

데이터 시각화

02. 데이터와 시각적 속성

최대영 교수



고려사이버대학교
THE CYBER UNIVERSITY OF KOREA



1

학습리뷰

1 데이터 분석과 시각화

시각화의 목적

- 인간의 인지를 확장시켜 통찰을 제공 (못 보던 것을 보게 함)
- 데이터 분석의 결과와 결론을 스토리로 전달하여 설득

시각화의 해석

- 어떤 데이터인가?
- 왜 시각화 하였는가?
- 시각화가 어떻게 설계되었는가?

2

1 데이터 분석과 시각화

📌 시각화의 특성

- 데이터의 특성, 시각화 목적 및 기술을 종합적으로 고려하여 다양한 방법으로 표현한 후 최적의 시각화 방법을 사람이 선택

2 그래픽 문법

📌 그래픽 문법이란?

- 시각화와 관련된 사항을 구조적으로 논의하기 위한 문법(언어, 틀)
- 시각화가 가진 공통의 구성요소(데이터, 시각적 속성 등)를 표현

📌 그래픽 문법의 7가지 요소

- | | |
|--------------------------|------------------------------|
| ■ 데이터(data) | ■ 시각적 속성(aesthetics) |
| ■ 스케일(scale) | ■ 기하학적 객체(geometric objects) |
| ■ 통계치(statistics) | ■ 면분할(facets) |
| ■ 좌표계(coordinate system) | |

● 학습목표

- 시각화 맵핑(mapping)에 대해 설명할 수 있다.
- 시각화의 요소 중 색에 대해 설명할 수 있다.
- 시각화의 요소 중 좌표계와 축에 대해 설명할 수 있다.

5

● 학습내용

- 1 시각화 맵핑
- 2 색
- 3 좌표계와 축

6



시각화 맵핑

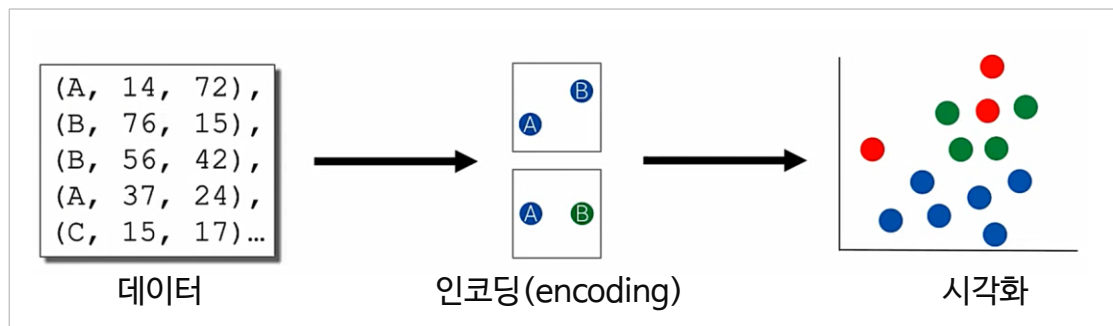
7

1. 시각화 맵핑

≡ 시각화 맵핑(mapping)이란?

☞ 맵핑(mapping): 하나의 공간에서 다른 공간으로 변환(translate)하는 것

☞ 시각화 맵핑: 데이터 공간(data space)에서 시각적 공간(visual space)으로 변환



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

8

1. 시각화 맵핑

≡ 데이터 공간(data space)

📁 데이터셋은 **차원(dimension, attribute)**으로 구성됨

차원

↓

성명	연봉(만원)	나이	야근시간	직급	성별
최윤희	7,700	29	3	부장	남
최다인	7,500	40	2	과장	여
유강민	3,000	24	14	사원	남
⋮	⋮	⋮	⋮	⋮	⋮
김세현	9,500	47	2	부장	남

9

1. 시각화 맵핑

≡ 데이터 공간(data space)

📁 각각의 차원은 **유형(type)**을 가짐

유형(연속형 숫자)

↓

성명	연봉(만원)	나이	야근시간	직급	성별
최윤희	7,700	29	3	부장	남
최다인	7,500	40	2	과장	여
유강민	3,000	24	14	사원	남
⋮	⋮	⋮	⋮	⋮	⋮
김세현	9,500	47	2	부장	남

10

1. 시각화 맵핑

≡ 데이터 공간(data space)

📌 유형은 맵핑과 관련 있음

- 범주 비순서형(categorical): 성별(남, 여)
- 범주 순서형(ordinal): 직급(사원, 과장, 부장)
- 연속형 숫자(continuous): 연봉(5,000 6,000 7,000 8,000)



11

1. 시각화 맵핑

≡ 데이터 유형(data type)

(1/3)

변수 유형	예시	적정 스케일	설명
정량적/ 연속형 숫자	1.3, 5.7, 83	연속형	<ul style="list-style-type: none">• 임의의 숫자• 유리수, 실수
정량적/ 이산형 숫자	1, 2, 3, 4	이산형	<ul style="list-style-type: none">• 이산적인 단위의 숫자• 대부분 정수이나 예외인 숫자도 있음• 예를 들어 0.5, 1.0, 1.5는 중간값이 없음

12

1. 시각화 맵핑

≡ 데이터 유형(data type)

(2/3)

변수 유형	예시	적정 스케일	설명
정성적/ 범주 비순서형	개, 고양이, 물고기	이산형	<ul style="list-style-type: none"> • 순서를 매길 수 없는 범주 • 데이터 값 사이에 서열 속성이 없는 고유의 범주 • 요인(factor)
정성적/ 범주 순서형	좋음, 보통, 나쁨	이산형	<ul style="list-style-type: none"> • 이산적이고 데이터 값 사이에 순서가 있는 범주 • 예를 들어, '보통'은 항상 '좋음'과 '나쁨' 사이에 위치

