



AI



AR/VR



Big Data



Cloud



ICT 융합기술



ICT 인프라기술



Media



Security

# Visualization & BI

## Case 로 실습하는 데이터 분석▪시각화

## Case 로 실습하는 데이터 분석▪시각화 (Visualization & BI) :

### 목적 :

- 데이터 분석가(Data Analyst)를 위한  
Data Literacy (데이터를 읽고 해석해서 현업에 활용할 수 있는 능력) 배양
- 오픈 데이터 리소스를 구하는 방법부터 시각화를 통한 insight 도출을  
10가지 실습으로 데이터분석 및 현업활용을 빠르고 쉽게 할 수 있도록 한다
- 대부분의 국내 회사들이 사무용 컴퓨터 오피스 툴로 쓰는 Microsoft Office 프로그램인 엑셀과  
파워BI로 데이터분석 및 시각화를 하면서 코딩 없이 빠르고 쉽게 비즈니스 인사이트를 찾는 방법을  
익혀 데이터 분석 실무에 즉시 활용해보도록 한다
- 도메인날리지 (경력) 이 있는 비전공자가 BA (data driven Business Analyst) 가 되기 위한 BI 능력배양



## Case 로 실습하는 데이터 분석▪시각화 (Visualization & BI) :

### [ 효 과 ]

- 복잡한 데이터도 한 눈에 읽기 쉽게 시각화 하는 능력 배양
- 실질적인 성과를 가져다 주는 데이터 분석 과정
- 데이터를 통한 가장 합리적인 비즈니스 의사결정 능력 배양
- 다양한 탐색(EDA)을 통한 데이터 시각화 및 인사이트 파악 능력  
(판매 금액 데이터와 같은 수치형 자료, 범주형 자료의 활용과 시계열 분석까지 포함)
- 자체 머신러닝 분석 기능을 내장한 BI툴을 활용 한 시간 효율적인 데이터분석
- 파이썬 코딩 없이 크롤링 하는 무료 방법 3가지로 현업/논문 데이터 수집에 즉시활용
- R 코딩 없이 R 시계열적 예측 (Time series)

## 시각화 개체 사용 툴(Tool) :

- Microsoft Power BI 데스크 탑 사전설치 필요

(설치 방법 사전 공지)

- Microsoft os 에서 구현, excel 필요.

(애플노트북에서 ios 만 있을 시

파워bi 웹 으로 참여 가능하나 파워bi 데스크탑 (다운로드 한 버전) 과 다를 수 있음

# 데이터 과학자가 커뮤니케이션 하기 가장 좋은 방법 - 시각화:

시각화는 커뮤니케이션(설명/설득) 하기 좋고

인사이트를 줄 수 있는 중요한 수단

시각화는 데이터로부터 유용한 정보와 인사이트를 얻어내기 위한 툴로서  
데이터에 숨겨진 패턴을 파악하거나 자료 사이에 숨겨진 관계를 찾아내기도 하고  
지표를 시각적으로 표현함으로써 빠르게 의사 결정을 내릴 수 있도록 하는 장점 있음

데이터를 분석하는데 있어 평가 지표를 만들어야

KPI(Key Performance Indicator)가 자주 활용되고 있으며

다른 변수들과의 관계를 통해 목적 달성을 위한 다양한 방법을 찾아낼 수 있기도 하다.





