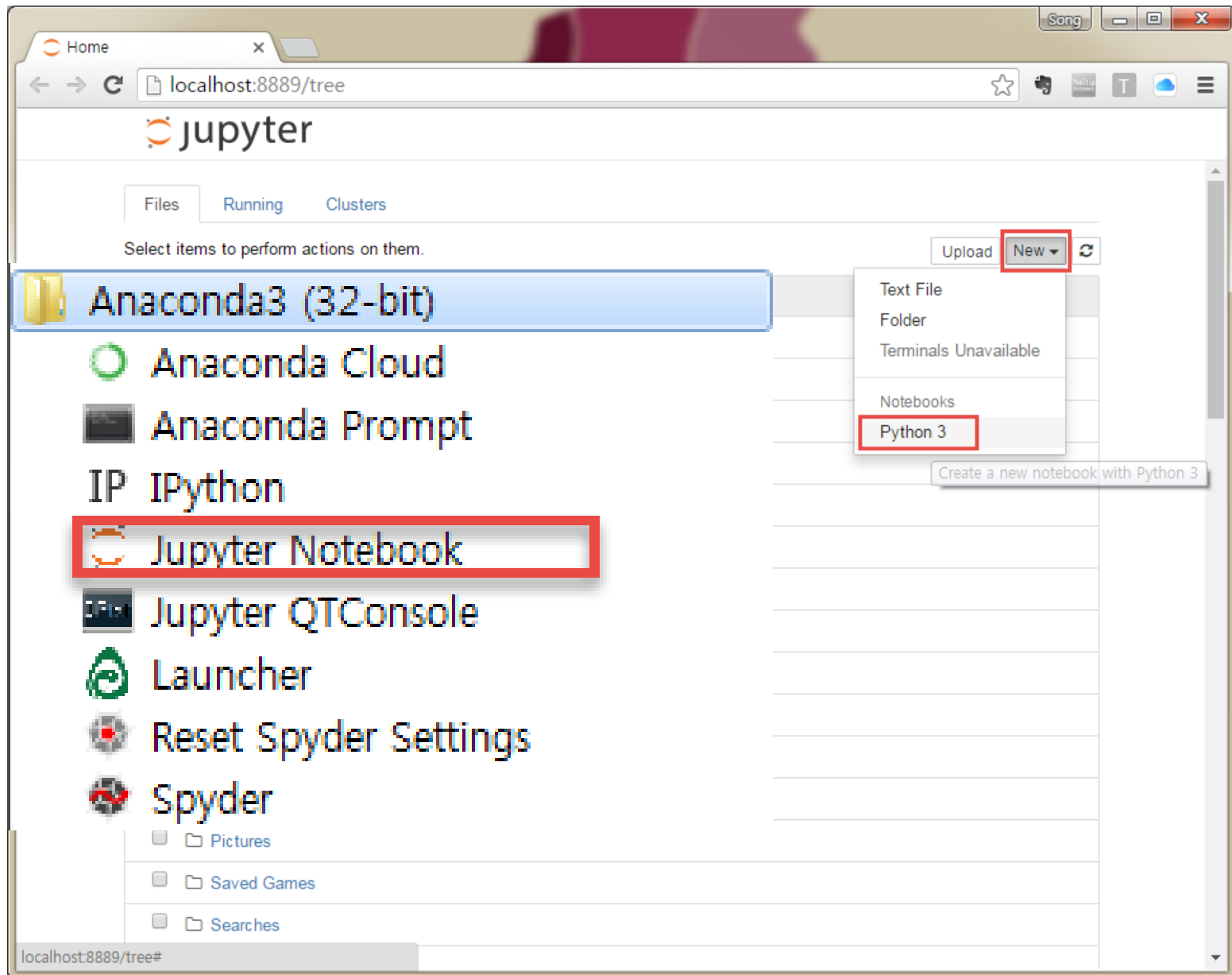


[illegible]

Jupyter Notebook 실행하기



A word cloud on a dark gray background. The central text is '공공데이터?' (Public Data?) in large white font. Surrounding it are various terms in different shades of brown and gold, including: Programming, Information, Learning, Systems, Data, Analysis, Applications, Reasoning, Verification, Games, Search, Neural, Process, Control, Machine, Networks, Modelling, Problems, Complexity, Resource, Stochastic, Web, Network, Constraint, Methods, Spaces, Computational, Software, Algorithms, Engineering, Optimisation, Foundations, Embedded, Logic, Flow, Models, Inference, Design, Theory, Multicore, Refinement, Support, Constraints, Efficient, Checking, and Model.

