3. Pandas dataFreame으로 변환하여 출려해보시오 4. '영등포구' 법정도 코드 번호를 확인하시오 실습과제-2 1. 서울시 열린 데이터 광장 공공데이터에서 관심있는 데이터를 선정하고, 데이터 값을 읽어와 데이터프레임에 저장하기 예 : 상권 분석, 공공 자전거 대여 현황 등 1. streamlit을 활용하여 해당 데이터에 적절한 홈 화면을 만드세요 LAB 1: 서울시 부동산 실거래가 데이터 가져오기 data.seoul.go.kr 서울 열린데이터 광장 데이터 가져오기 서울 열린 데이터 광장 데이터는 최대 1,000개 데이터만 가져올 수 있음 • 최근 1000개 데이터를 수집하여 seoul_real_estate_1000.csv에 저장하기 In []: # 실행 시점 기준 1~1000 데이터를 수집 import requests import json import pandas as pd API_KEY = '744d6d597773796a38397a75677550' def main() : # 'tbLnOpendataRtmsV'' 데이터세트의 최근 1000*11 데이터 요청 url = f'http://openapi.seoul.go.kr:8088/{API_KEY}/json/tbLn0pendataRtmsV/1/1000' print(url) req = requests.get(url) content = req.json() con = content['tbLnOpendataRtmsV']['row'] data = pd.DataFrame(con) data = data.reset_index(drop=True) data['DEAL_YMD'] = pd.to_datetime(data['DEAL_YMD'], format=("%Y%m%d")) data.to_csv('seoul_real_estate_1000.csv', index=False) print(data.head()) print(len(data), type(data)) if __name__ == "__main__": main() (실습과제) 최근 10,000개 데이터를 수집하여 seoul_real_estate.csv에 저장하기 In []: # 실행 시점 기준 1~11000 인덱스의 데이터를 수집 import requests import json import pandas as pd API_KEY = '744d6d597773796a38397a75677550' def main(): data = pd.DataFrame()): # 'tbLnOpendataRtmsV'' 데이터세트의 최근 1000*11 데이터 요청 url = f'http://openapi.seoul.go.kr:8088/{API_KEY}/json/tbLnOpendataRtmsV/{ }/{ print(url) req = requests.get(url) content = req.json() con = content['tbLnOpendataRtmsV']['row'] result = pd.DataFrame(con) data = (data = data.reset_index(drop=True) data['DEAL_YMD'] = pd.to_datetime(data['DEAL_YMD'], format=("%Y%m%d")) data.to_csv('seoul_real_estate.csv', index=False) print(data.head()) print(len(data), type(data)) if __name__ == "__main__": main() ETL (Extract 추출 / Transform 변환 / Load 적재) • Extract(추출): 데이터베이스, 스프레드시트, 인터넷 등 다양한 데이터 소스로 부터 데이터를 검색하고 수집 • Transdform(변환) :데이터를 효과적으로 분석할 수 있도록 정리, 필터링 및 구조화하는 작업 • Load(적재) : 변환된 데이터를 분석하기 위해 엑세스 할 수 있는 데이터 Lake와 같은 대상 시스템으로 이동 데이터 적재 단계에서 대시보드 데이터로 데이터가 연결됨 Streamlit 설치 1. ananconda Prompt 창 열기 2. streamlit 패키지 설치 • (패키지 설치) > pip install streamlit 3. spyder 실행후, 파이썬 프로그램 편집(app.py) • app.py 편집 후, 자신의 폴더 선택 후 저장 • ananconda Prompt 창에서 app.py 저장한 폴도로 이동 (> cd C:\Users\Jay\11-빅데이터프로젝트) 4. ananconda Prompt 창에서 streamlit 실행 • (8051 포드로 실행시) > streamlit run app.py • (8080 포트로 실행시) > streamlit run app.py --server.port 8080 • (주의) app.py 저장한 폴도로 이동해서 살해해야함 LAB 2 : Streamlit으로 홈 화면 꾸미기 1. 