

# 데이터 시각화

## 05. 시각적 인식

최대영 교수



**고려사이버대학교**  
THE CYBER UNIVERSITY OF KOREA



Data Visualization

데이터 시각화

# 시각적 인식

05주차

최대영 교수



1

## 학습리뷰

### 1 시각화 과업과 과업 끌어내기

#### 📌 시각화 과업

- 데이터를 시각적 표현으로 인코딩하여 사용자의 목적을 달성하기 위한 활동
- 시각화는 단순히 데이터를 시각적으로 표현에 그치는 것이 아니라 사용자가 시각적 표현을 이용하여 하고자 하는 과업을 포함

#### 📌 과업 끌어내기

- 데이터 시각화 개발자가 사용자로부터 과업에 대한 정보를 얻어내는 것
- 목적, 수단, 데이터의 특성, 목표 데이터와 범위, 작업 흐름, 역할

2

## 2 디자인 방법

### ✍ Five Design Sheet 방법

- 5장의 sheet를 만들어 이해관계자와 긴밀히 협력하여 시각화 모형을 만드는 방법

### ✍ Design Study Methodology

- 전제조건, 핵심, 분석의 3개의 범주로 구분
- 9단계로 구분되며 순서대로 진행되고 이전 단계로 돌아가 반복 가능

3

## 3 실습(파이썬으로 데이터 다루기)

### ✍ pandas

- 행과 열로 이루어진 데이터를 처리하는데 주로 쓰이는 라이브러리

### ✍ NumPy

- 다차원 배열의 연산, 정렬 등을 처리하는데 주로 쓰이는 라이브러리

4

## ● 학습목표

- 시각적 인식과 사전 주의에 대해 설명할 수 있다.
- 시각적 주의와 검색에 대해 설명할 수 있다.
- matplotlib의 기본 문법을 이해하고 활용할 수 있다.

5

## ● 학습내용

- 1 시각적 인식과 사전 주의
- 2 시각적 주의와 검색
- 3 실습(matplotlib 기본 문법)

