

데이터 시각화

01. 시각화 개요

최대영 교수



고려사이버대학교
THE CYBER UNIVERSITY OF KOREA



Data Visualization

데이터 시각화

시각화 개요

01주차

최대영 교수



1

● 학습목표



시각화의 목적, 해석방법 및 특성에 대해 설명할 수 있다.



시각화의 그래픽 문법에 대해 설명할 수 있다.

2

● 학습내용

1 데이터 분석과 시각화

2 그래픽 문법



3

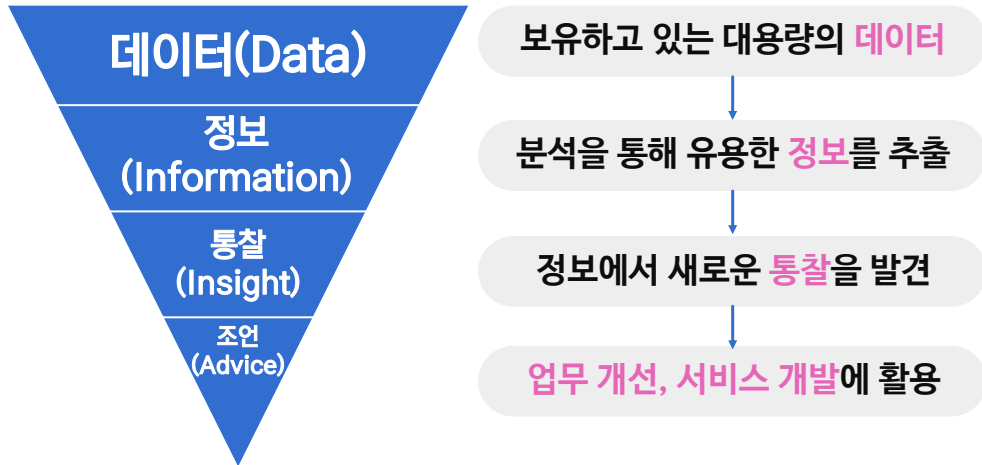


데이터 분석과 시각화

4

1. 데이터 분석과 시각화

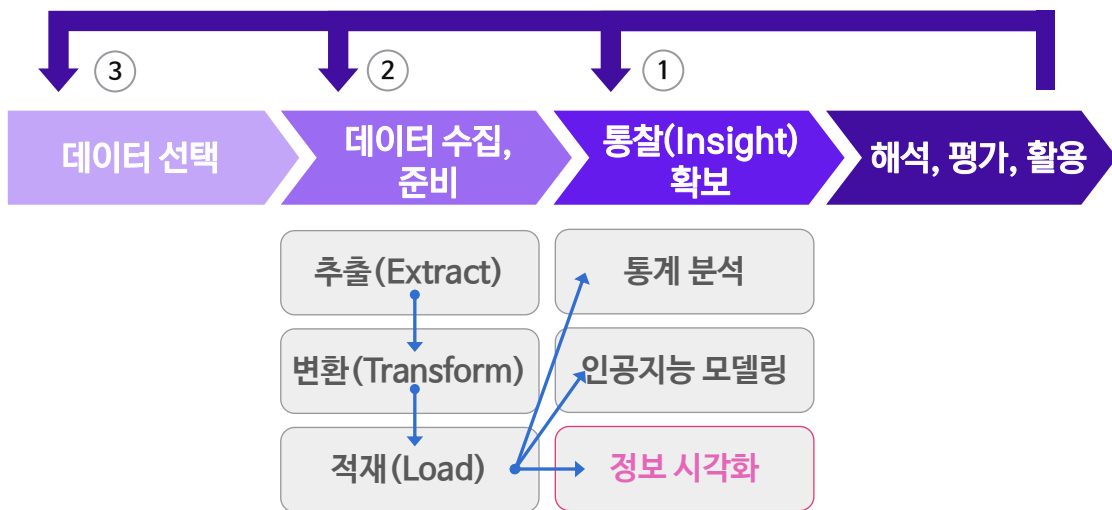
데이터 분석의 목적



5

1. 데이터 분석과 시각화

데이터 분석 과정



6

1. 데이터 분석과 시각화

≡ 공개데이터 예제

01 Enron Email Dataset

	Message-ID	Date	From	To	Subject	Mime-Version	Content-Type	Content-Transfer-Encoding	X-From
0	<18782981.1075855378110.JavaMail.evans@thyme>	Mon, 14 May 2001 16:39:00 -0700 (PDT)	phillip.allen@enron.com	tim.belden@enron.com		1.0	text/plain; charset=us-ascii	7bit	Phillip K Allen
1	<15464986.1075855378456.JavaMail.evans@thyme>	Fri, 4 May 2001 13:51:00 -0700 (PDT)	phillip.allen@enron.com	john.lavorato@enron.com	Re:	1.0	text/plain; charset=us-ascii	7bit	Phillip K Allen
2	<24216240.1075855687451.JavaMail.evans@thyme>	Wed, 18 Oct 2000 03:00:00 -0700 (PDT)	phillip.allen@enron.com	leah.arsdall@enron.com	Re: test	1.0	text/plain; charset=us-ascii	7bit	Phillip K Allen
3	<13505866.1075863688222.JavaMail.evans@thyme>	Mon, 23 Oct 2000 06:13:00 -0700 (PDT)	phillip.allen@enron.com	randall.gay@enron.com		1.0	text/plain; charset=us-ascii	7bit	Phillip K Allen
4	<30922949.1075863688243.JavaMail.evans@thyme>	Thu, 31 Aug 2000 05:07:00 -0700 (PDT)	phillip.allen@enron.com	greg.piper@enron.com	Re: Hello	1.0	text/plain; charset=us-ascii	7bit	Phillip K Allen