기본 화면 만들기 app.py (app1.py) > streamlit run app.py import streamlit as st from streamlit_option_menu import option_menu from utils import load_data from home import run_home def main() : total_df = load_data() with st.sidebar: selected = option_menu('데시보드 메뉴', ['홈', '탐색적 자료분석', '부동산 예측'], icons=['house', 'file-bar_graph', 'graph-up-arrow'], menu_icon='cast', default_index=0) if selected == '홈': run_home(total_df) elif selected == '탐색적 자료분석' : elif selected == '부동산 예측' : pass else: print('error') if __name__ == "__main__": main() 서울시실거래가분석하기 app - Streamlit C 🕀 localhost:8501 X ☑ 데시보드 메뉴 대시보드개요 습 홈 탐색적 자료분석 ∠ 부동산 예측 LAB 3: 자치구 월별 거래 최소가격, 최대가격 표시하기 • 'seoul_real_estate.csv' 파일을 읽어와서 데이터프레임으로 만드는 파이썬 프로그램 작성 • load_data() 함수를 별도의 'utils.py' 파일에 작성 In []: # # utils.py import pandas as pd def load_data() : data = pd.read_csv('seoul_real_estate.csv') return data • '홈' 버튼을 선택한 경우 실행 할 함수 run_home()를 작성 • run_home() 함수를 별도의 'home.py' 파일에 작성 run_home() 함수에서 load_data() 함수를 호출 • 숫자 단위가 커지면 천 단위 콤마(,)를 적용할 수 있도록 prettify 함수 사용하기 위해 millify 패키지 설치 pip install millify In []: # # home.py (home1.py) import pandas as pd import streamlit as st from millify import prettify def run_home(total_df) : st.markdown("## 대시보드 개요 ₩n" "본 프로젝트는 서울시 부동산 실거래가를 알려주는 대시보드 입니다." "여기에 추가하고 싶은 내용을 추가하면 됩니다") total_df['DEAL_YMD'] = pd.to_datetime(total_df['DEAL_YMD'], format="%Y-%m-%d") total_df['month'] = total_df['DEAL_YMD'].dt.month total_df['year'] = total_df['DEAL_YMD'].dt.year total_df = total_df.loc[total_df['HOUSE_TYPE'] == '아파트', :] st.dataframe(total_df) sgg_nm = st.sidebar.selectbox("자치구", sorted(total_df['SGG_NM'].unique())) acc_year = st.sidebar.selectbox("년도", [2023, 2024]) month_dic = { '1월 ' : 1, '2월 ' : 2, '3월 ' : 3, '4월 ' : 4, '5월 ' : 5, '6월 ' : 6, '7월': 7, '8월': 8, '9월': 9, '10월': 10, '11월': 11, '12월': 12 selected_month = st.sidebar.radio("확인하고 싶은 월을 선택하시오 ", list(month_dic.keys())) st.markdown("<hr>", unsafe_allow_html=True) st.subheader(f'{sgg_nm} {acc_year}년 {selected_month} 아파트 가격 개요') st.markdown('자치구와 월을 클릭하면 자동으로 각 지역구에서 거래된 **최소가격**, **최대가격**을 확인할 수 있습니다') col1, col2 = st.columns(2)filtered_month = total_df[total_df['month'] == month_dic[selected_month]] filtered_month = filtered_month[filtered_month['year'] == acc_year] filtered_month = filtered_month[filtered_month['SGG_NM'] == sgg_nm] march_min_price = filtered_month['OBJ_AMT'].min() march_max_price = filtered_month['OBJ_AMT'].max() with