

## 2025-1 빅데이터분석프로그래밍 중간고사 (20241519 조예성)

시리즈의 속성 5개:

각 값의 횟수를 반환하는 메소드: s.

- 각 값의 구간을 n개로 분할하여 개수를 반환하는 속성:

최소, 최대, 평균, 중간값을 반환하는 메소드: s. (), s. (), s. (), s. ()

행의 이름은?

열의 이름은?

	이름	나이	성별	점수
0	철수	23	남	85
1	영희	21	여	92
2	민수	25	남	78
3	지민	20	여	88
4	수지	22	여	95

위 df에서 나이가 22살 이상인 사람만 선택하는 코드:

데이터프레임의 행이나 컬럼에 인덱스 값으로 접근하는 코드: df.

- 해당 코드의 반환 값은:
- 해당 코드를 수정하여 DataFrame 형식으로 반환을 하고 싶으면: df.

데이터프레임의 행이나 컬럼에 레이블로 접근하는 코드: df.

- 데이터프레임에서 이름이 가인 행에 접근: df.
- 데이터프레임에서 이름이 학번인 행에 접근: df.
- 해당 코드의 반환 값은:

데이터프레임에서 각 열의 데이터 타입을 반환하는 속성: df.

Csv를 읽어오는 코드:

데이터프레임의 속성을 반환하는 메소드: df.

데이터프레임 값을 앞에서부터 n개 반환하는 메소드: df.

데이터프레임 값을 뒤에서부터 n개 반환하는 메소드: df.

데이터프레임 값을 열 기준으로 합산하는 메소드: df.

- 행 기준으로 합산하는 속성:

데이터프레임 열 기준으로 누적합을 구하는 메소드: df.

데이터프레임 평균 구하는 메소드: df.

모든 셀의 평균을 구하는 속성:

데이터프레임에서 질의하는 메소드: df.

- 해당 메소드에서 변수 사용하는 방법:

데이터프레임 정렬하는 명령어: df.

- 내림차순 속성: =
- 특정 열 기준으로 정렬: =
- 여러 개 하려면: =

없는 값은:

결측값 확인하는 메소드: pd. ( ) / . ( ) ||

각 열의 nan 빼고 개수 출력: df.

시리즈를 데이터프레임으로 변경하는 메소드: series.

Na가 있는 행 제거한 df 반환: df.

Na를 특정 값으로 채우는 메소드: df.

- 반환뿐이 아닌 저장까지 하도록 하는 메소드: =
- 뒤가 Nan이면 앞 값을 이어서 채우기: =

5. 다음 코드와 동일한 기능을 메소드 isin()을 사용해 구현하기 위한 빈 부분을 완성하세요.(4)

```
mpg = pd.read_csv('mpg.csv') # mpg 데이터 불러오기
mpg['size'] = np.where((mpg['category'] == 'compact') |
                       (mpg['category'] == '2seater'),
                       'small', 'large')

mpg['size'] = np.where(
    ,
    'small', 'large')
```

다양한 집계연산을 하는 데이터프레임 함수는? df.

그룹화 하는 메소드: df.

데이터프레임 2개를 합치는 명령어: pd. /

- 병합하는 방식을 변경하는 속성: = , , , ,
- 공통 컬럼 지정: = ' '