

# 데이터분석 전문가 가이드

## 과목 4. 데이터 시각화 제 2장 시각화 디자인

---

출처 : 데이터분석 전문가 가이드, 한국데이터베이스진흥원

참조 : [HTTP://WWW.SLIDESHARE.NET/NEOFUTURE/SDS-N2](http://www.slideshare.net/neofuture/sds-n2)

[HTTP://WWW.SLIDESHARE.NET/NEOFUTURE/SDS-2](http://www.slideshare.net/neofuture/sds-2)

[HTTP://WWW.SLIDESHARE.NET/NEOFUTURE/4-36648309](http://www.slideshare.net/neofuture/4-36648309)

# 제 1절 시각화의 정의

---

## 1. 데이터 시각화의 중요성

- 데이터 시각화는 “매우 광범위하게 분산된 방대한 양의 자료를 분석해 한눈에 볼 수 있도록 도표나 차트 등으로 정리한 것 ” 을 말한다.
  - 빅데이터를 잘 처리해 필요한 정보를 얼마나 신속하게 얻어내는냐가 기업의 경쟁력이고
  - 빅데이터에서 통찰력을 얻으려면 시각화가 필요
  - 따라서, 시각화는 통찰력을 부여하고 경쟁력을 좌우함.
- 데이터 시각화의 2가지 목적
  - 데이터분석
  - 의사소통 : 자료로부터 얻은 스토리를 이해하고, 이를 다른 사람과 공유함으로써 그 스토리를 발전시킬 수 있음.
- 데이터 시각화의 다양한 효과
  - 시간 절감
  - 흥미 유발과 빠른 확산
  - 기억하기 쉬움( 기억효과 )

# 제 1절 시각화의 정의

## 2. 시각 이해와 시각화

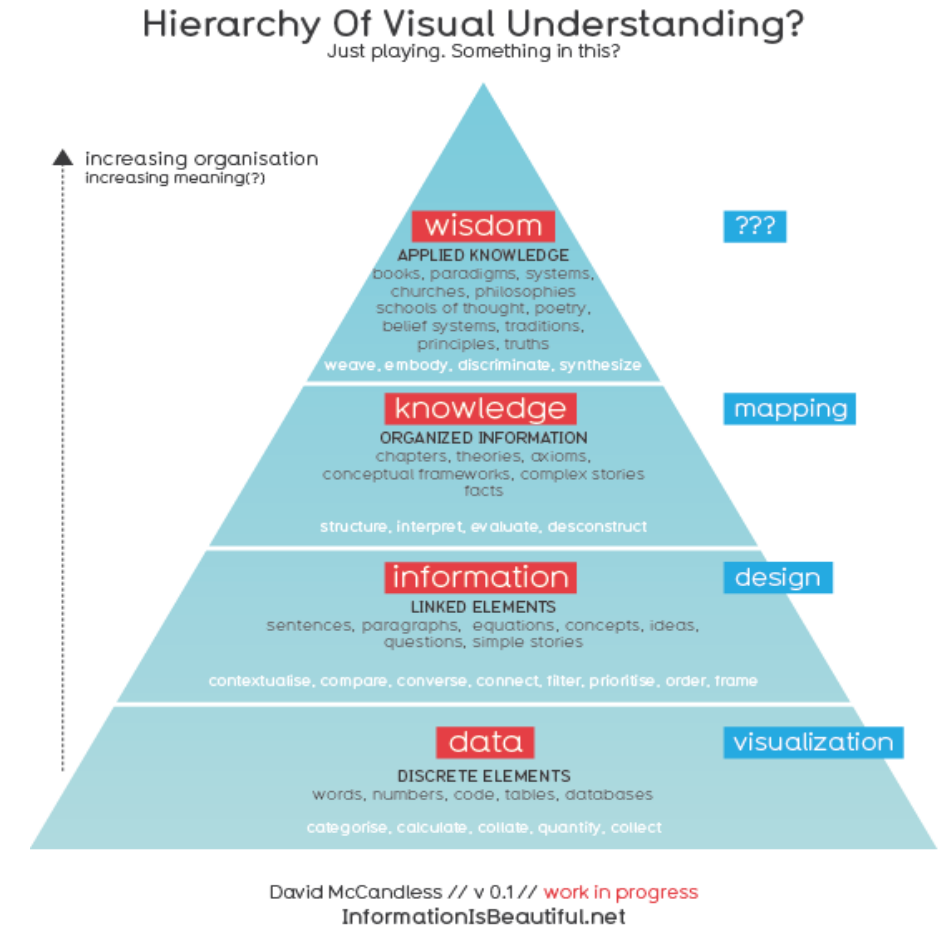
- “데이터 위계 모델”을 기초로 그린 “시각 이해의 위계”

지혜 ————— 정의되지 않은 것( ??? )

지식 ————— 매핑

정보 ————— 디자인

데이터 ————— 시각화



### 3. 시각화 분류와 구분

- 데이터의 시각적 표현의 연구영역을 말하며, 원 데이터를 효과적으로 전달하기 위함.

- [illegible]

# 제 1절 시각화의 정의

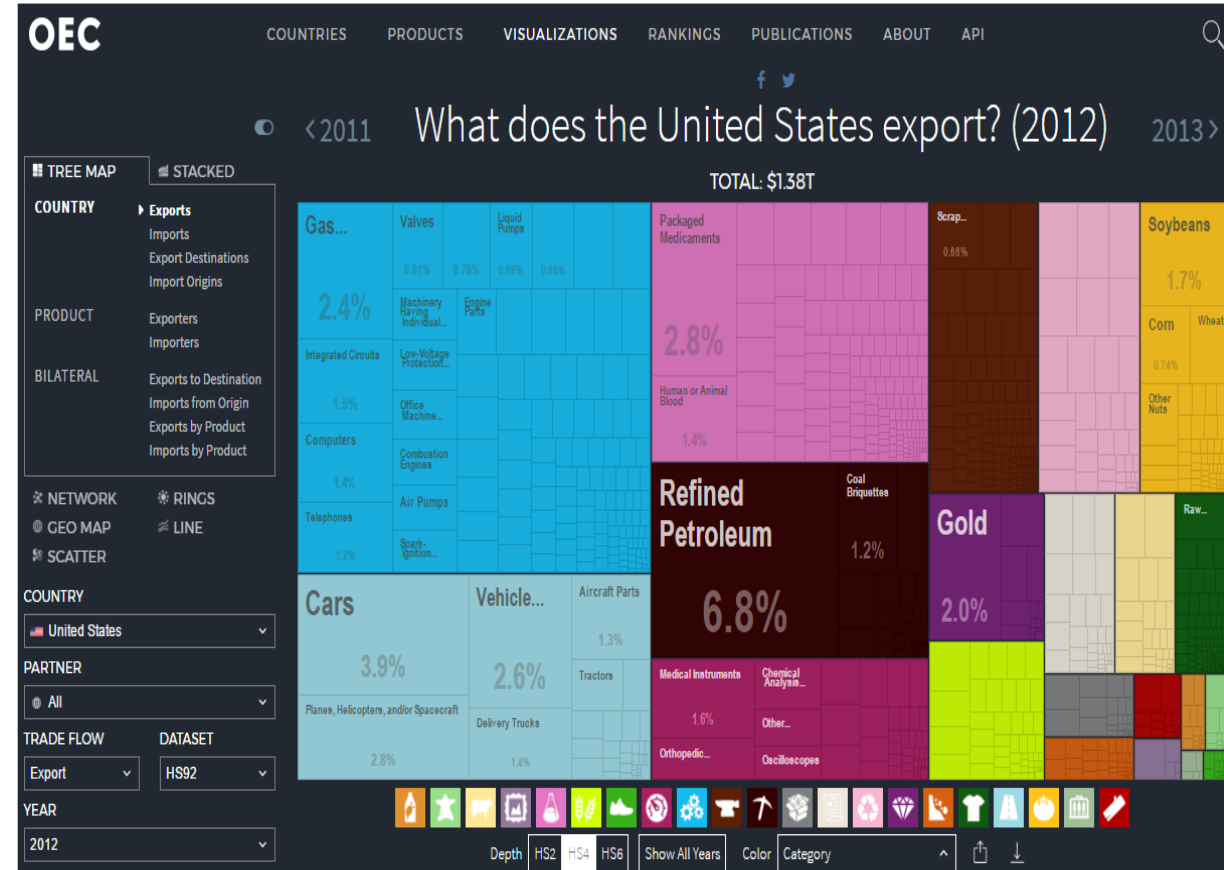
## 3. 시각화 분류와 구분

### 나. 정보 시각화

비수치 정보와 인터넷의 네트워크 관계 등  
큰 범위의 집합에 대한 시각적 표현방법

예) 분기도, 수지도, 히트맵

옆의 트리맵은 하버드-MIT 경제 복잡성의  
전망대에서 개발



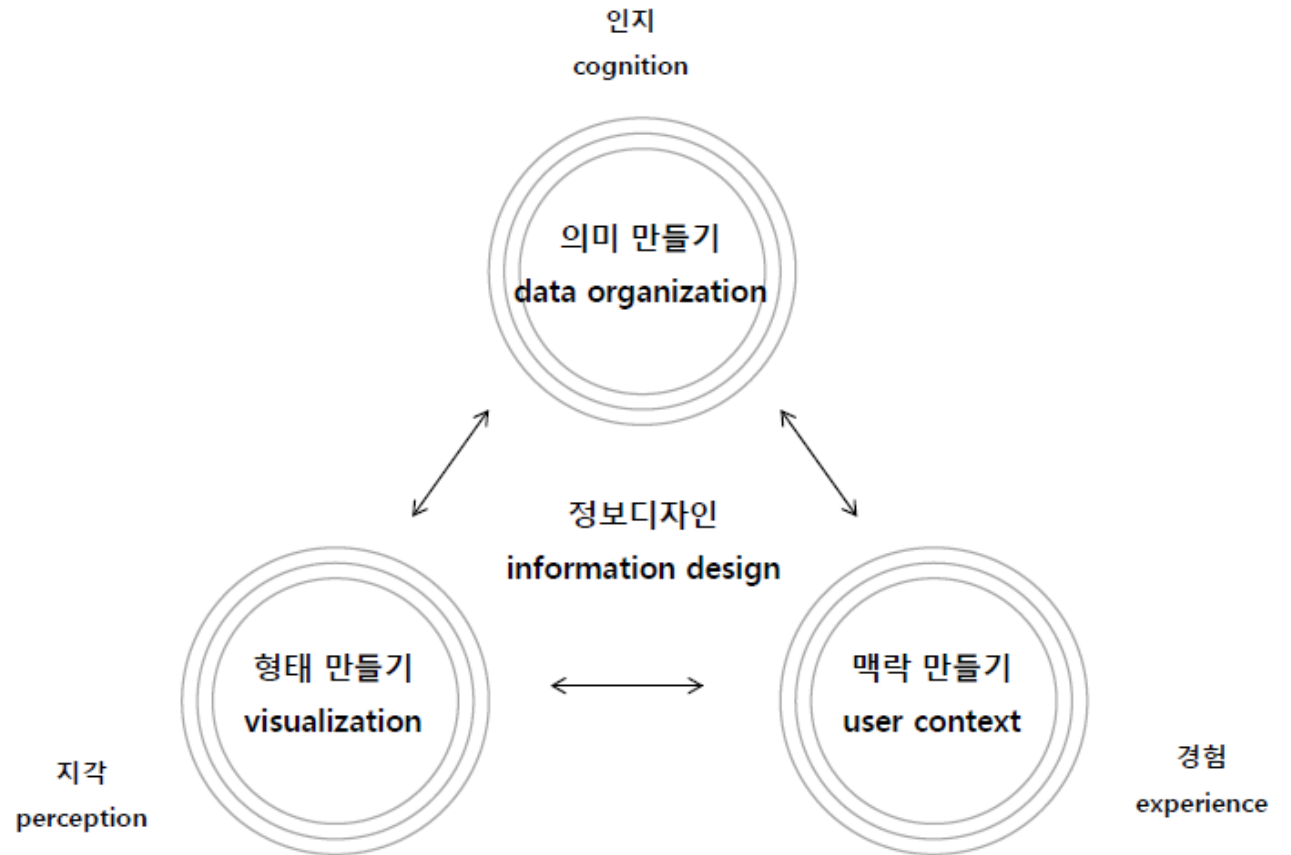
[http://atlas.media.mit.edu/en/visualize/tree\\_map/hs92/export/usa/all/show/2012/](http://atlas.media.mit.edu/en/visualize/tree_map/hs92/export/usa/all/show/2012/)

# 제 1절 시각화의 정의

## 3. 시각화 분류와 구분

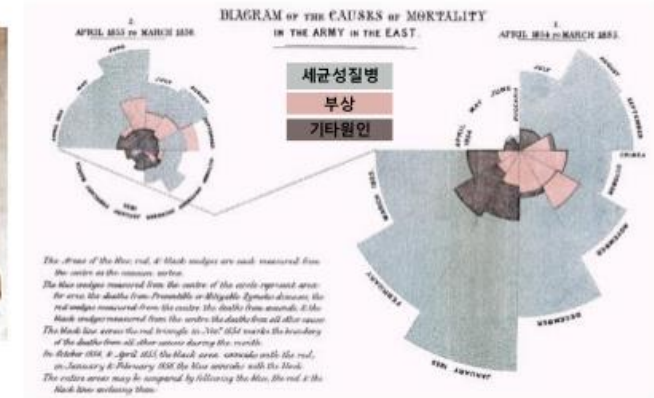
### 다. 정보 디자인

- 사람이 사용할 수 있는 효과적인 정보와 복잡하고 구조적이지 않은 기술 데이터를 시각적으로 표현하는 방법으로 좀 더 명확하게 의미를 이해할 수 있게 함
- 데이터 시각화, 정보 시각화도 정보 디자인의 범위에 속함
- 인포그래픽도 정보 디자인의 한 유형



### 3. 시각화 분류와 구분

#### 다. 정보 디자인



나이팅게일은 이러한 인사이트를 데이터를 통해 검증 로즈 다이어그램으로 당국을 설득하는 데 성공함



# 제 1절 시각화의 정의

## 3. 시각화 분류와 구분

### 라. 인포그래픽

- 정보디자인에서 목적에 따라서 2가지로 나눌 수 있음. => 정보형 메시지와 설득형 메시지
- 지하철 노선도는 실제 지형 지도를 왜곡하지만 보기 쉽도록 개념적으로 구현함.





# 제 1절 시각화의 정의

## 3. 시각화 분류와 구분

### 라. 인포그래픽

- 옆의 “사회 체제 묘사”은 설득적 메시지에 주력함.
- 사회 계층을 암시하는 옷을 가로로 재단하고 상하 계층별로 다시 연결해 한벌의 셔츠로 표현함.
- 인포그래픽은 원데이터는 취급하지 않고 실용적을 전달하기 위한 다양한 차트, 다이어그램, 일러스트 레이션을 사용.



# 제 1절 시각화의 정의

## 4. 빅데이터 시각화 영역



# 제 2절 시각화 프로세스

## 1. 정보 디자인 프로세스

가. 1단계 : 데이터 수집

나. 2단계 : 모든것을 읽기

다. 3단계 : 내러티브 찾기

- 데이터세트를 단순명료화, 프로세스 설명, 트렌트 강조
- 제공하고자 하는 정보를 담은 스토리를 만들어 내는 것이 가능한가?
- 이 주제에 관심이 가는가?
- 주목할 만한 사실 또는 가치를 말하고 있는가?



