

# 코랩과 구글드라이버 연동

400제로 배우는 파이썬 입문 부록

# 판다스를 사용할 수 있는 여러 방식

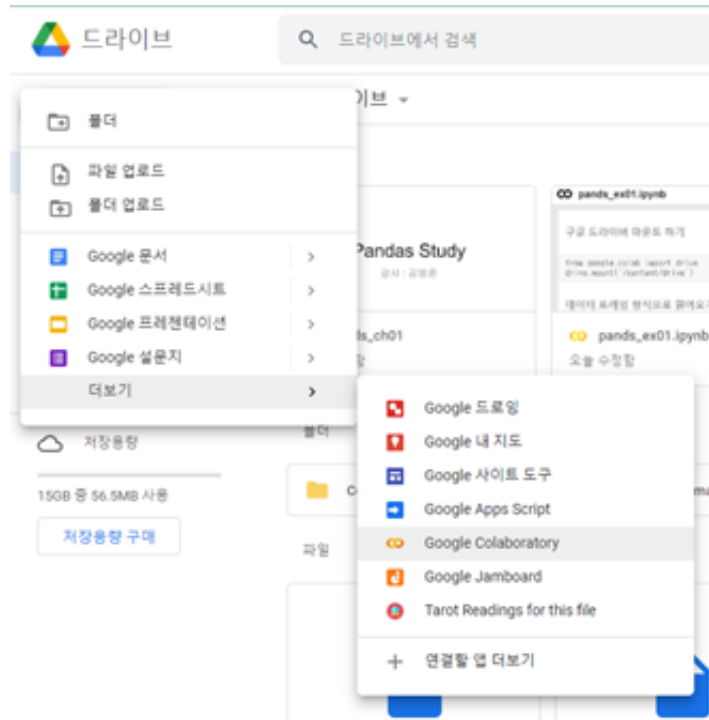
1. PC에 파이썬 설치 후 파이썬 사용
2. 파이참 **IDE** 설치 후 파이썬 사용
3. PC에 아나콘다 설치 후 파이썬 사용
4. 구글 **colab**과 구글 드라이브를 연동해서 파이썬 사용

# 구글 코랩과 구글 드라이브 연동해서 사용하기

- 구글 코랩은 구글에서 **colab**을 검색 해서 바로 사용 가능합니다.
- 코랩을 사용하는 다른 방법으로 구글 드라이브에서 새로만들기 버튼을 눌러서 코랩 파일을 생성해서 사용할 수 있습니다.
- 구글 드라이브와 구글 **colab**은 구글 회원 가입만 하면 무료로 사용 가능합니다.
- 이 문서는 구글 코랩에서 구를 드라이브를 마운트해서 사용하는 방법을 알려 줍니다.
- 구글 드라이브를 코랩에서 마운트하면 구글 드라이브에 파일 저장 및 파일 불러오기가 가능 해 집니다.

# 구글 드라이브에서 코랩 파일 생성

새로만들기 > 더보기 > Google Colaboratory



# 구글 Colab 사용하기

1. 구글 회원가입 (이미 가입되었다면 기존 계정 사용)
2. 구글 드라이브 사용
  - a. 파일 관리
  - b. 구글 드라이브의 유용한 기능들
  - c. 구글 워크스페이스
3. 구글 코랩 사용
  - a. 새 노트
  - b. 구글 드라이브에서 확인
  - c. 구글 드라이브 mount
  - d. 구글 드라이브 파일 i/o
  - e. 구글 드라이브 데이터 파일 사용

# Google에서 colab 검색

구글에서 colab을 검색해서 바로 사용 가능 (<https://colab.research.google.com/>)

The image shows a Google search for 'colab'. The search results page displays the Google Colab website as the top result. The website content includes a welcome message, a brief description of Colab's capabilities (importing datasets, training classifiers, and evaluating models), and links to various resources like 'Google Colab', 'Using Google Colab with GitHub', and 'Overview Notebook'. The sidebar on the right contains a section titled 'Collaboratory란?' (What is Collaboratory?) which explains that Colab is a cloud-based environment for running Python code. It lists features such as '구성 필요하지 않음' (No setup required), 'GPU 무료 액세스' (Free GPU access), and '간편한 공유' (Easy sharing). The sidebar also includes a '시작하기' (Get started) section with instructions on how to use Colab, mentioning that it can be used from a web browser or a local IDE like Command/Outliner. At the bottom of the sidebar, there are two code snippets demonstrating basic operations: creating a list and adding an element to it.

Google

colab

전체 이미지 동영상 쇼핑 지도 더보기

도구

검색결과 약 10,900,000개 (0.33초)

<https://research.google.com/collaboratory>

### Welcome To Colaboratory - Google Research

With Colab you can import an image dataset, train an image classifier on it, and evaluate the model, all in just a few lines of code. Colab notebooks execute...

