

매뉴얼

<목차>

1. 사용 tool 버전에 맞게 설치하기
 - 1-1. Python
 - 1-2. Geocoding
 - 1-3. Qgis

2. 데이터 수집 및 전처리
 - 2-1. 제공데이터
 - 2-2. 공공데이터
 - 2-3. 결과도출용 데이터 제작
 - 2-4. 데이터 병합

3. 모델링
 - 3-1. LightGBM 패키지 설치
 - 3-2. 아파트 모델 및 결과도출용 데이터 적용 후 저장
 - 3-3. 감정가 모델

4. Qgis시각화
 - 4-1. Qgis적용 파일 제작
 - 4-2. 모델 시각화
 - 4-3. 두 모델 비교
 - 4-4. 아파트모델 결과도출용 데이터 적용 시각화

1. 사용 tool 버전에 맞게 설치하기

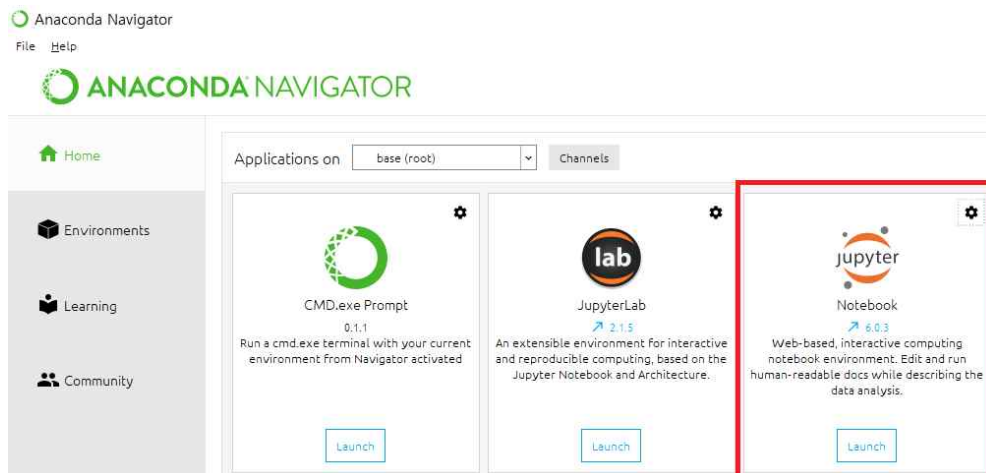
1-1. Python

<아나콘다 3.8 설치>

<https://www.anaconda.com/products/individual>



<Jupyter Notebook 6.0.3 설치>



1-2. Geocoding

<GeocodingTool 설치>

http://www.biz-gis.com/index.php?mid=pds&document_srl=187250

Go > 자료실 > 일반 자료실

16
2016-Mar

2. 프로그램 및 스크립트 소스 지오코딩 툴 정식 버전 - 최종 업데이트 일자:
작성자: 김한국 IP ADDRESS: *.12.216.187 조회 수: 48438

>> 최신버전 <<
2020년 10월 15일 버전 다운로드

구글 드라이브
32비트(구글 드라이브): 링크
64비트(구글 드라이브): 링크

※ 공공기관 고객은 구글 드라이브 차단 문제로 인해 담당자분께 별도 URL 보내드렸으니 확인 부탁드립니다.

수정사항:
- 2020년 6월 31일 기준 데이터 업데이트(마이너 업데이트)
- 드래그 & 드랍 기능 추가
- 최신 건물(주로 아파트) 검색 로직 추가

PC사양에 맞게 설치하기

1-3. Qgis

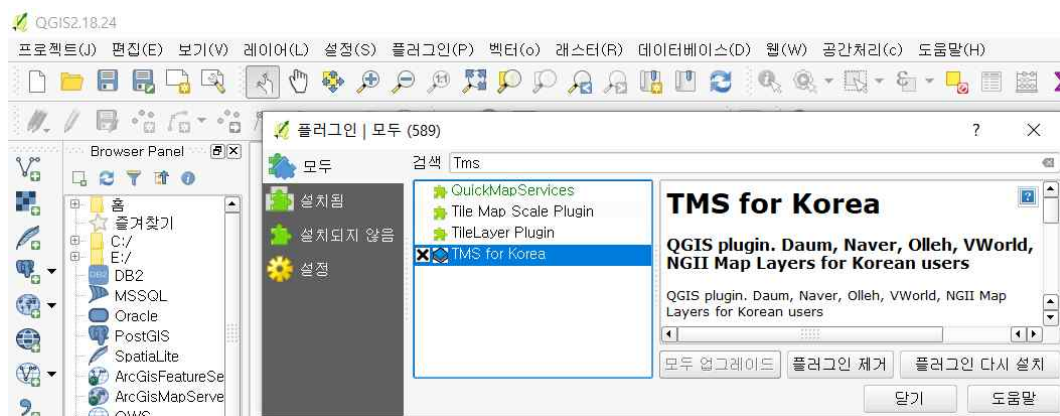
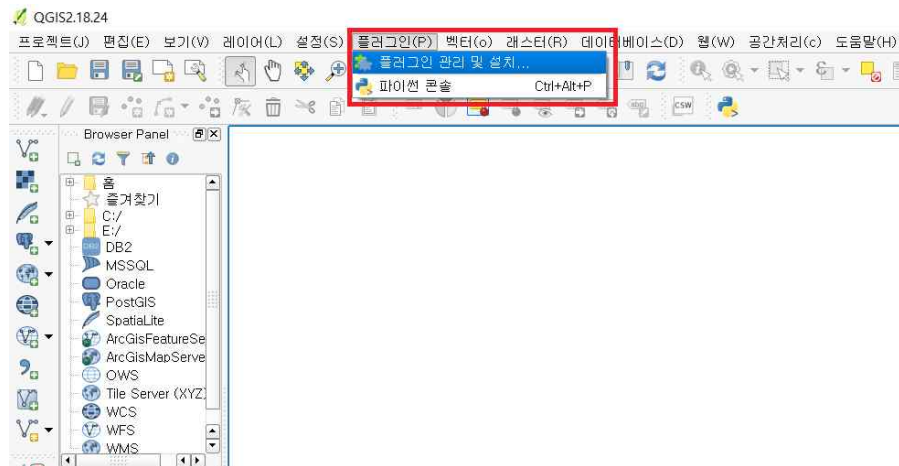
<Qgis 2.18.24 다운로드 및 세팅>

<https://qgis.org/downloads/>

사양에 맞게 다운로드

← → ↻	qgis.org/downloads/	
	QGIS-OSGeo4W-2.18.23-1-Setup-x86.exe	18-Aug-2018 22:49 339M
	QGIS-OSGeo4W-2.18.23-1-Setup-x86.exe.md5sum	18-Aug-2018 22:11 71
	QGIS-OSGeo4W-2.18.23-1-Setup-x86_64.exe	18-Aug-2018 23:04 400M
	QGIS-OSGeo4W-2.18.23-1-Setup-x86_64.exe.md5sum	18-Aug-2018 22:33 74
	QGIS-OSGeo4W-2.18.24-1-Setup-x86.exe	15-Sep-2018 15:30 339M
	QGIS-OSGeo4W-2.18.24-1-Setup-x86.exe.md5sum	15-Sep-2018 15:30 71
	QGIS-OSGeo4W-2.18.24-1-Setup-x86_64.exe	15-Sep-2018 15:46 401M
	QGIS-OSGeo4W-2.18.24-1-Setup-x86_64.exe.md5sum	15-Sep-2018 15:46 74
	QGIS-OSGeo4W-2.18.25-1-Setup-x86.exe	28-Oct-2018 12:25 344M

플러그인에서 TMS 설치



2. 데이터 수집 및 전처리

2-1. 제공 데이터

<아파트 일반정보 및 시세.csv (기ung)>

	A	B	C	D	E	F	G	H	I	J	K	L
1	기준년월	매매지역명	우편번호	아파트명	아파트구	아파트시	아파트동	아파트면적	최근매매	최근매매	최근전세	최근전세
	텍스트 오름차순 정렬(S)			강원도 삼척시	360	3	59.83	80500000	20190414	65000000	20190304	
	텍스트 내림차순 정렬(O)			강원도 동해시	114	3	84.99	1.08E+08	20190402	80000000	20190305	
	색 기준 정렬(D)			강원도 동해시	76	3	60	80000000	20190319	80000000	20190322	
	“매매지역명”에서 필터 해제(O)			서울특별시	28	1	50.49	5.4E+08	20181029	4E+08	20181030	
	색 기준 필터(I)			서울특별시	180	6	59.92	7E+08	20180825	4.1E+08	20190330	
	텍스트 필터(F)			서울특별시	72	6	114.67	8.5E+08	20180826	3E+08	20181128	
	검색			서울특별시	164	6	84.92	8E+08	20180912	4.5E+08	20190322	
	<input checked="" type="checkbox"/> (모두 선택)			서울특별시	110	2	84.95	7.22E+08	20180718	4E+08	20190311	
	<input type="checkbox"/> 강원도			서울특별시	110	2	112.54	7.99E+08	20180801	6.5E+08	20190309	
	<input type="checkbox"/> 대구광역시			서울특별시	193	5	94.51	1.05E+09	20190319	6.9E+08	20190328	
	<input checked="" type="checkbox"/> 서울특별시			서울특별시	28	5	136.4	1.13E+09	20180305	8.5E+08	20180518	
	확인			서울특별시	28	5	131.44	1.19E+09	20190119	8.5E+08	20190329	
				서울특별시	28	5	160.85	1.26E+09	20180228	9.5E+08	20181004	
				서울특별시	82	5	146.92	1.19E+09	20190130	8.5E+08	20181015	
				서울특별시	22	5	126.34	1.15E+09	20180904	4.5E+08	20170418	
				서울특별시	26	5	70.8	7.47E+08	20170822	6.5E+08	20181010	
				서울특별시	26	5	121.37	1.04E+09	20180813	7E+08	20180612	
				서울특별시	105	5	159.01	1.5E+09	20190304	9E+08	20181117	

