



# 데이터 시각화 과목 소개



# 데이터시각화 이유

---

“

데이터시각화?

”

“

시각화, 시각화 하는데 왜 해야  
할까?

”



# 데이터 시각화 이유

Never trust summary statistics alone;  
always visualize your data !!

- Alberto Cairo -



# 데이터 시각화 이유



아래 표에는 1쌍으로 이루어진 4개 그룹의 숫자가 있다.  
아래 표에서 무엇이 보이는가?  
숫자들 간에 패턴이나 추세의 차이를 구별할 수 있는가?

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8,04	10	9,14	10	7,46	8	6,58
8	6,95	8	8,14	8	6,77	8	5,76
13	7,58	13	8,74	13	12,74	8	7,71
9	8,81	9	8,77	9	7,11	8	8,84
11	8,33	11	9,26	11	7,81	8	8,47
14	9,96	14	8,1	14	8,84	8	7,04
6	7,24	6	6,13	6	6,08	8	5,25
4	4,26	4	3,1	4	5,39	19	12,5
12	10,84	12	9,13	12	8,15	8	5,56
7	4,82	7	7,26	7	6,42	8	7,91
5	5,68	5	4,74	5	5,73	8	6,89

# 데이터 시각화 이유

Q

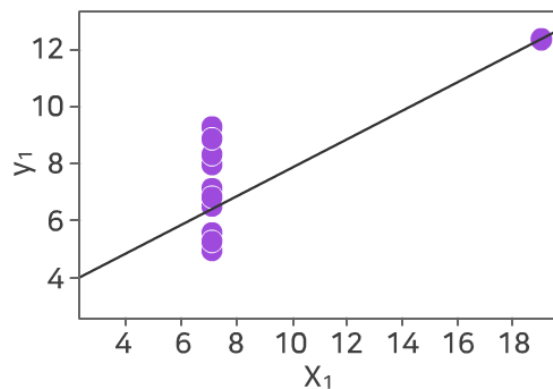
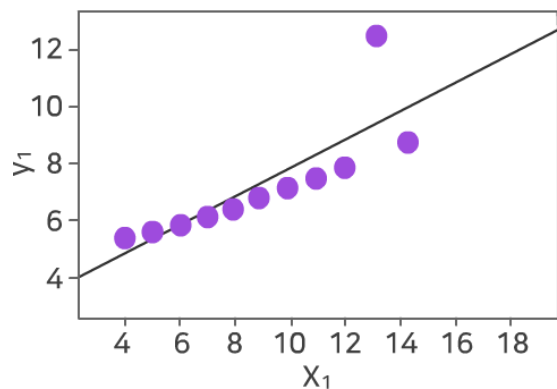
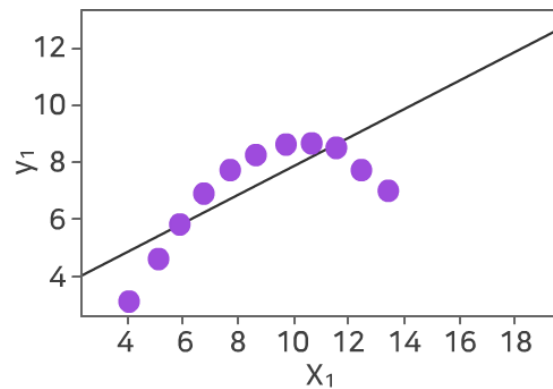
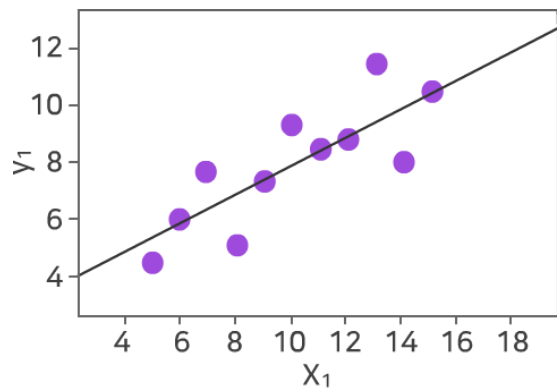
통계적 속성을 고려하여 데이터를 확인할 수 있다.  
아래 표는 X, Y가 각각 합계, 평균, 표준 편차가 같다.  
같은 패턴의 데이터라고 볼 수 있는가?

	I		II		III		IV	
	x	y	x	y	x	y	x	y
	10	8,04	10	9,14	10	7,46	8	6,58
	8	6,95	8	8,14	8	6,77	8	5,76
	13	7,58	13	8,74	13	12,74	8	7,71
	9	8,81	9	8,77	9	7,11	8	8,84
	11	8,33	11	9,26	11	7,81	8	8,47
	14	9,96	14	8,1	14	8,84	8	7,04
	6	7,24	6	6,13	6	6,08	8	5,25
	4	4,26	4	3,1	4	5,39	19	12,5
	12	10,84	12	9,13	12	8,15	8	5,56
	7	4,82	7	7,26	7	6,42	8	7,91
	5	5,68	5	4,74	5	5,73	8	6,89
SUM	99,00	82,51	99,00	82,51	99,00	82,50	99,00	82,51
AVG	9,00	7,50	9,00	7,50	9,00	7,50	9,00	7,50
STDEV	3,32	2,03	3,32	2,03	3,32	2,03	3,32	2,03

# 데이터 시각화 이유

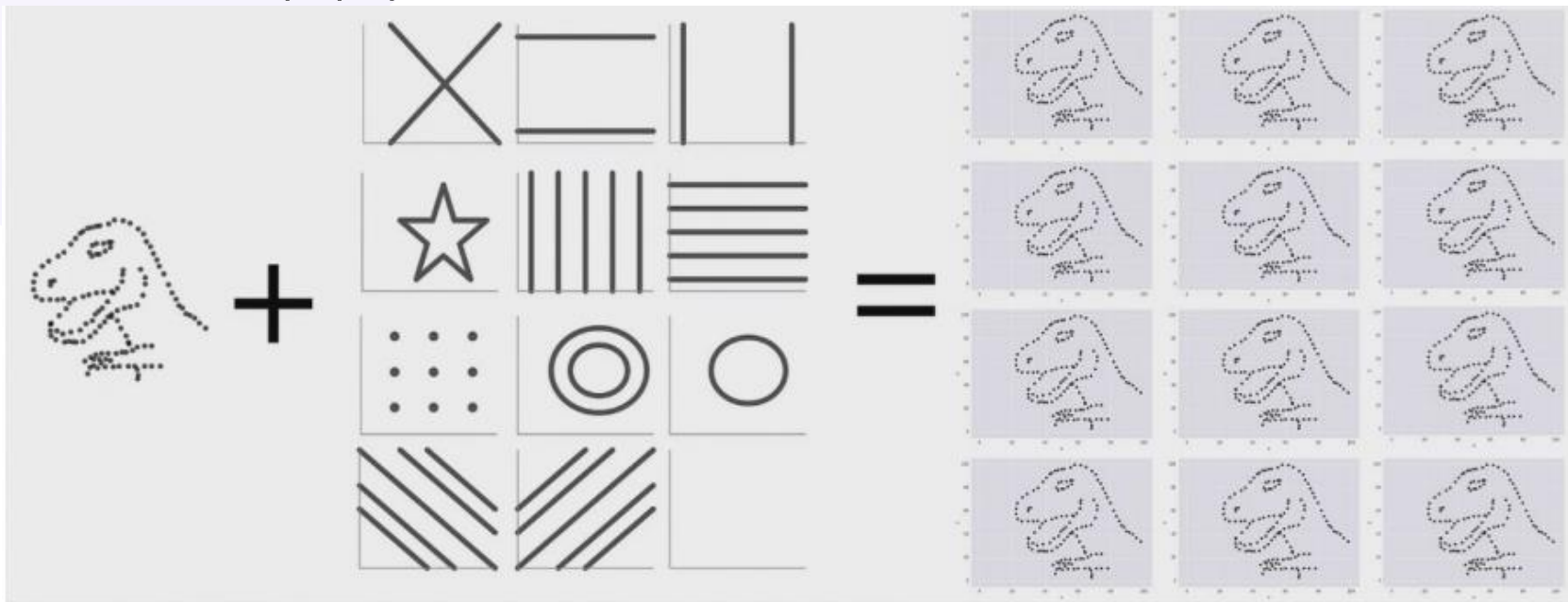
데이터를 차트로 옮겨 시각화해보면  
통계 값으로 알 수 없었던 결과를 확인할 수 있음

- ◆ 이 데이터는 앤스콤의 ‘4가지 그룹’이라는 숫자 세트임
- ◆ “수치 계산이 정확하며, 그래프는 대략 적”이라는 일반적인 관념에 맞섬



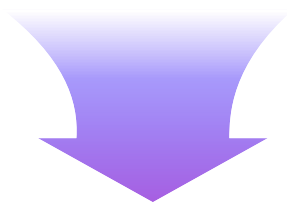
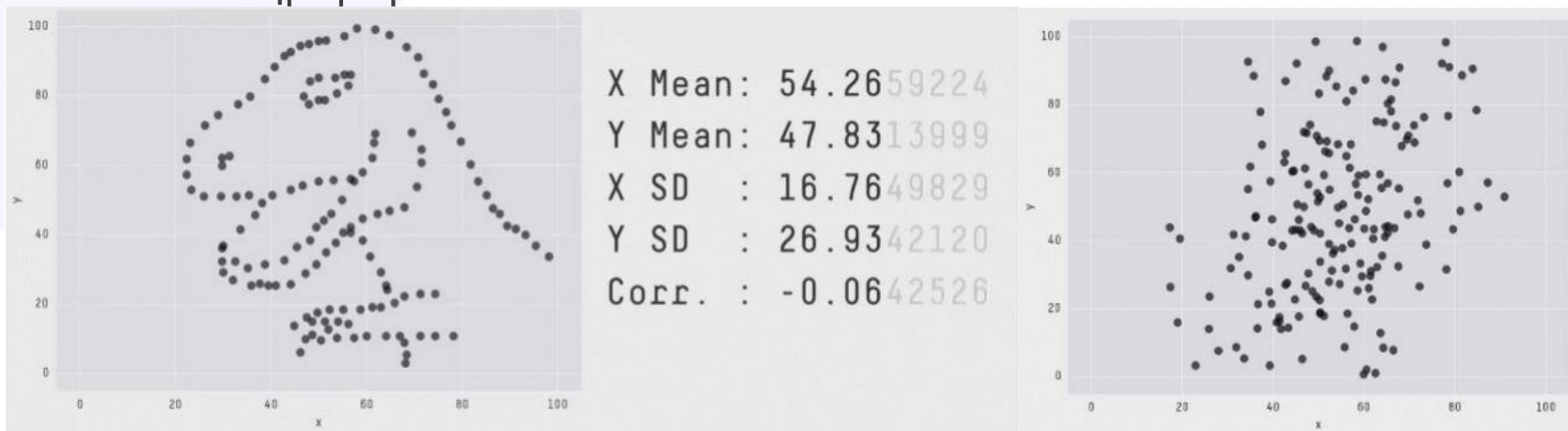
# 데이터 시각화 이유

