

빅데이터 분석 결과 시각화

데이터 시각화 정의

학습내용

- 데이터 시각화 정의
- 데이터 시각화 환경 설정
- 데이터 시각화 자료 수집

학습목표

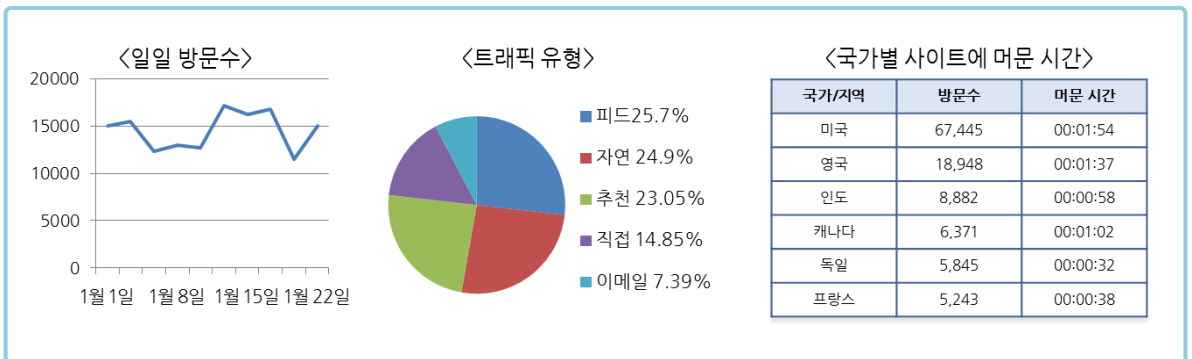
- 데이터 시각화에 대한 정의를 설명할 수 있다.
- 데이터 시각화를 위한 웹서버 설치를 통해 시각화를 도출할 수 있다.
- 데이터 시각화를 위한 자료를 수집하는 방법을 알 수 있으며, 활용 분야에 대해 설명할 수 있다.

● 데이터 시각화 정의

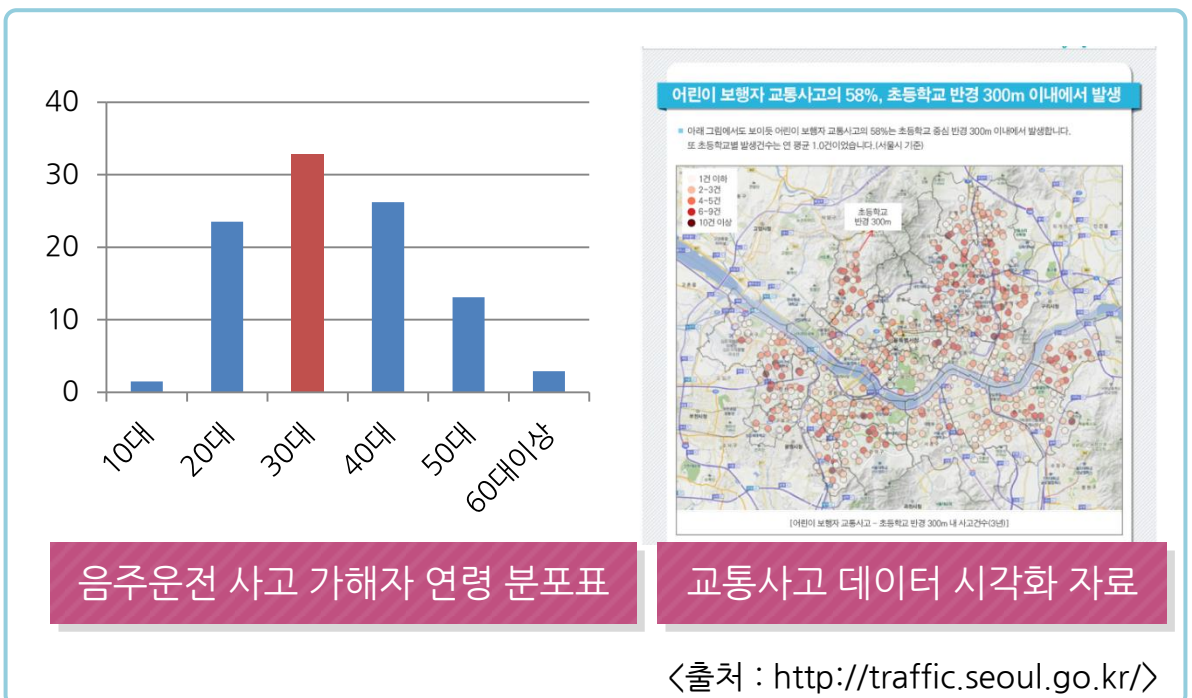
1. 데이터 시각화 개념

◆ 데이터 시각화란?