공공데이터의 제공 및 이용 활성화에 관한 법률

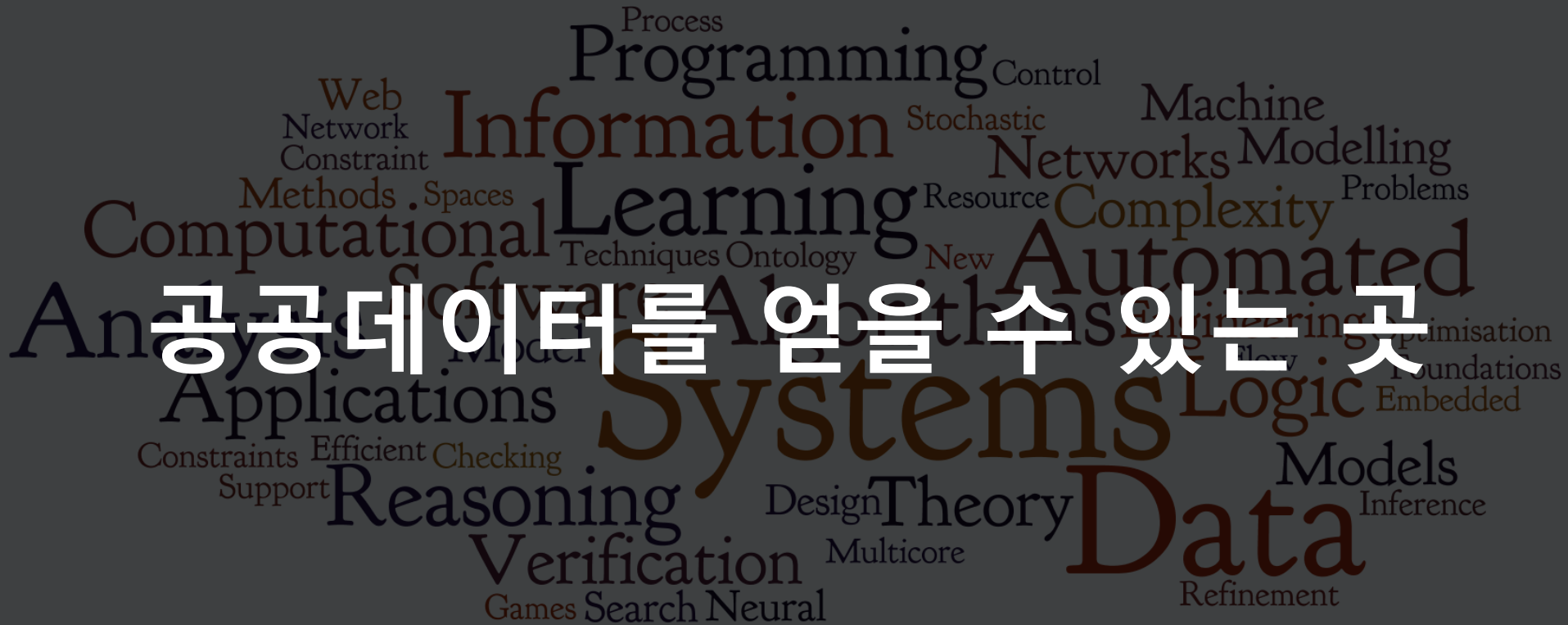
제1조 목적

이 법은 공공기관이 보유·관리하는 데이터의 제공 및 그 이용 활성화에 관한 사항을 규정함으로써 국민의 공공데이터에 대한 이용권을 보장하고, 공공데이터의 민간 활용을 통한 삶의 질 향상과 국민경제 발전에 이바지함을 목적으로 한다.

공공데이터의 제공 및 이용 활성화에 관한 법률

제2조 정의

“공공데이터”란 데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리된 자료 또는 정보를 말한다.





FILE DATA



OPEN API



STANDARD DATA



교육



국토관리



공공행정



재정금융



산업고용



사회복지



식품건강



문화관광



보건의료



재난안전



교통물류



환경기상



과학기술



농축수산



통일외교안보



법률

국가 중점개방 데이터

국민의 손으로 직접 선정한 "국가 중점개방 데이터"
36대 분야를 대용량 데이터로 개방합니다.