AI



AR/VR



Big Data



Cloud



ICT 융합기술



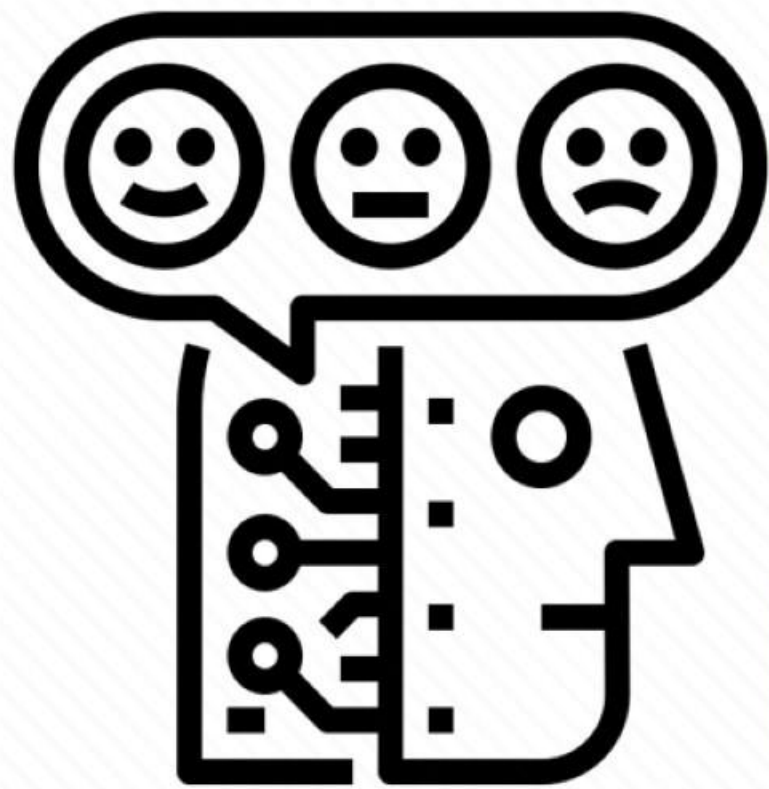
ICT 문화기술



Media

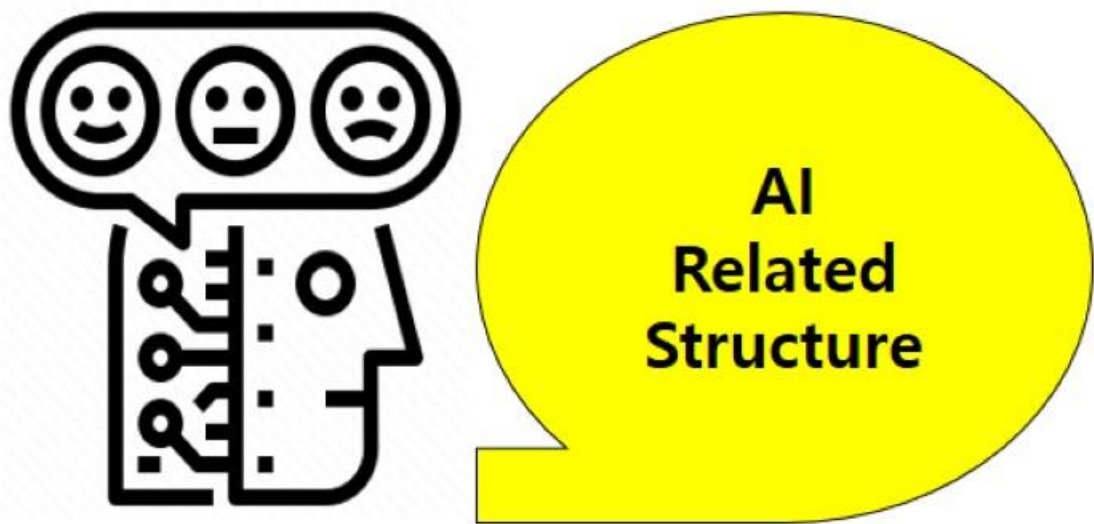


Security



데이터과학  
(Data science) 은  
현실의 문제를 해결하고  
이해하는 데 데이터를  
사용하는 과정

# 조직으로 본 데이터 사이언스 관련 직무

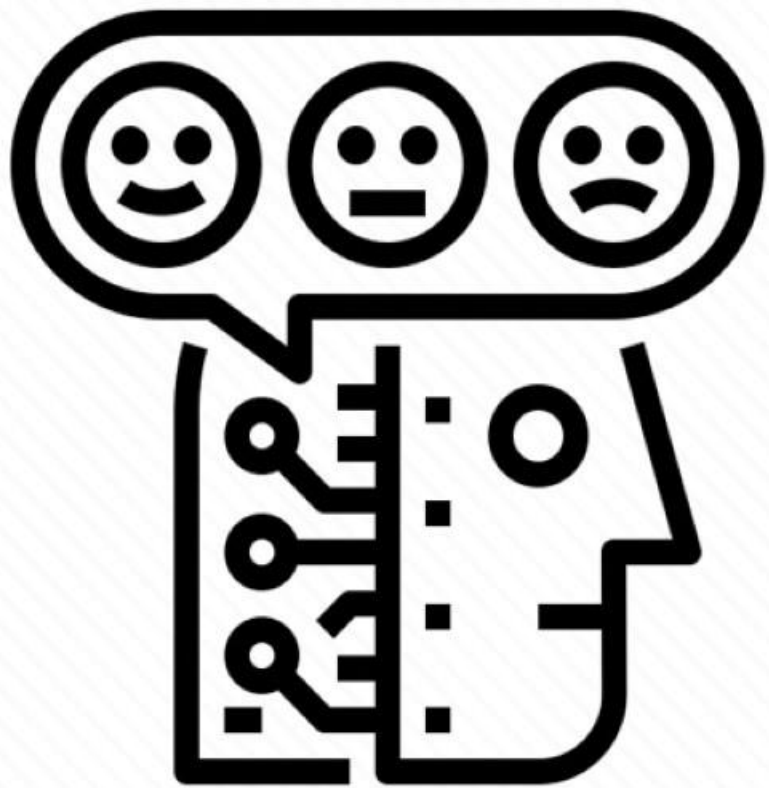


## 조직 구성

- Data Intelligence Lab
- Vision Lab
- Language Lab
- Business Analysis Division
