

Python을 활용한 언어학 통계 입문

제1강 Python 시작하기

제1강 Python 시작하기

박수민

서울대학교 인문대학 연계전공 인문데이터과학 / 언어학과

2025년 8월 11일 월요일



강좌 소개

기분 정보
일정

파이썬 맛보기

- 파이썬 소개
- 개발 환경
 - 텍스트 편집기
 - Google Colaboratory
- 기초 개념

마무리



강사

이름 박수민

소속 서울대학교 인문대학 연계전공 인문데이터과학 / 언어학과

전공 컴퓨터언어학

연락처 <mailto:mam3b@snu.ac.kr>
<https://github.com/suparklingmin>

수강생

- 1 컴퓨터에서 웹 브라우저를 사용할 수 있다.
- 2 한국어 및 영어로 된 글을 읽고 이해할 수 있다.
- 3 언어학 전공 과목을 수강했거나 관련 논문을 읽어 본 적이 있다.
- 4 개수나 횟수를 세거나 척도의 값을 측정한 적이 있다.

웹 브라우저 Google Chrome, Safari, ...

박수민

강좌 소개

기본 정보

일정

파이썬 맛보기

파이썬 소개

개발 환경

텍스트 편집기

기초 개념

마무리



2025 LSK-시몬느
언어학 학교

심주희. (2023). 〈식당 내 호칭 ‘저기요’, ‘여기요’ 를 대상으로 한 사회언어학적 변이 연구〉. 《담화와인지》 30(1), 64.

<표 7> 성별과 응답 간 상관관계

	빈도(%)		전체
성별	저기요	여기요	
여성	1038(46.1%)	1213(53.9%)	2251(100%)
남성	799(54.3%)	673(45.7%)	1472(100.0%)
$\chi^2(p)$	23.750(0.000)*		

$$p^* < 0.05, p^{**} < 0.01, p^{***} < 0.001$$

(N=3723)



연구 사례: 척도의 값을 측정하는 경우

김수한·문성민·고연숙. (2022). <이름의 성별에 따라 나타나는 음성상징성: 동물의 숲 주민 캐릭터 이름을 중심으로>. 《음성음운형태론연구》 28(1), 46.

첫 번째로 각 이름에서 포함하는 전체 자음과 모음의 개수를 나타내는 분절음과 홀로 발음할 수 있는 최소의 소리 단위를 나타내는 음절에 대한 분석 결과, 성별 간에 음절 수는 차이가 있으나, 성별 간의 분절음 수는 차이가 없는 것으로 나타났다. 그 이유는 선행연구인 Slater and Feinman (1985)에서 사용한 영어의 음절구조와 달리, 1.2 절에서 언급한 한국어의 음절구조의 경우 하나의 모음을 기준으로 두음과 말음에 최대 1 개의 자음만 위치할 수 있다. 그래서 분절음 수가 선행연구의 결과와 다른 결과가 나온 것으로 보인다.

표 7. 이름에 들어간 음절 수와 분절음 수에 따른 성별 간의 차이의 t 검정 결과

	등분산성	t	df	p-value
음절 수	등분산	-2.570	390	<.05 *
분절음 수	등분산	-0.290	390	.772

<표 8>은 성별에 따른 음절 수와 분절음 수의 평균 값을 나타내고 있으며 이를 통해 여성 캐릭터의 음절 수와 분절음 수의 평균이 남성 캐릭터 보다 더 높기 때문에 결과에 차이가 난 것을 확인할 수 있다.

표 8. 성별에 따른 음절 수와 분절음 수의 평균

	남성	여성
음절 수	2.50	2.71
전체 분절음 수	5.53	5.58

3.2.2 공명음, 파열음, 마찰음, 파찰음, 모음의 비율

이 강좌의 기대 효과

이 논문에 나온 말을 이해할 수 있다.



