

Visual Python 소개서

Contents

1. 프로젝트 추진 배경	02
1) 빅데이터 분석 시장	
2) 빅데이터 분석의 중요성	
3) 디지털트랜스포메이션과 데이터 분석 교육	
4) 데이터 분석과 사회문제의 해결	
5) 빅데이터 분석을 위한 프로그래밍 언어	
6) Python의 장점과 활용 이유	
2. 분석 솔루션 현황	09
1) 프로그래밍의 기술 장벽	
2) 셀프 서비스 비즈니스 인텔리전스(SSBI)의 필요성	
3) 셀프 서비스 비즈니스 인텔리전스(SSBI)의 한계	
4) Low-Code와 No-code의 차이	
3. Visual Python 소개	14
1) 솔루션 개요	
2) 개발 배경 및 목적	
3) 핵심 기능 및 활용 분야	
4) 개발 현황	
5) 솔루션 차별점	
4. Visual Python 현황	21
1) 사용자 현황	
2) 커뮤니티 현황	
3) 활용 사례	

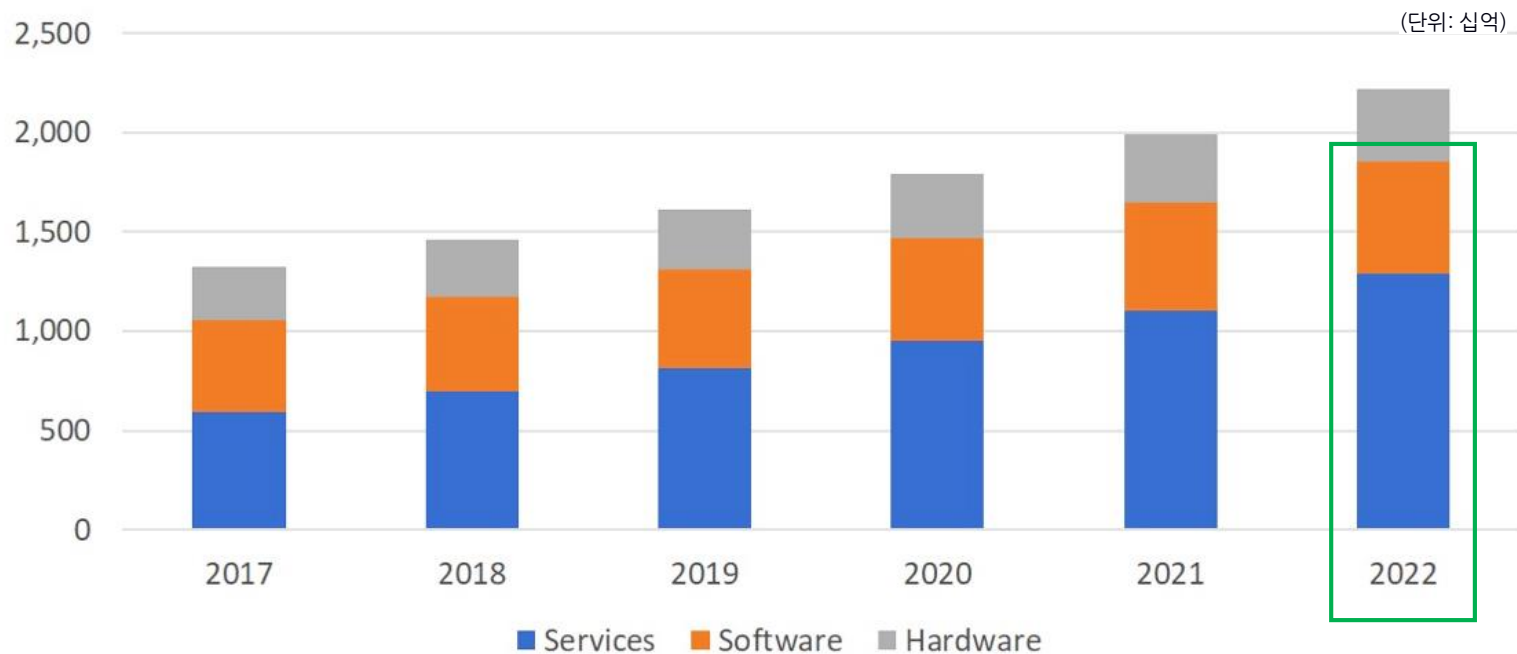
1. 프로젝트 추진 배경

1. 프로젝트 추진배경

1) 빅데이터 분석 시장

빠르게 변화하는 국내 빅데이터 및 분석 시장 5년간 연평균 10%대로 성장, 2022년에는 약 2조 2천억원의 규모에 이를 것으로 전망: 빠르게 변화하는 비즈니스 환경에 대처하면서 데이터 흐름을 실시간으로 수집하고 분석하는 것이 과거에 비해 중요한 화두가 되고 있기 때문.

