



한국 IT 교육원
이 두 화



2021-07-30

그로링으로 수집한

전국코로나현황 데이터와

Tableau를 활용한 시각화

1. 개요

2. 데이터 수집

3. 시각화

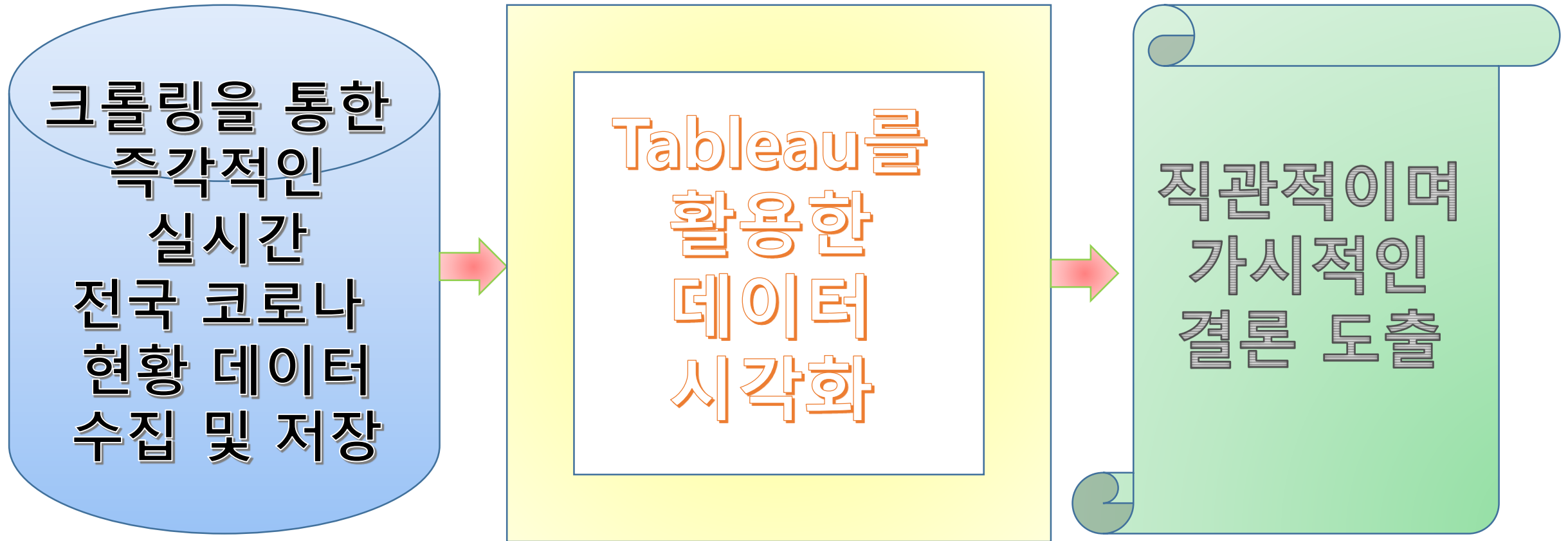
Q & A

4. 결론

A field of purple flowers, possibly lavender, with a bright sun in the background creating a bokeh effect. The text "1. 개요" is overlaid in the center.

1. 개요

1. 개요



수집된 데이터 정의

데이터명	의미	데이터 타입
지역	지역명	STR
누적확진자	전체 누적확진자 수	INT
전일확진자	전일 확진자 수	INT
위도	해당지역의 위도정보	float
경도	해당지역의 경도정보	float

파이썬을 활용한 데이터의 수집과정

```
71 코로나현황합계 = pd.DataFrame(코로나현황합계)
72 코로나현황1 = 코로나현황.append(코로나현황합계)
73 코로나현황2 = 코로나현황1.reset_index(drop=True, inplace=False) # 인덱스 재설정
74 좌표 = 위도_경도()
75 전국코로나현황 = pd.concat([코로나현황2, 좌표 ], axis=1)
76 전국코로나현황.to_excel('./xlsx/전국코로나현황.xlsx')
77 print(전국코로나현황)
78
79
80
81 if __name__ == '__main__':
82     코로나()
83     위도_경도()
84
85
86 # 데이터 프레임을 데이터베이스에 넣기
```

PROBLEMS

OUTPUT

터미널

디버그 콘솔

```
op\LDH\gitblog\py\포트폴리오\코로나전국현황_크롤링_코드.py
지역    누적확진자    전일확진자    위도    경도
0 서울    63883    487    37.5666    127.0780
1 경기    55386    522    37.4376    127.5172
2 대구    11623    62    35.8722    128.6025
3 인천    8964    119    37.4563    126.7052
4 부산    8134    78    35.1794    129.0755
5 경남    7105    99    35.4677    128.2147
6 경북    5437    37    36.3346    128.8789
7 충남    4883    37    36.6990    126.8004
8 강원    4454    20    37.8710    128.1657
9 대전    4103    85    36.3511    127.6850
10 충북    3831    37    37.0243    127.7010
11 광주    3378    25    35.1597    126.8530
12 울산    3189    18    35.5388    129.3166
13 전북    2753    23    35.7459    127.1530
14 전남    2022    19    34.8824    126.9929
15 제주    1706    18    33.5000    126.5166
16 세종    724    4    36.4875    127.2816
17 합계    191575    1690    36.3750    128.0230
```