공공데이터 활용사례



우리문화지킴이 (우문지)

대한민국 전통우리문화 인터넷 포털사이트! 우리의
전통문화를 계승, 발전시키고자 인터넷 포털사이...공공데이터
시각화서비스 오픈



열린데이터를 검색하세요



일반행정



문화관광



환경



보건



산업경제



도시관리



복지



교통



안전

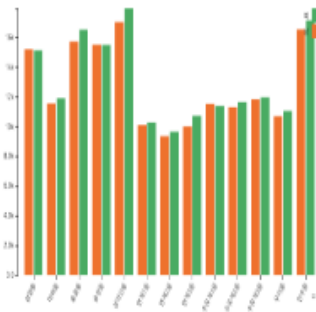


교육

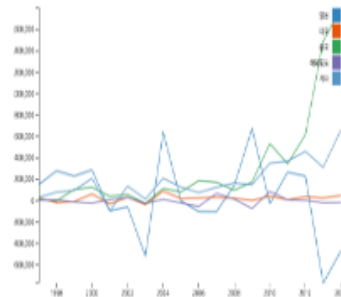
인기검색어

- 1 지하철 ↑ 1
- 2 동별 ↓ 1
- 3 인구 ↑ 2
- 4 shape ↑ new
- 5 구별 ↓ 2
- 6 청소년 ↑ new

시각화 서비스



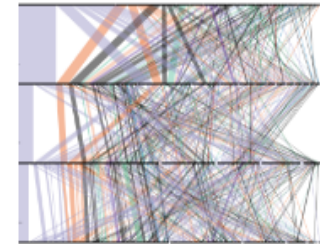
인구



국적별 외국인 관광객 변화



서울시 자치구별 장난감...



한국사학사 연구 사례2

활용갤러리



검색어로 알아보는 대한민국

검색어로 오늘을, 그리고 트렌드를 확인할 수 있습니다.



TV오락



책



영화



공연



자동차



게임



싱글남



싱글녀



주부



대학생



청소년

싱글녀

- 1 김창렬
- 2 하현우
- 3 김세아
- 4 김현수
- 5 또 오해영
- 6 현대상선
- 7 청와대 인사
- 8 신안군청
- 9 예지원
- 10 나인뮤지스

2016.06.08 13:00기준 ?

주부

- 1 또 오해영
- 2 민원24 가족관계증명
- 3 백희가 돌아왔다
- 4 꽃놀이패 정국
- 5 대구은행
- 6 복지로
- 7 의학용어사전
- 8 정기예금 이자 높은 은행
- 9 한국상담심리학회
- 10 임신 초기증상 나타나는...

2016.06.08 13:00기준 ?

대학생

- 1 오마이걸
- 2 페이스북
- 3 라붐
- 4 독립기념관 휴가
- 5 리그오브레전드 인벤
- 6 양성우
- 7 아이폰7 스펙
- 8 던파
- 9 음악대장
- 10 오버워치

2016.06.08 13:00기준 ?

청소년

- 1 리그오브레전드
- 2 도움이 필요한 이웃의 ...
- 3 아이오아이
- 4 최근에 일어난 지진과 ...
- 5 플래시 게임
- 6 꽃놀이패 조세호
- 7 나누고 싶은 따뜻한 이...
- 8 정보표현단위 바이트
- 9 최근에 일어난 화산폭발
- 10 백현 브이앱