“매매지역명”에서 서울특별시 필터링

	A	B	C	D	E	F	G	H	I	J
1	기준년월	매매지역명	우편번호	아파트명	아파트주소	동	아파트면적	아파트동	아파트면적	최근매매
5	201904	서울특별시	3092	창신	창신3동	156	7	54.7	4.1E+08	
6	201904	서울특별시	3092	창신	창신3동	284	7	64.66	4.78E+08	
7	201904	서울특별시	3092	창신	창신3동	283	7	79.87	5.35E+08	
8	201904	서울특별시	3092	창신	창신3동	112	7	106.62	5.37E+08	
9	201904	서울특별시	3092	창신	창신3동	84	7	115.53	6.29E+08	
10	201905	서울특별시	3092	창신	창신3동	156	7	54.7	4.1E+08	
11	201905	서울특별시	3092	창신	창신3동	284	7	64.66	4.78E+08	
12	201905	서울특별시	3092	창신	창신3동	283	7	79.87	5.35E+08	
13	201905	서울특별시	3092	창신	창신3동	112	7	106.62	5.37E+08	
14	201905	서울특별시	3092	창신	창신3동	84	7	115.53	5E+08	
15	201904	서울특별시	3092	창신	창신3동	84	7	115.53	6.29E+08	
16	201904	서울특별시	3092	창신	창신3동	112	7	106.62	5.37E+08	
17	201904	서울특별시	3092	창신	창신3동	283	7	79.87	5.35E+08	
18	201904	서울특별시	3092	창신	창신3동	284	7	64.66	4.78E+08	
19	201904	서울특별시	3092	창신	창신3동	156	7	54.7	4.1E+08	
20	201905	서울특별시	3092	창신	창신3동	156	7	54.7	4.1E+08	
21	201905	서울특별시	3092	창신	창신3동	284	7	64.66	4.78E+08	
22	201905	서울특별시	3092	창신	창신3동	283	7	79.87	5.35E+08	
23	201905	서울특별시	3092	창신	창신3동	112	7	106.62	5.37E+08	

“아파트주소” 오름차순 정렬 후 옆에 “동” 변수 생성, 이후 생성한 내용 복사 후 새로운 엑셀 파일에 붙여 넣고 “매매지역명” 변수 제거
아파트주소 값에 맞게 동 입력하기

=IF(G2>=H2,1,0)											
	A	B	C	D	E	F	G	H	I	AP	AQ
1	아파트명	아파트주소	동	아파트세대	아파트동수	아파트면적	최근매매실	매매일반	평균		
2	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
3	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
4	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
5	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
6	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
7	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
8	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
9	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
10	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0		
11	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.25E+08	0		
12	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.3E+08	0		
13	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.35E+08	0		
14	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.4E+08	0		
15	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.3E+08	0		
16	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.33E+08	0		
17	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.33E+08	0		
18	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.4E+08	0		
19	창신쌍용2	서울특별시	창신제3동	284	7	64.66	4.78E+08	4.55E+08	1		

사용할 변수만 추출 “아파트명, 아파트주소, 동, 아파트세대수, 아파트면적값,
최근매매실거래가격, 매매일반평균가격”

최근매매실거래가격이 매매일반평균가격보다 크거나 같으면 1, 아니면 0을 반환하는
변수 Y 제작.

이후 매매지역명 변수 제거하고 추출된 값 result1.csv로 저장

<매각결과.csv (한국감정평가사협회)>

	A	B	C	D	E	F	G	H	I	J	K	L
1	기준년월	법원명	사건번호	물건번호	소재지명	비고내용	감정평가	최소매각	매각내용	수집일자	적재일자	센터코드
19	201905	대구지방법원	2018타경31기	타	대구광역시 일괄매각	6.91E+08	6.91E+08	경매1계20	20190509	2.02E+13	BBP19	

“법원명”을 필터링하여 서울중앙지방법원만 추출

	A	B	C	D	E	F	G	H	I	J
1	기준년월	법원명	사건번호	물건번호	소재지및내역내용	동	비고내용	감정평가	최소매각	매각내용
170	202005	서울중앙지	2019타경41상가오피	사	텍스트 오름차순 정렬(S)	돈의동		20400000	20400000	경매7계20
226	202005	서울중앙지	2019타경41상가오피	사	텍스트 내림차순 정렬(Q)	돈의동		20400000	20400000	경매7계20
227	202005	서울중앙지	2019타경41상가오피	사	색 기준 정렬(O)	돈의동		20400000	20400000	경매7계20
228	202005	서울중앙지	2019타경41상가오피	사	*소재지및내역내용*에서 필터 해제(O)	돈의동		20400000	20400000	경매7계20
5620	202005	서울중앙지	2019타경41상가오피	사	색 기준 필터(U)	돈의동		20400000	20400000	경매7계20
5621	202005	서울중앙지	2019타경41상가오피	사	텍스트 필터(F)	돈의동		20400000	20400000	경매7계20
5622	202005	서울중앙지	2019타경41상가오피	사	검색	돈의동		20400000	20400000	경매7계20
5623	202004	서울중앙지	2020타경41다세대	사	<input checked="" type="checkbox"/> (모두 선택)	이화동		1.6E+08	1.6E+08	경매8계20
5624	202004	서울중앙지	2020타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 돈화문로5가길	이화동		1.6E+08	1.6E+08	경매8계20
5625	202004	서울중앙지	2020타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 동충4길 29	이화동		1.6E+08	1.6E+08	경매8계20
5626	202004	서울중앙지	2020타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 동충4길 54-6	이화동		1.6E+08	1.6E+08	경매8계20
5627	202004	서울중앙지	2019타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 세검정로7나길	이화동		1.6E+08	1.6E+08	경매8계20
5628	202004	서울중앙지	2019타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 수표로 96, 4길	이화동		1.6E+08	1.6E+08	경매8계20
5629	202004	서울중앙지	2019타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 수표로 96, 5길	이화동		1.6E+08	1.6E+08	경매8계20
5630	202004	서울중앙지	2019타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 옥인4길 2-3	이화동		1.6E+08	1.6E+08	경매8계20
5631	202004	서울중앙지	2019타경41다세대	사	<input checked="" type="checkbox"/> 서울특별시 종로구 옥인4길 2-3	이화동		1.6E+08	1.6E+08	경매8계20
5632	202004	서울중앙지	2020타경41다세대	사	확인	이화동		1.6E+08	1.6E+08	경매8계20
5633	202004	서울중앙지	2019타경41다세대	사	취소	이화동		1.6E+08	1.6E+08	경매8계20
5634	202004	서울중앙지	2020타경41다세대	사	서울특별시 양도구 동충4길 29, 에이동	이화동		1.6E+08	1.6E+08	경매8계20

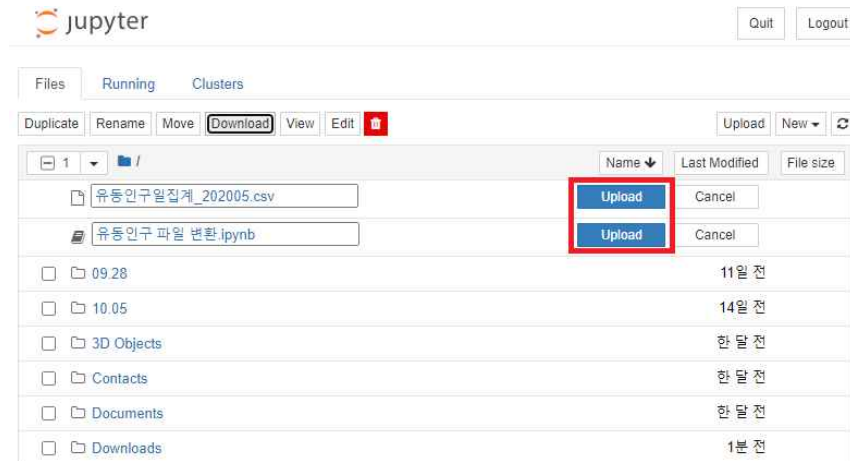
“소재지및내역내용” 오름차순 정렬후 옆에 “동” 변수 생성, 이후 생성한 내용 복사후 새로운 엑셀파일에 붙여넣고 “법원명” 변수 제거
주소 값에 맞게 동 입력하기

G2										
=IF(F2>=E2,1,0)										
	A	B	C	D	E	F	G	AN	AO	AP
1	물건번호	용도	소재지및내	동	주소	감정평가금	최소매각	Y		
2	1다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
3	1다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
4	1다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
5	1다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
6	1다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
7	1다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
8	2다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
9	2다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
10	2다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
11	2다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
12	2다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
13	2다세대	서울특별시 사직동	서울시 중	91500000	19189000	0				
14	2다세대	서울특별시 사직동	서울시 중	91500000	15351000	0				
15	2다세대	서울특별시 사직동	서울시 중	91500000	15351000	0				
16	2다세대	서울특별시 사직동	서울시 중	91500000	15351000	0				
17	2다세대	서울특별시 사직동	서울시 중	91500000	15351000	0				
18	2다세대	서울특별시 사직동	서울시 중	91500000	15351000	0				
19	2다세대	서울특별시 사직동	서울시 중	91500000	15351000	0				

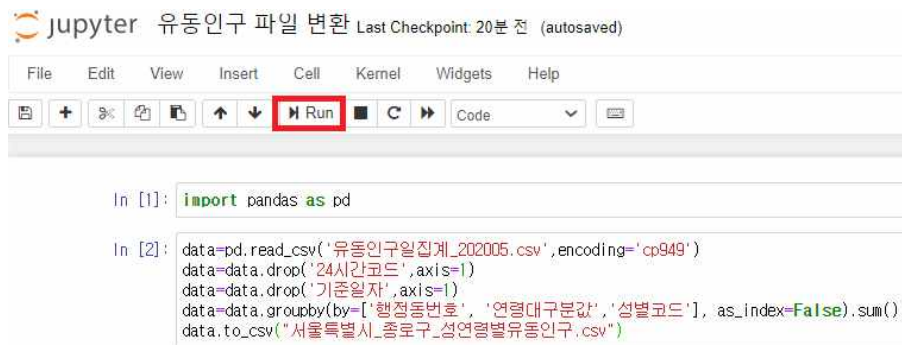
사용할 변수만 추출 “물건번호용도내용, 소재지및내역내용, 동, 감정평가금액, 최소매각가격”
최소매각가격이 감정평가금액 보다 크거나 같으면 1, 아니면 0을 반환하는
변수 Y 제작.

