

The background of the slide features a stylized illustration of a large group of runners in various shades of gray and brown, all in motion. In the foreground, a single runner in a dark red silhouette is shown from behind, with arms raised in a 'V' shape, indicating they have just crossed a finish line. A thin, wavy red line arches over the runner's head, representing the finish line banner. The overall scene conveys a sense of achievement and competition.

Page 1. 머신러닝 2일만에 완주하기

Index

Part -1. 오늘 사용할 데이터 소개

Part 0. 나만의 Python 작업공간 만들기 [Python Editor 정복기]

Part 1. 원하는거 다 가져와보자 [Crawling 1시간에 끝내기]

Part 2. 시각화가 그렇게 어려워? 응, 어려워

Data Scientist가 해야 할 일



Project 폴더 구성

 code
 dataset
 ipynb

Part -1. 오늘 사용할 데이터 소개

야구 [Baseball]

야구 [Baseball]

야구장 관중객 수가 미치는 영향은?

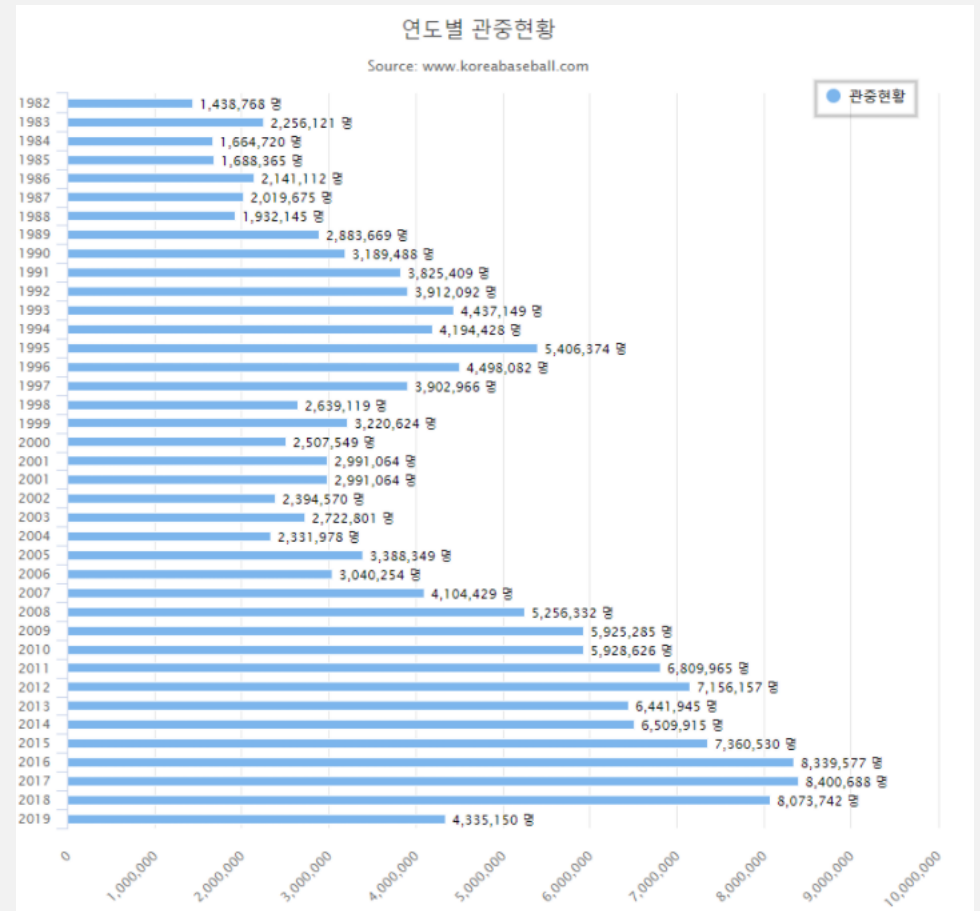
한국프로야구 관중 수는 1980년대 초반 약 140만 명을 시작으로

2018년에는 약 800만명으로, 한국 스포츠의 주요 종목으로 성장

한국프로야구의 입지를 증명했습니다.

이처럼 한국 프로 스포츠 발전의 가장 중요한 관심사 중 하나.

관중 수는 구단의 경영 측면과 직결되는 요소로써, 입장료부터 기타 시설의 이용료 등 다양한 수입을 통해 구단의 안정적인 재정 운영을 가능케합니다. 또한, 수요예측은 한국 프로 스포츠의 마케팅과 구단의 예산 전략 수립에 활용될 수 있습니다.

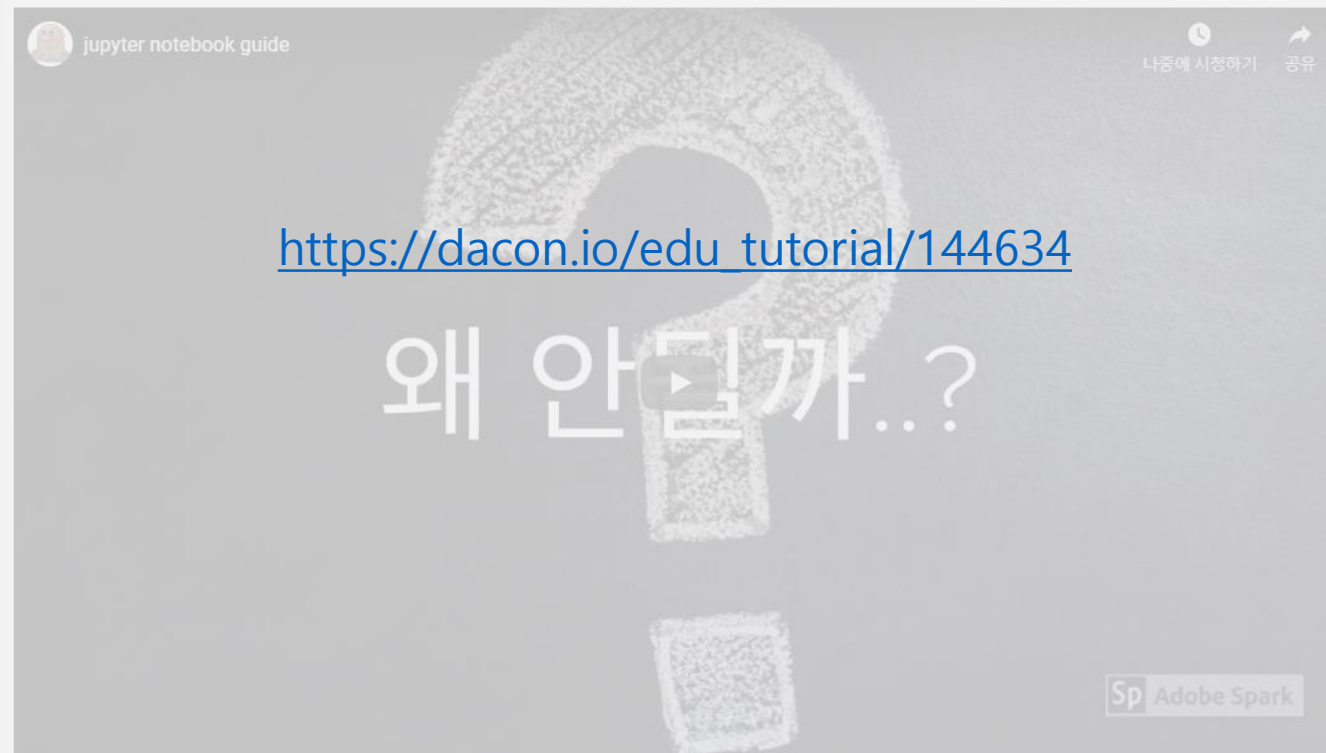


Part 0. 나만의 Python 작업공간 만들기 [Python Editor 정복기]

1. Jupyter Notebook

1. Jupyter Notebook

- 설치



1. Jupyter Notebook

- Nbextension

설치 방법

➤ 명령 프롬프트에서 아래 명령어 실행

```
pip install jupyter_contrib_nbextensions
```

```
jupyter contrib nbextensions install --sys-prefix --skip-running-check
```

1. Jupyter Notebook

- Nbextension

Configurable nbextensions

☒ disable configuration for nbextensions without explicit compatibility (they may break your notebook environment, but can be useful to show for nbextension development)

filter: by description, section, or tags

<input type="checkbox"/> (some) LaTeX environments for Jupyter	<input type="checkbox"/> 2to3 Converter	<input checked="" type="checkbox"/> AddBefore	<input type="checkbox"/> Autopep8
<input type="checkbox"/> AutoSaveTime	<input checked="" type="checkbox"/> Autoscroll	<input type="checkbox"/> Cell Filter	<input type="checkbox"/> Code Font Size
<input type="checkbox"/> Code prettify	<input checked="" type="checkbox"/> Codefolding	<input type="checkbox"/> Codefolding in Editor	<input checked="" type="checkbox"/> CodeMirror mode extensions
<input checked="" type="checkbox"/> Collapsible Headings	<input type="checkbox"/> Comment/Uncomment Hotkey	<input checked="" type="checkbox"/> contrib_nbextensions_help_item	<input type="checkbox"/> datestamper
<input type="checkbox"/> Equation Auto Numbering	<input type="checkbox"/> ExecuteTime	<input type="checkbox"/> Execution Dependencies	<input type="checkbox"/> Exercise
<input type="checkbox"/> Exercise2	<input type="checkbox"/> Export Embedded HTML	<input type="checkbox"/> Freeze	<input type="checkbox"/> Gist-it
<input checked="" type="checkbox"/> Help panel	<input type="checkbox"/> Hide Header	<input type="checkbox"/> Hide input	<input type="checkbox"/> Hide input all
<input type="checkbox"/> Highlight selected word	<input type="checkbox"/> highlighter	<input type="checkbox"/> Hinterland	<input type="checkbox"/> Initialization cells
<input checked="" type="checkbox"/> ipyparallel/main	<input type="checkbox"/> isort formatter	<input checked="" type="checkbox"/> jupyter-js-widgets/extension	<input type="checkbox"/> Keyboard shortcut editor
<input checked="" type="checkbox"/> Launch QTConsole	<input type="checkbox"/> Limit Output	<input type="checkbox"/> Live Markdown Preview	<input type="checkbox"/> Load TeX macros
<input checked="" type="checkbox"/> Move selected cells	<input type="checkbox"/> Navigation-Hotkeys	<input checked="" type="checkbox"/> Nbextensions dashboard tab	<input checked="" type="checkbox"/> Nbextensions edit menu item
<input type="checkbox"/> nbTranslate	<input type="checkbox"/> Notify	<input checked="" type="checkbox"/> plotlywidget/extension	<input type="checkbox"/> Printview
<input type="checkbox"/> Python Markdown	<input checked="" type="checkbox"/> qgrid/extension	<input type="checkbox"/> Rubberband	<input type="checkbox"/> Ruler
<input type="checkbox"/> Runtools	<input type="checkbox"/> Scratchpad	<input type="checkbox"/> ScrollDown	<input type="checkbox"/> Select CodeMirror Keymap
<input type="checkbox"/> SKILL Syntax	<input checked="" type="checkbox"/> Skip-Traceback	<input checked="" type="checkbox"/> Snippets	<input checked="" type="checkbox"/> Snippets Menu
<input type="checkbox"/> spellchecker	<input checked="" type="checkbox"/> Split Cells Notebook	<input checked="" type="checkbox"/> Table of Contents (2)	<input checked="" type="checkbox"/> table_beautifier
<input type="checkbox"/> Toggle all line numbers	<input type="checkbox"/> Tree Filter	<input checked="" type="checkbox"/> Variable Inspector	<input type="checkbox"/> zenmode

