|  |  |
| --- | --- |
| 교육 제목 | 판다스의 시리즈와 데이터프레임 |
| 교육 일시 | 2021.11.08 |
| 교육 장소 | 영우글로벌러닝 |
| **교육 내용** | |
| 오전 | ◇  판다스 들어가기  Help > pandas > 시작을 하거나 유저 가이드를 볼 수 있다  판다스는 형식적으로 여러 유형의 데이터를  공통으로 포멧하는데 의의가 있다  ◇  시리즈는 동일한 데이터가 들어간다  시리즈는 판다스 시리즈(pd.series)를 넣고 딕셔너리나 리스트를 넣는다  키는 인덱스가 되고 값은 value가 되며 인덱스와 값이 전부 저장된다  인덱스와 인덱스명으로 접근 가능하며  인덱스명으로 접근시 start : end 둘다 포함된다  여러개의 값 출력시 [[인덱스명 또는 인덱스]]  ◇  데이터프레임은 시리즈가 여러개 묶여있으며  행 인덱스와 열 이름을 나뉜다  drop을 이용하여 하나 또는 여러개의 인덱스를 지울 수 있다  ※  inplace=True 옵션을 추가해야 기존 데이터 변경이 가능하다  axis=0 - 기본 변수로 되어있으며 인덱스 삭제에 쓴다 생략이 가능(행)  axis=1- 칼럼 삭제(열)  loc[인덱스, 칼럼명]를 이용하여 특정 인덱스의 검색이 가능하다  정수의 경우 iloc[정수인덱스]를 이용한다  :(콤마)는 전부, 그 이후의 전부를 나타낸다  transpose()를 이용하여 칼럼과 인덱스의 위치변경이 가능하다 |
| 오후 | ※  df.set\_index(칼럼명) - 인덱스 설정  df.reindex(인덱스배열) - 인덱스 재배치  df.reset\_index() - 인덱스 제거  inplace=True - 기존의 인덱스변경  df.sort\_index - 행기준  df.sort\_values - 열기준  head() - 처음 5라인을 보여준다  tail() - 마지막 5라인을 보여준다  info() - 데이터프레임 속성정보와 데이터갯수를 보여준다  describe() - 통계정보를 보여준다 ex)합계, 평균값  df.to\_csv(파일명)을 이용하여 csv 파일로 저장이 가능하다 |