6



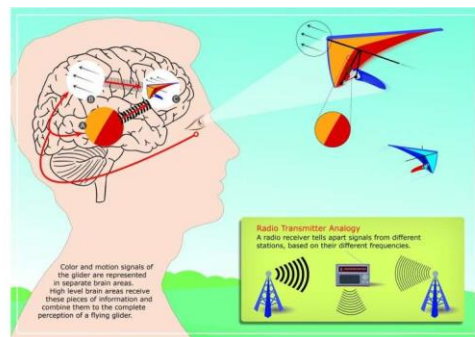
# 시각적 인식과 사전 주의

7

## 1. 시각적 인식

### ≡ 시각적 인식(Visual Perception)

사물에 반사된 가시 스펙트럼을 이용하는 시각을 통해  
주변 환경을 해석하는 능력



[출처] MedicalXpress

8

## 1. 시각적 인식

### Anscombe's quartet(앤스컴 콰르텟)

4개의 데이터셋은 같은 통계적 특성(평균, 분산, 상관계수 등)을 가짐

데이터셋 1

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$

데이터셋 2

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$

데이터셋 3

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$

데이터셋 4

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$

[출처] 위키피디아

9

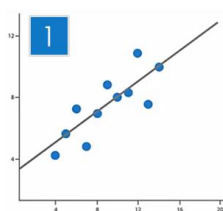
## 1. 시각적 인식

### Anscombe's quartet(앤스컴 콰르텟)

그러나, 시각화하면 매우 다른 특성의 데이터라는 것을 확인할 수 있음

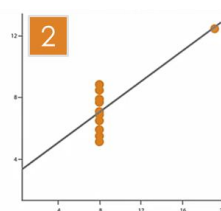
데이터셋 1

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$



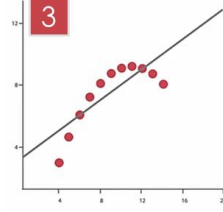
데이터셋 2

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$



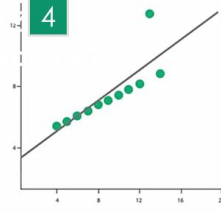
데이터셋 3

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$



데이터셋 4

특성	값
x의 평균	9.0
x의 분산	11.0
y의 평균	7.5
y의 분산	4.12
상관계수	0.816
선형회귀	$y = 0.5x + 3$



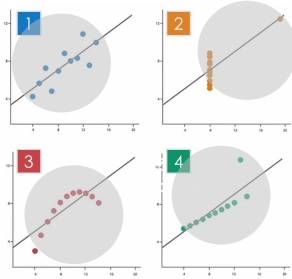
[출처] 위키피디아

10

## 1. 시각적 인식

### 시각 신호 입력부터 통찰을 얻기까지의 과정

☞ 빛을 이용해서 4가지 다른 색깔과 점의 위치를 인식

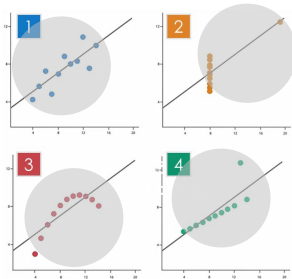


11

## 1. 시각적 인식

### 시각 신호 입력부터 통찰을 얻기까지의 과정

☞ 일정 시간을 들여 뇌에서 시각 정보를 처리하여 통찰을 얻어냄



데이터셋의 통계치는  
같지만 데이터셋의  
특징은 다름

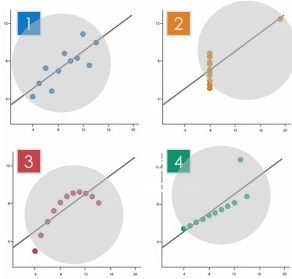


12

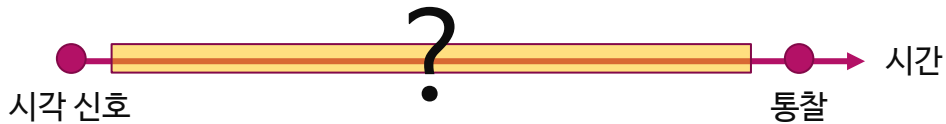
## 1. 시각적 인식

### 시각 신호 입력부터 통찰을 얻기까지의 과정

시각 정보를 받아들여 통찰까지 이어지는 **인식 과정**에 대한 이해 필요



데이터셋의 통계치는  
같지만 데이터셋의  
특징은 다름



13

## 1. 시각적 인식

### 시각적 인식의 중요성

사람이 **데이터를 해석하는 메커니즘(mechanism)**을 설명

- 메커니즘을 이해하면 사람에게 보여주고 싶은 시각화를 생성할 수 있음

시각화를 생성하기 위한 **예측(predictive)적인 가이드라인** 제공

- 사람이 특정 디자인을 어떻게 해석할지 예측할 수 있음

시각화를 위한 **정량적인 모델과 개념적인 모델** 생성

- 간단한 과업과 디자인 → 정량적인 모델
- 복잡한 과업과 디자인 → 개념적인 모델

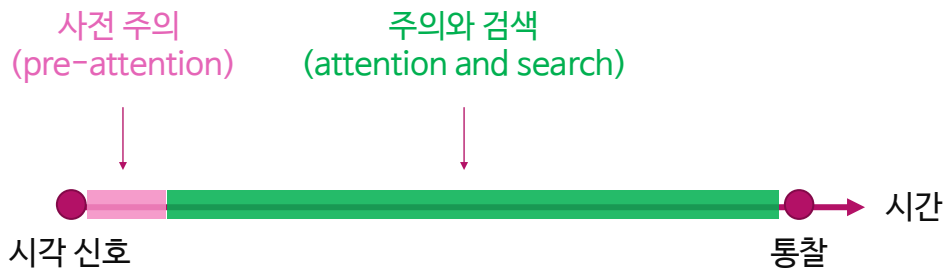
14



## 1. 시각적 인식

### 시각 신호 입력부터 통찰을 얻기까지의 과정

시각 정보를 받아들여 통찰까지 이어지는 **인식 과정**에 대한 이해 필요



15

## 2. 사전 주의

### 시각화 첫눈에 보기

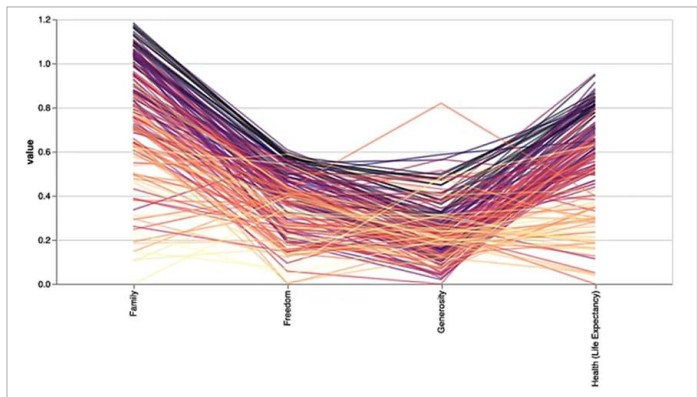
요지(Gist)

- 시각화를 볼 때 첫눈에 얻어지는 특징적인, 의미상의 정보

2가지 가장 강한 특징

Spatial Envelop(윤곽)

Ensemble Codes(전체)



[출처] Fundamentals of Data Visualization, University of Colorado Boulder

16

## 2. 사전 주의

### 시각화 첫눈에 보기

#### 요지(Gist)

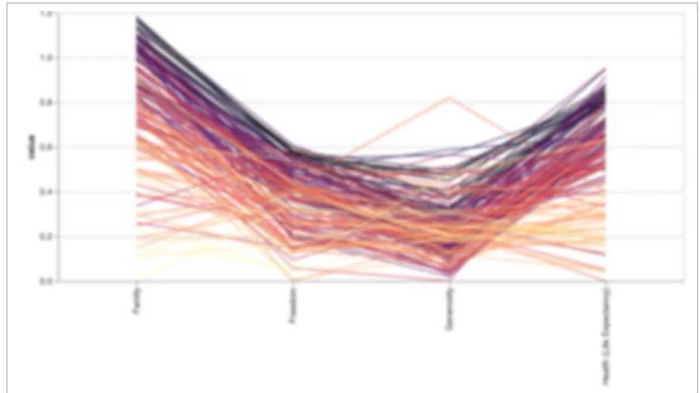
- 시각화를 볼 때 첫눈에 얻어지는 특징적인, 의미상의 정보

#### 2가지 가장 강한 특징

Spatial Envelop(윤곽)

Ensemble Codes(전체)

#### 시각화를 흐리게(blur)하여 gist 확인

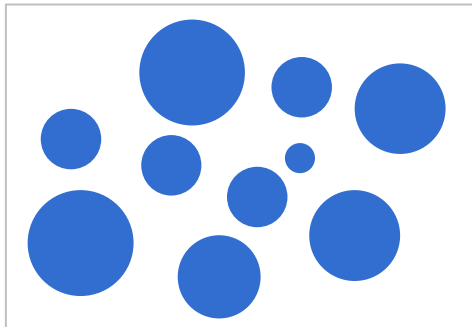


[출처] Fundamentals of Data Visualization, University of Colorado Boulder

17

## 2. 사전 주의

### Ensemble Coding



[출처] Fundamentals of Data Visualization, University of Colorado Boulder

18

## 2. 사전 주의

### Ensemble Coding

☞ A와 B 중 이전 그림에 있었던 원은?



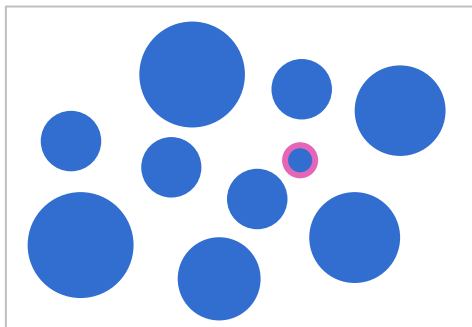
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

19

## 2. 사전 주의

### Ensemble Coding

☞ B를 선택하는 이유: 대략적으로 원의 평균 크기와 같음



[출처] Fundamentals of Data Visualization, University of Colorado Boulder

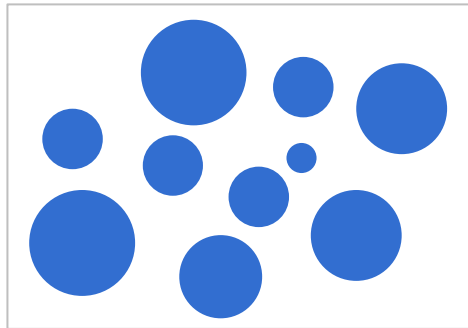
20

## 2. 사전 주의

### Ensemble Coding

☞ B를 선택하는 이유: 대략적으로 원의 평균 크기와 같음

- Plot을 힐끗 보면 전체의 평균적인 특징을 파악
- 평균적인 크기, 크기의 변화, 전체적인 분포



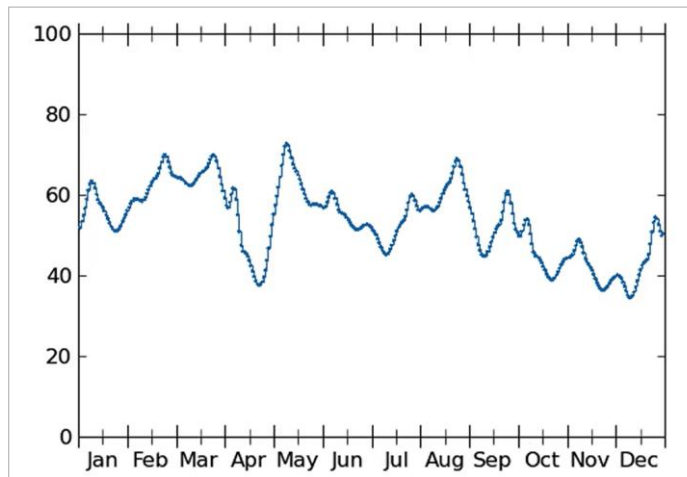
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

21

## 2. 사전 주의

### Encoding 방법과 인식

☞ 가장 판매가 많은 날이 있는 달은?