- 지식 정보 사회에 필요한 데이터 분석 결과를 누구나 쉽게 이해할 수 있도록 의미 있는 형태의 이미지나 그래프로 표현하는 것



- 점점 방대해지고 의미의 연관성이 높아지는 정보에 대해 복잡함을 줄여주고 패턴과 상관관계를 추론할 수 있음

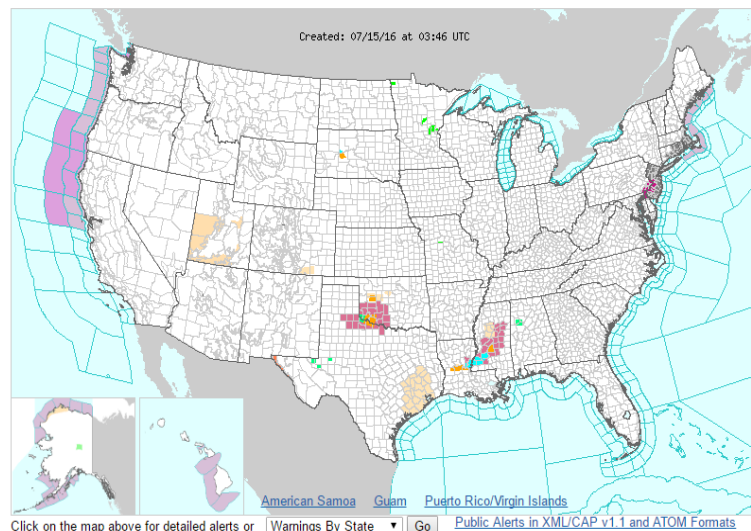


● 데이터 시각화 정의

1. 데이터 시각화 개념

◆ 데이터 시각화 활용 예시

- 미국 국립해양대기청(NOAA) - 전국기상서비스(National Weather Service)
 - 6시간 간격으로 미국 전역의
홍수, 가뭄, 허리케인, 뇌우, 화재 등 24가지의 경보를 제공함
 - 정확한 날씨 예측을 위해서는
각 지역, 배, 비행기 등을 통해 관찰된
각종 기후정보와 이미지, 수치 등
엄청난 양의 데이터를 저장하고 분석해야 함



미국 국립해양대기청 홈페이지

<출처 : <http://www.weather.gov>, 2016.07>

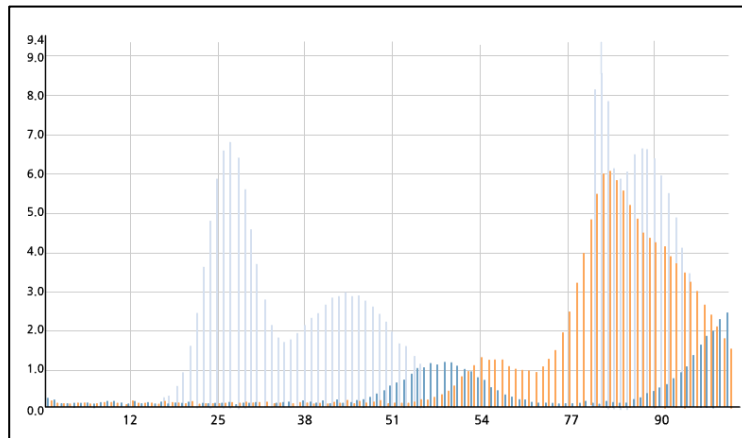
● 데이터 시각화 정의

1. 데이터 시각화 개념

◆ 데이터 시각화 활용 예시

● 데이터 시각화 종류

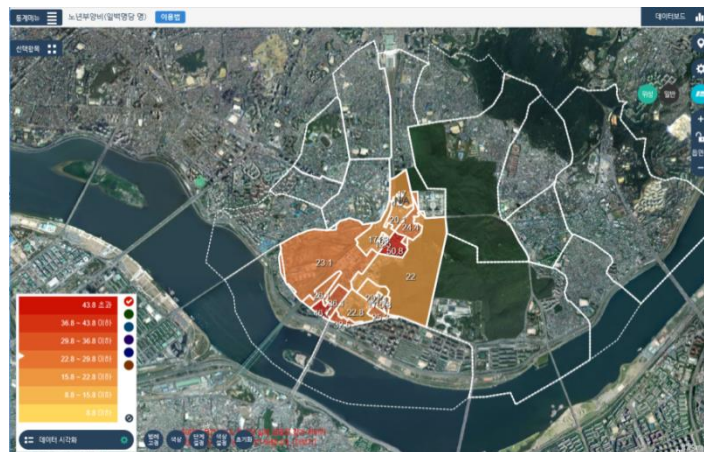
- 정보디자인 : 보통 대규모 데이터를 색채, 통계(도표, 그래프 등), 이미지 등을 활용해 요약적으로 표현하는 것