〈국내 빅데이터 및 분석 시장 전망 2017-2022〉



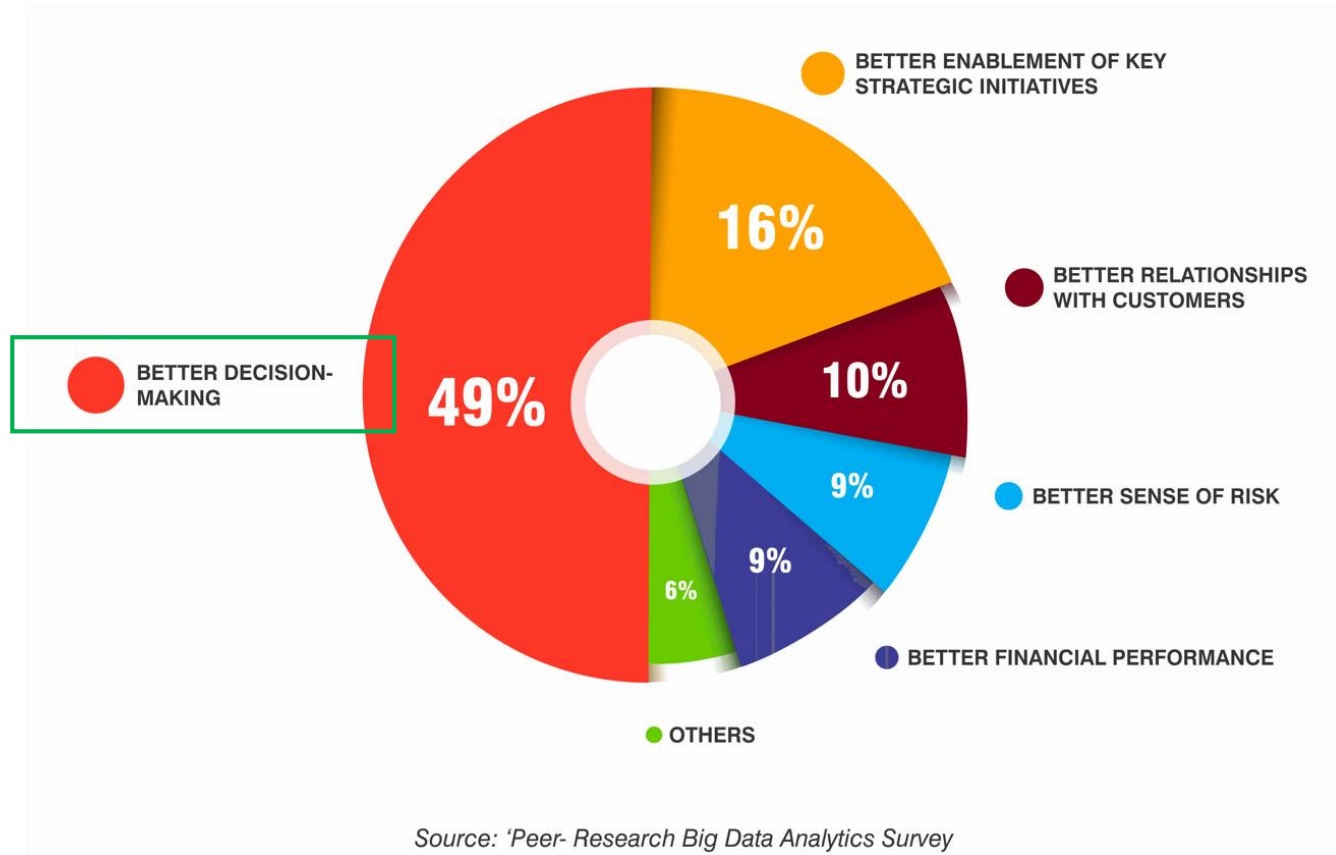
Source: IDC, 2019

1. 프로젝트 추진배경

2) 빅데이터 분석의 중요성

빅데이터(Big Data) 분석을 통한 기업 가치의 증대: 데이터 중심 의사 결정에 의존하는 회사는 평균적으로 5% 더 생산성이 높고 6%는 경쟁 업체보다 수익성이 높음.

※참고 논문: Strength in Numbers: How Does Data-Driven Decision making Affect Firm Performance?



1. 프로젝트 추진배경

3) 디지털 트랜스포메이션과 데이터 분석 교육

빅데이터 활용 기술은 모든 분야에서 문제 해결을 위한 필수 역량이 되어가는 추세. 이제는 일반 사무직도 업무 효율화 (반복 업무 자동화) 등을 위해 코딩을 배우고 있음.

KISTI-법무부, 외국인행정 빅데이터 분석 교육훈련

주 박정환 기자 | © 입력 2022.04.24 13:44 | © 수정 2022.04.24 13:50 | 0 댓글



데이터 분석 교육받고 공공기관서 실무...청년 350명 모집

송고시간 | 2022-04-10 12:00



(서울=연합뉴스) 김병규 기자 = 행정안전부는 '2022년 데이터 분석 청년인재 양성사업' 참가 희망자 350명을 오는 11일~다음달 2일 모집한다고 10일 밝혔다.

이 사업은 데이터 분석 교육을 받고 공공기관에서 실무경험을 쌓을 수 있게 돕는다.

만 19세~34세 청년이면 이 사업의 홈페이지(<https://niadata.recruitcenter.kr>)를 통해 참가 신청을 할 수 있다.

인/공무원

[NOW] 직장에도 부는 '코딩교육' 바람

최예지 조선예듀 기자

조선예듀 기사 작성일 : 2019.12.03 06:00



{기업 전사 차원으로 교육해 인재 양성
업무 효율화' 목적으로 배우는 직장인도



/하나금융그룹 제공

알록달록한 블록을 모아 붙이니 프로그램 하나가 금세 완성됐다. 어린이용 블록코딩 프로그램인 '스크래치'를 활용해 코딩하고 있는 이들은 다름 아닌 직장인들. 간단한 프로그램을 만들고 좋아하는 모습은 영락없는 학생이다. 최근 하나금융은 연말까지 그룹 관계사 전 임직원을 대상으로 코딩교육을 진행하기로 했다. 임원들은 모바일 애플리케이션을 직접 만들어보는 심화 교육도 받는다.

1. 프로젝트 추진배경

4) 빅데이터 분석과 사회문제의 해결

빅데이터는 다양한 사회 문제 해결의 열쇠. 기술이 사회 문제 해결을 위해 존재해야 한다는 인식이 확산되며, 기업들도 빅데이터 및 인공지능 기술을 사회공헌 프로젝트에 활용하는 사례가 많아짐.