col1: st.metric(label = f'{sgg_nm} 최소가격(만원)', value = prettify(march_min_price)) with col2: st.metric(label = f'{sgg_nm} 최대가격(만원)', value = prettify(march_max_price)) • app.py에서 '홈' 버튼 선택시 run_home() 호출하도록 수정 • home.py 파일의 run_home() 함수를 호출할 수 있도록 'from home import run_home' 추가 In []: # # app.py (app1.py) > streamlit run app.py import streamlit as st from streamlit_option_menu import option_menu from utils import load_data from home import run_home def main() : total_df = load_data() with st.sidebar: selected = option_menu('데시보드 메뉴', ['홈', '탐색적 자료분석', '부동산 예측'], icons=['house', 'file-bar_graph', 'graph-up-arrow'], menu_icon='cast', default_index=0) if selected == '홈': run_home(total_df) elif selected == '탐색적 자료분석' : elif selected == '부동산 예측' : pass else : print('error') if __name__ == "__main__": main() 서울시실거래가분석하기 🍟 app - Streamlit ⑤ 39 (4) 탐색적 자료분석 ∠ 부동산 예측 대시보드 개요 자치구 본 프로젝트는 서울시 부동산 실거래가를 알려주는 대시보드 입니다.여기에 추가하고 싶은 내용을 추 강북구 > 가하면 됩니다 년도 2023 ~ 강북구 2023년 12월 아파트 가격 개요 확인하고 싶은 월을 선택하시오 ○ 1월 자치구와 월을 클릭하면 자동으로 각 지역구에서 거래된 **최소가격**, **최대가격**을 확인할 수 있습니다 ○ 2월 ○ 3월 강북구 최대가격(만원) 강북구 최소가격(만원) ○ 4월 6,400 115,215 ○ 5월 ○ 6월 ○ 7월 ○ 8월 ○ 9월 ○ 10월 ○ 11월 🚺 12월 실습과제-1 - 자치구 월별 거래건수를 함께 표시하도록 수정하시오(home2) 서울시실거래가분석하기 🍟 app - Streamlit + localhost:8501 데시보드 메뉴 대시보드 개요 탐색적 자료분석 본 프로젝트는 서울시 부동산 실거래가를 알려주는 대시보드 입니다.여기에 추가하고 싶은 내용을 추 ☑ 부동산 예측 가하면 됩니다 자치구 강남구 강남구 2023년 11월 아파트 가격 개요 > 년도 자치구와 월을 클릭하면 자동으로 각 지역구에서 거래된 최소가격, 최대가격을 확인할 수 있습니다 2023 > 강남구 거래건수(건) 강남구 최소가격(만원) 강남구 최대가격(만원) 11,800 3,250,000 226 확인하고 싶은 월을 선택하시오 ○ 1월 ○ 2월 실습과제-2 - 자치구 월별 거래 아파트 가격 상위 3개와 하위 3개를 테이블 형태로 보여주시오(home3) • 테이블에는 SGG_NM(자치구명), BJDONG_NM(법정동명), BLDG_NM(건물명), BLDG_AREA(건물면적), OBJ_AMT(물건금액)을 정렬하여 보여주시오 서울시실거래가분석하기 🏻 🕶 app - Streamlit 区 🕝 구로구 2023년 12월 아파트 가격 개요 ☑ 데시보드 메뉴 자치구와 월을 클릭하면 자동으로 각 지역구에서 거래된 최소가격, 최대가격을 확인할 수 있습니다 습 홈 구로구 최소가격(만원) 구로구 최대가격(만원) 구로구 거래건수(건) 6,000 165,000 128 탐색적 자료분석 ∠ 부동산 예측 아파트 가격 상위 3위 자치구 구로구 ~ SGG_NM BJDONG_NM BLDG_NM BLDG_AREA OBJ_AMT 년도 신도림동 대림e-편한 0 구로구 152.07 165,000 1 구로구 신도림동 대림2 2023 134.99 128,800 2 구로구 신도림동 동아1 103,000 84.91 확인하고 싶은 월을 선택하시오 ○ 1월 ○ 2월 ○ 3월 ○ 4월 아파트 가격 하위 3위 ○ 5월 ○ 6월 SGG_NM BJDONG_NM BLDG_NM BLDG_AREA OBJ_AMT ○ 7월 0 구로구 궁동 ○ 8월 삼융빌라4기 6,000 33.96 ○ 9월 오류동 1 구로구 아크로펠리 7,500 18 ○ 10월 2 구로구 오류동 썬앤빌 7,800 14.06 ○ 11월 🚺 12월

In []:

2주차 복습 과제

1. 자치구 코드(시군구 코드)는 '서울시 건축물대장 법정동 코드정보'에서 확인

2. '서울시 건축물대장 법정동 코드정보' Open API를 할용하여 30개 데이터를 가져와서

실습과제-1