그래프형 정보디자인

<출처 : <http://nvd3.org/>, 2016.07>

- 과학적 시각화 : 실험결과나 시뮬레이션 데이터 등 복잡한 데이터를 쉽게 탐색할 수 있도록 3차원 그래픽 기술 등을 활용하여 시각화하는 기술



3차원 그래픽 기술

<출처 : <http://sgis.kostat.go.kr/>, 2016.07>

● 데이터 시각화 정의

1. 데이터 시각화 개념

◆ 데이터 시각화 활용 예시

● 데이터 시각화 종류

- 인포그래픽 : 인포메이션과 그래픽의 합성어로, 복잡한 수치나 글로 표현되어 있는 다량의 정보를 차트, 지도, 다이어그램, 로고, 일러스트레이션 등을 활용하여 한눈에 파악할 수 있도록 하는 디자인



● 인포그래픽의 유형

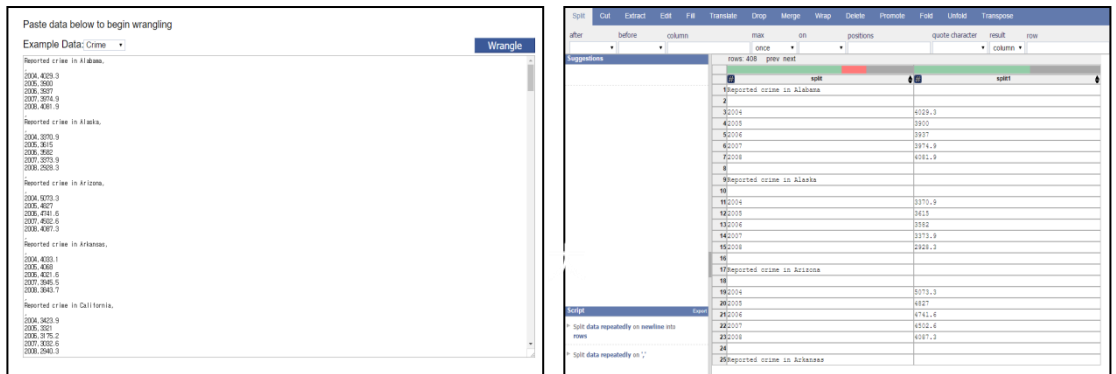
- 통계 기반 인포그래픽 : 표나 그래프를 이용하여 만들어진 유형
예) 미세먼지, 흡연율
- 타임라인 기반 인포그래픽 : 시간순서대로 정보를 보여주는 형태
예) 한국 역사 분석, 기업 제품 분석
- 프로세스 기반 인포그래픽 : 복잡한 업무 처리 과정을 한 장으로 보여줄 때 효과적
예) 심리테스트, 조립설명서
- 위치 / 지리 기반 인포그래픽 : 정보를 지도 위에 표현
예) 질병감지, 성씨분포
- 비교 형식 기반 인포그래픽 : 서로 상반되는 데이터 대조 및 비교 시 유용
예) 남녀 시각차, 정치 성향 분석
- 스토리 기반 인포그래픽 : 이야기 기반으로 진행되는 시각적 의미 전달
예) 하루 일과표, 다이어트 식단과 운동