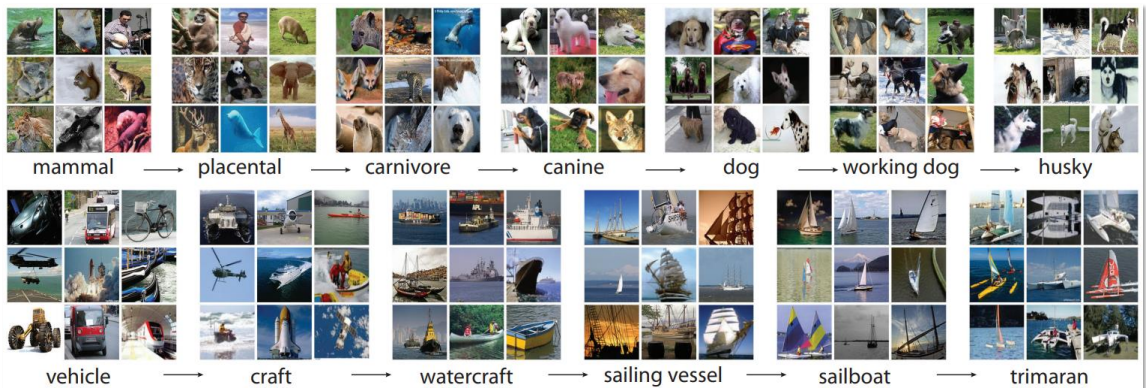
[출처] Enron Email Dataset

7

1. 데이터 분석과 시각화

≡ 공개데이터 예제

02 ImageNet Dataset



[출처] Deng et al., ImageNet: A Large-Scale Hierarchical Image Database

8

1. 데이터 분석과 시각화

업무데이터 예제

- ☞ 데이터의 **생성과정**에 대해 알 수 없는 경우가 많고 잘못된 데이터에 대한 **이유 확인이 어려움**
- ☞ 데이터에 대한 **기본적인 이해**가 먼저 필요, 실제로 데이터를 생성하는 사람과의 **협업 필요**
- ☞ **누락된 값, 알 수 없는 값(99999, NA), 형식(Type) 불일치 값 존재**

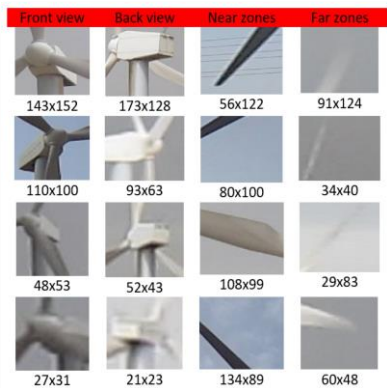


9

1. 데이터 분석과 시각화

업무데이터 예제

- ☞ 크기가 다르거나 다른 각도에서 찍은 사진, 라벨(label)이 없는 사진 존재



[출처] Bahaghighat et al., Estimation of Wind Turbine Angular Velocity Remotely Found on Video Mining and Convolutional Neural Network
Shin et al., AI-assistance for predictive maintenance of renewable energy systems

10

1. 데이터 분석과 시각화

시각화의 필요성

인간은 텍스트나 숫자로 된 대량의 데이터를 직관적으로 이해하는데 **인지의 한계**를 가짐

time	windSpeed	windSpeedMin	windSpeedMax	windSpeedSD	absWindDir	relWindDir	activePower	rotorSpeed
2016-02-01 AM 12:00:05	8.08718	5.61	12.81	1.029	300.14343	4.8971	802.15076	14.75116
2016-02-01 AM 12:05:00	8.80226	5.25	13.27	1.278	302.98285	-0.02261	1106.49963	15.69041
2016-02-01 AM 12:10:00	8.88634	4.89	13.92	1.444	299.37222	-5.82768	1061.32789	15.61896
2016-02-01 AM 12:15:00	9.02426	5.14	13.29	1.236	299.06009	-2.14379	1020.55994	15.59639
2016-02-01 AM 12:20:00	7.2168	2.61	11.35	1.724	303.97052	0.92548	658.26361	13.7001
2016-02-01 AM 12:25:00	8.58957	3.36	14.89	1.904	301.19699	-2.20294	1061.39111	15.30118
2016-02-01 AM 12:30:00	8.45736	4.93	12.37	1.216	305.72885	2.32893	1018.98462	15.37628
2016-02-01 AM 12:35:00	7.91647	4.74	12.13	1.286	304.2272	0.82734	828.81958	14.95707
2016-02-01 AM 12:40:00	7.95421	3.69	13.76	1.752	308.95898	5.22947	804.43066	14.72393
2016-02-01 AM 12:45:00	8.33013	4.2	13.25	1.177	316.94943	9.90101	1006.52078	15.60788
2016-02-01 AM 12:50:00	8.85471	4.48	12.95	1.494	300.76953	-7.43041	1085.7196	15.62712
2016-02-01 AM 12:55:00	8.99052	3.88	13.66	1.567	307.5842	2.31873	1199.40479	15.68338
2016-02-01 AM 01:00:00	9.44123	6.1	14.19	1.189	309.55099	5.551	1317.92689	16.00022
2016-02-01 AM 01:05:00	9.78405	5.59	15.93	1.358	315.17853	6.96985	1471.45923	16.1235
2016-02-01 AM 01:10:00	9.22982	5.28	13.06	1.256	319.92609	5.4686	1224.9458	15.89056
2016-02-01 AM 01:15:00	9.15731	5.73	13.59	1.393	305.97928	-8.52691	1138.42603	15.70146
2016-02-01 AM 01:20:00	9.26506	5.73	13.83	1.237	308.7341	3.92089	1312.1799	16.02081
2016-02-01 AM 01:25:00	8.95563	4.1	13.32	1.36	311.25037	5.44914	1119.38489	15.67789
2016-02-01 AM 01:30:00	8.54358	4.25	12.97	1.338	309.85477	4.03144	995.12561	15.36687
2016-02-01 AM 01:35:00	9.21237	6.5	12.84	1.002	309.9476	3.95008	1267.00896	15.97017
2016-02-01 AM 01:40:00	9.04621	6.05	12.66	0.992	310.72894	1.32883	1177.60364	15.94948
2016-02-01 AM 01:45:00	8.55168	4.29	14.37	1.254	304.08011	-2.37072	1045.70374	15.60725
2016-02-01 AM 01:50:00	9.13649	5.9	13.19	1.033	309.38002	4.18102	1295.22803	16.00655
2016-02-01 AM 01:55:00	8.92376	4.26	13.9	1.301	303.53784	-1.66224	1116.77563	15.79138
2016-02-01 AM 02:00:00	9.43881	6.18	13.43	0.917	308.06735	2.86701	1360.96472	16.07355
2016-02-01 AM 02:05:00	8.81884	5.01	13.5	1.325	306.79556	0.58212	1041.85535	15.70312
2016-02-01 AM 02:10:00	9.24028	5.7	15.76	1.335	300.41827	-7.18108	1141.36206	15.86309
2016-02-01 AM 02:15:00	8.83925	4.97	14.14	1.458	303.50128	-4.09872	1089.87512	15.42291
2016-02-01 AM 02:20:00	7.7338	3.57	13.37	1.628	299.98425	-7.6156	759.35394	14.22095

