeExtension)

## 2.2 R

### 2.2.1. install Chrome Extension

1. access [https://github.com/choiyooung/summaryhelp/](https://github.com/choiyooung/summaryhelp/tree/feature/chromeExtension)

## 2.2 KNIME

### 2.2.1. install Chrome Extension

1. access [https://github.com/choiyooung/summaryhelp/](https://github.com/choiyooung/summaryhelp/tree/feature/chromeExtension)

# Vamap

## 1. Team, position

## 2. Architecture

## 3. Workflow

### 1. github

Chrome extension

### 2. personal

Chrome extension

## 4. Use case

### 1. Social facility: Streetlight

Coverage:

Other data:

### 2. Convenience store: CU

Coverage:

Other data:

### 3. Cultural facility: Library

Coverage:

Other data:

## 5. 3D map result

## 6. implementation

### 1. Parser.

1. Chrome extension

### 2. dataproc.

1. Chrome extension

### 3. scoring.

1. Chrome extension

## 7. Implementation after MVP

### 1. Hyeonjae Lee.

1. Chrome extension

### 2. Sangmin Lee

1. Chrome extension

### 3. Seongjae Min.

1. Chrome extension

### 4. Seokcheon Ju

1. Chrome extension

# performance

## 4.1Difference

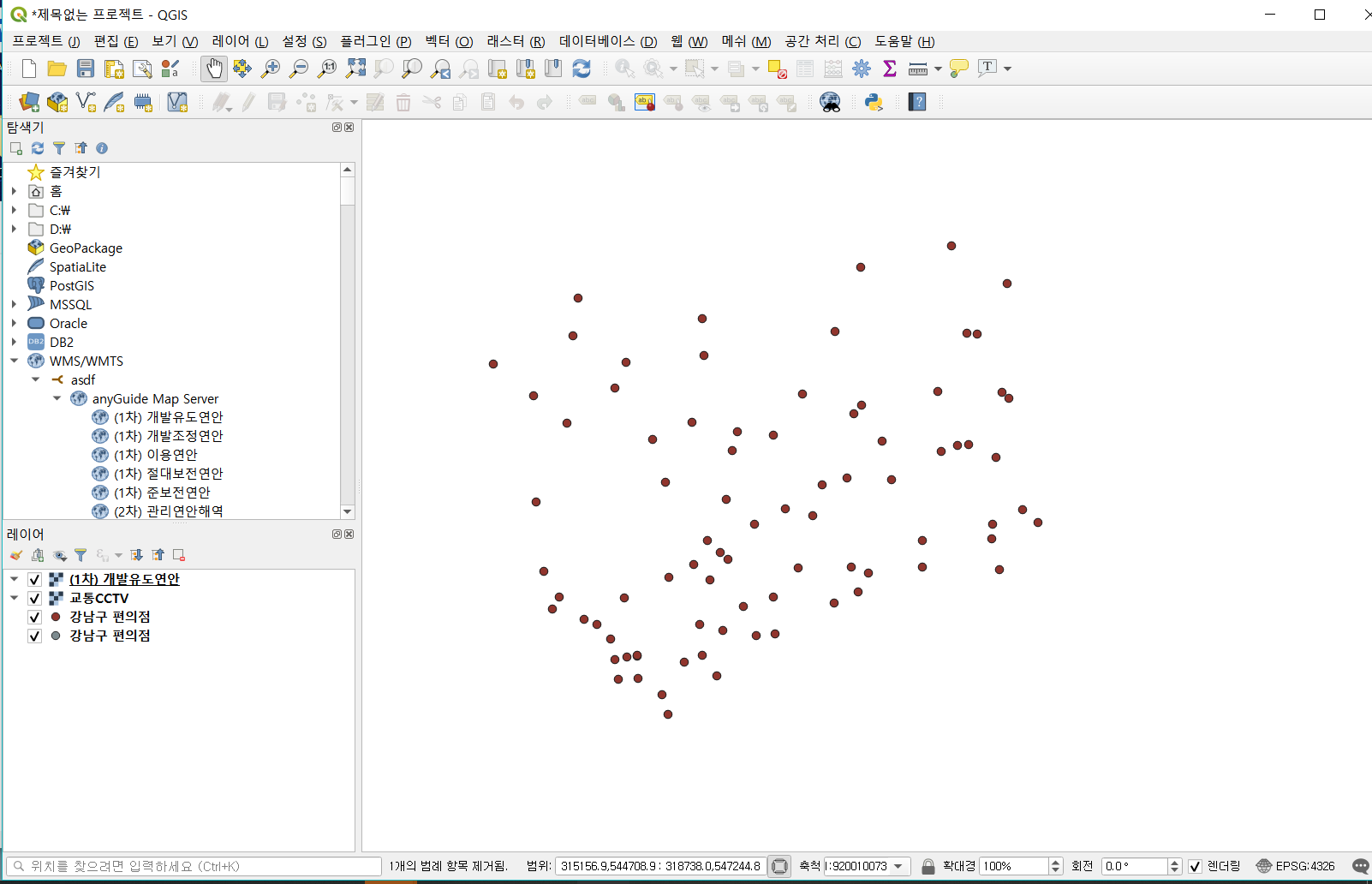
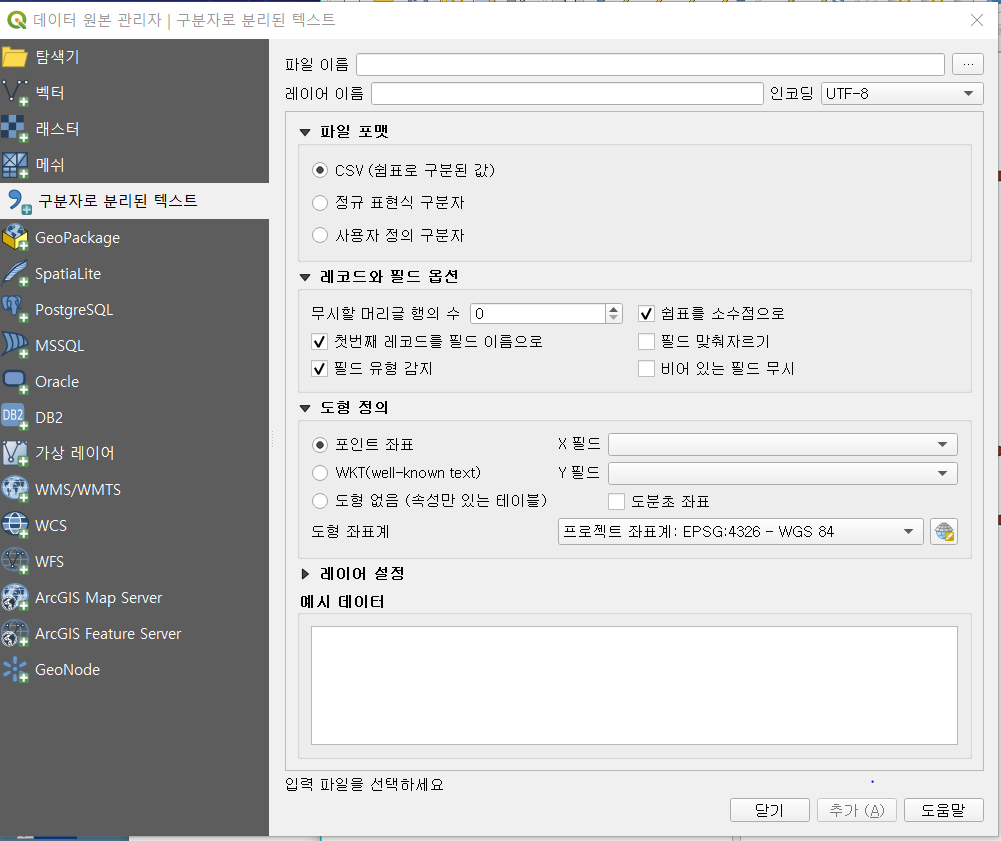
### 4.1.1 QGIS

### 4.1.1.1 Feature

.csv 파일을 추가하거나 다른 데이터베이스의 테이블을 가져오는 것이 가능하다.