13

1. 시각화 맵핑

≡ 데이터 유형(data type)

(3/3)

변수 유형	예시	적정 스케일	설명
날짜, 시간	2018년 1월 1일, 오전 8시	연속형/ 이산형	<ul style="list-style-type: none"> • 특정 날짜나 시간 • 연도가 특정되지 않아도 됨
텍스트	날썩 갈색 여우가 게으른 개를 뛰어넘다	해당 없음/ 이산형	<ul style="list-style-type: none"> • 자유 형식의 문자 • 필요에 따라 범주형으로 처리할 수 있음

14

1. 시각화 맵핑

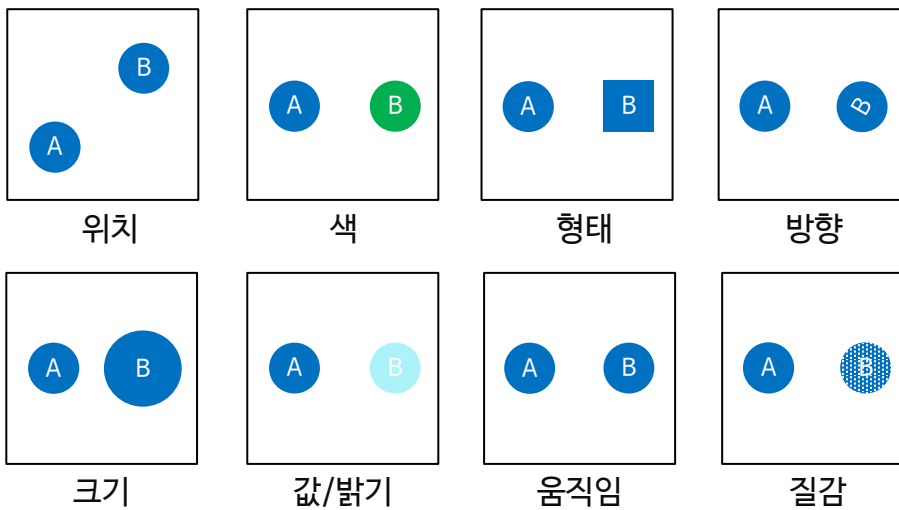
데이터 유형 예시

월	일	지역	관측소 ID	기온(화씨)
1월	1일	시카고	USW00014819	25.6
1월	1일	샌디에이고	USW00093107	55.2
1월	1일	휴스턴	USW00012918	53.9
1월	1일	데스밸리	USW00042319	51.0
1월	2일	시카고	USW00014819	25.5
1월	2일	샌디에이고	USW00093107	55.3
1월	2일	휴스턴	USW00012918	53.8
1월	2일	데스밸리	USW00042319	51.2

15

1. 시각화 맵핑

시각적 공간(visual space)

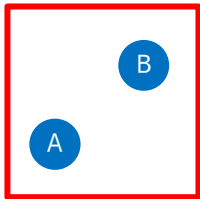


[출처] Fundamentals of Data Visualization, University of Colorado Boulder

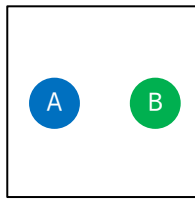
16

1. 시각화 맵핑

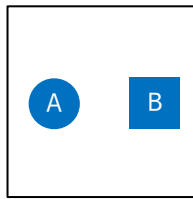
시각적 공간(visual space)



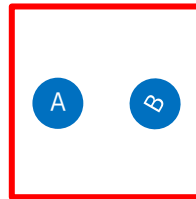
위치



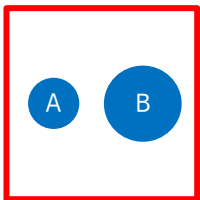
색



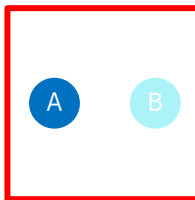
형태



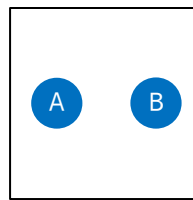
방향



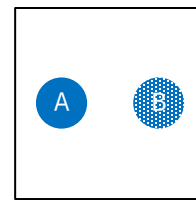
크기



값/밝기



움직임



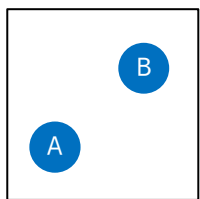
질감

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

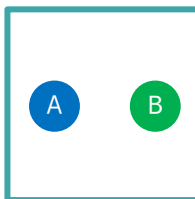
17

1. 시각화 맵핑

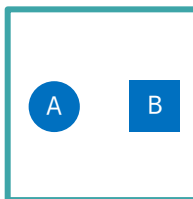
시각적 공간(visual space)



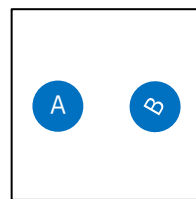
위치



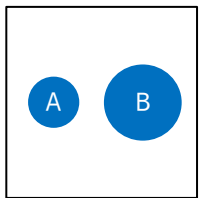
색



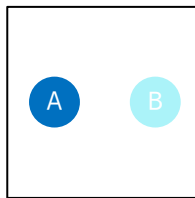
형태



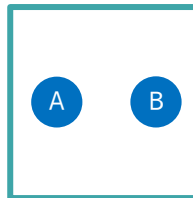
방향



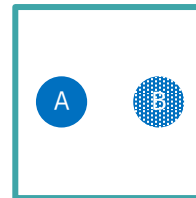
크기



값/밝기



움직임



질감

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

18

1. 시각화 맵핑

시각적 공간(visual space)

시각화로 데이터 값에 얼마나 차이가 있는지는 정확히 알 수 없으나
직관적으로 차이가 있다는 것은 알 수 있음



[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

19

1. 시각화 맵핑

시각화 맵핑 방법의 결정

시각화 맵핑은 데이터를 **시각적 표현(visual representation)**으로
변환하는 것

두 가지 질문

- 어떤 시각적 표현을 쓸 것인가? 최적의 방법은 무엇인가?
- 데이터 값(value)으로 부터 어떻게 시각적 표현을 계산해 낼 것인가?