▶ 풍력발전기 데이터

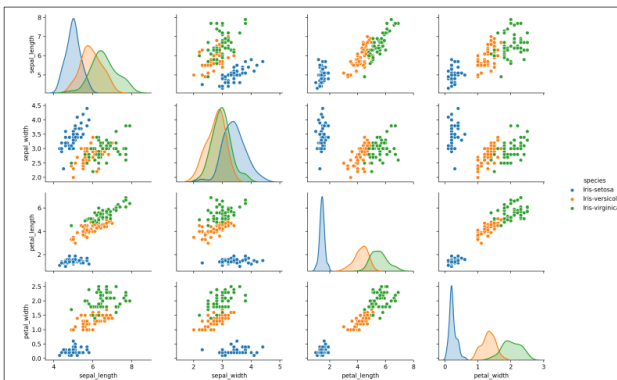
11

1. 데이터 분석과 시각화

시각화의 목적

01 인간의 **인지**를 확장시켜 통찰을 제공(못 보던 것을 보게 함)

■ 시각적 요소를 이용하여 정보를 인간이 효율적으로 이해할 수 있도록 표현



[출처] MAKESHIFTINSIGHTS

▶ 붓꽃 데이터(Iris dataset)의 분포

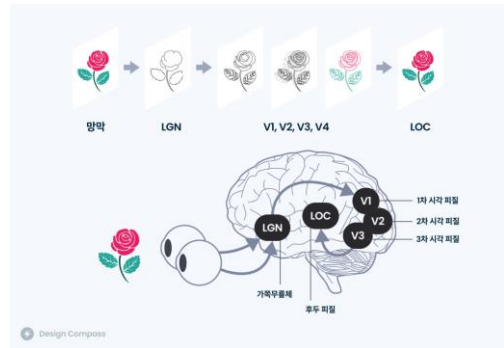
12

1. 데이터 분석과 시각화

시각화의 유용성

01 시각화가 통계치나 숫자 자체 보다 더 많은 통찰을 제공

- 인간의 뇌로 전달되는 정보의 약 90%가 시각적 정보
- 인간의 뇌는 문자 정보보다 시각 정보를 60,000배 빠르게 처리
- 인간의 뇌는 13~80ms안에 관찰된 시각 정보를 처리



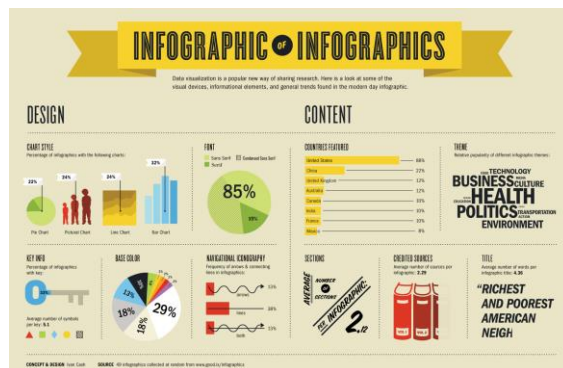
[출처] 디자인나침반레터

15

1. 데이터 분석과 시각화

시각화의 유용성

02 시각화의 결과물인 차트와 그래프를 정보를 더 빠르고 효율적으로 주고 받기 위한 언어로 활용



[출처] STC(Society for Technical Communication)

16

1. 데이터 분석과 시각화

Anscombe's quartet(앤스컴 콰르텟)

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

1		2		3		4	
x	y	x	y	x	y	x	y
10.0	8.04	8.0	6.58	10.0	9.14	10.0	7.46
8.0	6.95	8.0	5.76	8.0	8.14	8.0	6.77
13.0	7.58	8.0	7.71	13.0	8.74	13.0	12.74
9.0	8.81	8.0	8.84	9.0	8.77	9.0	7.11
11.0	8.33	8.0	8.47	11.0	9.26	11.0	7.81
14.0	9.96	8.0	7.04	14.0	8.10	14.0	8.84
6.0	7.24	8.0	5.25	6.0	6.13	6.0	6.08
4.0	4.26	19.0	12.50	4.0	3.10	4.0	5.39
12.0	10.84	8.0	5.56	12.0	9.13	12.0	8.15
7.0	4.82	8.0	7.91	7.0	7.26	7.0	6.42
5.0	5.68	8.0	6.89	5.0	4.74	5.0	5.73

[출처] 위키피디아

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

17

1. 데이터 분석과 시각화

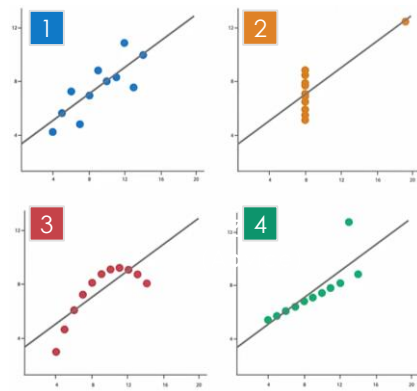
Anscombe's quartet(앤스컴 콰르텟)

그러나 시각화 해보면 매우 다른 특성의 데이터라는 것을 확인할 수 있음

시각화의 중요성, 이상치(outlier)의 영향에 대해서 알 수 있는 예제

1		2		3		4	
x	y	x	y	x	y	x	y
10.0	8.04	8.0	6.58	10.0	9.14	10.0	7.46
8.0	6.95	8.0	5.76	8.0	8.14	8.0	6.77
13.0	7.58	8.0	7.71	13.0	8.74	13.0	12.74
9.0	8.81	8.0	8.84	9.0	8.77	9.0	7.11
11.0	8.33	8.0	8.47	11.0	9.26	11.0	7.81
14.0	9.96	8.0	7.04	14.0	8.10	14.0	8.84
6.0	7.24	8.0	5.25	6.0	6.13	6.0	6.08
4.0	4.26	19.0	12.50	4.0	3.10	4.0	5.39
12.0	10.84	8.0	5.56	12.0	9.13	12.0	8.15
7.0	4.82	8.0	7.91	7.0	7.26	7.0	6.42
5.0	5.68	8.0	6.89	5.0	4.74	5.0	5.73

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

18

1. 데이터 분석과 시각화

시각화의 해석

01 어떤 데이터인가?

- 입력 데이터(input)는 무엇인가?
- 데이터가 어떻게 가공되었는가?

02 왜 시각화 하였는가?

- 시각화의 목적이 무엇인가?
- 정보를 보여주기 위해 왜 이 시각화 방법을 선택하였는가?

03 시각화가 어떻게 설계되었는가?

- 어떤 시각적 표현이 사용되었는가?
- 어떤 상호작용이 사용되었는가?

19

1. 데이터 분석과 시각화

시각화의 해석 예시

01 어떤 데이터인가?