● 데이터 시각화 정의

2. 데이터 시각화 분석 툴 소개

◆ 데이터 시각화 분석 툴

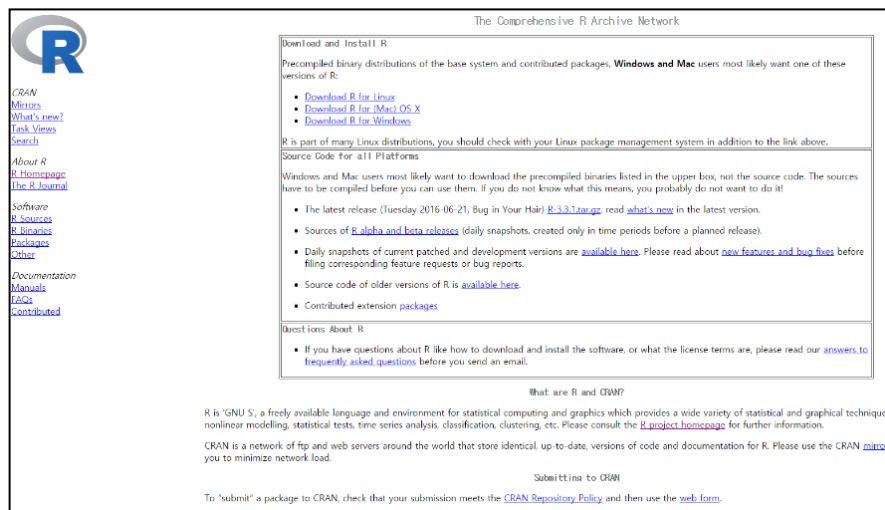
- DataWrangler
 - 웹 기반의 데이터 정제 서비스



DataWrangler

<출처 : <http://vis.stanford.edu/>, 2016.07>

- R
 - 윈도우, 리눅스, 매킨토시 통계를 활용한 분석 서비스 프로그램 제공



R 홈페이지

<출처 : <https://www.r-project.org/>, 2016.07>

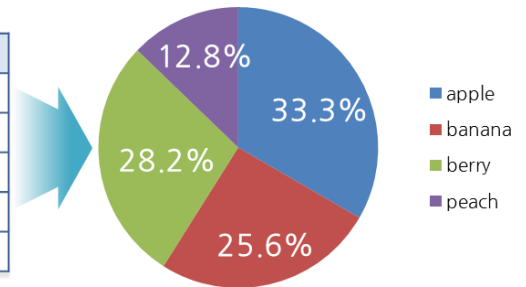
● 데이터 시각화 정의

2. 데이터 시각화 분석 툴 소개

◆ 데이터 시각화 분석 툴

- 구글 퓨전 테이블
 - 구글에 데이터를 올려 분석해 주는 웹서비스

1	year	name	qty	price
2	2019	Apple	13	13000
3	2019	Banana	10	5000
4	2019	Peach	5	2500
5	2019	berry	11	1100
...



과일 수량(qty) 데이터 시각화

<출처 : <https://fusiontables.google.com/>, 2016.07>

- D3js
 - JavaScript 라이브러리를 이용한 시각화 툴 서비스



D3 홈페이지

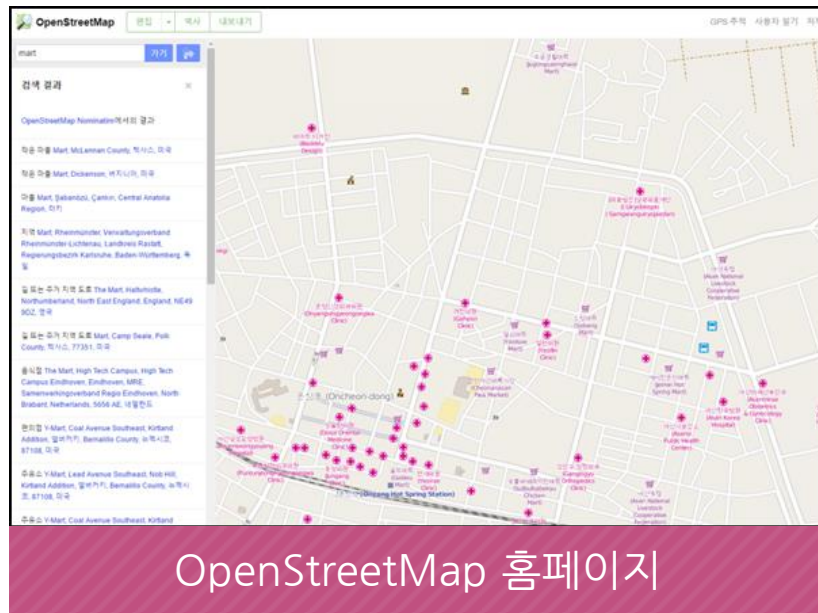
<출처 : <https://d3js.org/>, 2016.07>

● 데이터 시각화 정의

2. 데이터 시각화 분석 툴 소개

◆ 데이터 시각화 분석 툴

- OpenStreetMap
 - 검색을 통한 GIS 맵 서비스



<출처 : <http://www.openstreetmap.org/>, 2016.07>

● 데이터 시각화 환경설정

1. 데이터 시각화에 필요한 환경 설정

◆ APMSETUP 설치

- <http://dev.naver.com/projects/apmsetup/download> 으로 접속
(아파치 / MySQL / PHP 설치를 한꺼번에 할 수 있음)
- APMSETUP 7 파일 다운로드
- APMSETUP 7 설치
- htdocs폴더에 html 및 javascript, css 파일을 넣음