[Google Colab](#) [Colab](#) [Using Google Colab with GitHub](#) [Overview Notebook](#)

<https://colab.sandbox.google.com/notebooks>

### Colab - Google

Colab은 다음과 같은 이점을 자랑합니다. 구성이 필요하지 않음, GPU 무료 액세스, 간편한 공유, 학생이든, 데이터 과학자든, AI 연구원이든 Colab으로 업무를...

<https://colab.research.google.com>

### Google Colab

Sign in.

<https://colab.research.google.com/notebooks/intro>

### Google Colab Notebook Intro

Sign in.

### Collaboratory란?

줄여서 'Colab'이라고도 하는 Collaboratory를 사용하면 브라우저에서 Python을 작성하고 실행할 수 있습니다. Colab은 다음과 같은 이점을 자랑합니다.

- 구성이 필요하지 않음
- GPU 무료 액세스
- 간편한 공유

학생이든, 데이터 과학자든, AI 연구원이든 Colab으로 업무를 더욱 간편하게 처리할 수 있습니다. [Colab 소개 영상](#)에서 자세한 내용을 확인하거나 아래에서 시작해 보세요.

### 시작하기

지금 읽고 계신 문서는 어떻게 할까요? 아니면 코드를 작성하고 실행할 수 있는 대화형 환경인 Colab 예제집입니다.

예를 들어 다음은 값을 계산하여 변수로 저장하고 결과를 출력하는 간단한 Python 스크립트가 포함된 코드 셀입니다.

```
1 | second_half = 2 + 5 * 60
   second_half
None
```

위 셀의 코드를 실행하려면 셀을 클릭하여 선택한 후 코드 셀의 실행 버튼을 누르거나 단축키 'Command/Ctrl+Enter'를 사용하세요. 셀을 클릭하면 코드 수정을 바로 시작할 수 있습니다.

특정 셀에서 정의한 변수를 나중에 다른 셀에서 사용할 수 있습니다.

```
1 | second_half = 7 + second_half
   second_half
None
```

Colab 예제집을 사용하면 실행 코드와 시작되는 텍스트를 여백지, HTML, LaTeX 등과 함께 하나의 문서로 통합할 수 있습니다. Colab 예제집을

# 구글 드라이브 마운트

- 코랩에서 구글 드라이브를 마운트 하기 위해서는 인증 코드가 필요 함.

```
from google.colab import drive  
drive.mount('/content/drive')
```

... Go to this URL in a browser: <https://accounts.google.com/ConnectApp?authuser=1&continue=https%3A%2F%2Fcolab.research.google.com%2F%3Fipynop%3D1>

Enter your authorization code:


Google

로그인

이 코드를 복사하여 애플리케이션으로 전환한 다음 붙여넣으세요.

4/1AX4XfWju0ZEYlQwCWiHTZMTW\_iq1P2ySQgVBD2\_dN0f  
FHh3kf9A5a97B9Jo 

```
[ ] from google.colab import drive  
drive.mount('/content/drive')
```

 Mounted at /content/drive

# 파이썬 파일 입출력 기능으로 구글 드라이브에 저장

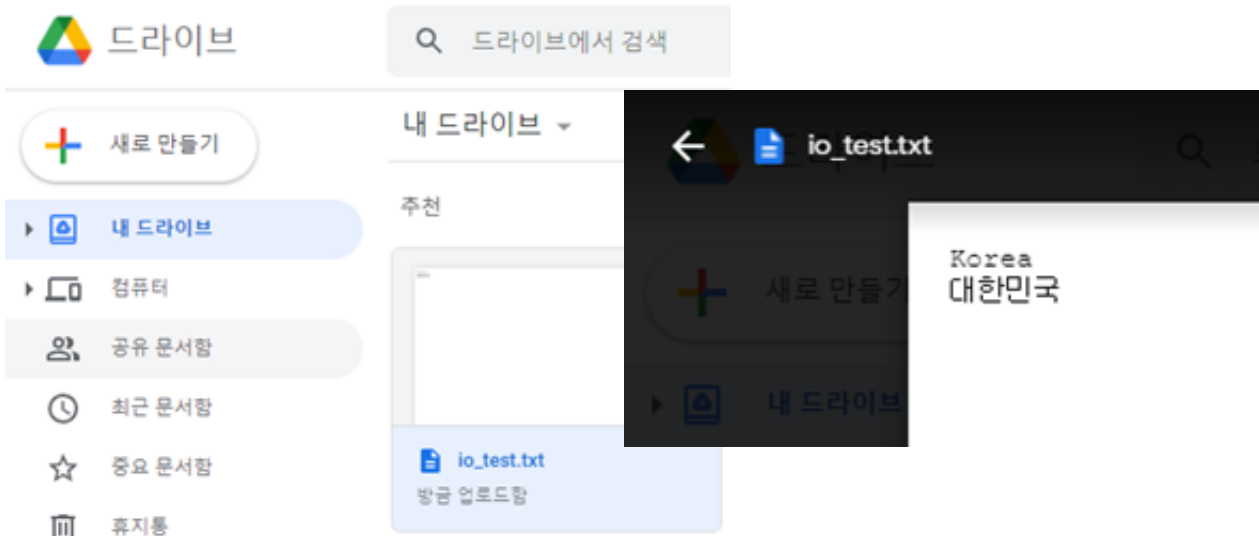
```
from google.colab import drive  
drive.mount('/content/drive')
```

```
fp = open('/content/drive/MyDrive/io_test.txt', 'w', encoding='utf-8')
```

```
fp.write('Korea\n')
```

```
fp.write('대한민국\n')
```

```
fp.close()
```