맵핑을 정하기 위해 **사람의 인식(perception)**을 사용

- 어떤 시각적 패턴을 강조하고 싶은가?

20

1. 시각화 맵핑

≡ 표현성(expressiveness)과 효과성(effectiveness)

☞ **표현성**: 데이터의 차원(속성)에 포함된 정보를 표현해야 함



☞ **효과성**: 차원(속성)의 중요도는 시각적 표현의 현저함(salience)과 일치해야 함



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

21

1. 시각화 맵핑

≡ 시각적 속성의 고려사항

☞ **크기(size)**: 길이(length)와 면적(area)

☞ **형태(shape)**: 각도(angle), 모서리(edge), 정성적인 형태(qualitative shape)

☞ **움직임(motion)**: 속도(velocity), 가속도(acceleration), 방향(direction)

☞ **질감(texture)**: 밀도(density), 변화(variation), 패턴(pattern)

☞ **색(color)**: 색조(hue), 채도(chroma), 밝기(luminance)

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

22

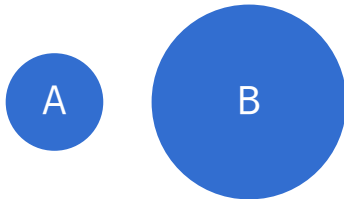
1. 시각화 맵핑

시각적 속성의 고려사항

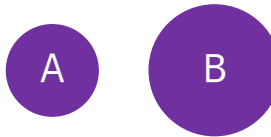
크기(size): 길이(length)와 면적(area)

예

그룹 A는 10,000명, 그룹 B는 20,000명



■ **지름**으로 값을 인코딩
➔ 그룹 B가 **4배** 큼



■ **면적**으로 값을 인코딩
➔ 그룹 B가 **2배** 큼

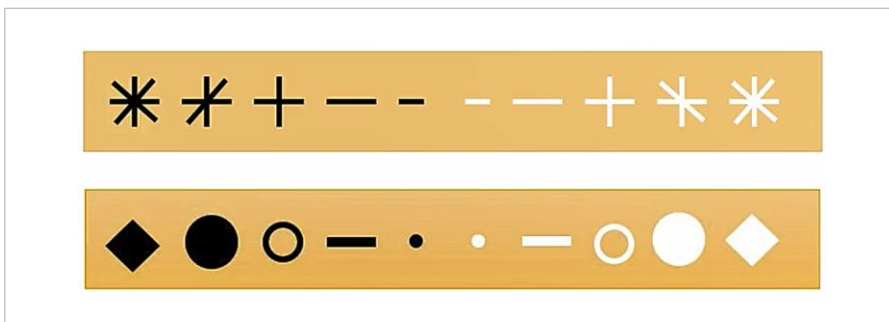
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

23

1. 시각화 맵핑

시각적 속성의 고려사항

형태(shape): 각도(angle), 모서리(edge),
정성적인 형태(qualitative shape)



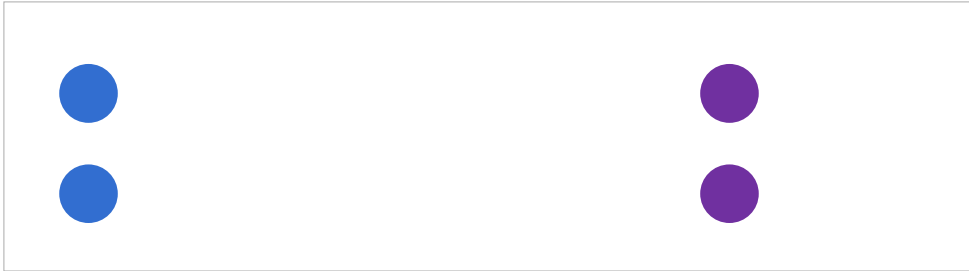
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

24

1. 시각화 맵핑

시각적 속성의 고려사항

☞ 움직임(motion): 속도(velocity), 가속도(acceleration), 방향(direction)



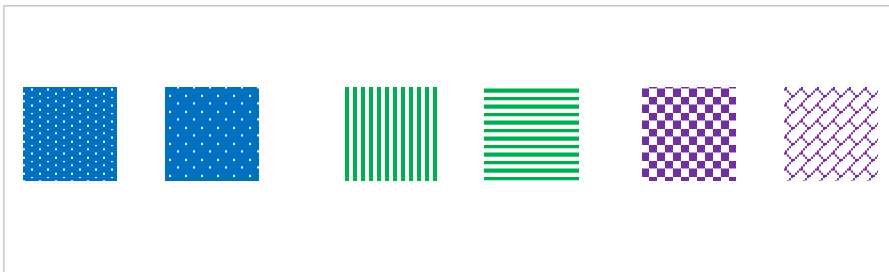
[출처] Fundamentals of Data Visualization, University of Colorado Boulder