1. Jupyter Notebook

- Nbextension

알아 두면 좋은 기능

Configurable nbextensions

☒ disable configuration for nbextensions without explicit compatibility (they may break your notebook environment, but can be useful to show for nbextension development)

filter: by description, section, or tags

<input type="checkbox"/> (some) LaTeX environments for Jupyter	<input type="checkbox"/> 2to3 Converter	<input checked="" type="checkbox"/> AddBefore	<input type="checkbox"/> Autopep8
<input type="checkbox"/> AutoSaveTime	<input checked="" type="checkbox"/> Autoscroll	<input type="checkbox"/> Cell Filter	<input type="checkbox"/> Code Font Size
<input type="checkbox"/> Code prettify	<input checked="" type="checkbox"/> Codefolding	<input type="checkbox"/> Codefolding in Editor	<input checked="" type="checkbox"/> CodeMirror mode extensions
<input checked="" type="checkbox"/> Collapsible Headings	<input type="checkbox"/> Comment/Uncomment Hotkey	<input checked="" type="checkbox"/> contrib_nbextensions_help_item	<input type="checkbox"/> datestamper
<input type="checkbox"/> Equation Auto Numbering	<input type="checkbox"/> ExecuteTime	<input type="checkbox"/> Execution Dependencies	<input type="checkbox"/> Exercise
<input type="checkbox"/> Exercise2	<input type="checkbox"/> Export Embedded HTML	<input type="checkbox"/> Freeze	<input type="checkbox"/> Gist-it
<input checked="" type="checkbox"/> Help panel	<input type="checkbox"/> Hide Header	<input type="checkbox"/> Hide input	<input type="checkbox"/> Hide input all
<input type="checkbox"/> Highlight selected word	<input type="checkbox"/> highlighter	<input type="checkbox"/> Hinterland	<input type="checkbox"/> Initialization cells
<input checked="" type="checkbox"/> ipyparallel/main	<input type="checkbox"/> isort formatter	<input checked="" type="checkbox"/> jupyter-js-widgets/extension	<input type="checkbox"/> Keyboard shortcut editor
<input checked="" type="checkbox"/> Launch QTConsole	<input type="checkbox"/> Limit Output	<input type="checkbox"/> Live Markdown Preview	<input type="checkbox"/> Load TeX macros
<input checked="" type="checkbox"/> Move selected cells	<input type="checkbox"/> Navigation-Hotkeys	<input checked="" type="checkbox"/> Nbextensions dashboard tab	<input checked="" type="checkbox"/> Nbextensions edit menu item
<input type="checkbox"/> nbTranslate	<input type="checkbox"/> Notify	<input checked="" type="checkbox"/> plotlywidget/extension	<input type="checkbox"/> Printview
<input type="checkbox"/> Python Markdown	<input checked="" type="checkbox"/> qgrid/extension	<input type="checkbox"/> Rubberband	<input type="checkbox"/> Ruler
<input type="checkbox"/> Runtools	<input type="checkbox"/> Scratchpad	<input type="checkbox"/> Scroll to window	<input type="checkbox"/> Select CodeMirror Keymap
<input type="checkbox"/> SKILL Syntax	<input checked="" type="checkbox"/> Skip-Traceback	<input checked="" type="checkbox"/> Snippets	<input checked="" type="checkbox"/> Snippets Menu
<input type="checkbox"/> spellchecker	<input checked="" type="checkbox"/> Split Cells Notebook	<input checked="" type="checkbox"/> Table of Contents (2)	<input checked="" type="checkbox"/> table_beautifier
<input type="checkbox"/> Toggle all line numbers	<input type="checkbox"/> Tree Filter	<input checked="" type="checkbox"/> Variable Inspector	<input type="checkbox"/> zenmode

1. Jupyter Notebook

- 테마 바꾸기

설치 방법

➤ 명령 프롬프트에서 아래 명령어 실행

```
pip install jupyterthemes
```

```
jt -l
```

```
jt -t chesterish
```

1. Jupyter Notebook

- Markdown

Markdown 왜 배워야 할까?

1. 설명력은 높이기 위해서
2. 설명력을 높이기 위해서
3. 설명력을 높이기 위해서

1. Jupyter Notebook

- Markdown

Markdown 사용 실습

Part 0. 나만의 Python 작업공간 만들기 [Python Editor 정복기]

2. Pycharm 정복기

2. Pycharm

- 사용하는 이유

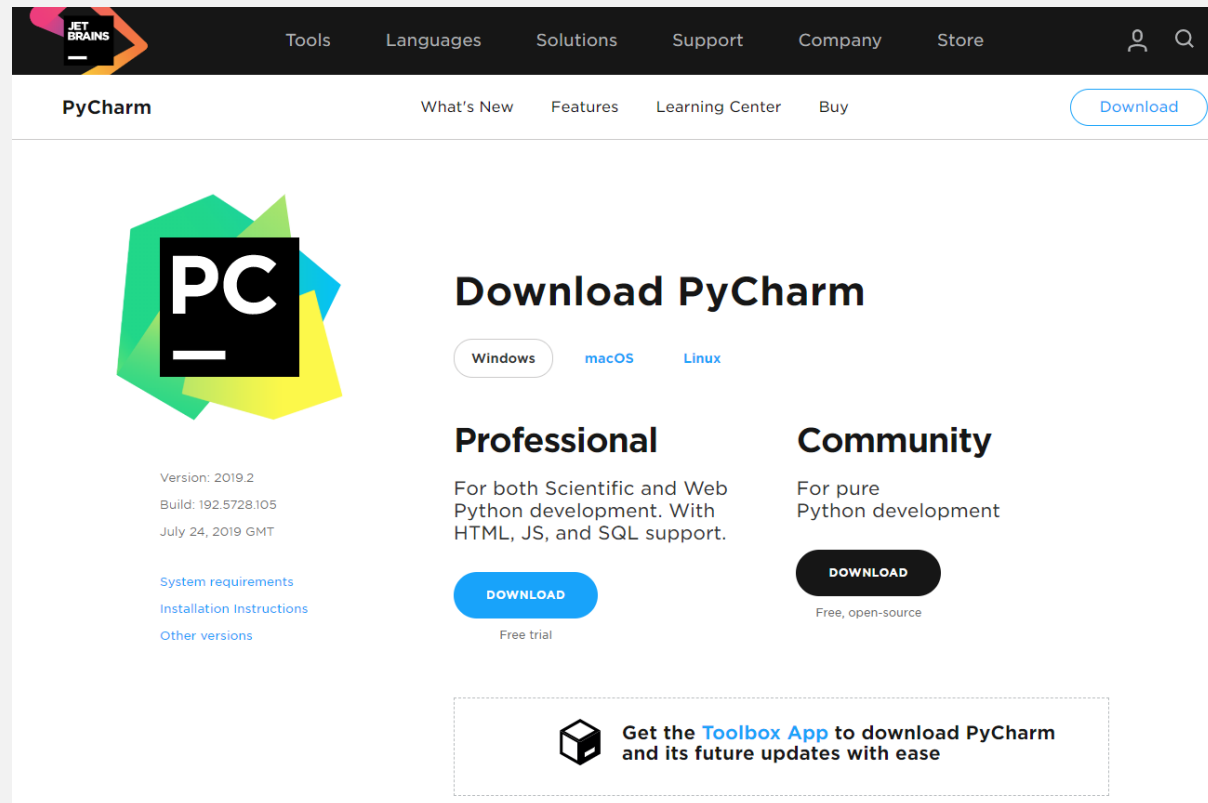
바로 결과 확인 가능

그러나 자동화는 어려움

2. Pycharm


- 설치

구글에 pycharm download 검색



The screenshot shows the official PyCharm download page. At the top is a dark navigation bar with the JetBrains logo and links for Tools, Languages, Solutions, Support, Company, and Store. Below this is a white sub-header with 'PyCharm', 'What's New', 'Features', 'Learning Center', 'Buy', and a 'Download' button. The main content area features the PyCharm logo (a colorful hexagon with 'PC' in a black square) on the left. To its right, the heading 'Download PyCharm' is followed by tabs for 'Windows', 'macOS', and 'Linux'. Below these are two columns: 'Professional' (described as 'For both Scientific and Web Python development. With HTML, JS, and SQL support.') and 'Community' (described as 'For pure Python development'). Each column has a 'DOWNLOAD' button. Under the Professional button is the text 'Free trial'. Under the Community button is 'Free, open-source'. On the left side, below the logo, are links for 'System requirements', 'Installation Instructions', and 'Other versions', along with version and build information. At the bottom, a dashed box contains a message about the 'Toolbox App'.

PyCharm What's New Features Learning Center Buy [Download](#)



Version: 2019.2
Build: 192.5728.105
July 24, 2019 GMT

[System requirements](#)
[Installation Instructions](#)
[Other versions](#)

Download PyCharm

[Windows](#) [macOS](#) [Linux](#)

Professional

For both Scientific and Web Python development. With HTML, JS, and SQL support.

[DOWNLOAD](#)


Free trial

Community

For pure Python development

[DOWNLOAD](#)