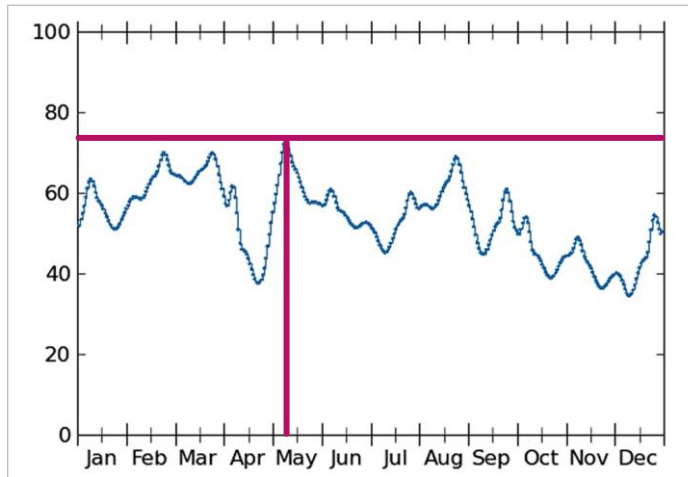
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

22

## 2. 사전 주의

### Encoding 방법과 인식

가장 판매가 많은 날이 있는 달은?



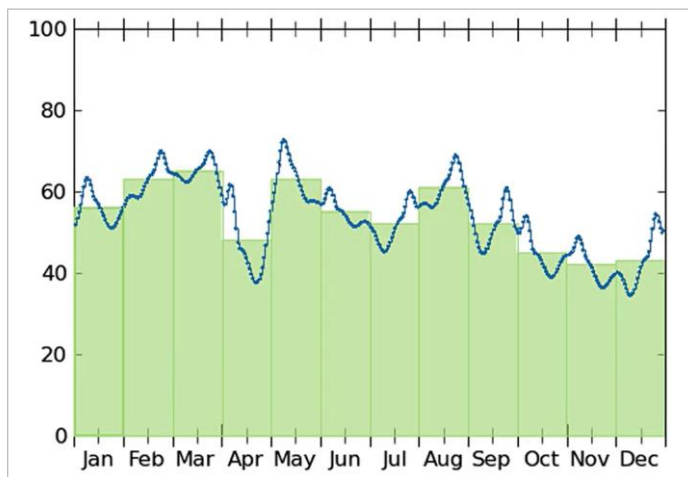
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

23

## 2. 사전 주의

### Encoding 방법과 인식

평균 판매가 가장 많은 달은?



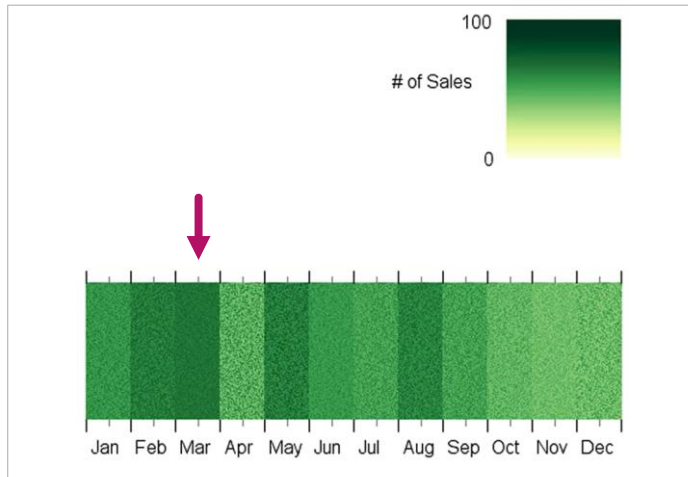
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

24

## 2. 사전 주의

### Encoding 방법과 인식

☞ 평균 판매가 가장 많은 달은?



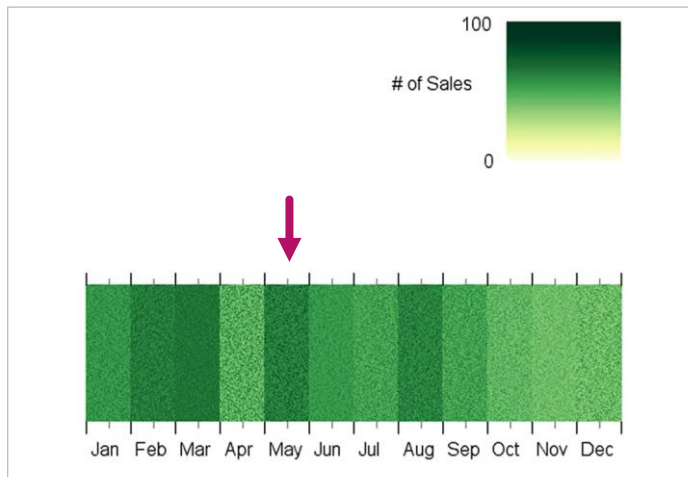
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

25

## 2. 사전 주의

### Encoding 방법과 인식

☞ 가장 판매가 많은 날이 있는 달은?



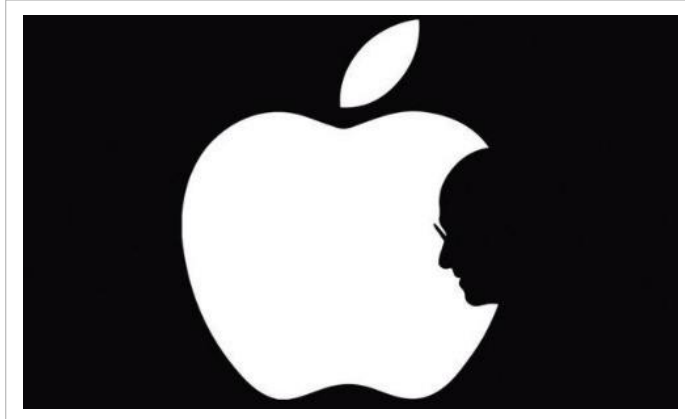
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