# 제 2절 시각화 프로세스

---

## 1. 정보 디자인 프로세스

### 라. 4단계 : 문제의 정의

- 이야기가 내포한 결과에 대한 현실성을 검토 필요.

### 마. 5단계 : 계층구조 만들기

- 중요한 것을 주인공으로 만들고 나머지는 보조적인 요소로 배열

### 바. 6단계 : 와이어프레임 그리기

- 데이터가 손질되고 흥미로운 사실들이 선택되어 계층구조가 결정되어 와이어프레임이 창조됨.

### 사. 7단계 : 포맷 선택하기

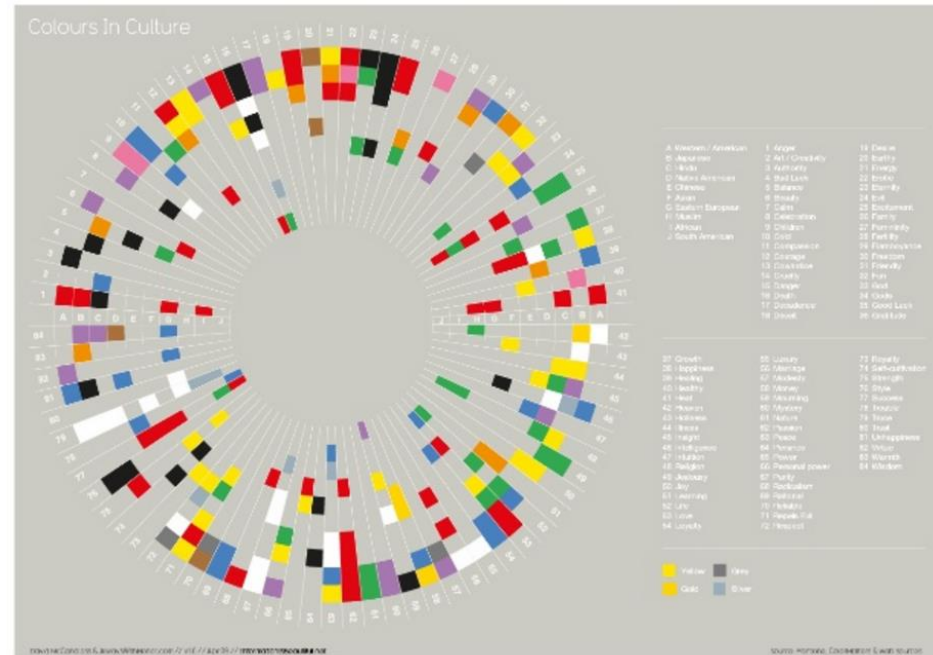
- 정보를 표현의 방법으로 전통적인 차트와 그래프 등을 선택

# 제 2절 시각화 프로세스

## 1. 정보 디자인 프로세스

### 아. 8단계: 시각 접근 방법 결정하기

- 데이터의 아름다움을 만들어내는 과정



맥캔들레스의 문화에 따른 색상

<http://www.informationisbeautiful.net/visualizations/colours-in-cultures/>

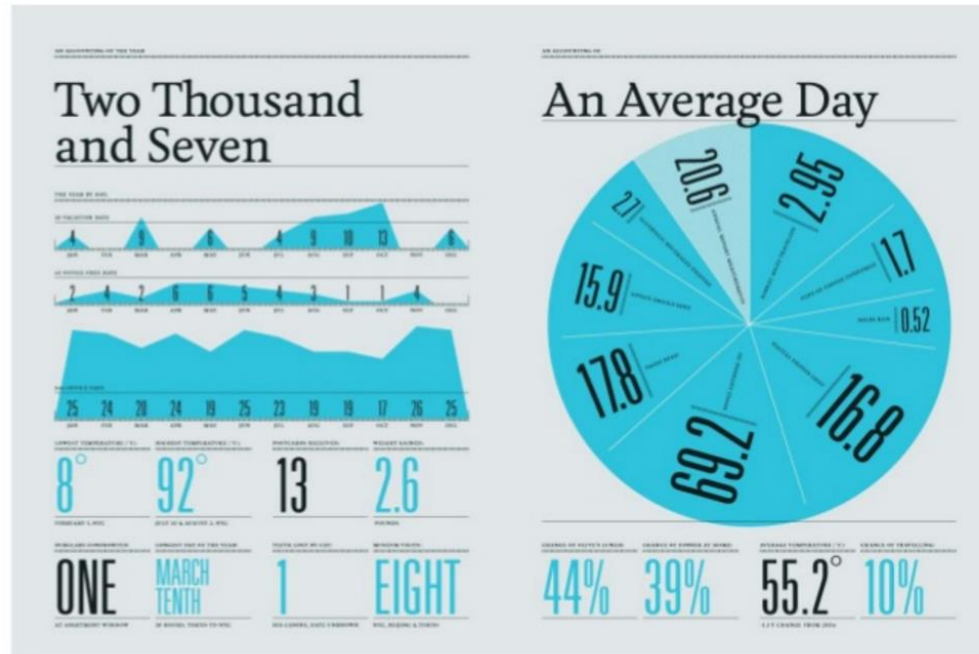
- 맥캔들레스는 색상의 의미를 문화권별로 구분한 색상환을 제작
- 원형 바깥의 숫자들은 오른쪽 범례에 있는 단어들의 의미를 보여주고, 원 안은 바깥부터 서구권, 동양권, 이슬람권 등으로 구분된 순서로 색상을 나열

# 제 2절 시각화 프로세스

## 1. 정보 디자인 프로세스

### 아. 8단계: 시각 접근 방법 결정하기

- 데이터의 아름다움을 만들어내는 과정



펠튼 2007년 연차 보고서

[http://feltron.com/ar07\\_02.htm](http://feltron.com/ar07_02.htm)

- 니콜라스 펠튼은 최근 매년 자신의 삶의 모든 정량적 측면을 문서화
- 2000 조각의 한정판으로 주문하는 개인 연차 보고서(Annual Report)를 인쇄해 발표해오고 있음
- 2007년 그가 살았던 한 해의 데이터를 수치화해 시각적으로 표현한 것



# 제 2절 시각화 프로세스

## 1. 정보 디자인 프로세스

### 아. 8단계: 시각 접근 방법 결정하기

- 데이터의 아름다움을 만들어내는 과정

#### ■ 8단계: 시각 접근 방법 결정하기(2)

- 피터 온토트나 스코트 스토클 처럼 일러스트레이션이나 메타포를 이용하는 방법



피터 온토트의 콘텍스트에서 인포메이션 그래픽

<http://www.behance.net/gallery/Information-graphics-in-context/924345>

- 피터 온토트의 사진을 이용한 시각화 방법은 은유적인 측면을 콘텍스트에 담아 정보를 극명하게 보여줌
- 덴마크 공공직업에서 종교 복장을 입는 것에 대한 덴마크 사람들의 윤리의식을 조사한 것



# 제 2절 시각화 프로세스

---

1. 정보 디자인 프로세스

자. 9단계 : 정제와 테스트

차. 10단계 : 세상에 선보이기

# 제 2절 시각화 프로세스

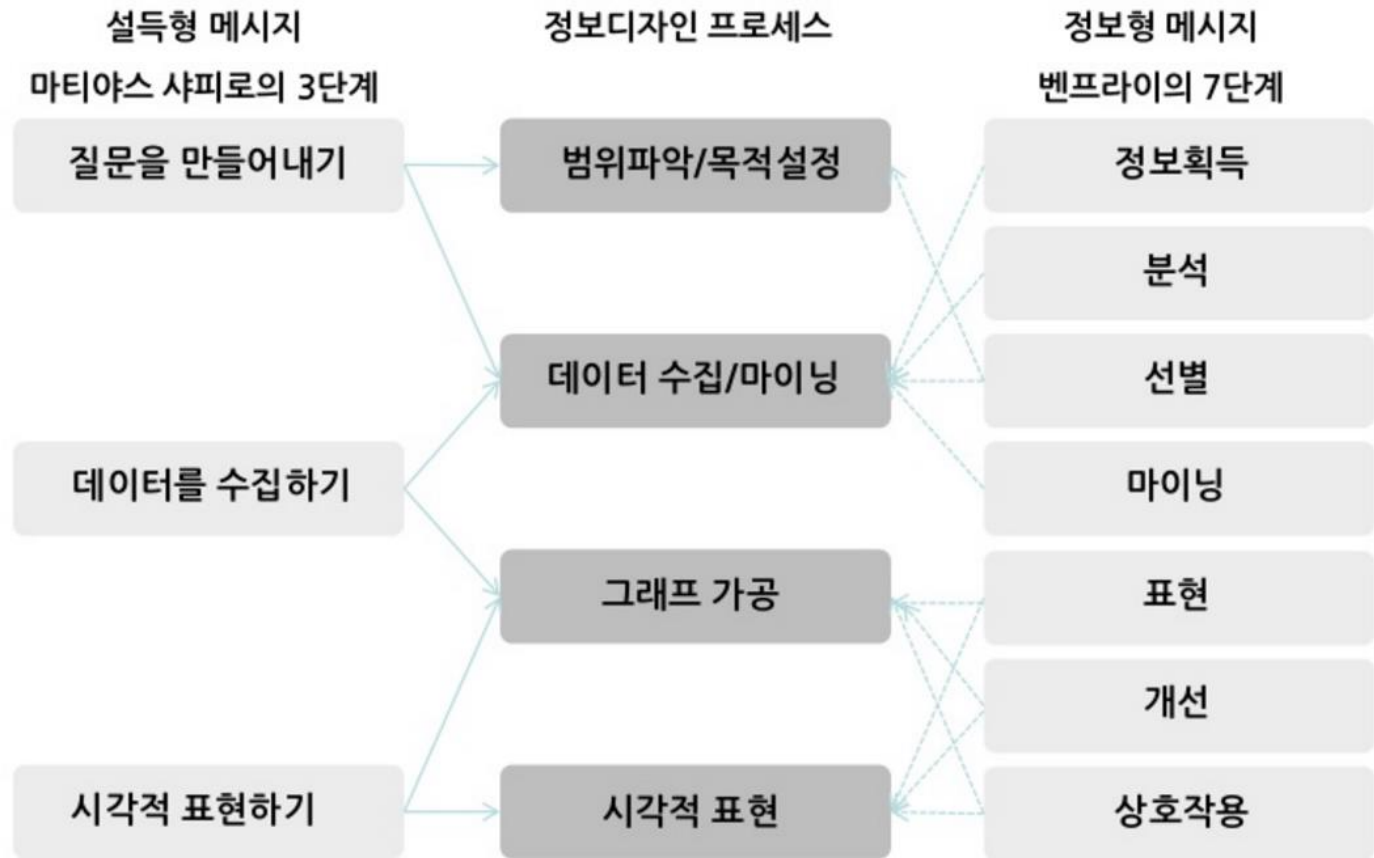
---

## 2. 빅데이터 시각화 프로세스

- 우선 빅데이터 시각화를 위한 다양한 시각화 방법을 살펴보면, 정보 디자인 교과서에서는 정보 시각화 프로세스를 조직화된 데이터(organized data), 시각적 매핑(visual mapping), 시각적 형태(visual form), 전달방식(representation)이라는 4단계로 제시
- 마티아스 샤피로는 질문 만들기, 데이터 수집하기, 시각적 표현 적용하기라는 3개의 핵심 단계를 제시, 빅데이터 시각화에서는 주어진 범위의 데이터에서 어떻게 데이터들을 먼저 갖고 질문을 만들어 낼 것인가의 순서가 바뀌어 나타날 수 있음.
- 프로세싱(Processing)을 개발하고 다수의 데이터 시각화 작업을 수행한 벤 프라이의 7단계에서는 정보 시각화를 만드는 데 필요한 정보획득(acquire), 분해(parse), 선별(filter), 마이닝(mine), 표현(represent), 정제(refine), 상호작용(interact)이라는 7단계를 제시

# 제 2절 시각화 프로세스

## 2. 빅데이터 시각화 프로세스



# 제 3절 시각화 방법

빅데이터 시각화는 정보 구조화, 정보 시각화, 정보 시각표현의 3단계로 진행

정보의 구조화를 위해서는 정보를 분류, 조직화, 재배열하는 기본 원리를 익히고, 정보시각화를 위해서는 각종 툴에서 일반적으로 제공하는 그래프 스타일의 원리와 쓰임새를 아는 것이 중요

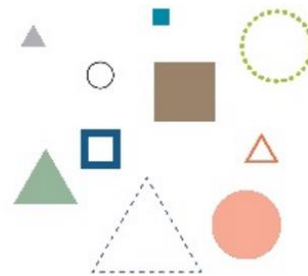


# 제 3절 시각화 방법

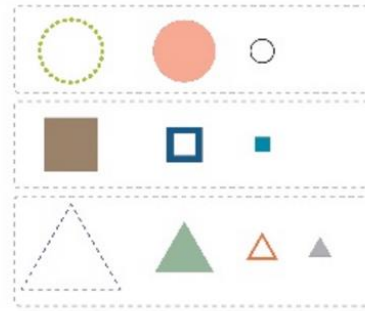
## 1. 정보 구조화

- 빅데이터의 정보 디자인의 성패는 빅데이터에서 다루는 데이터를 어떻게 통합했는지, 데이터 고유의 특성이 훼손되지 않으면서 비교되는 통계 품질이 어떠한지, 의도한 정보 디자인을 위한 통계 데이터의 도출이 적절하게 됐는지 의해 좌우됨.