- Fundamental Research Lab



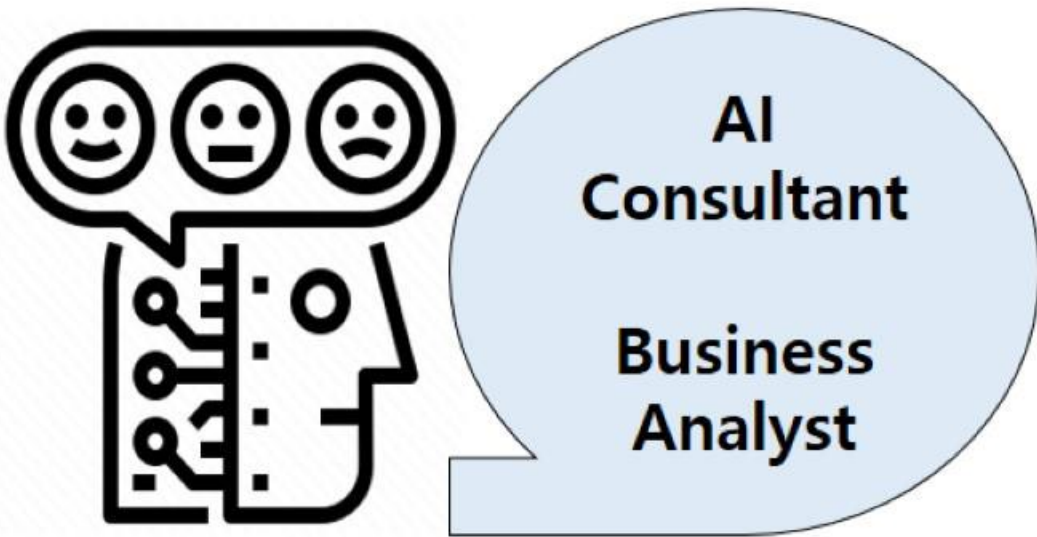
잠깐 !!



Data Engineering,  
머신러닝 엔지니어,  
Data Analytics,  
Data infra 를 하는  
엔지니어 직군 만이  
데이터 관련 직군은 아니다



## ■ 직무 관련 : AI Business Dev.Div



### 목표 :

Biz Domain 지식 및 최신 AI 기술에 대한 이해를 바탕으로 여러 조직들과 함께 난제를 발굴 및 문제 정의를 통해 검증하고, 더 나아가 AI 기술 기반의 새로운 사업 기회 탐색 및 사업 모델 개발을 통해 사업화를 추진하는 것을 목표로 함