26

## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

📁 Figure-Ground: 강조 요소(figure)와 배경(background)을 구분하여 인식



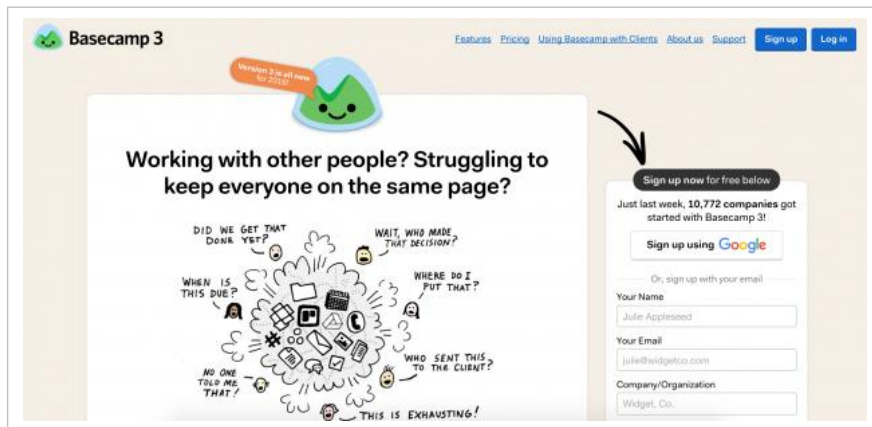
[출처] A Dwarf Named Warren

27

## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

📁 Figure-Ground: 강조 요소(figure)와 배경(background)을 구분하여 인식



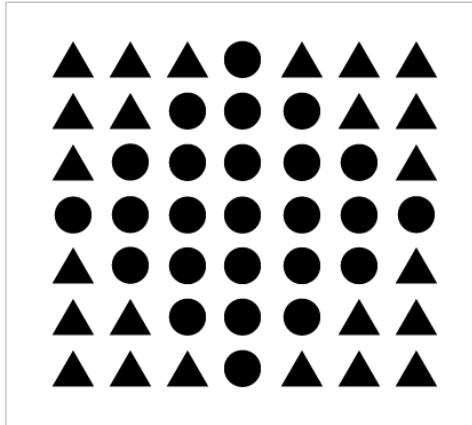
[출처] User Testing

28

## 2. 사전 주의

### 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Similarity: 비슷한(similar) 시각화 요소를 그룹으로 인식



[출처] Creative Beacon

29

## 2. 사전 주의

### 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Similarity: 비슷한(similar) 시각화 요소를 그룹으로 인식



[출처] Andy Rutledge

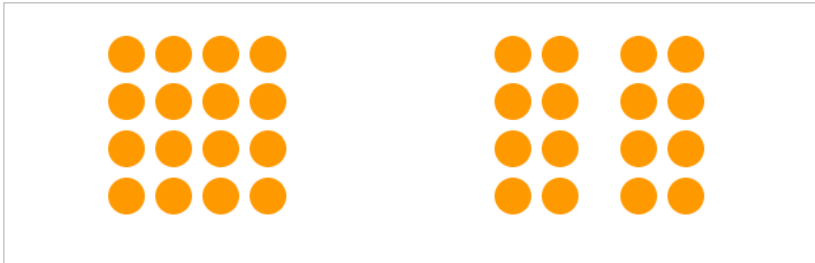
30



## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Proximity(근접): 가까이 있는 시각적 요소를 더 관련 있다고 인식



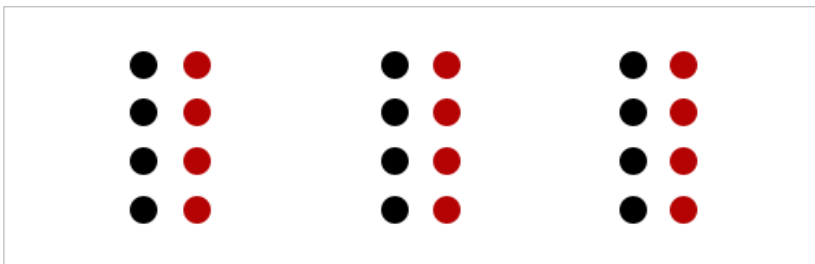
[출처] Andy Rutledge

31

## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Proximity(근접): 가까이 있는 시각적 요소를 더 관련 있다고 인식



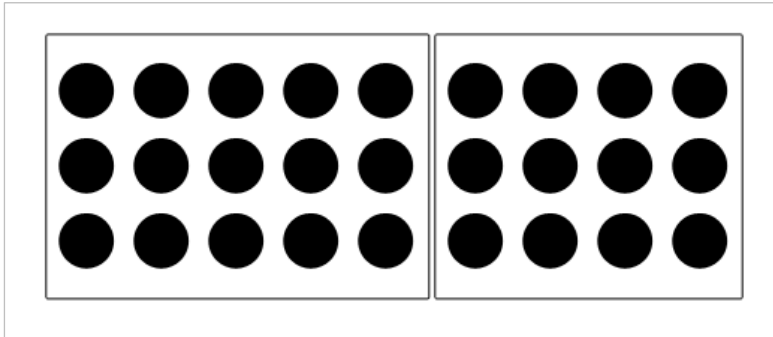
[출처] Steven Bradley

32

## 2. 사전 주의

### 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Common Region: 같은 닫힌 공간에 있는 시각적 요소를 그룹으로 인식



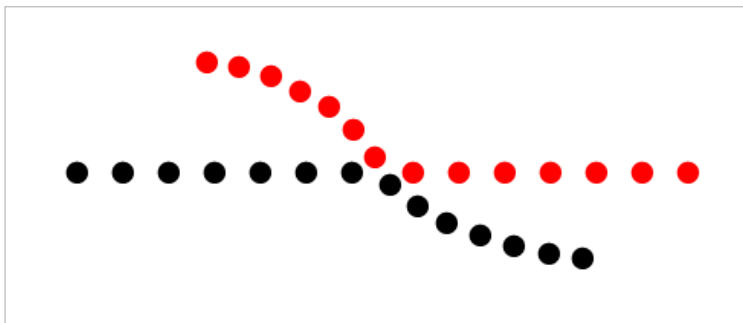
[출처] Smashing Magazine

33

## 2. 사전 주의

### 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Continuity: 선을 따라 정렬되어 있는 시각적 요소를 더 관련 있는 것으로 인식



[출처] Smashing Magazine

34

## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Closure: 복잡하게 시각적 요소가 나열되어 있을 때 단일한 패턴으로 인식