# 구글 드라이브에서 데이터 읽어오기

```
filepath = '/content/drive/MyDrive/gapminder.tsv'
```

```
import pandas as pd
```

```
df=pd.read_csv(filepath, '\t')
```

```
print(df.head())
```

```
print(df.tail())
```

	country	continent	year	lifeExp	pop	gdpPercap
0	Afghanistan	Asia	1952	28.801	8425333	779.445314
1	Afghanistan	Asia	1957	30.332	9240934	820.853030
2	Afghanistan	Asia	1962	31.997	10267083	853.100710
3	Afghanistan	Asia	1967	34.020	11537966	836.197138
4	Afghanistan	Asia	1972	36.088	13079460	739.981106
	country	continent	year	lifeExp	pop	gdpPercap
1699	Zimbabwe	Africa	1987	62.351	9216418	706.157306
1700	Zimbabwe	Africa	1992	60.377	10704340	693.420786
1701	Zimbabwe	Africa	1997	46.809	11404948	792.449960
1702	Zimbabwe	Africa	2002	39.989	11926563	672.038623
1703	Zimbabwe	Africa	2007	43.487	12311143	469.709298

# 분석한 데이터 구글 드라이브로 저장

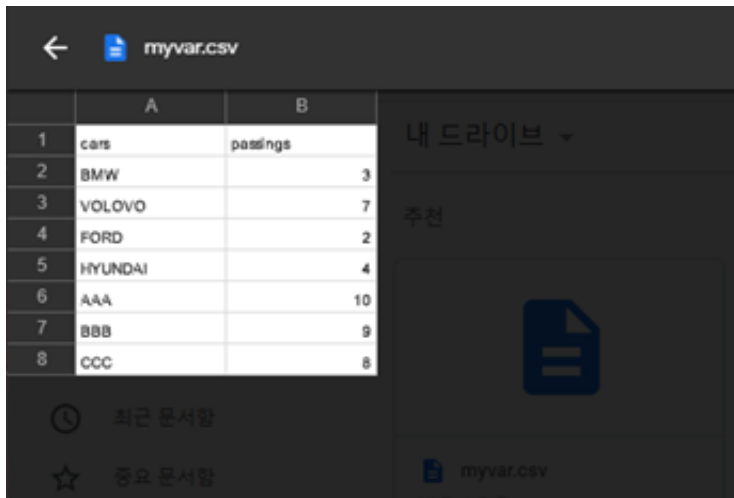
```
from google.colab import drive
drive.mount('/content/drive')

import pandas as pd

mydataset = {
    'cars': ['BMW', 'VOLOVO', 'FORD', 'HYUNDAI', 'AAA', 'BBB', 'CCC'],
    'passings': [3, 7, 2, 4, 10, 9, 8]
}

#int(data), float(data), string(data)
df = pd.DataFrame(mydataset)
print(df.head())

savepath = '/content/drive/MyDrive/myvar.csv'
df.to_csv(savepath, index=False, encoding='utf-8')
```



The screenshot shows the Google Drive interface. At the top, there is a back arrow and a file icon next to the name 'myvar.csv'. Below this is a table with two columns, 'A' and 'B'. The table contains 8 rows of data. To the right of the table, there is a sidebar with the text '내 드라이브' (My Drive) and a dropdown arrow. Below this, there is a section labeled '추천' (Recommendations) with a large blue document icon. At the bottom of the sidebar, there is a section labeled '최근 문서함' (Recent Documents) and '중요 문서함' (Important Documents).