Free, open-source

 Get the [Toolbox App](#) to download PyCharm and its future updates with ease

2. Pycharm

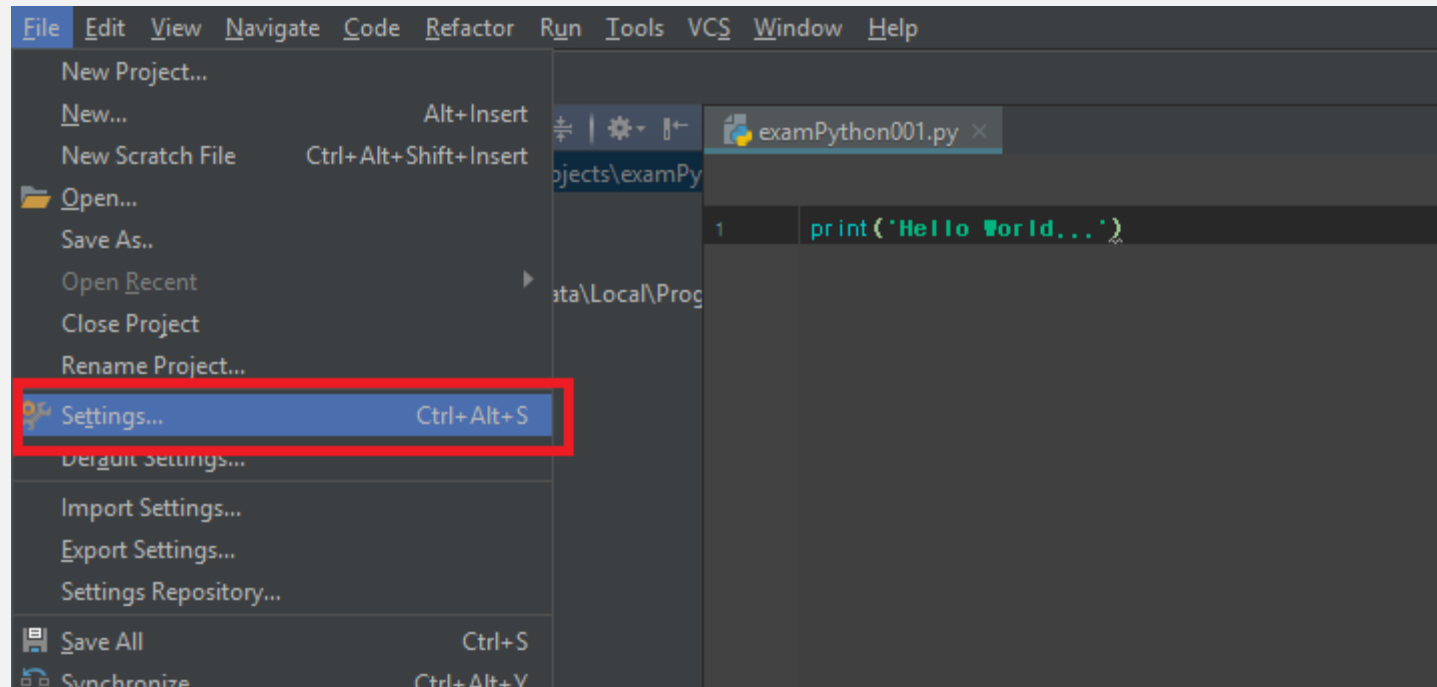
- 설치

학생 계정이 있으면

Professional 버전 무료사용 가능

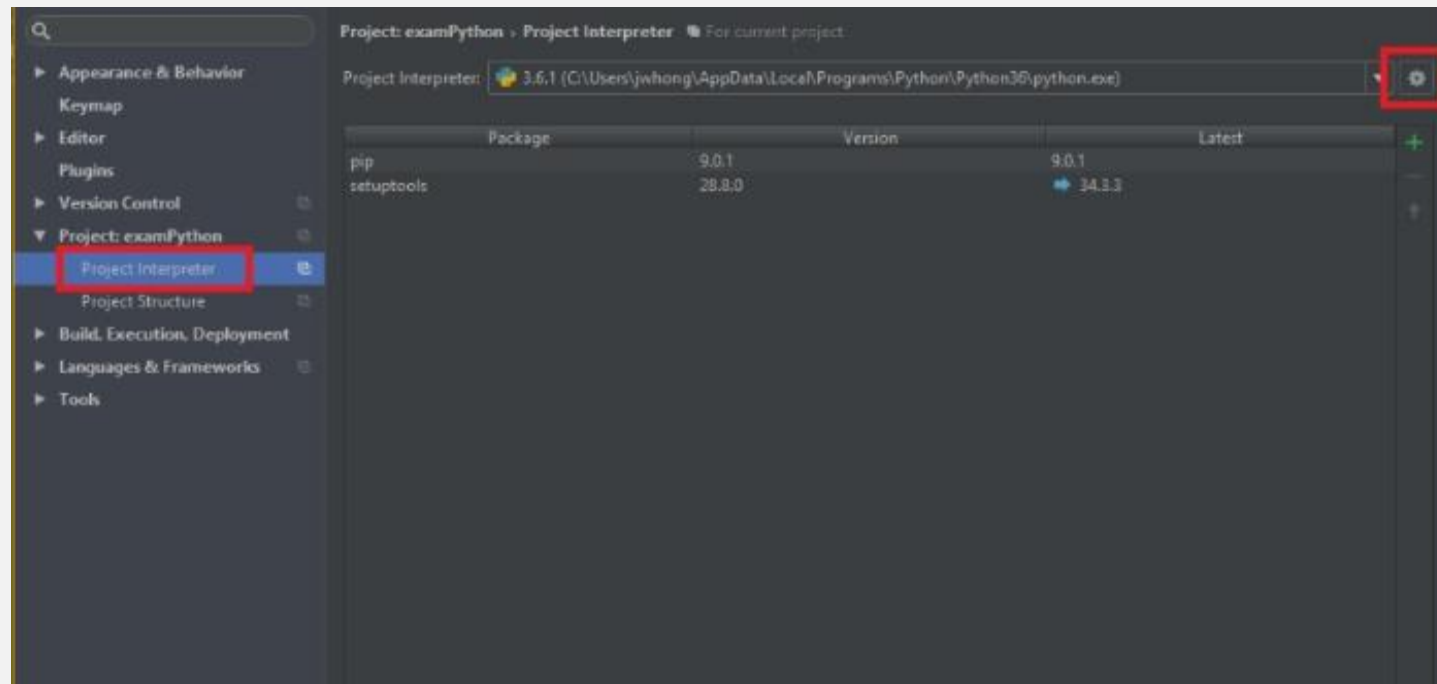
2. Pycharm

- 환경설정



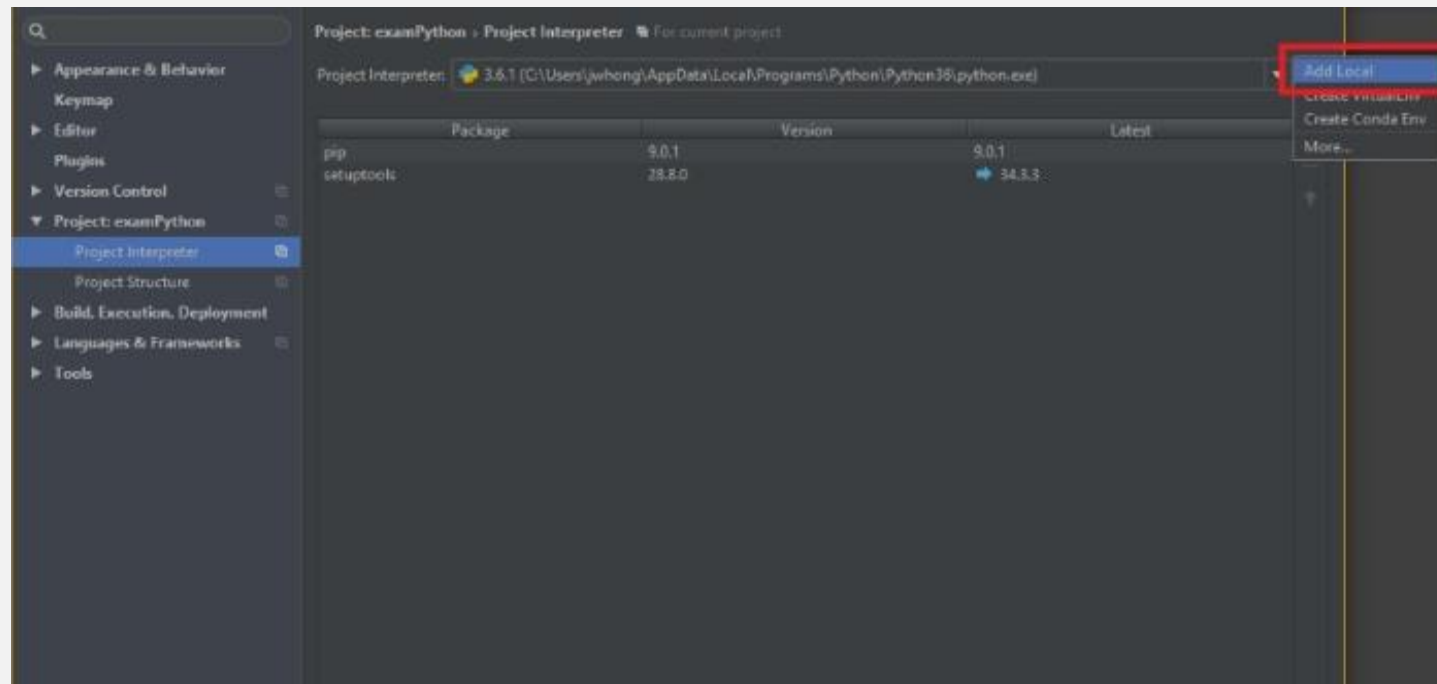
2. Pycharm

- 환경설정



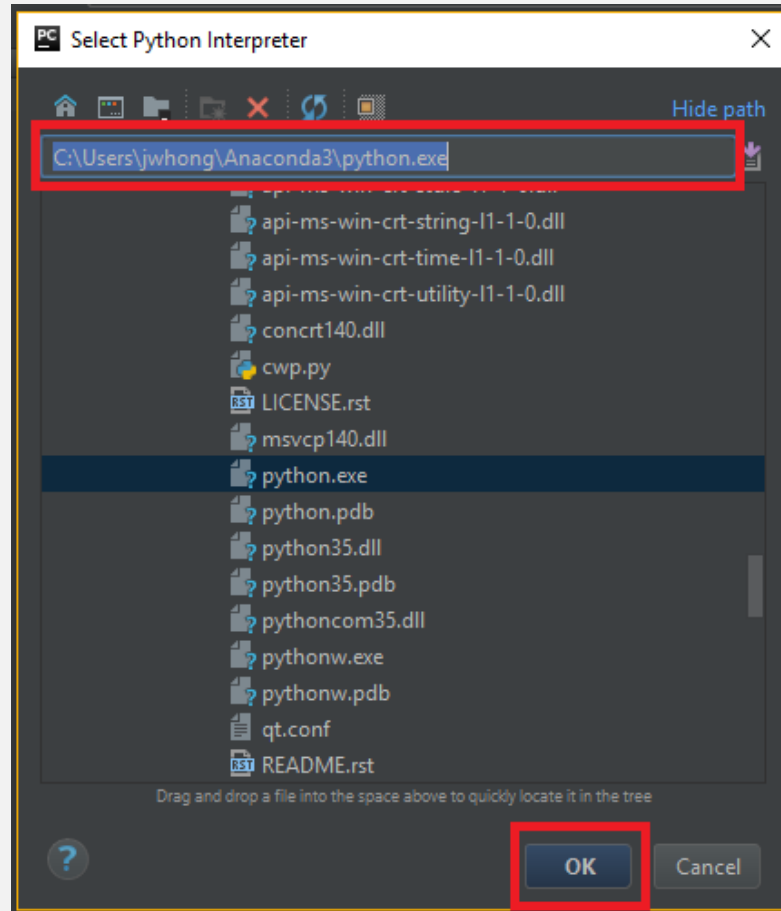
2. Pycharm

- 환경설정



2. Pycharm

- 환경설정



2. Pycharm

- Jupyter Notebook

#%% - 코드

#%% md - 마크다운

Part 1. 원하는거 다 가져와보자 [Crawling 1시간에 끝내기]

1. HTML

1. HTML

- 구조 파악하기

Hyper Text Markup Language

???

1. HTML

- 구조 파악하기

웹 페이지는 구성할 수 있는 언어

1. HTML

- 구조 파악하기

기본 구조

```
<html>  
  <head>  
    <p> 여기는 머리 </p>  
  </head>  
  <body>  
    <p> 여기는 바디 </p>  
  </body>  
</html>
```

1. HTML

- 구조 파악하기

실제 웹페이지는?

F12 또는 페이지 소스보기를 통해 확인가능

www.naver.com

1. HTML & CSS

- XML / XPath

Extensible Markup Language

???

1. HTML & CSS

- XML / XPath

XML은 데이터를 구조화한다

???

1. HTML & CSS

- XML / XPath

Xpath란 XML Path Language를 의미한다.

???

1. HTML & CSS

- XML / XPath

크롤링할 때 경로를 파악하기 유용하다.

Part 1. 원하는거 다 가져와보자 [Crawling 1시간에 끝내기]
2. robots.txt

2. robots.txt

- 사용법

이거 막 가져와도 되나?

➤ 안되지 않을까?

➤ **상관없지**

2. robots.txt

- 사용법

이거 막 가져와도 되나?

➤ 안되지 않을까?

➤ 상관없지 -> 패쓰

2. robots.txt

- 사용법

이거 막 가져와도 되나?