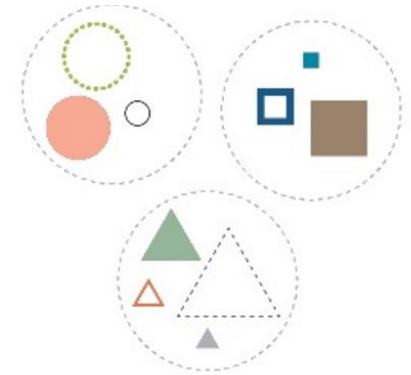
정보의 조직화 과정



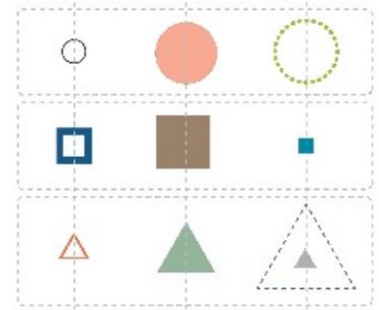
[데이터 수집] 데이터 자료 또는 논거를 수집



[배열] arranging 값의 의미에 따라 배치



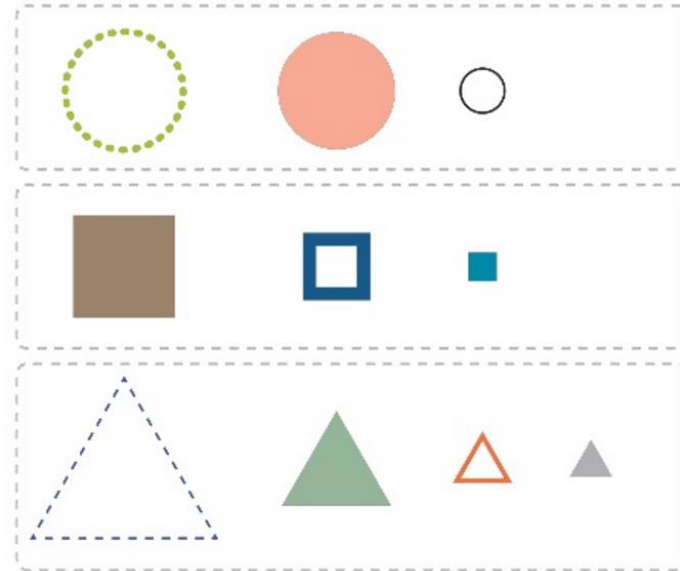
[분류] classifying 분류해 유사한 것끼리 묶음



[관계 맺기] organizing 데이터 사이의 관계를 만듦

# 제 3절 시각화 방법

## 1. 정보 구조화



### 배열

- 리처드 솔 위먼의 저서인 『정보 욕구(Information Anxiety)』에서는 정보의 조직화를 위한 래치(LATCH) 방법 제시
- 위치(Location), 알파벳(Alphabet), 시간(Time), 카테고리(Category), 위계(Hierarchy) 이상 5가지가 정보를 정리 또는 조직화하는 기준이 됨

# 제 3절 시각화 방법

## 2. 정보 시각화

정보 시각화 방법				
시간 시각화	분포 시각화	관계 시각화	비교 시각화	공간 시각화
막대그래프 누적 막대그래프 점그래프	파이차트 도우넛차트 트리맵 누적연속그래프	스캐터 플롯 버블차트 히스토그램	히트맵 체르노프 페이스 스타차트 평행 좌표계 다차원 척도법	지도 매핑



# 제 3절 시각화 방법

---

## 2. 정보 시각화

### 가. 시간 시각화

- 시간에 따른 데이터의 변화를 표현하며, 시계열 데이터의 가장 특징적인 요소는 트렌트, 즉 경향성으로 **장시간에 걸쳐 진행되는 변화 또는 트렌트를 추적하는데 주로 사용**
- 시간 데이터는 분절형과 연속형을 분류 가능
  - 분절형 : 데이터의 특정 시점 또는 특정 시간 구간 값
  - 연속형 : 기온 변화 같은 데이터

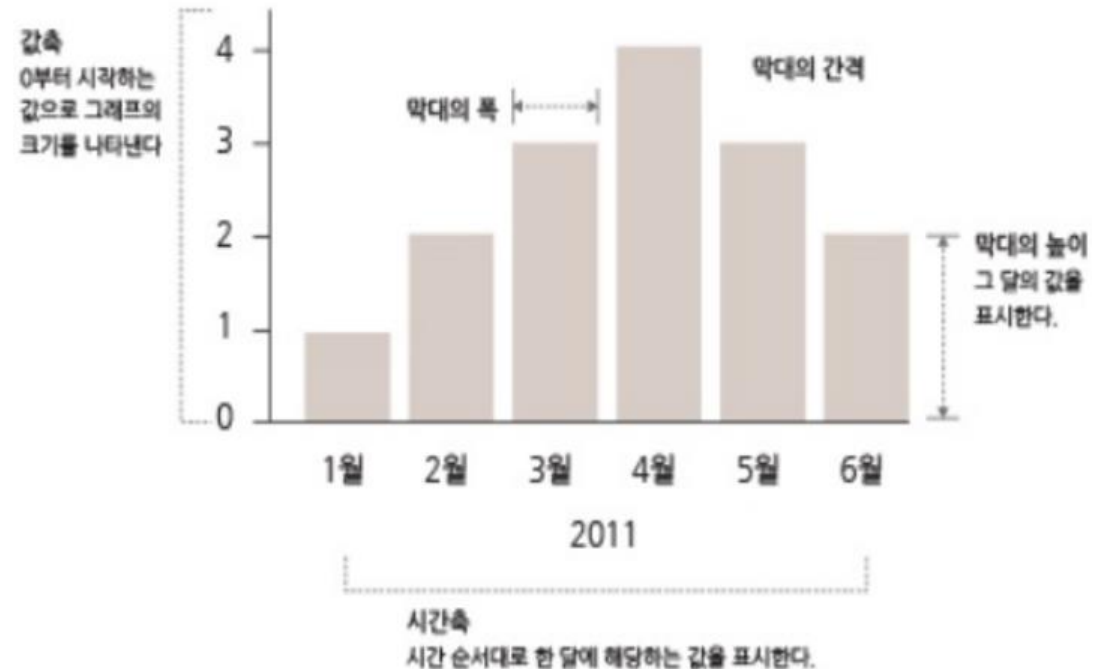
# 제 3절 시각화 방법

## 2. 정보 시각화

### 가. 시간 시각화

#### ① 막대 그래프

- **값들이 뚜렷한 차이를 보이는 경우로 수치를 길이로 표현하여 상대적인 차이를 한눈에 알아보는 것이 가능하다**
  - 시간축(가로)은 시간 순서대로 정렬된 시간의 특징 시점을 나타냄.
  - 값축(세로)은 그래프의 크기 범위를 나타냄
- **막대 값들의 차이가 미미할때는 비교가 쉽지 않은데, 막대에 다양한 색상을 적용할 수 있는데, 이때 색상은 특정 상태나 범위를 나타냄.**



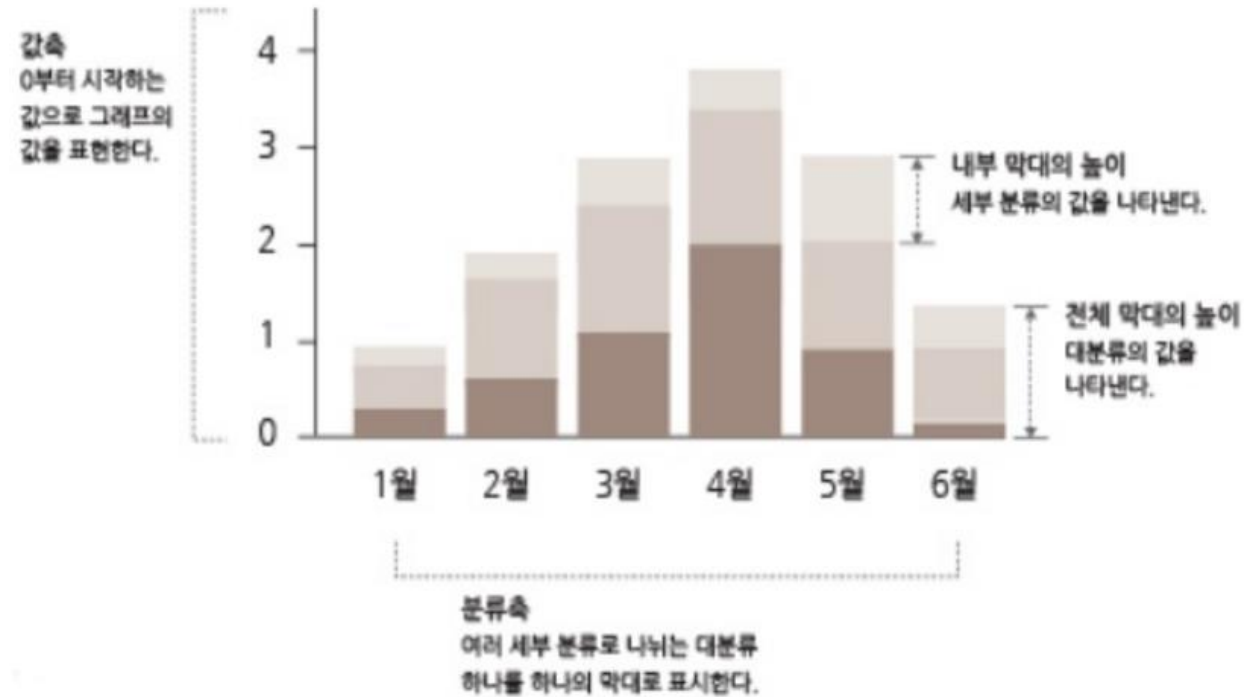
# 제 3절 시각화 방법

## 2. 정보 시각화

### 가. 시간 시각화

#### ② 누적 막대 그래프

- 누적 막대그래프의 구성은 막대 그래프와 거의 비슷, 차이점은 **한 구간에 해당하는 막대가 누적**된다는 점
- 한 구간이 몇 개의 세부 항목으로 나뉘면서도 전체의 합이 의미가 있을 때 누적 막대그래프 사용
- 한 구간의 각 세부항목은 질감 또는 색상으로 구분



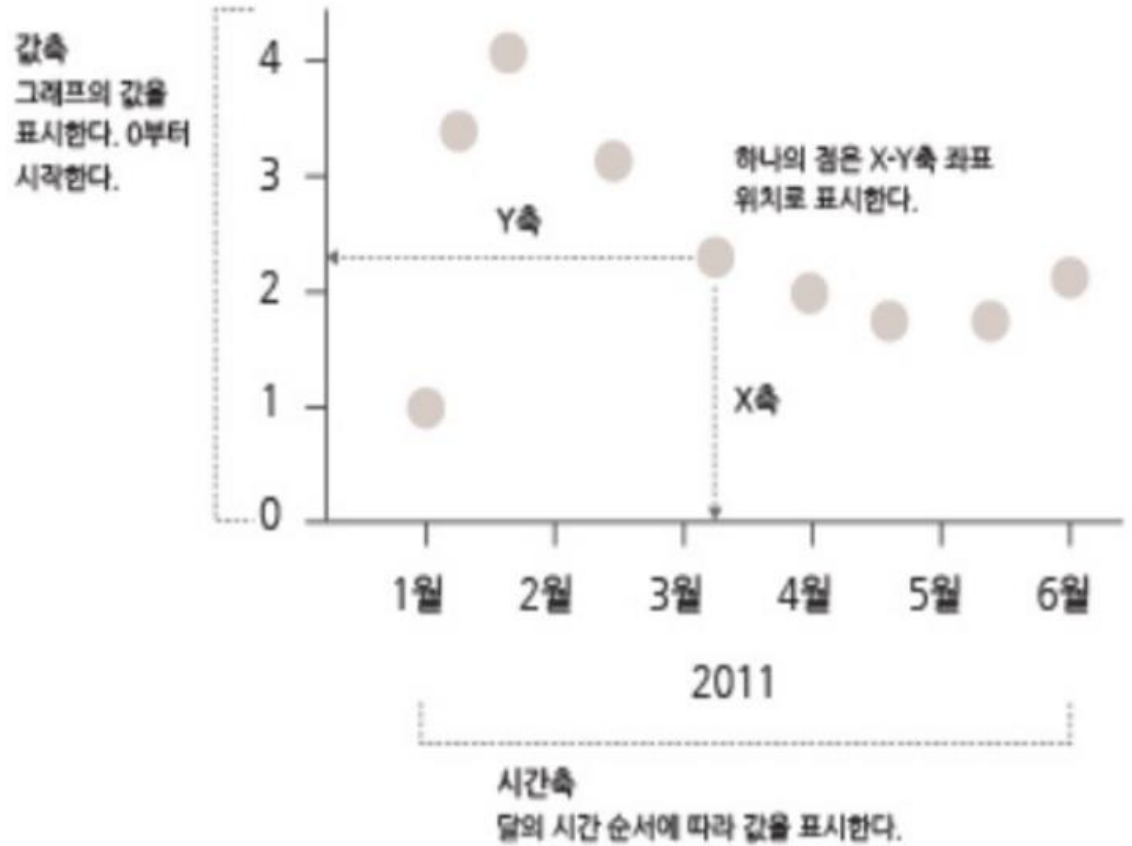
# 제 3절 시각화 방법

## 2. 정보 시각화

### 가. 시간 시각화

#### ③ 점그래프

- 점그래프는 면적을 표시할 필요가 없기 때문에 **더 적은 공간**에 그릴 수 있고, 한점에서 다음점으로 변하는 **점의 집중 정도와 배치에 따라 흐름을 파악**하기 용이.
- 일반적으로 **두 변수의 연관 관계**를 보여줄 때 많이 쓰임.



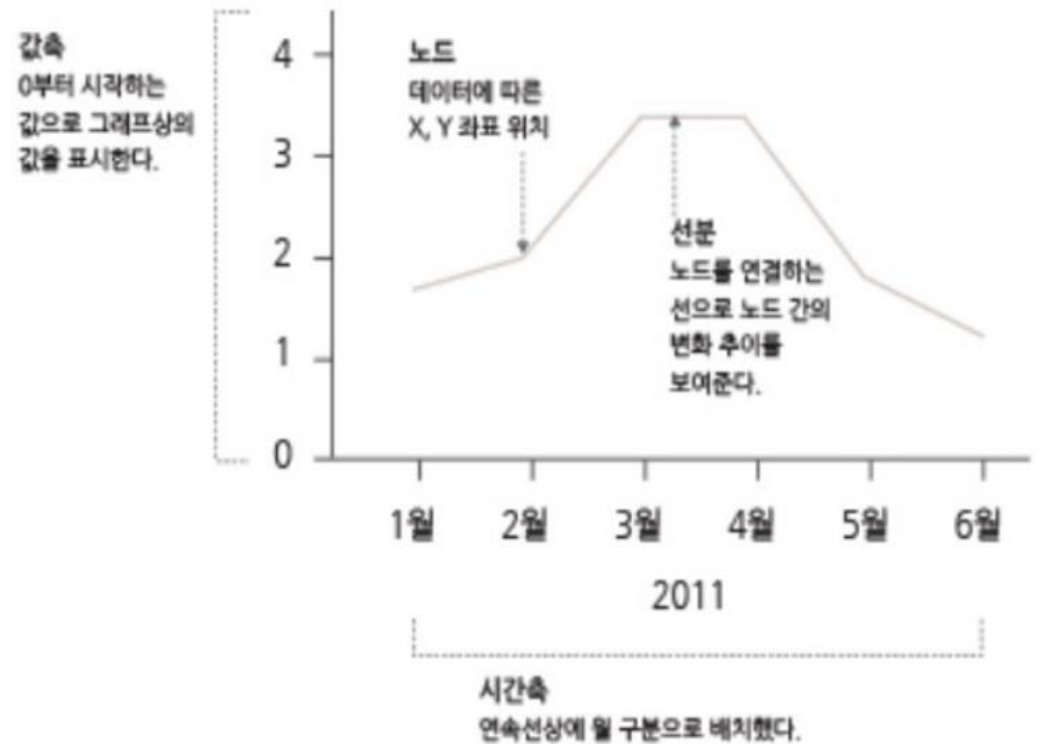
# 제 3절 시각화 방법

## 2. 정보 시각화

### 가. 시간 시각화

#### ④ 연속형 데이터: 연결된 점.선 그래프

- 연속 시계열 그래프는 점그래프와 거의 같은데, **점 사이를 선으로 연결한다는 점이 차이**
- 선으로 표현되는 연속적인 데이터의 끊임없이 변화하는 현상의 추이를 볼 수 있으며, **변수의 변화를 명확히 표시할 필요가 있거나 트렌딩 또는 변화율 정보가 중요한 경우 선그래프를 사용 가능**