빅데이터분석으로 복잡한 사회문제 분석·해결 [신종 사회문제, 기술로 푼다] '착한 ICT'가 뜬다

▲ 남연우 수습 기자 | ✉ nkssooel@HelloDD.com | ⌚ 입력 2019.01.30 09:43 | ⌚ 수정 2019.01.30 18:50 | 💬 댓글 0

노경조·정명섭 기자 | 입력 : 2020-11-16 05:55

KISTI 이슈브리프 제 5호 발간, 관찰·예측으로 사회문제 진단·발굴
사회문제 전과정의 빅데이터 분석을 일관성있게 추진해야



공유하기



이통3사 친환경 화두...감염병 예방도 '척척'
네이버·카카오 사회문제 해결 강조...AI 필수
"지속가능 기업 위해 사회문제 관심 필수"

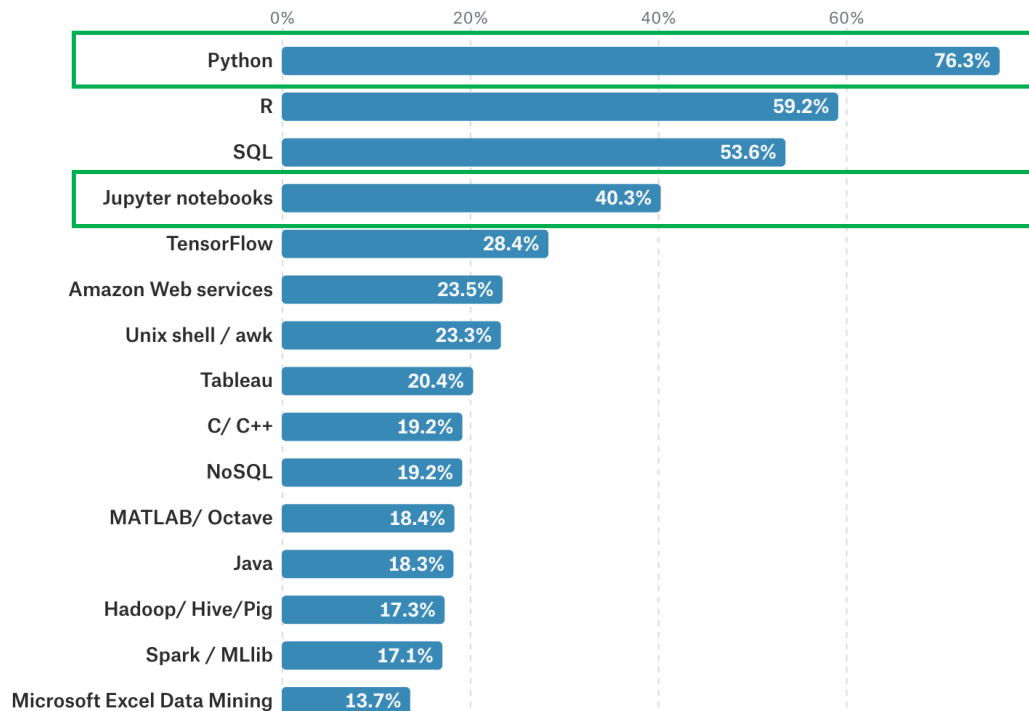
5G와 A·B·C(인공지능·빅데이터·클라우드) 기술의 발전은 우리 경제와 사회 전반에 혁신적인 변화를 일으키고 있다. 첨단 기술의 발전은 기업의 수익모델로 직결되기도 하지만, 단순히 기업의 수익만을 따질 수 없는 시대가 됐다. "사회문제를 어떻게 해결할 것인가"라는 지속가능개발목표(SDGs)와 환경·사회·지배구조(ESG)라는 관점에서 기술이 개발돼야 한다는 목소리가 높아지면서 기술이 사회문제 해결을 위해 존재해야 한다는 인식이 퍼지기 시작했다. 이러한 시대적 흐름을 간파한 국내외 IT 기업들은 각자가 보유한 기술로 사회를 더 나은 방향으로 변화시키겠다는 '착한 ICT'를 지향하고 나섰다.

1. 프로젝트 추진배경

5) 빅데이터 분석을 위한 프로그래밍 언어

Python(파이썬)은 데이터 사이언스/데이터 프로세싱, 웹 & 비즈니스 앱 분야에서 가장 많이 사용 & 선호 되는 리딩 프로그래밍 언어. 사용자 친화성과 오픈소스 커뮤니티로부터 엄청난 양의 최신 기술이 Python으로 공유되기 때문.

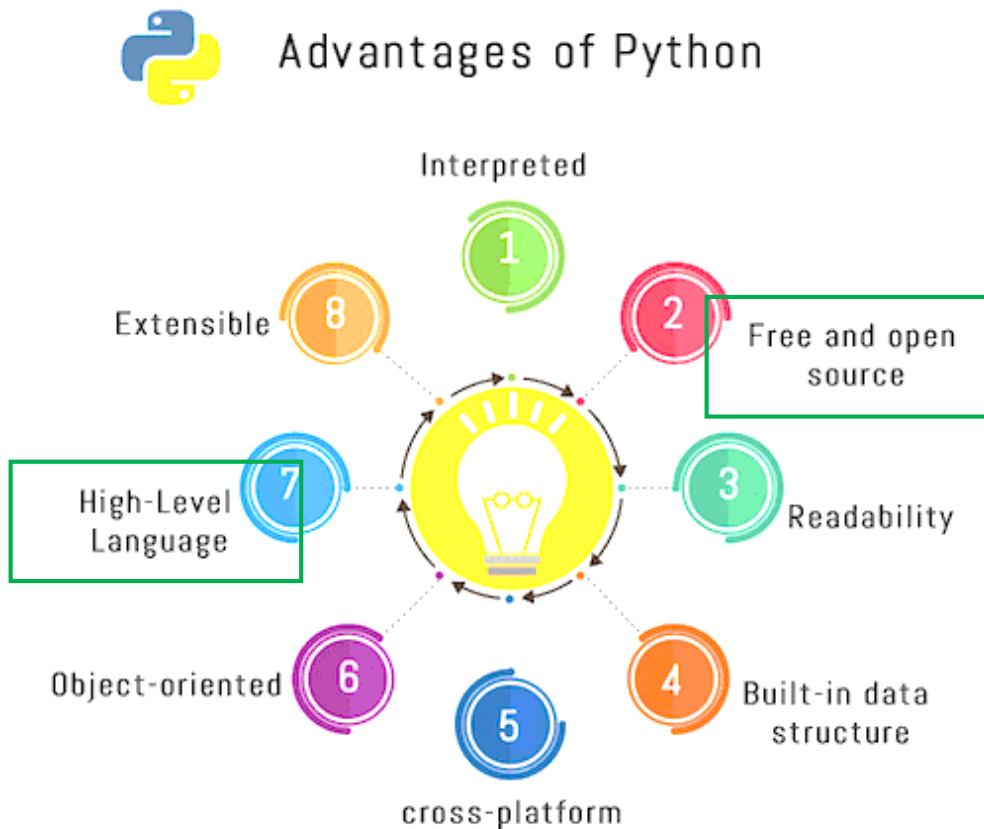
Programming Languages of Data Science



1. 프로젝트 추진배경

6) Python의 장점과 활용 이유

Python은 무료 오픈 소스로 누구나 무료로 사용 할 수 있고 간결하고 활용 범위가 큰 프로그래밍 언어로 현재 전세계 820만이 사용하고 있으며, 사용자 수는 꾸준히 증가 추세