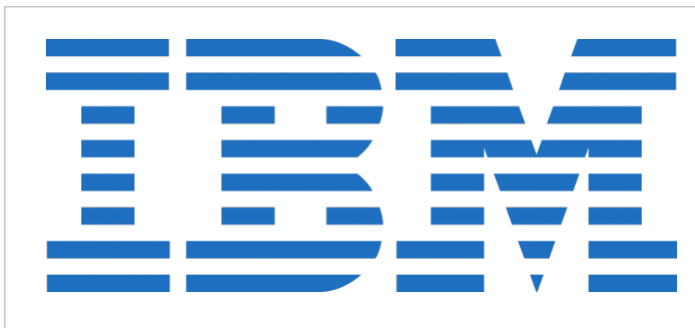
[출처] Eduard Volianskyi

35

## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

☞ Closure: 복잡하게 시각적 요소가 나열되어 있을 때 단일한 패턴으로 인식

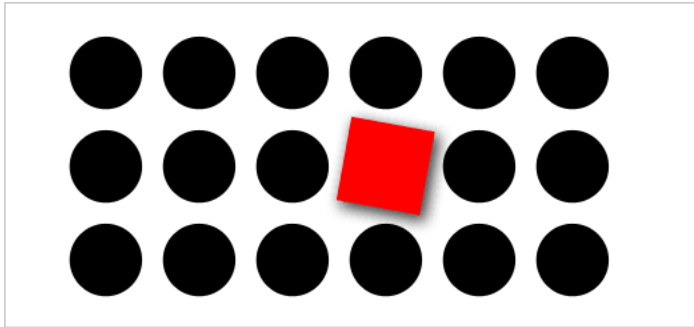


36

## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

📌 Pop-out(focal point): 가장 두드러진 시각적 요소가 가장 먼저 주의를 끄



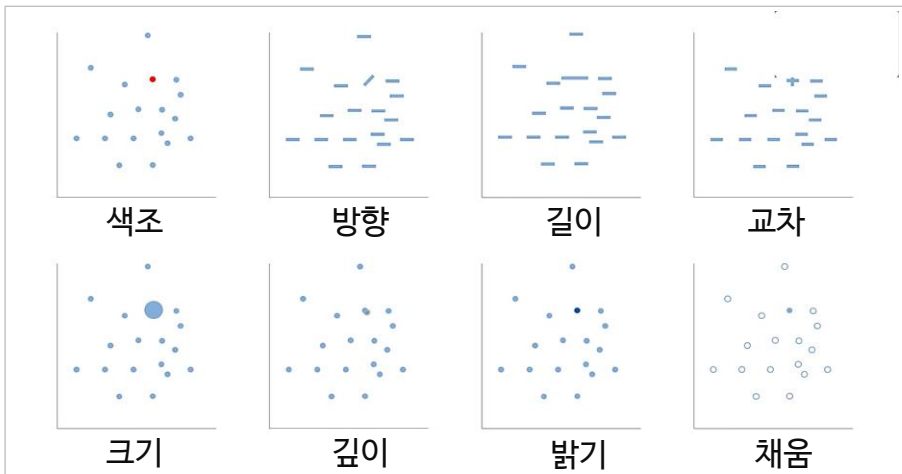
[출처] Smashing Magazine

37

## 2. 사전 주의

### ≡ 게슈탈트 원칙(Gestalt Principles of Visual Perception)

📌 Pop-out(focal point): 가장 두드러진 시각적 요소가 가장 먼저 주의를 끄



[출처] Fundamentals of Data Visualization, University of Colorado Boulder

38



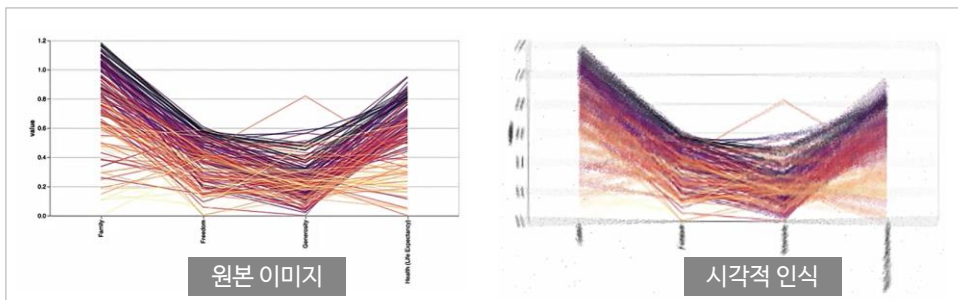
# 시각적 주의와 검색

39

## 1. 시각적 주의

### ≡ 시각적 주의(Visual Attention)란?

- ☞ 간단히 정의하면 ‘사람이 보고 있는 곳’
- ☞ 사람은 시야에 들어온 모든 부분을 선명하다고 느끼지만 이는 일부에 그침
  - 중심부: 실제로 집중하여 해상도가 높은 정보를 얻음
  - 주변부: 시각화에서 대략적이고 흐릿한 정보를 얻음



[출처] Fundamentals of Data Visualization, University of Colorado Boulder

40

## 1. 시각적 주의

### 시각적 주의 설계 방법

#### Top Down 방법

- 지식(knowledge)이 사람이 보고자 하는 것을 안내
- 사람이 관심있어 하는 정보를 유심히 관찰하여 찾을 때 사용

#### Bottom Up 방법

- 시각화의 색, 모양, 크기 등의 특징이 사람이 보고자 하는 것을 안내
- 데이터를 탐색적으로 분석하는 과정에서 주로 사용

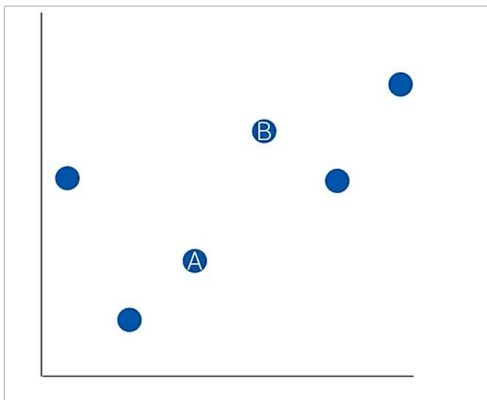
41

## 1. 시각적 주의

### Top Down 방법의 예시

☞ B가 A보다 대략적으로 얼마나 클까?

- **지식**이 사람이 보고자 하는 것을 안내 → A와 B 찾기, A와 B 비교하기



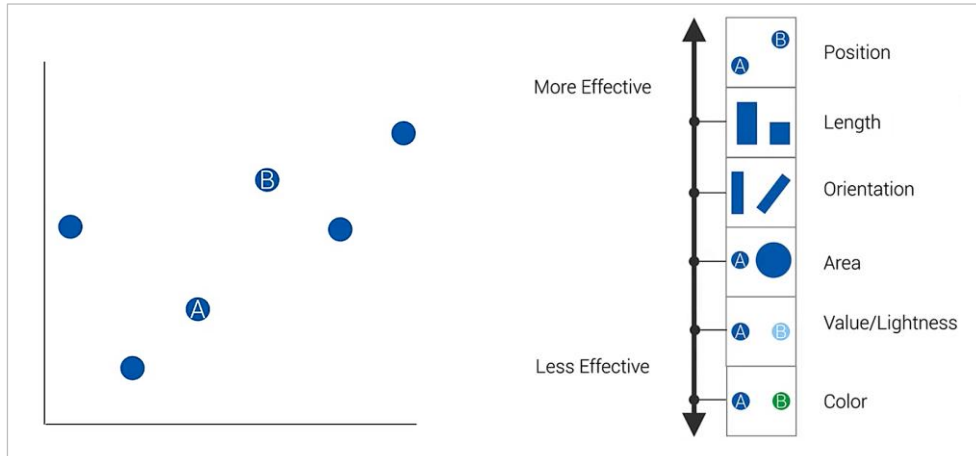
[출처1] Cleveland and McGill, Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods, 1984  
[출처2] Fundamentals of Data Visualization, University of Colorado Boulder

42

## 1. 시각적 주의

### Top Down 방법의 예시

대답을 기준으로 어떤 시각화가 더 효율적인지 순위를 매김



[출처1] Cleveland and McGill, Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods, 1984  
[출처2] Fundamentals of Data Visualization, University of Colorado Boulder

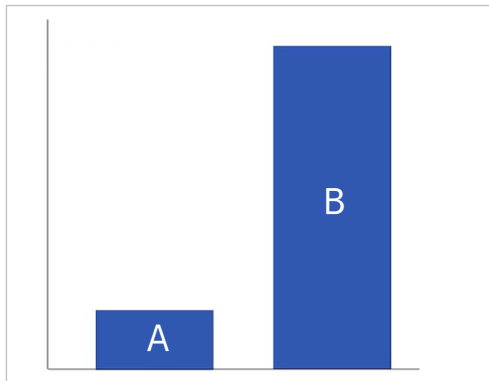
43

## 1. 시각적 주의

### Bottom Up 방법의 예시

보고 있는 그래프를 설명하세요.