- Python을 클라우드 환경에서 사용하기: Google Colaboratory
- Python을 계산기로 써먹기: 산술 연산

- 데이터프레임 잘 만들기: “tidy” data vs. “messy” data
- 데이터의 유형 이해하기: 수치형(이산형, 연속형), 범주형(명목형, 순서형)

- 기술통계량: 평균, 표준편차, 중앙값, 범위, 상관관계 등
- 시각화: 산점도, 히스토그램, 상자그림 등



- 이론적 배경: 현실의 모델링(모형화), 반응변수(종속변수)와 설명변수(독립변수)
- 선형회귀분석: 반응변수가 연속형 자료인 경우의 모델링

- 로지스틱 회귀분석: 반응변수가 명목형 자료인 경우(이진 분류)의 모델링
- 푸아송 회귀분석: 반응변수가 이산형 자료인 경우(0 이상의 횟수)의 모델링

- 혼합 효과 모형: 반복 측정 데이터에서 개체 간 변이를 고려하는 모델링
- 모델 선택

강좌 소개

기본 정보

일정

파이썬 맛보기

파이썬 소개

개발 환경

텍스트 편집기

기초 개념

마무리



2025 LSK-시몬느
언어학 학교

7 통계적 추론 (1) (2025-08-14)

- 이론적 배경: 가설 검정, 1종 오류, 2종 오류, 통계적 유의미
- t -검정: 두 집단의 평균 차이가 통계적으로 유의미한지 알아보기

8 통계적 추론 (2) (2025-08-14)

- 카이제곱검정: 범주형 데이터에서 두 변수의 관계가 통계적으로 유의미한지 알아보기
- 분산분석: 세 집단 이상에서 평균 차이가 통계적으로 유의미한지 알아보기



파이썬 공식이 말하는 파이썬

박수민

기본 정보
일정

파이썬 소개

개발 환경

텍스트 편집기

기초 개념

마무리

<https://www.python.org>



2025 LSK-시몬느
언어학 학교

C: 총 6행; Java: 총 5행

박수민

일정

파이썬 소개

개발 환경

텍스트 편집기

기초 개념

마무리

```
1 #include<stdio.h>
2 int main(int argc, char * argv[])
3 {
4     printf("Hello, World!\n");
5     return 0;
6 }
```

```
1 public class HelloWorld {
2     public static void main(String[] args) {
3         System.out.println("Hello, World!");
4     }
5 }
```



2025 LSK-시몬느
언어학 학교

파이썬 공식이 말하는 파이썬

“파이썬은 인터프리터 언어입니다. 컴파일과 링크 단계가 필요 없으므로 개발 시간을 상당히 단축해줍니다. 인터프리터는 대화형으로 사용할 수 있어서, 언어의 기능을 실험하거나, 쓰고 버릴 프로그램을 만들거나, 바닥부터 프로그램을 만들어가는 동안 함수들을 테스트하기 쉽습니다. **간편한 탁상용 계산기**이기도 합니다.”

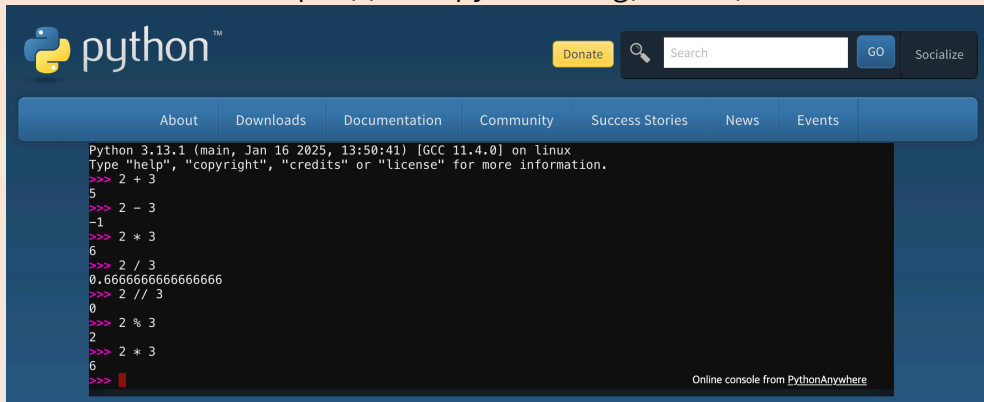
<https://docs.python.org/ko/3/tutorial/appetite.html>



