|  |  |
| --- | --- |
| 교육 제목 | 혼자 공부하는 파이썬 |
| 교육 일시 | 20210928 화요일 |
| 교육 장소 | 영우글로벌러닝 |
| **교육 내용** | |
| 오전 | #pandas를 활용하여 데이터 처리  import pandas as pd  dict\_data = {'a':1, 'b':2,'c':3}  ar = pd.Series(dict\_data) # dictinary 를 series 형식으로 변환  print(type(dict\_data))  print(type(ar))  ===  <class 'dict'>  <class 'pandas.core.series.Series'>  print(dict\_data)  print(ar)  print(dict\_data['a']," : ", ar["a"], " : ", ar[0]) # ar에서 'a' 인덱스명  print(ar[['a','c']])  ar[1:2] # 인덱스 철자로 접근하면 마지막 철자 위치 데이터 불포함  ar['b' : 'c'] # 인덱스명으로 접근하면 마지막 데이터 포함  #딕셔너리를 시리즈로 변경 : 딕셔너리의 키가 시리즈의 인덱명이 됨  # 접근은 인덱스명 또는 인덱스 첨자로 접근 가능  ===  {'a': 1, 'b': 2, 'c': 3}  a 1  b 2  c 3  dtype: int64  1 : 1 : 1  a 1  c 3  dtype: int64  b 2  c 3  dtype: int64  dict\_data = {'a':[1,2,3],'b':[4,5,6], 'c':[7,8,9]} # 'a'는 key 값, 데이터푸래임의 컬럼명  df = pd.DataFrame(dict\_data)  print(df)  #index명을 부여  df1 = pd.DataFrame(dict\_data, index=["idx1", "idx2", "idx3"])  print(df1)  a\_ver = [[1,2,3],[4,5,6],[7,8,9]]  df2 = pd.DataFrame(a\_ver, index=["idx1", "idx2", "idx3"],columns=['a1','b1','c1'])  df2  print(df2.rename(index={'idx1':'ix1','idx2':'ix2','idx3':'ix3'}, inplace=True))  df2  ===  a b c  0 1 4 7  1 2 5 8  2 3 6 9  a b c  idx1 1 4 7  idx2 2 5 8  idx3 3 6 9  None  a1 b1 c1  ix1 1 2 3  ix2 4 5 6  ix3 7 8 9 |
| 오후 | # 행인덱스명 변경 : object, index = 새로운 행 인덱스 배열  # 컬럼 인덱스명 변경 : object.columns = 새로운 컬럼명 배열  a\_ver = [[1,2,3],[4,5,6],[7,8,9]]  df = pd.DataFrame(a\_ver)  df  df.index = ['a','b','c']  df.columns = ['c1','c2','c3']  df1 = df  df1.rename(index = {'a':'ida','b':'idb'}, inplace=True)  print("df1 : :")  print(df)  #행 삭제 axis=0, 열 삭제 axis=1  df2 = df1  print(df1.drop('ida',axis=0,inplace=True))  df1  df2.drop('c1', axis=1)  df2  ===  df1 : :  c1 c2 c3  ida 1 2 3  idb 4 5 6  c 7 8 9  None  c1 c2 c3  idb 4 5 6  c 7 8 9  # 예제 1-7  import pandas as pd  # DataFrame() 함수로 데이터프레임 변환, 변수 df에 저장  exam\_data = {'수학' : [90, 80, 70],'영어' : [98,89,95],  '음악' : [85,95,100], '체육' : [100,90,90]}  df = pd.DataFrame(exam\_data,index=['서준','우현','인아'])  print(df)  print()  # 데이터프레임 df를 복제하여 변수 df2에 저장. df2의 1개 행(row) 삭제  df2 = df[:]  df2.drop('우현', inplace=True)  print(df2)  print()  #데이터프레임 df를 복제하여 변수 df3에 저장. df3의 2개 행(row) 삭제  df3 = df[:]  df3.drop(['우현','인아'], axis=0, inplace=True)  print(df3)  print()  df3.drop(['수학'], axis=1, inplace=True)  print(df3)  ===  수학 영어 