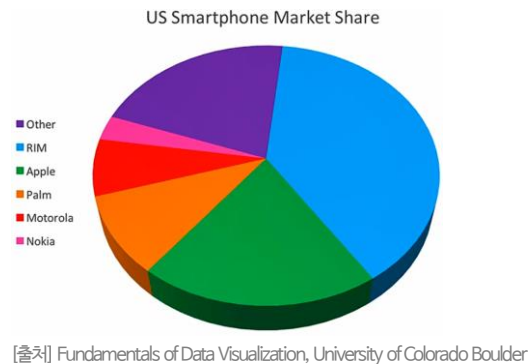
미국의 스마트폰 시장점유율

02 왜 시각화 하였는가?

관련된 사람들의 의사소통을 위해

03 시각화가 어떻게 설계되었는가?

3차원 파이차트(pie)를 사용하여 상대적인 점유율을 표현



20

1. 데이터 분석과 시각화

시각화의 해석 예시

01 어떤 데이터인가?

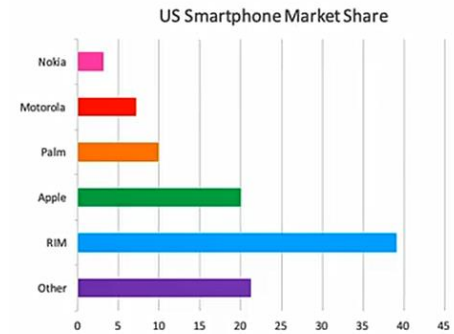
미국의 스마트폰 시장점유율

02 왜 시각화 하였는가?

관련된 사람들의 의사소통을 위해

03 시각화가 어떻게 설계되었는가?

3차원 파이차트(pie)를 사용하여 상대적인 점유율을 표현



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

21

1. 데이터 분석과 시각화

데이터 분석 방법과 시각화

기술적 방법(descriptive)

데이터를 관찰하여 전반적인 정보를 기술(describe)하는 방법, 탐색적 데이터 분석

진단적 방법(diagnostic)

데이터를 군집화(clustering) 등 가공하여 데이터에서 특이한 정보를 찾는 방법, 비지도학습

예측적 방법(predictive)

데이터로부터 변수 사이의 근본적인 관계성을 찾는 방법, 지도학습

처방/적응적 방법 (prescriptive/adaptive)

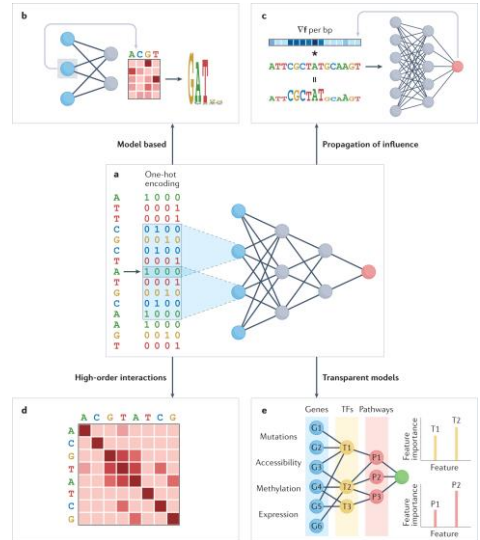
문제를 풀거나 목적함수(objective function)를 최적화하기 위해 시행과 보상받기를 반복하는 방법, 강화학습

22

1. 데이터 분석과 시각화

≡ 설명가능한 인공지능(explainable AI)

- ☞ 최근 인공지능 모델이 크고 복잡해짐에 따라 설명가능성(explainability)이 중요해짐
- ☞ 시각화를 통해 블랙박스(black box) 모델의 결과를 해석하려는 노력 진행
- ☞ 인공지능이 어떻게, 왜 그런 결과를 만들었는지 해석이 가능해야 활용 가능



[출처] Novakovsky et al., Obtaining genetics insights from deep learning via explainable artificial intelligence

23

1. 데이터 분석과 시각화

≡ 시각화의 특성

- ☞ 기술적(descriptive) 분석에만 한정되지 않고 분석의 전과정에 쓰임
- ☞ 항상 최적(최상)인 시각화 방법은 없음
- ☞ 시각화와 통계치, 숫자가 적절히 함께 사용될 때 정보 전달력을 높일 수 있음
- ☞ 데이터의 특성, 시각화 목적 및 기술을 종합적으로 고려하여 다양한 방법으로 표현한 후 최적의 시각화 방법을 사람이 선택

24



그래픽 문법

25

2. 그래픽 문법

≡ 그래픽 문법이란?

- ☞ 1990년대 월킨슨(Leland Wilkinson) 박사가 제안한 시각화의 구성요소와 관련된 프레임워크
- ☞ 시각화와 관련된 사항을 구조적으로 논의하기 위한 문법(언어, 틀)
 - 그래프를 이해하는 것 뿐만 아니라 정보를 전달하기 위한 최적의 시각화 방법을 선택 가능하게 함
 - 그래픽 문법으로 데이터의 처리부터 어떤 시각화 표현을 사용할 것인지까지의 시각화 전체에 대한 설명이 가능

26

2. 그래픽 문법

📖 그래픽 문법이란?

📌 시각화가 가진 공통의 구성요소(데이터, 시각적 속성 등)를 표현

- 다양한 시각화 도구(tool)는 그래픽 문법을 바탕으로 구현됨
- 파이썬의 matplotlib, R의 ggplot 등

matplotlib

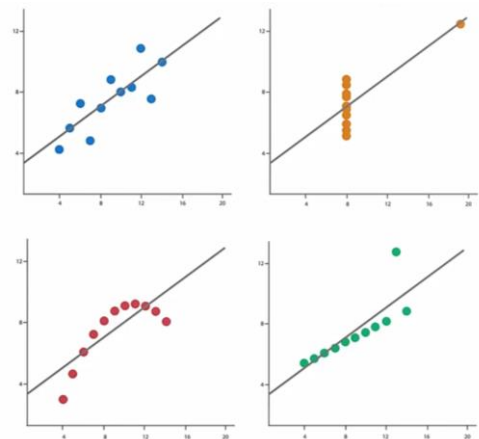


27

2. 그래픽 문법

📖 그래픽 문법의 7가지 요소

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



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

28

2. 그래픽 문법

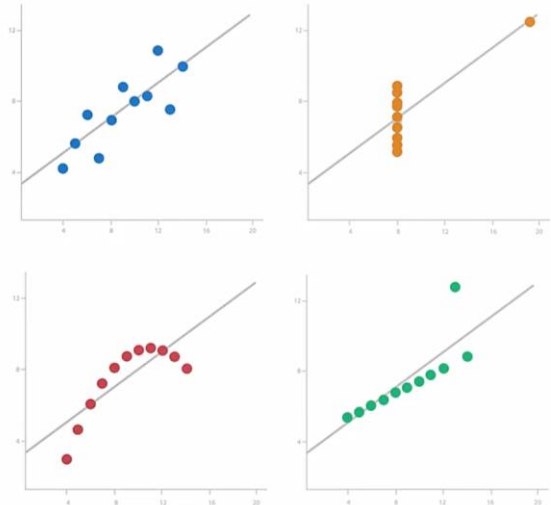
그래픽 문법의 7가지 요소

01 데이터(data)

어떤 데이터를 시각화하는가?