➤ 안되지 않을까?

➤ 상관없지

2. robots.txt

- 사용법

www.naver.com/robots.txt

2. robots.txt

- 사용법

User-agent / Disallow / Allow

User-agent : 다음 규칙이 적용되는 로봇의 이름

Disallow : 차단할 URL 경로

Allow : 차단 된 상위 디렉토리의 하위 디렉토리에있는 URL 경로이며 차단 해제 할 디렉토리

Part 1. 원하는거 다 가져와보자 [Crawling 1시간에 끝내기]

3. Selenium

3. Selenium

- 무엇?

웹드라이버를 통해 웹 자동화를 도와주는 라이브러리

3. Selenium

- 무엇?

여러 브라우저를 지원

Firefox / Internet Explorer / **Chrome** / Safari / Opera

3. Selenium

- 설치 방법

설치 방법

- 명령 프롬프트에서 아래 명령어 실행

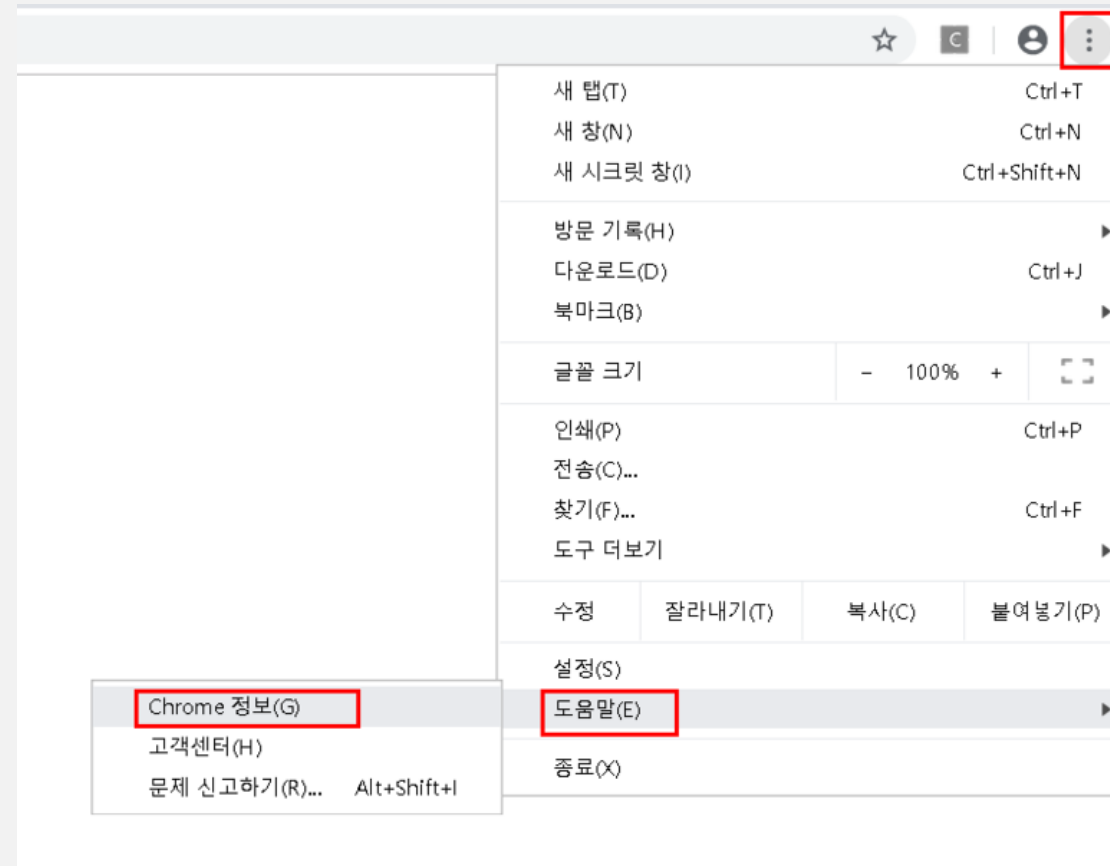
```
pip install selenium
```

Part 1. 원하는거 다 가져와보자 [Crawling 1시간에 끝내기]

4. webdriver

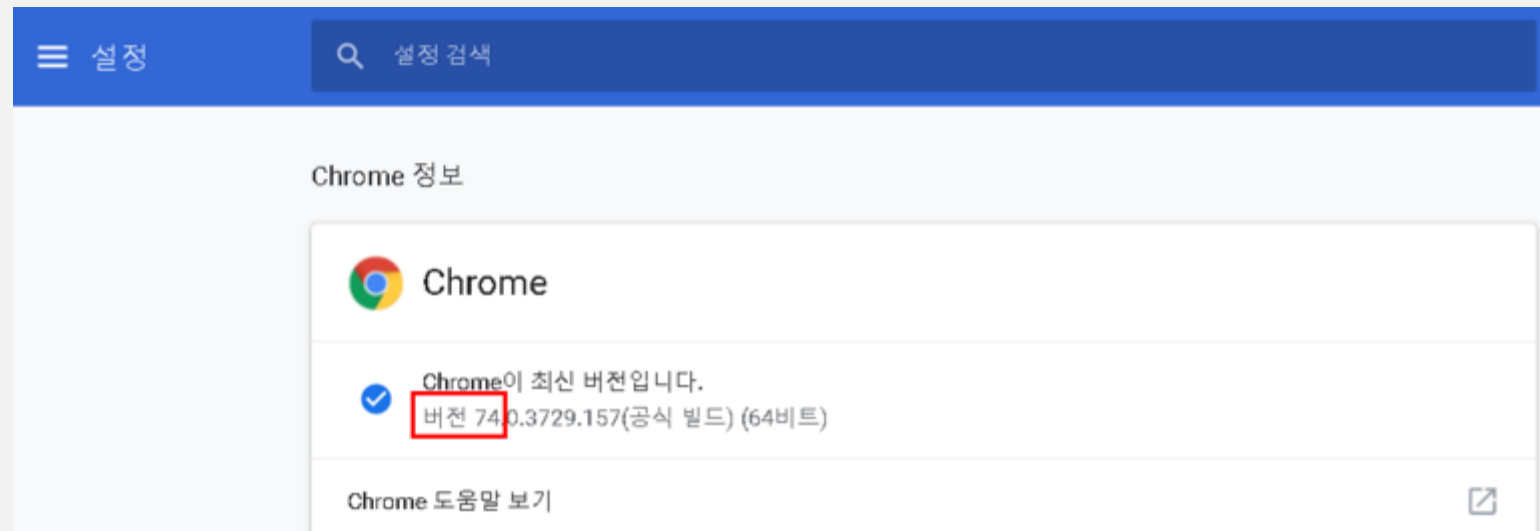
4. webdriver

- 설치



4. webdriver

- 설치



4. webdriver

- 설치

<https://sites.google.com/a/chromium.org/chromedriver/downloads>

https://sites.google.com/a/chromium.org/chromedriver/downloads

ChromeDriver - WebDriver for Chrome

Search this site

- CHROMEDRIVER
- CAPABILITIES & CHROME OPTIONS
- CHROME EXTENSIONS
- CHROMEDRIVER CANARY
- CONTRIBUTING
- DOWNLOADS**
 - VERSION SELECTION
- GETTING STARTED
 - ANDROID
 - CHROME OS
- LOGGING
 - PERFORMANCE LOG
- MOBILE EMULATION
- NEED HELP?
 - CHROME DOESN'T START OR CRASHES IMMEDIATELY
 - CHROMEDRIVER CRASHES
 - CLICKING ISSUES
 - DEVTOOLS WINDOW KEEPS CLOSING
 - OPERATION NOT SUPPORTED WHEN USING REMOTE DEBUGGING
- SECURITY CONSIDERATIONS

Downloads

Current Releases

- If you are using Chrome version 75, please download [ChromeDriver 75.0.3770.8](#)
- If you are using Chrome version 74, please download [ChromeDriver 74.0.3729.6](#)
- If you are using Chrome version 73, please download [ChromeDriver 73.0.3683.68](#)
- For older version of Chrome, please see below for the version of ChromeDriver that supports it.

If you are using Chrome from Dev or Canary channel, please download [ChromeDriver 75.0.3770.8](#). This is not officially supported, but in most cases it should work without major issues.

For more information on selecting the right version of ChromeDriver, please see the [Version Selection](#) page.

ChromeDriver 75.0.3770.8

Supports Chrome version 75

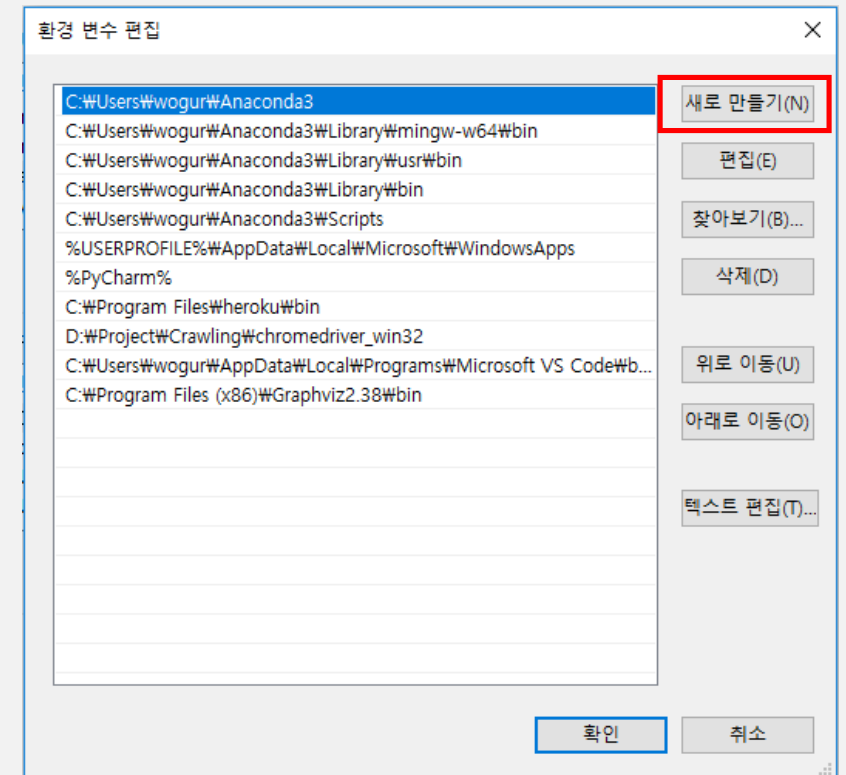
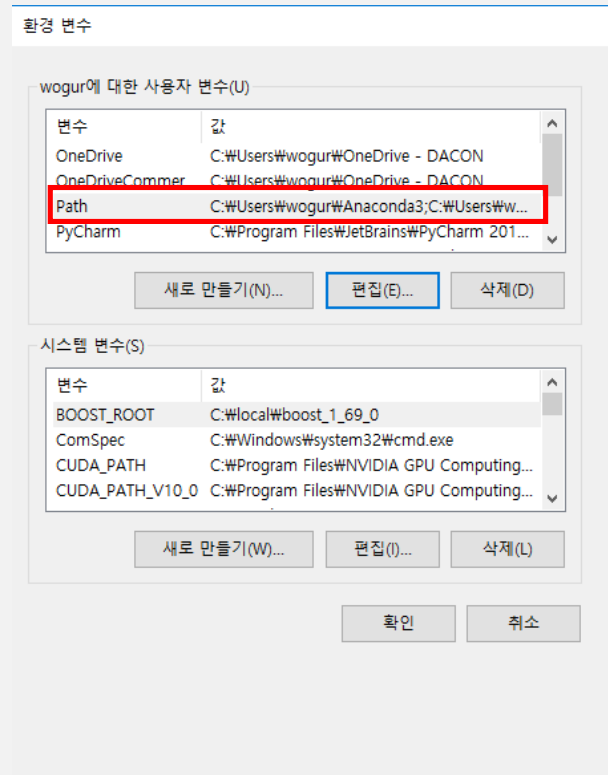
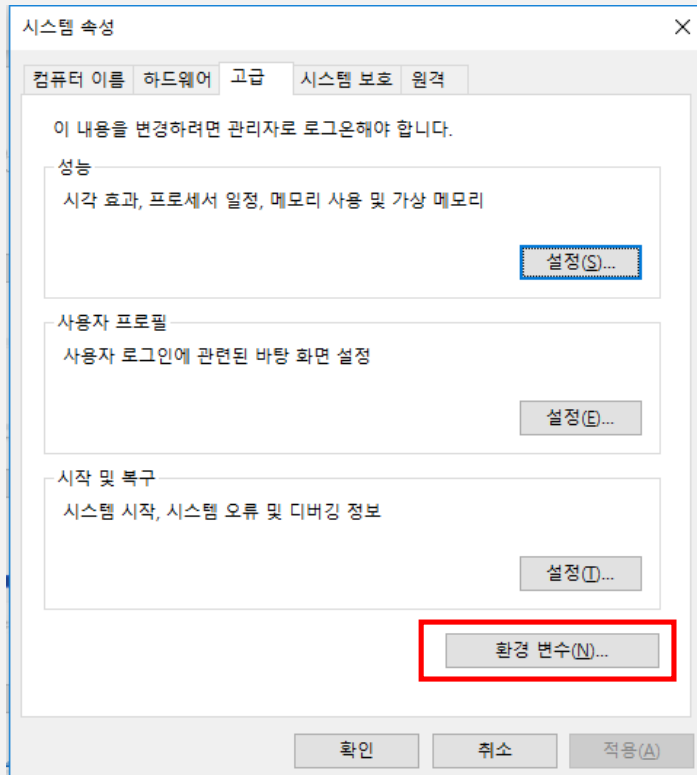
The most noticeable change is ChromeDriver now **runs in W3C standard compliant mode by default**. Other changes include:

