



즐거운  
프로그래밍  
경험 ★

# 모두의

# 데이터 분석

with 파이썬

실생활 예제로 시작하는  
데이터 분석 첫걸음

송석리, 이현아 지음

- 파이썬 기초 문법만 알아도 OK!
- 공공·오픈데이터를 활용한 다양한 프로젝트 수록!
- 실생활 데이터 분석을 통해 나만의 인사이트 발견!



## 모두의 데이터 분석 with 파이썬

Data Analysis for Everyone

초판 발행 · 2019년 4월 25일

지은이 · 송석리, 이현아

발행인 · 이종원

발행처 · (주)도서출판 길벗

출판사 등록일 · 1990년 12월 24일

주소 · 서울시 마포구 월드컵로 10길 56(서교동)

대표전화 · 02)332-0931 | 팩스 · 02)323-0586

홈페이지 · [www.gilbut.co.kr](http://www.gilbut.co.kr) | 이메일 · [gilbut@gilbut.co.kr](mailto:gilbut@gilbut.co.kr)

기획 및 책임편집 · 김윤지([yunjikim@gilbut.co.kr](mailto:yunjikim@gilbut.co.kr)) | 디자인 · 배진웅 | 제작 · 이준호, 손일순, 이진혁

영업마케팅 · 임태호, 전선하, 지운집, 박성용 | 영업관리 · 김명자 | 독자지원 · 송혜란, 홍혜진

교정교열 · 황진주 | 전산편집 · 도설아 | 출력 및 인쇄 · 예림인쇄 | 제본 · 신정문화사

- 잘못된 책은 구입한 서점에서 바꿔 드립니다.
- 이 책에 실린 모든 내용, 디자인, 이미지, 편집 구성의 저작권은 (주)도서출판 길벗과 지은이에게 있습니다.  
허락 없이 복제하거나 다른 매체에 옮겨 실을 수 없습니다.

ISBN 979-11-6050-781-2 93000

(길벗 도서번호 007029)

정가 18,000원

독자의 1초를 아껴주는 정성 길벗출판사

(주)도서출판 길벗 [www.gilbut.co.kr](http://www.gilbut.co.kr)

페이스북 · [www.facebook.com/gbitbook](https://www.facebook.com/gbitbook)



베타테스터의 한마디 · 3 | 머리말 · 4 | 책 소개 · 5 | 이 책의 구성과 활용법 · 6

첫째  
마당

## 기온 공공데이터

13

<b>UNIT 01</b> 기온 데이터 분석 시작하기	<b>14</b>
1   기온 공공데이터 살펴보기.....	14
2   CSV 파일이란.....	18
3   데이터 분석에 필요한 환경 만들기.....	21
<b>UNIT 02</b> 서울의 기온 데이터 분석하기	<b>28</b>
1   CSV 파일에서 데이터 읽어오기.....	28
2   데이터 출력하기.....	36
3   헤더 저장하기.....	39
4   기온 공공데이터에 질문하기.....	41
<b>UNIT 03</b> 서울이 가장 더웠던 날은 언제였을까	<b>43</b>
1   질문 다듬기.....	43
2   문제 해결 방법 구상하기.....	44
3   파이썬 코드로 구현하기.....	45

<b>UNIT 04</b> 기본 그래프 그리기	<b>54</b>
1   'matplotlib 라이브러리'란?	54
2   기본 그래프 그리기	55
3   그래프에 옵션 추가하기	58
<b>UNIT 05</b> 내 생일의 기온 변화를 그래프로 그리기	<b>64</b>
1   데이터에 질문하기	64
2   데이터 시각화하기	67
3   날짜 데이터 추출하기	69
<b>UNIT 06</b> 기온 데이터를 다양하게 시각화하기	<b>77</b>
1   데이터에 질문하기	77
2   히스토그램	78
3   기온 데이터를 히스토그램으로 표현하기	82
4   기온 데이터를 상자 그림으로 표현하기	86

<b>UNIT 07</b> 우리 동네 인구 구조 시각화하기	<b>96</b>
1   인구 공공데이터 내려받기	96
2   인구 데이터 살펴보고 질문하기	100
3   우리 동네 인구 구조 시각화하기	101

<b>UNIT 08</b> 인구 구조를 다양한 형태로 시각화하기	<b>109</b>
1   막대그래프 그리기.....	111
2   향아리 모양 그래프 그리기.....	114
<b>UNIT 09</b> 우리 동네 인구 구조를 파이 차트로 나타내기	<b>124</b>
1   제주도에 여성 비율이 더 높을까.....	125
2   혈액형 비율 표현하기.....	129
3   제주도의 성별 인구 비율 표현하기.....	136
<b>UNIT 10</b> 우리 동네 인구 구조를 산점도로 나타내기	<b>140</b>
1   꺾은선 그래프로 표현하기.....	141
2   막대그래프로 표현하기.....	143
3   산점도로 표현하기.....	144
4   scatter() 함수로 표현하기.....	146
5   버블 차트로 표현하기.....	147
6   제주도의 연령대별 성별 비율을 산점도로 표현하기.....	153