◀ 공룡 모양의 데이터를 다른 모양으로 변형한 데이터 ▶



# 데이터 시각화 이유

◀ 공통모양의 데이터를 다른 모양으로 변형한 데이터 ▶

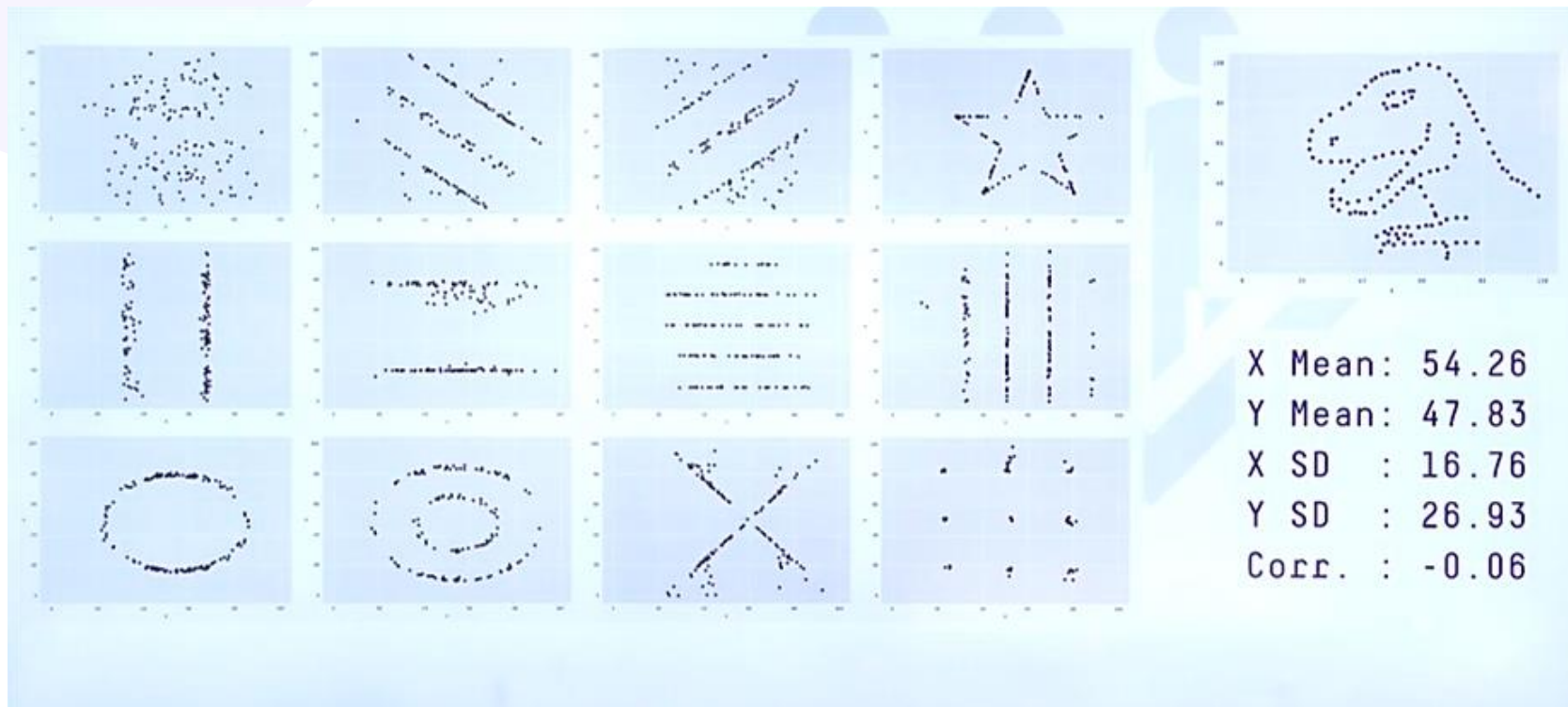


데이터의 패턴은 다르지만 통계 값이 같음



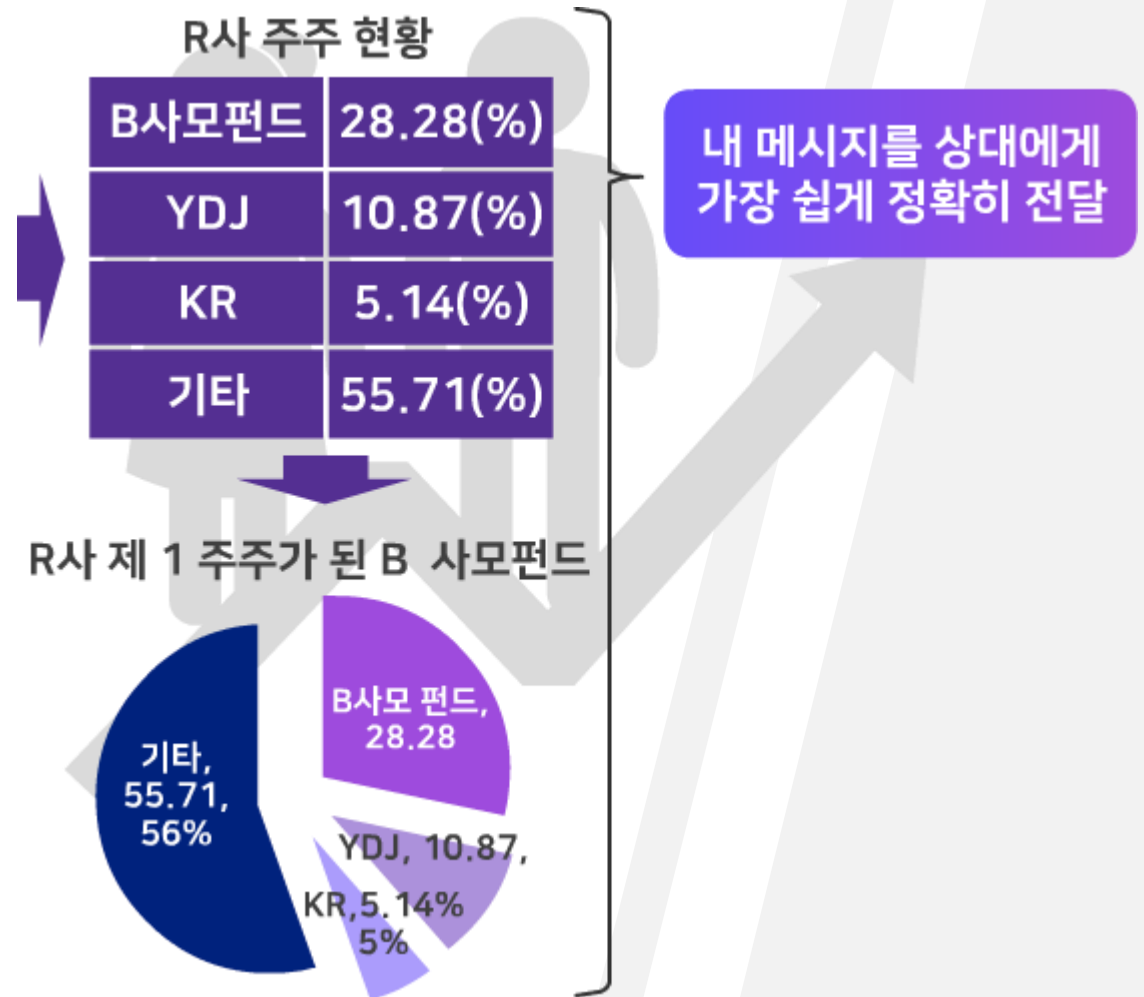
# 데이터 시각화 이유

요약 통계만으로 데이터 집합을 신뢰할 수 없고 눈으로 봐야 함

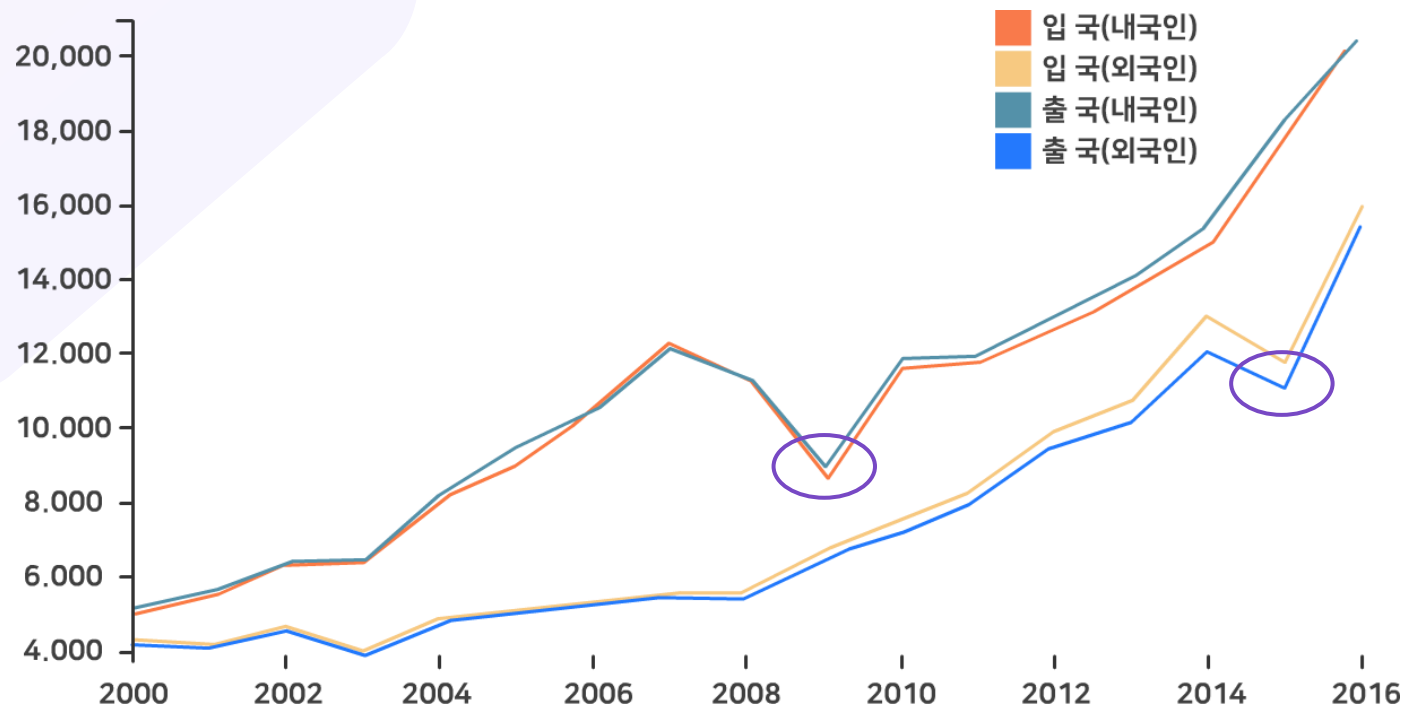


# 데이터시각화 이유

- B사모투자전문회사가 현 주가보다 비싼 가격으로 유상증자 및 전환사채(CB) 인수에 참여해 R의 최대주주로 올라섰다.  
R은 지난 28일 공시를 통해 유상증자 대금 납입으로 최대주주가 YDJ에서 B사모투자전문회사로 변경됐다고 밝혔다. 보고펀드는 R의 제3자 배정 유상증자에 약 423억 원을 출자, 730만 7,478주(28.28%)를 확보하게 됐다. 기존 최대주주인 YDJ 대표의 지분율은 10.87%로 낮아졌다.
- 보고펀드는 이와는 별도로 84억 원 규모의 CB도 인수했다. 코리아글로벌펀드도 15억 원 규모의 CB를 사들였다.  
그러나 이들 투자자의 R 신주 인수가는 주당 5,790원으로 현 주가인 5,500원보다 5.2% 높다. 또 CB의 전환가는 무려 7,300원으로 역시 현 주가에 비해 32.7%나 할증된 가격이다.
- R사 주가가 지난해 12월 11일 유상증자 및 CB 발행 결정 당시 7,130원까지 치솟았지만 이후 내리막길을 걸었기 때문이다. R은 양 대표가 계속 경영권을 갖게 되며, 보고펀드는 2명의 비상근 이사를 파견해 경영에 참여할 예정이다.



# 데이터시각화 이유



- ◆ 2008년~2009년 내국인의 출입국 수 감소
- ◆ 2014년~2015년 외국인 출입국 수 감소
- ◆ 특이사항을 찾으면 결과에 대한 원인을 찾기 위해 드릴다운 분석을 해야 함



# 데이터시각화 개념

데이터시각화란 데이터 분석 결과를 쉽게 이해할 수 있도록  
시각적으로 표현하고 전달되는 과정을 말함



데이터 시각화는 도표(Graph)라는 수단을 통해  
정보를 명확하고 효과적으로 전달하는 것을 목적으로 함

# 데이터시각화 개념

## 데이터시각 화

- 데이터를 분석하고 가공해 사용자가 시각적으로 데이터를 보게 되는 전체 과정임
- 숫자가 아닌 시각적 요소로 바꾸어서 받아들이는 사람이 보다 쉽고 빠르게 이해할 수 있도록 만드는 기술이자 커뮤니케이션 도구임



# 데이터시각화 개념

- 고려사항

01

방대한 데이터를 인지할 수 있는 형태로 시각화하기 위해서는 많은 처리과정과 계산이 필요함

02

데이터시각화의 효과적인 측면을 위해서 언제, 어디까지, 누구를 위해 데이터를 시각화 해야 하는지 그 범위와 대상을 최대한 좁혀서 구체적으로 만들어야 함

03

차트를 보여 주는 것이 목적이 되면 안 되고 어떤 활동 때문에 차트의 선이 오르내리는지를 보여줘야 함



# 데이터시각화 개념

- 인포그래픽이란 무엇인가?

인포그래픽

어휘 자체는 정보를 나타내는 Information과 Graphic의 합성어 정보를 시각화하는 것임



# 데이터시각화 개념

- 대시보드란 무엇인가?