- 구체적인 차원(dimension), 컬럼(column), 속성(attribute)

각각 13개의 데이터 포인트로 이뤄진 4개의 데이터셋



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

29

2. 그래픽 문법

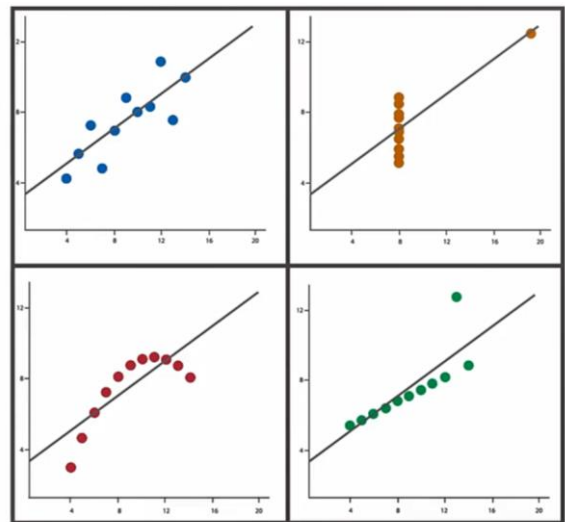
그래픽 문법의 7가지 요소

02 시각적 속성(aesthetics)

시각화 요소들은 무엇인가?

- 축(axis), 상대적 위치(relative position) 등

1개의 격자(grid) 안에 4개의 도표(plot)



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

30

2. 그래픽 문법

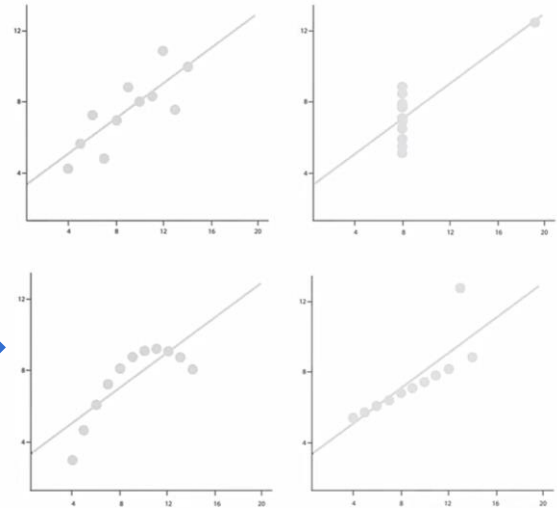
그래픽 문법의 7가지 요소

02 시각적 속성(aesthetics)

시각화 요소들은 무엇인가?

- 축(axis), 상대적 위치(relative position) 등

각각의 도표에는 각각 2개의 축



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

31

2. 그래픽 문법

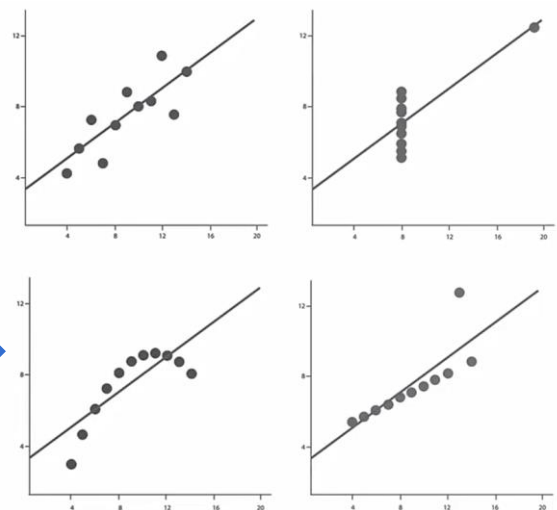
그래픽 문법의 7가지 요소

02 시각적 속성(aesthetics)

시각화 요소들은 무엇인가?

- 축(axis), 상대적 위치(relative position) 등

값(value)의 크기를
위치(position)로 표현



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

32

2. 그래픽 문법

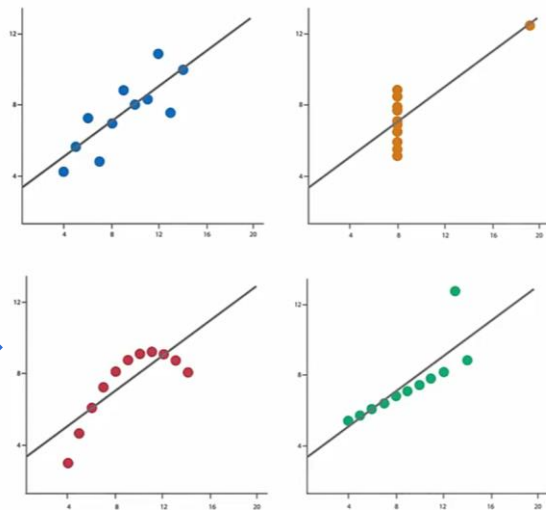
그래픽 문법의 7가지 요소

02 시각적 속성(aesthetics)

시각화 요소들은 무엇인가?

- 축(axis), 상대적 위치(relative position) 등

서로 다른 데이터셋을 표현하기 위해 색조(hue)를 사용



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

33

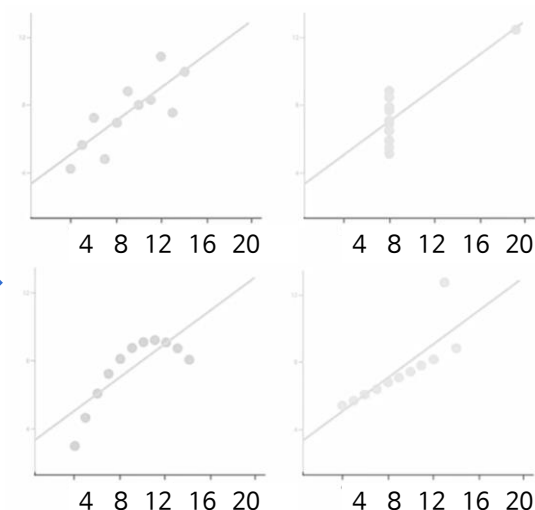
2. 그래픽 문법

그래픽 문법의 7가지 요소

03 스케일(scale)

어떤 범위의 값이 표현되는가?