# 제 3절 시각화 방법

---

## 2. 정보 시각화

### 나. 분포 시각화

분포 데이터의 일반적인 특성은 최대, 최소, 전체분포로 분류

분포 그래픽에서 가장 주목해야 할 것은 분포 정도

분포 데이터는 부분을 전부 합치면, 1 또는 100% 가 됨

# 제 3절 시각화 방법

## 2. 정보 시각화

### 나. 분포 시각화

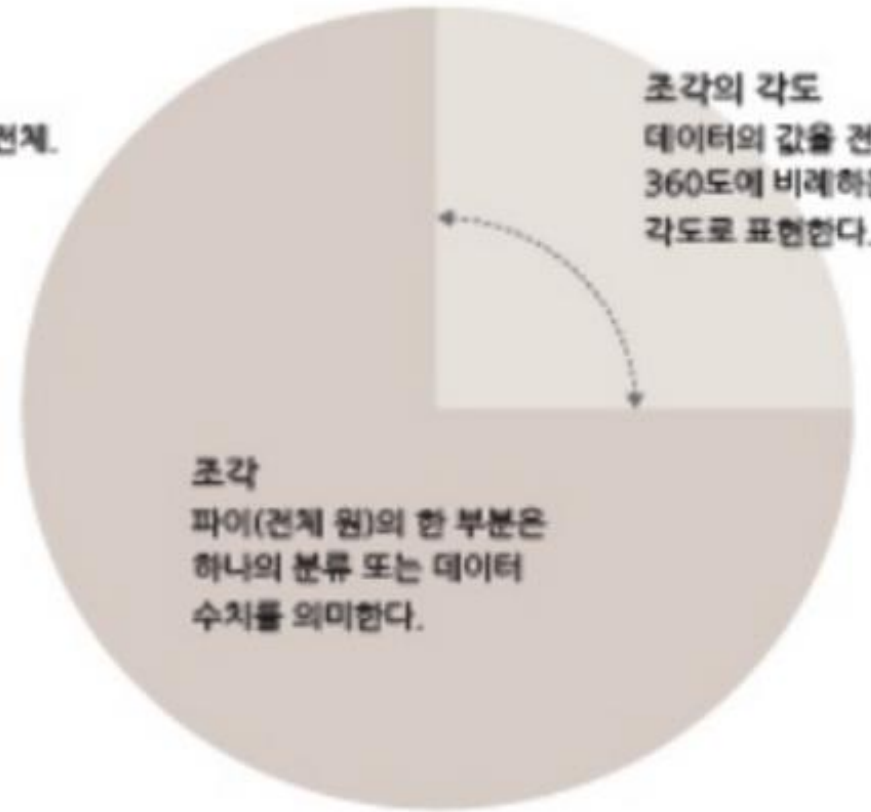
#### ① 원 그래프

- 원그래프는 부분과 부분 간의 비율을 알아보는 데 사용되는 방법
- 최대한 구성 요소를 제한하고 내용 설명하기 위한 텍스트와 퍼센티지를 포함
- 원그래프는 면적으로 값을 보여주며, 수치를 각도로 표시

전체의 부분  
모든 조각의 합은 전체  
100%다.

조각의 각도  
데이터의 값을 전체  
360도에 비례하는  
각도로 표현한다.

조각  
파이(전체 원)의 한 부분은  
하나의 분류 또는 데이터  
수치를 의미한다.





# 제 3절 시각화 방법

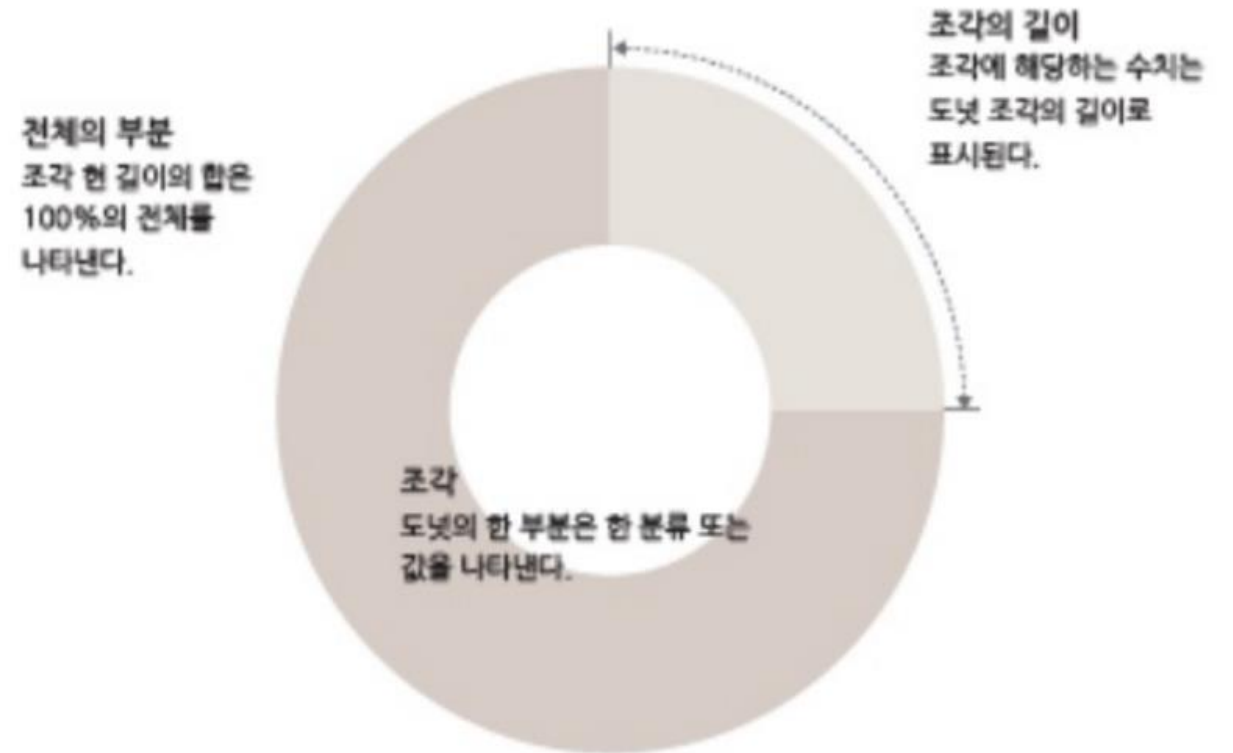
## 2. 정보 시각화

### 나. 분포 시각화

#### ② 도넛 차트

도넛 차트는 원그래프와 마찬가지로 수치를 각도로 표시

중심의 구멍 때문에 조각에 해당하는 수치는 **조각의 면적이 아닌 길이로 표시**



# 제 3절 시각화 방법

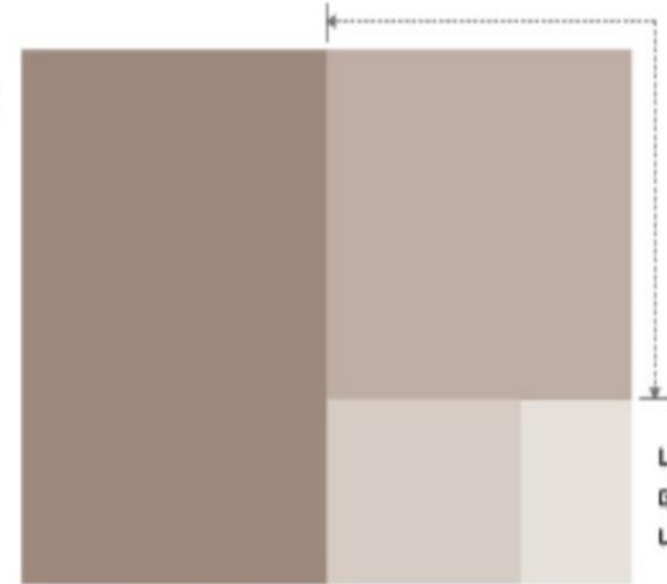
## 2. 정보 시각화

### 나. 분포 시각화

#### ③ 트리맵

- 트리맵은 영역 기반의 시각화로, **각 사각형의 크기** **수치**를 나타냄
- 한 사각형을 포함하는 있는 바깥의 영역은 그 사각형이 포함된 대분류를, 내부의 사각형은 내부적인 세부 분류를 의미
- 트리맵은 분류별 분포시각 기법이지만, **위계 구조가 있는 데이터나 트리 구조의 데이터**를 표시할때 활용 가능

전체의 부분  
모든 사각형의 면적은  
전체 100%를  
나타낸다.



면적  
사각형의 크기(면적)은  
전체 값의 비율을  
의미한다.

내부 사각형  
데이터의 위계 구조를  
나타낸다.



뉴스맵

<http://newsmap.jp/>

# 제 3절 시각화 방법

## 2. 정보 시각화

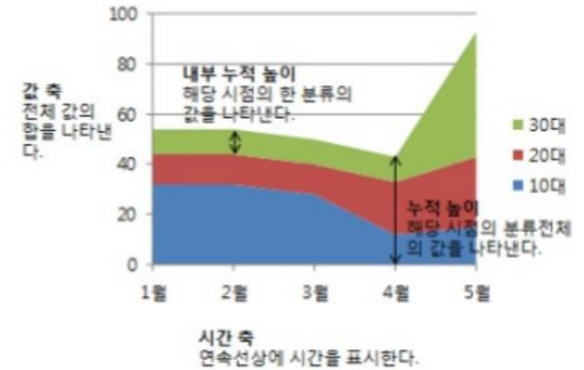
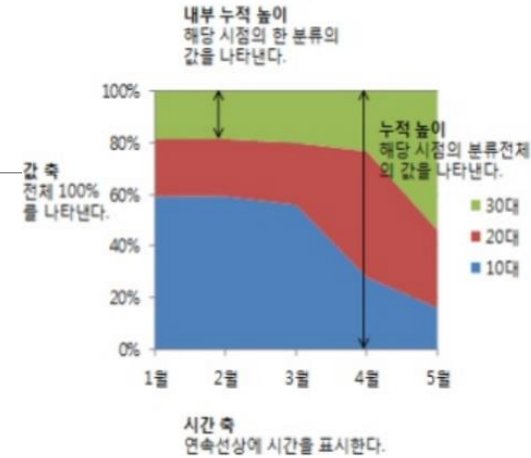
### 나. 분포 시각화

#### ④ 누적 연속 그래프

몇 개의 시계열 그래프를 차곡차곡 쌓아 올려  
그려 빈 공간을 채워가는 것

가로축은 시간, 세로축은 데이터 값

누적 연속 그래프는 한 시점의 세로 단면을 가  
져오면 그 시점의 분포를 볼 수 있음



네임보이저는  
100년이 넘는  
아기 이름  
데이터베이스

# 제 3절 시각화 방법

---

## 2. 정보 시각화

### 다. 관계 시각화

상관관계를 알면 한 수치의 변화를 통해 다른 수치의 변화를 예측 가능

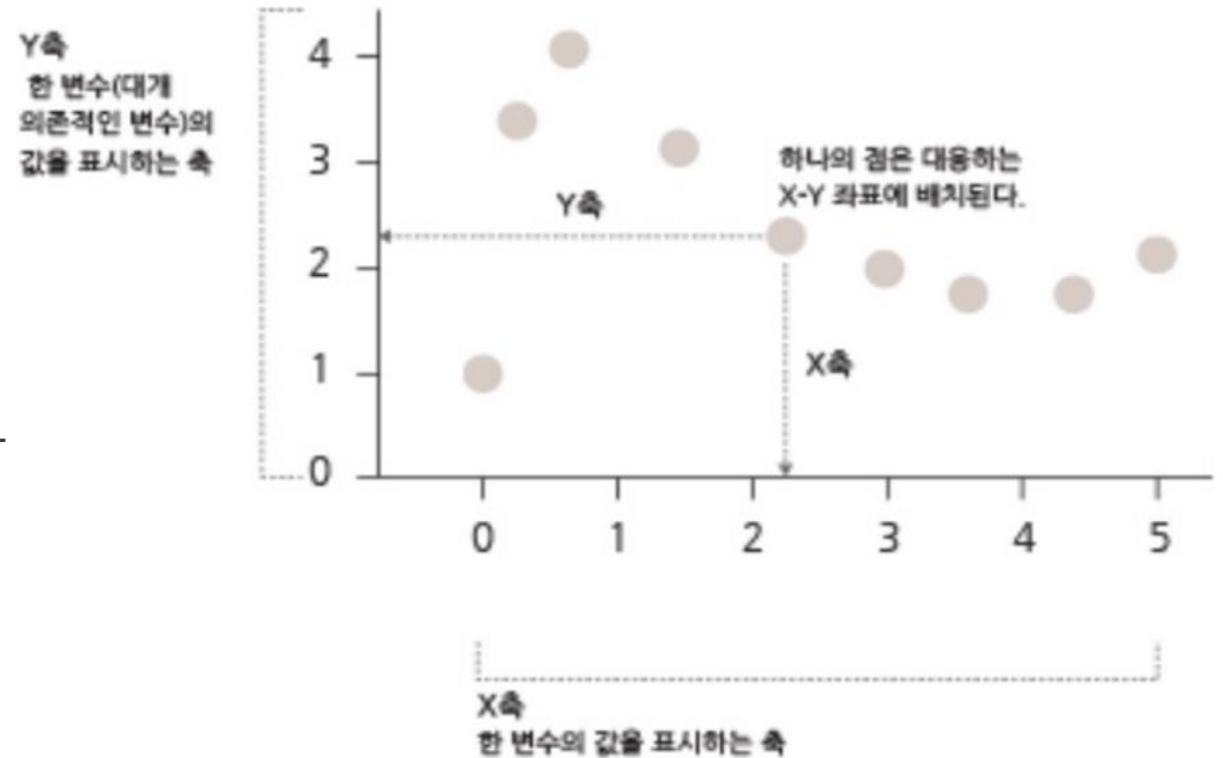
# 제 3절 시각화 방법

## 2. 정보 시각화

### 다. 관계 시각화

#### ① 산점도

- 두 데이터 항목의 공통 변이를 나타내는 2차원 도표
- 데이터 포인트가 많을 때 특히 유용한 반면, 데이터 포인트의 수가 적은 경우에는 막대 그래프나 일반 표가 효과적임.