2. 데이터 시각화 맛보기

◆ 데이터 시각화 실행

- <http://d3js.org>에 접속 후 d3.zip 파일 다운
- 샘플 예제 테스트를 위해 See More Example 메뉴 선택
- Github 사이트에 있는 Bubblechart 예제 선택
- index.html과 flare.json 파일을 마우스로 드래그하여 각각 파일을 만들
- localhost / bubblechart.html 파일 실행

● 데이터 시각화 자료 수집

1. 국내 데이터 자료 수집

◆ 정부3.0 공공데이터포털

- 국민의 손으로 직접 선정한 국가 중점개방 데이터 36대 분야를 대용량 데이터로 개방



<출처 : <https://www.data.go.kr>, 2016.07>

◆ 경기데이터드림

- 지도 기반으로 우리 지역 관련 공공데이터를 손쉽게 찾으실 수 있고, 멀티미디어 데이터부터 도내 타기관 서비스까지 다양한 데이터를 제공함



<출처 : <http://data.gg.go.kr/>, 2016.07>

● 데이터 시각화 자료 수집

1. 국내 데이터 자료 수집

◆ TourAPI

- 국문뿐만 아니라, 영문, 일문, 중문간체, 중문번체, 독일어, 프랑스어, 스페인어, 러시아어 등 국내 유일의 다국어 관광정보를 제공함



<출처 : <http://api.visitkorea.or.kr/>, 2016.07>

◆ 농림축산식품 공공데이터 포털

- 농림축산식품부에서 제공하는 분석데이터를 그리드 형태, 엑셀, 한글, txt문서 등의 파일 다운로드 형태, 링크 형태, 오픈API, 원시데이터의 형태로 제공함



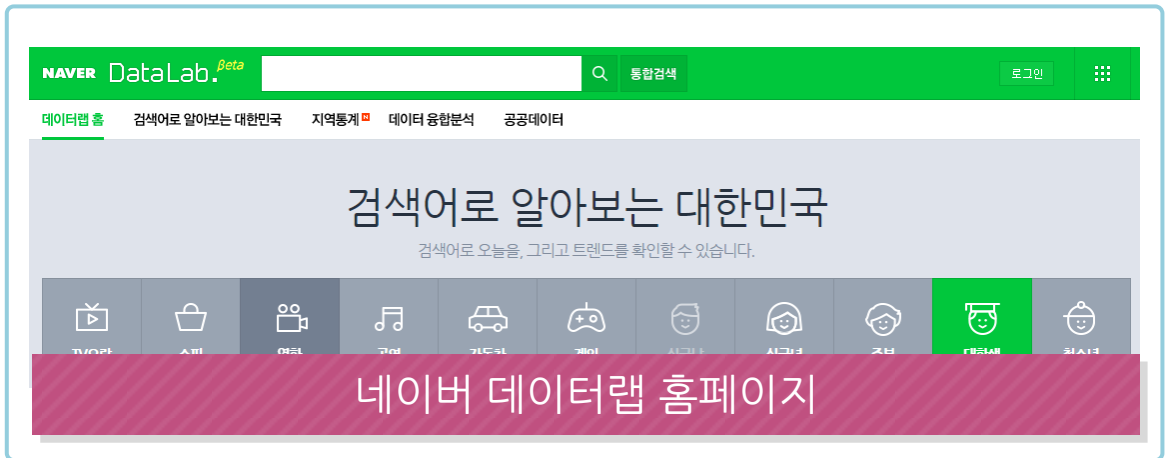
<출처 : <http://data.mafra.go.kr/>, 2016.07>