x는 0~21 범위



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

34

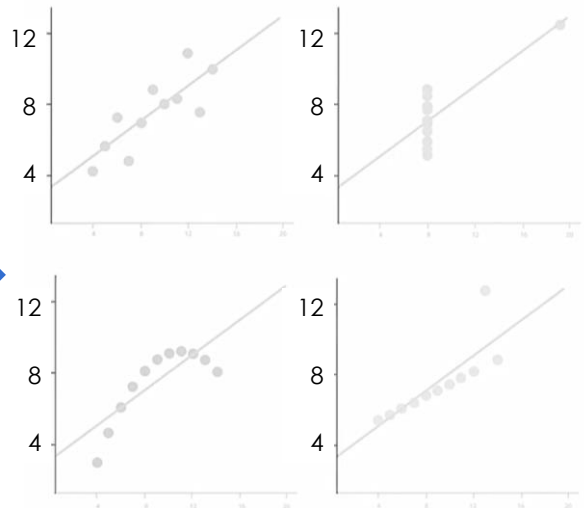
2. 그래픽 문법

그래픽 문법의 7가지 요소

03 스케일(scale)

어떤 범위의 값이 표현되는가?

y는 0~13 범위



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

35

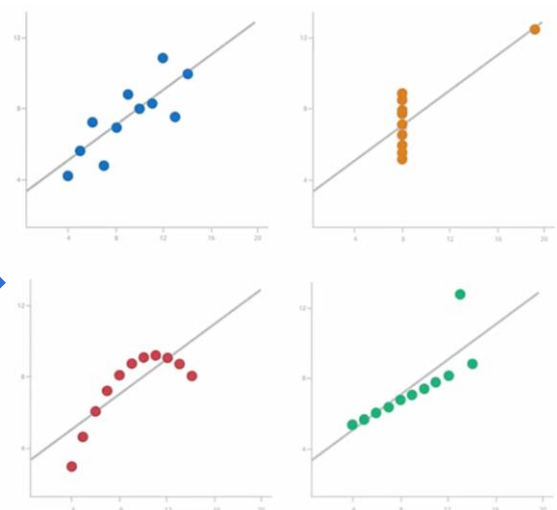
2. 그래픽 문법

그래픽 문법의 7가지 요소

03 스케일(scale)

어떤 범위의 값이 표현되는가?

파란색, 주황색, 빨간색, 녹색



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

36

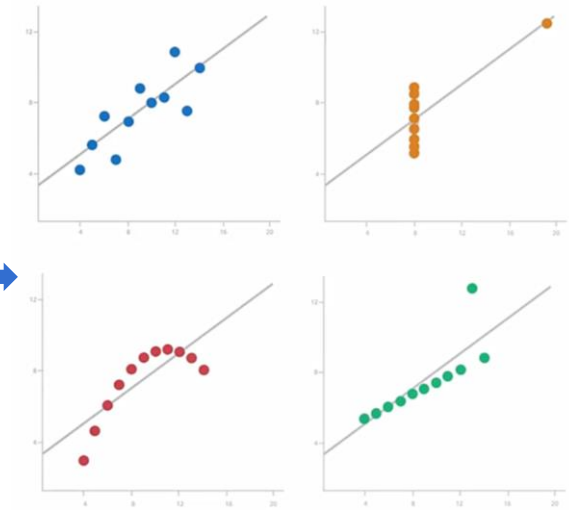
2. 그래픽 문법

그래픽 문법의 7가지 요소

04 기하학적 객체 (geometric object)

마크(mark)가 어떤 모양인가?

점(point)과 원(circle)으로
데이터를 표현



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

37

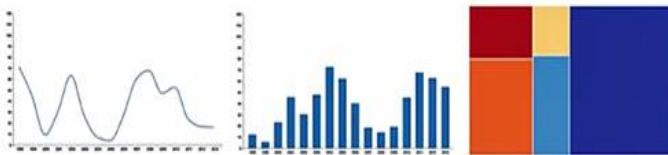
2. 그래픽 문법

그래픽 문법의 7가지 요소

04 기하학적 객체 (geometric object)

마크(mark)가 어떤 모양인가?

- 대안: 선(line), 막대(bar), 면(area)



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

38

2. 그래픽 문법

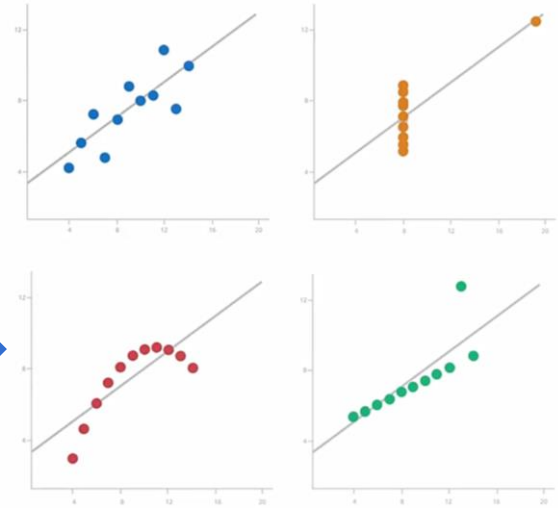
≡ 그래픽 문법의 7가지 요소

05 통계치(statistics)

☞ 어떤 계산된 정보가 표현되는가?

- 데이터가 어떻게 그룹화(grouped)되었는지도 포함

원시 데이터를 표현



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

39

2. 그래픽 문법

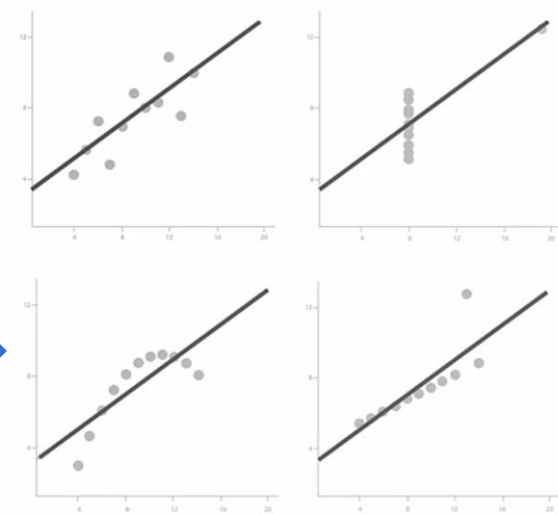
≡ 그래픽 문법의 7가지 요소

05 통계치(statistics)

☞ 어떤 계산된 정보가 표현되는가?

