

데이터 전처리를 위한 Python Quiz

주어진 날짜별 코스피 상장사 주가 데이터(kospi.csv)에서 아래의 문항에 맞는 코드를 작성하세요.

(1) kospi_df 의 변수 이름으로 csv 파일을 데이터 프레임으로 불러오세요.

(2) 파일에 있는 유니크한 날짜수(dt)와 기업수(name)를 출력하세요.

(3) 2019년 12월 08일 이후의 날짜별 평균 상승률을 소수 셋째자리까지 리스트의 딕셔너리 형태로 데이터를 출력하세요.

(dt : 날짜, cr : 상승률)

(출력 포맷 예시 : [{'dt': 20191109, 'cr': -1.234}, {'dt': 20191110, 'cr': 5.512}])

(4) 2019년 10월 21일 상승률 상위 5개회사를 출력하세요. (cr : 상승률)

(5) 2019년 11월 30일에서 2019년 12월 5일까지 일별 상승률 상위 5개 회사의 평균 상승율을 소수점 셋째 자리까지 출력하세요. (cr : 상승률)

(6) 전체 기간에서 평균 등락이 높은 3개의 회사이름과 평균 등락 수치를 출력하세요. (cv: 등락)

(7) NAVER와 한국전력의 종가 데이터의 상관계수를 소수둘째자리까지 출력하세요. (ncv: 종가)
상관계수 수식

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \cdot \sum(y-\bar{y})^2}}$$