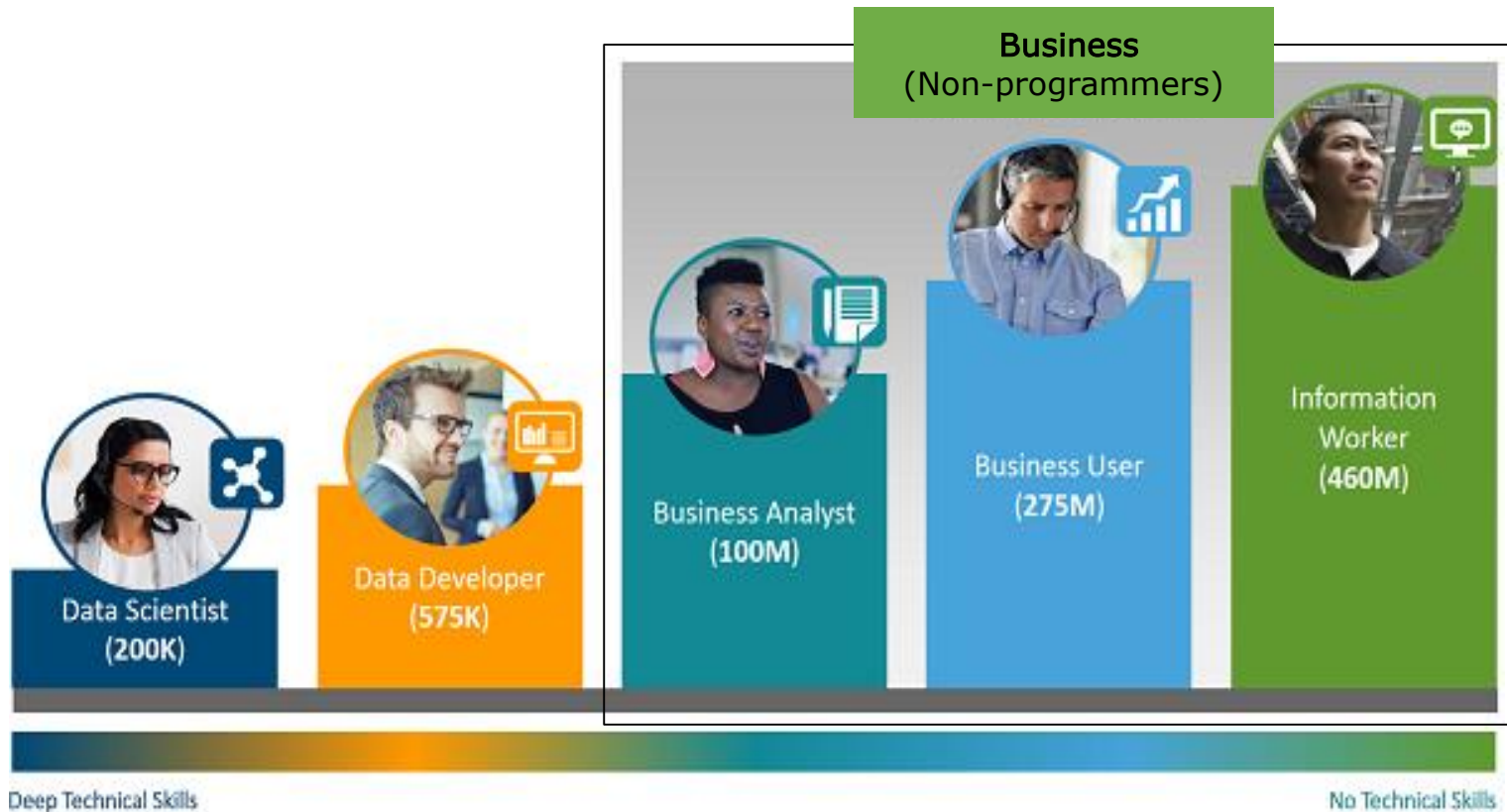


2. 분석 솔루션 현황

2. 분석 솔루션 현황

1) 프로그래밍의 기술장벽

비즈니스 조직이, IT 조직에 의존하지 않고 직접 데이터를 분석하고자 하는 니즈가 확대 되었지만 비전문가가 단기간에 데이터 분석을 위한 **프로그래밍 언어**를 습득 및 **비즈니스에 활용하기**에는 높은 기술적 장벽이 존재



(Source: Paxata)

2. 분석 솔루션 현황

2) 셀프 서비스 비즈니스 인텔리전스(SSBI)의 필요성

그래픽 사용자 인터페이스로(GUI)를 기반으로 직관적인 데이터 분석 즉, **코딩 없이도**(또는 최소한의 코딩 기술로) **데이터 분석이 가능한** 셀프 서비스 비즈니스 인텔리전스(SSBI)의 필요가 증가하게 됨.



2. 분석 솔루션 현황

3) 셀프 서비스 비즈니스 인텔리전스(SSBI)의 한계

현재 시중에서 판매되는 SSBI 솔루션은 다음과 같은 특징/한계를 가지고 있음:

- 어플리케이션 설치 필요
- 따라서, **특정 제품에 Lock-in** → 공동 작업이 어려움 (호환성 떨어짐)
- 자동화, 시각화 등 성능 고도화에 집중 됨
- No-code 솔루션으로 code를 확인 할 수 없음