## 직무

- 사내 AI 난제 발굴, 문제 정의 및 Consulting
  - 1) AI 니즈 발굴 및 문제 정의 리딩
  - 2) 개발 조직과 AI 구현 협업
  - 3) 도메인 Data 분석 및 AI Training 위한 데이터 발굴 지원
- 초거대 AI를 활용하여 신규 사업 기회 발굴 (도메인 영역 : 리테일, 에듀, 금융, 헬스, 패션 외)
  - 1) 고객의 초거대 AI 니즈 발굴 및 문제 정의
  - 2) 서비스 컨셉, 신규 사업 기회 발굴 및 Proof of Concept 진행
  - 3) 사업계획 수립 및 Research Lab과 구현 협업
  - 4) 외부 Partnership
  - 5) 고객 Data 분석 및 초거대 AI Training 위한 데이터 가이드

## 자격요건

[Business Data Analyst]

- 리테일/에듀/금융/헬스/패션 사업 영역에서 고객, 서비스, 엔터프라이즈 데이터 분석 및 인사이트 발굴 프로젝트 리딩 (7년 이상)
- Deep Learning/Machine Learning 등 AI 관련 지식 보유(관련 학위자 우대)
- SQL 등 활용하여 데이터수집/가공/정제/저장 경험
- Python, R 등을 활용한 데이터 프로세싱 경험
- Multi-pier 프로젝트 협업 경험
- Tableau, PowerBI 등 데이터 시각화 경험

[AI Consultant]

- 유통/교육/금융/헬스/패션 사업 영역에 대한 디지털 컨설팅 경험
- 현업과 문제 발굴, 정의, 해결을 위한 일련의 과정을 리딩하고 Multi-pier와 원활한 커뮤니케이션 및 Management Reporting

## 전공

- 전공 무관

실무에서의 데이터 분석은 사실 어렵지 않습니다

중요한 것은

요약 데이터의 해석!



## AI Consultant, Business Analyst

로써 필요한 능력은

① Domain Knowledge

② 분석력

③ 설명력

④ 설득력



비즈니스 필드 에서

EDA 와 데이터 시각화 (Visualizaiton) 는 ?

# 시각화가 중요한 이유

- ① 많은 양의 데이터를 한눈에 볼 수 있음
- ② 도메인 지식이 부족해도, 쉽게 시각화로 인사이트를 찾을 수 있음
- ③ 보다 정확하게 데이터를 이해하는 데 필요
- ④ 데이터 인사이트를 공유하는 데 효과적

(특히 비즈니스에서는 의사결정권자를 설득하는 일과 협업자 들의 동의를 끌어내는 데 필수)

- ⑤ 데이터가 존재하는 다양한 분야에서 활용 가능



# 데이터 시각화 (BI) Tool 의 동향

## ( BI = Business Intelligence )



# 데이터 시각화 Tool 소개

Flourish

동적 차트

<https://flourish.studio/>

Flourish는 Canva 제품군의 일부

Python  
(seaborn, Matplotlib 등)

R

Google 데이터 스튜디오

실시간 차트, 가성비 갑

태블로

시각적으로 우수(컬러), 유료 (무료버전은 1년)

Microsoft 파워BI

실시간 차트, 가성비 갑

엑셀/스프레드시트/파워포인트

기본표, 기본차트

# 데이터 시각화 (BI) Tool 소개

- Tableau



- 시각화 전문 솔루션

- 데이터 연결의 유연성

- 유료 라이선스

- 2019년 salesforce.com에 의해 인수 되면서 (157억 달러, 약 18조 5,500억 원)

- 세일즈 포스 모든 고객 데이터를 태블로로 통합된 뷰 제공"

# TABLEU : 교육 동영상

무료)

[https://help.tableau.com/current/pro/desktop/en-us/dataview\\_examples.htm](https://help.tableau.com/current/pro/desktop/en-us/dataview_examples.htm)

<https://www.tableau.com/ko-kr/learn/training/20212>

[https://public.tableau.com/s/resources?qt-overview\\_resources=1#qt-overview\\_resources](https://public.tableau.com/s/resources?qt-overview_resources=1#qt-overview_resources)

<https://greatified.com/2019/03/20/how-to-build-an-animated-bar-chart-race-in-tableau-software/>

