* 머리말

데이터 시각화(data visualization)는 자료에 내포된 정보를 내보이는 그림이다.

예컨대 파이 차트, 막대 그림, 산점도, 히스토그램과 같은 것이다. 이런 것들은 오래 전부터 써왔었는데 무슨 이유 로 다시 살리는 것일까?

근래의 컴퓨터 기술 혁신이 데이터 시각화에 관한 모든 것을 바꾸고 있기 때문이다.

데이터의 구조와 크기, 그래픽 작업 방식, 보편적 활용 등에서 그렇다.

이 책의 주제는 데이터의 기술적 시각화(descriptive visualization)이다. 분석적

시각화(analytical visualization)는 다루지 않는다. 1) 그러니까 이 책 은 수학이 없는 초급 시각화에 집중하며2) 범사회적 그래프 활용수준의 제고를 지향한다.

이 책은 25종의 그림을 절(節)에 나누어 보인다. 절의 순서는 의 미가 없다 책 보듯이 뒤적이다가 마음에 와 닿은 그래프가 있으면 꼼꼼히 읽기 바란다.

그림

25 종 그림을 데이터 또는 변 수 유형 에 따라 몇 개 그룹으로 나누면 다움과 같다.

(큰 괄호 안 숫자는 절 번호) 그룹별로 그래프를 접해보는 것도 한 방법이다.

- 시계열 자료 : 채색 달력 [l]. 시도표 [2 2]. 지나온 길 [23 ]. 영역 그림 [17].

- 산점도 : 이변량 밀도 [2]. '큰’ 자료의 산점도 [4]. 흩뜨림 [ 6]. 버블 차트 [18].

매개변 수 곡선 [1 6]. 조건부 플롯 [24].

- 3차원 공간 : 이변량 히스토그램 [3]. 설명변수 2 개의 회귀모형 [5], 움직이는

산점도 [25].

- 범주형 자료: 사각 타일 [7]. 테이 블 그림 [8]. 모자이크 플 롯 [9], 나무 지 도

[l 이. 막대 그림 [11]. 파이 차트 [12], 단어 클라우드 [ 2 이 .

- 연속형 변수 : 대시보 드 [1 3] . 상자 그 림 [19] .

- 기타 미분류 : R 의 컬 러 [21]. 행렬의 이미지 [1 4]. 몬테 칼로 모의시행 [15] .

- 움직이는 그림 : [S ]. [16]. [ 25 ] .

1) 기 술 적(記述的) 시각화는 데이터의 생김생김에 대 한 스케치 를 의미한다 . 이와 대비하여

분석적(分析的) 시 각화는 데이터 를 수학적으로 분석 한 결과를 그 림으로 표현해낸 다 . 주

성분 분석과 관련하여 행 렬도(bipl ot) 가 그 런 예이다 .

2) 수학(