## 대시보드

상황을 모니터링 하거나, 이해를 돕거나,  
둘 다에 사용하는 데이터의 시각적 표시

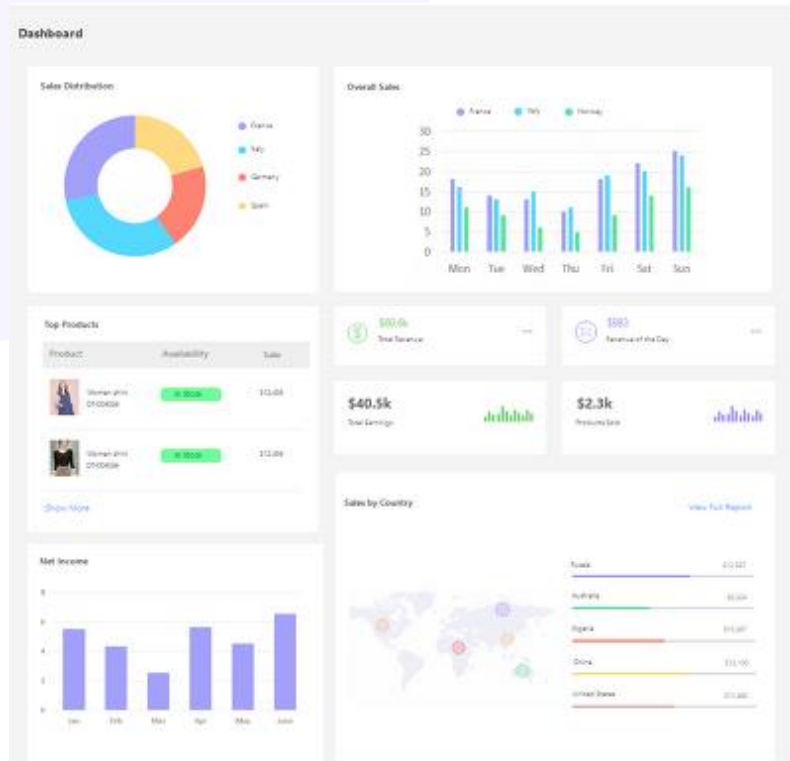
예

- 지역별, 산업별, 신체 부위별 노동자의 보상 청구를 파악할 수 있는 인터랙티브한 도면
- 매주 월요일 아침 임원에게 이메일로 전송되는 핵심지표 PDF 파일
- 실시간으로 지원 센터 통계를 보여주는 대형 벽걸이 현황판
- 영업 관리자가 지역별 성과를 검토하고 전년 대비 올해의 영업 성과를 비교할 수 있는 모바일 앱
- 신규 온라인 서비스 개발을 위해 진행하는 A/B 테스트 결과



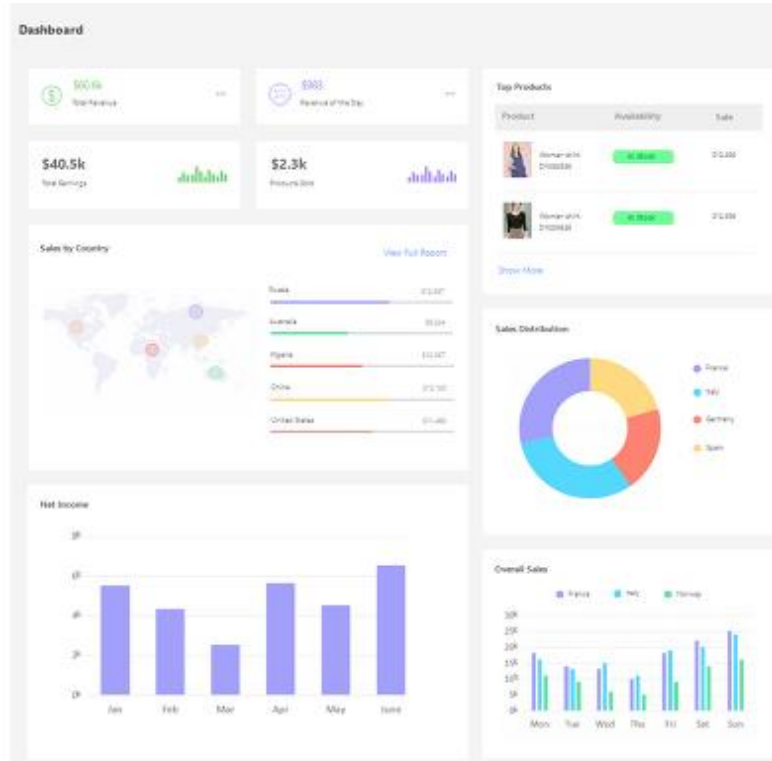
# 데이터시각화 개념

- 대시보드란 무엇인가?



## Don't

Don't display key stats at a random place in the dashboard.



## Do

Display key stats or summarized data at a position where it gets user attention at first glance. The top left is the best position to display important

# 데이터시각화 개념

- 차트란 무엇인가?

차트  
(Chart)

데이터를 그래픽적으로 표현한 것임

Supported chart types



# 데이터시각화 개념

- 차트는 '왜' 사용하나요?

Comparison  
and Ranking  
(비교, 순위)

Part to whole  
(전체 중에 어떤  
파트에 속하는가?)

Trend  
(경향, 추세)

Correlation  
(상관관계)

Distribution  
(범위, 분배,  
퍼진 정도)



# 수업 계획 및 평가

- 수업 계획

본 수업은 파이썬과 파워BI를 통해 시각화 과정을 학습하는 과정임

- 파이썬, 주피터노트북으로 실습(판다스 활용)
- 마이크로소프트 파워BI 실습

수업은 데이터시각화 관련 이론 수업과 실습으로 진행되며 과목 특성상 실습 위주로 진행됨

- 출석을 매우 중시하며, 출석 100%(지각, 결석 없음)일 경우에 한하여 과제 점수에 가산점 부여함

# 수업 계획 및 평가

- 수업 평가

01

출석(20%)

02

참여도(10%)

- 자유게시판, 학습자료실, 수업질의응답에 글 게시 횟수로 시스템에서 자동 카운트되며, 출석 여부도 역시 자동으로 점수가 부여됨

# 수업 계획 및 평가

- 수업 평가

03

중간고사 - 과제형  
평가(30%)

- 필기시험 대체 과제형 시험임
- 파이썬을 활용한 데이터 분석 시각화 파일 제출
- 학습자가 분석하고자 하는 데이터 직접 선정, 학습한 분석방법 활용하여 데이터 분석 후 시각화 파일 작성 및 제출
- 추가 제출 기한(원 제출기한에서 1개월 내)에 제출 시, 취득 점수에서 10%(2주 이내), 20%(2주~1개월) 이내 감점

# 수업 계획 및 평가

- 수업 평가

03

기말고사 - 필기

시험(15%)

- 각 주차별 제공되는 형성평가와 80% 유사하게 출제
- 15문항(객관식 15문항)
- 유형
  - 파이썬과 파워BI 주요 기능
  - 데이터 시각화 관련 데이터 분석 방법론 등
  - 주관식 단답형 : 파이썬 코드 문항 2문제

# 수업 계획 및 평가

- 수업 평가

05

## 실습 과제(25%)

- 매주 수업한 내용을 학습자 스스로 정리하는 자가보고서
- 양식 제공
  - 한글 또는 워드 파일 제공 - 맑은 고딕체, 폰트 12pt
  - 양식에 벗어난 파일 제출 시, 교수 능멸 죄 적용 (50% 감점 처리함)
- 주차 별 실습
  - 총 12회(2~6주, 9주~14주) 주차 별 2점
- 제출 시
  - 자가보고서(1페이지 이상, 정리한 내용 포함), 실습파일
  - 둘 중 하나만 제출 시 1점으로 처리함
  - 100% 출석자 1점 가산점 부여



# 수업 계획 및 평가

- 수업 평가

06

## 실습 과제(25%)

- 기말고사 전 주까지 제출 가능(단, 취득한 원 점수의 10% 감점)
- 제출 시, 반드시 다음 방식으로 제출해야 함  
(방식 어긋 시, 교수 무시 죄 적용 50% 감점)
  - 학번 이름 X주차 실습 보고서  
예 • 20200101\_양진욱\_2주차 실습 보고서
  - 20200101\_양진욱\_2주차 실습 파일



자세한 사항은 데이터처리와 분석 수업계획서  
참조

# 수업 계획 및 평가

- 무엇을 배우나요?

## 1주차

### 강의소개/데이터시각화란 무엇인가

- 데이터시각화 개념, 기능적이면서도 아름다운 대시보드 설계 및 보고서 작성법
- 데이터 시각화 이론(핵심차트 10) 및 텍스트 시각개체

## 2주차

### 데이터프레임(Data Frame)과 Subset Observation

- 쥬피터 노트북 설치
- 판다스 데이터프레임 생성, 판다스 데이터 프레임 비교연산자로 색인, 일부 컬럼 기준으로 데이터 가져오기

# 수업 계획 및 평가

- 무엇을 배우나요?

## 3주차

### 데이터 요약, 결측치 다루기, 새로운 컬럼 만들기

- 기본 통계(Value\_counts, Sum, Mean, Median),  
판다스 Apply 함수, Lambda 익명함수, 결측치 다루기,  
Assign으로 새로운 컬럼 만들기

## 4주차

### 데이터 재형성, 데이터 세트 합치기, 데이터 집계(sql 문) 활용

- Reshape 함수, melt, Pivot으로 깔끔한 데이터 만들기,  
pd.concat 시리즈
- Merge로 데이터 합치기, Left, Right, Inner, Outer Option 사용,  
Groupby 활용 데이터 집계 활용하기

# 수업 계획 및 평가

- 무엇을 배우나요?

## 5주차

### 시계열 데이터 분석과 데이터 시각화 plotting 기능 활용(1)

- 시계열 데이터 분석하기, 판다스로 막대그래프와 히스토그램, 상자수염그림과 산점도 그리기

## 6주차

### 데이터시각화 plotting 기능 활용(2), 실제 데이터 실습하기(1)

- Scatter Matrix Plot 산점도와 커널 밀도함수 그리기
- 판다스로 서울 코로나 19 발생동향 분석하기(1)

## 7주차

### 실제 데이터 실습하기(1)

- 서울시 코로나 19 발생동향 분석하기 (2)

# 수업 계획 및 평가

- 무엇을 배우나요?

## 8주차

### 중간시험 대체 과제

- 파이썬 활용한 시각화 파일 제출(30%)

## 9주차

### 파워BI 데스크탑, 데이터 가져오기 및 변환, 데이터 모델링

- 파워BI 소개 및 기본 이해, BI와 Self Service BI에 대한 기본 이해
- 파워BI 화면구성 및 옵션, 데이터베이스 연결 및 관계설정, 데이터 변환

## 10주차

### 데이터시각화와 보고서 게시, 데이터원본

- 보고서 페이지 설정 및 시각화 순서, 대시보드 생성 및 공유, 데이터베이스와 SQL, EXCEL 파일과 폴더, 웹데이터 가져오기

# 수업 계획 및 평가

- 무엇을 배우나요?



## 11주차

### 복합모델, 데이터변환, 시각화

- 복합모델과 다대다 관계, 파워쿼리 편집기 활용 및 보고서 작성, 클러스터와 예측 분석

## 12주차

### DAX 소개, 슬라이서와 필터

- DAX 소개 및 활용, DAX 함수, 맵 시각화(도형 맵, ArcGIS 맵), 슬라이서와 필터

## 13주차

### 책갈피와 선택, 단추, 대시보드 테마, 사용자 지정 시각적 개체

# 수업 계획 및 평가

- 무엇을 배우나요?