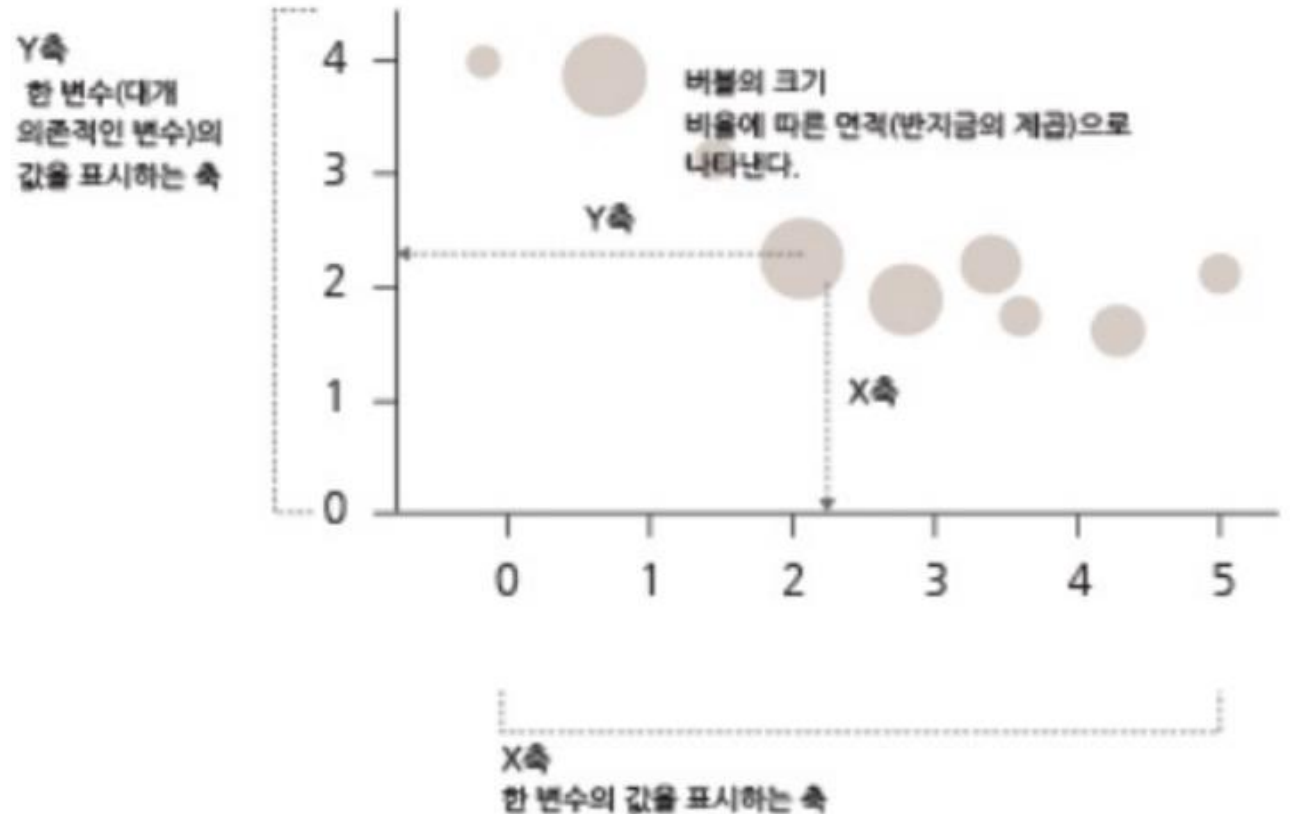
# 제 3절 시각화 방법

## 2. 정보 시각화

### 다. 관계 시각화

#### ② 버블차트

- 세 가지 요소의 상관관계를 표현
  - 가로축은 변수, 세로축의 변수, 버블의 크기
- 특정 값들을 다양한 크기의 버블로 시각적으로 표현하고자 할때 이 방식을 사용
- 데이터의 값은 원의 지름이 아닌 원의 면적



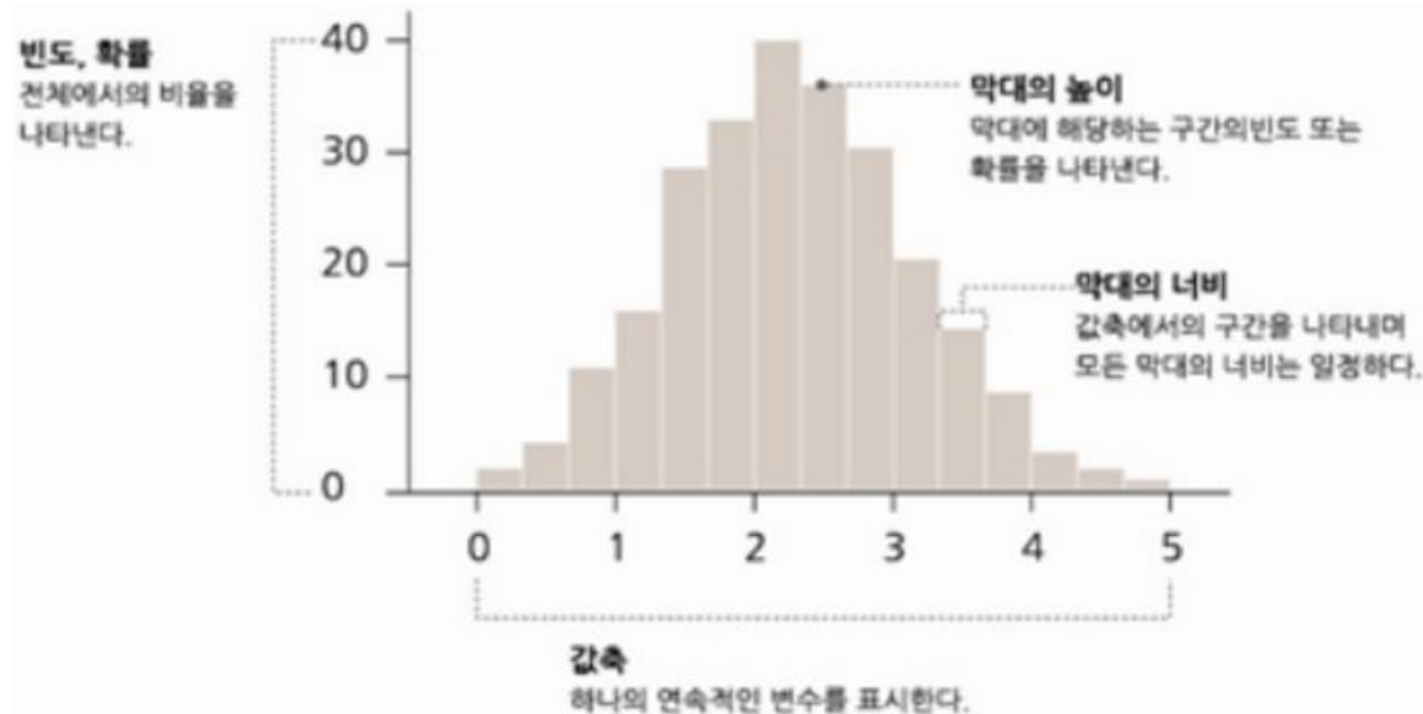
# 제 3절 시각화 방법

## 2. 정보 시각화

### 다. 관계 시각화

#### ③ 히스토그램

- 히스토그램(분포그래프) : 평균값을 중심으로 양옆이 점진적인 감소 모양을 보이는 종 곡성
- 막대의 높이는 빈도를 나타내고, 폭은 의미가 거의 없음.





# 제 3절 시각화 방법

---

## 2. 정보 시각화

### 라. 비교 시각화

- 여러 개의 변수를 다뤄야 할 때 마주하는 첫 번째 난관은 시작점을 찾는 것으로 자신에게 있는 데이터를 끊임없이 생각하다 보면 너무 많은 변수와 세부 분류에 압도되곤 함.
- 때로는 모든 데이터를 한번 훑어본 다음 흥미로운 점을 짚고 다른 점을 찾아가는 과정이 더 도움이 될 수 있음.

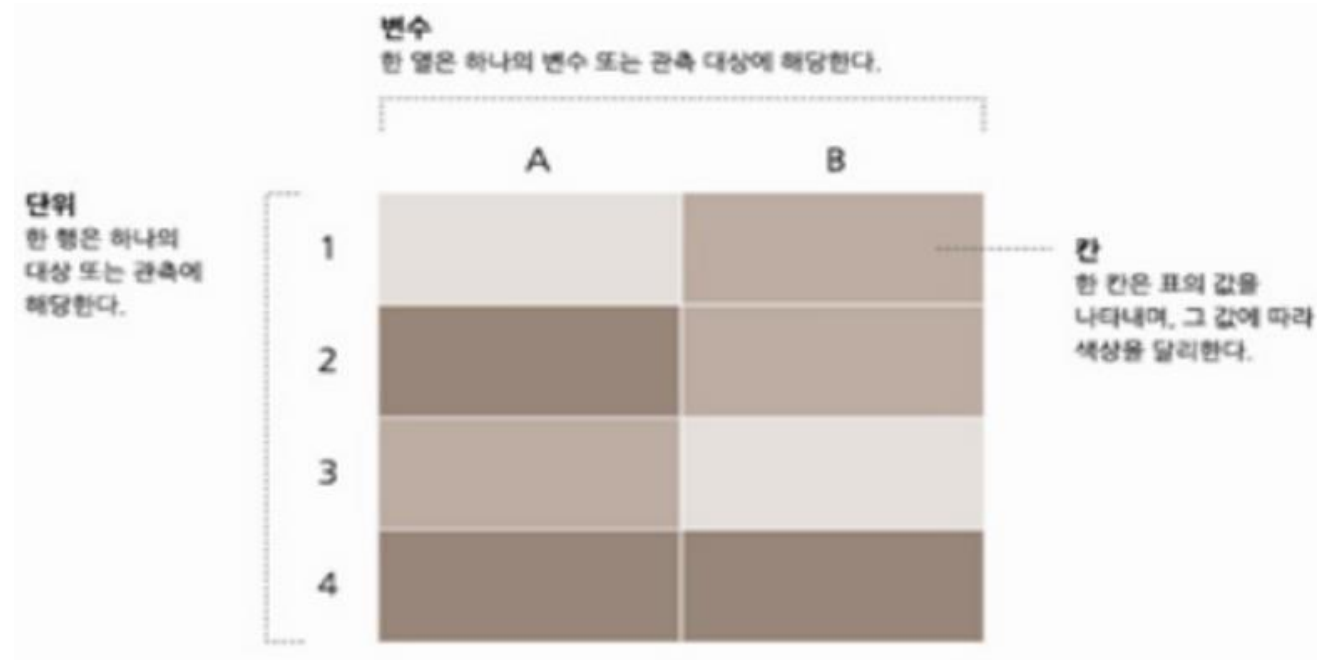
# 제 3절 시각화 방법

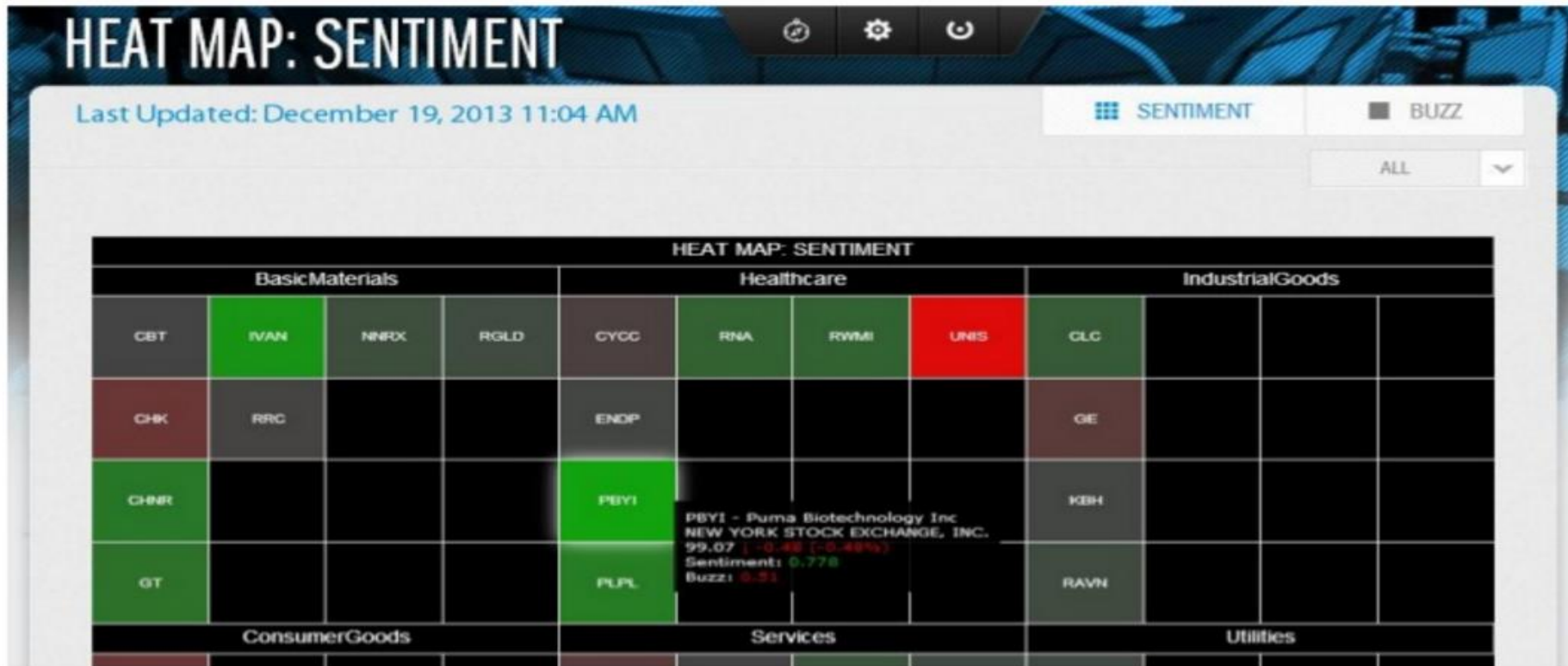
## 2. 정보 시각화

### 라. 비교 시각화

#### ① 히트맵

- 시각화 기법에서 가장 많이 유용하게 쓰이는 그래프 중 하나로 여러 가지 변수를 비교 가능
- 한칸의 색상으로 데이터 값을 표현
- 하나의 대상에 해당하는 한 행을 왼쪽으로 오른쪽을 보면서 모두 변수를 파악할 수 있고, 하나의 변수에 대응하는 한 열을 위해서 아래로 읽을 수 있음.
- 데이터가 지나치게 많은 경우엔 더 혼란스러울 수 있으니 적당한 색상을 선택하고 약간의 정렬 과정을 거쳐야 함.





## 감정 히트맵

<http://www.marketprophit.com/heat-maps/>

- 주식시장에 대한 전망을 SNS의 사회적 주식 지수로 보여줌
- 트위터에서 많이 이야기되는 주식이 블록에 나타나며 해당 주식에 대해 어떠한 감정으로 이야기하는 지가 색상으로 표시