- 데이터 시각화 자료 수집

1. 국내 데이터 자료 수집

- ◆ 기타 국내 데이터 자료 수집

- 네이버 데이터랩 : 네이버와 공공데이터를 융합한 지역별 통계정보를 제공



● 데이터 시각화 자료 수집

2. 공공정보 활용 서비스

◆ 미국 공공데이터 포털

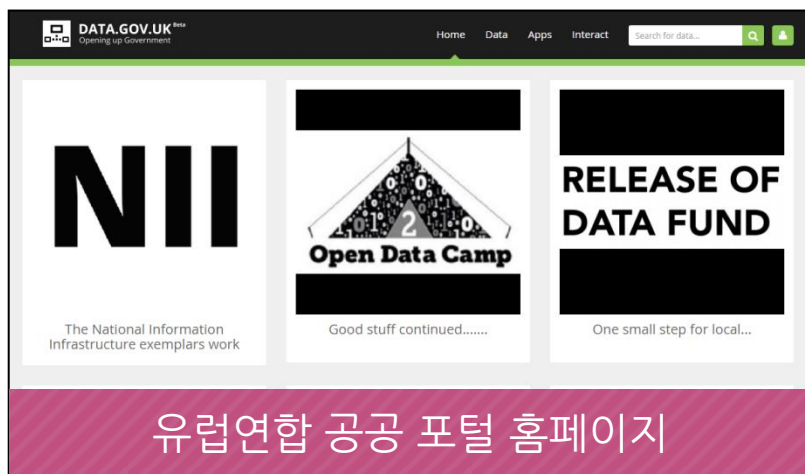
- 정부부처 및 공공기관이 참여하는 ‘Open Government Initiative’를 만들고 시민 참여와 민관 협력을 증진시키는 것이 목표



<출처 : <https://www.data.gov>, 2016.07>

◆ 유럽연합 공공 포털

- 유럽연합은 2011년 12월 경제성장 및 사회현안 해결, 과학기술 선도, 행정의 투명성과 효율성 향상 등을 목적으로 EU기구와 27개 회원국 공공기관의 모든 공공데이터의 온라인 개방을 의무화함



<출처 : <https://data.gov.uk>, 2016.07>

● 데이터 시각화 자료 수집

2. 공공정보 활용 서비스

◆ 미국 통계국

- 공공기관, 기업들이 협력하여 시각화 및 정보제공을 위해 만들어진 사이트로 Data/Data visualization 을 선택하면 데이터 시각화 자료를 볼 수 있음



<출처 : <http://www.census.gov/library/>, 2016.07>

◆ 기타 공공정보 활용 서비스

- NHN 가상박물관 서비스 : 다양한 공공기관으로부터 공공데이터를 받고, 사용자들이 이용하기 좋은 형태로 가공해 제공

핵심요약

1. 데이터 시각화 정의

■ 데이터 시각화란?

- 지식 정보 사회에 필요한 데이터 분석 결과를 누구나 쉽게 이해할 수 있도록 의미 있는 형태의 이미지나 그래프로 표현하는 것

■ 데이터 시각화의 종류

- 정보디자인
- 과학적시각화
- 인포그래픽

■ 인포그래픽의 유형

- 통계 기반 인포그래픽
- 타임라인 기반 인포그래픽
- 프로세스 기반 인포그래픽
- 위치/지리 기반 인포그래픽
- 비교 형식 기반 인포그래픽
- 스토리 기반 인포그래픽

■ 데이터 시각화 분석툴

- DataWrangler
- R
- 구글 퓨전 테이블
- D3js
- OpenStreetMap

2. 데이터 시각화 환경설정

■ 데이터 시각화에 필요한 환경 설정

- <http://dev.naver.com/projects/apmsetup/download>으로 접속
(아파치 / MySQL / PHP 설치를 한꺼번에 할 수 있음)
- APMSETUP 7 파일 다운로드
- APMSETUP 7 설치
- htdocs폴더에 html 및 javascript, css 파일을 넣음

■ 데이터 시각화 맛보기

- <http://d3js.org>에 접속 후 d3.zip 파일 다운
- 샘플 예제 테스트를 위해 See More Example 메뉴 선택
- Github 사이트에 있는 Bubblechart 예제 선택
- index.html과 flare.json 파일을 마우스로 드래그하여 각각 파일을 만들
- localhost / bubblechart.html 파일 실행

3. 데이터 시각화 자료수집

■ 국내 데이터 자료 수집

- 정부3.0 공공데이터포털
- 경기데이터드림
- 네이버 데이터랩
- TOUR API
- 농림축산식품 공공데이터 포털

■ 공공정보 활용 서비스

- NHN 가상박물관 서비스
- 미국 공공데이터 포털
- 유럽 연합 공공포털
- 미국 통계국
- 호주 통계청