|  |  |
| --- | --- |
| 교육 제목 | 개발기초 역량교육\_디지털 트랜스포메이션\_2 |
| 교육 일시 | 2021.09.29 |
| 교육 장소 | 영우글로벌러닝 |
| **교육 내용** | |
| 오전 | 1. 데이터프레임 복습, 인덱스의 활용   1-1. 메소드 loc는 늘 대괄호 쓴다. 주의.  1-2. set\_index는 그 자체로 데이터프레임을 리턴한다. 그러니까 그 왼쪽에 df=을 쓸 필요가 없다.즉 df.set\_index(“이름”,inplace=True) 이런 식으로 사용가능  1-3 대괄호를 사용할 때는 인덱스로 각각 선택할 수도 있고, 콜론으로 연결되면 마지막을 포함하지 않는 연속된 인덱스를 의미한다.  1-4 set\_index()메소드를 사용하여 데이터프레임의 특정 열을 행 인덱스로 설정한다. 기본은 원본 데이터프레임을 바꾸지 않으며, 원본을 바꾸려면 inplace=True값을 추가로 더 넣어줘야 한다. 재귀적으로 데이터프레임을 다룰 때 사용하면 편리하다.  1-5행 인덱스를 재배열할때는 reindex를 사용하며, 빈 값은 NaN(Not a Number)값을 반환한다. 에러를 유발할 수 있으니 0으로 변경해주자,.  1-6..reset\_index로 배열을 초기화할수 있으며 사용시 기존의 인덱스를 초기화한다.  1-8 정렬은 행 인덱스를 기준으로, 혹은 열에 해당하는 값에 따라, key값을 함수로 다양한 방법으로 정렬할 수 있다.  1-9 데이터프레임도 사칙연산을 만족한다.요소에 NaN이 있으면 NaN을 리턴함에 유의  1-10 클래스의 메소드, 데이터프레임의 행과 열 표현, 데이터프레임의 메소드 표현 등에 익숙해져야 한다. 상당히 복잡해 보이지만 익숙해지면 별거없다. |
| 오후 | 1. 데이터 입출력   2-1.파이썬이 하나의 플랫폼처럼 있어 여러 형태의 파일을 비슷하게 읽고 저장할 수 있도록 지원한다.  2-2 csv, html, excel, json에 대해 read\_를 전치하여 읽을 수 있고, to\_가 전치되면 쓸 수 있게 된다.  2-3. 처음에 오타가 없는데도 새로운 기능이 작동하지 않으면 pip install이 잘 됐나 확인하자. html이라 방심했는데 lxml을 pip install해야만 했다. 꽤 오래 오탈자 찾다가 잘 안돼서 파이썬 디버그 기능을 활용하여 알아냈는데, 오탈자도 없고 데이터도 제 위치에 잘 있다면 무조건 디버그 기능을 잘 활용해 보자  2-4. json은 DataFrame과 구조가 동일하고, csv는 2차원 배열이 기본형인데 이를 DataFrame으로 변경해서 사용한다.  2-5 다양한 형태가 막 다르고 어렵다기보다 낯설 수 있으니 이런것도 있구나 하면서 익숙해지자. 어차피 데이터를 판다스로 옮기기만 하면 익숙한 형태로 쓸 수 있다. |