2. 분석 솔루션 현황

4) Low-Code와 No-code의 차이

Low-Code

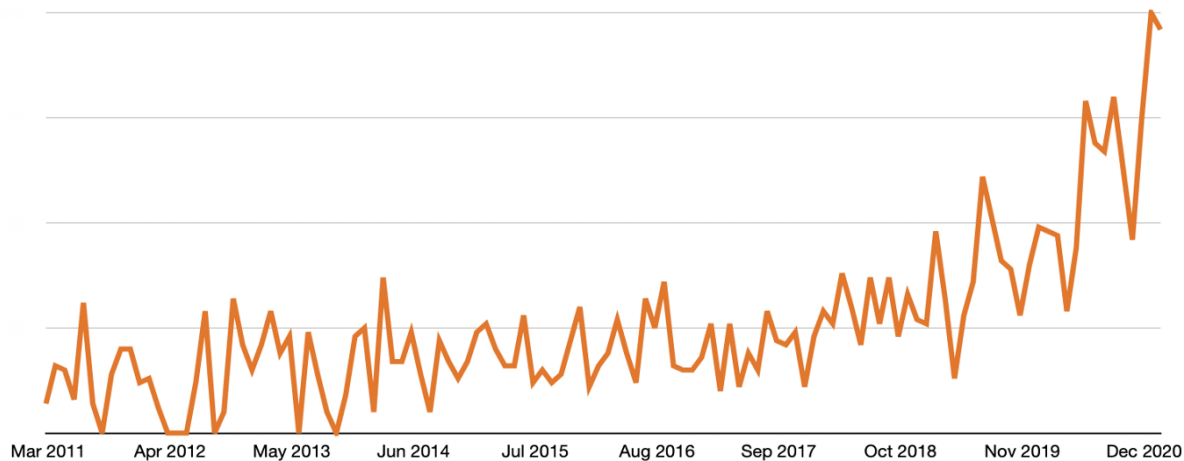
- 사용자가 **부분적으로 코드를 수정할 수 있음**
- 기본적인 코딩 개념이 필요하지만 데이터 처리 속도가 빠름
- 사용자 기능 커스터마이징이 가능

No-Code

- 코드가 드러나지 않고 **수정할 수 없음**
- 코딩이 필요 없어 사용이 쉽지만 데이터 처리 속도가 느림
- 미리 만들어진 기능에 제한됨



Google Searches for "no code low code", Global

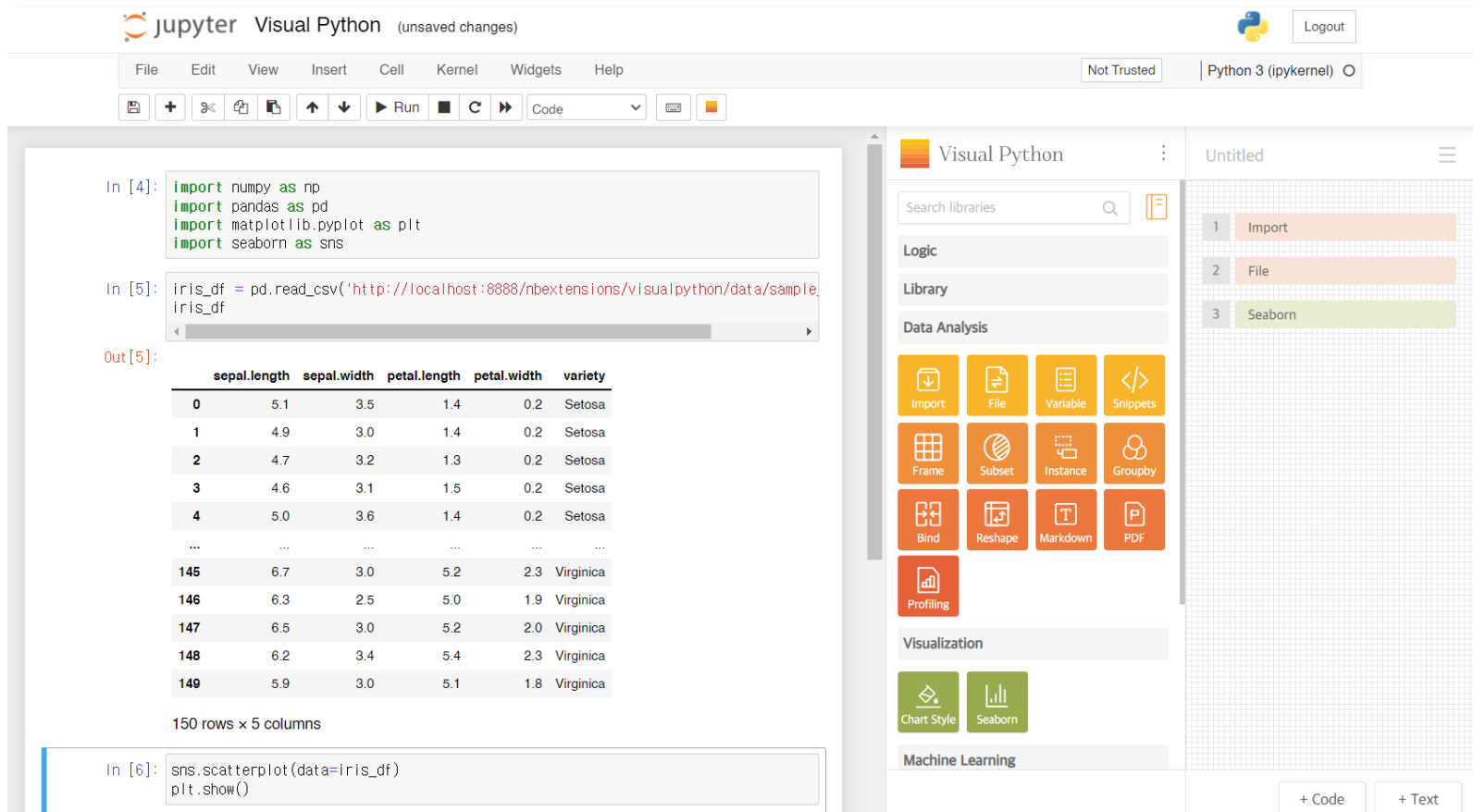


3. Visual Python 소개

3. Visual Python 소개

1) 솔루션 개요

Visual Python 은 빅데이터 분석 분야에서 가장 많이 사용 되는 파이썬(Python) 프로그래밍 언어의 code를 그래픽 사용자 인터페이스 기반으로 생성해주는 GUI-based Code Generator.