이후 법원명 변수 제거하고 추출된 값 result2.csv로 저장

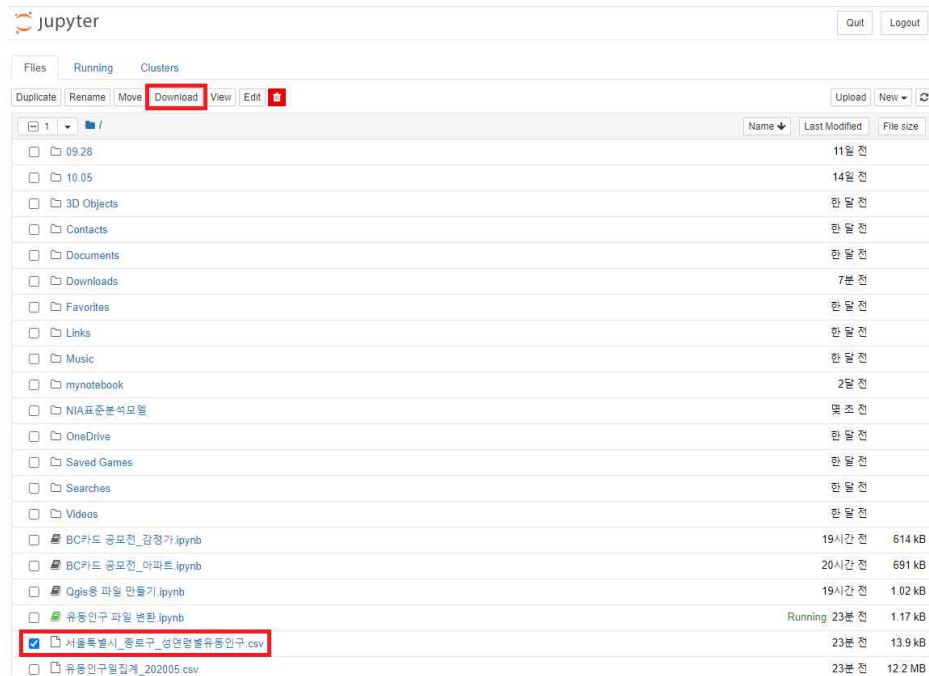
<유동인구일집계_202005.csv (케이티)>



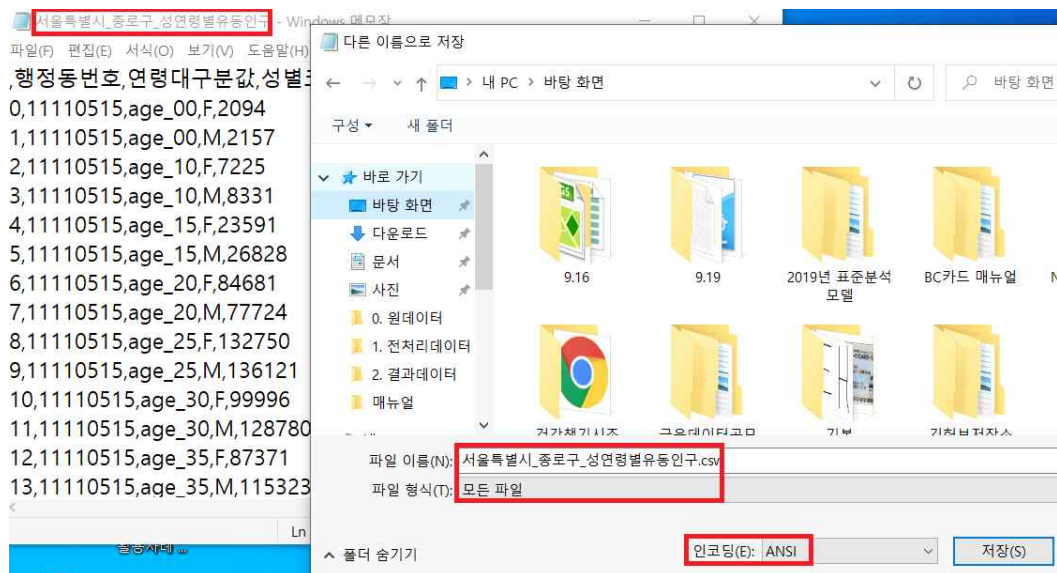
유동인구일집계_202005.csv 파일과 유동인구 파일 변환.ipynb 파일을 주피터 노트북상에 드래그 하여 올린 후 Upload



유동인구 파일 변환 코드 실행하기



생성된 서울특별시_종로구_성연령별유동인구.csv 파일 다운로드



서울특별시_종로구_성연령별유동인구.csv 파일을 메모장으로 읽어온후 인코딩을 ANSI로 지정하고 서울특별시_종로구_성연령별유동인구.csv 로 저장

<행정동번호를 동으로 바꾸기>

https://www.mois.go.kr/frt/bbs/type001/commonSelectBoardArticle.do?bbsId=BBSMS TR_000000000052&nttId=79418

- ※ 첨부된 zip파일 실행시 압축이 해제되는 파일형식은 TEXT, 엑셀이며, 윈도우(OS)에 기본 설치된 "워드패드"로 확인할 수 있습니다.
- ※ 파일 상단에 항목제목 추가되었습니다.
 - KiKcd_H.20200814→ 행정기관코드(행정동)
 - KiKcd_B.20200814 → 법정동코드(실제주소)
 - KiKmix.20200814 → 행정기관코드 + 관할 법정동코드

첨부파일	 주민등록업무 행정기관 및 관할구역 변경내역(세종시).hwp [41.5 KB]	바로보기
	 주민등록 행정구역코드 Layout.hwp [11.5 KB]	바로보기
	 jscore20200814.zip [1.9 MB]	
	 jscore20200814(말소코드포함).zip [4.3 MB]	

※시스템관리를 위해 10MB가 넘으면 바로보기가 제한 됩니다.

jscore20200814.zip 다운로드 받은후 압축풀고 KIKcd_H.20200814.xlsx 파일 추출

	A	B	C	D	E	F	G	H	I
1	행정동	시도명	시군구명	읍면동명	생성일자	말소일자			
	텍스트 오름차순 정렬(S)				19880423				
	텍스트 내림차순 정렬(O)				20081101				
	색 기준 정렬(I)				직동	19880423			
	"시군구명"에서 필터 해제(C)				청동	19880423			
	색 기준 필터(I)				암동	19880423			
	텍스트 필터(F)				창동	19880423			
	검색				약동	19880423			
	<input type="checkbox"/> 정읍시				남동	19880423			
	<input type="checkbox"/> 제주시				회동	19880423			
	<input type="checkbox"/> 제천시				로1.2.3.4	19981201			
	<input checked="" type="checkbox"/> 종로구				로5.6가	19880423			
	<input type="checkbox"/> 중구				화동	19880423			
	<input type="checkbox"/> 중구영종출장소				화동	19880423			
	<input type="checkbox"/> 중구용유출장소				신제1동	19880423			
	<input type="checkbox"/> 중랑구				신제2동	19880423			
	<input type="checkbox"/> 중평군				신제3동	19880423			
	<input type="checkbox"/> 지도구				인제1동	19880423			
					인제2동	19880423			

“시군구명” 변수에서 종로구만 필터링 후 추출

12												
	A	B	C	D	E	F	G	H	I	J	K	L
1	행정동번호	동	연령대구분	성별코드	인구수			행정동코드	행정동번호	시도명	시군구명	읍면동명
2	11110515	#N/A	age_00	F	2094			1111000000	11110000	서울특별시	종로구	
3	11110515		age_00	M	2157			1111051500	11110515	서울특별시	종로구	청운효자동
4	11110515		age_10	F	7225			1111053000	11110530	서울특별시	종로구	사직동
5	11110515		age_10	M	8331			1111054000	11110540	서울특별시	종로구	삼청동
6	11110515		age_15	F	23591			1111055000	11110550	서울특별시	종로구	부암동
7	11110515		age_15	M	26828			1111056000	11110560	서울특별시	종로구	평창동
8	11110515		age_20	F	84681			1111057000	11110570	서울특별시	종로구	무악동
9	11110515		age_20	M	77724			1111058000	11110580	서울특별시	종로구	교남동
10	11110515		age_25	F	132750			1111060000	11110600	서울특별시	종로구	가회동
11	11110515		age_25	M	136121			1111061500	11110615	서울특별시	종로구	종로1,2,3,4동
12	11110515		age_30	F	99996			1111063000	11110630	서울특별시	종로구	종로5,6가동
13	11110515		age_30	M	128780			1111064000	11110640	서울특별시	종로구	이화동
14	11110515		age_35	F	87371			1111065000	11110650	서울특별시	종로구	혜화동
15	11110515		age_35	M	115323			1111067000	11110670	서울특별시	종로구	창신제1동
16	11110515		age_40	F	76893			1111068000	11110680	서울특별시	종로구	창신제2동
17	11110515		age_40	M	105881			1111069000	11110690	서울특별시	종로구	창신제3동
18	11110515		age_45	F	89461			1111070000	11110700	서울특별시	종로구	송인제1동
19	11110515		age_45	M	129379			1111071000	11110710	서울특별시	종로구	송인제2동

추출한 부분을 서울특별시_종로구_성연령별유동인구.csv 파일을 열어서 붙여넣기.

기존의 “행정동번호” 옆에 “동” 변수 제작 후 붙여넣은 “행정동코드” 옆에 “행정동번호” 변수 제작후 행정동코드/100 값으로 채워넣기.

B2												
	A	B	C	D	E	F	G	H	I	J	K	L
1	행정동번호	동	연령대구분	성별코드	인구수			행정동코드	행정동번호	시도명	시군구명	읍면동명
2	11110515	청운효자동	age_00	F	2094			1111000000	11110000	서울특별시	종로구	
3	11110515	청운효자동	age_00	M	2157			1111051500	11110515	서울특별시	종로구	청운효자동
4	11110515	청운효자동	age_10	F	7225			1111053000	11110530	서울특별시	종로구	사직동
5	11110515	청운효자동	age_10	M	8331			1111054000	11110540	서울특별시	종로구	삼청동
6	11110515	청운효자동	age_15	F	23591			1111055000	11110550	서울특별시	종로구	부암동
7	11110515	청운효자동	age_15	M	26828			1111056000	11110560	서울특별시	종로구	평창동
8	11110515	청운효자동	age_20	F	84681			1111057000	11110570	서울특별시	종로구	무악동
9	11110515	청운효자동	age_20	M	77724			1111058000	11110580	서울특별시	종로구	교남동
10	11110515	청운효자동	age_25	F	132750			1111060000	11110600	서울특별시	종로구	가회동
11	11110515	청운효자동	age_25	M	136121			1111061500	11110615	서울특별시	종로구	종로1,2,3,4
12	11110515	청운효자동	age_30	F	99996			1111063000	11110630	서울특별시	종로구	종로5,6가동
13	11110515	청운효자동	age_30	M	128780			1111064000	11110640	서울특별시	종로구	이화동
14	11110515	청운효자동	age_35	F	87371			1111065000	11110650	서울특별시	종로구	혜화동
15	11110515	청운효자동	age_35	M	115323			1111067000	11110670	서울특별시	종로구	창신제1동
16	11110515	청운효자동	age_40	F	76893			1111068000	11110680	서울특별시	종로구	창신제2동
17	11110515	청운효자동	age_40	M	105881			1111069000	11110690	서울특별시	종로구	창신제3동
18	11110515	청운효자동	age_45	F	89461			1111070000	11110700	서울특별시	종로구	송인제1동
19	11110515	청운효자동	age_45	M	129379			1111071000	11110710	서울특별시	종로구	송인제2동

동변수에 Vlookup함수 활용해서 행정동번호별 동 이름 채워 넣기.