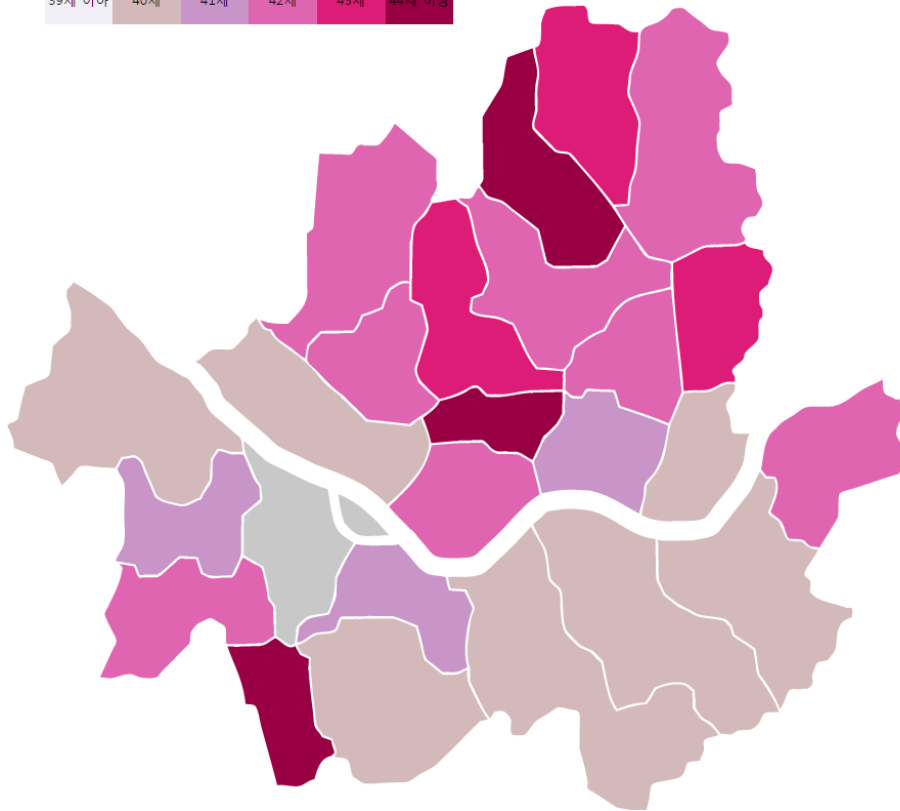
2016.06.08 13:00기준

오늘의 프로젝트 순서

[구상 단계]

0. 프로젝트 구상하기

:::: 서울시 구별 중위연령 분포도 ::::



〈중위연령〉

총인구를 연령순으로 나열할 때
중앙에 있게 되는 사람의 연령

오늘의 프로젝트 순서

[준비 단계]

1. 서울시 지도 가져오기

- <https://commons.wikimedia.org>

2. 공공데이터 찾아보기

- <http://data.seoul.go.kr/>

3. 데이터 다듬기

- 엑셀 파일 -> csv 파일로 변환 -> 손질하기

오늘의 프로젝트 순서

[프로그래밍 단계]

4. 관련 라이브러리 import하기 : csv 관련, svg 관련
5. 필요한 변수 세팅하기
6. 파일 불러오기
7. * 데이터 처리하기
8. * 데이터 시각화를 위한 코딩하기

[시각화 단계]

9. 결과물 확인하기
10. 결과물 다듬기

4. 오늘 사용되는 라이브러리 소개

```
import csv
```

```
# csv 파일을 파이썬으로 읽어오기 위한 라이브러리
```

```
from bs4 import BeautifulSoup
```

```
# svg 파일(지도)을 가공하기 위한 라이브러리
```

```
# svg 파일 중 필요한 부분을 읽어와서 수정해줄 예정
```

5. 파일 불러오기

```
svg = open('Seoul_districts.svg','r').read()  
# Seoul_districts.svg 파일을 svg 변수에 저장
```

```
data = csv.reader(open('age.csv','r'),delimiter=",")  
# age.csv 파일을 ','로 구분하여 data 변수에 저장
```

[실행결과를 확인해봅시다]

```
for row in data :  
    print(row)
```

6. 필요한 변수 세팅하기

```
age = {}
```

```
# Dictionary : { '키' : '값' }
```

```
result = []
```

```
# 리스트이름.append(내용) : 내용을 리스트에 추가
```

[실행결과를 확인해봅시다]

```
In [50]: age['Seodaemun-gu'] = 42.2  
age
```

```
Out [50]: {'Seodaemun-gu': 42.2}
```

```
In [51]: result.append(age)  
result
```

```
Out [51]: [{'Seodaemun-gu': 42.2}]
```


7. 데이터 처리하기

1) data에서 한 줄씩 읽어오기

`for row in data :`

2) 첫 번째 항목(구 이름) district 변수에 저장하기

`district = row[0]`

3) 세 번째 항목(연령 값, 실수형) 저장하기

`count = float(row[3])`

4) 구별 연령 값 age dictionary에 저장

`age[district] = count`

5) result 리스트에 age 추가하기

`result.append(age)`

8. 데이터 시각화를 위한 코딩하기

```
soup = BeautifulSoup(svg)
```

```
# svg 파일을 변환하기 BeautifulSoup 객체로 만들기
```

```
paths = soup.findAll('path')
```

```
# soup 객체에서 'path' 부분을 뽑은 paths 만들기
```

```
# path 부분은 각 구 이름이 입력된 부분
```