	A	B	C	D	E	F
1		지역	누적확진자	전일확진자	위도	경도
2	0	서울	63883	487	37.5666	127.0780
3	1	경기	55386	522	37.4376	127.5172
4	2	대구	11623	62	35.8722	128.6025
5	3	인천	8964	119	37.4563	126.7052
6	4	부산	8134	78	35.1794	129.0755
7	5	경남	7105	99	35.4677	128.2147
8	6	경북	5437	37	36.3346	128.8789
9	7	충남	4883	37	36.6990	126.8004
10	8	강원	4454	20	37.8710	128.1657
11	9	대전	4103	85	36.3511	127.6850
12	10	충북	3831	37	37.0243	127.7010
13	11	광주	3378	25	35.1597	126.8530
14	12	울산	3189	18	35.5388	129.3166
15	13	전북	2753	23	35.7459	127.1530
16	14	전남	2022	19	34.8824	126.9929
17	15	제주	1706	18	33.5000	126.5166
18	16	세종	724	4	36.4875	127.2816

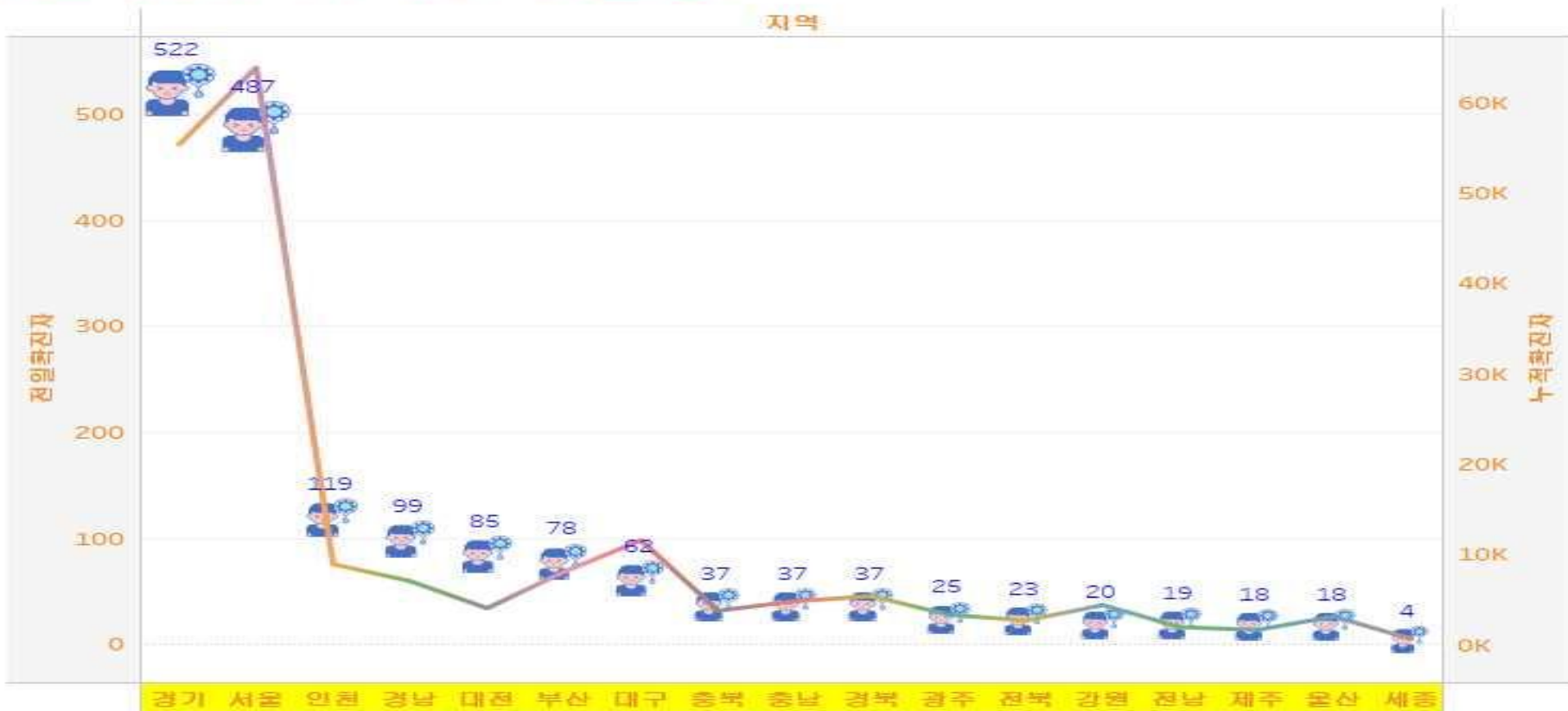
크롤링한
데이터가
엑셀파일로
저장된 모습



지역별로
확진자 현황을
파악하기 쉽도록
맵형식으로
구성한 차트
(확진자 수에 따라
아이콘의
크기가 달라짐)



전일확진자와 누적확진자 비교 그래프



전일확진자와 누적확진자를
비교해 볼 수 있도록 해놓은 그래프차트
(평소보다 확진자수가 늘었는지 줄었는지 파악이 쉬움)

필터를 사용하면 특정지역만 볼 수 있음

지역별 필터(전체보기는 우측에 흰색 공백부분 클릭)

강원 경기 경남 경북 광주 대구 대전 부산 서울 세종 울산 인천 전남 전북 제주 충남 충북

전일확진자와 누적확진자 비교 그래프





결론

크롤링을 통한 실시간 데이터 수집과
오라클 데이터베이스와 연계하여
수집된 데이터를 데이터베이스화 하고
Tableau를 사용한 시각화를 통해
보다 직관적이고 실용적인
결과물을 도출할수 있다.





블로그

Tableau 스토리보드

오라클 데이터베이스



QnA