25

1. 시각화 맵핑

시각적 속성의 고려사항

☞ 질감(texture): 밀도(density), 변화(variation), 패턴(pattern)



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

26

1. 시각화 맵핑

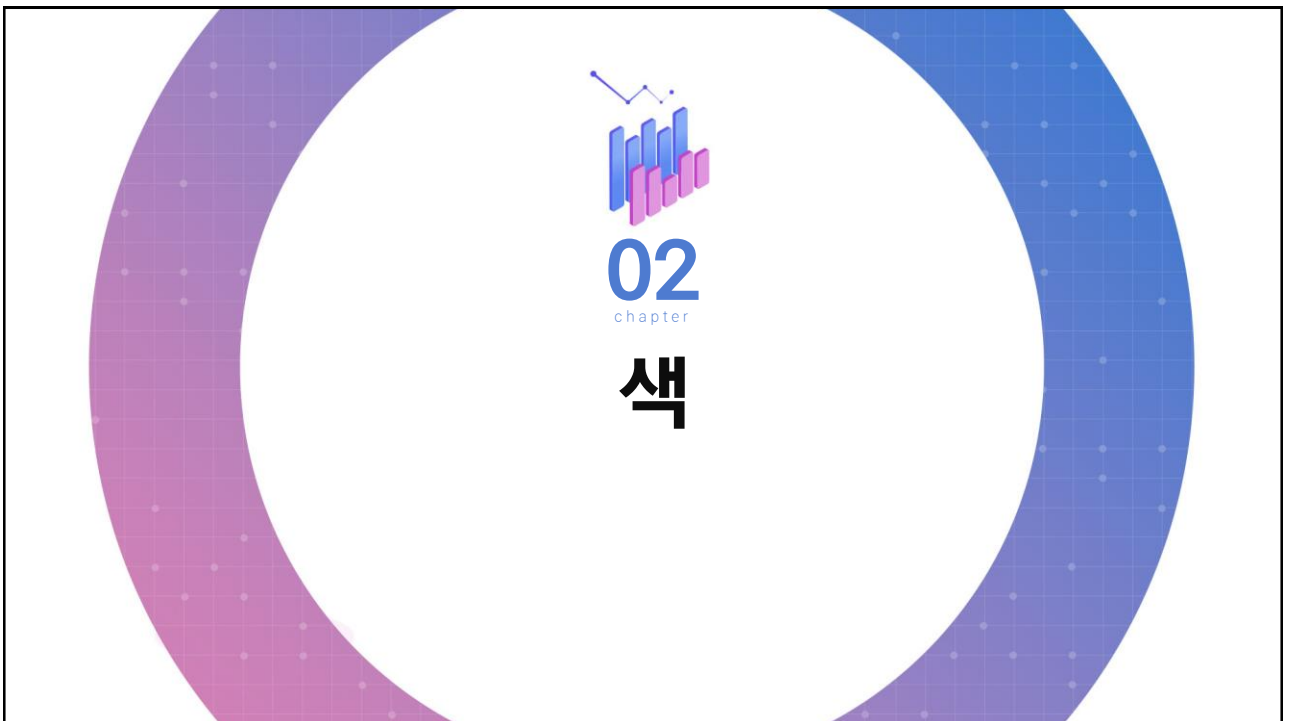
시각적 속성의 고려사항

색(color): 색조(hue), 채도(chroma), 밝기(luminance)

hue
chroma
luminance

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

27



28

2. 색

≡ 색의 인코딩 방법

☞ 순서형 인코딩

- 밝기, 채도 사용
- 연속적으로 구분이 가능해야 함



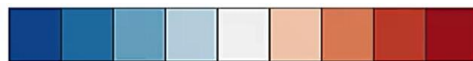
☞ 범주형 인코딩

- 색조 사용
- 범주별로 구분이 가능해야 함



☞ 발산형 인코딩

- 다른 그룹에서는 범주형으로 구분
- 같은 그룹에서는 순서형으로 구분



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

29

2. 색

≡ 좋은 인코딩을 위한 원칙

☞ 순서형, 발산형 인코딩

- 인지적으로 선형성을 가져야 함
- 자연스러운 순서를 가져야 함
- 밝기는 한방향(monotonic)으로 변해야 함
- 발산형 인코딩의 0의 위치 조절