14주차

작업영역, 데이터 흐름, 행수준 보안, 관리포털

기말 시험

필답고사

- 파이썬과 파워BI
  - 15%, 객관식 15문항





# 대시보드 설계와 보고서 구성법





# 대시보드 설계와 보고서 구성법

- 대시보드 설계

시각화 및 게시 프로세스  
이해

데이터 가져오기 및 정리



보고서 시각화하기(엑셀, R, 파이썬, 파워BI)



반응형 보고서 만들기(파워BI) – 슬라이서, 필터



POWER BI SERVICE와 POWER BI MOBILE로 게시

# 대시보드 설계와 보고서 구성법

- 대시보드 설계

01

데이터  
가져오기

- EXCEL, CSV file



# 대시보드 설계와 보고서 구성법

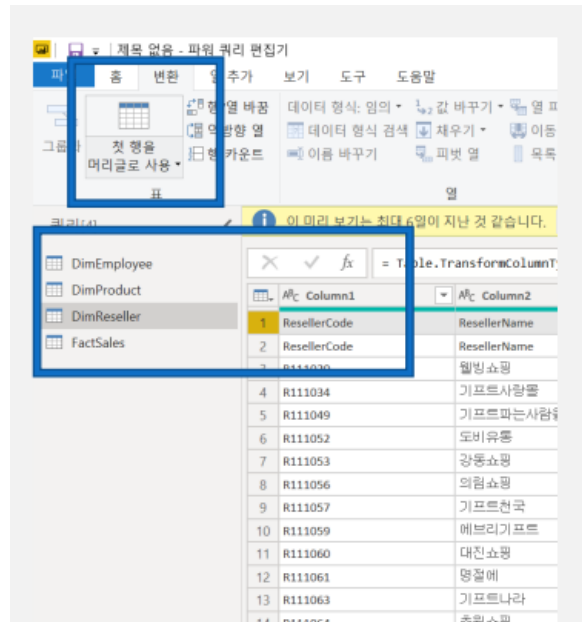
- 대시보드 설계

02

데이터 사전 작업(가장 시간 소요가 되면서 중요한 단계)

- 파이썬 판다스(Pandas)
- 파워BI : 파워 쿼리 편집기 이용, 데이터 모델 편집기 이용

Pandas



# 대시보드 설계와 보고서 구성법

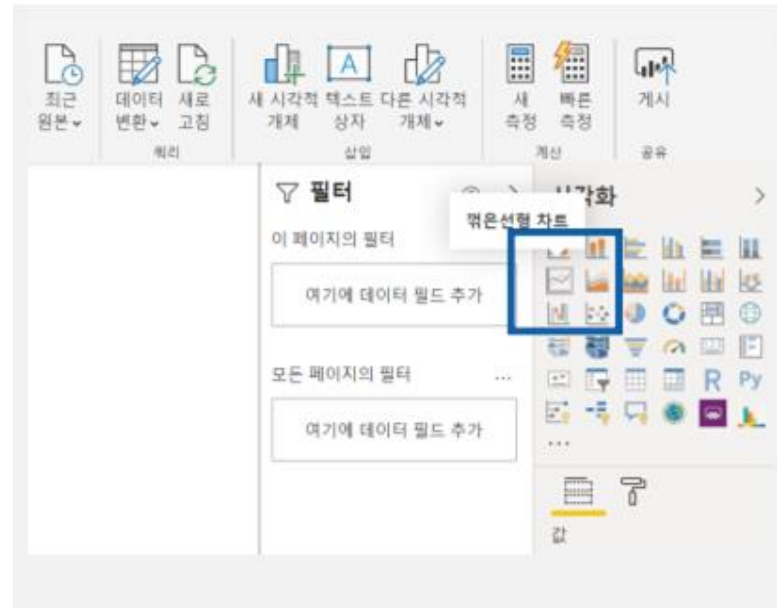
- 대시보드 설계

03

보고서 작성(시각화 개체 활용)

- 파이썬 MATPLOTLIB
- 파워BI 슬라이서, 필터

matplotlib



# 대시보드 설계와 보고서 구성법

- 대시보드 설계

03

보고서 작성(시각화 개체  
활용)

슬라이서



시각적 개체로 제공되는  
필터

필터



시각적 개체의 특성으로  
제공되는 데이터 필터

# 대시보드 설계와 보고서 구성법

- 대시보드 설계

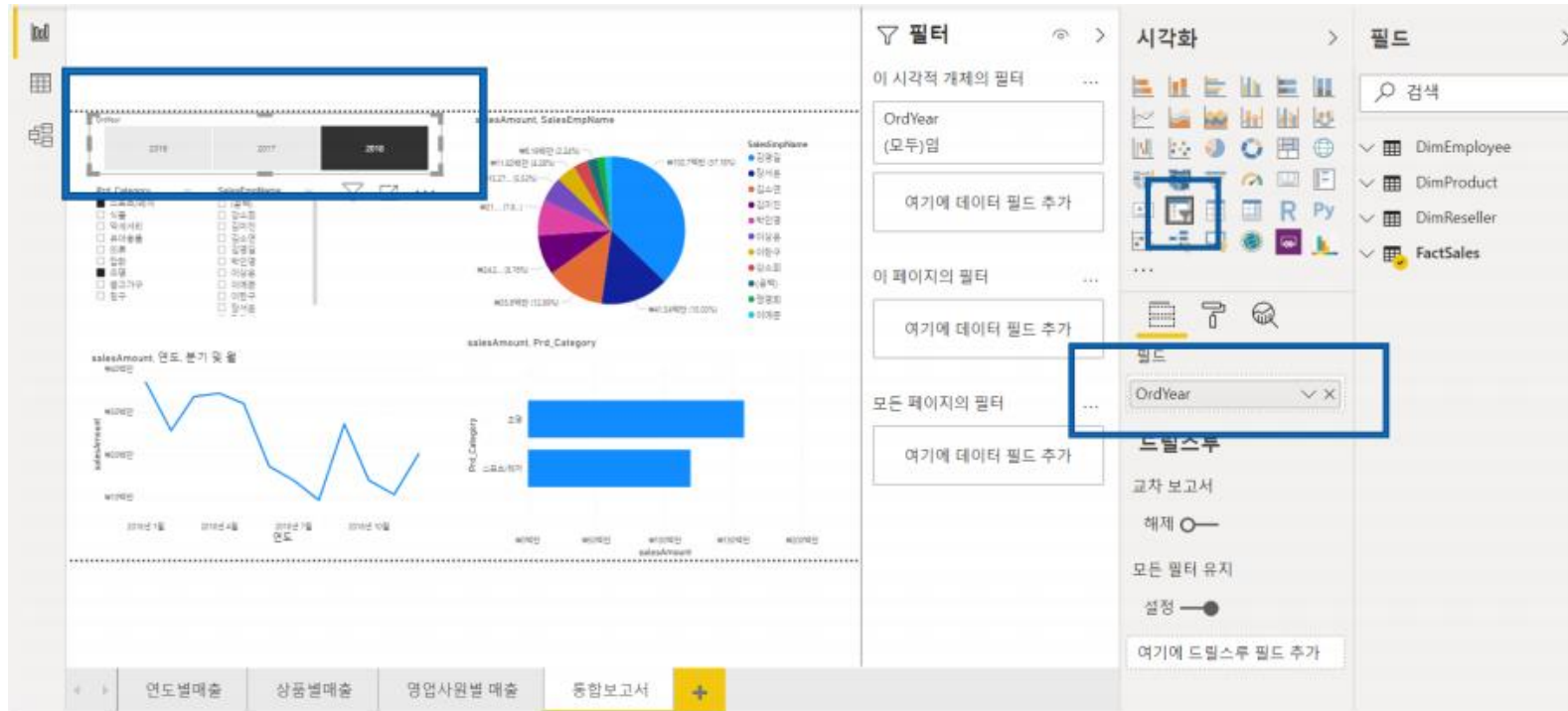
슬라이서

시각적 개체로 제공되는 필터



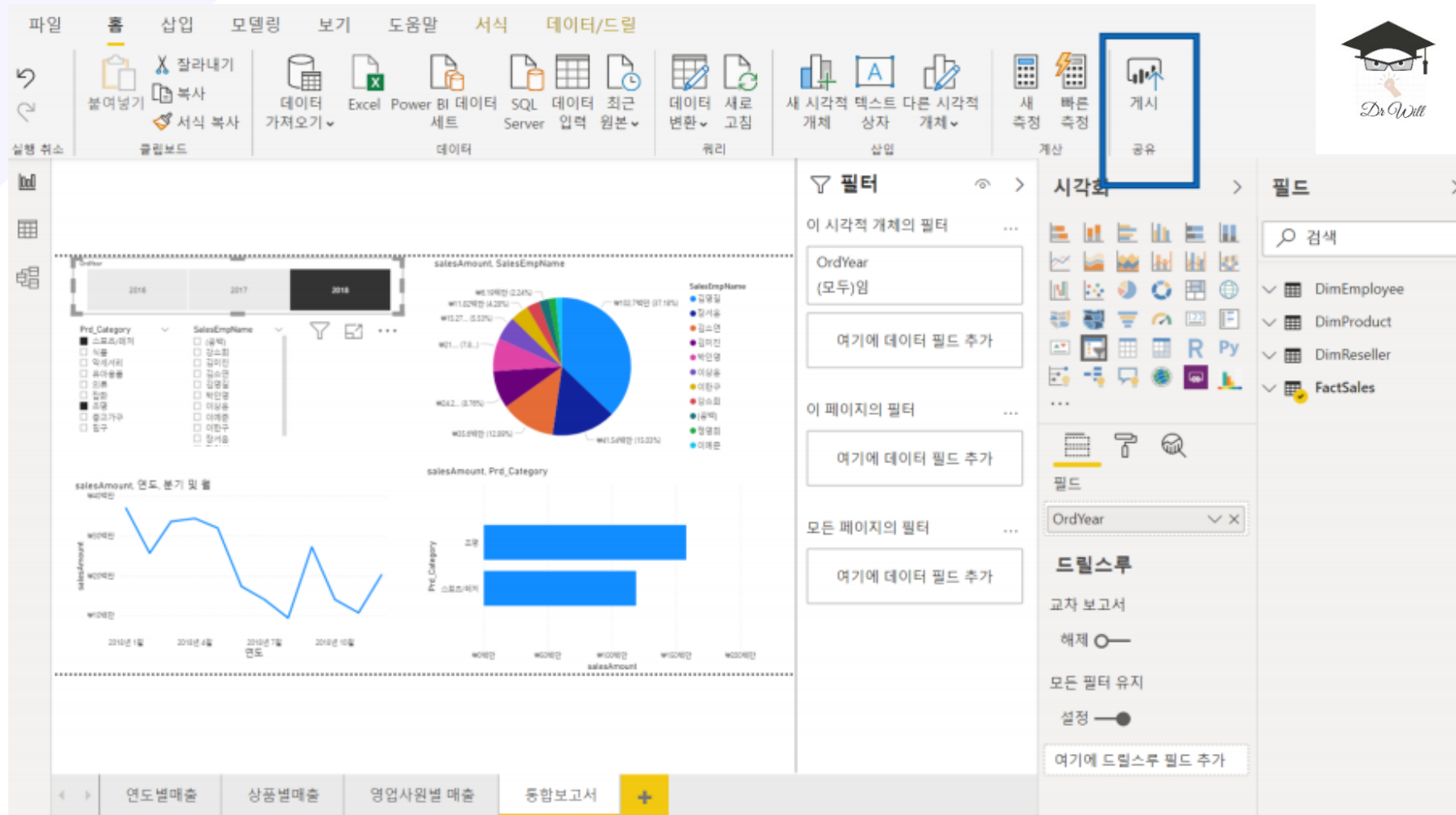
필터

시각적 개체의 특성으로 제공되는 데이터 필터



# 대시보드 설계와 보고서 구성법

- 보고서 구성법
  - 데이터의 이해





# 데이터 시각화 이론





# 데이터 유형

- 데이터 유형별 시각화

시각화를 하기 위해서는 데이터 유형에 맞는 조합을 해야 함

데이터는 3가지 유형이 있음

범주  
형  
데이  
터

순서  
형  
데이  
터

정량  
적  
데이  
터

# 데이터 유형

- 데이터 유형별 시각화

범주형 데이터

순서형 데이터

정량적 데이터

- 사물을 나타내며, 이 사물은 어떤 숫자 값도 없는 상호 배타적인 레이블을 의미함
- 명목형 데이터(Nominal)라고도 함
- 범주형 변수를 그 값의 형태나 의미에 따라 세부 유형으로 나눌 수 있음
  - 텍스트(때에 따라 Ordinal)
  - 지역(Geographic)
  - 날짜(Temporal)
  - 숫자(Numeric)로 구분