- 시각화 요소의 특징이 사람이 보고자 하는 것을 안내 → 막대, 선



[출처1] Zacks and Tversky, Bars and lines: A study of graphic communication, 1999  
[출처2] Fundamentals of Data Visualization, University of Colorado Boulder

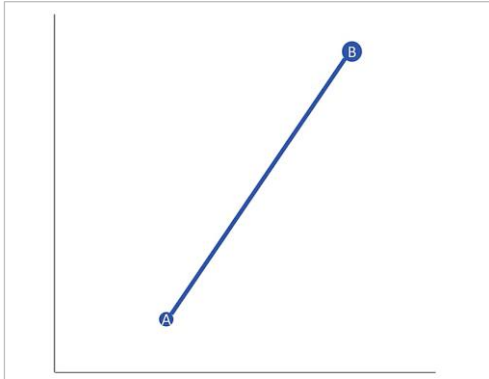
44

## 1. 시각적 주의

### Bottom Up 방법의 예시

☞ 보고 있는 그래프를 설명하세요.

- 시각화 요소의 특징이 사람이 보고자 하는 것을 안내 → 막대, 선



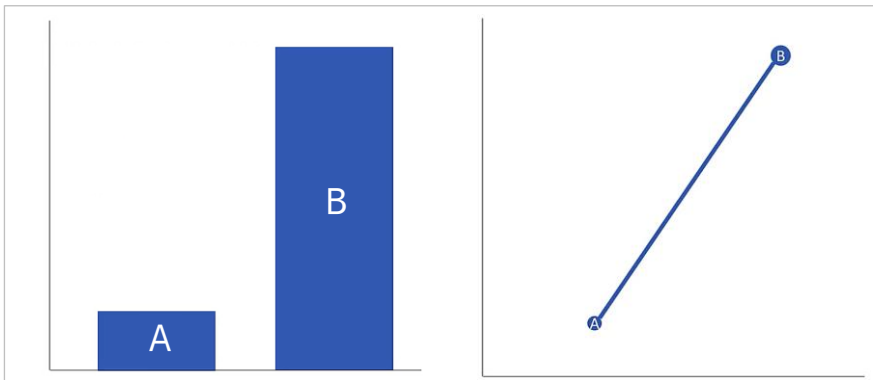
[출처1] Zacks and Tversky, Bars and lines: A study of graphic communication, 1999  
[출처2] Fundamentals of Data Visualization, University of Colorado Boulder

45

## 1. 시각적 주의

### Bottom Up 방법의 예시

☞ 데이터를 어떻게 표현하는가가 결론을 유도



[출처1] Zacks and Tversky, Bars and lines: A study of graphic communication, 1999  
[출처2] Fundamentals of Data Visualization, University of Colorado Boulder

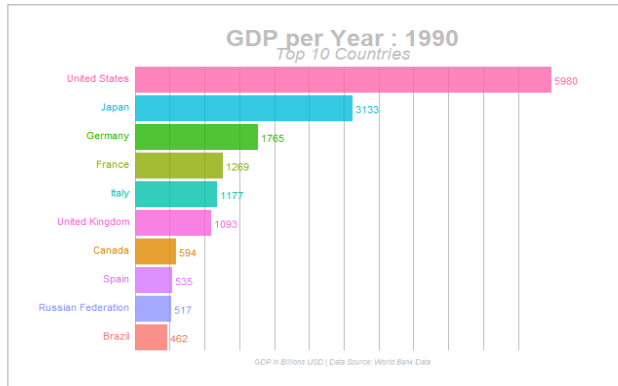
46



## 1. 시각적 주의

### ≡ 변화 맹시(Change Blindness)

- ☞ 연속 제시되는 장면들에서 어느 한 부분의 변화가 있음에도 불구하고 그 변화를 탐지해 내기 어려운 현상



[출처] R-bloggers

47

## 1. 시각적 주의

### ≡ 변화 맹시(Change Blindness)

- ☞ 선택적 주의 테스트(Selective Attention Test)



[출처] Daniel Simons, 2010.03.10

48

## 1. 시각적 주의

### ≡ 변화 맹시(Change Blindness)

- ☞ 사람은 데이터가 시간에 따라 변화하는 방식(animation)보다 공간에 고정되어 있는 방식을 비교하는 것을 더 쉽게 생각
  - 사람의 기억(memory)과 주의를 제한된 자원
- ☞ 애니메이션 시각화 방법을 사용할 때는 주의를 기울여야 함
  - 하나의 시각화에 적은 수의 정보를 담도록 설계

49

## 2. 시각적 검색

### ≡ 시각적 검색(Visual Search)란?

흥미로운 것을 찾기 위해 시각적으로 주의를 기울여  
유심히 살펴보는 것

- 시각화 요소의 현저함(saliency)이 클수록 더 주의를 끌어냄 → 검색이 쉬움



[출처] BuzzFeed News

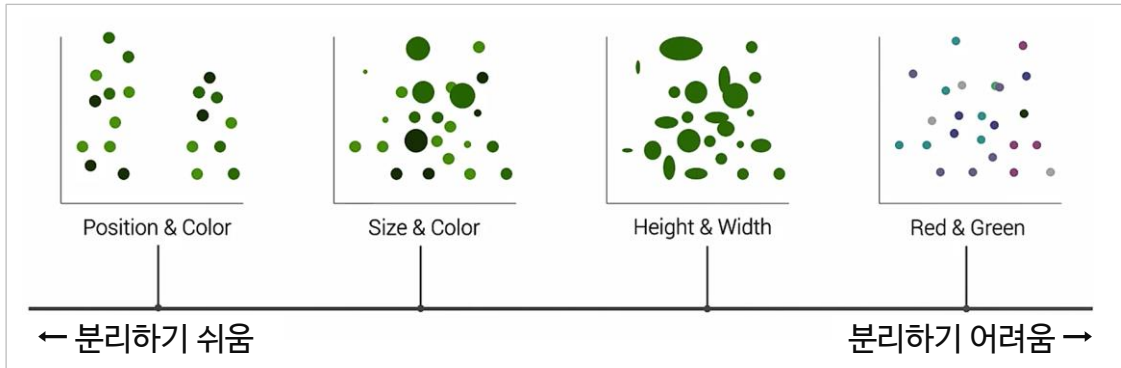
50

## 2. 시각적 검색

### ≡ 분리가능성(Separability)

☞ 여러 시각화 요소가 섞여 있을 때 분리를 위한 시각적 검색이 일어남

- 분리가 쉬운(분리가능성이 높음) 시각화 요소부터 한번에 하나씩 분리



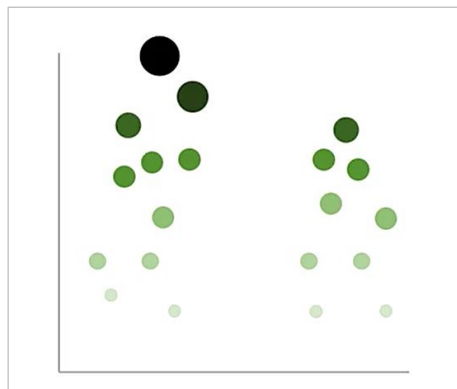
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