# 제 3절 시각화 방법

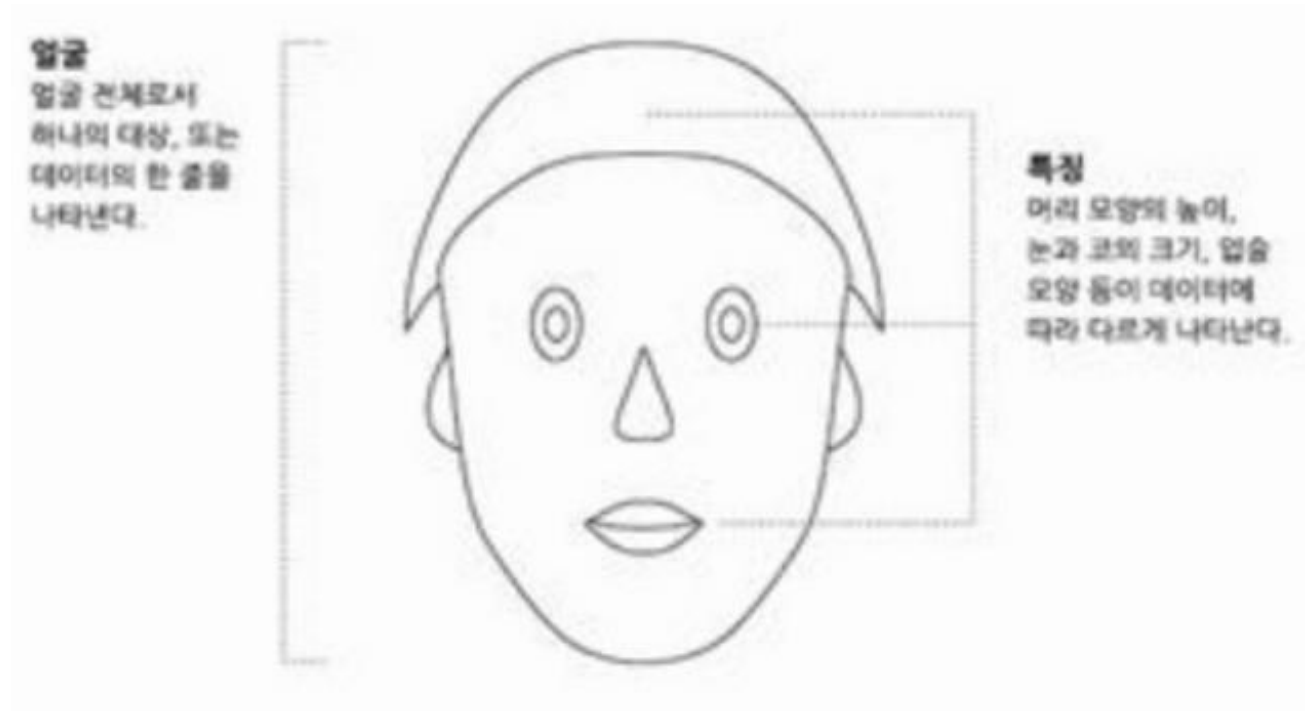
## 2. 정보 시각화

### 라. 비교 시각화

#### ② 체로노프 페이스

데이터를 사람의 얼굴 이미지로 표현하는 방법

유용성보다는 전문가의 흥미가 주목적



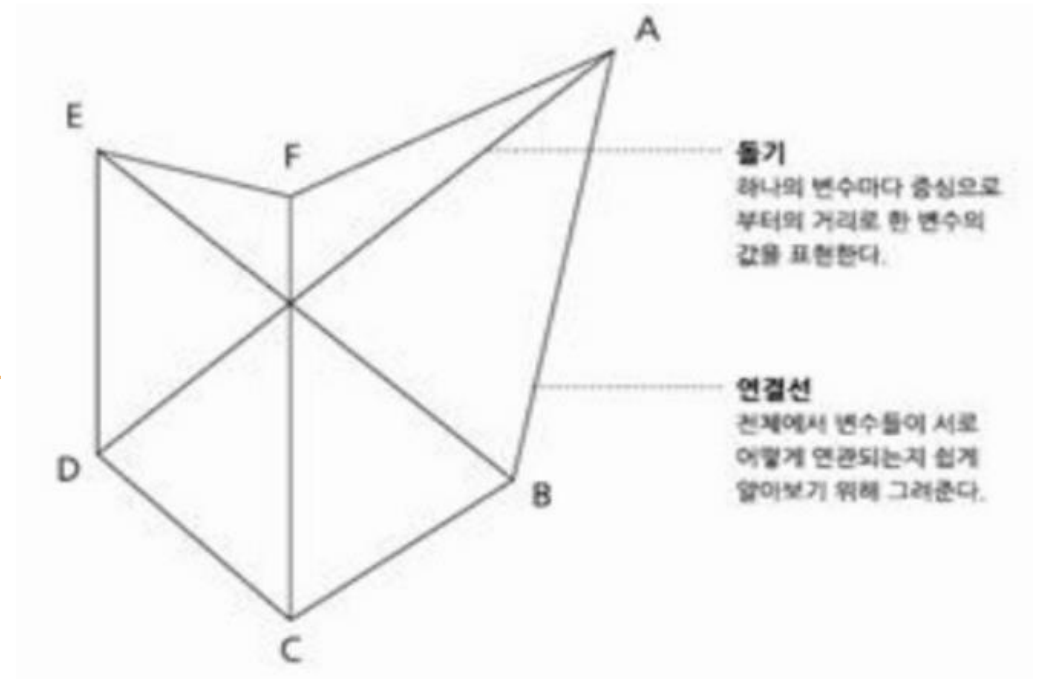
# 제 3절 시각화 방법

## 2. 정보 시각화

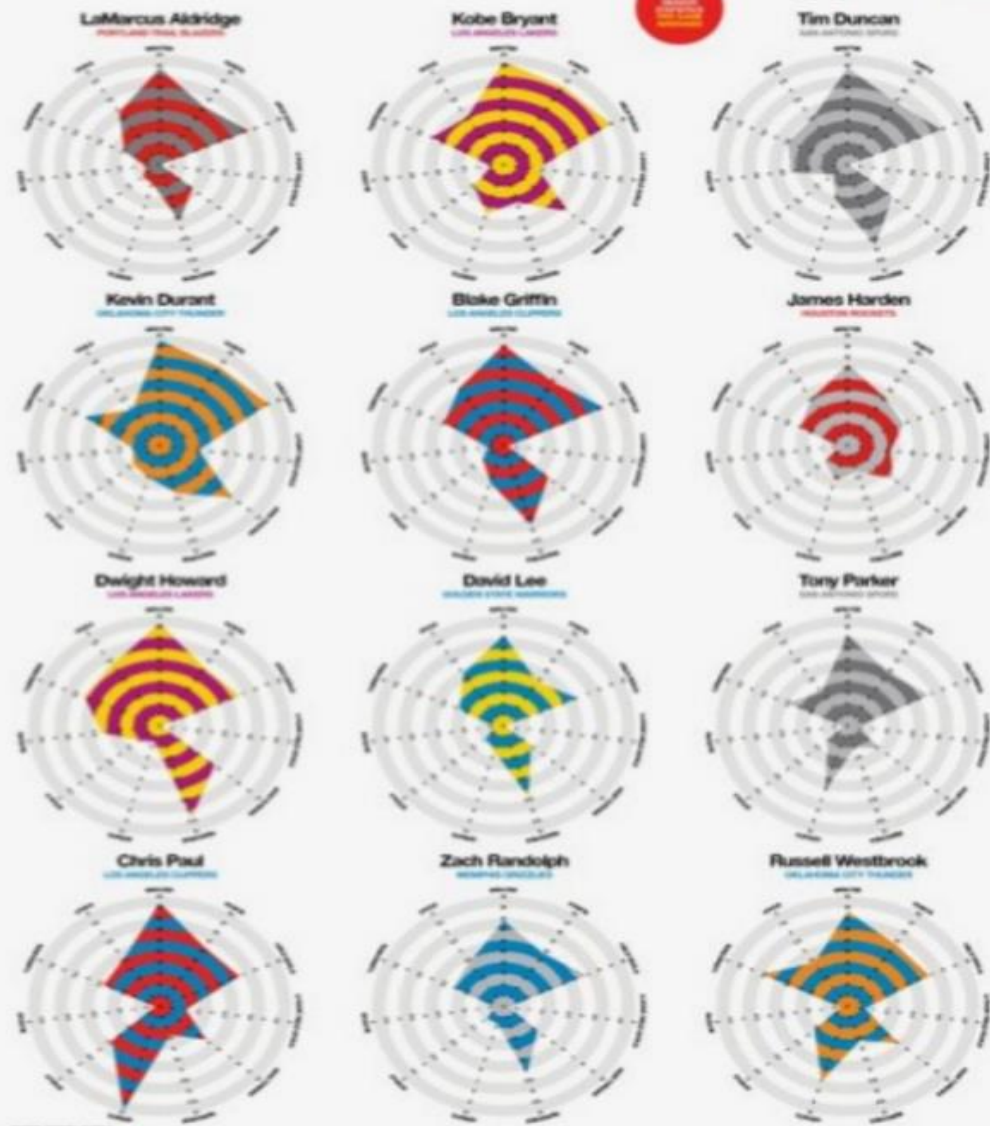
### 라. 비교 시각화

#### ③ 스타차트

- 스타차트는 거미줄 차트 또는 방상형 차트
- 중앙에서 외부 링까지 이어지는 몇 개의 축을 그리고, 전체 공간에서 하나의 변수마다 축 위의 중앙으로부터의 거리로 수치를 나타냄
- 각 변수를 라인 위에 표시한 지점을 연결해 연결선을 그리며 그 결과는 별 모양의 도형으로 나타냄
  - 중심점 : 축이 나타내는 값의 최소값
  - 바깥점 : 가장 먼 끝은 최대값을 나타냄



# 2013 NBA ALL-STARs WEST



- 미국 서부지역 NBA 농구 선수들의 강점을 방사형 차트로 표현한 것이다. 각 색상코드는 선수들이 속한 구단의 색상을 의미하며, 각 축은 농구선수들의 기록을 나타낸다.

2013 NBA 올스타 서부지역

<http://ramimo.com/2013-NBA-All-Stars>

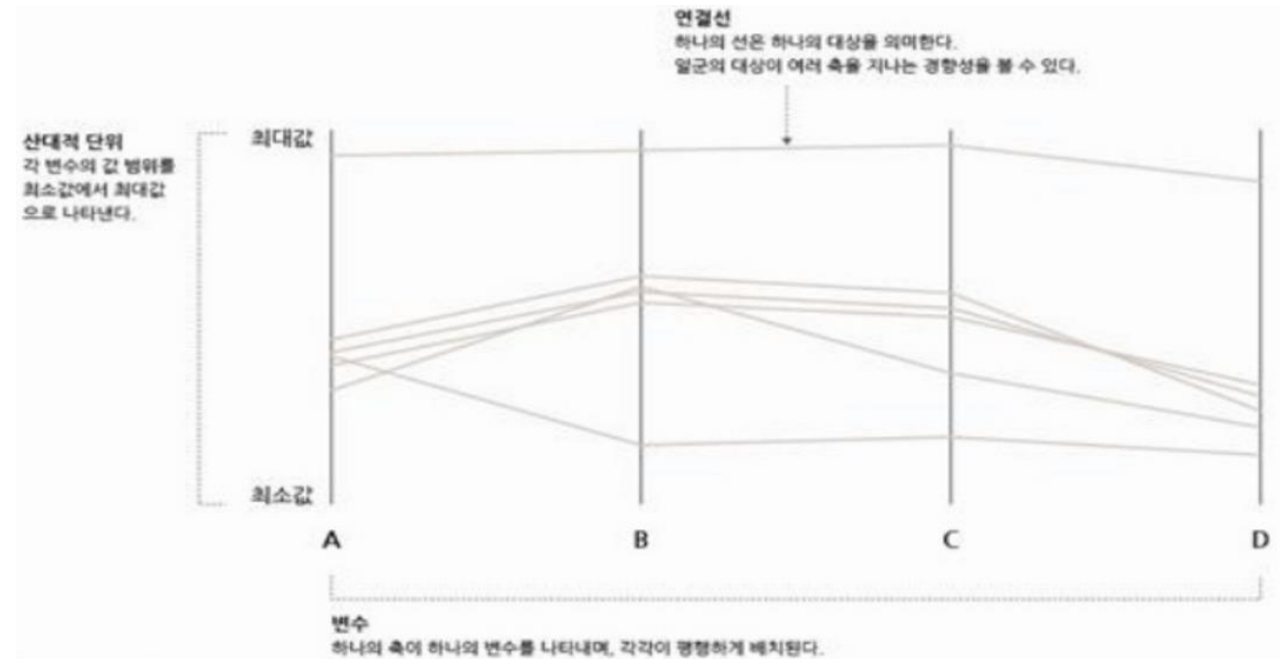
# 제 3절 시각화 방법

## 2. 정보 시각화

### 라. 비교 시각화

#### ④ 평행 좌표계

- 대상이 많은 데이터에서 집단적인 경향성을 쉽게 알아볼 수 있게 해줌.
  - 여러 축을 평행으로 배치해서 만듦.
  - Y축에서 윗부분은 변수값 범위의 최대값, 아래는 변수값 범위의 최소값
  - 측정 대상은 변수 값에 따라 위아래로 이어지는 연결선으로 그려짐.





## Nutrient Contents - Parallel Coordinates

An interactive visualization of the [USDA Nutrient Database](#). For information on parallel coordinates, read this [tutorial](#).

Hide Ticks

Dark

Shadows

Opacity: 10%

Per 100g of Food

Selected 1153 rows

Keep

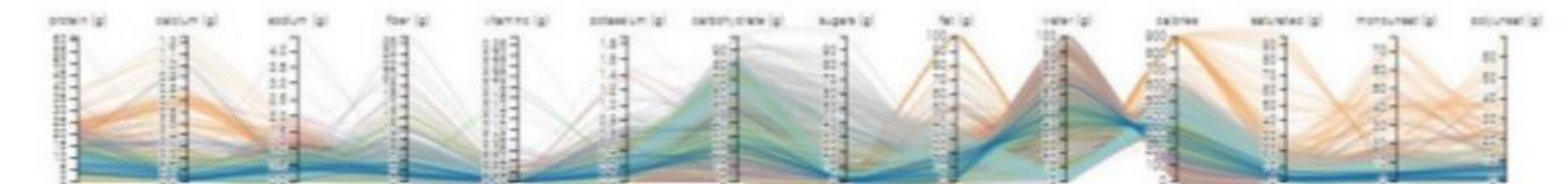
Remove

Export

Group Breakdown

Total Selected

Dairy and Egg Products
Pale and Ole
Poultry Products
Soups, Sauces, and Gravies
Vegetables and Vegetable Products
Sausages and Luncheon Meats
Breakfast Cereals
Fruits and Fruit Juices
Nut and Seed Products
Beverages
For Fish and Shellfish Products
Legumes and Legume Products
Baked Products
Sweets
Cereal Grains and Pasta
Fast Foods
Steak, Entrées, and Side Dishes
Snacks
Restaurant Foods



name	group	protein (g)	calcium (g)	sodium (g)	fiber (g)	vitamin...	potassiu...	carbohy...	sugars (g)	fat (g)	water (g)	calories
Cheese, brie	Dairy and Egg Products	20.75	0.184	0.629	0	0	0.152	0.45	0.45	27.68	48.42	334
Cheese, camembert	Dairy and Egg Products	19.8	0.388	0.842	0	0	0.187	0.46	0.46	24.26	51.8	300
Cheese, cheddar	Dairy and Egg Products	24.9	0.721	0.621	0	0	0.098	1.28	0.52	33.14	36.75	403
Cheese, cheshire	Dairy and Egg Products	23.37	0.643	0.7	0	0	0.095	4.78	0.52	30.6	37.65	387
Cheese, colby	Dairy and Egg Products	23.76	0.685	0.604	0	0	0.127	2.57	0.52	32.11	38.2	394
Cheese, cottage, creamed, large or small curd	Dairy and Egg Products	11.12	0.083	0.364	0	0	0.104	3.38	2.67	4.3	79.79	98
Cheese, cream	Dairy and Egg Products	5.93	0.098	0.321	0	0	0.138	4.97	3.21	34.24	54.44	342
Cheese, edam	Dairy and Egg Products	24.99	0.731	0.965	0	0	0.188	1.43	1.43	27.8	41.56	357

카이 창의 영양소 콘텐츠

<http://exposedata.com/parallel/>

- 카이 창(Kai Chang)의 영양소 콘텐츠 : USDA의 영양소 데이터베이스를 시각화
- 각 식품 100g당 들어있는 영양성분을 비교
- 카이 창이 개발해 배포하는 D3.js의 평행좌표계 - Parallel Coordinates(0.2.2)를 이용해 직접 개발 가능