예

남자, 여자/국어, 수학, 과학/성공,  
실패

# 데이터 유형

- 데이터 유형별 시각화

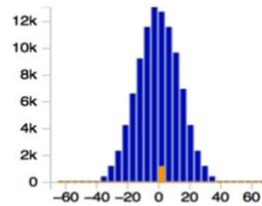
범주형 데이터

순서형 데이터

정량적 데이터

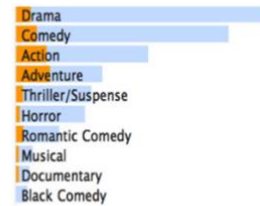
1D

Numeric



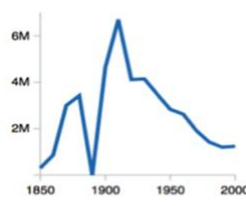
Histogram

Ordinal



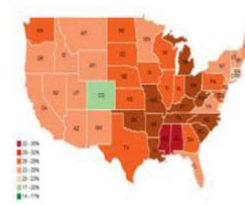
Bar Chart

Temporal



Line Graph /  
Area Chart

Geographic



Choropleth Map

# 데이터 유형

- 데이터 유형별 시각화

범주형 데이터

순서형 데이터

정량적 데이터

- 분명한 **순서**가 있다는 점을 제외하면 범주형 데이터와 유사함
- 각 값 사이의 간격은 일정하지 않음

예

나이 기준

- 영아, 유아, 어린이, 청소년, 청년, 장년, 노인

성적 기준

- A, B, C, D, E, F

# 데이터 유형

- 데이터 유형별 시각화

범주형 데이터

순서형 데이터

정량적 데이터

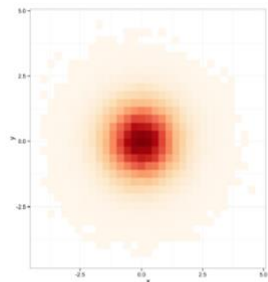
**Numeric**

**Ordinal**

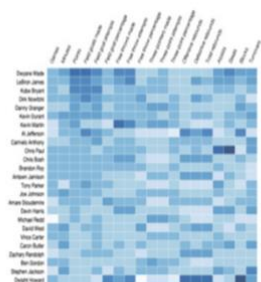
**Temporal**

**Geographic**

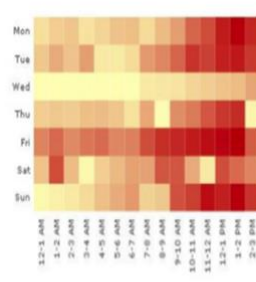
2D



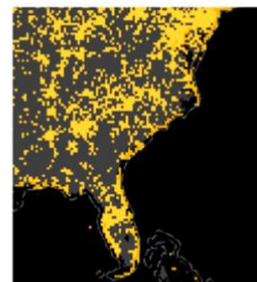
Binned  
Scatter Plot



Heatmap



Temporal  
Heatmap



Geographic  
Heatmap

# 데이터 유형

- 데이터 유형별 시각화

범주형 데이터

순서형 데이터

정량적 데이터

- 정량적(수치) 데이터는 숫자로서 측정, 분류 가능한 데이터

예 키, 몸무게, 나이,  
매출액

- 이산 데이터와 연속 데이터의 2가지 방법으로 표현할 수 있음

- 이산 데이터는 미리 정의된 정확한 포인트로 제시되어  
데이터 값 '사이'에는 아무것도 없음

예 나이 : 20,21,22,23,24,25(O)/25.4(X)

- 연속 데이터는 가능한 중간 값의 무한한 수가 있음

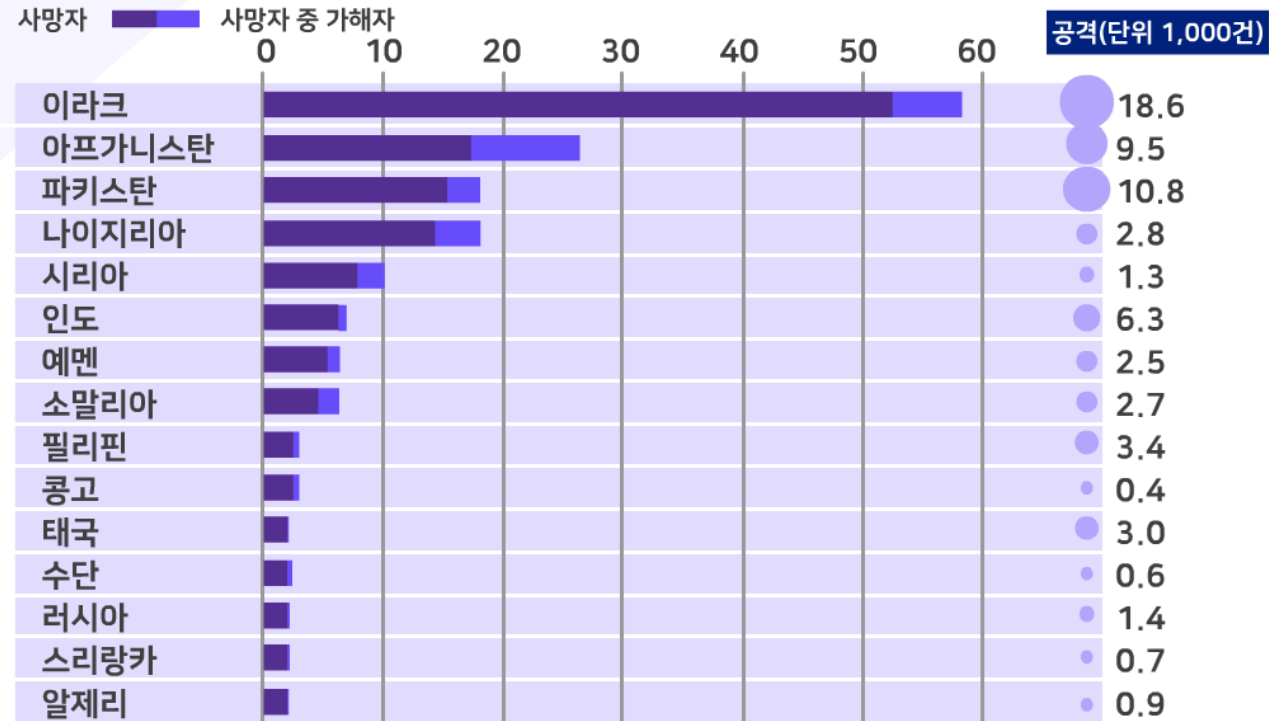
예 키 : 150.5, 160.44, 170.4345....

# 데이터 유형

## • 데이터 유형별 시각화

끔찍한 기록

2003~2015년 테러 사망자가 가장 많은 나라(단위: 1,000명)



데이터	데이터 유형	규약화	참고
나라	범주형 데이터	위치	나라마다 행별로 구분된 막대로 표시되어 있다(사망자 수 집계로 정렬).
사망자	정량적 데이터	길이	막대 길이는 사망자 수를 보여준다.
사망 유형	범주형 데이터	색	짙은 파란색은 희생자의 사망 수, 하늘색은 가해자의 사망 수를 나타낸다.
공격	정량적 데이터	크기	우측 원은 공격 건수에 따라 크기가 다르다.



# 데이터시각화-색상

색상은 데이터시각화에서 이해해야 할  
가장 중요한 요소 중 하나임에도 오용하는 경우가 많음

**순차적**  
색상을 낮음에서 높음으로 정렬



**발산형**  
중립적인 중간점이 있는 순차적인  
2가지 색상



**범주형**  
개별 비교를 위한 색상 대비



**강조**  
원가를 강조하기 위한 색 사용



**경고**  
독자에게 경고하기 위한 색 사용



- 생기를 불어넣는 목적만을 위해서 색을 사용해서는 안됨
- 색상은 의도를 갖고 사용해야 함
  - 색으로 독자의 관심을 끌거나
  - 데이터 일부를 강조하거나
  - 다양한 카테고리를 구별함



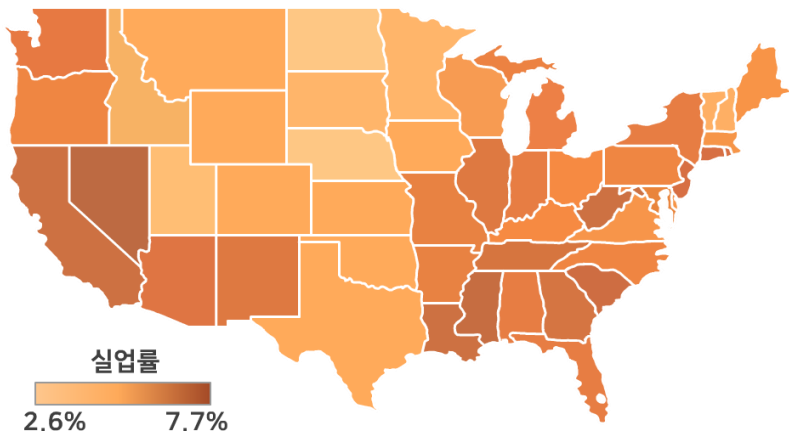
# 데이터시각화-색상

- 순차적 색상

밝은 색부터 어두운 색까지 단색 사용

◀ 순차적 색 체계를 사용한 주별 실업률 ▶

주별 실업률



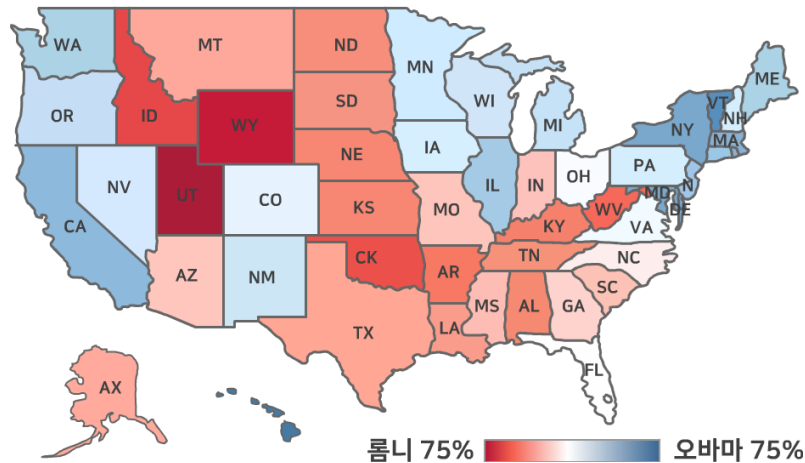
- 주황색이 어두울수록 실업률이 높음
- 밝을수록 실업률이 낮아지는 것 표현함

# 데이터시각화-색상

- 발산형 색상

중간점에서 발생하는 범위를 표시할 때 사용함

## ◀ 각 주별 투표자 감정 등급 ▶



- 상반되는 2가지 측정 범위를 나타내거나 2가지 카테고리 중에서 1가지 측정 범위를 나타낼 때 쓰임
- 미국 주별로 유권자가 민주당이나 공화당 중 어느 정당에 투표했는지를 보여주는 그래프임