<https://www.youtube.com/watch?v=PdwOWILPa9o>

[https://help.tableau.com/current/pro/desktop/ko-kr/forecast\\_how\\_it\\_works.htm](https://help.tableau.com/current/pro/desktop/ko-kr/forecast_how_it_works.htm)

<https://www.youtube.com/watch?v=um-puo3jeVU>

<https://www.youtube.com/watch?v=RXda5d2QGA8>



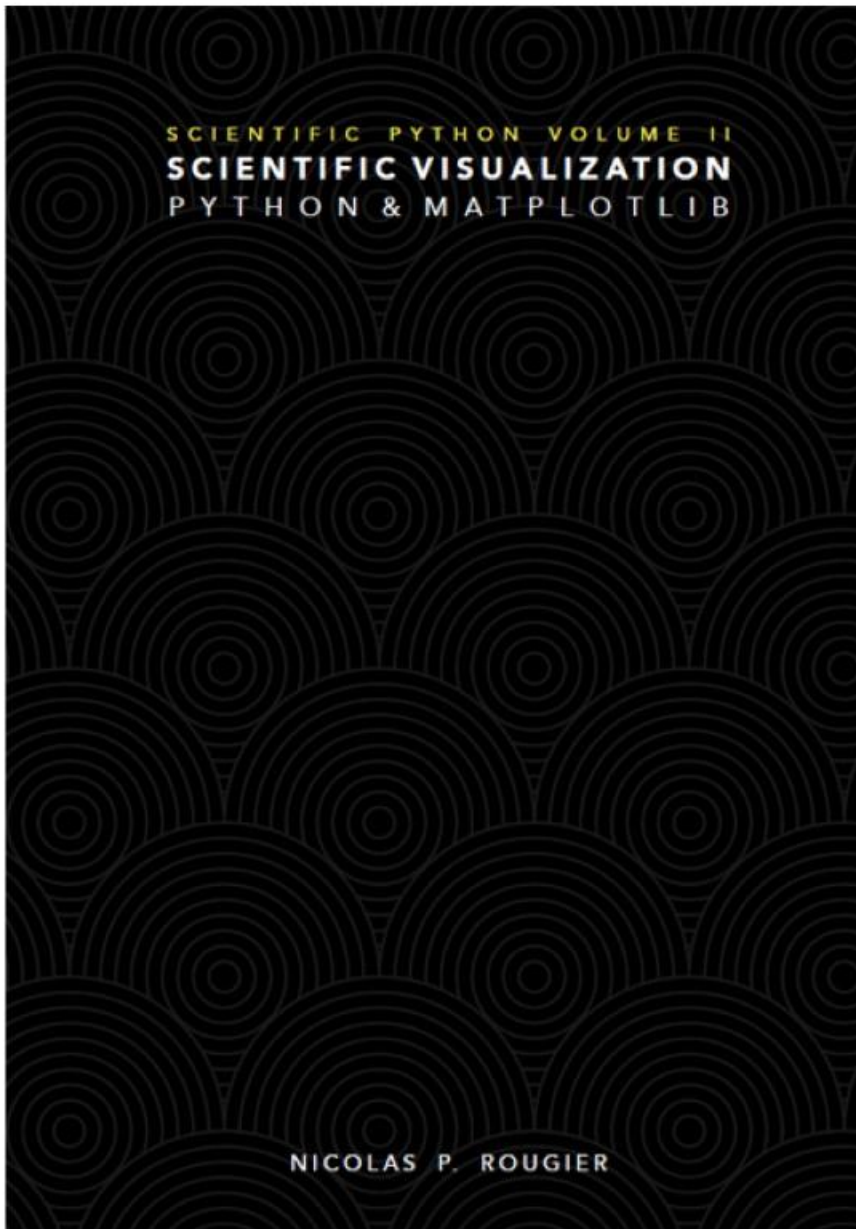
# 데이터 시각화 (BI) Tool 소개



# Python

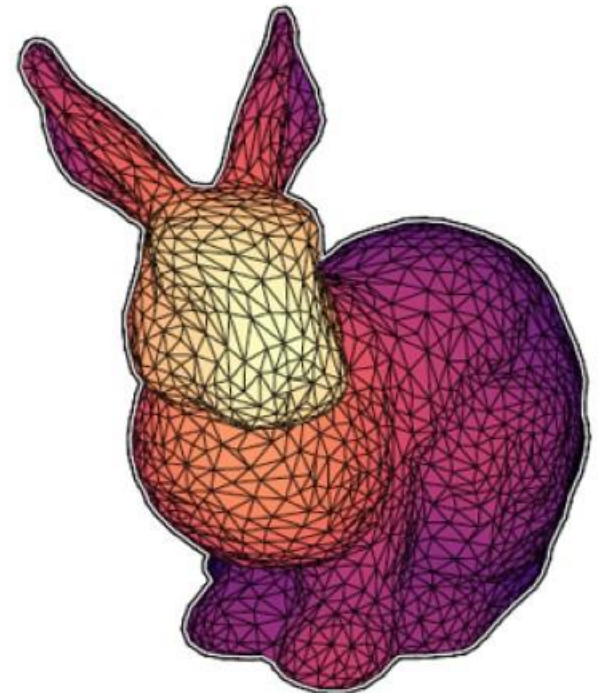


# 파이썬 시각화 에 대해 더 공부하고 싶다면 Scientific Visualization 도서 참고



Preface	ix
Acknowledgments	xiii
Introduction	xvii
I Fundamentals	
1 Anatomy of a figure	3
2 Coordinate systems	19
3 Scales & projections	33
4 Elements of typography	45
5 A primer on colors	55
II Figure design	
6 Ten simple rules	71
7 Mastering the defaults	85
8 Size, aspect & layout	95
9 Ornaments	109
III Advanced concepts	
10 Animation	121
11 Going 3D	135
12 Architecture & optimization	149
13 Graphic library	159
IV Showcase	
Filled contours with dropshadows	177
