|  |  |
| --- | --- |
| 교육 제목 | **Pandas 입문** |
| 교육 일시 | 2021.11.08 |
| 교육 장소 | 영우 글로벌 러닝 |
| **교육 내용** | |
| 오전 | 1. **판다스 입문**   -좌측에 index가 상부에 column들로 구성된 데이터를 다룬다.  Index의 이름이 지정 안 될시 0부터 1씩 진행된다.  리스트와 딕셔너리를 pd.series(인덱스,데이타)로 시리즈를 구성할 수 있다.  [start:end]에서 index의 이름으로 지정할 시 end까지 포함한 범위로 출력된다.  \*데이터가 순차적으로 나열된 1차원 배열을 시리즈라 한다.  -DataFrame의 경우 pd.DataFrame(데이터)로 실행하며 딕셔너리 형태의 경우 key 가 column이 되며 value값은 반드시 list화 되있어야 한다.  DataFrame(데이터,index=[인덱스명…],columns=[칼럼명….]으로 인덱스 이름과 칼럼 이름을 정의 할 수 있다.  \*이미 칼럼이 정해진 딕셔너리 형태의 데이터를 넣고 칼럼명을 재정의 하면 값이 정의 되어 있지 않기 때문에 nan이 출력된다.  -행,열 삭제  DataFrame 객체.drop(행 인덱스 또는 배열,axis=0)  DataFrame 객체.drop(columns명 또는 배열,axis=1)  \*inplace=True를 해야 원본 데이터에 편집이 적용된다.  -행 선택  DataFrame 객체.loc(index명)  DataFrame 객체.iloc(index 정수)  로 행을 선택 할 수 있다.  -열 선택  DataFrame 객체[칼럼명]  -행과 열로 특정 데이터 추출하기  DataFrame 객체.iloc[행열 위치]  DataFrame 객체.loc[인덱스명,칼럼명]  -새로운 행과 열추가  DataFrame 객체.loc[인덱스명]=[데이터]  DataFrame 객체[칼럼명]=[데이터]  -행과 열 위치 바꾸기  DataFrame 객체.transpose()  DataFrame 객체.T |
| 오후 | 1. **Pandas 입문**   다차원 배열이나 행렬 수학 함수 지원해주는 모듈이다.  -인덱스 활용  DataFrame객체.set\_index(열 이름)-특정 열을 인덱스로 설정  Set\_index를 사용하면 기존 인덱스는 삭제된다.  DataFrame객체.reindex(열 이름,fill\_value=값)-새로운 인덱스 배열과 값 지정  DataFrame객체.reset\_index()-행 인덱스를 정수형으로 초기화  DataFrame객체.sort\_index  DataFrame객체.sort\_values(ascending=)-ascending값이 True면 오름차 False면 내림차로 데이터를 정렬한다.  -산술연산  Nan 값과 연산시 출력되는 값은 nan 이므로 연산 메소드 사용시 fill\_value를 이용하여서 숫자 0을 지정해주어 계산을 진행한다.  \* DataFrame객체.head() 꼭대기 5개의 데이터 출력  DataFrame객체.tail() 하부 5개의 데이터 출력  \*seaborn 그래프 생성에 도움되는 라이브러리로 그 외에 좋은 샘플 데이터를 제공한다.  2. 데이터 입출력  -csv 파일가져오기  Pd.read\_csv(경로와 파일명)  Header=None 첫라인을 data로 취급  Index\_col 인덱스로 사용할 칼럼의 번호 또는 이름을 사용  Names= 열이름으로 사용할 문자열의 리스트  Skiprows 생략할 줄을 설정  DataFrame객체.to\_csv(경로,파일명) csv파일로 저장 |