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

30

2. 색

≡ 좋은 인코딩을 위한 원칙

☞ 순서형, 발산형 인코딩

- 인지적으로 **선형성**을 가져야 함

RGB 값의 차이와 인지적 차이는 다름



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

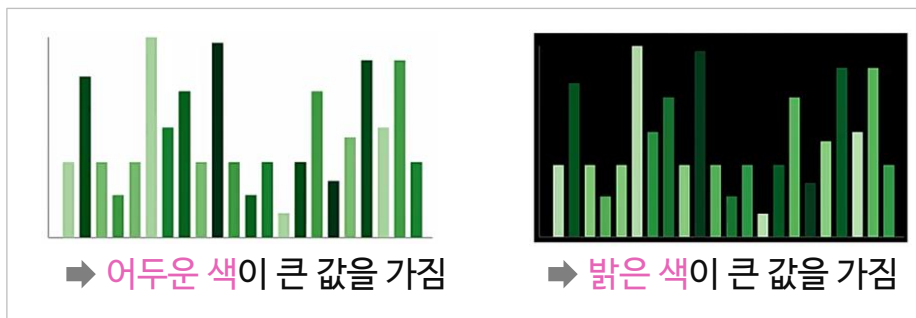
31

2. 색

≡ 좋은 인코딩을 위한 원칙

☞ 순서형, 발산형 인코딩

- 자연스러운 **순서**를 가져야 함



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

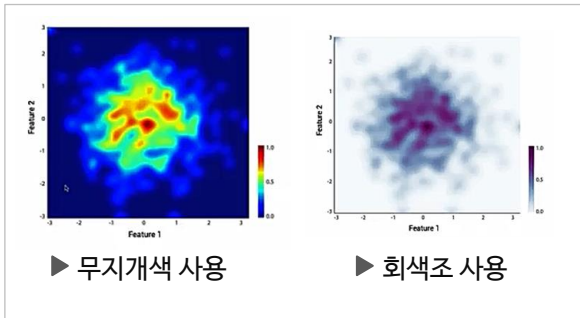
32

2. 색

≡ 좋은 인코딩을 위한 원칙

☞ 순서형, 발산형 인코딩

- 밝기는 **한방향(monotonic)**으로 변해야 함



무지개색 사용을 피해야 하는 이유

- 01 인위적인 값의 경계가 형성됨
- 02 직관적이지 않은 밝기

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

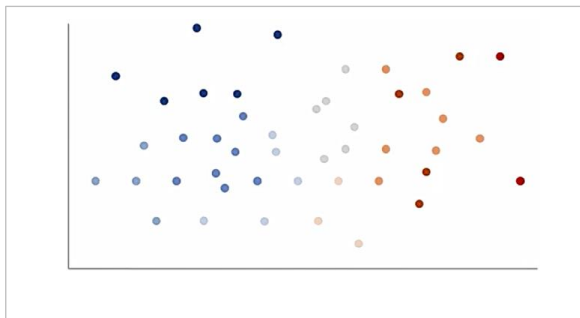
33

2. 색

≡ 좋은 인코딩을 위한 원칙

☞ 순서형, 발산형 인코딩

- 발산형 인코딩의 **0의 위치** 조절



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

34

2. 색

좋은 인코딩을 위한 원칙

범주형 인코딩

- 색으로 범주 구분이 명확해야 함
- 맥락에 맞는 색을 사용해야 함
- 색의 현저함은 서로 비슷해야 함



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

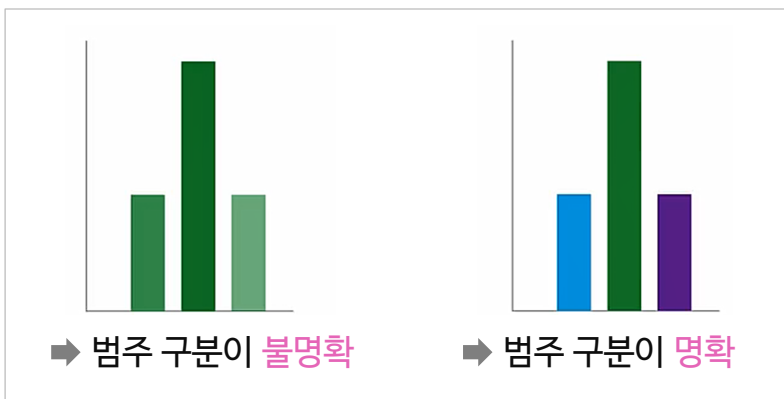
35

2. 색

좋은 인코딩을 위한 원칙

범주형 인코딩

- 색으로 범주 구분이 명확해야 함



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

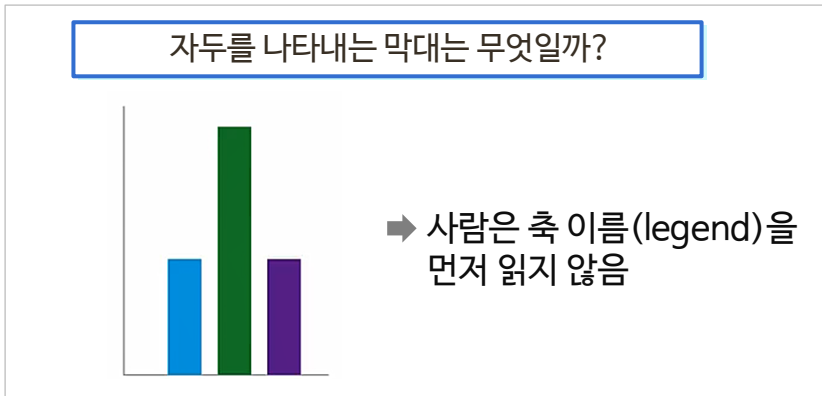
36

2. 색

좋은 인코딩을 위한 원칙

범주형 인코딩