- 데이터가 어떻게 그룹화(grouped)되었는지도 포함

회귀선(regression line)을 표현



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

40

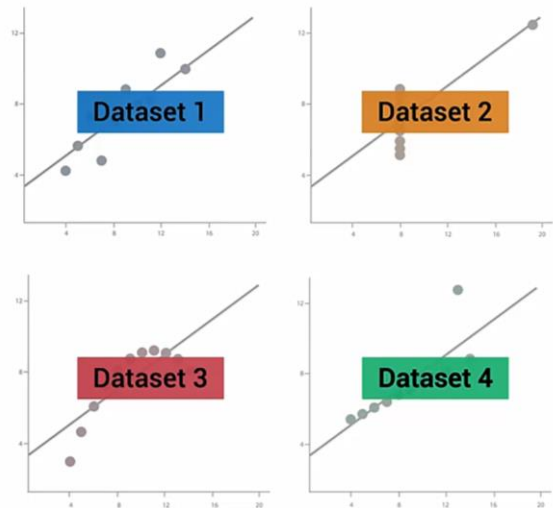
2. 그래픽 문법

그래픽 문법의 7가지 요소

06 면분할 (facets)

- ☑ 종종 데이터의 차원을 나눠 서로 다른 그래프로 표현 (x와 y, y와 z, x와 z)

데이터가 데이터셋의 종류에 따라 분할됨



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

41

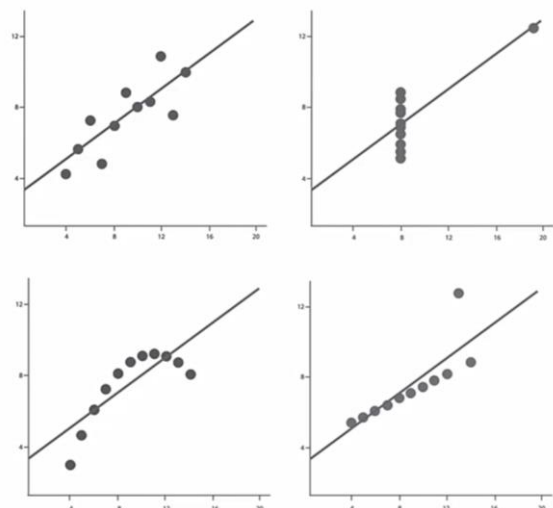
2. 그래픽 문법

그래픽 문법의 7가지 요소

07 좌표계 (coordinate system)

- ☑ 데이터 값이 어떻게 시각적 요소에 맵핑 (mapping) 되는가?

데이터가 축이 2개인 직교좌표계 (Cartesian, 데카르트)에 맵핑



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

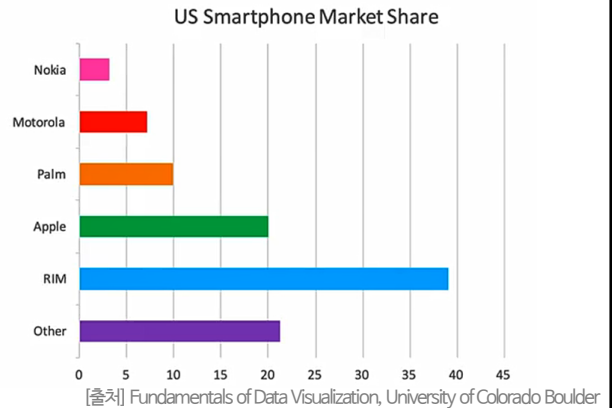
42

2. 그래픽 문법

그래픽 문법 예시(1)

7가지 요소 분석

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



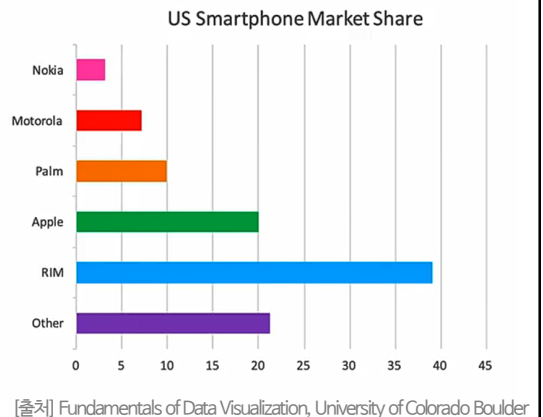
43

2. 그래픽 문법

그래픽 문법 예시(1)

7가지 요소 분석

요소	내용
데이터	미국 스마트폰 시장점유율(%)
시각적 속성	<ul style="list-style-type: none"> • 표이름(title), 축이름(label) • x축은 연속형 숫자 • y축은 범주형 문자 • 색은 범주 표현
스케일	<ul style="list-style-type: none"> • x: 0~45 • y: 위치 1~6 • 색: ● ● ● ● ●



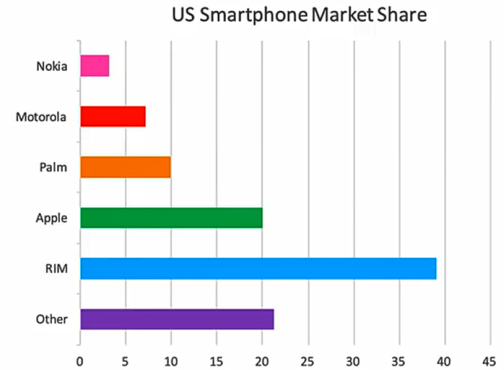
44

2. 그래픽 문법

그래픽 문법 예시(1)

7가지 요소 분석

요소	내용
기하학적 객체	막대
통계치	원시 데이터
면분할	1개의 차트에 전체 데이터 표현
좌표계	직교 좌표계



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

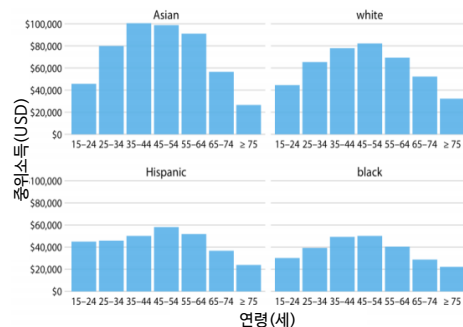
45

2. 그래픽 문법

그래픽 문법 예시(2)

7가지 요소 분석

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



[출처] Fundamentals of Data Visualization, Claus O. Wilke

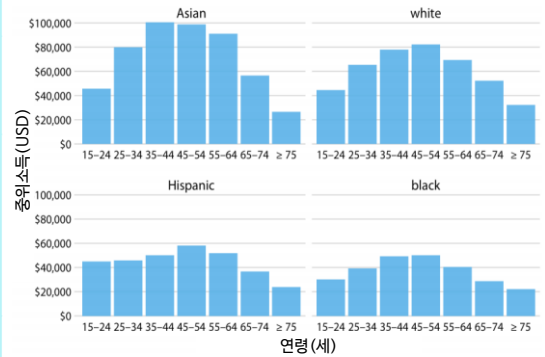
46

2. 그래픽 문법

≡ 그래픽 문법 예시(2)

