|  |  |
| --- | --- |
| 교육 제목 | **Pandas - DataFrame 복습 및 데이터 입출력** |
| 교육 일시 | 21.09.29 |
| 교육 장소 | 오프라인 (영우글로벌러닝) |
| **교육 내용** | |
| 오전 | **▣ DataFrame**  -------------------------------------------------------------------------------------------  **∇ 행 찾기 (행은 범위가 있음)**  \* df.loc['행 이름'] or df.loc['행 이름' : '행 이름'] (마지막 포함)  \* df.iloc[숫자] or df.iloc[행 번호 : 행 번호] (마지막 미포함)  -------------------------------------------------------------------------------------------  **∇ 열 찾기 (열은 범위가 없음)**  \* df['열 이름']  -------------------------------------------------------------------------------------------  **∇ 행 열 동시에 찾기**  \* df.loc['행 이름', '열 이름'] or df.loc['행 이름' : '행 이름', '열 이름' : '열 이름']  or df.loc[['행 이름', '행 이름'], ['열 이름', '열 이름']]  \* df.iloc[행 번호, 열 번호] or df.iloc[행 번호 : 행 번호, 열 번호 : 열 번호]  or df.iloc[[행 번호, 행 번호], [열 번호, 열 번호]]  -------------------------------------------------------------------------------------------  **∇ 행 추가하기**  \* df.loc[행 숫자] = 데이터 값 or df.loc[행 숫자] = [데이터 값, ... ,값 ]  \* df.loc['추가하려는 행 이름'] = 데이터 값 or df.loc['추가하려는 행 이름'] = [데이터 값, ... ,값 ]  -------------------------------------------------------------------------------------------  **∇ 열 추가하기**  \* df['추가하려는 열 이름'] = 데이터 값 or df['추가하려는 열 이름'] = [데이터 값, ... ,값 ]  \* df['추가하려는 열 이름'] = 데이터 값  ------------------------------------------------------------------------------------------- |
| 후 | **▣ 데이터 입출력**  -------------------------------------------------------------------------------------------  **∇ 행과 열 바꾸기**  \* df.T or df.transpoed() -> 원본은 바뀌지 않음  -------------------------------------------------------------------------------------------  **∇ index 재배치 하기**  \* new\_index = ['r0', 'r1', 'r2', 'r3', 'r4'] # 추가되는 인덱스 이름 데이터를 리스트로 먼저 설정  \* ndf = df.reindex(new\_index) # df.reindex(인덱스 리스트 명),ill\_value=0 가 없으면 NaN값으로 대체됨.  \* ndf = df.reindex(new\_index, fill\_value=0) # fill\_value=0 가 포함되면 0값으로 대체됨.  -------------------------------------------------------------------------------------------  **∇ .reset\_index()**  \* 인덱스를 초기화 해주고 index를 columns명으로 해서 데이터값으로 들어감  -------------------------------------------------------------------------------------------  **∇ .sort\_index()**  \* 행의 오름차순, 내림차순 또는 특정 열의 오름차순, 내림차순으로 재배열  \* df.sort\_index(ascending = True) # 행의 오름차순으로 재배열  \* df.sort\_index(ascending = False) # 행의 내림차순으로 재배열    \* df.sort\_values(by= '특정 열', ascending = True/False)  # 특정 열 기준의 오르차순 내림차순으로 재배열  \* df.sort\_values(by= ['특정 열', '특정열'] ascending = True/False)  # 특정 열 기준의 오르차순 내림차순으로 재배열  -------------------------------------------------------------------------------------------  **∇ Series 연산**  \* '+, -, \*, /'' 사용하기  -> 각각의 일치하는 인덱스이름끼리 계산, 불일치 인덱스는 NaN으로 출력  \* add, sub, mul, div 사용하기.  시리즈1.add(시리즈2, fill\_value = x), 시리즈1.sub(시리즈2, fill\_value = x),  시리즈1.mul(시리즈2, fill\_value = x), 시리즈1.div(시리즈2, fill\_value = x)  -> 불일치 인덱스는 fill\_value 사용.  -------------------------------------------------------------------------------------------  **∇ DataFrame 연산**  \* 기존 모듈에서 dataframe을 불러들임  -------------------------------------------------------------------------------------------  **∇ seaborn module**  \* titanic 등의 데이터가 있음.  -------------------------------------------------------------------------------------------  **∇ 데이터 입출력하기**  \* 파일변수 = "./파일이름.csv" # 파일변수 생성  \* df = pd.read\_csv(파일변수) # 가져온 파일변수를 DataFrame 변수로 생성  \* df  -------------------------------------------------------------------------------------------  **∇ pd.read 옵션들**  \* path = 파일의 위치를 포함한 파일명  \* sep = 필드를 구분하는 구분자  \* header = 헤더가 정의되어 있는지 None  \* index\_col = 인덱스로 사용될 컬럼명  \* names = 칼럼 이름으로 사용 될 문자열 리스트  \* skipprows: 처음 행 부터 skip 하고자하는 행 수  \* skip\_footer = 마지막 행 부터 skip 하고자하는 행 수  \* encoding = 텍스트 인코딩 종류를 지정 'utf-8'  -------------------------------------------------------------------------------------------  **∇ 확장자 별로 파일 읽기 및 저장하기**  \* 읽기  변수.read\_csv, 변수.read\_html, 변수.read\_json, 변수.read\_excel  \* 저장하기  변수.to\_csv("파일 명"), 변수.to\_html("파일 명"), 변수.to\_json("파일 명"), 변수.to\_excel("파일 명")  ------------------------------------------------------------------------------------------- |