51

## 2. 시각적 검색

### ≡ 중복 인코딩(Redundant Encoding)

☞ 검색을 돕기 위해 하나의 데이터를 여러가지 방법으로 인코딩



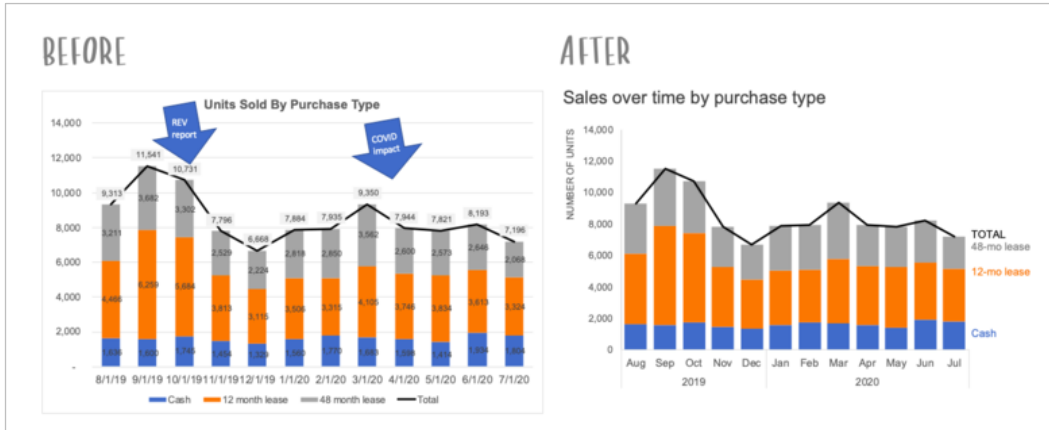
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

52

## 2. 시각적 검색

### 시각적 어수선했(Visual Clutter)

아이템, 표현방법, 정렬방법이 많거나 복잡하면 인식의 효율이 떨어짐



[출처] storytelling with data

53

## 2. 시각적 검색

### 데이터-잉크 비율(Data-Ink Ratio)

가능한 적은 양의 픽셀로 가능한 많은 양의 데이터를 표현

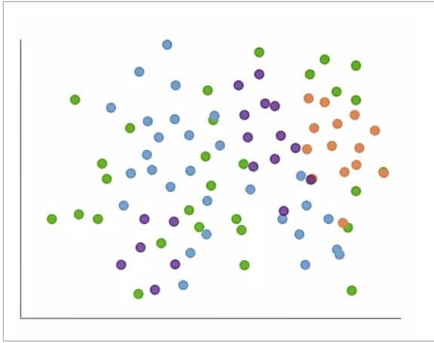
$$\text{데이터 잉크 비율} = \frac{\text{시각화한 데이터}}{\text{사용된 픽셀수}} \rightarrow \text{최대화}$$

54

## 2. 시각적 검색

### ≡ 종합(Aggregation)

☞ 데이터를 그룹지어 좀 더 간결하게 시각화하는 방법은?



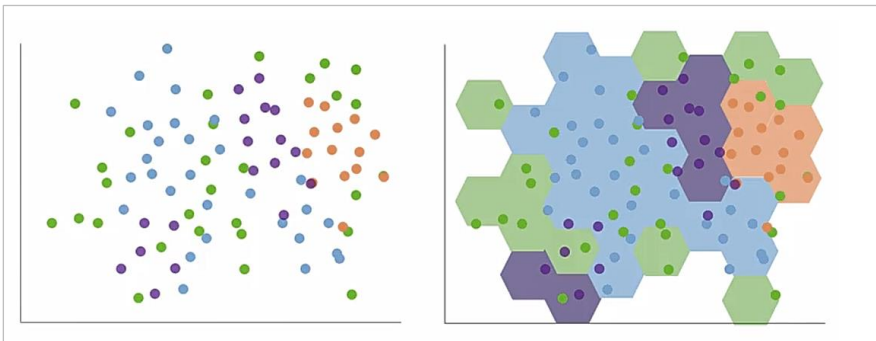
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

55

## 2. 시각적 검색

### ≡ 종합(Aggregation)

☞ 데이터를 육각형의 구역으로 나눠 가장 많은 점의 색으로 색칠



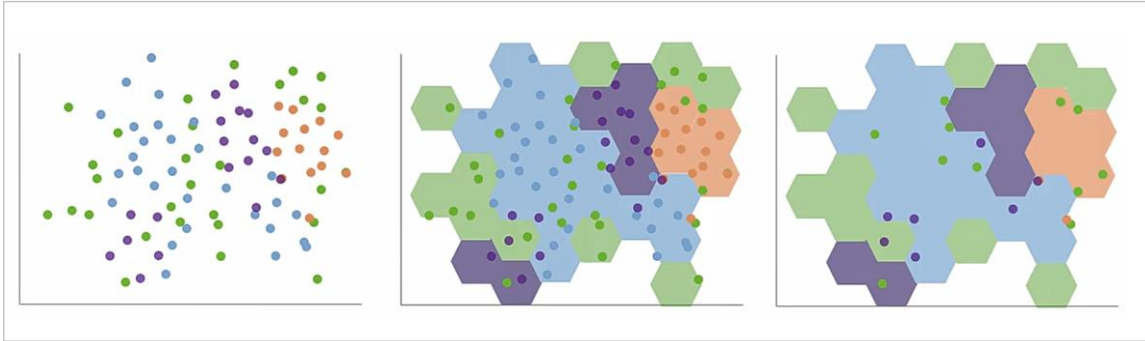
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

56

## 2. 시각적 검색

### ≡ 종합(Aggregation)

☞ 이상치(outlier)를 제외하고는 원래 점을 제거하여 간략화



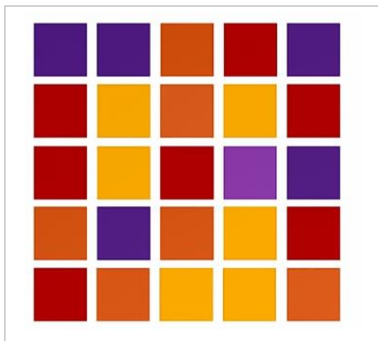
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

57

## 2. 시각적 검색

### ≡ 분류(Sorting)와 정리(Organization)

☞ 히트맵(heatmap)에서 한가지 색만 있는 사각형은?