- 맥락에 맞는 색을 사용해야 함



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

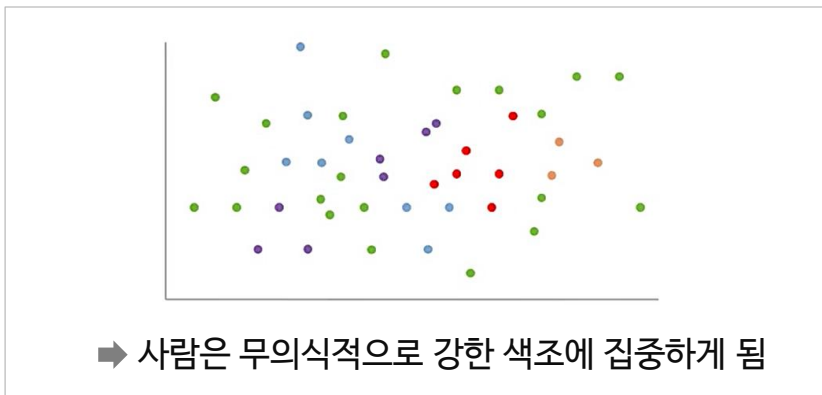
37

2. 색

좋은 인코딩을 위한 원칙

범주형 인코딩

- 색의 현저함은 서로 비슷해야 함



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

38



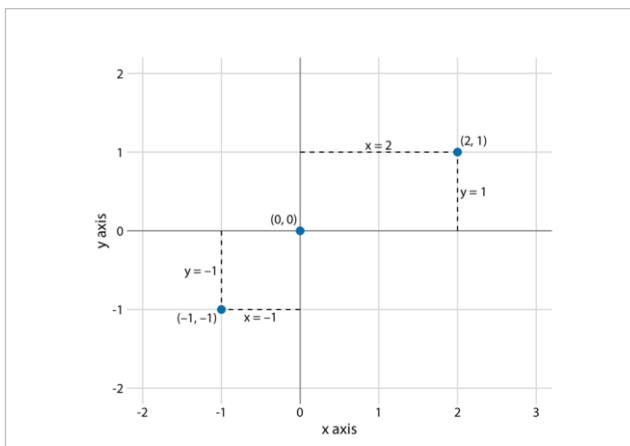
좌표계와 축

39

3. 좌표계와 축

≡ 좌표계(coordinate system)

☞ 위치 스케일을 모은 하나의 집합과 이 스케일들의 기하학적인 배열 관계



[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

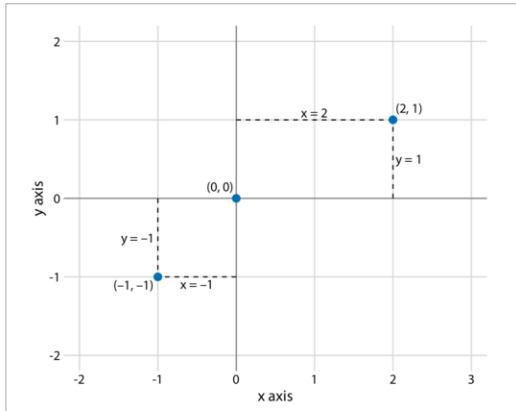
40

3. 좌표계와 축

직교좌표계(Cartesian coordinate system)

✎ x축과 y축이 서로 직각을 이루며 축(axis)을 따라 일정한 간격으로 데이터가 위치

■ 프랑스의 수학자인 데카르트의 이름을 따 데카르트 좌표계라고도 함



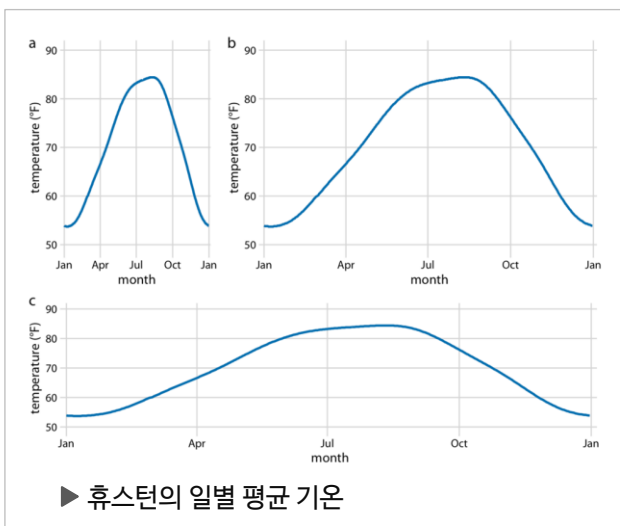
[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

- x축과 y축이 직각을 이룸
- 연속형 위치 스케일이며 양수와 음수를 모두 나타냄
- 각 축의 데이터를 아우르는 범위를 표시

41

3. 좌표계와 축

직교좌표계 예시



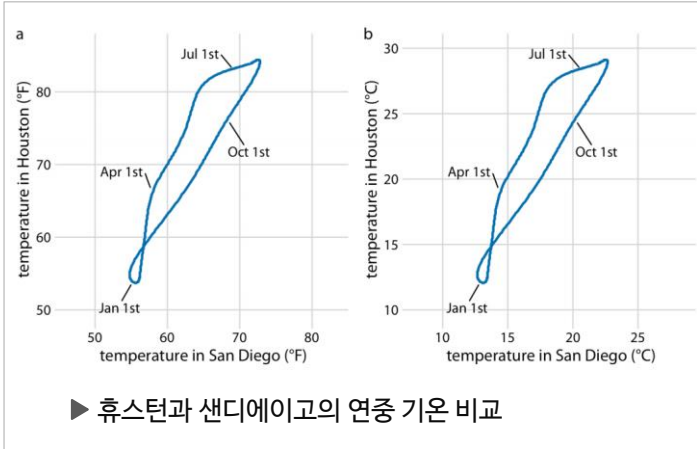
[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

- 축에는 데이터의 단위를 함께 표시 (온도, 거리 등)
- 두 축의 단위가 다른 경우 한 축의 데이터가 다른 축의 값에 비례하여 늘거나 줄어야 함
- 전달하려는 메시지에 따라 '가로 세로 비율(aspect ratio)'을 조절해야 함

42

3. 좌표계와 축

직교좌표계 예시



- x축과 y축이 같은 단위를 사용한다면 두 축의 격자선의 간격은 같아야 함
- 단위를 바꾸면 선형적인 변경 작업 필요 (a는 화씨, b는 섭씨 단위 사용)