📁 7가지 요소 분석

요소	내용
데이터	미국의 연령, 인종별 연간 중위 가계소득
시각적 속성	<ul style="list-style-type: none"> 축이름 x축은 범위가 있는 숫자 y축은 연속형 숫자 색은 의미 없음
스케일	<ul style="list-style-type: none"> x: 15~24, ..., 75~ y: 0~100,000



[출처] Fundamentals of Data Visualization, Claus O. Wilke

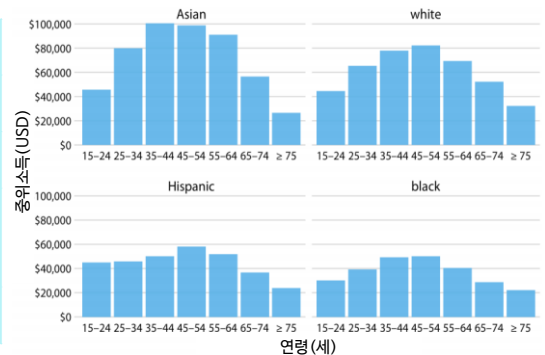
47

2. 그래픽 문법

≡ 그래픽 문법 예시(2)

📁 7가지 요소 분석

요소	내용
기하학적 객체	막대
통계치	원시 데이터
면분할	1개의 차트에 4개 데이터를 인종별로 표현
좌표계	직교 좌표계



[출처] Fundamentals of Data Visualization, Claus O. Wilke

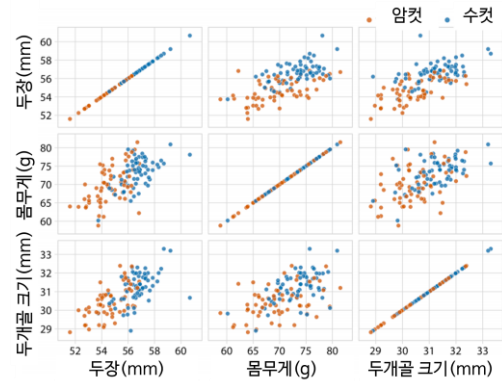
48

2. 그래픽 문법

그래픽 문법 예시(3)

7가지 요소 분석

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



[출처] Fundamentals of Data Visualization, Claus O. Wilke

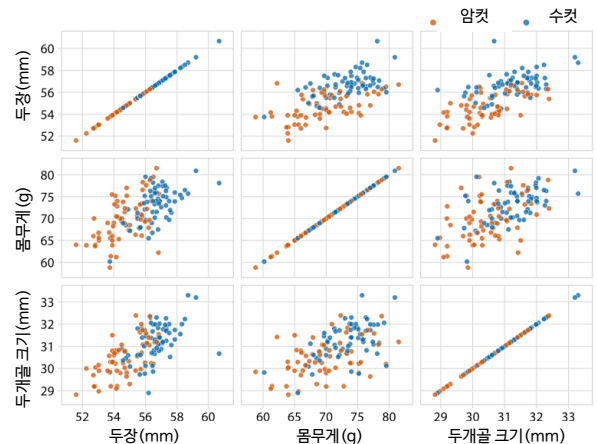
49

2. 그래픽 문법

그래픽 문법 예시(3)

7가지 요소 분석

요소	내용
데이터	큰어치 123마리의 두장(머리 길이), 몸무게, 두개골 크기
시각적 속성	<ul style="list-style-type: none"> • 축이름 • x축은 연속형 숫자 • y축은 연속형 숫자 • 색은 범주 표현(암/수컷)



[출처] Fundamentals of Data Visualization, Claus O. Wilke

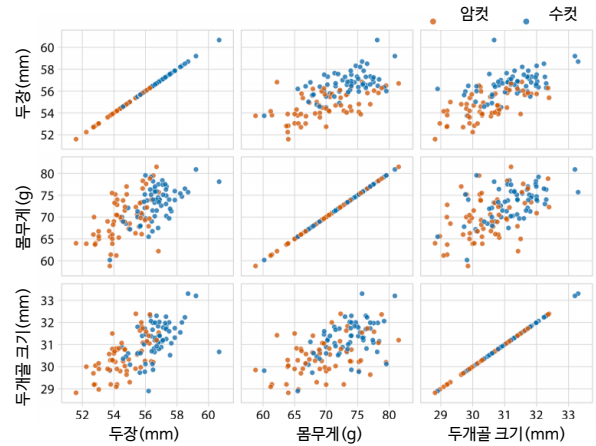
50

2. 그래픽 문법

≡ 그래픽 문법 예시(3)

📌 7가지 요소 분석

요소	내용
스케일	<ul style="list-style-type: none"> • x_1: 50~62, x_2: 55~85, x_3: 28~34 • y_1: 50~62, y_2: 55~85, y_3: 28~34 • 색: ● ●
기하학적 객체	점



[출처] Fundamentals of Data Visualization, Claus O. Wilke

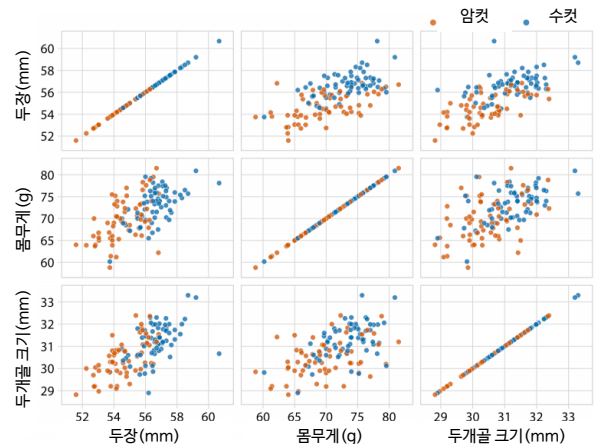
51

2. 그래픽 문법

≡ 그래픽 문법 예시(3)

📌 7가지 요소 분석

요소	내용
통계치	원시 데이터
면분할	9개의 차트에 1개의 같은 데이터를 반복적으로 표현
좌표계	직교 좌표계



[출처] Fundamentals of Data Visualization, Claus O. Wilke

52

● 학습정리

1 데이터 분석과 시각화

📌 시각화 목적

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

📌 시각화 해석

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

53

● 학습정리

1 데이터 분석과 시각화

📌 시각화의 특성

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

54

● 학습정리

2 그래픽 문법

📝 그래픽 문법이란?

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

55

● 학습정리

2 그래픽 문법

📝 그래픽 문법의 7가지 요소

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

56

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.