이후 해당 동 변수 옆에 새로운 열 생성 후 값으로 붙여넣기.

이후 이전에 붙인 행정동코드~읍면동명 열 제거

E2					=C2&" "&D2				
	A	B	C	D	E	F	G	H	I
1	행정동번호동		연령대구분	성별코드	성연령	인구수			
2	11110515	청운효자동	age_00	F	age_00_F	2094			
3	11110515	청운효자동	age_00	M	age_00_M	2157			
4	11110515	청운효자동	age_10	F	age_10_F	7225			
5	11110515	청운효자동	age_10	M	age_10_M	8331			
6	11110515	청운효자동	age_15	F	age_15_F	23591			
7	11110515	청운효자동	age_15	M	age_15_M	26828			
8	11110515	청운효자동	age_20	F	age_20_F	84681			
9	11110515	청운효자동	age_20	M	age_20_M	77724			
10	11110515	청운효자동	age_25	F	age_25_F	132750			
11	11110515	청운효자동	age_25	M	age_25_M	136121			
12	11110515	청운효자동	age_30	F	age_30_F	99996			
13	11110515	청운효자동	age_30	M	age_30_M	128780			

성별코드 열 옆에 성연령 열 생성 후 연령대구분과 성별코드를 합친 변수 생성하기.
이후 새로운 열에 성연령을 값으로 붙여 넣은 후 연령대구분과 성별코드 열 제거

	A	B	C	D	E	F	G	H	I	J
1	행정동번호동		성연령	인구수						
2	11110515	청운효자동	age_00_F	2094						
3	11110515	청운효자동	age_00_M	2157						
4	11110515	청운효자동	age_10_F	7225						
5	11110515	청운효자동	age_10_M	8331						
6	11110515	청운효자동	age_15_F	23591						
7	11110515	청운효자동	age_15_M	26828						
8	11110515	청운효자동	age_20_F	84681						
9	11110515	청운효자동	age_20_M	77724						
10	11110515	청운효자동	age_25_F	132750						
11	11110515	청운효자동	age_25_M	136121						
12	11110515	청운효자동	age_30_F	99996						
13	11110515	청운효자동	age_30_M	128780						
14	11110515	청운효자동	age_35_F	87371						
15	11110515	청운효자동	age_35_M	115323						
16	11110515	청운효자동	age_40_F	76893						
17	11110515	청운효자동	age_40_M	105881						
18	11110515	청운효자동	age_45_F	89461						
19	11110515	청운효자동	age_45_M	129379						

선택하여 붙여넣기

붙여넣기

☒ 모두(A)
☐ 원본 테마 사용(H)
☐ 수식(E)
☐ 테두리만 제외(X)
☐ 값(V)
☐ 열 너비(W)
☐ 서식(I)
☐ 수식 및 숫자 서식(B)
☐ 메모(C)
☐ 값 및 숫자 서식(U)
☐ 유효성 검사(N)
☐ 조건부 서식 모두 병합(G)

연산

☒ 없음(O)
☐ 곱하기(M)
☐ 더하기(D)
☐ 나누기(I)
☐ 빼기(S)

☐ 내용 있는 셀만 붙여넣기(B)
☒ 행/열 바꿈(E)

연결하여 붙여넣기(L)

확인

취소

성연령과 인구수를 행/열 바꿈으로 E번 열에 붙여넣기.
이후 동별로 하나의 행만 남기고 이후 제거.
다음 동 인구수 복사 후 행/열 바꿔서 붙여넣기 반복.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	동	Flow_00_F	Flow_10_F	Flow_20_F	Flow_30_F	Flow_40_F	Flow_50_F	Flow_60_F	Flow_70_F	Flow_00_N	Flow_10_N	Flow_20_N	Flow_30_N
2	가회동	1099	17735	163563	145299	106951	102236	58679	20298	852	17361	121615	2446
3	교남동	3247	39353	234340	216784	201887	244302	145353	54096	2014	38263	188901	2556
4	무악동	2307	23555	107173	102785	105714	144143	95267	38348	1523	22940	105945	1123
5	부암동	3459	52066	205831	210900	221411	275813	150842	58089	3175	43859	225561	3113
6	사직동	3245	58309	641481	601051	443061	393182	212146	75018	2402	44617	476943	7697
7	삼청동	806	13647	120371	114500	86060	89313	46758	15615	906	12515	108385	1444
8	송인제1동	1786	26234	148532	123812	120924	151449	90801	37648	1913	24461	137958	1657
9	송인제2동	2565	32076	275373	230331	225512	270423	161804	59746	1849	32503	236947	3097
10	이화동	4523	81642	821255	576035	380374	392262	242900	96446	3453	57177	586229	6119
11	종로1.2.3.4	5223	103695	1543479	1107977	750779	761057	459552	170419	4284	79515	1237276	16389
12	종로5.6가동	2789	39748	469202	366410	301219	386375	260346	96942	1908	34234	373612	5078
13	창신제1동	1527	20617	205315	173283	156317	216387	136222	46072	1392	21517	194951	2595
14	창신제2동	881	14802	62045	42055	44201	74654	42747	17758	768	12532	61690	592
15	창신제3동	919	8212	46901	43345	36920	50080	27388	12406	580	9371	44055	553
16	청운효자동	2094	30816	217431	187367	166354	170222	89480	36362	2156	35159	213845	2441
17	평창동	6413	60184	195208	215116	242252	315380	181626	65390	4796	54372	230572	3289
18	해화동	3490	83967	586653	279324	226580	237003	135364	58926	2574	75569	485272	2963

성연령별 유동인구수 변수 새로 생성

ex) Flow_10_F = age_10_F+age_15_f -> 10대 유동인구

이후 기존 성연령 변수 제거.

결과물을 서울특별시_종로구_성연령별유동인구.csv 파일로 저장하기.

2-2. 공공데이터

<서울특별시_종로구_성연령별거주인구.csv 만들기>

http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=INH_1B04005N_11&vw_cd=MT_ZTITLE&list_id=201_20103&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE

일괄설정 +

항목 [2/3]

행정구역(동읍면)별 [...]

5세별 [21/22]

시점 [1/107] [...]

현재까지 선택된 데이터는 항목(2) × 분류(18×21) × 시점(1) = 756 셀 [20,000 (조회) | 200,000 (다운로드) | 다운로드 불가]

선택:

▶ 행정구역(동읍면)별 [전체레벨 3]

1 레벨 전체선택

☐ 서울특별시

2 레벨 전체선택 ?

☐ 서울특별시

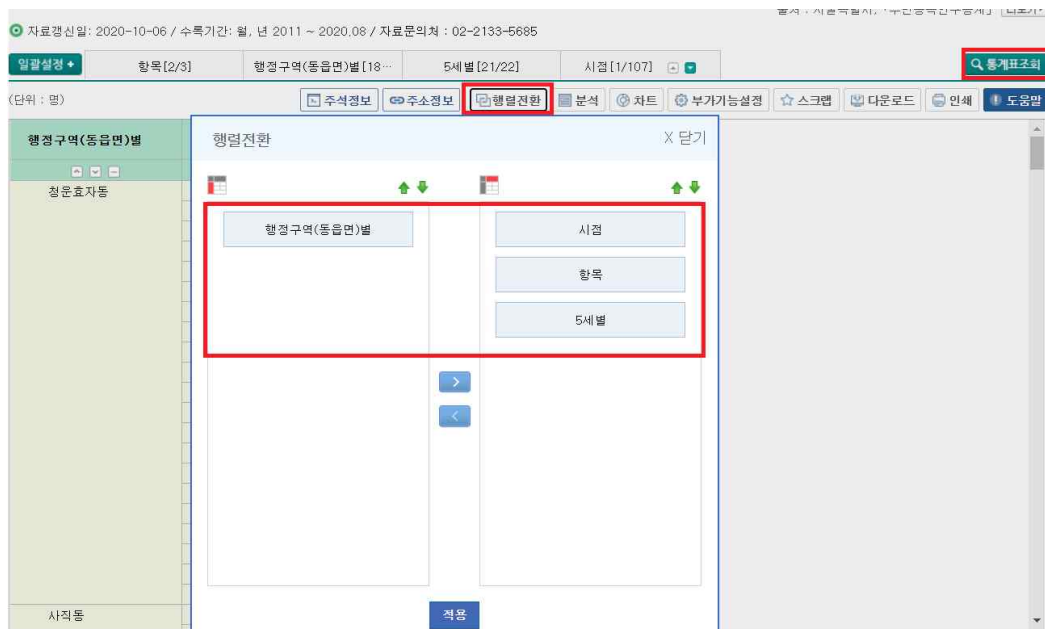
☐ 종로구
☐ 중구
☐ 용산구
☐ 성동구
☐ 광진구
☐ 동대문구
☐ 중랑구

3 레벨 전체선택 ?

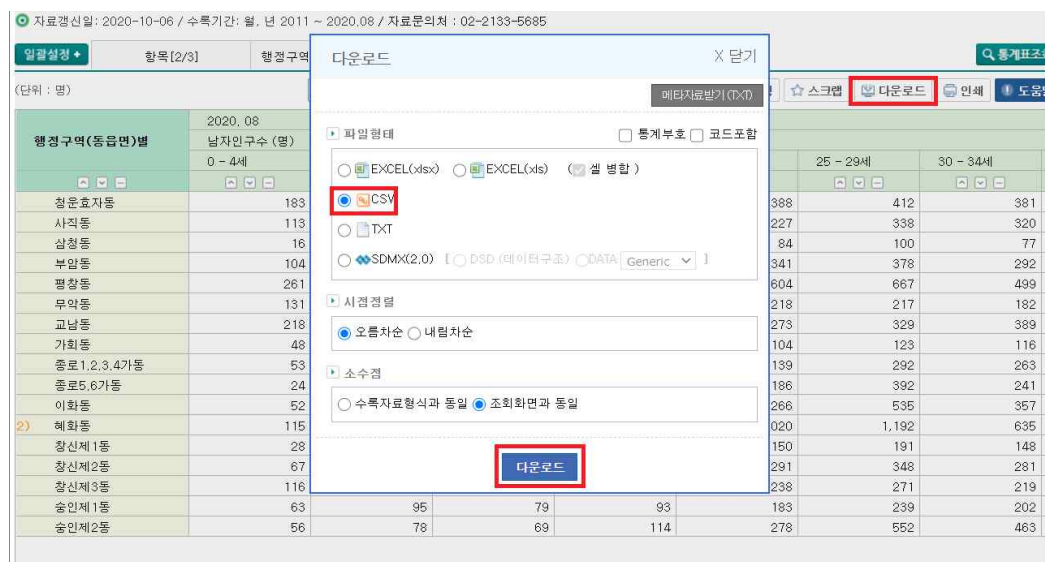
☒ 종로구

☒ 청운효자동
☒ 사직동
☒ 삼청동
☒ 부암동
☒ 평창동
☒ 무악동
☒ 교남동

행정구역에서 종로구에 해당하는 동 모두 선택.



