**# 패키지 import**

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

**# 코로나 데이터 결측치 제거 및 월별 통합 전처리**

df = pd.read\_csv('./covid\_day\_.csv')

df.info()

df\_date = df['date'].unique()

df.iloc[:, 1:] = df.iloc[:, 1:].replace('-', '0')

df['total'] = df['total'].str.replace(',', '')

df['kor'] = df['kor'].str.replace(',', '')

df\_sum = []

for i in range(1,10):

df\_sum.append(df[df['date'] == f'2020. 0{i}'].iloc[:, 1:].sum().values)

for i in range(10, 13):

df\_sum.append(df[df['date'] == f'2020. {i}'].iloc[:, 1:].sum().values)

for i in range(1, 10):

df\_sum.append(df[df['date'] == f'2021. 0{i}'].iloc[:, 1:].sum().values)

df\_date = pd.DataFrame(df\_date)

df\_sum = pd.DataFrame(df\_sum)

df\_date.columns = [‘date]

df\_sum.columns = ['total', 'kor', 'foreign', 'death']

total = [df\_date, df\_sum]

corona = pd.concat(total, axis=1)

corona.to\_csv('./covid\_month.csv', index=False)

**# Naver click data 결측치 제거 및 월별 통합 전처리**

df = pd.read\_csv('./click\_elec\_online.csv')

df\_date = df['date'].unique()

df\_sum = []

for i in range(3,10):

df\_sum.append(df[df['date'] == f'2020. 0{i}'].iloc[:, 1:7].sum().round(1).values)

for i in range(10, 13):

df\_sum.append(df[df['date'] == f'2020. {i}'].iloc[:, 1:7].sum().round(1).values)

for i in range(1, 8):

df\_sum.append(df[df['date'] == f'2021. 0{i}'].iloc[:, 1:7].sum().round(1).values)

df\_date = pd.DataFrame(df\_date)

df\_sum = pd.DataFrame(df\_sum)

df\_date.columns = ['date']

df\_sum.columns = ['elec\_age10-19','elec\_age20-29', 'elec\_age30-39','elec\_age40-49', 'elec\_age50-59','elec\_age60-',

'gro\_age10-19', 'gro\_age20-29', 'gro\_age30-39', 'gro\_age40-49', 'gro\_age50-59', 'gro\_age60-',

'kids\_age10-19', 'kids\_age20-29', 'kids\_age30-39', 'kids\_age40-49', 'kids\_age50-59', 'kids\_age60-',

'trav\_age10-19', 'trav\_age20-29', 'trav\_age30-39', 'trav\_age40-49', 'trav\_age50-59', 'trav\_age60-',

'cul\_age10-19', 'cul\_age20-29', 'cul\_age30-39', 'cul\_age40-49', 'cul\_age50-59', 'cul\_age60-',

'offi\_age10-19', 'offi\_age20-29', 'offi\_age30-39', 'offi\_age40-49', 'offi\_age50-59', 'offi\_age60-'

]

total = [df\_date, df\_sum]

click = pd.concat(total, axis=1)

click.to\_csv('./click\_elec.csv', index=False)