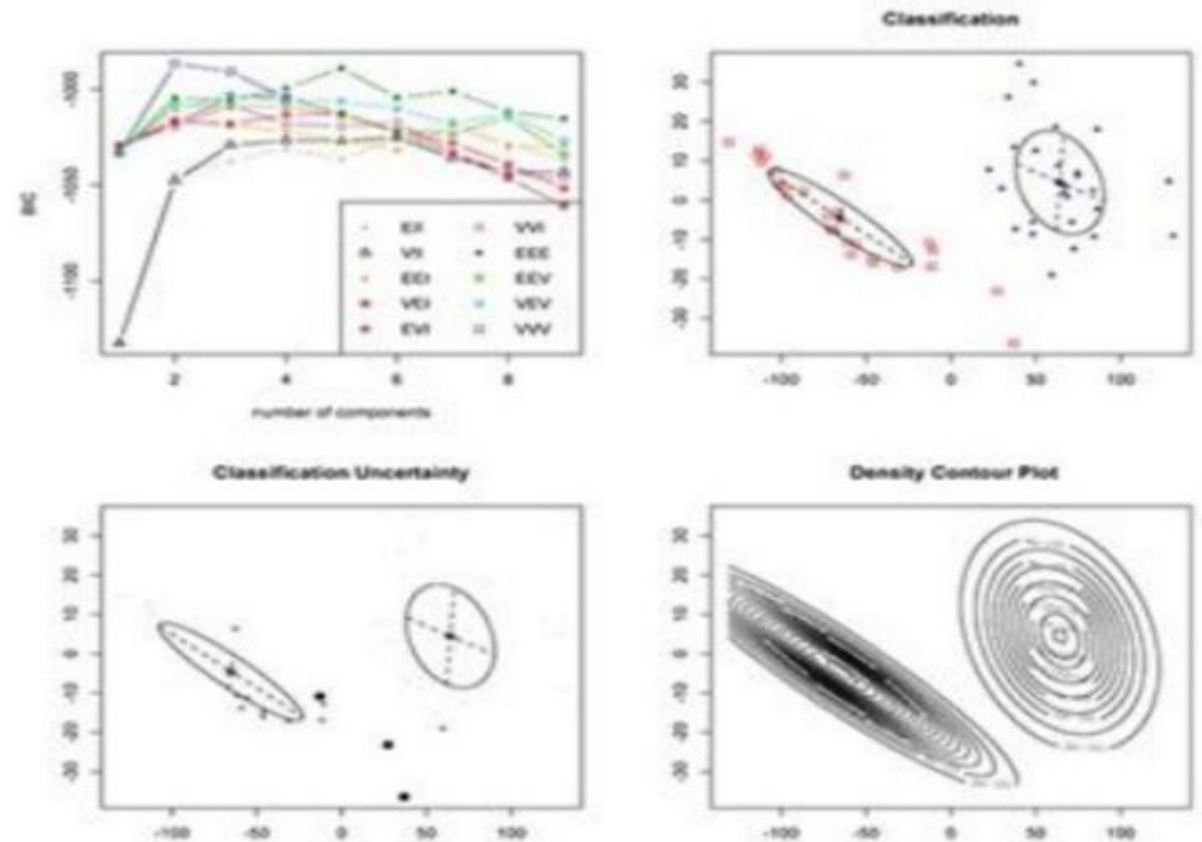
# 제 3절 시각화 방법

## 2. 정보 시각화

### 라. 비교 시각화

#### ⑤ 다차원 척도법

- 데이터세트상의 개별 데이터 간의 유사도를 바탕으로 시각화하는 방법으로 대상간의 유사도 측도에 의거해서 대상을 다차원 공간 속에 배치시키는 방법
- 표현하고자 하는 객체 간 간격이 발생하는, 즉 거리행렬을 포함하는 데이터의 시각화에 유용
- 유사성이 작은 대상끼리는 멀리, 유사성이 큰 대상끼리는 가까이 위치시킴



# 제 3절 시각화 방법

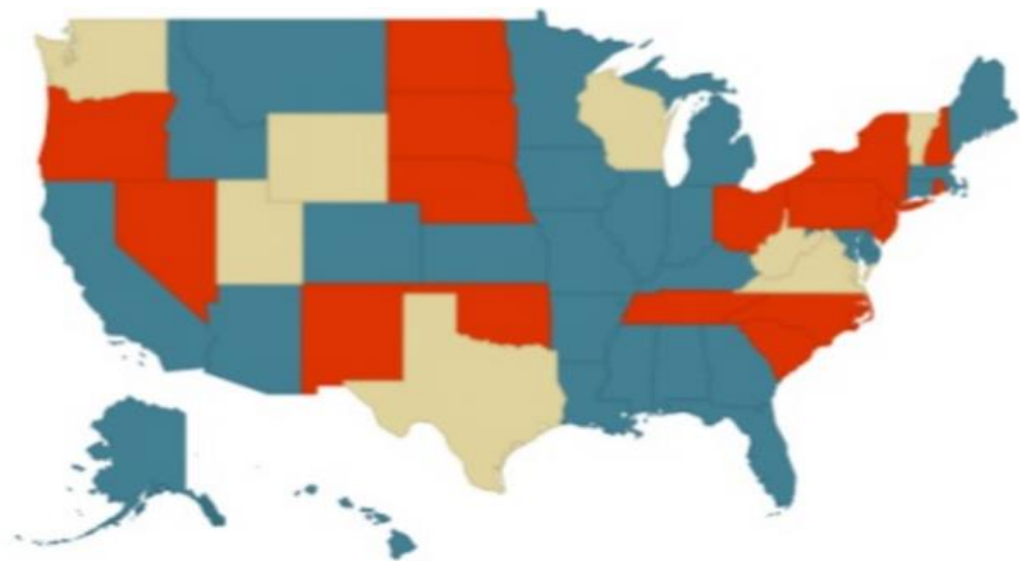
## 2. 정보 시각화

### 라. 비교 시각화

#### ⑤ 공간 시각화

- 지도의 한 위치를 다른 위치와 비교해보는 것
- 지도를 만들때, 위치를 정확히 배치해야 하며, 색상 구분도 정확해야 하고, 라벨이 위치를 가려서는 안 되며, 정확한 투시 방법을 선택해야 함.
- 하나의 지도는 시간상의 한 지점, 한 순간의 현실만을 반영하고 있지만, 여러 장의 지도를 통해 시간의 여러 단면을 표현 가능





Obama Romney Undecided Click any state to see a custom message here!

#### 인터랙티브 미국 선거 지도

<http://cmoreira.net/interactive-world-maps-demo/usa-custom>

- 위도와 경도의 위치 값을 이용해 지도 위에 정확한 매핑 포인트를 표시해야 하지만 구글 차트의 지오차트(Geo Chart)는 이 값을 모르고 지명만 알아도 시각화 작업을 가능
- 영역 모드와 마커 모드를 갖고 나라별, 대륙별 또는 지역의 지도를 쉽게 그려볼 수 있음
- 영역 모드 : 국가, 지방 또는 국가와 같은 전체 지역을 채색
- 마커 모드 마크 : 사용자가 지정한 값에 따라 조정되는 것으로 버블 차트를 사용해 영역을 지정가능

# 제 3절 시각화 방법

---

## 3. 정보 시각표현

### 그래픽적으로 디자인을 완성시키는 단계

정보 시각 표현에 사용되는 정보 디자인의 그래픽 요소는 **정보의 내용을 시각적 형식으로 전환해 전달 과정에서 사용자의 감각을 자극**하여 정보가 쉽게 이해하도록 시각적 안내 역할을 함.








# 제 3절 시각화 방법

## 3. 정보 시각표현

### 가. 정보표현을 위한 그래픽 요소

- 시각 정보를 보기 위해 뇌는 패턴을 감지하고 이를 해석한 뒤에 학습
- **자크 베르탱의 그래픽 7요소**
- 그래픽 요소는 수치적(양적), 질적 데이터 간의 차이, 순서, 비율의 관계를 시각적으로 전환할 수 있는 가장 효율적인 표현의 수단이며, 사용자가 이를 통해 정보를 시각적으로 지각하고 이해하게 하는 절대적인 역할을 한다.

- 자크 베르탱(Jacques Bertin)의 그래픽 7요소

Bertin's Original Visual Variables	
<b>Position</b> changes in the x, y location	
<b>Size</b> change in length, area or repetition	
<b>Shape</b> infinite number of shapes	
<b>Value</b> changes from light to dark	
<b>Colour</b> changes in hue at a given value	
<b>Orientation</b> 기울기, 정위 changes in alignment	
<b>Texture</b> 질감 variation in 'grain'	

# 제 3절 시각화 방법

---

## 3. 정보 시각표현

### 나. 시각화를 위한 그래픽 디자인 기본 원리

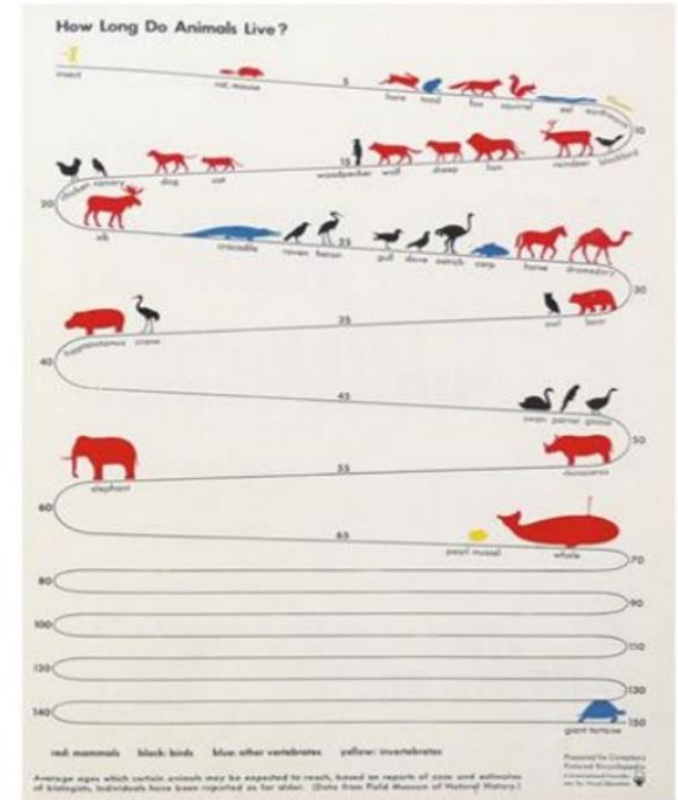
- ① 타이포 그래피 : 2가지의 서체 또는 1가지의 서체에 크기나 스타일의 변화를 주는 것이 중요
- ② 색상 : 여러 색상을 함께 쓸때 채도를 달리하면서 균형을 깨는것
- ③ 그리드 : 그리드를 이용해 블록 레이아웃을 잡고 그 위에 요소를 효율적으로 올려 전체적인 조화 추구
- ④ 아이소타이프 : 많은 양의 데이터를 쉽게 지각할 수 있도록 도와주는 시각표현 방법

# 제 3절 시각화 방법

## 3. 정보 시각표현

### 아이소타이프 사례

- 1960년부터 활동한 미국 디자이너 리처드 솔 워먼
- '이해를 결정하는 것은 사물 그 자체가 아니라 우리가 사물과 함께 연상하는 의미나 패턴'
- 이미지를 활용해 정보를 명확하게 전달하는 방식
- 아이소타입은 그래프를 표현하는 수단으로도 자주 활용
- 최근 들어 아이콘 등으로 발전



동물은 얼마나 오래 사는가?

<http://visual.ly/how-long-do-animals-live>



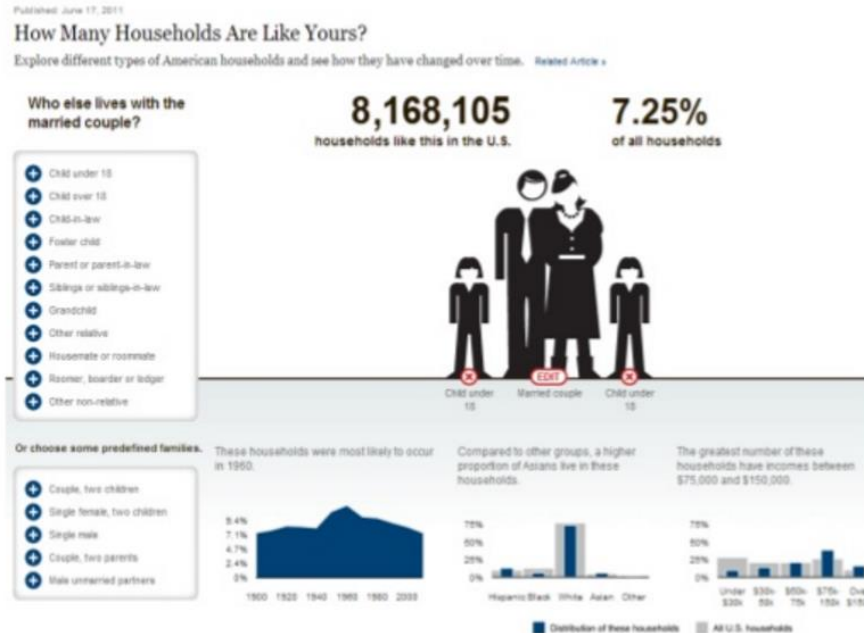
# 제 3절 시각화 방법

## 3. 정보 시각표현

### 다. 인터랙션

사용자 스스로 정보를 필터링하고 탐색하는 과정에서 인사이트를 얻어갈 수 있는 기회 제공을 함.

- 그 예로 '당신과 같은 가구는 얼마나 있는가?(How Many Households Are Like Yours?)'에서 사용자는 가족구성원을 더하거나 빼면서 기본 가구타입을 선택할 수 있다. 변경의 경우에, 즉각적으로 반영되는 시각화된 정보가 보여진다. 이러한 인터랙션은 수많은 가구타입을 데이터를 한꺼번에 보지 않아도 많은 다른 종류의 가구 타입의 데이터를 보는 것을 가능하게 한다.



당신과 같은 가구는 얼마나 있는가?

[http://www.nytimes.com/interactive/2011/06/19/nyregion/how-many-households-are-like-yours.html?\\_r=0](http://www.nytimes.com/interactive/2011/06/19/nyregion/how-many-households-are-like-yours.html?_r=0)



# 제 3절 시각화 방법

## 3. 정보 시각표현

### 라. 시각 정보 디자인 7원칙

- 위의 정보 구조화, 정보 시각화, 정보 시각표현을 거쳐 최종적으로 시각화가 완성된다고 할 수 있다. 에드워드 터프티(Edward Tofte)는 다음과 같은 시각정보 디자인 7원칙을 제시한다.

터프티의 시각정보 디자인 7원칙

1. 시각적 비교를 강화하라	연과된 변수와 트렌드를 비교할 수 있는 도구를 제공해야 한다. 정보는 비교를 통해 사용자에게 더욱 가치가 올라간다.
2. 인과관계를 보여라	정보를 디자인할 때 원인과 결과를 명쾌하게 제시하라.
3. 다중변수를 표시하라	여러 개의 연관된 변수를 활용해 정보를 표현하는 데이터도 있다.
4. 텍스트, 그래픽, 데이터를 조화롭게 배치하라	라벨과 범례가 도표에 녹아 있는 다이어그램이 더욱 효과적이다.
5. 콘텐츠의 질과 연관성, 진실성을 분명히 하라	보여주려는 정보가 과연 사용자가 특정 목적을 달성하는 데 도움이 되는지를 고민하라
6. 시간순이 아닌 공간순으로 나열하라	시간보다는 공간에 따라 나열할 때 사용자의 이해가 쉬워진다.
7. 정량적 자료의 정량성을 제거하지 마라	트렌드를 나타내기 위해 정량적 자료를 그래프나 도표로 표현할 수 있다. 정량적인 정보를 한 눈에 파악할 수 있기 때문이다.