통계표 조회후 행렬전환에서 위와 같이 설정후 적용.



결과물 csv파일로 다운로드 받기.

서울특별시_읍면동별_5세별_주민등록인구_다운로드받은 시간.csv 으로 저장 됨

D4												
	A	B	C	D	E	F	G	H	I	J	K	L
1	행정구역(시/군/구)	2020. 08	2020. 08		2020. 08	2020. 08		2020. 08	2020. 08		2020. 08	2020. 08
2	행정구역(시/군/구)	남자인구수	남자인구수 (명)		남자인구수	남자인구수 (명)		남자인구수	남자인구수 (명)		남자인구수	남자인구수 (명)
3	행정구역(시/군/구)	0 - 4세	5 - 9세	male_00	10 - 14세	15 - 19세	male_10	20 - 24세	25 - 29세	male_20	30 - 34세	35 - 39세
4	청운효자동	183	269	452	284	344		388	412		381	
5	사직동	113	168	281	167	154		227	338		320	
6	삼청동	16	29	45	60	58		84	100		77	
7	부암동	104	185	289	245	249		341	378		292	
8	평창동	261	357	618	367	482		604	667		499	
9	무악동	131	188	319	223	250		218	217		182	
10	교남동	218	243	461	222	194		273	329		389	
11	가회동	48	58	106	92	88		104	123		116	
12	종로1,2,3,4가	53	51	104	60	58		139	292		263	
13	종로5,6가	24	35	59	40	70		186	392		241	
14	이화동	52	80	132	95	113		266	535		357	
15	해화동	115	227	342	311	384		1020	1192		635	
16	창신제1동	28	46	74	47	80		150	191		148	
17	창신제2동	67	76	143	100	164		291	348		281	
18	창신제3동	116	98	214	122	140		238	271		219	
19	송인제1동	63	95	158	79	93		183	239		202	
20	송인제2동	56	78	134	69	114		278	552		463	

1, 2 번 행 제거 해주고 연령별로 사이사이에 위와 같이 male_00, male_10 변수 새로 만들어줌 female_70 까지 반복 70~100세이상 까지는 70대이상으로 합쳐준다.
 새로 만들어진 변수에는 10대의 경우 10-14세와 15-19세를 합치는 식으로 제작 결과물을 서울특별시_종로구_성연령별거주인구.csv 파일로 저장

2-3. 결과도출용 데이터 제작

	A	B	C	D	E	AL	AM	AN	AO	AP
1	아파트명	아파트주소동		아파트세디	아파트면적값					
2	금강파크빌리지	서울시 중구 평창동		168	84					
3	갑을	서울시 중구 평창동		99	30.46					
4	형우럭스빌	서울시 중구 평창동		21	103.07					
5	평창롯데캐슬로	서울시 중구 평창동		112	190.74					
6	르트랑시망 1동	서울시 중구 평창동		19	135.17					
7	동대문맨션	서울시 중구 창신제2동		56	95					
8	두산	서울시 중구 창신제2동		529	86					
9	동대문상가	서울시 중구 창신제1동		120	49.75					
10	창신쌍용 1단지	서울시 중구 창신제3동		585	79.87					
11	창신쌍용 2단지	서울시 중구 창신제3동		919	106.62					
12	종로1가대성스키	종로1가대	종로1,2,3,4	54	137.55					
13	로얄팰리스스위	서울시 중구 종로1,2,3,4		21	36.85					
14	운현신화타워	서울특별시 종로1,2,3,4		64	47.61					
15	낙원	서울특별시 종로1,2,3,4		146	110.15					
16	현대드레비앙	서우리 중구 종로1,2,3,4		291	45.5					
17	도시그린빌라트	서울시 중구 송인제1동		10	66.96					

종로구에 부동산 매물로 올라온 아파트들중 일부 선택하여 result1.csv 파일과 동일한 형태로 제작하기.

data_new.csv 로 저장하기

2-4. 데이터 병합

Excel spreadsheet showing data for apartment listings and demographic information. The formula bar shows: `=VLOOKUP(C2:C1216,B2:B18,3,FALSE)`

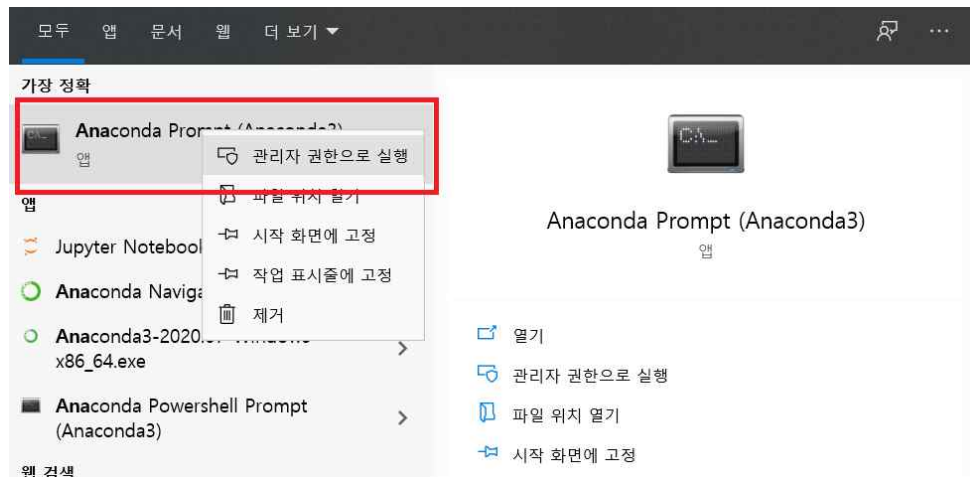
	A	B	C	D	E	F	G	H	I	AP	AQ	BH	BI	BJ	BK	BL
1	아파트명	아파트주소	아파트세대	아파트동수	아파트면적	최근매출	매매일반표	Y		male_0	male_10	행정구역(시/군/구)	male_0	male_10	male_20	male_30
2	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214	262	성년영별거주인구	452	628	800	804
3	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		사직동	281	321	565	622
4	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		삼청동	45	118	184	176
5	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		부암동	289	494	719	573
6	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		평창동	618	849	1271	1055
7	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		무악동	319	473	435	437
8	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		교남동	461	416	602	829
9	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		가회동	106	180	227	279
10	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.1E+08	4.13E+08	0	214		종로1.2.3.4	104	118	431	518
11	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.25E+08	0	214		종로5.6가	59	110	578	412
12	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.3E+08	0	214		이화동	132	208	801	600
13	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.35E+08	0	214		혜화동	342	695	2212	1122
14	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.4E+08	0	214		창신제1동	74	127	341	308
15	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.3E+08	0	214		창신제2동	143	264	639	531
16	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.33E+08	0	214		창신제3동	214	262	509	479
17	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.33E+08	0	214		송인제1동	158	172	422	438
18	창신쌍용2	서울특별시	창신제3동	156	7	54.7	4.17E+08	4.4E+08	0	214		송인제2동	134	183	830	833
19	창신쌍용2	서울특별시	창신제3동	284	7	64.66	4.78E+08	4.55E+08	1	214						
20	창신쌍용2	서울특별시	창신제3동	284	7	64.66	4.78E+08	4.55E+08	1	214						

- ①result1.csv 파일에 서울특별시_종로구_성연령별거주인구.csv 파일의 내용을 복사해서 붙여 넣는다.
- ②붙여 넣은 내용에서 성연령 변수들의 이름과 동일한 변수들을 만들어 준다.
- ③해당 변수의 값을 vlookup함수를 활용해서 동별로 채워 넣는다.
- ④male_0부터 female_70까지 반복한다.
- ⑤이후 작업한 값을 새로운 열에 값으로 붙여 넣기 후 이전에 서울특별시_종로구_성연령별거주인구.csv에서 가져온 열들을 제거한다.
- ⑥저장 한 후에 같은 작업을 서울특별시_종로구_성연령별거주인구.csv에서 서울특별시_종로구_성연령별유동인구.csv 파일로 변경한 후 반복한다.
- ⑦결과파일을 result1.csv 로 저장한다.

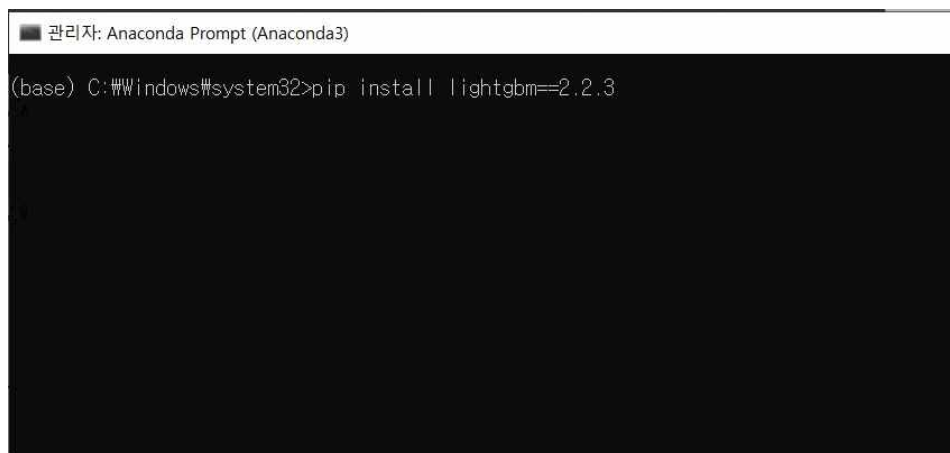
①~⑦작업에서 result1.csv 파일을 result2.csv 와 data_new.csv 로 바꾼후 똑같이 반복해 준다.