Domain coloring	179
Escher like projections	181
Self-organizing maps	183
Waterfall plots	185
Streamlines	187
Mandelbrot set	189
Recursive Voronoi	191
3D Heightmap	193
Voronoi mosaic	195
Text shadow	197
Text spiral	199
V Conclusion	
VI Appendix	
A External resources	207
B Quick References	211
Bibliography	221

```
V, F = [], []
with open("bunny.obj") as f:
    for line in f.readlines():
        if line.startswith('#'): continue
        values = line.split()
        if not values: continue
        if values[0] == 'v':
            V.append([float(x) for x in values[1:4]])
        elif values[0] == 'f':
            F.append([int(x) for x in values[1:4]])
V, F = np.array(V), np.array(F)-1
```



# 데이터 시각화 (BI) Tool 소개



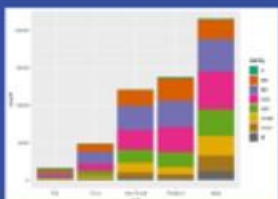
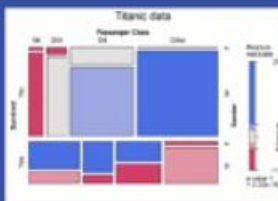
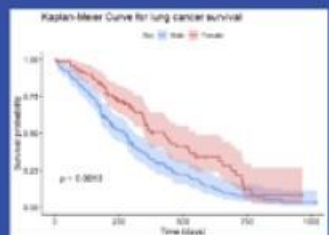
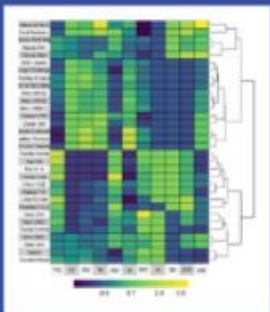
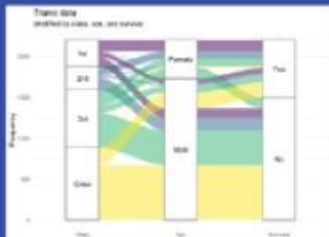
R