The screenshot displays the Visual Python Jupyter environment. The main workspace shows two input cells and their output. The first cell imports necessary libraries, and the second cell reads a CSV file. The output of the second cell is a table of Iris dataset data.

Code Input [4]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Code Input [5]:

```
iris_df = pd.read_csv('http://localhost:8888/nbextensions/visualpython/data/sample_iris.csv')
iris_df
```

Output [5]:

	sepal.length	sepal.width	petal.length	petal.width	variety
0	5.1	3.5	1.4	0.2	Setosa
1	4.9	3.0	1.4	0.2	Setosa
2	4.7	3.2	1.3	0.2	Setosa
3	4.6	3.1	1.5	0.2	Setosa
4	5.0	3.6	1.4	0.2	Setosa
...
145	6.7	3.0	5.2	2.3	Virginica
146	6.3	2.5	5.0	1.9	Virginica
147	6.5	3.0	5.2	2.0	Virginica
148	6.2	3.4	5.4	2.3	Virginica
149	5.9	3.0	5.1	1.8	Virginica

150 rows x 5 columns

Code Input [6]:

```
sns.scatterplot(data=iris_df)
plt.show()
```

The right sidebar contains a 'Visual Python' panel with a search bar and categorized tool icons: Logic, Library, Data Analysis (Import, File, Variable, Snippets, Frame, Subset, Instance, Groupby, Bind, Reshape, Markdown, PDF, Profiling), Visualization (Chart Style, Seaborn), and Machine Learning. The 'Untitled' workspace on the far right shows a sequence of three steps: 1. Import, 2. File, and 3. Seaborn.

3. Visual Python 소개

2) 개발 배경 및 목적

Problem Definition

SSBI 솔루션을 사용한다

- 대부분 유료, 높은 금액대
- 타 도구와 호환 불가 (특정 제품에 Lock-in)
- 자유도가 낮고 제한적인 기능

프로그래밍 언어를 배운다

- 비개발자에게 높은 기술 진입장벽
- 전문적인 교육이 필요
- 개발자도 모든 코드를 외우지 못함



Solution

Jupyter Notebook 기반 (기존 환경을 활용)

무료 오픈 소스 (누구나 무료로 사용 가능한)

Low-code SSBI (최소한의 코딩 개념으로 쉽고 빠르게 분석)

※ Python은 데이터 사이언스, 데이터 프로세싱, 웹 & 비즈니스에서 가장 많이 사용 & 선호되는 프로그래밍 언어

3. Visual Python 소개

3) 핵심 기능 및 활용 분야

Key Features

그래픽 UI를 기반으로 Python code 자동 생성

데이터 분석 및 머신 러닝 중심 Python 기능 구현

데이터 분석 프로세스 저장 및 공유

Application



데이터 분석 교육

Visual Python 활용한
AI·빅데이터 교육

- 중고등학교 및 대학 교육
- 기술 소외계층 교육
- 기업 전문 인력 양성 교육



기업 솔루션

비즈니스 데이터 분석
솔루션으로 활용

- 고객 분석, 상품 개발&관리
- 키워드 및 소셜데이터 분석
마케팅 등



개인 사용

연구 과제 및 논문 데이터
분석에 활용

- 데이터 전처리 및 시각화
- 통계 분석

3. Visual Python 소개

4) 개발 현황

Visual Python 은 현재 데이터 전처리(Pre-processing) 및 시각화(Visualization)를 위한 기능을 갖추고 있으며, 2022년 내에 딥러닝(Deep Learning)과 통계 분석(Statistical analysis)을 위한 기능 등이 추가될 예정.

기본 라이브러리	데이터 전처리	시각화	머신러닝
Python	DataFrame Editor	Pandas Plot	Regression
NumPy	Subset	Matplotlib	Classification
Pandas	Groupby	Seaborn	Clustering
	Combine & Reshape		Dimension Reduction

추가 예정

딥러닝	통계 분석	데이터 수집/관리
텍스트 분석	Regression	SQL
이미지 분석	Correlation	Open API
	Bayesian	Crawling
	Hypothesis Testing	

3. Visual Python 소개

5) 솔루션 차별점

시중에는 Visual Python과 유사한 개념의 소프트웨어가 이미 판매되고 있으나 Visual Python은 그들과 차별되는 몇 가지 특징을 가짐.

bamboolib

Mito 

 Visual Python

구분	Bamboolib	Mito	Visual Python
주요기능	GUI Python for Data Analysis	GUI Python for Data Analysis	GUI Python for data science
실행환경	Jupyter Notebook/Lab	Jupyter Notebook	Jupyter Notebook / Google Colab*
실행방식	Jupyter 위젯 (셀 내에서 작업)	Jupyter 위젯 (셀 내에서 작업)	Jupyter 익스텐션 (Notebook 전체 활용)
사용범위	데이터 분석	데이터 분석	데이터 분석 + 머신러닝 & 딥러닝
결과물	코드	코드	코드 + UI 옵션 값 + 프로세스
가 격	무료 유료 (Databricks에서 사용 가능)	무료 유료 (\$10/month)	무료
시작연도	2019	2020	2020

*(적용 예정)

3. Visual Python 소개

5) 솔루션 차별점

Bamboolib

```
In [3]: df
# bamboolib code export - don't write below - rerun cell to execute code
split_df = df["Name"].str.split(',', expand=True)
split_df.columns = [f"Name_{id}" for id in range(len(split_df.columns))]
df = pd.merge(df, split_df, how="left", left_index=True, right_index=True)
split_df = df["Name_1"].str.split(',', expand=True)
split_df.columns = [f"Name_1_{id}" for id in range(len(split_df.columns))]
df = pd.merge(df, split_df, how="left", left_index=True, right_index=True)
df = df.rename(columns={'Name_1_1': 'Title'})
```

Show static HTML

History Export ☒ Live Code Export

Add Python Code Select/Drop columns Filter Sort

New column formula Aggregate/Group by Join Replace value

Set/Update values OneHotEncoder Pivot Unpivot/Melt

Visualize DataFrame

891 rows × 26 columns

i	SibSp	i	Parch	o	Ticket	f	Fare	o	Cab
1	0	0	A/5 21171	25	nan	^			
1	0	0	PC 17599	12833	C85				
0	0	0	STON/O2 31012...	525	nan				

