

데이터 다루기 미니 프로젝트

❖ 필요 데이터

- 서울 열린데이터 광장(<https://data.seoul.go.kr/index.do>)
 - 자치구 년도별 CCTV 설치 현황
 - 서울통계서비스 > 주민등록인구(구별) 통계

❖ CCTV 데이터 정리하기

- 데이터 파일 읽기
- Null data 있는지 확인하기
- 1000 단위 구분기호 없애고 정수로 변환하기
- 2021년을 제외한 최근 3개년간의 CCTV 증가율 구하기
- 구분, 총계, 최근증가율 컬럼만 추출하고 컬럼명을 구별, CCTV댓수로 변경하기
- 구 이름에서 공백지우기
- 첫번째 행(계) 지우기

데이터 다루기 미니 프로젝트

❖ 인구 데이터 정리하기

- 데이터 파일 읽기
- Null data 있는지 확인하기
- '자치구','계','계.1','계.2','65세이상고령자' 컬럼만 추출하고 이름을 '구별','인구수','내국인','외국인','고령자'로 변경
- 1000 단위 구분기호 없애고 정수로 변환하기
- 첫번째 행 지우기
- '외국인비율'과 '고령자비율' 컬럼 만들기

❖ CCTV 데이터와 인구 데이터 합치고 저장하기

- 데이터 합치기
- '구별' 컬럼을 인덱스로 만들기
- 합친 데이터를 cctv_pop.csv로 저장하기