3. 모델링

3-1. 라이트쥬비엠 설치

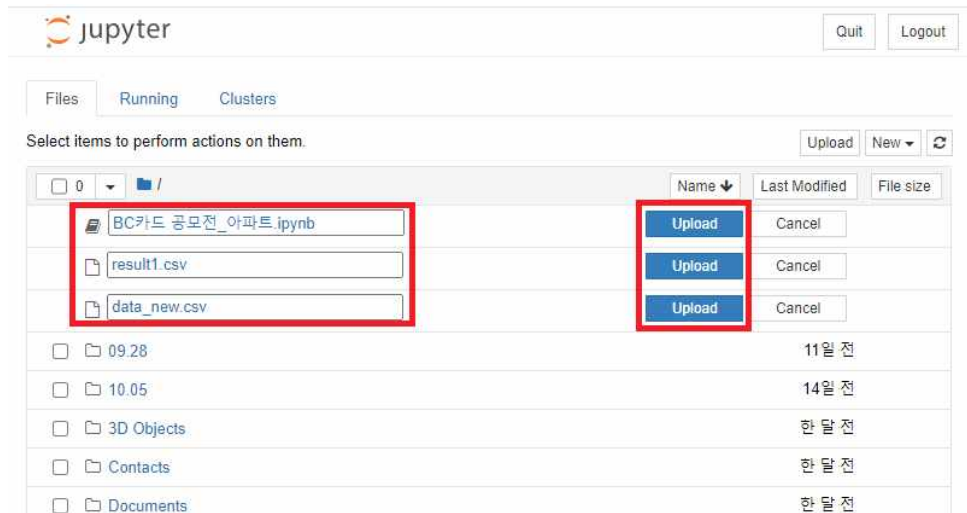


Anaconda Prompt 창 관리자권한으로 실행

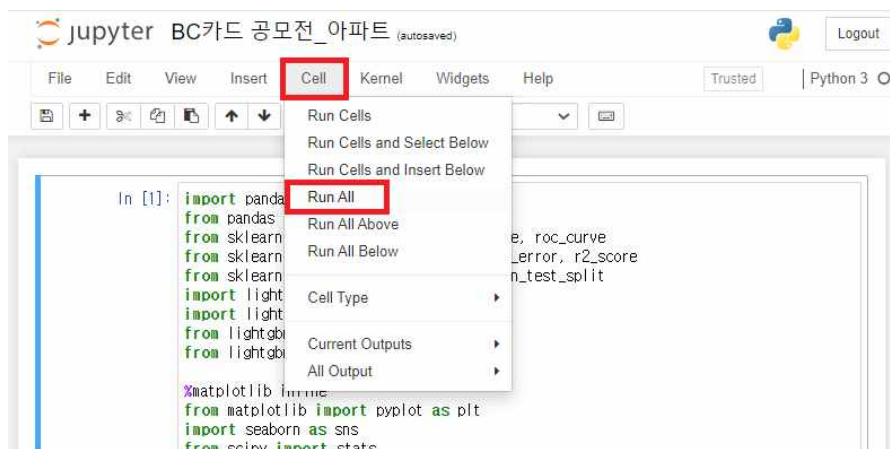


pip install lightgbm==2.2.3 입력 (2.2.3 버전을 설치해야 한다.)

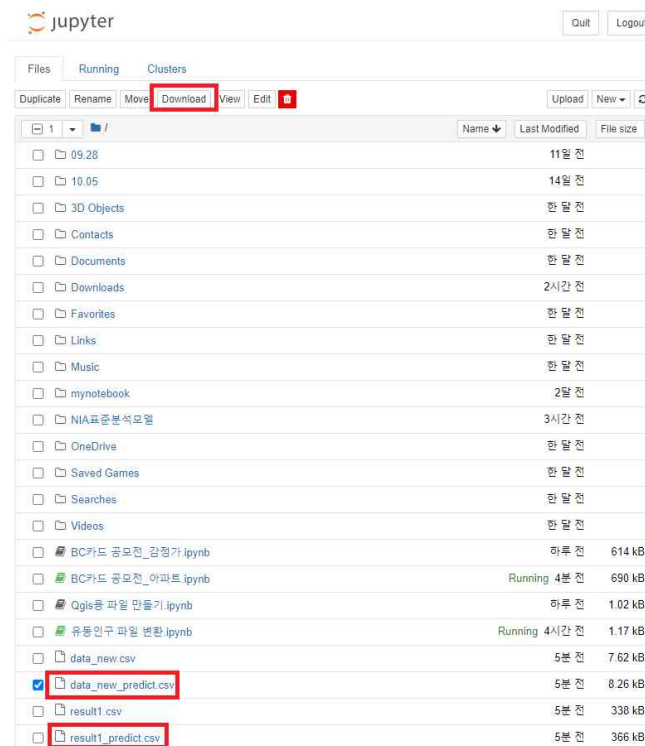
3-2. 아파트 모델



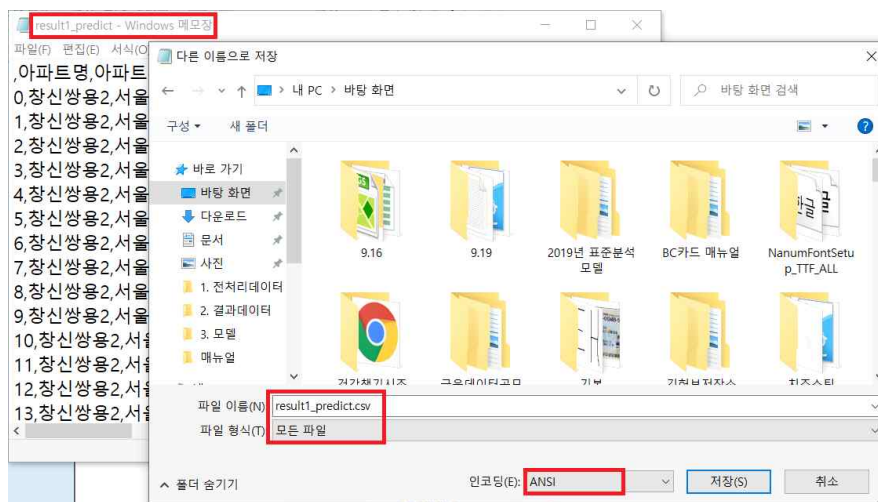
주피터 노트북을 실행후 홈페이지에 result1.csv, data_new.csv 파일과 BC카드_공모전_아파트.ipynb를 업로드 해준다.(ipynb파일이 업로드 안될 경우 새 파이썬 파일을 만든후 html파일을 열어서 복사 붙여넣기 해주기)



BC카드_공모전_아파트 코드를 열어서 Cell의 Run All로 실행해 준다.



주피터 노트북 홈페이지에 생성된 data_new_predict.csv 파일과 result1_predict.csv 파일을 각각 다운로드 받아준다.



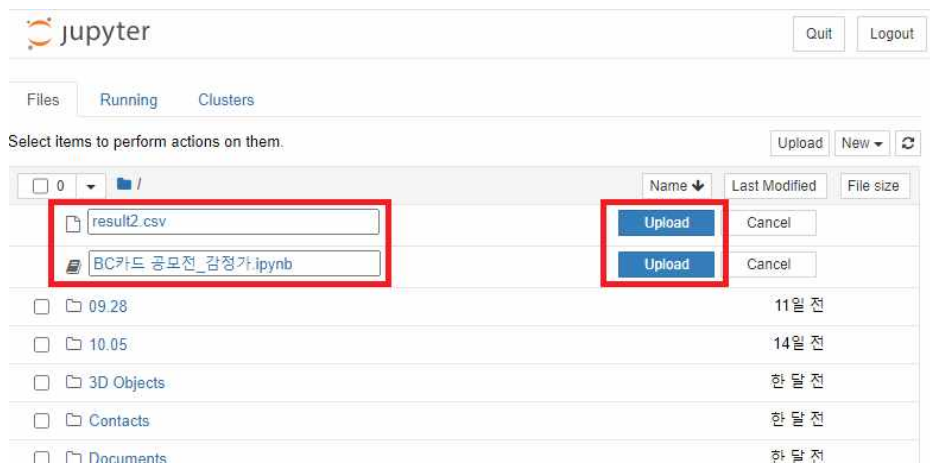
①result1_predict.csv 파일을 메모장으로 읽어온후 인코딩을 ANSI로 설정하고 result1_predict.csv로 저장해 준다.

	A	B	C	D	E	F	AN	AO	AP	AQ
1		아파트명	아파트주소동		아파트세디	아파트동수	Flow_50_M	Flow_60_M	Flow_70_M	Predict
2	0	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
3	1	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
4	2	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
5	3	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
6	4	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
7	5	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
8	6	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
9	7	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
10	8	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
11	9	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
12	10	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
13	11	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
14	12	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
15	13	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903
16	14	창신쌍용2	서울특별시	창신제3동	156	7	56397	40368	15657	0.245903

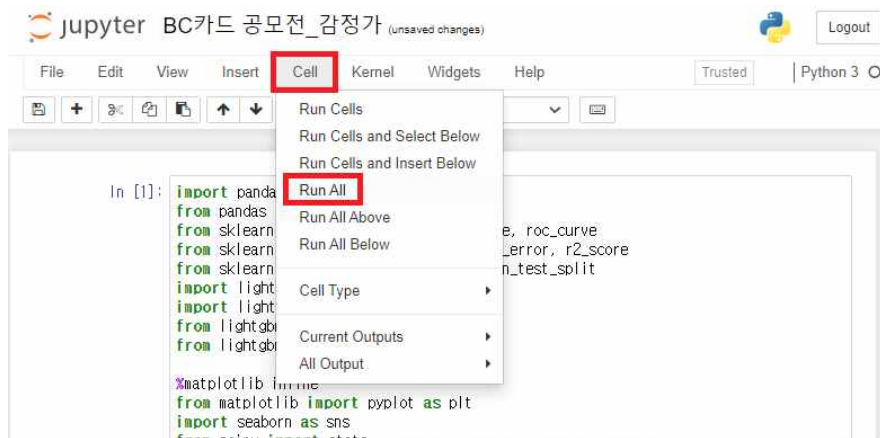
②result1_predict.csv 파일을 열어서 제일 첫 번째 열을 제거하고 가장 끝에 있는 열의 이름이 '0'으로 되어있는데 이를 Predict로 바꿔준다.

①, ② 과정(이하 인코딩 변환과정)을 data_new_predict.csv파일로도 진행한다.