Transformations history

Manipulate strings of 'Name' and perform a split on ','

Manipulate strings of 'Name_1' and perform a split on ','

Rename column Name_1_1 to Title

Undo last step Recover last step

- Jupyter Notebook 셀 내에 포함된 UI
- 1개의 데이터를 개별로 처리
- 모든 코드가 1개의 셀에 누적
- Pandas(데이터 분석) 기능에 국한

Visual Python

jupyter Visual Python Tutorial Last Checkpoint 2시간 전 (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Not Trusted Python 3

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

In [2]: iris_df = pd.read_csv('http://localhost:8888/notebooks/visualpython/data/iris_iris.csv')

Out[2]:
```

	sepal.length	sepal.width	petal.length	petal.width	variety
0	5.1	3.5	1.4	0.2	Setosa
1	4.9	3.0	1.4	0.2	Setosa
2	4.7	3.2	1.3	0.2	Setosa
3	4.6	3.1	1.5	0.2	Setosa
4	5.0	3.6	1.4	0.2	Setosa
5	5.4	3.9	1.7	0.4	Setosa
6	4.6	3.4	1.4	0.3	Setosa
7	5.0	3.4	1.5	0.2	Setosa
8	4.4	2.9	1.4	0.2	Setosa
9	4.9	3.1	1.5	0.1	Setosa
10	5.4	3.7	1.5	0.2	Setosa