- 에드워드 터프티(Edward Tofte)는 ‘훌륭한 시각 디자인은 시각적으로 표현된 명쾌한 생각’이라고 말했다. 보는 이의 인지적인 과업과 디자인 원칙을 깊이 이해해야 훌륭한 시각화 작업을 할 수 있다.

# 제 4절 빅데이터와 시각화 디자인

## 1. 빅데이터와 시각화 이슈

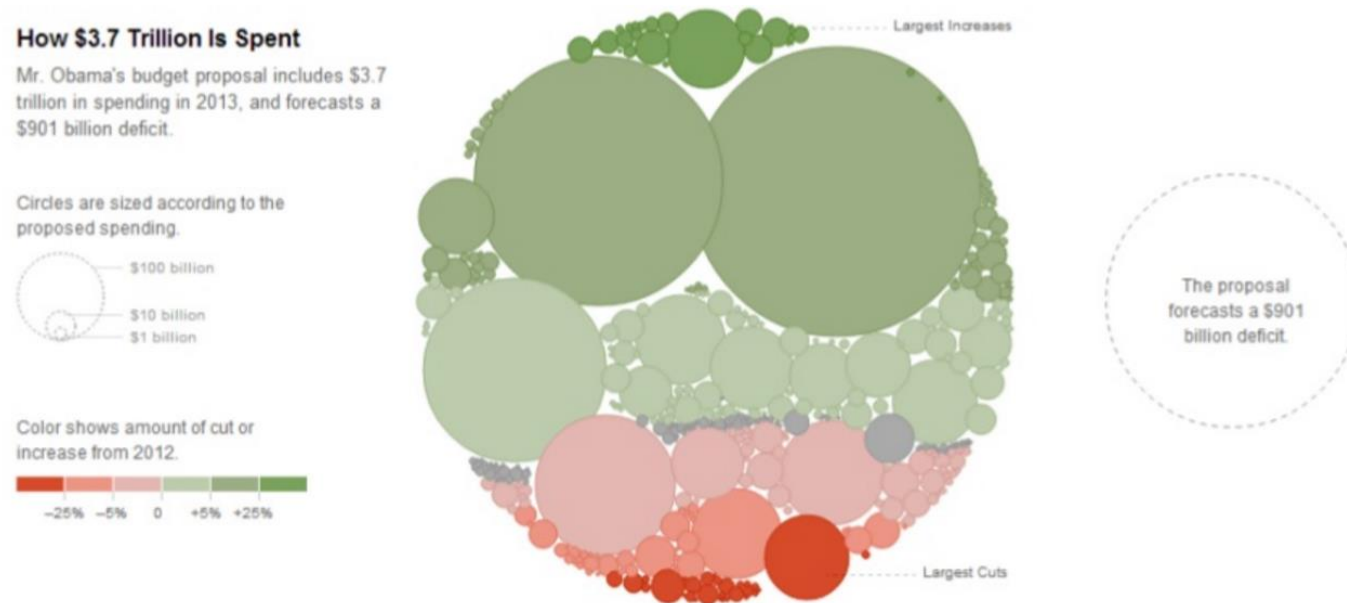


디지털 조선

위의 이미지는 디지털조선의 인포그래픽 사이트다. 디지털조선의 인포그래픽은 국내 뉴스 미디어의 대표적인 사례로 2D 이미지 위주의 인포그래픽을 제공한다. 대부분의 인포그래픽은 기사 텍스트에 보조적인 역할을 할 뿐, 인포그래픽만의 특성을 살린 시각화는 찾아보기 어렵다.

# 제 4절 빅데이터와 시각화 디자인

## 1. 빅데이터와 시각화 이슈



뉴욕타임즈의 시각화 서비스 [http://www.nytimes.com/interactive/2012/02/13/us/politics/2013-budget-proposal-graphic.html?\\_r=2&](http://www.nytimes.com/interactive/2012/02/13/us/politics/2013-budget-proposal-graphic.html?_r=2&)

반면 ‘뉴욕타임즈’의 시각화 서비스는 다양한 인터랙션의 형태를 띄고 있으며, 다양한 각도에서 탐색·분석해 놓은 이 시각화 결과물 자체만으로도 하나의 훌륭한 정보 서비스가 되고 있다.

# 제 4절 빅데이터와 시각화 디자인

## 2. 빅데이터와 시각화 디자인 사례



페이스북의 인기

<http://sweb12.hciresearch.org/content/data-visualization-age-big-data>

‘페이스북의 인기’라는 시각화 작업에서는 데이터 테이블 대신, 페이스북 사용 패턴을 보여주는 시각적 이미지를 제시하고 있다. 빛의 강도는 페이스북의 인기도를 나타낸다. 전 세계의 페이스북 사용현황을 한번에 보여주는 빅픽처 역할을 한다.

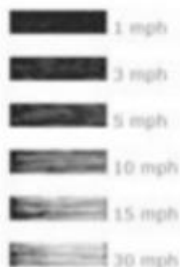
## ■ 모션영상

- 모션 영상에서는 데이터 시각화를 순차적으로 보여주는 영상 또는 자동 애니메이션을 제공하는 경우가 많다. 의도된 스토리에 따라 제공하는 모션 인포그래픽도 있다.

### wind map

Nov. 25, 2013  
6:55 pm EST  
(time of forecast download)

top speed: 27.7 mph  
average: 6.4 mph



윈드 맵

<http://hint.fm/wind/>

- ‘윈드 맵’은 미국 내 풍향 및 풍속 데이터를 받아 지도 위에 시각화한 프로젝트다. 비트맵 애니메이션이 아닌 프로그래밍 기법으로 실제로 바람이 부는 것처럼 표현했다. 사이트에 들어가면 자동으로 재생되며, 미국 지도 위에 풍속의 변화를 볼 수 있다. 미국 국립 디지털 예보 데이터베이스를 이용했다.

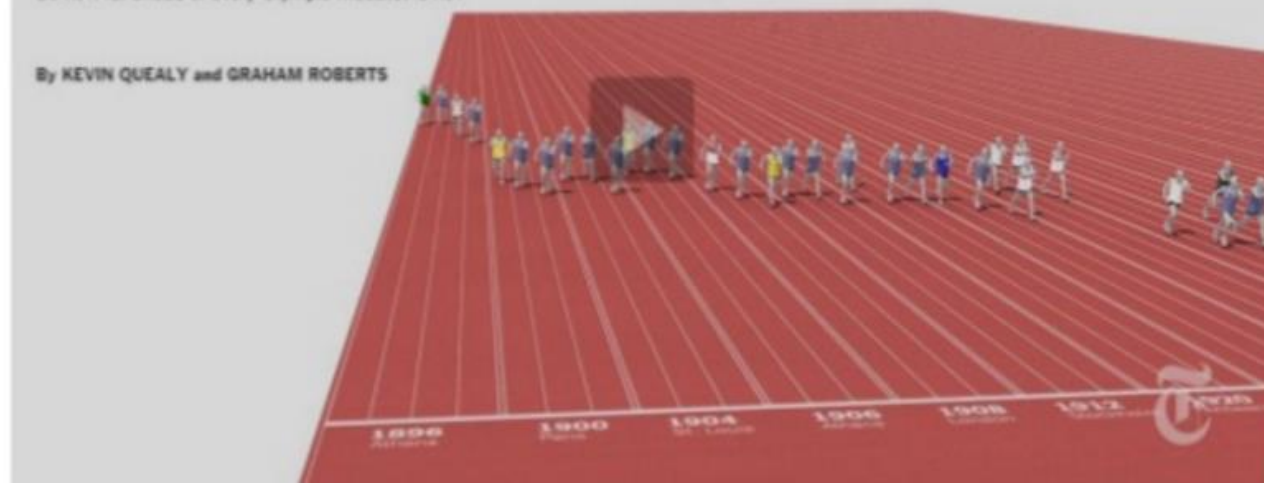


All the Medalists: **Men's 100-Meter Sprint**

## One Race, Every Medalist Ever

Usain Bolt's 9.63 set an Olympic record in the 100.  
So how far ahead of every Olympic medalist is he?

By KEVIN QUEALY and GRAHAM ROBERTS



Sources: "The Complete Book of the Olympics" by David Walsh and John Lucky, International Olympic Committee, Amateur Athletic Association; Photographs: Chang W. Lee/The New York Times.

FACEBOOK TWITTER GOOGLE+ EMAIL SHARE

모든 메달리스트들이 한 경기에

[http://www.nytimes.com/interactive/2012/08/05/sports/olympic-100-meter-dash-one-race-every-medalist-ever.html?\\_r=0](http://www.nytimes.com/interactive/2012/08/05/sports/olympic-100-meter-dash-one-race-every-medalist-ever.html?_r=0)

‘뉴욕타임즈’는 1896년 그리스 아테네 올림픽부터 런던 올림픽에 이르기까지 모든 데이터를 이용해 100미터 육상, 수영, 장애물 넘기를 3D 모델로 구현했다. 우사인 볼트의 9.63초 100미터 육상경기와 그가 어떻게 빨리 먼저 나가는지, 과거의 육상경기 우승자들의 모습과 비교해 시각화한 동영상도 돋보인다는 동영상 제공되는 사이트에서 여러 선수에 대한 정보와 데이터를 비교한 그래프 등을 볼 수 있다.



#### Great British Class Survey 테스트 결과 화면

<http://styleandthensome.wordpress.com/2013/04/07/im-an-emergent-service-worker-and-proud-what-class-are-you/>

영국의 GBCS(Great British Class Survey)는 16만 명의 영국인들에게 설문조사를 한 후 21세기 영국의 새로운 계층을 7가지로 분류한 결과를 BBC의 웹 사이트에서 분류법에 따라 제공한다. 자신이 어느 계층에 속하는지를 알아볼 수 있다. 규모 있는 빅데이터의 분류 결과를 바탕으로 이뤄지는 테스트 과정에서 사용자는 질문을 클릭할 때마다 즉각적으로 변하는 인터랙티브 인포그래픽을 체험하게 된다. 자신의 어떤 성향이 특정 계층으로 분류는지를 바로 확인할 수 있다. 페이스북이나 트위터 같은 SNS로 결과를 공유해 대중의 평가를 유도해 향후 업데이트에 반영할 수 있는 흥미로운 인포그래픽이다.



### GE의 에코매지네이션 아이디어 챌린지 시각화

[http://visualization.geblogs.com/visualization/ecomagination\\_challenge/](http://visualization.geblogs.com/visualization/ecomagination_challenge/)

전 세계에서 6만 5000명 이상의 유저들이 GE의 에코지메이션 챌린지(Ecomagination Challenge)에 로그인해 제출한 **3500개 이상의 아이디어와 투표, 댓글을 시각화한 결과물**이다. 방문자가 좋은 아이디어를 좀 더 쉽게 찾을 수 있도록, GE는 해당 사이트를 시각화해 보여주고 있다.





### 야후! 코어 웹사이트

<http://visualize.yahoo.com/core/>

2012년 2월, 야후는 전 세계 사용자들의 관심 콘텐츠와 트래픽 등 이른바 '빅데이터'를 분석해 시각화해 보여주는 웹사이트인 '야후! C.O.R.E(코어) 데이터 시각화' 사이트를 개설했다. 사용자는 야후의 빅데이터 서비스를 통해 **자신의 행동방식과 기사 검색취향 등을 얘기하고 다른 사람들의 정보도 참고**할 수 있게 된다. 이용자는 다른 이용자와 자신을 밀접하게 연관시켜 열린 의사소통을 체험하게 된다..

[Build a New York ENERGY STAR® Certified Home and save energy year-round. Learn more »](#)

## New York Crimes

[Overview](#)[Market Trends](#)[Schools](#)[Crime](#)[Community Info](#)[Experian: Credit Report & Score \\$1](#)

트룰리아: 뉴욕 범죄 히트맵

[http://www.trulia.com/real\\_estate/New\\_York-New\\_York/crime/](http://www.trulia.com/real_estate/New_York-New_York/crime/)

지역 부동산 실시간 데이터 서비스를 제공하는 트룰리아(trulia.com)는 다양한 데이터를 복합 분석해 부동산 구매를 유도하는 서비스를 제공하고 있다. 트룰리아가 제공하는 통계자료 중 ‘가격 인하’된 자료가 많은 관심을 받고 있다. **시장에 나온 매물 가운데 가격을 한 차례 이상 인하한 매물을 인하 비율별로 분류해 사용자들이 검색할 수 있도록 한 시스템**이다. 그림에 표시된 범죄 데이터는 스팟크라임과 크라임리포트에서 갖고 온다. 트룰리아는 **범죄에 외에도 다양한 시각화 정보를 제공한다.**

# 연습문제

---

문제1. 다음 설명 중 틀리는 것은 무엇인가?

- ① 데이터는 정보를 만들기 위한 원자재와 같은 것이다.
- ② 정보는 생산자와 사용자의 관점에 구애 받지 않고 같게 전달될 수 있으며, 형태와 형식을 갖고 있다.
- ③ 지식은 다른 영역의 정보가 자기 조직화에 획득할 수 있다.
- ④ 지혜는 자기 내면화한 지식으로 명시적인 언어로 상대방에게 전달하기 어렵다.

# 연습문제

---

문제2. 다음 설명 중 일반적인 시각화 프로세스로 적절한 것은?

- ① 시각접근 방법 결정하기 – 와이어프레임 그리기 – 데이터 수집과 읽기 – 내러티브 찾기 – 문제정의하기 – 계층구조 만들기 – 포맷 선택하기 – 정제와 테스트
- ② 데이터 수집과 읽기 – 내러티브 찾기 – 문제 정의하기 – 계층구조 만들기 – 와이어프레임 그리기 – 포맷 선택하기 – 시각적 접근 방법 결정하기 – 정제와 테스트
- ③ 포맷 선택하기 – 와이어프레임 그리기 – 정제와 테스트 – 데이터 수집과 읽기 – 내러티브 찾기 – 문제 정의하기 – 계층구조 만들기 – 시각접근 방법 결정하기
- ④ 문제 정의하기 – 와이어프레임 그리기 – 시각접근 방법 결정하기 – 데이터 수집과 읽기 – 내러티브 찾기 – 계층구조 만들기 – 포맷 선택하기 – 정제와 테스트

# 연습문제

---

문제 3. 리처드 솔 워먼이 제시한 정보를 조직화하는 LATCH에 해당하지 않은 것은 무엇인가?

- ① Category
- ② Location
- ③ Hour
- ④ Alphabet

# 연습문제

문제 4. 다음의 그래프가 속하는 네이션 야우의 정보 시각화 방법의 종류는 어느 것인가?

- ① 공간시각화
- ② 관계시각화
- ③ 시간시각화
- ④ 분포시각화



뉴스맵

<http://newsmap.jp/>



# 연습문제

---

문제 5. 정보 시각 표현을 위한 자크 배르탱의 그래픽 7요소가 아닌 것은?

- ① 기울기
- ② 그리드
- ③ 명도
- ④ 모양

# 연습문제

---

문제 6. 에드워드 터프티가 제시한 시각 정보 디자인 7원칙에 해당하지 않은 것은 무엇인가?

- ① 공간 순이 아닌 시간순으로 나열하라
- ② 콘텐츠의 질과 연관성, 진실성을 분명히 하라
- ③ 텍스트, 그래픽, 데이터를 한 화면에 조화롭게 배치하라
- ④ 원인과 결과를 명쾌하게 제시하라.