

Jin-Soo Kim
(jinsoo.kim@snu.ac.kr)

Systems Software &
Architecture Lab.

Seoul National University

Jan. 6 – 17, 2020

Python for Data Analytics

BeautifulSoup



Lab 1. 네이버 실시간 검색어

- <https://datalab.naver.com/> 의 실시간 검색어를 추출
- `def get_rank()`의 반환값인 `rank`는 아래와 같이 출력
 - 날짜별로 검색어의 순위와, 검색어의 정보를 추출

출력 예시 :

...

```
['2020.01.01.(수)', {'1': '리버클레시',  
'2': '아비쏘구스다운', '3':  
'내셔널지오그래픽숏패딩', '4': 'mlb패딩',  
'5': '다이나핏패딩', '6':  
'내셔널지오그래픽롱패딩', '7':  
'로맨틱크라운패딩', '8': '원피스', '9':  
'롱패딩', '10': '여성숏패딩'}]...
```

쇼핑인사이드 쇼핑 분야별 클릭 추이와 분야별 검색어 현황을 확인할 수 있습니다.

분야별 인기 검색어 인기분야

패션의류 일간

2019.12.29.(일)	2019.12.30.(월)	2019.12.31.(화)	2020.01.01.(수)
1 리버클레시	1 리버클레시	1 리버클레시	1 리버클레시
2 아비쏘	2 아비쏘	2 아비쏘구스다운	2 아비쏘구스다운
3 팬필드후삭	3 내셔널지오그래픽숏패딩	3 mlb패딩	3 내셔널지오그래픽숏패딩
4 하프크라이즈패딩	4 mlb패딩	4 내셔널지오그래픽숏패딩	4 mlb패딩
5 내셔널지오그래픽숏패딩	5 다이나핏패딩	5 다이나핏패딩	5 다이나핏패딩
6 mlb패딩	6 로맨틱크라운패딩	6 내셔널지오그래픽롱패딩	6 내셔널지오그래픽롱패딩
7 로맨틱크라운패딩	7 원피스	7 로맨틱크라운패딩	7 로맨틱크라운패딩
8 다이나핏패딩	8 여성숏패딩	8 원피스	8 원피스
9 여성숏패딩	9 내셔널지오그래픽롱패딩	9 롱패딩	9 롱패딩
10 원피스	10 헛드메이드코트	10 여성숏패딩	10 여성숏패딩

검색어 통계 자세히 보기 >

Lab 1. 네이버 실시간 검색어

- 결과 값은 ranking data는 각 순위 별로 dictionary 형태로 저장
 - [data , { '1' : 'data', '2' : 'data', '3' : 'data' , ... }] ← just example
 - data의 형태는 관계없음, html parsing해서 원하는 정보를 가져오는 것이 목적
 - Hint : find_all() , get_text(), list comprehension(of for loop)
 - Skeleton code 제공
 - get_rank() return : 얻은 데이터