[실행결과를 확인해봅시다]

```
In [61]: soup = BeautifulSoup(svg)
paths = soup.findAll('path')
#svg에서 path부분을 모아서 paths 만들기
print(paths)
```

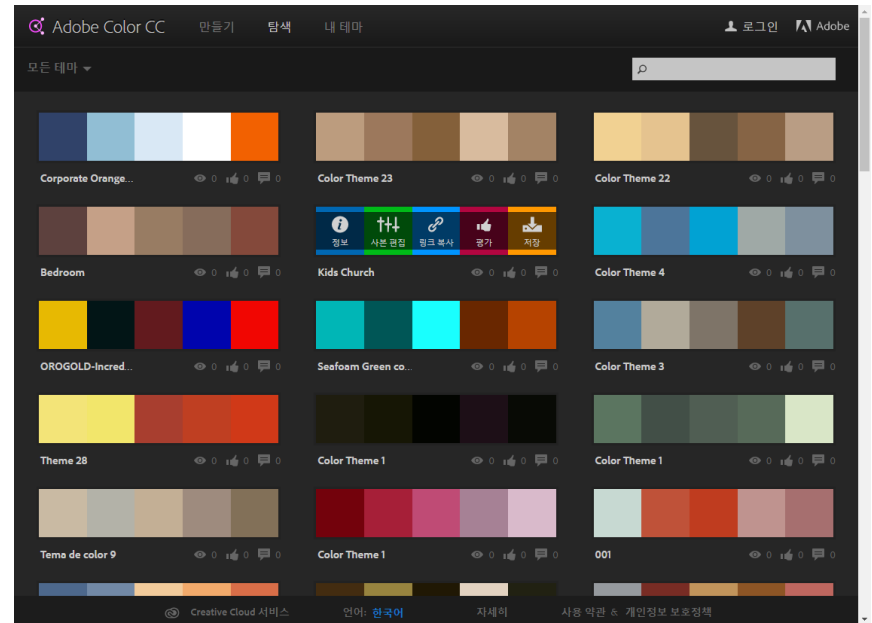
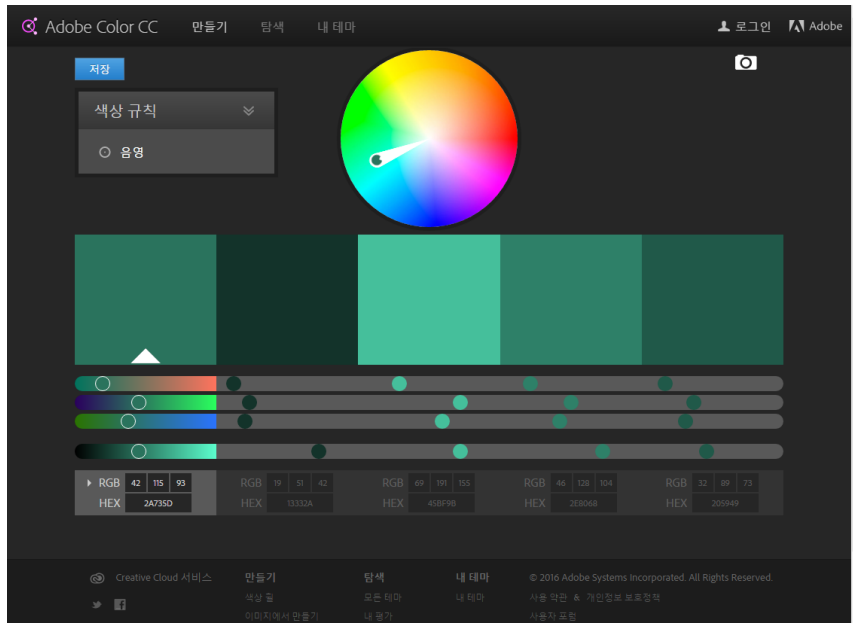
```
c1.134-3.341,0.625-8.999,1-13.503c2.579-30.969,5.016-65.498,8.002-94.521c11.516,0.432,24.255,0.549,34.508,1
c3.694,0.163,8,1.385,11.503,0c7.246-2.866,6.368-17.586,19.004-13.503c0.693,1.141,3.481,0.187,5.001,0.5
C569.141,330.705,572.308,330.705,575.476,330.705z" fill="#C8C8C8" fill-rule="evenodd" id="Eunpyeong-gu"></path>, <path clip-rule
="evenodd" d="M809.029,188.173
```