이 파일이나 테이블속에서 좌표를 의미하는 X필드, Y필드를 설정해주면

프로그램상에 각 좌표가 찍히게 된다.

  
다만 바로 지도위에 각 객체들이 찍히는 것이 아니고 지도를 따로 가져와야 한다.

지도를 가져오는 방법은 API를 가져오거나 다른 사람이 등록한 플러그인에 있는 파일을 추가하는 방식이다.

또한 좌표를 기준으로 찍힌 객체, 지도, 행정구역 경계등이 이 프로그램에서는 각각 ‘레이어’ 라고 불리는데 이 레이어들을 한꺼번에 보고싶다면 동기화하는 과정도 필요하다.

QGIS 또한 오픈소스이기 때문에 많은 유저들이 등록해 놓은 플러그인을 가져다 쓸 수 있다는

장점이 있지만 애초에 프로그램 자체에서 데이터 객체를 지도상에 띄우는 것이 굉장히 번거롭다.

또한 VAMAP의 차별성이라고 할 수 있는 시각화 뿐만이 아닌 각 객체가 가지는 가치를 계산하는 기능은 QGIS에 있지 않다.

### 4.1.1.2 Advantage

GUI가 구현된 프로그램에서 파일을 추가하는 과정이 간편하며 데이터베이스도 추가 가능

오픈소스이기에 다른 유저들이 만든 플러그인을 사용하는 것이 가능

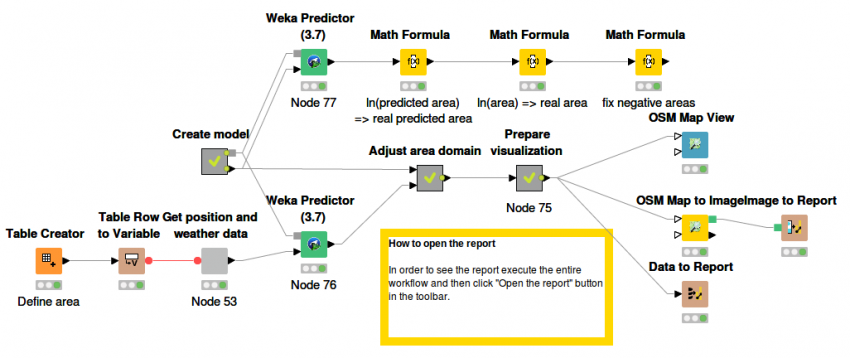
### 4.1.1.3 Disadvantage

데이터의 객체들을 지도위에 띄우는 과정이 데이터 객체 띄우기와 지도를 띄우는 각각의 과정으로 되어있음. 또한 이를 한꺼번에 보려면 동기화하는 과정도 필요하여 VAMAP과 비교하여 굉장히 번거로움.

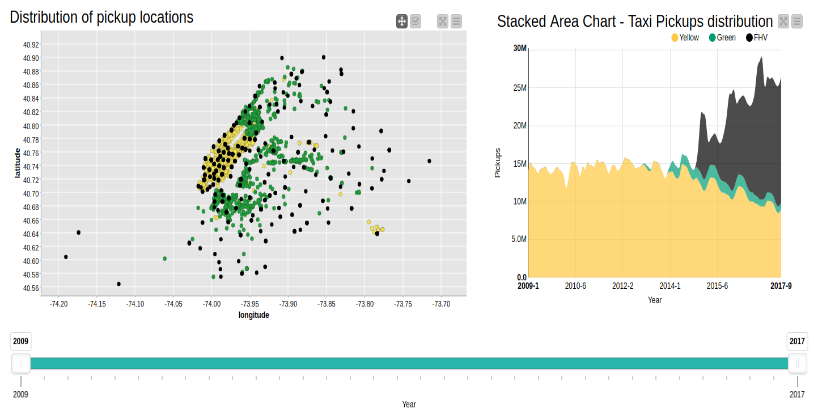
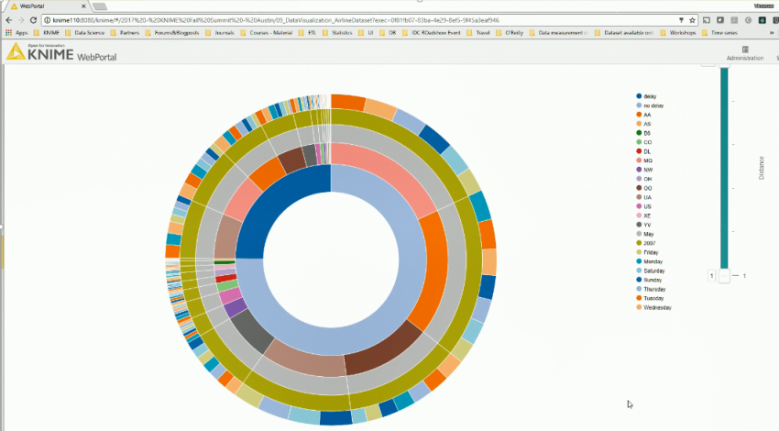
또한 여러가지 다른 기능도 제공하지만 따로 공부를 하거나 알아보지 않으면 사용하거나 활용하기가 너무 어렵다.