```
In [1]: from sklearn.datasets import load_boston

ldata = load_boston()
# Create DataFrame
df_ldata = pd.DataFrame(data=ldata.data, columns=ldata.feature_names)
df_ldata['target'] = ldata.target
df_ldata
```

Visual Python

Search Variables

Data Analysis

Import File del ID for ids in range() joins return

Machine Learning

Classifer Clustering Dimensionality Reduction AutoML

+ Code + Text

- Jupyter Notebook과 분리된 UI
- 여러 데이터를 통합적으로 처리
- 기능 별로 구분된 코드 및 셀 생성
- 데이터 분석 + 머신러닝 + 딥러닝

4. Visual Python 현황

4. Visual Python 현황

1) 사용자 현황

Visual Python은 2021년 1월 GitHub에 공식 오픈소스 프로젝트로 공개되었으며 (2023년 5월 기준) 누적 124K 다운로드를 기록하였으며, 월 평균 5K 회 다운로드 중

visualpython

Summary

PyPI link

<https://pypi.org/project/visualpython>

Total downloads

124,723

Total downloads - 30 days

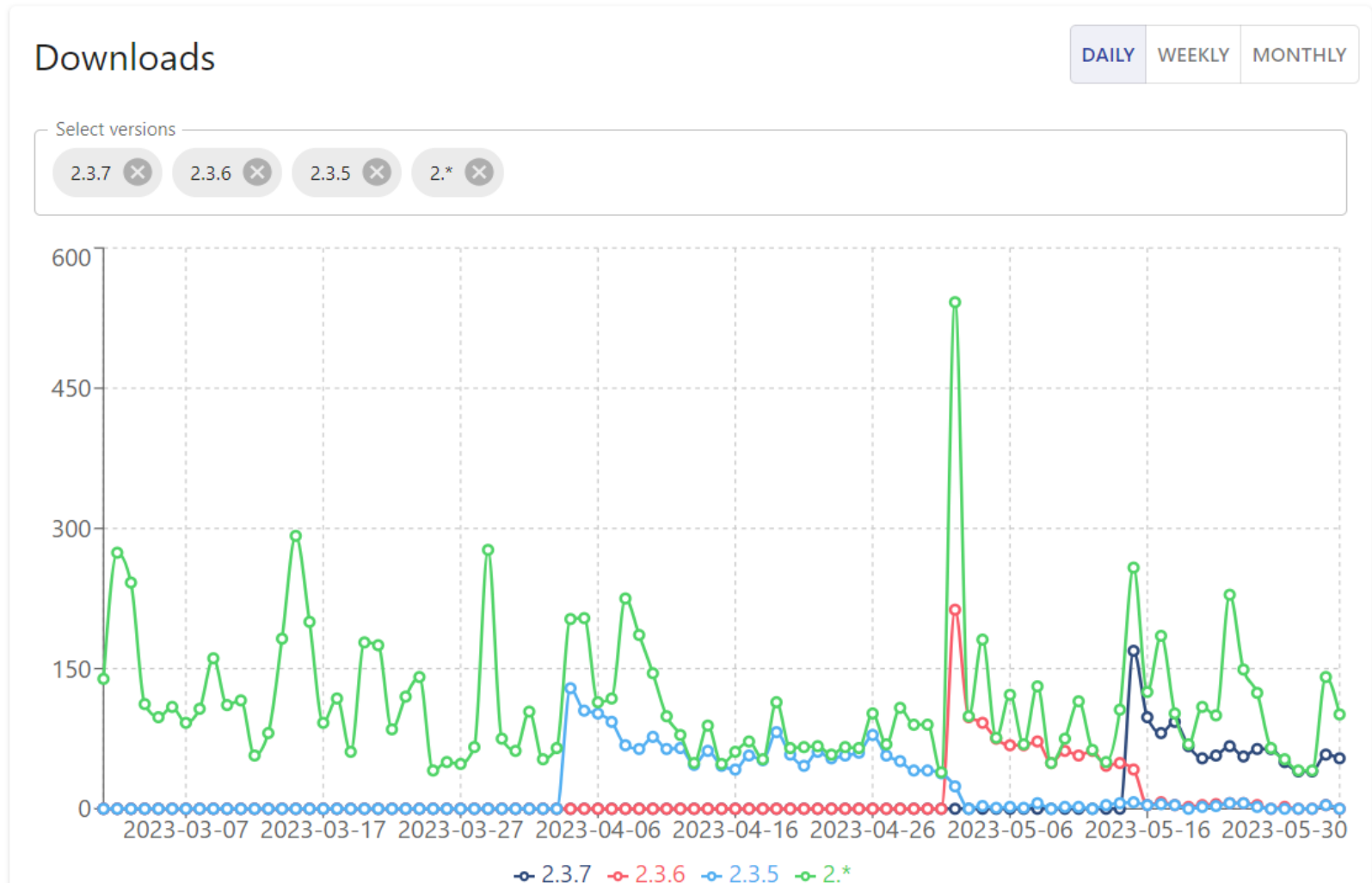
5,615

Total downloads - 7 days

851

4. Visual Python 현황

1) 사용자 현황



4. Visual Python 현황

2) 외부 소개 현황

해외의 블로그 및 다양한 web page에서 Visual Python을 소개하고 있음

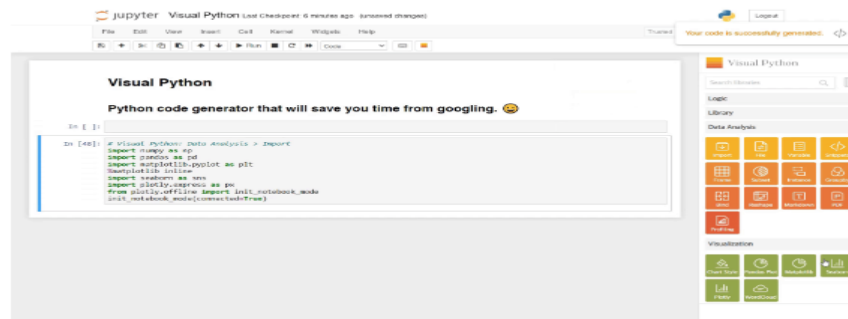
visualpython



Ismailouahbi · Follow
3 min read · Jan 28

81

1



Visual Python 2.2.8.gif (1742x980) (raw.githubusercontent.com)

It's time to discover **visualpython** a GUI for easy dealing with repetitive data science tasks with python in addition to its powerful support for various python data science libraries(pandas, seaborn..etc)

A step-by-step guide:

Try Visual Python if you would like to:

manage big data with minimal coding skills.

help students / business analysts / researchers to overcome learning barriers for

4. Visual Python 현황

3) 커뮤니티 채널 현황

Visual Python은 오픈소스 프로젝트로 사용자와 기여자의 적극적인 참여가 중요. 따라서, 사용자 확대 및 네트워크 형성을 위해 GitHub, YouTube, Twitter, LinkedIn 등 다양한 커뮤니티 채널을 운영하고 있음.

주요 커뮤니티 채널:

Visual Python Hub

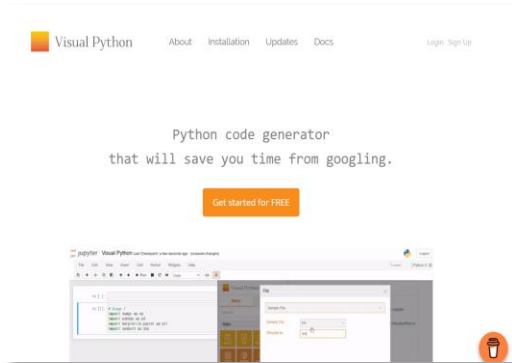
- Visual Python과 관련된 기본 정보
- VP Note, Snippets 등 자료 제공

GitHub

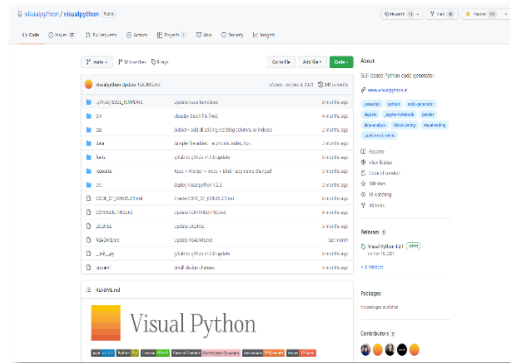
- 버전 업데이트 및 소스 공개
- 버그 및 기타 문의사항에 대하여 소통

YouTube

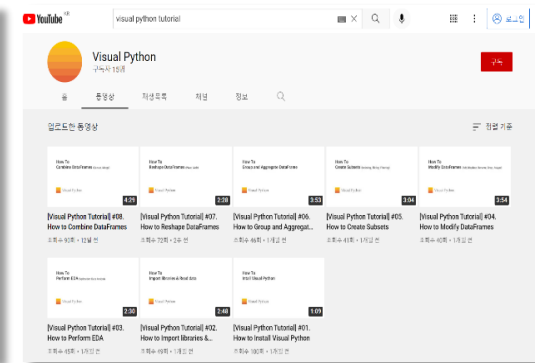
- Visual Python Tutorial 영상
- 사용 방법 및 TIP 제공



[Visit Visual Python Hub](#)



[Visit GitHub: Visual Python](#)



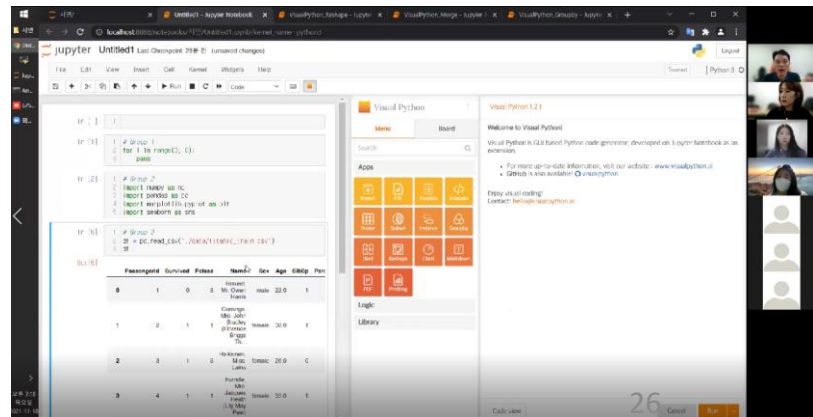
[Visit YouTube: Visual Python](#)

4. Visual Python 현황

3) 활용 사례

Visual Python은 현재 여러 기업 및 기관에서 데이터 분석 및 머신 러닝 교육을 위한 도구로 사용되고 있으며, 청년들의 오픈소스 소프트웨어 개발에 참여를 돕기 위한 프로그램 2022 오픈소스 컨트리뷰션 아카데미의 사례로 선정된 바 있음.

주요 교육 대상 기업/기관:



End of Document