8. 데이터 시각화를 위한 코딩하기

```
colors = ["#F1EEF6", "#D4B9BA", "#C994C7", "#DF65B0",  
"#DD1C77", "#980043"]
```

결과물의 색깔 정하기

<https://color.adobe.com>



8. 데이터 시각화를 위한 코딩하기

```
path_style = 'fill:'
```

```
# svg 파일에 색을 입히기 위한 준비 작업
```

```
# path_style + color => fill: #F1EEF6
```

8. 데이터 시각화를 위한 코딩하기

```
for p in paths:
    if p['id']:
        count = age[p['id']]
        if count >= 43 :
            color_class = 5
        elif count >= 42:
            color_class = 4
        else :
            color_class = 0
    color = colors[color_class] # colors리스트에서 색깔 저장
    p['style'] = path_style + color # style 속성에 fill:#F1EEF6
```

8. 데이터 시각화를 위한 코딩하기

```
print(soup.prettify())
```

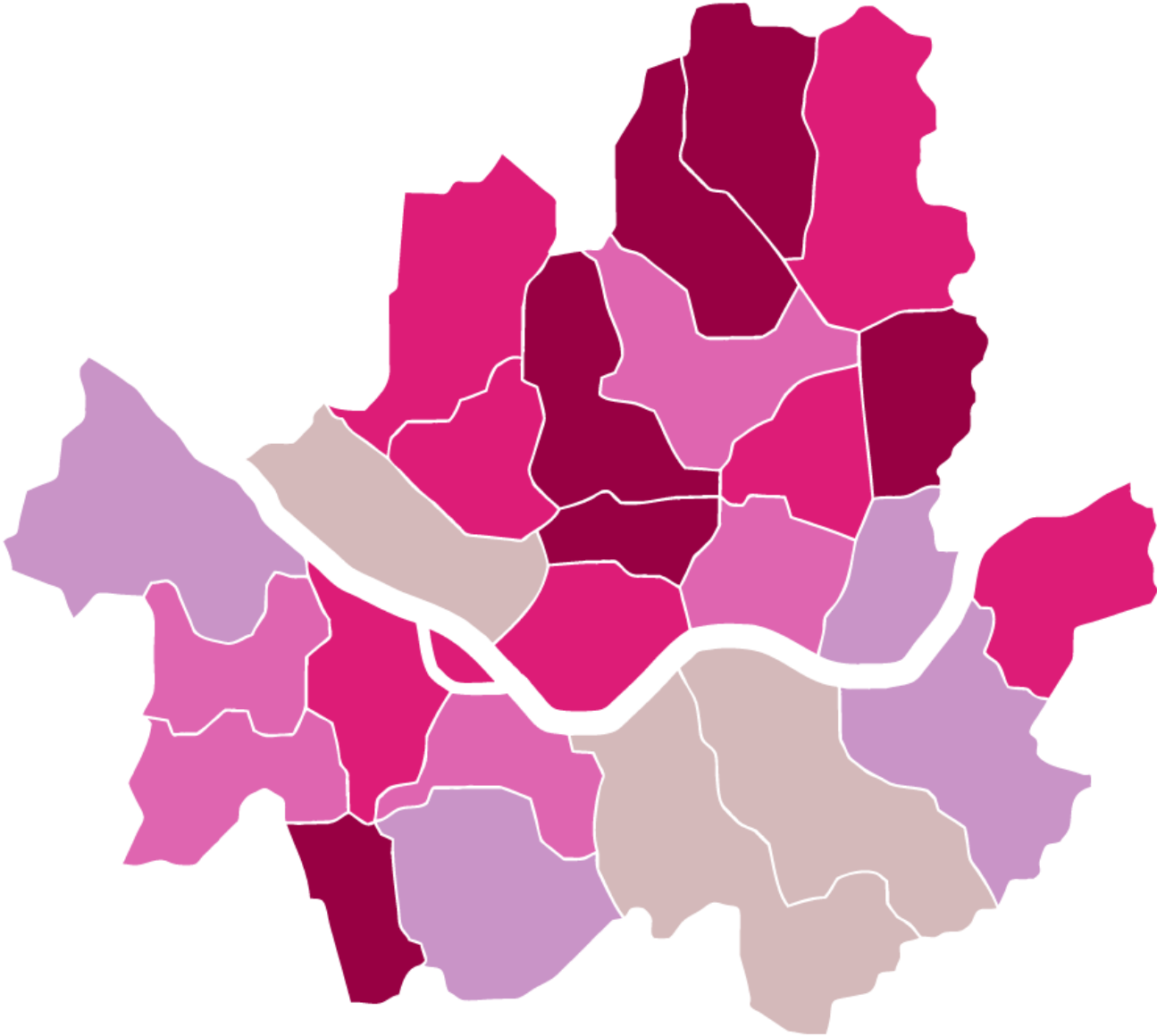
```
# 수정된 svg 파일 출력
```

```
# 복붙해서 “result.html”로 저장하기
```

```
print(soup.prettify())
```

```
<?xml version="1.0" encoding="utf-8"?>
<!-- Generator: Adobe Illustrator 15.1.0, SVG Export Plug-In . SVG Version: 6.00 Build 0) -->
<!DOCTYPE svg PUBLIC "-//W3C//DTD SVG 1.1//EN" "http://www.w3.org/Graphics/SVG/1.1/DTD/svg11.dtd">
<html>
  <body>
    <svg enable-background="new 0 0 1400 1400" height="1400px" id="Layer_1" version="1.1" viewBox="0 0 1400 1400" width="1400px" x="0px" xml:space="preserve" xmlns="http://www.w3.org/2000/svg" xmlns:xlink="http://www.w3.org/1999/xlink" y="0px">
      <path clip-rule="evenodd" d="M964.064,164.667
        c-1.447,9.018-0.285,18.105-2.002,27.506c-2.068,11.332-9.018,22.101-11.502,33.507c-0.867,3.979-0.977,
        9.201-1.5,14.003
        c-0.508,4.656-1.969,10.129-1.5,14.003c0.779,6.456,5.756,14.04,8.502,21.005c2.287,5.805,7.385,15.948,
        8.002,21.505
        c0.539,4.856-0.953,11.628-1.502,17.504c-0.547,5.879-1.484,11.904-2,17.504c-1.135,12.346-0.799,25.368-
        3,36.008
        c-1.582,7.641-5.57,14.402-7.002,21.505c-1.725,8.558-1.271,18.438-3,27.506c-6.813,0.022-13.387,0.283-1
        9.506,1
        c-14.793-19.111-31.705-39.509-48.51-58.013c-4.902-5.398-11.217-16.078-17.504-18.504c-7.016-2.707-17.6
        23,0.042-22.006-7.001
        c-4.459,4.076,0.127,0.544,14.003,14.003c-0.149,1.02,1.254,0.00,1.502,2.2,2.059,4.976,0.127,0.544,12.50
```