# R 시각화 에 대해 더 공부하고 싶다면 Data Visualization with R 도서 참고

## Data Visualization with R



Rob Kabacoff, Ph.D

Quantitative Analysis Center  
Wesleyan University

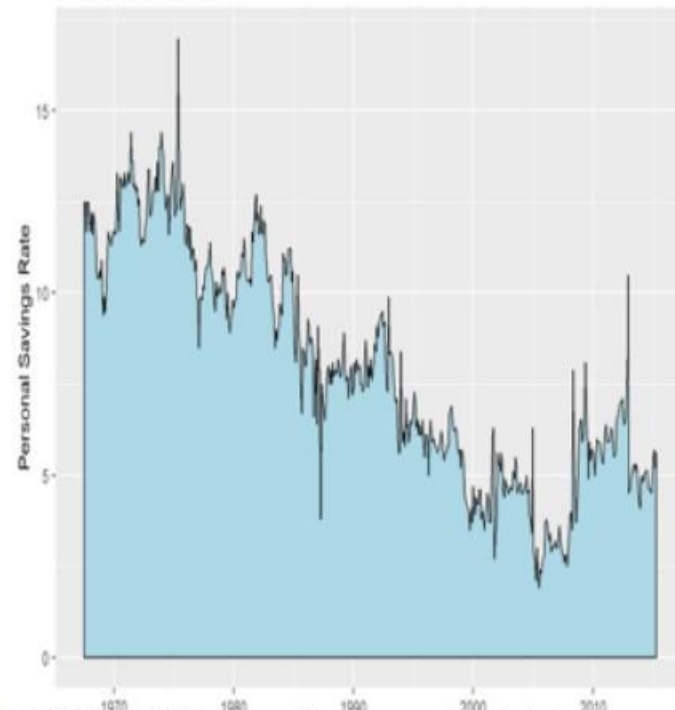
Chapter	Description
Ch 1	provides a quick overview of how to get your data into R and how to prepare it for analysis.
Ch 2	provides an overview of the <code>ggplot2</code> package.
Ch 3	describes graphs for visualizing the distribution of a single categorical (e.g. race) or quantitative (e.g. income) variable.
Ch 4	describes graphs that display the relationship between two variables.
Ch 5	describes graphs that display the relationships among 3 or more variables. It is helpful to read chapters 3 and 4 before this chapter.
Ch 6	provides a brief introduction to displaying data geographically.
Ch 7	describes graphs that display change over time.
Ch 8	describes graphs that can help you interpret the results of statistical models.
Ch 9	covers graphs that do not fit neatly elsewhere (every book needs a miscellaneous chapter).
Ch 10	describes how to customize the look and feel of your graphs. If you are going to share your graphs with others, be sure to skim this chapter.
Ch 11	covers how to save your graphs. Different formats are optimized for different purposes.
Ch 12	provides an introduction to interactive graphics.
Ch 13	gives advice on creating effective graphs and where to go to learn more. It's worth a look.
The Appendices	describe each of the datasets used in this book, and provides a short blurb about the author and the Wesleyan Quantitative Analysis Center.

There is no one right graph for displaying data. Check out the examples, and see which type best fits your needs.

A simple area chart is basically a line graph, with a fill from the line to the x-axis.

```
# basic area chart
ggplot(economics, aes(x = date, y = psavert)) +
  geom_area(fill="lightblue", color="black") +
  labs(title = "Personal Savings Rate",
        x = "Date",
        y = "Personal Savings Rate")
```

Personal Savings Rate



# 데이터 시각화 (BI) Tool 소개



## Flourish

# 데이터 시각화 (BI) Tool 소개



## Data Studio

Google Analytics Marketing Website



# 데이터 시각화 (BI) Tool 소개



## Microsoft Power BI



# Power BI, **가성비 높은 시각화 및 데이터 분석 솔루션**

1. 웹사이트 접속 <https://powerbi.microsoft.com/ko-kr/>
2. 회원 가입 (반드시 회사, 기관, 학교 이메일로 가입)
3. 가입 후 **Power BI 데스크탑** 다운로드 후 설치 (Windows OS만 지원)



- **Desktop** : 내 노트북과 PC 에서, 무료로 똑딱
- **Pro** : 공유와 공동 작업 가능
- **Premium** : AI 기능 활용 (autoML, 영향력 분석, Azure Cognitive Services 등)  
빅데이터 분석용 성능 10배향상으로 효율성 및 정확성 구현
- **Report Server** : 오늘은 온프레미스 보고,  
내일은 클라우드로 이동하는 유연성有
- **Mobile app** : 모든 디바이스를 위한 모바일앱 (Window,iOS,Android용) 으로 share,  
한번의 터치로 모바일 비즈니스 인텔리전스에 액세스



## Power BI Pro

## 사용자 기준

**₩11,200**

월 사용자 기준

현대적인 방식의 셀프서비스 분석으로 개별 사용자에게 라이선스를 부여하며 라이브 대시보드 및 보고서를 통해 데이터를 시각화하고 조직 전체에서 인사이트를 공유합니다.