- Fixed a bug that caused blur event to be raised twice during element clear
- Renamed capability loggingPrefs to googLoggingPrefs, as required by W3C standard
- Fixed error handling in W3C mode
- Correct handling of creating cookies without leading dot in the domain name
- Allows null for script timeout
- Fixed Element Send Keys command to file input with 'multiple' attribute

4. webdriver

- 환경 설정

환경 변수에서 경로 설정



4. webdriver

- 데이터 가져오기

야구 데이터 가져오기

1. 정답 데이터 가져오기

<https://www.koreabaseball.com/History/Crowd/GraphDaily.aspx>

4. webdriver

- 데이터 가져오기

야구 데이터 가져오기

2. 분석에 사용할 데이터 가져오기

Statiz : <http://www.statiz.co.kr/schedule.php>

날씨 : <https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36>

4. webdriver

- 데이터 가져오기

실습

Part 2. 시각화가 그렇게 어려워? 응, 어려워

1. 시각화의 기본 예절

1. 시각화의 기본 예절

- 시각화의 이유

데이터 시각화를 해야 하는 이유 5가지

1. 많은 양의 데이터를 한눈에 볼 수 있다.
2. 데이터 분석에 대한 전문 지식이 없어도, 누구나 쉽게 데이터 인사이트를 찾을 수 있다.
3. 요약 통계보다 정확한 데이터 분석 결과를 도출할 수 있다.
4. 효과적인 데이터 인사이트 공유로 데이터 기반의 의사결정을 할 수 있다.
5. 데이터 시각화를 활용할 수 있는 분야와 방법이 무궁무진하다.

1. 시각화의 기본 예절

- 시각화의 이유

데이터 시각화를 해야 하는 이유 5가지

1. 많은 양의 데이터를 한눈에 볼 수 있다.
- 2. 데이터 분석에 대한 전문 지식이 없어도, 누구나 쉽게 데이터 인사이트를 찾을 수 있다.**
3. 요약 통계보다 정확한 데이터 분석 결과를 도출할 수 있다.
4. 효과적인 데이터 인사이트 공유로 데이터 기반의 의사결정을 할 수 있다.
5. 데이터 시각화를 활용할 수 있는 분야와 방법이 무궁무진하다.

1. 시각화의 기본 예절

- 시각화의 이유

데이터 시각화를 해야 하는 이유 5가지

1. 많은 양의 데이터를 한눈에 볼 수 있다.

입문자 입장에서 한눈에 내용을 알 수 있게

2. 데이터 분석에 대한 전문 지식이 없어도, 누구나 쉽게 데이터 인사이트를 찾을 수 있다.

3. 요약 통계보다 정확한 데이터 분석 결과를 도출할 수 있다.

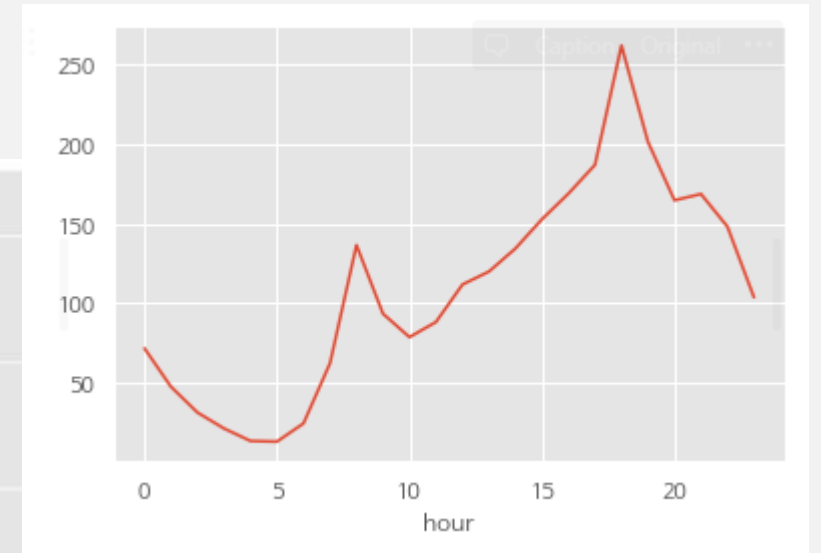
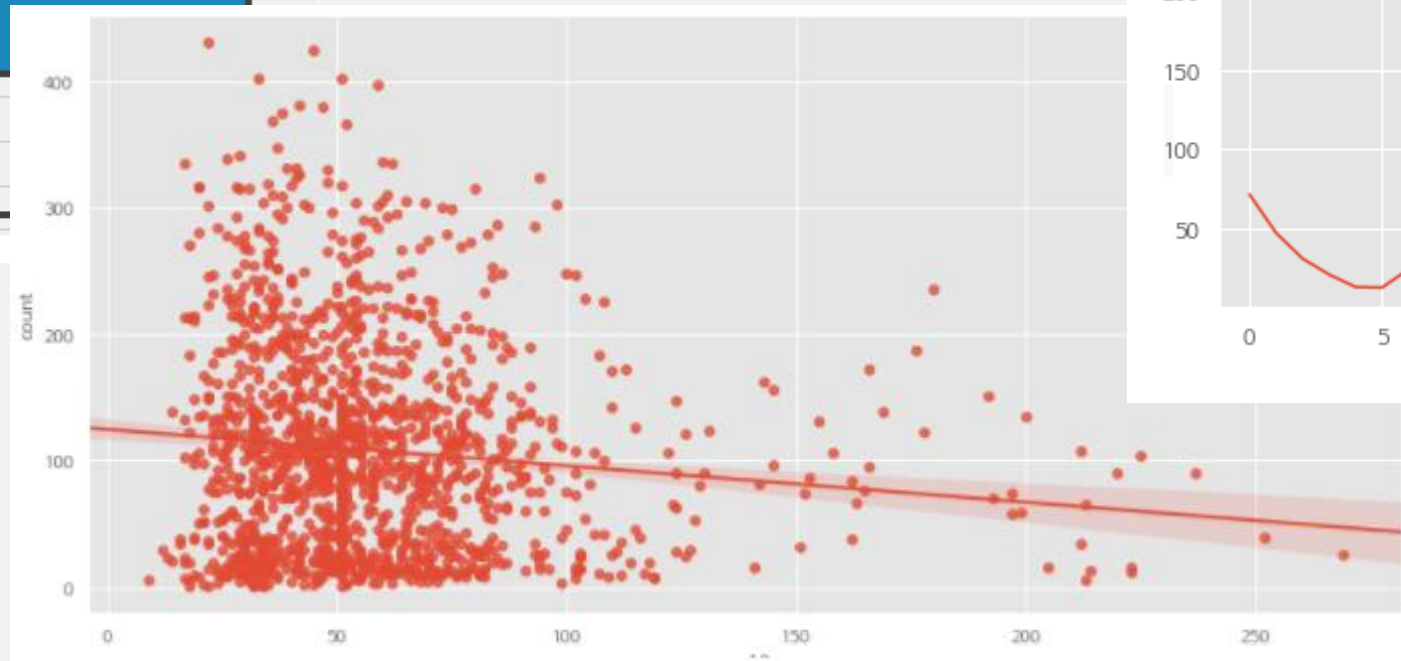
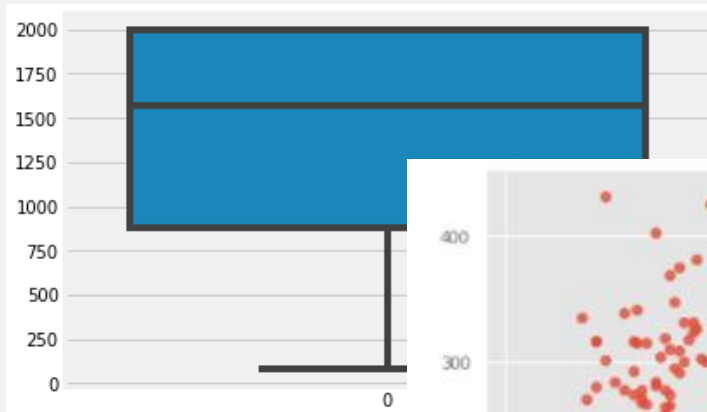
4. 효과적인 데이터 인사이트 공유로 데이터 기반의 의사결정을 할 수 있다.

5. 데이터 시각화를 활용할 수 있는 분야와 방법이 무궁무진하다.

