# -\*- coding: utf-8 -\*-

"""

Created on Thu Feb 27 16:46:58 2020

@author: sundooedu

"""

#%%

import pandas as pd

import numpy as np

# 데이터 불러오기

soil = pd.read\_csv('../data/토양정보.csv')

proportion = pd.read\_csv('../data/논밭비율.csv')

# 필요없는 컬럼 제거

soil.drop(['Unnamed: 0','년도'], axis=1, inplace=True)

proportion.drop(['Unnamed: 0','년도'], axis=1, inplace=True)

# 필요없는 행 제거

proportion.drop(index=proportion[proportion['시도']=='합계'].index, axis=0,inplace=True)

proportion.reset\_index(inplace=True)

proportion.drop(['index'], axis=1, inplace=True)

soil.drop(index=soil[soil['화학성']=='벤조(a)피렌'].index, axis=0, inplace=True)

soil.reset\_index(inplace=True)

soil.drop(['index'], axis=1, inplace=True)

# 결측치 바꾸기

soil.replace('-',0,inplace=True)

#데이터 프레임 합치기

df=pd.DataFrame()

for i,sido in enumerate(proportion.year\_sido):

x=proportion.iloc[i,2] #논비율

y=proportion.iloc[i,1] #밭비율

for j in list(range(0,170,17)):

df\_temp=soil.iloc[i:i+17,:]

temp\_list=[]

for k in list(range(17)):

temp=(float(df\_temp.iloc[k,1])\*x + float(df\_temp.iloc[k,2])\*y)/2

temp\_list.append(temp)

df\_temp['value']=np.array(temp\_list)

df\_tmp=pd.DataFrame(df\_temp['value'].values.reshape(1,-1),

columns=df\_temp['화학성'],

index=[sido])

df=pd.concat([df,df\_tmp], axis=0)

df.to\_csv('../data/soil\_proportion.csv')