9. HTML 출력 결과 확인하기 : result.html



10. 결과물 다듬기

<http://colorscripter.com/s/j7TbTAz>

```
<style>
span{
  position: absolute; top : 100px;
  display:inline-block; width:100px; height:50px;
  align : center; font-size:20;
  vertical-align:middle; padding:10px 0 0 0px; text-align:center;
}
.color1 {
  background-color: #F1EEF6; left : 100px;
}
.color2 {
  background-color: #D4B9BA; left : 200px;
}
.color3 {
  background-color: #C994C7; left : 300px;
}
.color4 {
  background-color: #DF65B0; left : 400px;
}
.color5 {
  background-color: #DD1C77; left : 500px;
}
.color6 {
  background-color: #980043; left : 600px;
}
</style>
<div style="position:absolute;left:100;top:50;font-size:30">::: 서울시 구별 중위연령 분포도 :::</div>
<span class="color1">39세 이하</span> <span class="color2">40세</span>
<span class="color3">41세</span> <span class="color4">42세</span>
<span class="color5">43세</span> <span class="color6">44세 이상</span>
```


개인 연습 프로젝트

0. 구상하기

1. 데이터 찾기

2. 코드 수정하기

3. 발표 자료 만들기

- 1) 왜 이 프로젝트를 하게 되었나?
- 2) 프로젝트 과정 소개
- 3) 프로젝트 결과 소개
- 4) 프로젝트를 통해 배우고 느낀 점

참고자료

참고강좌

- <http://visualize.tistory.com/47>

서울시 중위연령 데이터

- <http://data.seoul.go.kr>

서울시 지도 SVG 파일

- <https://commons.wikimedia.org>