	A	B
1	cars	passings
2	BMW	3
3	VOLOVO	7
4	FORD	2
5	HYUNDAI	4
6	AAA	10
7	BBB	9
8	CCC	8

# List형식의 데이터를 Series 형식으로 변환

```
a = [11, 22, 33, 44, 55]
print(type(a))
```

```
myvar = pd.Series(a)
print(type(myvar))
```

```
calories = {'day1': 420, 'day2': 380, 'day3': 290}
```

```
myvar = pd.Series(calories)
```

```
print(myvar)
```

```
a = [10, 70, 20]
```

```
myvar = pd.Series(a, index=['x', 'y', 'z'])
```

```
print(myvar)
```

```
myvar2 = pd.Series(calories, index=['day1', 'day2'])
```

```
print(myvar2)
```

# Dict 형식의 데이터를 DataFrame 형식으로 변환

```
mydataset = {  
    'cars': ['BMW', 'VOLVO', 'FORD', 'HYUNDAI', 'AAA', 'BBB', 'CCC'],  
    'passings': [3, 7, 2, 4, 10, 9, 8]  
}
```

```
#int(data), float(data), string(data)
```

```
myvar = pd.DataFrame(mydataset)
```

```
print(type(myvar))
```

```
#print(myvar.head())
```

```
print(myvar.loc[2])
```

```
<class 'pandas.core.frame.DataFrame'>  
cars      FORD  
passings      2  
Name: 2, dtype: object
```

# BeautifulSoup를 이용한 웹 크롤링(1)

```
from google.colab import drive
drive.mount('/content/drive')

import requests
from bs4 import BeautifulSoup

response = requests.get('https://www.naver.com/')
if response.status_code == 200 :
    soup = BeautifulSoup(response.text, 'html.parser')
    print(soup.select('ul.list_nav.'))
```

# BeautifulSoup를 이용한 웹 크롤링(2)

```
response = requests.get('https://www.naver.com/')
```

```
if response.status_code == 200 :
    soup = BeautifulSoup(response.text, 'html.parser')
    # print(soup.select('ul.list_nav.type_fix li a'))
    data_list = soup.select('ul.list_nav.type_fix li a')
    title_lis = []
    link_lis = []
    for i, item in enumerate(data_list) :
        title_lis.append(item.text)
        link_lis.append(item['href'])

#print(title_lis, link_lis)
dataset = {
    "title": title_lis,
    "link" : link_lis
}
```

	title	link
0	메일	<a href="https://mail.naver.com/">https://mail.naver.com/</a>
1	카페	<a href="https://section.cafe.naver.com/">https://section.cafe.naver.com/</a>
2	블로그	<a href="https://section.blog.naver.com/">https://section.blog.naver.com/</a>
3	지식IN	<a href="https://kin.naver.com/">https://kin.naver.com/</a>
4	쇼핑	<a href="https://shopping.naver.com/">https://shopping.naver.com/</a>

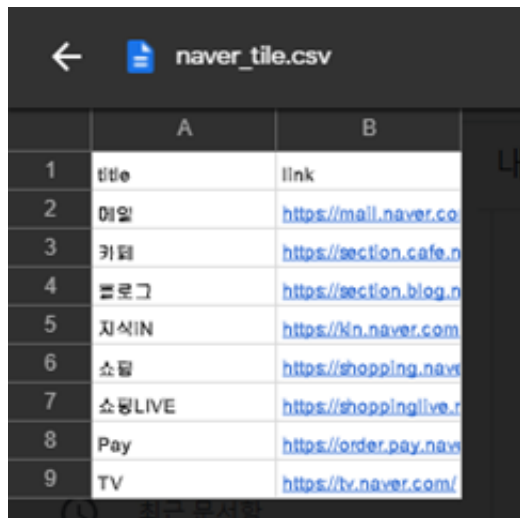
# 크롤링 한 데이터를 구글 드라이브로 저장

```
import pandas as pd
```

```
df = pd.DataFrame(dataset)
```

```
print(df.head())
```

```
df.to_csv('/content/drive/MyDrive/naver_tile.csv', index=False, encoding='utf-8')
```



	A	B
1	title	link
2	메일	<a href="https://mail.naver.co">https://mail.naver.co</a>
3	카페	<a href="https://section.cafe.n">https://section.cafe.n</a>
4	블로그	<a href="https://section.blog.n">https://section.blog.n</a>
5	지식IN	<a href="https://kin.naver.com">https://kin.naver.com</a>
6	쇼핑	<a href="https://shopping.naver">https://shopping.naver</a>
7	쇼핑LIVE	<a href="https://shoppinglive.r">https://shoppinglive.r</a>
8	Pay	<a href="https://order.pay.naver">https://order.pay.naver</a>
9	TV	<a href="https://tv.naver.com/">https://tv.naver.com/</a>