넷째  
마당

## 대중교통 데이터 프로젝트

157

<b>UNIT 11</b> 대중교통 데이터 시각화하기	<b>158</b>
1   대중교통 데이터 내려받기.....	158
2   지하철 유무임별 이용현황 데이터 정제하기.....	161
3   유임 승차 비율이 가장 높은 역은 어디일까.....	165
4   유무임 승하차 인원이 가장 많은 역은 어디일까.....	172
5   모든 역의 유무임 승하차 비율은 어떻게 될까.....	175
<b>UNIT 12</b> 지하철 시간대별 데이터 시각화하기	<b>180</b>
1   지하철 시간대별 이용 현황 데이터 정제하기.....	180

2   출근 시간대 사람들이 가장 많이 타고 내리는 역은 어디일까.....	185
3   밤 11시에 사람들이 가장 많이 타는 역은 어디일까.....	190
4   시간대별로 사람들이 가장 많이 타고 내리는 역은 어디일까.....	193
5   모든 지하철역에서 시간대별 승하차 인원을 모두 더하면.....	198

**다섯째  
마당**

**파이썬 데이터 분석 라이브러리를  
활용한 프로젝트**

203

<b>UNIT 13</b> 숫자 데이터를 쉽게 다루게 돕는 numpy 라이브러리	<b>204</b>
1   matplotlib 홈페이지.....	204
2   numpy 라이브러리 시작하기.....	207
3   numpy 라이브러리를 활용해 그래프 그리기.....	211
4   numpy array 생성하기.....	215
5   numpy array의 다양한 활용.....	220
<b>UNIT 14</b> numpy를 활용한 나만의 프로젝트 만들기	<b>227</b>
1   관심 있는 데이터 찾기.....	227
2   데이터 살펴보며 질문하기.....	229
3   질문을 명확한 문제로 정의하기.....	230
4   알고리즘 설계하기.....	231
5   알고리즘을 코드로 표현하기.....	232
<b>UNIT 15</b> 테이블 형태의 데이터를 쉽게 다루도록 도와주는 pandas 라이브러리	<b>251</b>
1   위키피디아 데이터 엑셀로 저장하기.....	252
2   pandas란.....	259
3   데이터 프레임 기초.....	260
4   pandas로 인구 구조 분석하기.....	268

1   출력과 입력 그리고 변수.....	275
2   연산자 사용하기.....	283
3   함수 불러오기.....	288
4   반복과 선택.....	290
5   순서 있는 저장 공간 리스트.....	296

마치는 글 · 301 | 찾아보기 · 302

# 누구나 파이썬으로

## 생활 속 데이터를 분석하고 활용할 수 있다!



시작

### 데이터 분석 실행 환경 만들기

데이터 분석에 필요한 아나콘다를 설치하고 환경 설정을 합니다.

상습

### 공공/오픈데이터를 활용한 실생활 프로젝트 실습하기

기존 공공데이터, 인구 공공데이터, 대중교통 데이터를 활용해서 질문을 던지고 답을 찾습니다. 또 꺾은선 그래프, 막대그래프, 향아리 모양 그래프, 파이 차트 등 다양한 형태로 데이터를 시각화하고 분석하면서 나만의 인사이트를 발견합니다.

활용

### 파이썬 라이브러리 활용하기

데이터 분석을 할 때 가장 많이 사용되는 파이썬 라이브러라인 `numpy`, `pandas`를 활용해 다양한 프로젝트를 진행합니다.



이 책을

먼저 본

네타스터의

원마디



이 책 한 권을 독파하면 데이터 분석의 진입 장벽이 확 낮아질 거예요. **김유안님**

의미 없어 보였던 데이터 영어라에서 숨겨진 가치를 찾아내는 과정이 매우 인상적이었어요. **이현서님**

일상생활 속 데이터에 질문을 던지고 프로젝트로 진행하는 방식이 흥미로웠어요. **서승희님**

데이터 분석을 어떻게 시작할지 고민했던 사람들에게 좋은 길라잡이가 되는 책이에요. **강경희님**

이 책 덕분에 파이썬이라는 간절한 고급 언어로 데이터를 쉽게 분석할 수 있다는 것을 배웠어요. **김현아님**

모드업

모두의 파이썬

모두의 데이터 분석 with 파이썬

모두의 데이터 리팩 with 파이썬

모두의 데이터 분석 with 파이썬  
DATA ANALYSIS FOR EVERYONE

정가 18,000원



9 791160 507812  
ISBN 979-11-6050-781-2