1. 시각화의 기본 예절

- 기본 예절

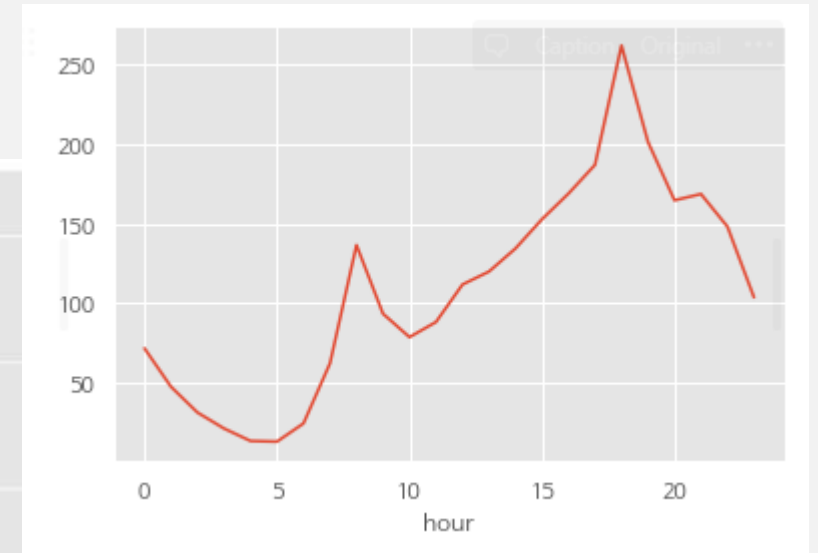
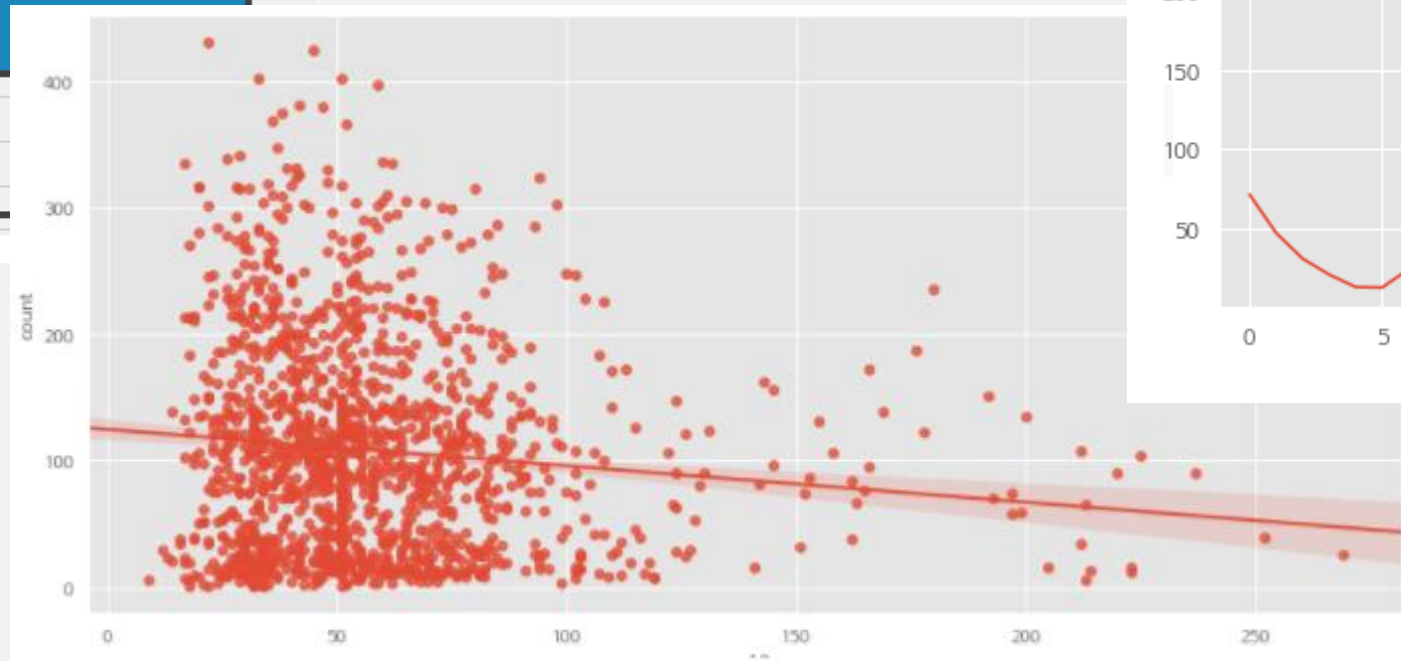
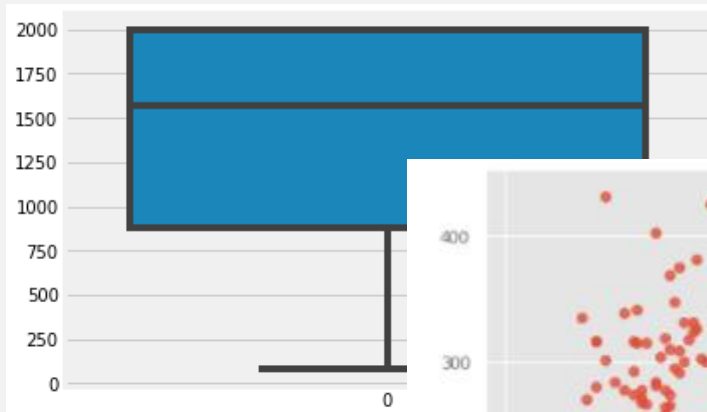
시각화의 좋은 예시



1. 시각화의 기본 예절

- 기본 예절

시각화의 ~~좋은~~ 예시



1. 시각화의 기본 예절

- 기본 예절

세 그래프의 공통점

제목이 없다

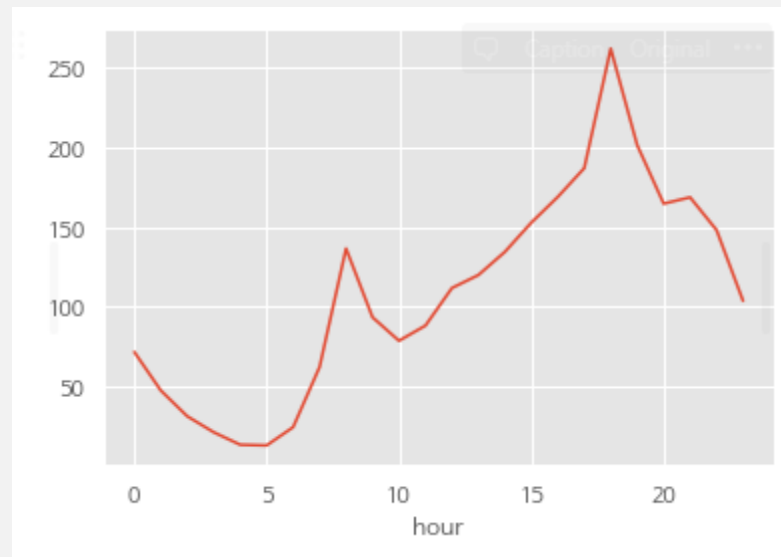
나타낸 값이 뭘까

무슨 의미일까

1. 시각화의 기본 예절

- 기본 예절

값이 무엇을 의미하는지

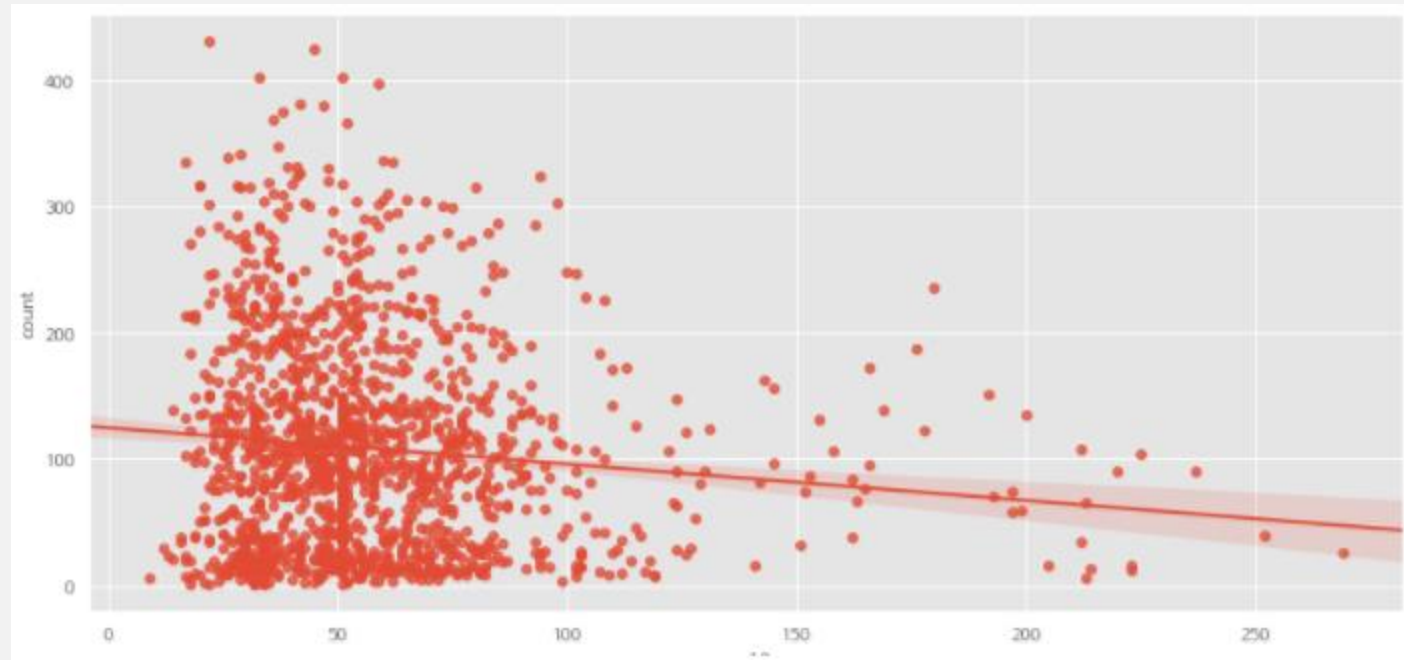


X 축이 나타내는 것

1. 시각화의 기본 예절

- 기본 예절

인사이트가 없는 건 제외



Part 2. 시각화가 그렇게 어려워? 응, 어려워
2. 그래프 어디까지 아니?