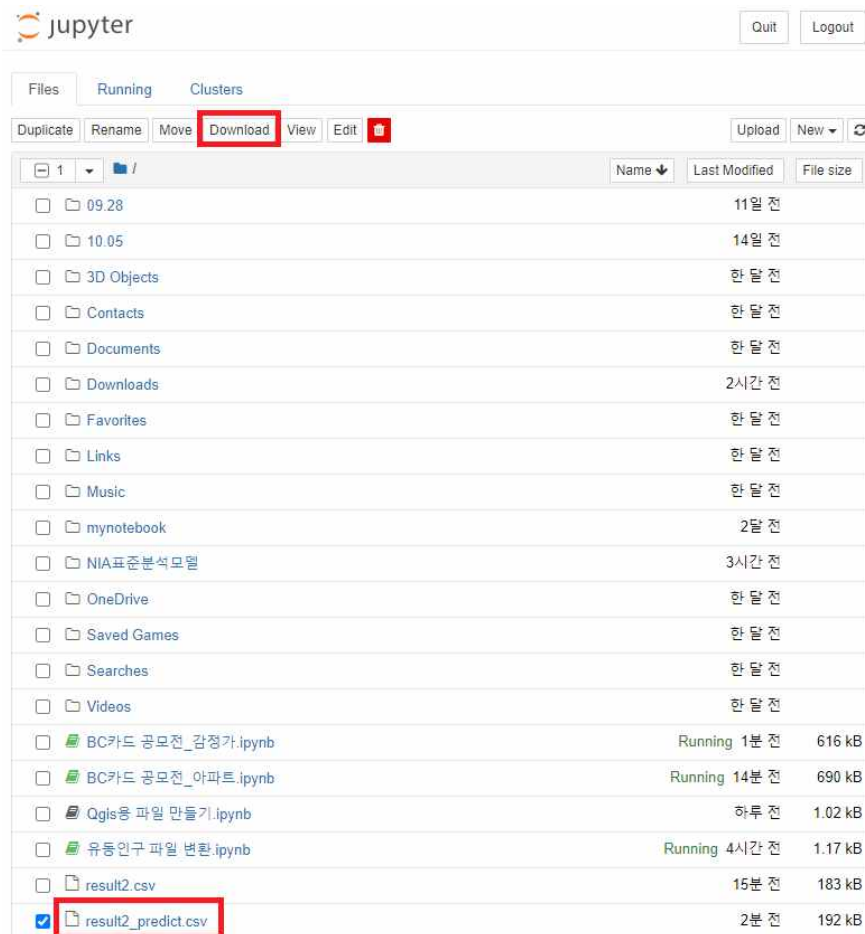
3-3. 감정가 모델



주피터 노트북을 실행후 홈페이지에 result2.csv 파일과 BC카드_공모전_감정가.ipynb를 업로드 해준다.(ipynb파일이 업로드 안될 경우 새 파이썬 파일을 만든후 html파일을 열어서 복사 붙여넣기 해주기)



BC카드_공모전_감정가 코드를 열어서 Cell의 Run All로 실행해 준다.

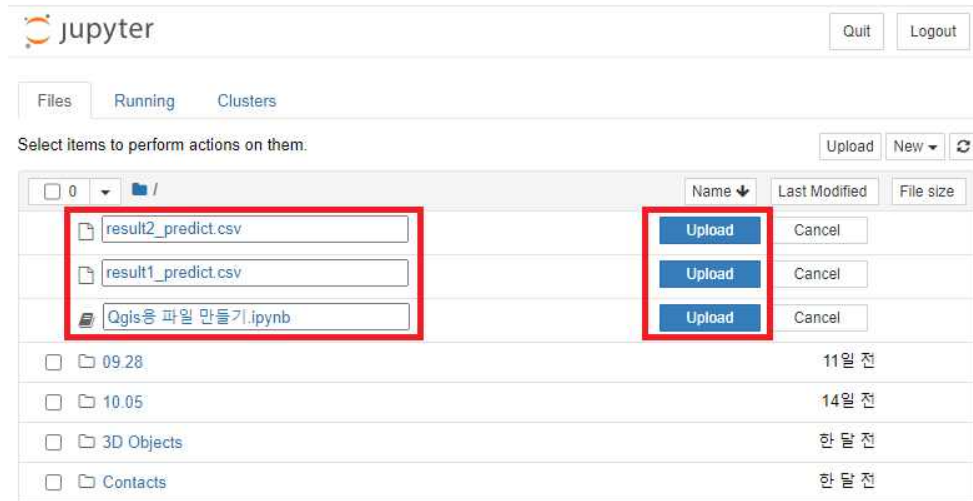


주피터 노트북 홈페이지에 생성된 result2_predict.csv파일을 다운로드 받아준다.
이전의 인코딩 변환과정을 result2_predict.csv 파일로도 진행한다.

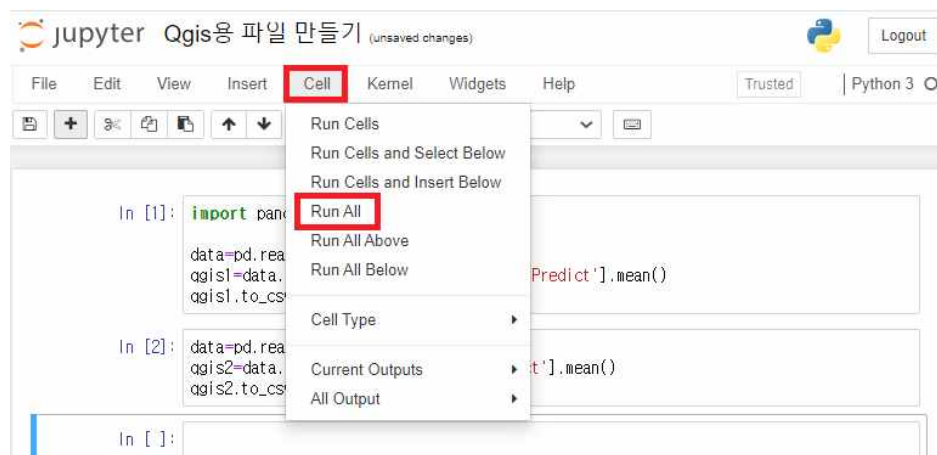
4. Qgis시각화

4-1. Qgis적용 파일 제작

<좌표값 생성하기 위한 파일 제작>



주피터 노트북을 실행하여 result1_predict.csv, result2_predict.csv, Qgis용 파일 만들기.ipynb 파일을 업로드 해준다.



Qgis용 파일 만들기 코드를 Cell - Run All 로 실행시켜 준다.

jupyter

Quit Logout

Files Running Clusters

Duplicate Rename Move Download View Edit

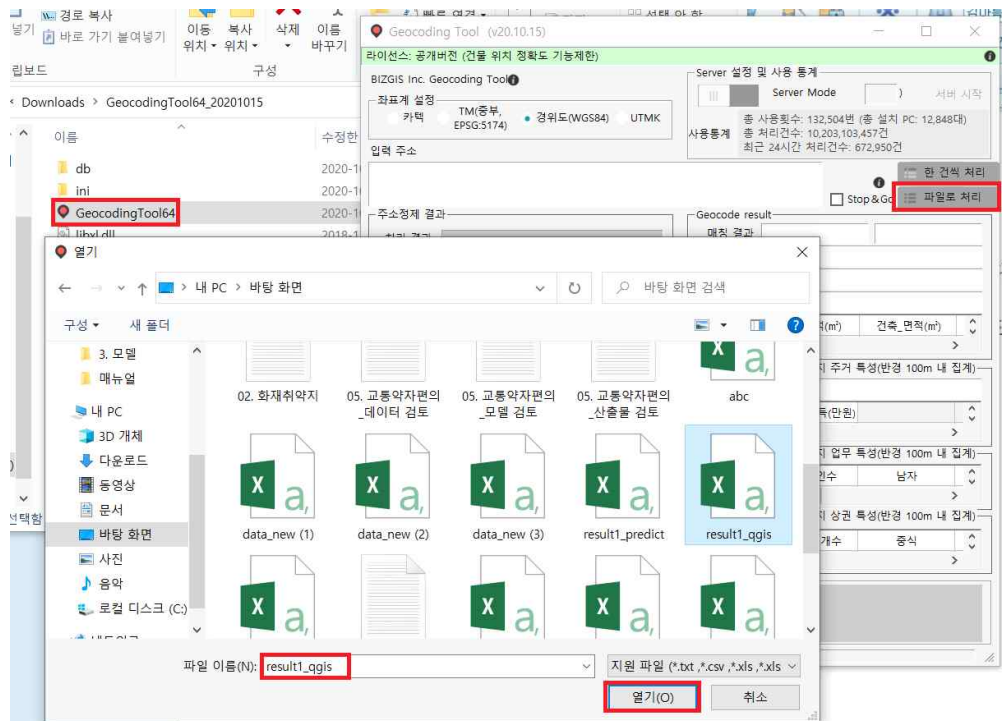
Upload New

	Name	Last Modified	File size
<input type="checkbox"/>	09.28	11일 전	
<input type="checkbox"/>	10.05	14일 전	
<input type="checkbox"/>	3D Objects	한 달 전	
<input type="checkbox"/>	Contacts	한 달 전	
<input type="checkbox"/>	Documents	한 달 전	
<input type="checkbox"/>	Downloads	15분 전	
<input type="checkbox"/>	Favorites	한 달 전	
<input type="checkbox"/>	Links	한 달 전	
<input type="checkbox"/>	Music	한 달 전	
<input type="checkbox"/>	mynotebook	2달 전	
<input type="checkbox"/>	NIA표준분석모델	4시간 전	
<input type="checkbox"/>	OneDrive	한 달 전	
<input type="checkbox"/>	Saved Games	한 달 전	
<input type="checkbox"/>	Searches	한 달 전	
<input type="checkbox"/>	Videos	한 달 전	
<input type="checkbox"/>	Qgis종 파일 만들기.ipynb	Running 몇 초 전	1.13 kB
<input type="checkbox"/>	result1_predict.csv	2분 전	353 kB
<input checked="" type="checkbox"/>	result1_qgis.csv	1분 전	1.33 kB
<input type="checkbox"/>	result2_predict.csv	2분 전	188 kB
<input type="checkbox"/>	result2_qgis.csv	1분 전	1.64 kB

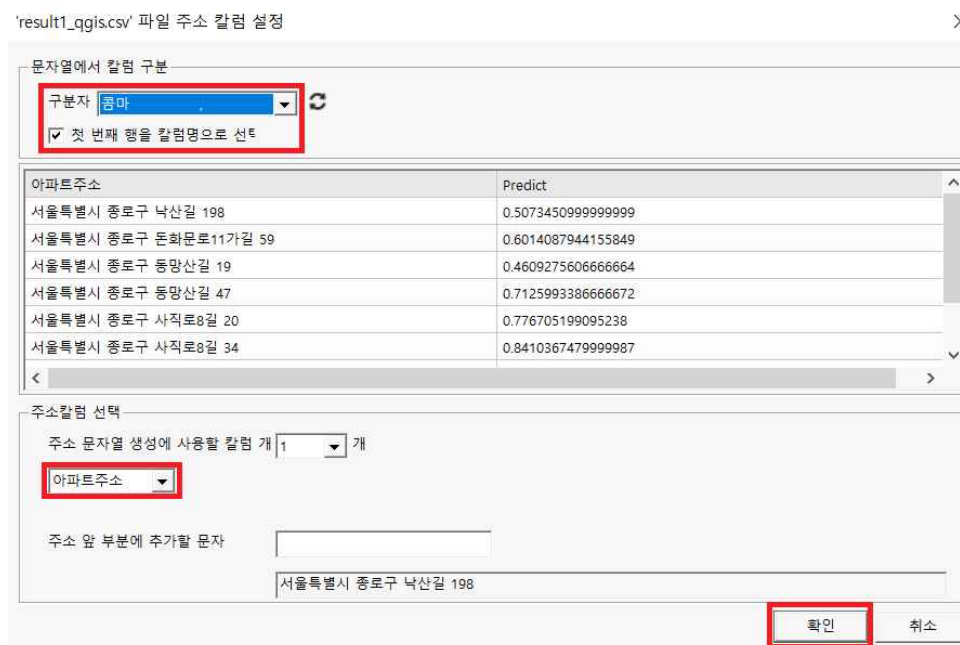
주피터 노트북 홈페이지 에서 result1_qgis.csv와 result2_qgis.csv 파일을 각각 다운로드 받아준다.

