|  |  |
| --- | --- |
| 교육 제목 | **Pandas** |
| 교육 일시 | 21년 09월 30일 |
| 교육 장소 | YGL |
| **교육 내용** | |
| 오전 | 데이터 내용 미리보기  -객체.head(): 데이터프레임의 위쪽에 있는 내용을 괄호안의 갯수만큼 보는 것이고 빈 괄호면 기본적으로 5개이다.  -객체.tail(): 데이터프레임의 아래쪽에 있는 내용을 괄호안의 갯수만큼 보는 것이고 빈 괄호면 기본적으로 5개이다.  데이터 요약 정보 확인하기  -객체.shape : 이 메소드는 데이터 프레임의 행과 열의 갯수를 보여준다.  -객체.info(): 이 메소드 데이터프레임에 관한 기본 정보를 화면에 출력하는데 클래스 유형, 행 인덱스의 구성, 열 이름의 종류와 개수, 각열의 자료형과 개수, 메모리 할당량에 관한 정보가 출력된다.  판다스 자료형(data type)  int64: int 정수형 데이터, fload64:float 실수형 데이터, object:string 문자열 데이터, datetime64/timedelta64: 파이썬에는 없고 datetime 라이브러리 활용 시간데이터  -객체.describe():이 메소드는 산술 데이터를 갖는 열에 대한 기술 통계 정보(평균, 표준편차, 최대값, 최소값, 중간값 등)을 요약해서 출력한다.  -객체.count():이 메소드는 데이터프레임의 각 열의 데이터 개수를 표시해 준다.  -객체[”열이름”].value\_counts(): 이 메소드는 시리즈 객체의 고유값(같은 값) 개수를 세는데 사용하는데, 데이터프레임의 열은 시리즈이므로 각 열의 고유값의 종류와 개수를 출력해준다.  -객체.mean/median/min/max/std/corr() 또는 객체[”열이름”].mean/median/min/max/std/corr() :이 메소드 들은 각각 평균값,중간값,최소값,최대값,표준편차,상관계수를 출력해준다.  판다스 내장 그래프 도구 활용  -객체.plot(kind = kind옵션): 판다스에 내장되어 있는 그래프들을 데이터 프레임에 적용해주는 메소드로 kind 옵션들로는 line 선그래프 , bar 수직 막대 그래프, barh 수평 막대 그래프, his 히스토그램, box 박스 플롯, kde 커널 밀도 그래프, area 면적 그래프, pie 파이 그래프, scatter 산점도 그래프, hexbin 고밀도 산점도 그래프가 있고, 빈 괄호로 두면 line / 선그래프가 출력이 된다.  \*데이터 전처리 : 컬럼의 데이터 타입 변경, 원하는 정보만 추출, 데이터 가공, NaN 데이터 처리 ...  Matplotlib 기본 그래프 도구  -파이썬과 판다스에 기본 내장 그래프가 있지만 다양하지가 않다, 그래서 맷플롯라이브러리는 2d평면 그래프에 관한 다양한 포맷과 기능을 지원해서 많이 유용하다.  -이건 판다스에 기본적으로 내장된 기능이 아니어서 그래프를 쓰려면 fig.ex = 객체와 그래프 출력 메소드 (여기서 fig는 matplotlib에서 도화지를 하나 깔아주는 키워드/객체이다) 와 같이 변수에 저장을 해야한다.  -객체.fillna(method='옵션') : 데이터를 데이터프레임으로 변환할때 적절한 값을 찾지 못해서 NaN데이터가 들어가 있는 경우가 있는데 이때 fillna 메소드에서 method='ffill' 옵션을 사용하면 누락데이터(NaN)가 들어 있는 행의 바로 앞에 위치한 행의 데이터 값으로 채운다.  -plt.title/xlabel/ylabel(”이름”) : 각각 그래프의 제목, x축,y축의 이름을 출력해준다  -plt.show():그래프를 출력해준다  -그래프 객체를 생성해서 서브 플롯을 생성할 수 있다  ex) fig = plt.figure(figsize=(20,10)  ax = fig.add\_subplot(행의 갯수, 열의 갯수, 위치) |
| 오후 | Seaborn 라이브러리  -seaborn은 matplotlib의 기능과 스타일을 확장한 파이썬 시각화 도구의 고급버전.  -regplot() 함수는 서로 다른 2개의 연속 변수 사이의 산점도르 그리고 선형회구분석에 의한 회귀선을 함께 나타낸다. fit\_reg=False 옵션을 설정하면 회귀선을 안보이게 할 수 있다.  Folium 라이브러리  지도 위에 시각화할 때 유용한 도구이고 세계 지도를 기본으로 지원하고 다양한 스타일의 이미지를 제공하고 있다.  -Map() 함수: 객체 = folium.Map(location=[위도,경도],옵션)으로 간단히 현실 세계의 지도를 만들 수 있다. |