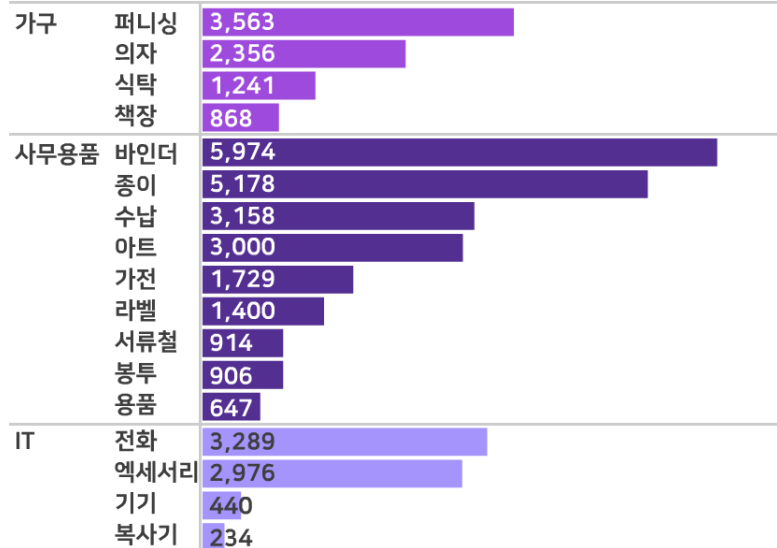
# 데이터시각화-색상

- 범주형 색상

카테고리(범주)마다 다양한 색조로 구별함

카테고리별  
수량

카테고리별 수량

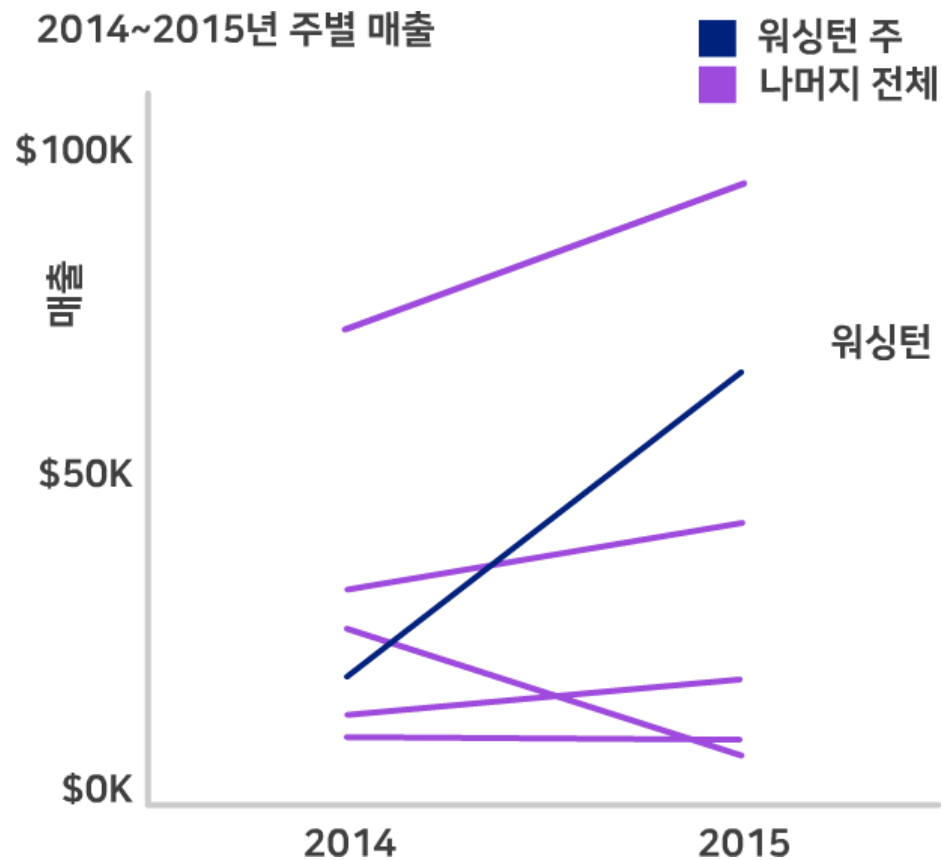


- 가구, 사무용품, IT 관련  
카테고리를 정해 표시

# 데이터 시각화-색상

- 강조

특정 데이터 포인트, 표 안의 텍스트, 꺾은선 차트의 특정 선, 막대 차트의 특정 막대를 강조할 때 사용

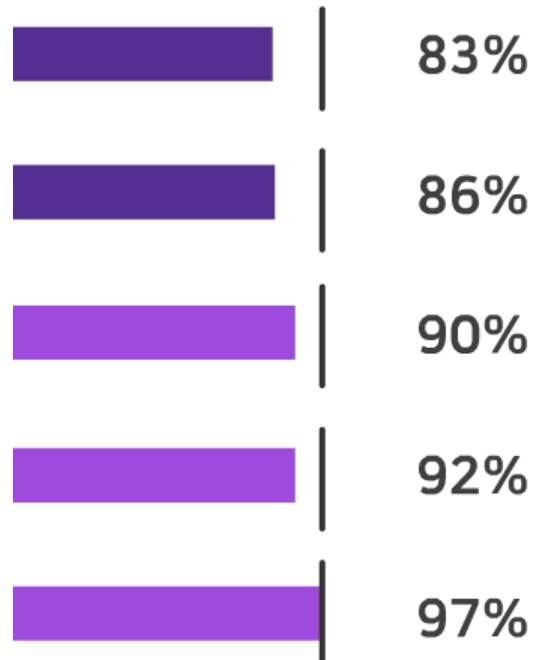


# 데이터시각화-색상

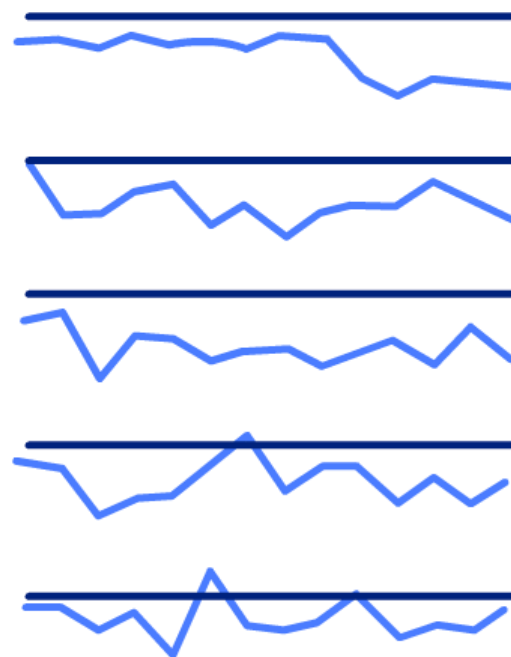
- 경고

사용자의 주목을 끌어야 할 때 사용할 때 사용하며,  
밝고 경종을 울리는 색 사용이 좋음

실제 대비 전망



14일간 추세



최근 조사일

3일 전  
19일 전 ●  
2일 전  
7일 전  
13일 전 ●

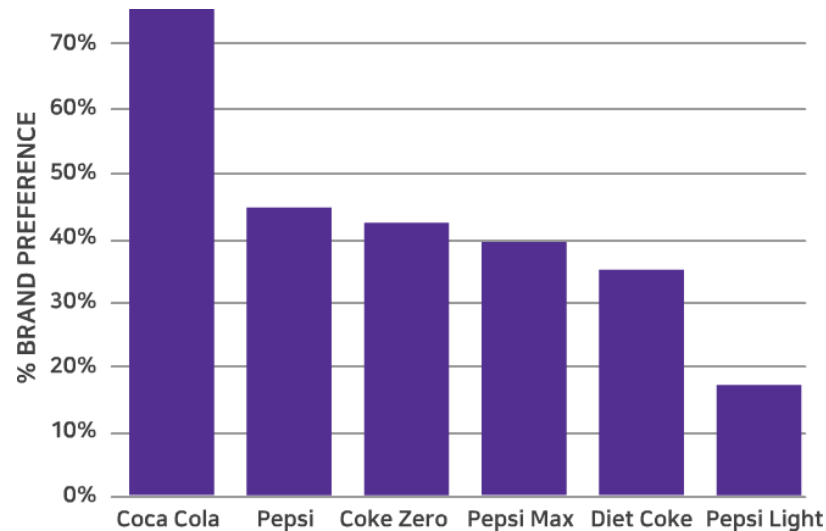


# 핵심차트

- 막대차트(Bar Chart)

막대 차트는 **길이로 척도**를 나타냄

- 사람은 공통적인 기준선으로부터 길이가 **아주 조금만 차이가 나더라도 빨리 인지**할 수 있음
- 막대 차트는 **카테고리를 비교하는 가장 효과적인 방법**으로 데이터 시각화에 널리 사용됨

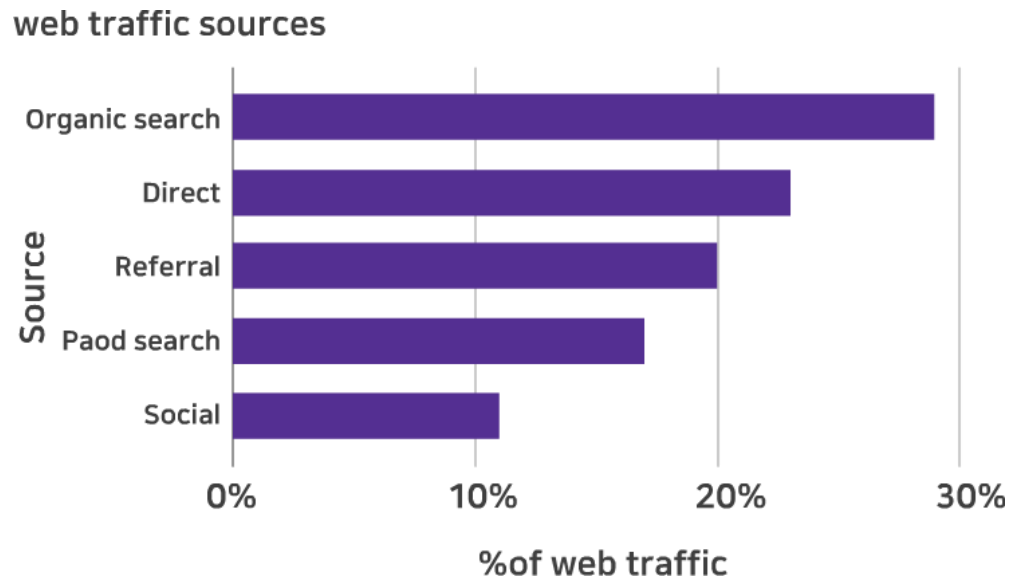


# 핵심차트

- 막대차트(Bar Chart)

막대 차트는 **길이로 척도**를 나타냄

- 수평이나 수직 방향으로 사용할 수 있음
- 막대 차트는 최대/최소 항목을 찾을 때 가장 흔히 사용하므로, 정렬해서 비교하면 매우 유용함



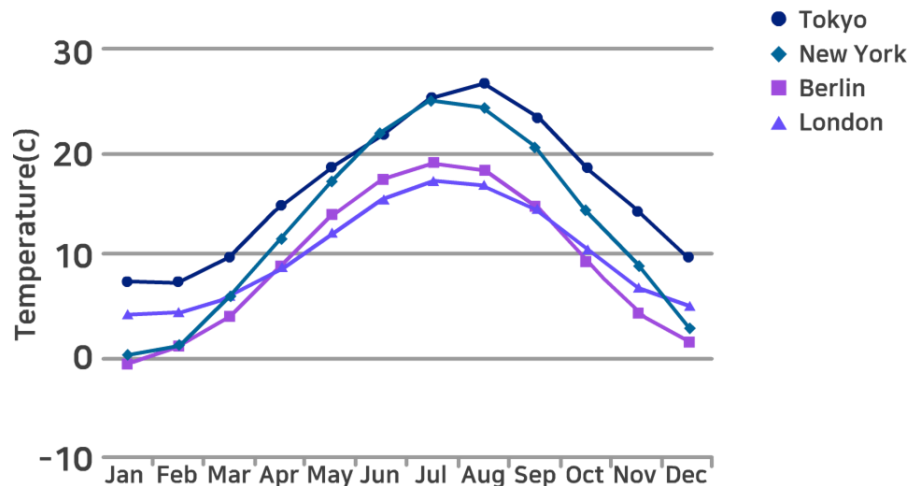
# 핵심차트

- 선도표(Line Chart, 시계열 꺾은선 차트)

직선 세그먼트로 연결된 '마커'라는 이름의 일련의 데이터 지점으로 정보를 표시하는 차트의 일종임

- 장기적인 변화를 볼 때 이용
- 수평인 X축은 시간을 나타내며, 수직인 Y축에는 척도를 표시
- 선의 높이와 경사로 추세를 볼 수 있음

## Monthly Average Temperature



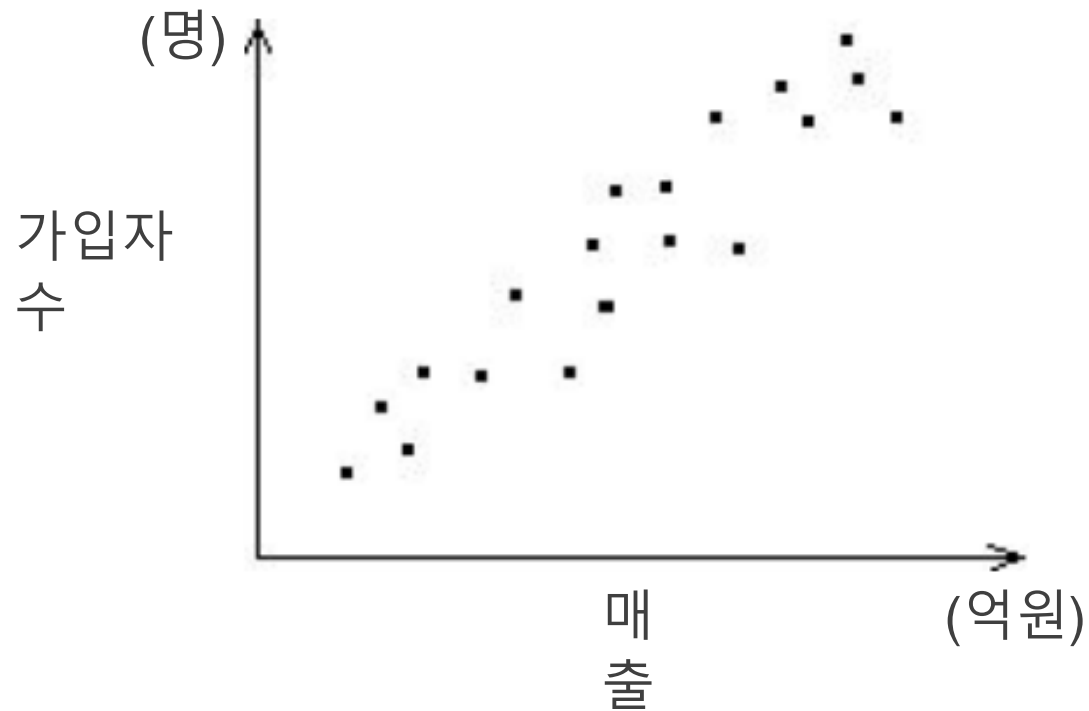


# 핵심차트

- 산포도 차트(Scatter Chart)