음악 체육  서준 90 98 85 100  우현 80 89 95 90  인아 70 95 100 90  수학 영어 음악 체육  서준 90 98 85 100  인아 70 95 100 90  수학 영어 음악 체육  서준 90 98 85 100  영어 음악 체육  서준 98 85 100  exam\_data = {'수학' : [90, 80, 70],'영어' : [98,89,95],  '음악' : [85,95,100], '체육' : [100,90,90]}  df = pd.DataFrame(exam\_data,index=['서준','우현','인아'])  print(df.loc['서준'])  print()  print(df.iloc[0])  print()  print(df.loc['서준':'인아']) #'인아' 의 자료도 포함  print()  print(df.iloc[0:2]) # 첨자 인덱스의 2의 자료는 포함하지 않음  print()  print(df['수학']) # 컬럼 인덱스로 검색  print()  print(df.수학)  print()  print(df['수학':'영어']) # 컬럼 인덱스로 검색  print()  df.loc["과학"] = [100,100,100,100] # row로 데이터를 추가하는 경우 df.loc[인덱스명] = [값, ... ]  df  ===  수학 90  영어 98  음악 85  체육 100  Name: 서준, dtype: int64  수학 90  영어 98  음악 85  체육 100  Name: 서준, dtype: int64  수학 영어 음악 체육  서준 90 98 85 100  우현 80 89 95 90  인아 70 95 100 90  수학 영어 음악 체육  서준 90 98 85 100  우현 80 89 95 90  서준 90  우현 80  인아 70  Name: 수학, dtype: int64  서준 90  우현 80  인아 70  Name: 수학, dtype: int64  Empty DataFrame  Columns: [수학, 영어, 음악, 체육]  Index: []  수학 영어 음악 체육  서준 90 98 85 100  우현 80 89 95 90  인아 70 95 100 90  과학 100 100 100 100  def input\_score():  scores = input("국어 영어 수학 점수 입력 :").split()  for i, score in enumerate(scores):  scores[i] = int(score)  return scores  # 성적 입력 프로그램 작성  import pandas as pd  df = pd.DataFrame(columns=["국어","영어","수학"])  while True:  name = input("이름 입력 quit 종료 : ")  if name == "quit":  break    score = input\_score()  df.loc[name] = score  df  search\_name = input("검색할 이름 입력 :")  for name in list(df.index):  if search\_name == name:  kor = df.loc[search\_name,'국어']  eng = df.loc[search\_name,'영어']  math = df.loc[search\_name,'수학']  print("{}:{},{},{}, total = {}".format(search\_name,kor,eng,math,kor+eng+math))  ===  이름 입력 quit 종료 : 홍  국어 영어 수학 점수 입력 :10 20 30  이름 입력 quit 종료 : 길  국어 영어 수학 점수 입력 :40 50 60  이름 입력 quit 종료 : 동  국어 영어 수학 점수 입력 :50 60 70  이름 입력 quit 종료 : quit  국어 영어 수학  홍 10 20 30  길 40 50 60  동 50 60 70  검색할 이름 입력 :홍  홍:10,20,30, total = 60  import pandas as pd  # DataFrame() 함수로 데이터프레임 변환. 변수 df에 저장  exam\_data = {'이름' : [ '서준', '우현', '인아'],  '수학' : [ 90, 80, 70],  '영어' : [ 98, 89, 95],  '음악' : [ 85, 95, 100],  '체육' : [ 100, 90, 90]}  df = pd.DataFrame(exam\_data)  # '이름' 열을 새로운 인덱스로 지정하고, df 객체에 변경사항 반영  df.set\_index('이름', inplace=True)  df  ===  수학 영어 음악 체육  이름  서준 90 98 85 100  우현 80 89 95 90  인아 70 95 100 90 |