파이썬 대화형 인터프리터

파이썬을 계산기처럼 사용하기

<https://www.python.org/shell/>



python™

Donate Search GO Socialize

About Downloads Documentation Community Success Stories News Events

```
Python 3.13.1 (main, Jan 16 2025, 13:50:41) [GCC 11.4.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> 2 + 3
5
>>> 2 - 3
-1
>>> 2 * 3
6
>>> 2 / 3
0.6666666666666666
>>> 2 // 3
0
>>> 2 % 3
2
>>> 2 * 3
6
>>>
```

Online console from [PythonAnywhere](#)

Python 시작하기

박수민

강좌 소개

기본 정보

일정

파이썬 맛보기

파이썬 소개

개발 환경

텍스트 편집기

Google Colaboratory

기초 개념

마무리



2025 LSK-시몬느
언어학 학교

문제

우리는 파이썬 코드를 어디에서 어떻게 작성해야 하는가?



파이썬 코드를 작성하는 “앱”이 따로 있나요?

문서를 다루는 응용프로그램의 종류

- 워드 프로세서
 - 한글, MS Word 등
- 텍스트 편집기
 - EmEditor, Sublime Text, Visual Studio Code 등

텍스트 편집기의 특징

- 문자(글자, 숫자, 문장부호, 특수 문자 포함)만 사용할 수 있다.
 - 워드 프로세서에서는 글자와 문단 모양을 바꾸고 도표를 삽입할 수 있다.
- 텍스트 파일은 모든 텍스트 에디터에서 읽고 쓸 수 있다.
 - 한 워드 프로세서 프로그램에서 작성한 문서는 다른 프로그램에서 열기 어렵다.

추천 Visual Studio Code(<https://code.visualstudio.com/>)



파이썬을 클라우드 환경에서 사용하기

Colab: Google Colaboratory

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

준비물 Google 계정(무료)

Python 시작하기

박수민

강좌 소개

기본 정보

일정

파이썬 맛보기

파이썬 소개

개발 환경

텍스트 편집기

Google Colaboratory

기초 개념

마무리



2025 LSK-시몬느
언어학 학교

파이썬을 클라우드 환경에서 사용하기

Python 시작하기

박수민

강좌 소개

기본 정보

일정

파이썬 맛보기

파이썬 소개

개발 환경

텍스트 편집기

Google Colaboratory

기초 개념

마무리

Colab 시작하기

설정 공유 Gemini

노트 열기

- 예 >
- 최근 사용 >
- Google Drive >
- GitHub >
- 업로드 >

노트북 검색

제목

- [Overview of Colab Features](#)
- [Markdown Guide](#)
- [Charts in Colab](#)

+ 새 노트

취소



2025 LSK-시몬느
언어학 학교

지금 바로 Colab에서 할 일

- 파이썬과 인사하기
- 파이썬을 계산기처럼 사용하기

The screenshot shows the Google Colab web interface. At the top, there's a header with the Colab logo, a title "[2025 언어학 학교] 1강 Python 시작하기", and various icons for file management, settings, and sharing. Below the header, there's a toolbar with options like "명령어", "코드", "텍스트", and "모두 실행". The main workspace contains two code cells. The first cell has the code `[1] 1 print('Hello, World!')` and its output is `Hello, World!`. The second cell has the code `1 print(2 + 3)` and its output is `5`. On the right side of the workspace, there's a status bar showing RAM and disk usage.



2025

LSK-시몬느 언어학 학교

2025 LSK-Simone Linguistic School