[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

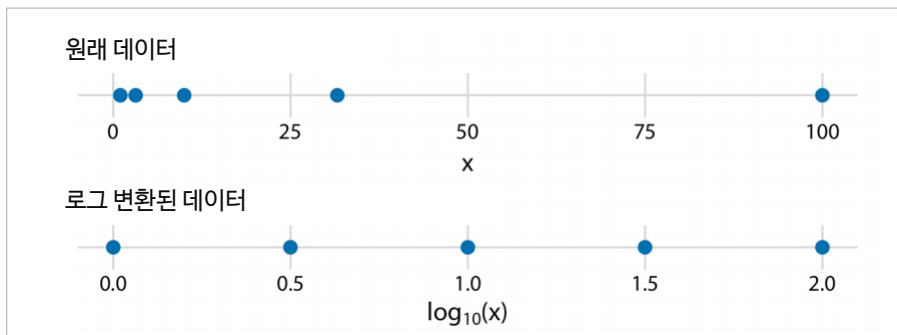
43

3. 좌표계와 축

비선형 축

☞ 격자선의 데이터의 단위 또는 격자의 간격이 비선형(일정하지 않은)적인 축

- 로그(log) 스케일은 대표적인 비선형 축으로 눈금의 단위가 고정된 값의 배수임
- 로그 스케일을 사용할 때는 축에 레이블(label)을 명확하게 표시해야 함

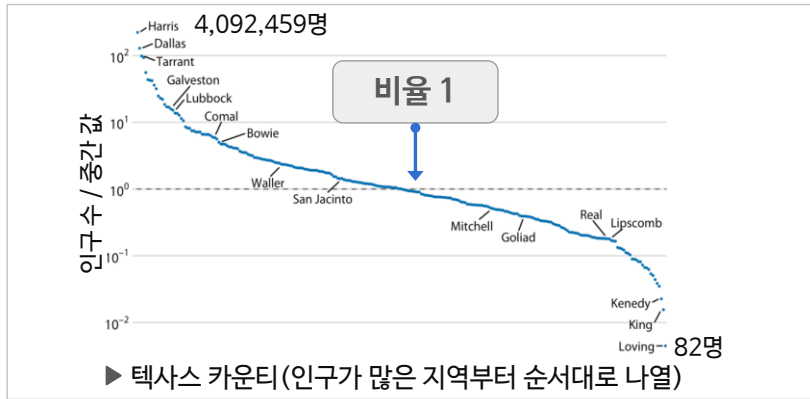


[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

44

3. 좌표계와 축

≡ 비선형 축 예시



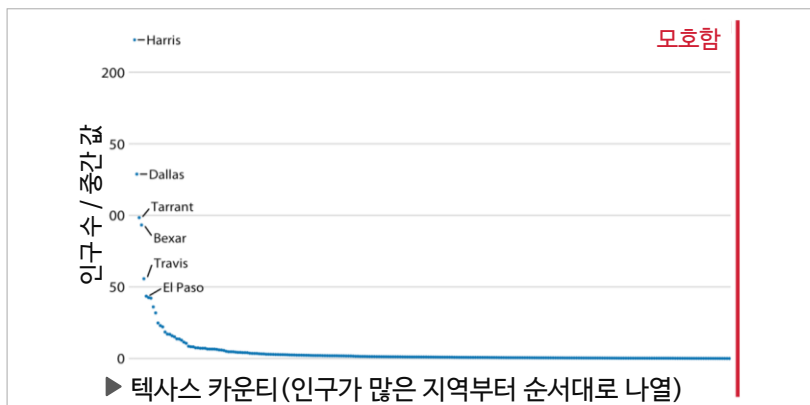
- 데이터의 비율 값은 **로그 스케일**로 표현하는 경우가 많음
- 텍사스의 카운티별 인구를 텍사스 전체 인구의 중간 값으로 나눈 비율

