

Python 데이터 분석 02

2021-2 KCA

김지환

데이터 분석에 자료구조/알고리즘이 필요할까?

더 찾아봤더니...

네이버, 카카오 등 유명 IT 기업들은 다른 직군과 마찬가지로 코딩테스트를 본다!

토스는 직무 인터뷰(사전 쿼리 테스트 포함) 라 하는데 무엇일까?

그러나 다른 기업들은 코테를 안 보는 경우, 경험을 중시하는 경우, 자격증 (데이터 분석가..)을 보는 경우 등 다양했다.

단, Python, R, Spark, SAS 다뤄본 경험 필수적으로 요구.

제 경험에 관하여..

- 100GB 가량의 빅데이터에서 원하는 행들을 뽑는 코드를 짤 때. (Search 속도)
이분 탐색 + 빠른 파일 입출력 => 속도를 거의 10배 가량 상승

⇒ 그런데 일은 컴퓨터가 한다. + 데이터 분석은 코드 최적화보다 무슨 의미를 찾는냐가 더 중요!

⇒ 코드 재사용이 적다면, 최적화 효율 떨어짐.

실제로는 잘 안 쓰이더라도, 배우는 것은 의미 있을 것 같다.
(자소서로든, 코테 대비로든...)

Google Colaboratory!



google 계정으로 무료 사용.
GPU 사용 가능.
jupyter notebook의 온라인 버전.

기본 라이브러리 엄청 많이 제공!!! => 환경 구축할 필요 없음!

1. Colab 검색
2. 접속
3. 파일 -> 새 노트

라이브러리(패키지) 추가 방법

- 파이썬은 pip (패키지 관리자) 존재!!
 - 간편하게 설치, 제거, 관리 가능!
 - 패키지 > 모듈 > 클래스, 함수, 변수
-

```
import 모듈  
...  
모듈.함수()
```

```
from 모듈 import 변수, 함수, 클래스  
...  
함수()
```

```
import 패키지.모듈  
...  
패키지.모듈.함수()
```

```
import 모듈 as 이름  
...  
이름.함수()
```

```
from 모듈 import 함수 as 이름  
...  
이름()
```

```
from 패키지.모듈 import 변수, 함수, 클래스  
...  
함수()
```

참고 : <https://dojang.io/mod/page/view.php?id=2441>
점프 투 파이썬 & 파이썬 코딩 도장 추천!!

[실습] 라이브러리(패키지) 추가 실습

- math 모듈 : 수학 관련

```
import math
```

```
...
```

```
math.pi
```

```
math.sqrt(4.0)
```

```
import 모듈 as 이름
```

```
...
```

```
이름.함수()
```

```
from 모듈 import 변수, 함수, 클래스
```

```
...
```

```
함수()
```

```
from 모듈 import 함수 as 이름
```

```
...
```

```
이름()
```

```
import 패키지.모듈
```

```
...
```

```
패키지.모듈.함수()
```

```
from 패키지.모듈 import 변수, 함수, 클래스
```

```
...
```

```
함수()
```