- Power BI Pro가 [Microsoft 365 E5](#)에 포함됩니다.
- 신용카드로 바로 구매 가능<sup>1</sup>

## Power BI Premium

## 사용자 기준

**₩22,500**월 사용자 기준<sup>2</sup>

고급 AI로 인사이트에 빠르게 액세스할 수 있도록 개별 사용자에게 라이선스를 부여하고 빅 데이터를 위해 셀프서비스 데이터 준비를 하고, 엔터프라이즈 규모로 데이터 관리 및 액세스를 간소화합니다.

- Power BI Pro의 모든 [기능](#) 제공.
- 신용카드로 바로 구매 가능<sup>1</sup>

## 용량 기준

기본

**₩5,617,600**

월 용량 기준

고급 AI로 인사이트에 빠르게 액세스하기 위한 용량 기준 라이선스를 조직에게 부여하고, 빅 데이터에 대한 셀프서비스 데이터 준비를 하고, 엔터프라이즈 규모로 데이터 관리 및 액세스를 간소화합니다. 콘텐츠 사용 소비자에 대해 사용자 기준 라이선스를 사용하지 않습니다.

- Power BI Premium(용량 기준)에 콘텐츠 게시를 위한 Power BI Pro 라이선스가 필요합니다.
- Azure 구독을 통해 [자동 크기 조정](#)을 설정하고 Power BI Premium 용량을 자동으로 확장합니다.



# Power BI : 가성비 높은 시각화 및 데이터 분석 솔루션

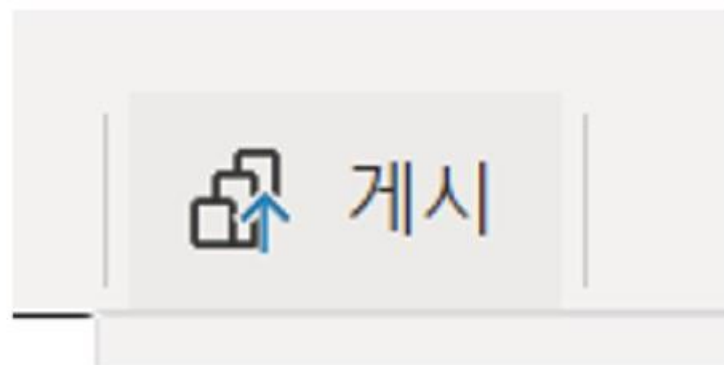
## 주 용도

- 데이터 가져오기 (Multiple data sources)  
엑셀, CSV, SQL, WEB 등 다양한 데이터 소스 불러오기 가능
- 데이터 연결 (Connecting to data)  
불러온 데이터를 각 연결 하고 관계형 DB 로 만들기 가능
- 데이터 정제/모델링 (Data Modeling)  
파워쿼리로 데이터 모델링 및 변환 가능, 수학적 연산 및 통계 가능
- 시각화 (Visualizing)  
220여개 종류의 다양한 시각적 개체로 보고서 만들기 가능
- 작성공유 (Sharing)  
작성보고서 Power BI 서비스로 게시 하기 가능



## Power BI

**“게시”** 를 통해, Web 에서 보고서 를 볼 수 있음



**애플 MAC 컴퓨터** 로는  
파워BI 데스크탑 (로컬 로 다운로드) 을 쓰지는 못하지만 (IOS 불가)

**웹** (파워BI서비스) 에서 는 **사용 가능**

(단, 데스크탑 (로컬)에 있는 데이터를 웹에 업로드 후에 가능하나 제약 많음)



# E.O.D