2. 다양한 그래프

- 그래프 종류 확인



2. 다양한 그래프

- 그래프 종류 확인



2. 다양한 그래프

- 그래프 종류 확인



2. 다양한 그래프

- 그래프 종류 확인



Part 2. 시각화가 그렇게 어려워? 응, 어려워
3. Python 그래프 라이브러리

3. Python 그래프 라이브러리

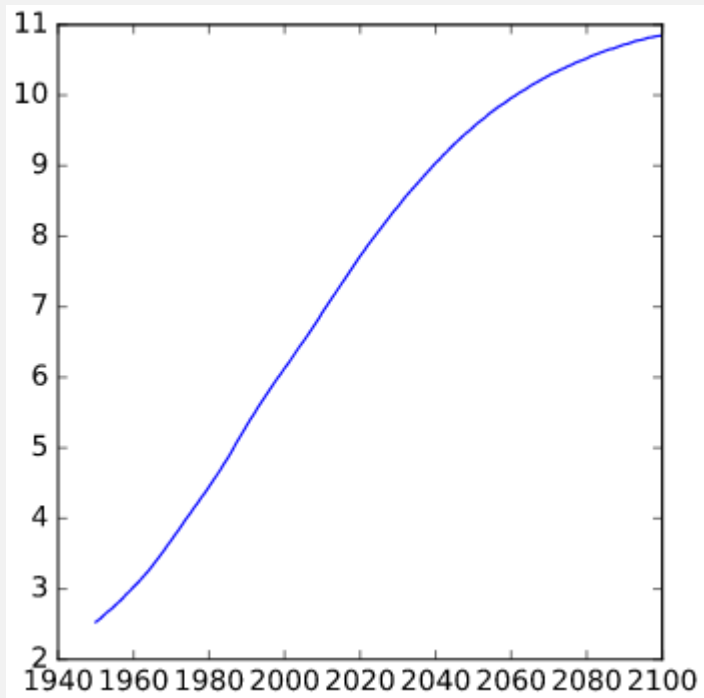
- matplotlib

데이터를 시각화(visualization)할 수 있게
도와주는 python의 기본적인 모듈 중 하나

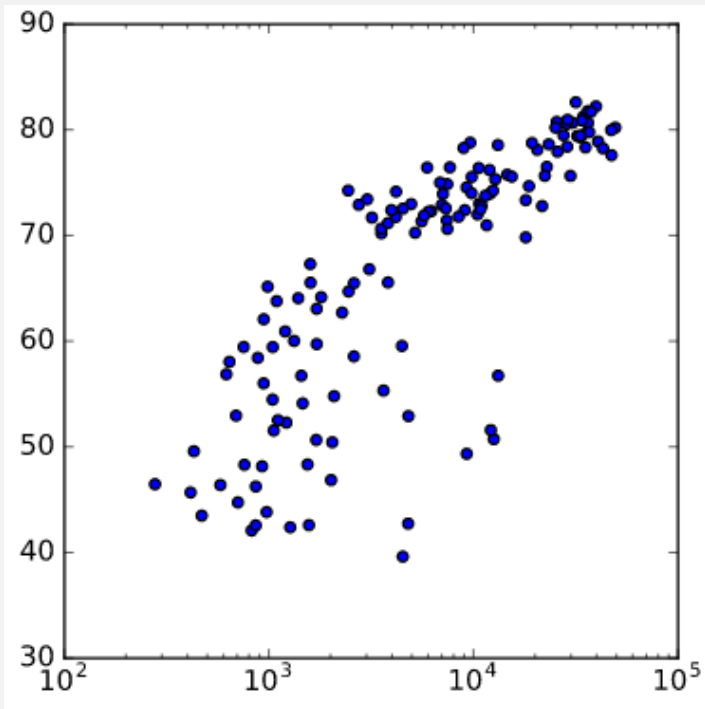
3. Python 그래프 라이브러리

- matplotlib

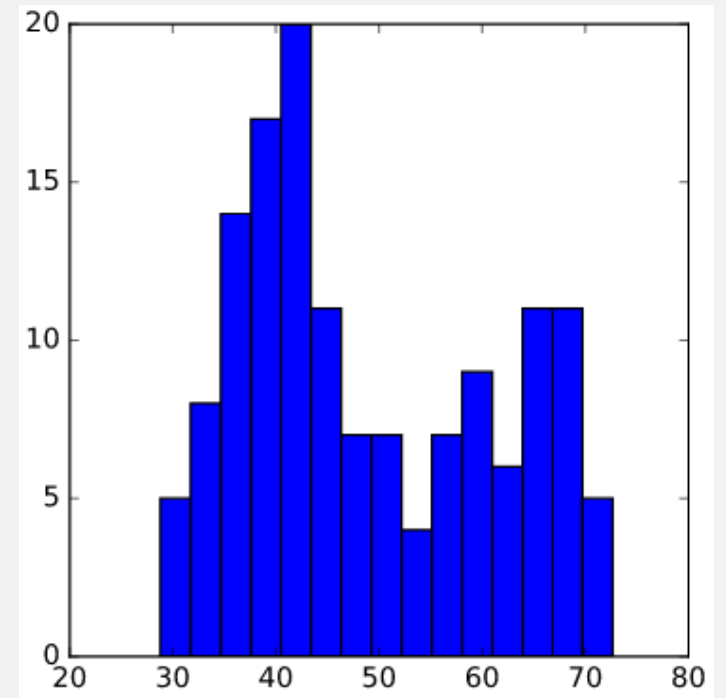
Line Plot



Scatter plot



Histogram



3. Python 그래프 라이브러리

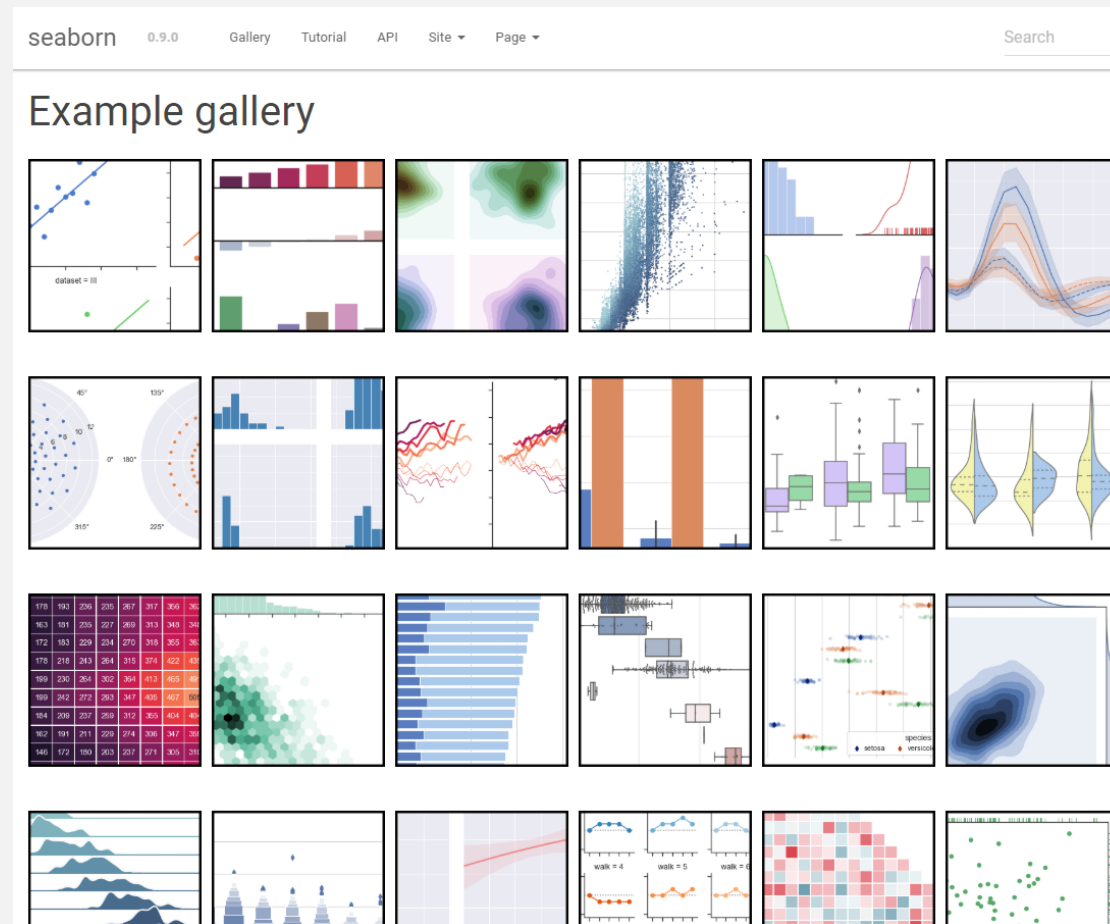
- seaborn

Matplotlib를 기반으로 다양한 색상과
차트 등의 기능이 추가된 라이브러리

3. Python 그래프 라이브러리

- seaborn

<https://seaborn.pydata.org/examples/index.html>