서로 다른 두 척도를 비교할 수 있음

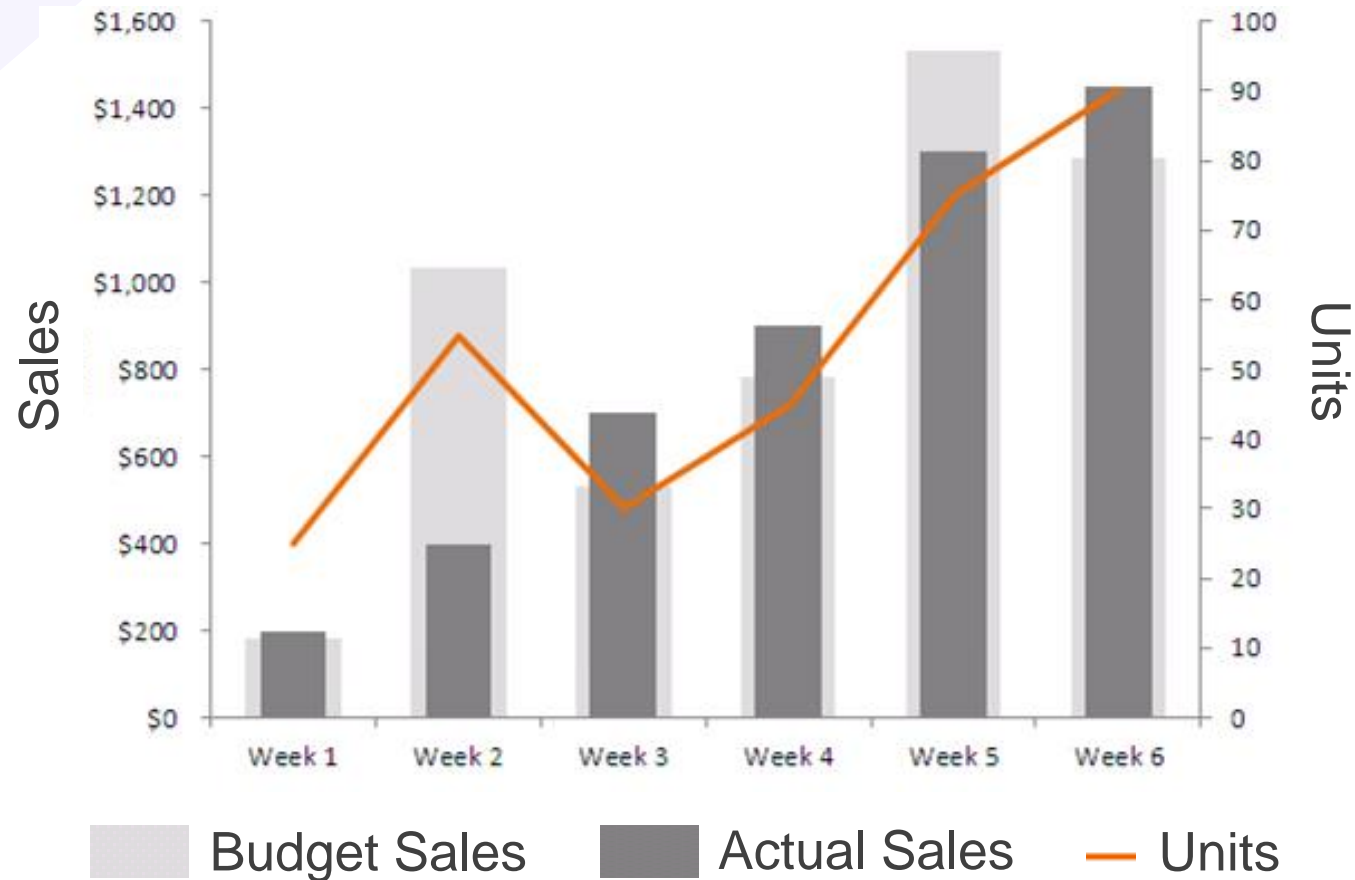
- 각 척도는 수평 축과 수직 축 상의 위치를 나타냄
- 산포도는 2가지 변수 간의 관계를 관찰할 때 유용함



# 핵심차트

- Dual axis Chart/Pareto Chart

다른 크기와 척도를 가진 두 변수간 관계를 보여주고자 할 때  
사용함



# 핵심차트

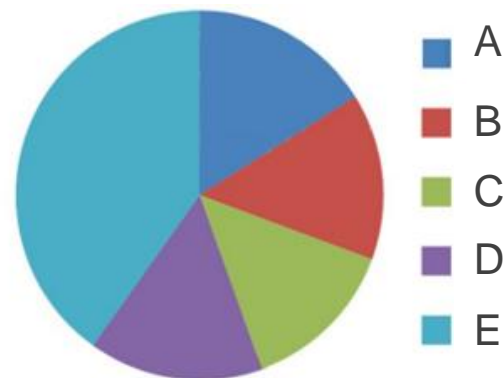
- 파이차트(Pie Chart)

각도, 영역, 아치로 데이터를 표현하여  
전체에서 부분이 차지하는 비교를 표시함

- 비율 등 표시에는 권장함
  - 단, 구분이 5개를 넘는 것은 권장하지 않음
- 특히 파이차트는 비슷한 크기가 있을 때 구분이 어려움
  - 시각 처리 오류 발생 가능



파란 영역은 몇 비중이 몇 퍼센트인가?



3번째로 큰 영역은 어떤  
카테고리인가?

# 핵심차트

- 버블 차트(Bubble Chart)

## 산포도의 변형

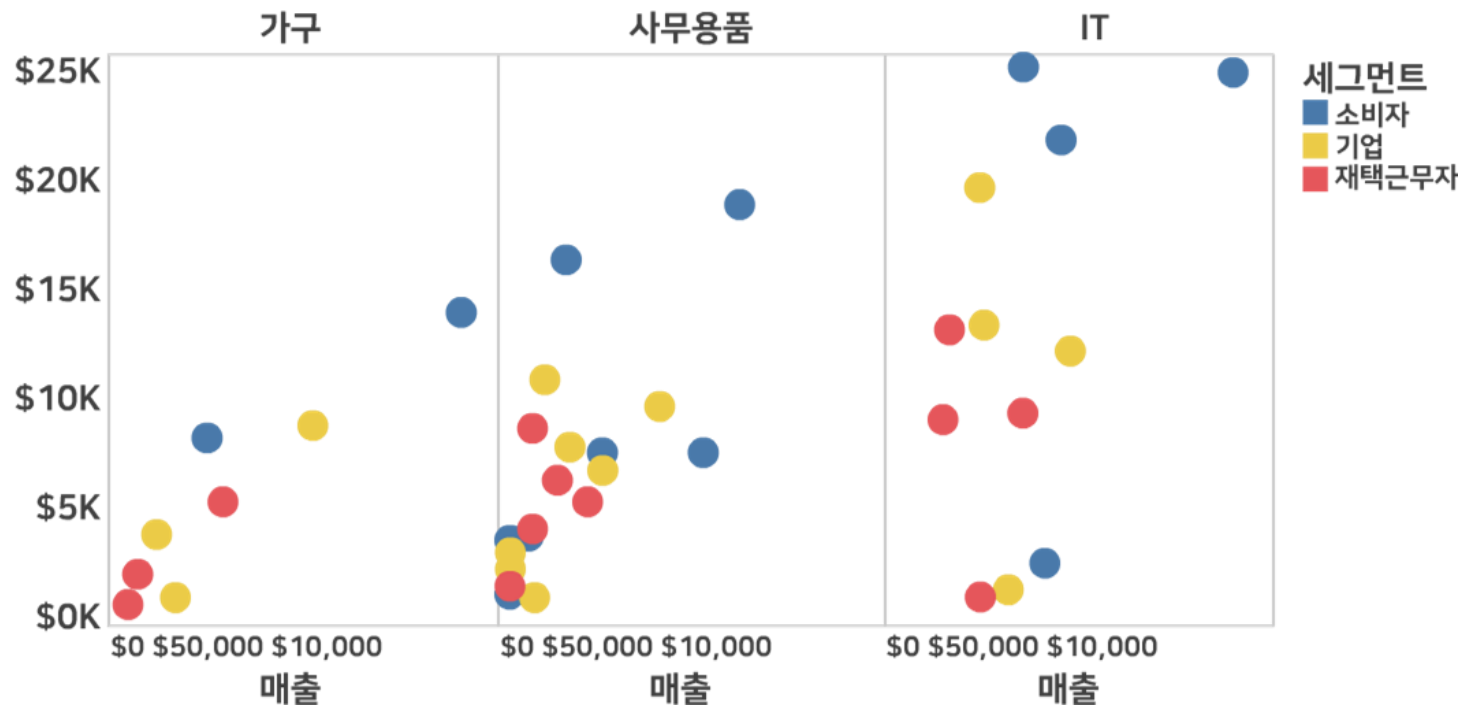
- 버블(원)의 사이즈 변화로 추가적인 정보를 제공함
- 정확한 정량 비교는 어렵다는 단점이 있음
  - 카테고리 별로 단일 산포도를 3개 패널로 그리면 이전 차트보다 결과가 눈에 띄게 보임
  - IT 제품이 매출대비 수익이 큰 것을 명확히 확인 가능

# 핵심차트

- 버블 차트(Bubble Chart)

## 산포도의 변형

카테고리와 세그먼트별 매출 및 수익

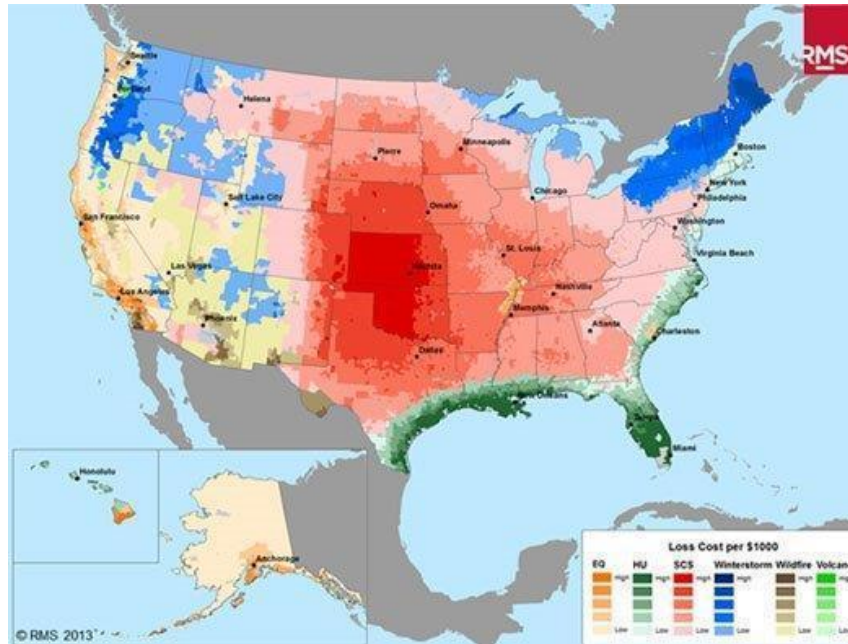


# 핵심차트

- 히트맵(Heat Map)

색으로 데이터를 표현하여 숫자 없이 표의 파이를 강조함

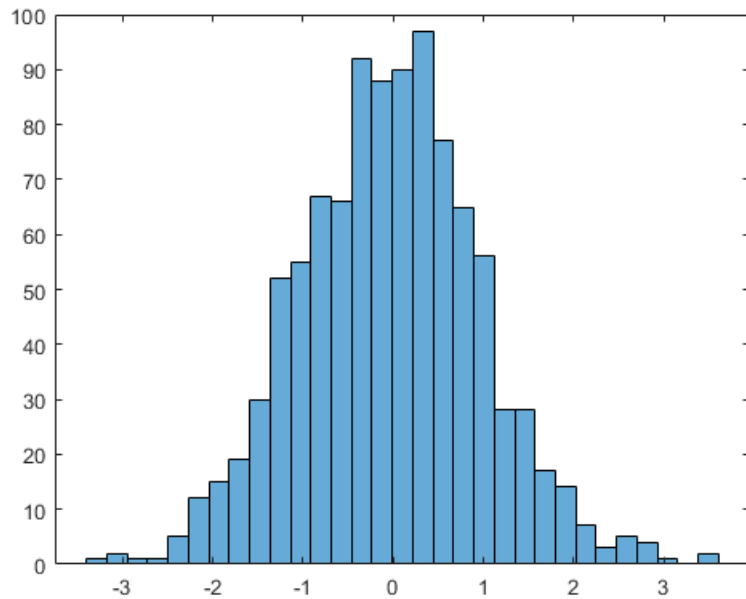
- 열을 뜻하는 Heat와 지도를 뜻하는 Map을 결합시킨 단어임
- 색상으로 표현할 수 있는 다양한 정보를 일정한 이미지 위에 열 분포 형태로 시각화한 것이 특징임