[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

45

3. 좌표계와 축

≡ 비선형 축 예시



- 데이터의 비율 값은 **로그 스케일**로 표현하는 경우가 많음
- 텍사스의 카운티별 인구를 텍사스 전체 인구의 중간 값으로 나눈 비율

[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

46

3. 좌표계와 축

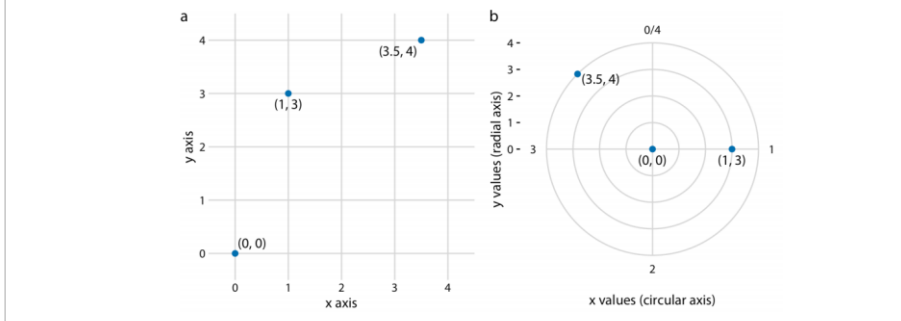
≡ 곡선 축이 있는 좌표계

📌 **극(polar) 좌표계**: 원점으로부터의 거리와 방향을 통해 위치를 표시하는 좌표계

- 눈금 한 쪽 끝의 데이터 값이 다른 쪽 끝과 논리적으로 이어지는 경우에 유용

예

1년 주기로 생성되는 데이터, 360도 방위로 표현되는 데이터

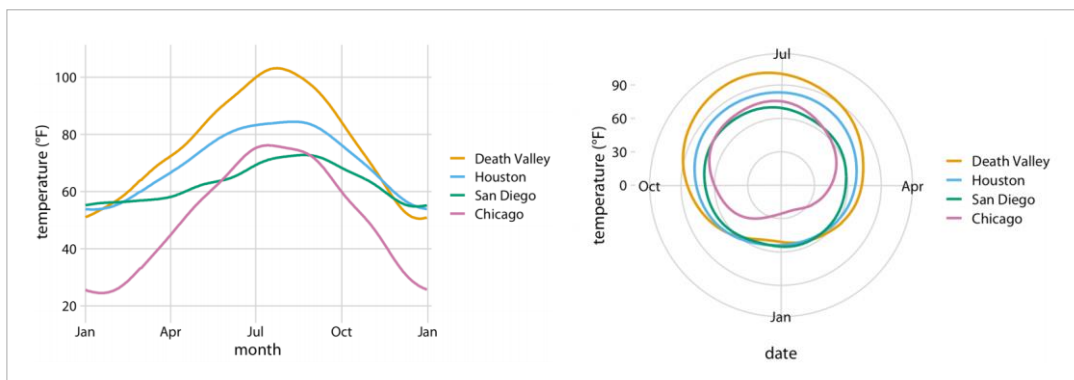


[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

47

3. 좌표계와 축

≡ 곡선 축이 있는 좌표계 예시



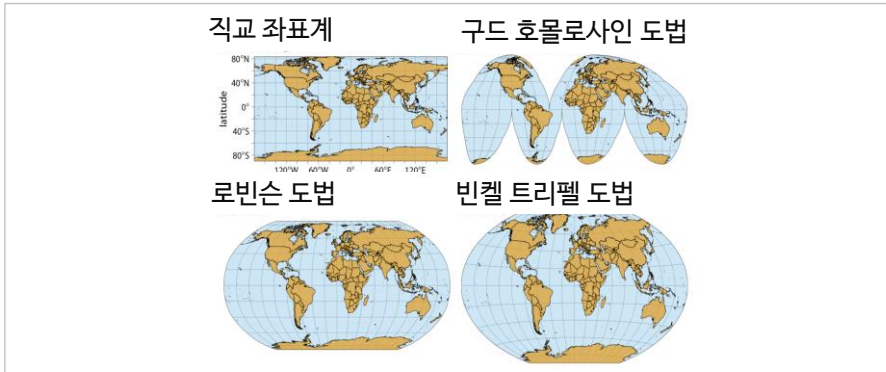
- 극 좌표계에 연중 기온을 나타내면 데이터의 순환을 강조할 수 있음

[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

48

3. 좌표계와 축

≡ 곡선 축이 있는 좌표계 예시



- 지도와 같은 지리공간적 데이터에도 곡선 축을 사용
- 지구는 둥글기 때문에 직교 좌표계에 위경도를 표시하는 것은 정보를 왜곡할 수 있음

[출처] Fundamentals of Data Visualization(데이터 시각화 교과서)

49

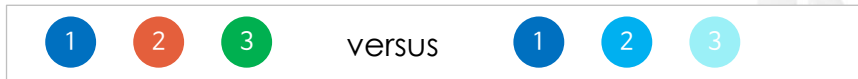
● 학습정리

1 시각화 맵핑

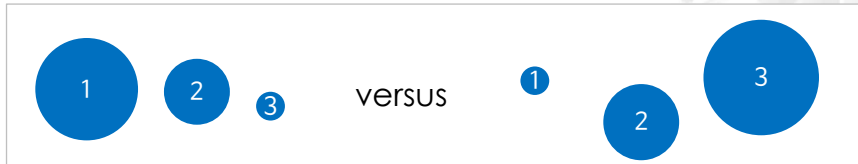
☞ 맵핑: 하나의 공간에서 다른 공간으로 변환(translate)하는 것

☞ 시각화 맵핑: 데이터 공간에서 시각적 공간으로 변환

☞ 표현성: 데이터의 차원(속성)에 포함된 정보를 표현해야 함



☞ 효과성: 차원(속성)의 중요도는 시각적 표현의 현저함과 일치해야 함



50

● 학습정리

2 색

☞ 순서형, 발산형 인코딩

- 인지적으로 선형성을 가져야 함
- 자연스러운 순서를 가져야 함
- 밝기는 한방향(monotonic)으로 변해야 함
- 발산형 인코딩의 0의 위치 조절

순서형 인코딩



발산형 인코딩



범주형 인코딩



51

● 학습정리

2 색

☞ 범주형 인코딩

- 색으로 범주 구분이 명확해야 함
- 맥락에 맞는 색을 사용해야 함
- 색의 현저함은 서로 비슷해야 함

순서형 인코딩



발산형 인코딩



범주형 인코딩



52

● 학습정리

3 좌표계와 축

📌 직교좌표계 (Cartesian coordinate system)

- x축과 y축이 서로 직각을 이루며 축을 따라 일정한 간격으로 데이터가 위치
- 두 축의 단위가 다른 경우 한 축의 데이터가 다른 축의 값에 비례하여 늘거나 줄어야 함

53

● 학습정리

3 좌표계와 축

📌 비선형 축

- 격자선의 데이터의 단위 또는 격자의 간격이 비선형(일정하지 않은)적인 축
- 로그 스케일을 사용할 때는 축에 레이블(label)을 명확하게 표시해야 함

54

참고문헌

- ☞ 「데이터 시각화 교과서」, Claus O. Wilke, 책만, 2020.
- ☞ 「Fundamentals of Data Visualization」, Claus O. Wilke, O'Reilly Media, 2019.
- ☞ 「Fundamentals of Data Visualization」, Danielle Albers Szair, University of Colorado Boulder(coursera course).

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

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.