Part 2. 시각화가 그렇게 어려워? 응, 어려워

4. 살아 움직이는 그래프

4. Interactive Graph

- Pyecharts 설치

https://pyecharts.org/#/en-us/charts_base

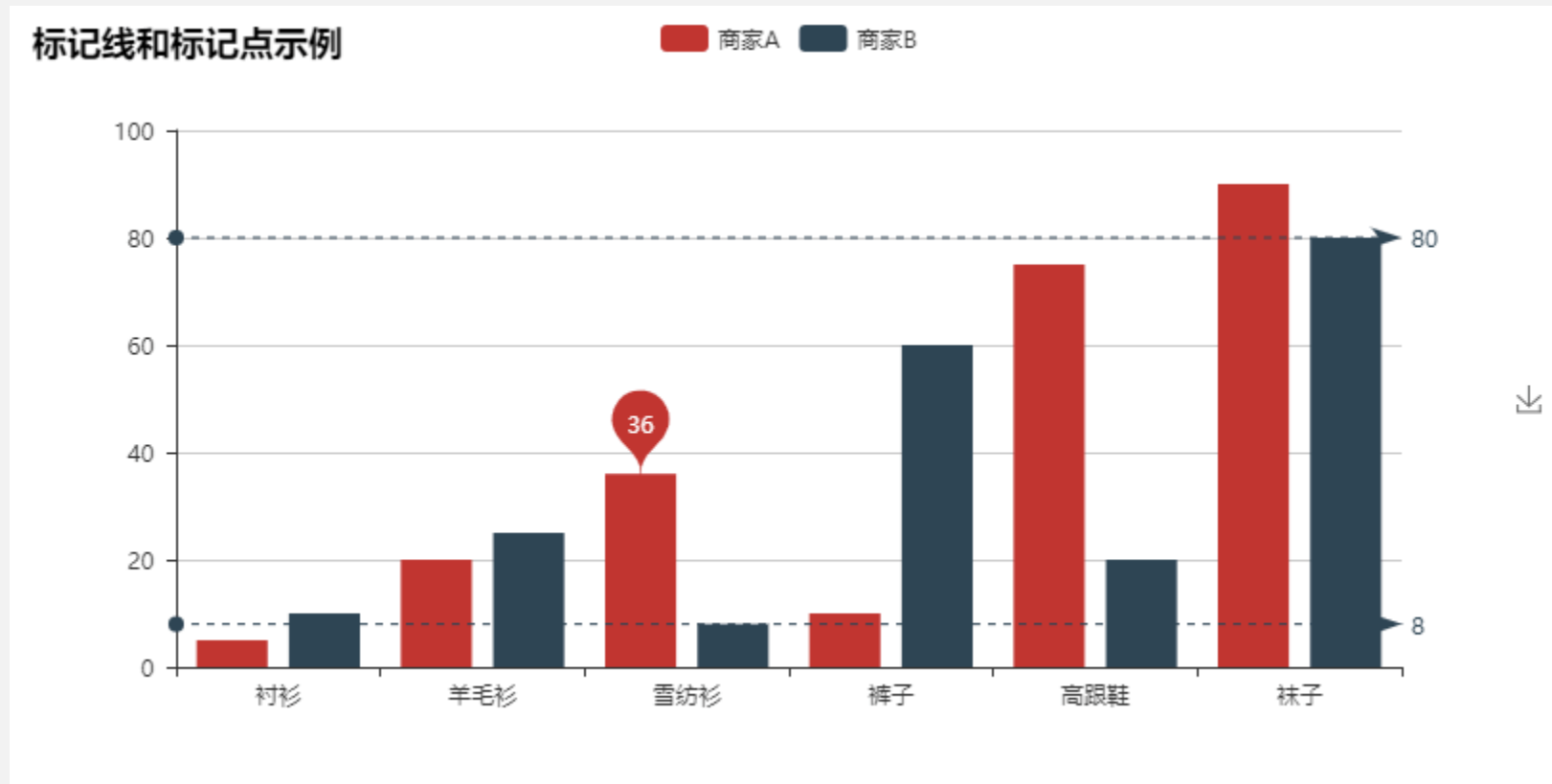
설치 방법

- 명령 프롬프트에서 아래 명령어 실행

```
pip install pyecharts==0.5.8
```

4. Interactive Graph

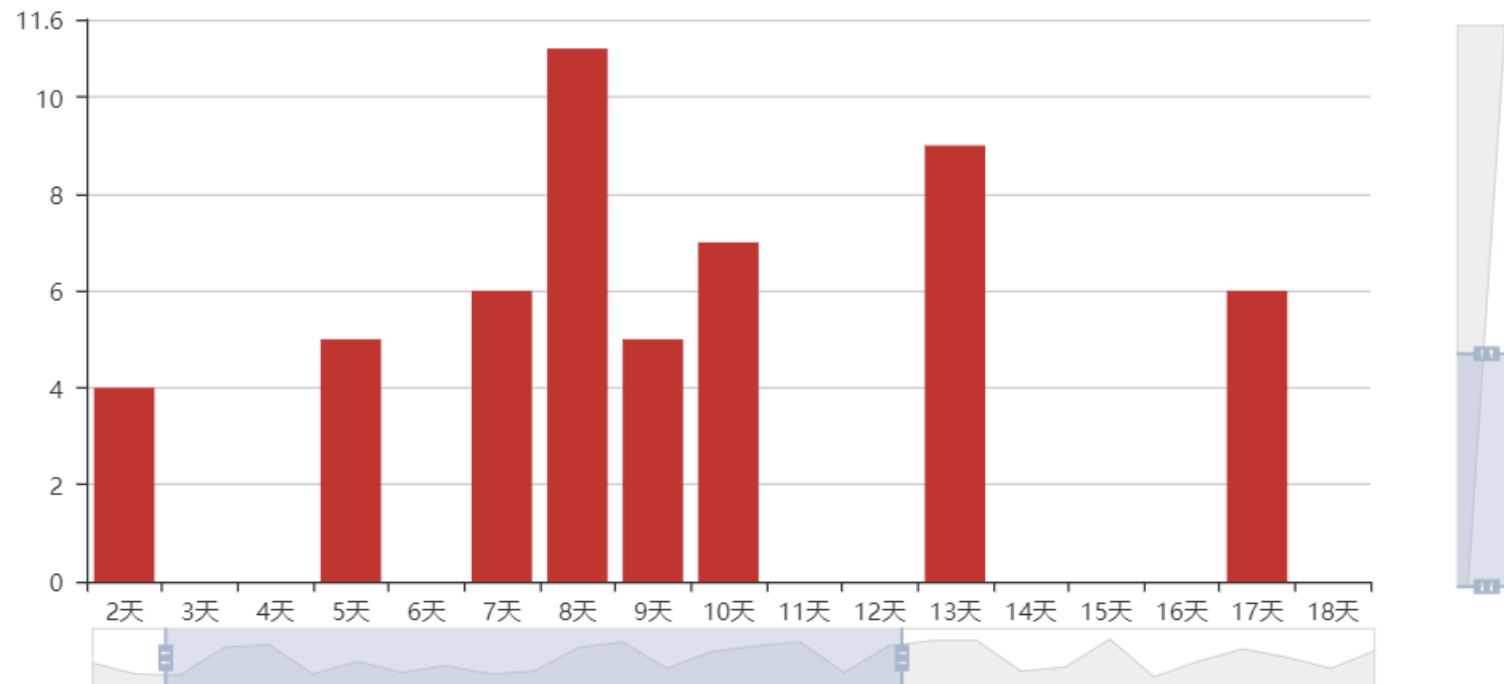
- Bar Plot



4. Interactive Graph

- Bar Plot

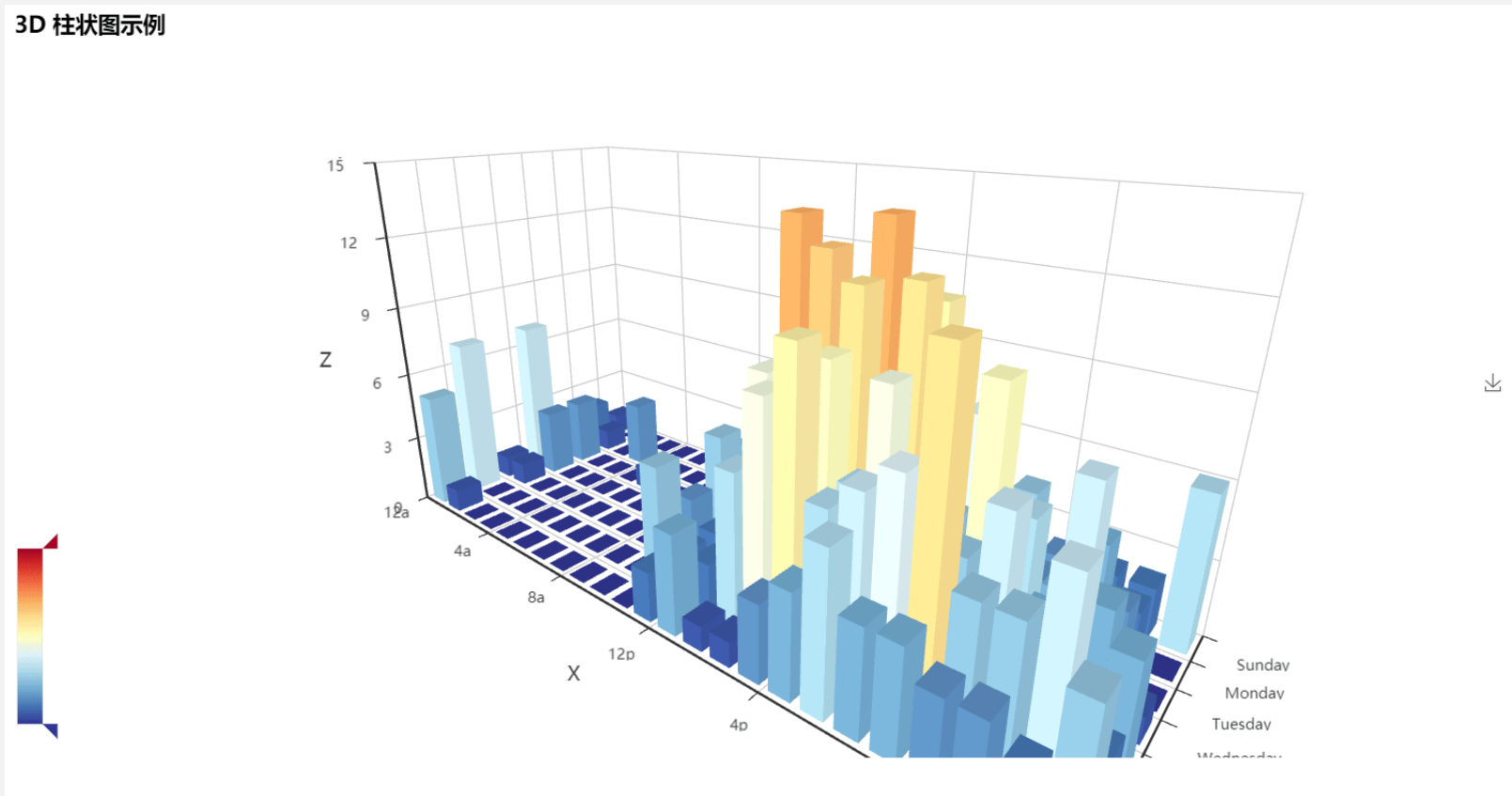
Bar - datazoom - xaxis/yaxis 示例



4. Interactive Graph

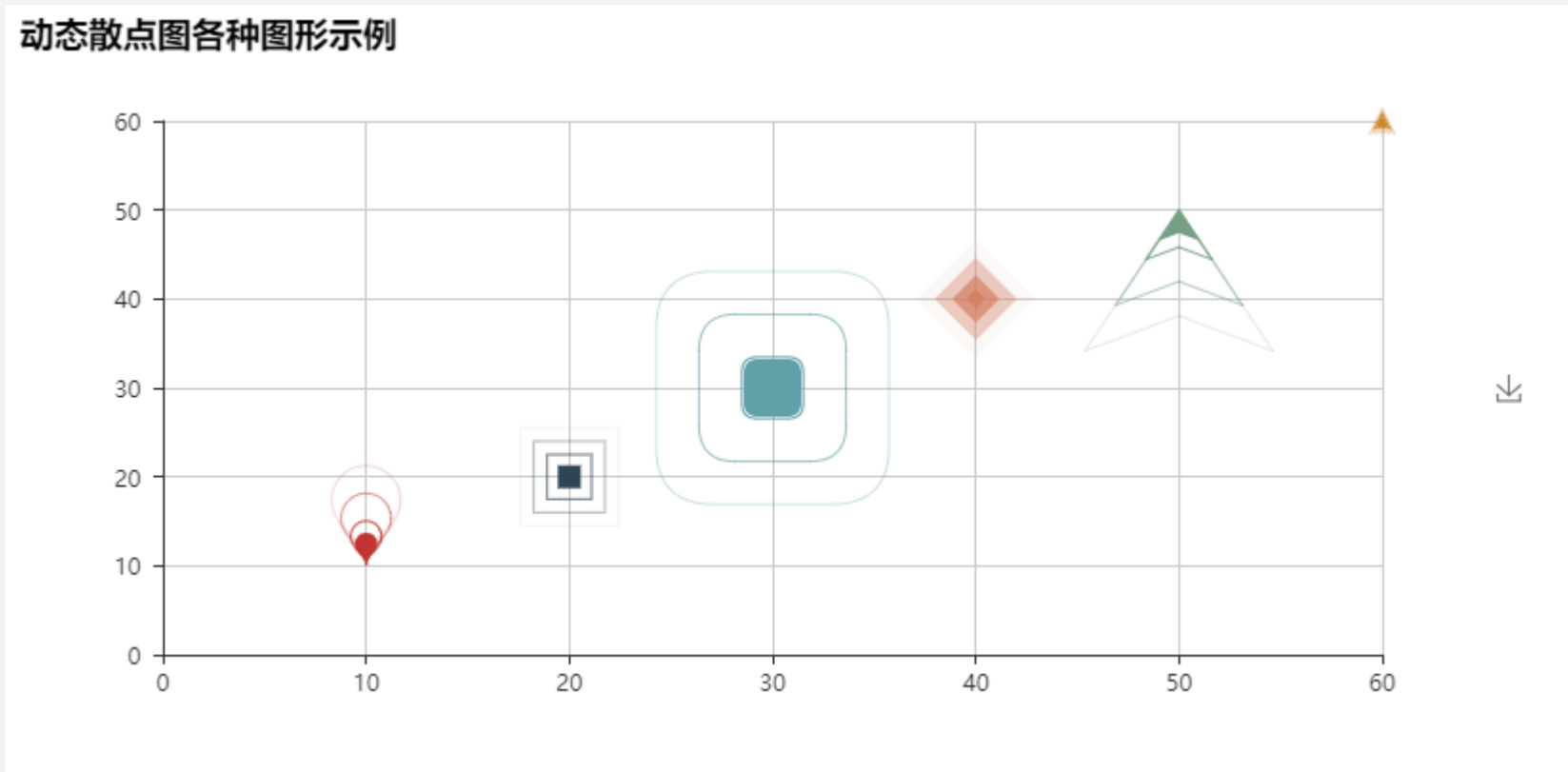
- 3D Bar Plot

3D 柱状图示例



4. Interactive Graph

- Scatter Plot



4. Interactive Graph

- Dot Map



4. Interactive Graph

- Liquid Plot

水球图示例