# 핵심차트

- 히스토그램(Histogram)

표로 되어 있는 도수 분포를 정보 그림으로 나타낸 것임



- 도수분포표를 그래프로 나타낸 것임
  - 연속형 데이터 표현
- 높이로 데이터를 표현하고 분포를 표시함

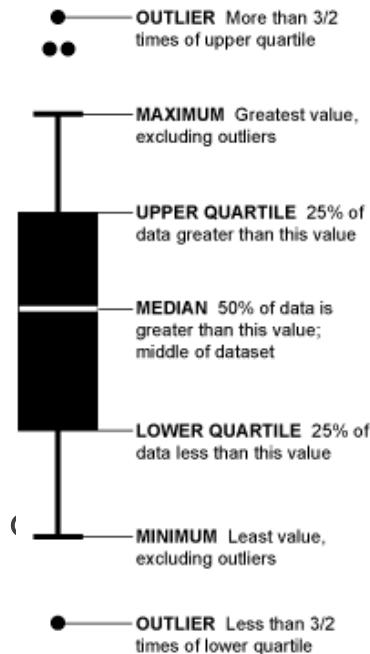
# 핵심차트

- 상자수염 그림, 박스 플롯(Box Plot)

기술통계학에서 5가지 중요한 수치(Big 5)를  
수치적 자료로 표현하는 자료임

- 5가지 수치 요약

- 최솟값 : 제 1사분위에서 1.5 IQR을 뺀 위치임
- 제 1사분위(Q1) : 25%의 위치를 의미함
- 제 2사분위(Q2) : 50%의 위치로  
중앙값(median)을 의미함(중위수)
- 제 3사분위(Q3) : 75%의 위치를 의미함
- 최댓값 : 제 3사분위에서 1.5 IQR을 더한 위치



사분위 범위 수(IQR) : 데이터의 중간 50%(Q3 - Q1)



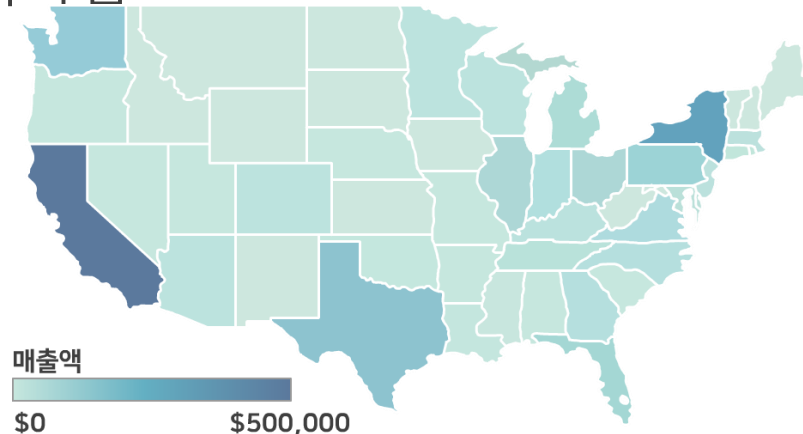


# 핵심차트

- 지역 지도 차트(Geo Chart)

A Geo Chart is made explore location based distributed data.

- Map of a country, a continent, or a region with areas identified in one of three ways
  - The region mode colors whole regions, such as countries, provinces, or states.
- 지역별로 미리 정의된 영역 내에서 음영이나 색의 차이로 그 영역의 값이나 카테고리를 나타냄



# 텍스트 시각 개체

- 텍스트 개체 사용하기

## 테이블과 행렬

테이블은 논리적으로 연속된 행과 열에서 관련된 데이터를 포함하는 표

Category	This Year Sales Status	Average Unit Price	Last Year Sales	This Year Sales	This Year Sales Goal
010-Womens	●	\$7.30	\$2,680,662	\$1,787,958	\$2,680,662
020-Mens	●	\$7.12	\$4,453,133	\$4,452,421	\$4,453,133
030-Kids	●	\$5.30	\$2,726,892	\$2,705,490	\$2,726,892
040-Juniors	●	\$7.00	\$3,105,550	\$2,930,385	\$3,105,550
050-Shoes	●	\$13.84	\$3,640,471	\$3,574,900	\$3,640,471
060-Intimate	●	\$4.28	\$955,370	\$852,329	\$955,370
070-Hosiery	●	\$3.69	\$573,604	\$486,106	\$573,604
080-Accessories	●	\$4.84	\$1,273,096	\$1,379,259	\$1,273,096
090-Home	●	\$3.93	\$2,913,647	\$3,053,326	\$2,913,647
100-Groceries	●	\$1.47	\$810,176	\$829,776	\$810,176
<b>Total</b>	●	<b>\$5.49</b>	<b>\$23,132,601</b>	<b>\$22,051,952</b>	<b>\$23,132,601</b>

# 텍스트 시각 개체

- 텍스트 개체 사용하기
  - 테이블을 사용하는 경우

01

자세한 데이터 및 정확한 값을 보고  
비교

- 시각적 표현 대신

02

테이블 형식으로 데이터를  
표시

03

숫자 데이터를 범주별로  
표시

# 텍스트 시각 개체

- 텍스트 개체 사용하기

## 행렬

여러 차원에 걸쳐(계단형 레이아웃 지원) 데이터를  
유의미하게 표시하기가 더 쉬움

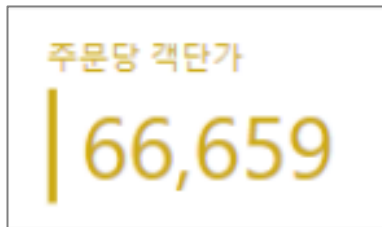
Region	East		West		Total	
Sales Stage	Opportunity Count	Revenue	Opportunity Count	Revenue	Opportunity Count	Revenue
Lead	114	\$473,887,837	52	\$256,159,114	166	\$730,046,951
Small	44	\$55,009,105	13	\$11,979,738	57	\$66,988,843
Medium	31	\$120,678,637	18	\$72,871,697	49	\$193,550,334
Large	39	\$298,200,095	21	\$171,307,679	60	\$469,507,774
Proposal	13	\$59,970,924	10	\$43,032,669	23	\$103,003,593
Small	3	\$4,770,862	3	\$3,720,287	6	\$8,491,149
Medium	6	\$25,607,581	5	\$21,456,937	11	\$47,064,518
Large	4	\$29,592,481	2	\$17,855,445	6	\$47,447,926
Finalize	5	\$30,696,428	4	\$21,176,185	9	\$51,872,613
Small	1	\$1,693,585			1	\$1,693,585
Medium			2	\$7,926,517	2	\$7,926,517
Large	4	\$29,002,843	2	\$13,249,668	6	\$42,252,511
Total	132	\$564,555,189	66	\$320,367,968	198	\$884,923,157

# 텍스트 시각 개체

- 텍스트 개체 사용하기
  - 한 줄 카드와 여러 줄 카드

## 한 줄 카드

하나의 타일에 한 가지  
중요한 숫자를 확인할 때  
사용



## 여러 줄 카드

하나의 타일에 여러 정보  
제공 시 사용



과목명	데이터시각화	주차명	01. 데이터시각화란 무엇인가?		페이지번호	05_01_01																
학습목차		학습평가				화면설명																
<div>들어가기<ul style="list-style-type: none"><li>- 인트로</li><li>- 학습개요</li></ul></div> <div>데이터시각화 과목 소개<ul style="list-style-type: none"><li>- 데이터시각화 이유</li><li>- 데이터시각화 개념</li><li>- 수업 계획 및 평가</li></ul></div> <div>대시보드 설계와 보고서 구성법<ul style="list-style-type: none"><li>- 대시보드 설계와 보고서 구성법</li></ul></div> <div>데이터시각화 이론<ul style="list-style-type: none"><li>- 데이터 유형</li><li>- 데이터시각화 - 색상</li><li>- 핵심차트</li><li>- 텍스트시각 개체</li></ul></div> <div>평가하기<ul style="list-style-type: none"><li>- 학습평가</li></ul></div> <div>정리하기<ul style="list-style-type: none"><li>- 학습정리</li></ul></div>		<div>학습한 내용을 바탕으로 다음 문제를 풀어 보세요.</div> <table><thead><tr><th>번호</th><th>문제</th><th>정답</th><th>해설</th></tr></thead><tbody><tr><td>1</td><td>데이터시각화 할 때 고려사항으로 맞지 <b>않는</b> 것을 고르시오.  1. 대량의 데이터를 시각화하기 위해서는 많은 처리과정과 계산이 필요하다. 2. 데이터시각화를 위해서는 그 범위와 대상을 최대한 좁혀서 구체적으로 만들어야 한다. 3. 차트는 보여주는 것이 목적이어야 한다. 4. 어떤 활동 때문에 차트 선이 변동되는 지 보여주어야 한다.</td><td>3</td><td>차트를 보여 주는 것이 목적이 되면 안 되고 어떤 활동 때문에 차트의 선이 오르내리는지를 보여줘야 한다.</td></tr><tr><td>2</td><td>색상으로 표현할 수 있는 다양한 정보를 일정한 이미지위에 열 분포 형태로 시각화한 것이 특징인 차트를 무엇이라 하는가?</td><td>히트맵</td><td>히트맵의 설명이다.</td></tr><tr><td>3</td><td>시각화 및 게시 프로세스에 해당되지 <b>않는</b> 것은?  1. 데이터 가져오기 2. 필터로 데이터 정제 3. 보고서 만들기 4. 게시 및 배포</td><td>2</td><td>파워BI의 슬라이서와 필터는 보고서 작성시 사용한다.</td></tr></tbody></table>				번호	문제	정답	해설	1	데이터시각화 할 때 고려사항으로 맞지 <b>않는</b> 것을 고르시오.  1. 대량의 데이터를 시각화하기 위해서는 많은 처리과정과 계산이 필요하다. 2. 데이터시각화를 위해서는 그 범위와 대상을 최대한 좁혀서 구체적으로 만들어야 한다. 3. 차트는 보여주는 것이 목적이어야 한다. 4. 어떤 활동 때문에 차트 선이 변동되는 지 보여주어야 한다.	3	차트를 보여 주는 것이 목적이 되면 안 되고 어떤 활동 때문에 차트의 선이 오르내리는지를 보여줘야 한다.	2	색상으로 표현할 수 있는 다양한 정보를 일정한 이미지위에 열 분포 형태로 시각화한 것이 특징인 차트를 무엇이라 하는가?	히트맵	히트맵의 설명이다.	3	시각화 및 게시 프로세스에 해당되지 <b>않는</b> 것은?  1. 데이터 가져오기 2. 필터로 데이터 정제 3. 보고서 만들기 4. 게시 및 배포	2	파워BI의 슬라이서와 필터는 보고서 작성시 사용한다.	<div>[학습평가 페이지]</div> <ul style="list-style-type: none"><li>페이지 퀴즈 컴포넌트 사용하여 페이지 개발</li></ul>
번호	문제	정답	해설																			
1	데이터시각화 할 때 고려사항으로 맞지 <b>않는</b> 것을 고르시오.  1. 대량의 데이터를 시각화하기 위해서는 많은 처리과정과 계산이 필요하다. 2. 데이터시각화를 위해서는 그 범위와 대상을 최대한 좁혀서 구체적으로 만들어야 한다. 3. 차트는 보여주는 것이 목적이어야 한다. 4. 어떤 활동 때문에 차트 선이 변동되는 지 보여주어야 한다.	3	차트를 보여 주는 것이 목적이 되면 안 되고 어떤 활동 때문에 차트의 선이 오르내리는지를 보여줘야 한다.																			
2	색상으로 표현할 수 있는 다양한 정보를 일정한 이미지위에 열 분포 형태로 시각화한 것이 특징인 차트를 무엇이라 하는가?	히트맵	히트맵의 설명이다.																			
3	시각화 및 게시 프로세스에 해당되지 <b>않는</b> 것은?  1. 데이터 가져오기 2. 필터로 데이터 정제 3. 보고서 만들기 4. 게시 및 배포	2	파워BI의 슬라이서와 필터는 보고서 작성시 사용한다.																			