### 4.1.2 KNIME

### 4.1.2.1 Feature

.

기본적으로 workflow 형태로 node와 edge를 포함하여 프로젝트를 구성한다.



Csv 파일과 데이터베이스 테이블, .arff 파일 등 여러가지 파일이 load 가능하다.

시각화에 특화된 프로그램 답게 다양한 시각화 방식을 제공한다.

여기에는 맵 위에서 특정 데이터에 대한 분포 등을 시각화해주는 것도 존재하는데,

이 데이터들 간의 상관관계나 특정 데이터의 가치 등을 계산하는 방식은 존재하지않는다.

### 4.1.2.2 Advantage

다양한 파일을 load 할 수 있고 다양한 시각화 방식을 확인해볼 수 있다

### 4.1.2.3 Disadvantage

역시나 숙련된 이용자가 아니라면 사용하기가 매우 어렵다. 다양한 시각화 기능을 제공하지만 데이터들의 분포를 지도상에서 확인하고 각각의 가치를 확인하고자 하는 목적이라면 크게 메리트가 없다. 또한 실제로 시각화를 할 때 시간이 많이 소요된다.

### 4.1.3 Difference

QGIS, KNIME 은 기본적으로 사용자에게 일정 수준 이상의 지식을 요구한다. 예를 들어 파일을 load 하거나 지도를 불러오는 것, 지도위에 데이터를 띄우는 것이나 시각화를 하는 모든 과정이 처음 사용하는 사용자에게는 매우 어려운 일이다.

반면 VAMAP은 먼저 Setup.txt에서 요구하는 대로 파일의 형태를 사용자가 수정하기만 하고 터미널에서 오직 ‘python main.py’ 만 실행하면 파일 내의 내용이 Setup.txt에서 요구하는 형식인지 확인하고 엑셀파일의 주소정보를 자동으로 좌표정보로 바꿔주며 지도상에 객체의 위치와 기본데이터들의 마커가 찍힌 html 파일이 자동으로 생성된다.

또한 두 프로그램에는 존재하지않는 유저데이터가 어떤 위치에서 얼마의 가치를 가지는지를 제공한다. 따라서 내가 가지고있는 데이터들이 지도상에 어디에 위치하며 그 위치에서 어떤 가치를 가지는지를 알고 싶은 저들에게는 VAMAP이 최선의 선택이 될 것이다.

## 4.2 Demo

# Github repository

## 5.1. commit, contributor, issue/pull request

1. You drag the article you want to summarize
2. if you do right-click and then click our icon of context menu, you can show that the article you dragged is stored in extension program.
3. Click ‘summary’ button and then you will show summarized article.

## 5.2. Other issue, pull request

1. You drag the article you want to summarize
2. if you do right-click and then click our icon of context menu, you can show that the article you dragged is stored in extension program.
3. Click ‘summary’ button and then you will show summarized article.

## 5.3. etcs

1. You drag the article you want to summarize
2. if you do right-click and then click our icon of context menu, you can show that the article you dragged is stored in extension program.
3. Click ‘summary’ button and then you will show summarized article.