[출처] Fundamentals of Data Visualization, University of Colorado Boulder

58

## 2. 시각적 검색

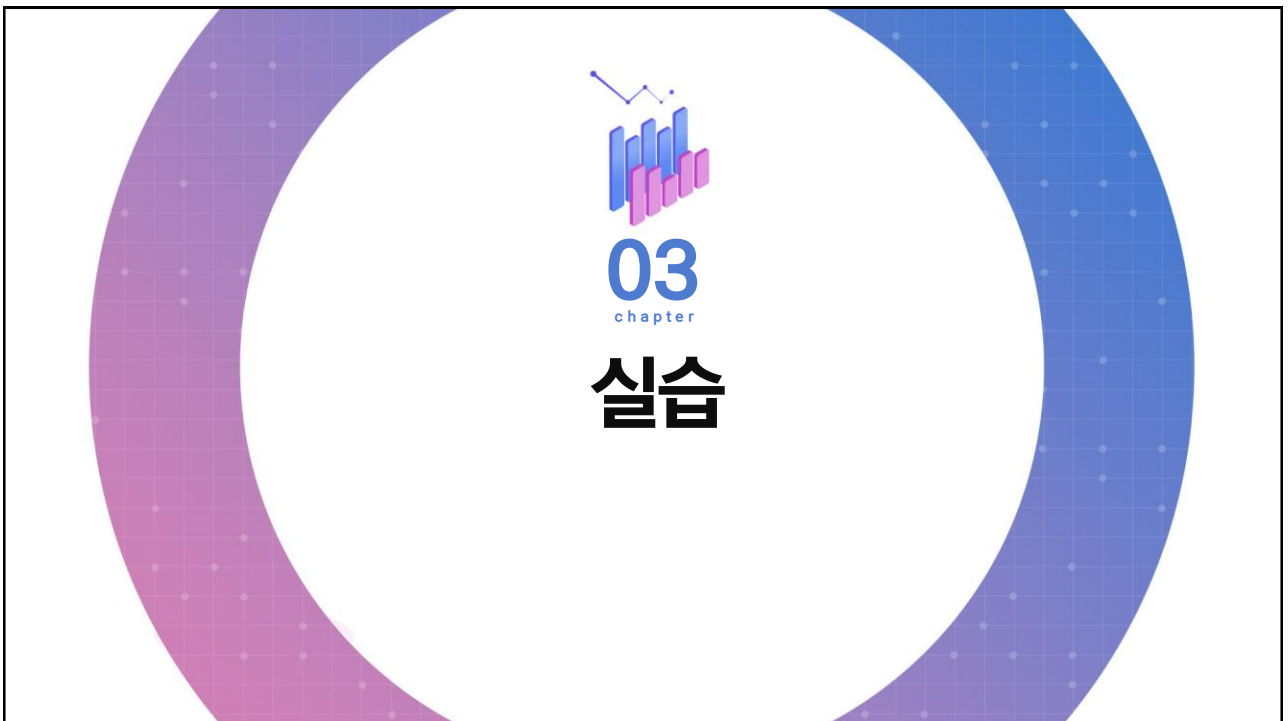
### ≡ 분류(Sorting)와 정리(Organization)

📌 히트맵(heatmap)에서 한가지 색만 있는 사각형은?



[출처] Fundamentals of Data Visualization, University of Colorado Boulder

59

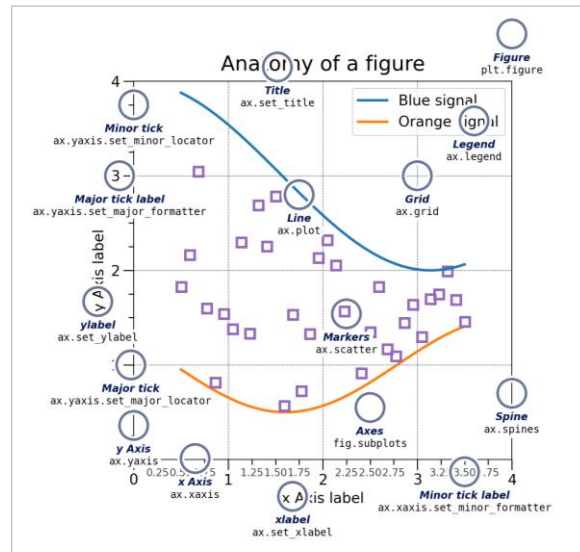


60

## 1. matplotlib 기본 문법

### matplotlib

- ☞ 파이썬에서 가장 많이 사용되는 시각화 라이브러리
- ☞ 정적, 동적, 상호작용이 있는 시각화 도표 제작 가능
- ☞ 다양한 파일형식으로 변환 가능



[출처] matplotlib

61

## ● 학습정리

### 1 시각적 인식과 사전 주의

#### ☞ 시각적 인식

- 사물에 반사된 가시 스펙트럼을 이용하는 시각을 통해 주변 환경을 해석하는 능력

#### ☞ 시각적 인식의 중요성

- 사람이 데이터를 해석하는 메커니즘(mechanism)을 설명
- 시각화를 생성하기 위한 예측(predictive)적인 가이드라인 제공
- 시각화를 위한 정량적인 모델과 개념적인 모델 생성

62



## ● 학습정리

### 1 시각적 인식과 사전 주의

#### 📌 요지(Gist)

- 시각화를 볼 때 첫눈에 얻어지는 특징적인, 의미상의 정보
- Spatial Envelop(윤곽), Ensemble Codes(전체)

#### 📌 게슈탈트 원칙(Gestalt principles of visual perception)

- Figure-Ground, Similarity, Proximity, Common Region, Continuity, Closure, Pop-out

63

## ● 학습정리

### 2 시각적 주의와 검색

#### 📌 시각적 주의(Visual Attention)

- 간단히 정의하면 ‘사람이 보고 있는 곳’
- 사람은 시야에 들어온 모든 부분을 선명하다고 느끼지만 일부에 그침

#### 📌 시각적 주의 설계 방법

- Top Down 방법: 지식이 사람이 보고자 하는 것을 안내
- Bottom Up 방법: 시각화의 색조, 모양, 크기 등의 특징이 사람이 보고자 하는 것을 안내

64

## ● 학습정리

### 2 시각적 주의와 검색

#### 🔗 시각적 검색(Visual Search)

- 흥미로운 것을 찾기 위해 시각적으로 주의를 기울여 유심히 살펴보는 것

#### 🔗 분리가능성(Separability)과 중복 인코딩(Redundant Encoding)

- 분리가능성: 분리가 쉬운 시각화 요소부터 한번에 하나씩 분리
- 중복 인코딩: 검색을 돕기 위해 하나의 데이터를 여러가지 방법으로 인코딩

#### 🔗 시각적 어수선했음(Visual Clutter)

- 데이터 잉크 비율(Data-Ink Ratio)을 최대화
- 종합(Aggregation), 분류(Sorting)와 정리(Organization)로 극복

65

## ● 참고문헌

- 📖 「Fundamentals of Data Visualization」, Danielle Albers Szair, University of Colorado Boulder(coursera course).
- 📖 「Attention and Visual Memory in Visualization and Computer Graphics」, Healey et al., IEEE Transactions on Visualization and Computer Graphics, 2012.

※ 서체 출처 | 넥슨Lv2고딕-(넥슨코리아)www.levelup.nexon.com / 나눔바른고딕(네이버)

66

## 저작권 안내

이 강의록은 저작권법에 의해 보호받는 저작물로서  
저작권자의 허락 없이 저작재산권 일체(복제권,  
배포권, 대여권, 공연권, 공중전송권, 전시권, 2차적  
저작물 작성권)를 침해 시 저작권법에 의거 처벌받을  
수 있습니다.