이후 인코딩 변환과정을 result1_qgis.csv와 result2_qgis.csv 파일에 대해 각각 진행한다.

<Geocoding으로 좌표값 생성하기>



①설치한 GeocodingTool64를 실행한 후 파일로 처리를 클릭하여 result1_qgis.csv 파일을 열어준다.



②옵션들을 확인한 이후 확인을 눌러준다

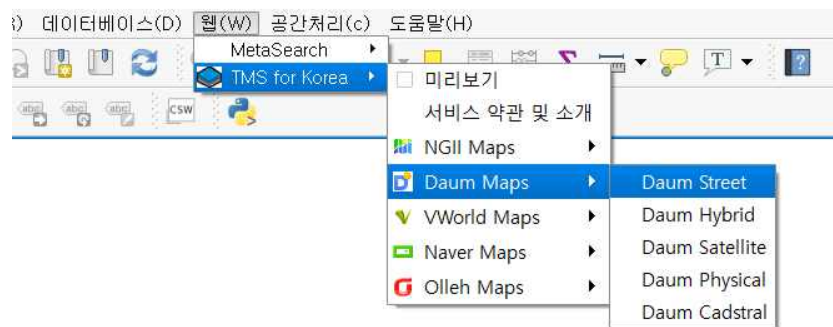
	A	B	C	D	E	F	G	H	I	J	K
1	아파트주소	Predict	입력주소	X	Y	CLSS	PNU	주소구분	표준신주소	표준구주소	우편번호
2	서울특별시	0.507345	서울특별시	127.0117	37.58031	정좌표	공개용버전 새주소	미지원	미지원		3092
3	서울특별시	0.601409	서울특별시	126.9901	37.57505	정좌표	공개용버전 새주소	미지원	미지원		3132
4	서울특별시	0.460928	서울특별시	127.0139	37.5804	정좌표	공개용버전 새주소	미지원	미지원		3093
5	서울특별시	0.712599	서울특별시	127.0161	37.5809	정좌표	공개용버전 새주소	미지원	미지원		3107
6	서울특별시	0.776705	서울특별시	126.9708	37.57428	정좌표	공개용버전 새주소	미지원	미지원		3174
7	서울특별시	0.841037	서울특별시	126.9718	37.57321	정좌표	공개용버전 새주소	미지원	미지원		3174
8	서울특별시	0.685626	서울특별시	126.9689	37.57446	정좌표	공개용버전 새주소	미지원	미지원		3168
9	서울특별시	0.704461	서울특별시	126.9722	37.57255	정좌표	공개용버전 새주소	미지원	미지원		3174
10	서울특별시	0.839751	서울특별시	127.0212	37.57593	정좌표	공개용버전 새주소	미지원	미지원		3112
11	서울특별시	0.390277	서울특별시	126.9987	37.5715	정좌표	공개용버전 새주소	미지원	미지원		3130
12	서울특별시	0.547572	서울특별시	127.0161	37.57359	정좌표	공개용버전 새주소	미지원	미지원		3113
13	서울특별시	0.939081	서울특별시	127.0152	37.57227	정좌표	공개용버전 새주소	미지원	미지원		3121
14	서울특별시	0.288209	서울특별시	127.0148	37.57791	정좌표	공개용버전 새주소	미지원	미지원		3095
15	서울특별시	0.684844	서울특별시	127.0141	37.57409	정좌표	공개용버전 새주소	미지원	미지원		3104
16	서울특별시	0.894731	서울특별시	126.9994	37.58553	정좌표	공개용버전 새주소	미지원	미지원		3075
17	서울특별시	0.680231	서울특별시	126.9608	37.57541	정좌표	공개용버전 새주소	미지원	미지원		3025

③생성된 파일중 result1_qgis.add 파일을 엑셀을 실행하여 열어준 다음에 X, Y 좌표 이후에 생성된 필요없는 열을 제거한다. (여기서는 F~K열)

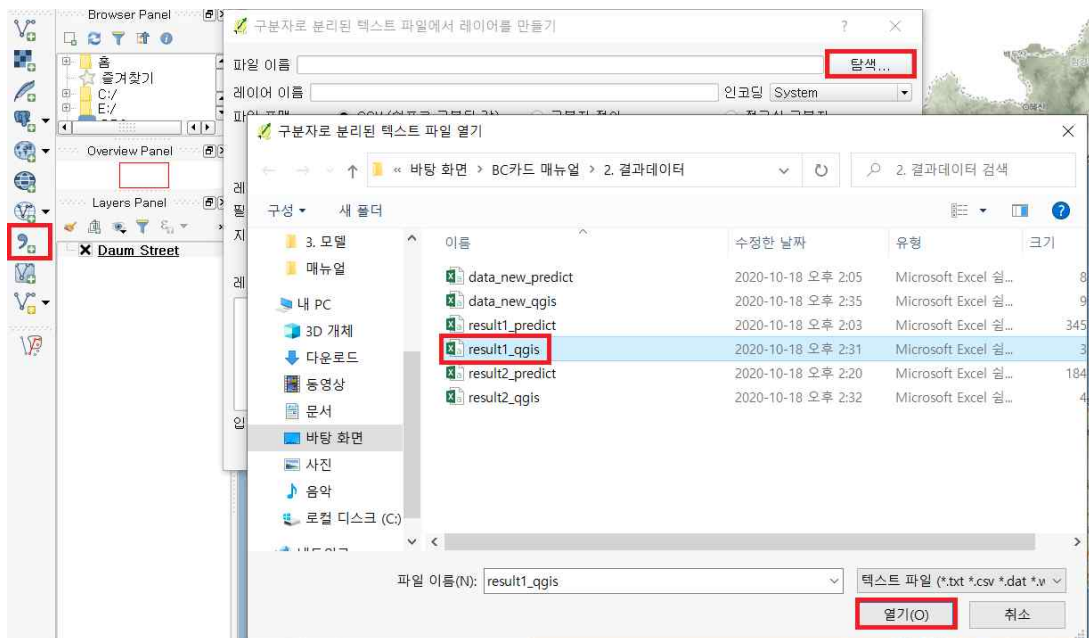
④결과물을 result1_qgis.csv 파일로 저장한다.

①~④의 과정을 result1_qgis.csv 파일과 data_new_predict.csv 파일에 대해서도 진행한다.
(data_new_predict.csv의 경우 ④에서 결과물을 data_new_qgis.csv 파일로 저장한다.)

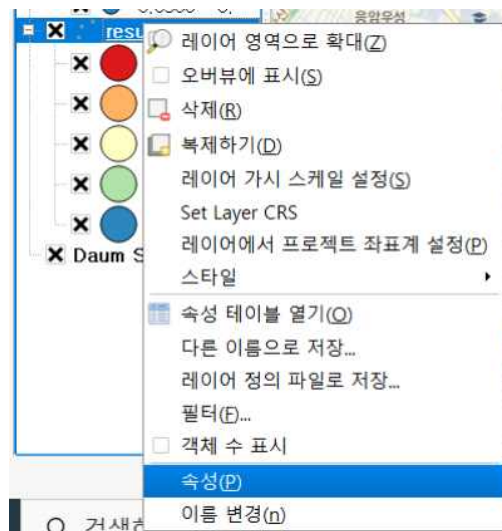
4-2. 모델 시각화



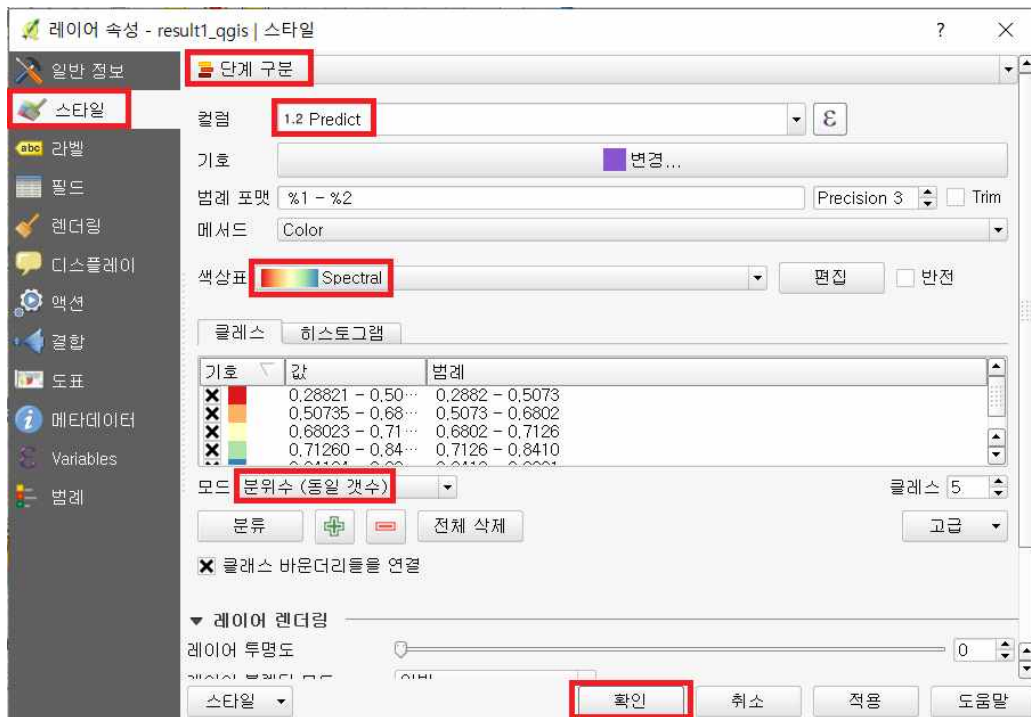
①설치한 Qgis를 실행한후 위와 같이 Daum Street Layer을 추가해 준다.



②“구분자로 분리된 텍스트 레이어를 추가”를 클릭한 이후 탐색에서 result1_qgis.csv 파일을 열어준다. 이 과정을 result2_qgis.csv, data_new_qgis.csv 파일에 대해서도 반복해 준다.



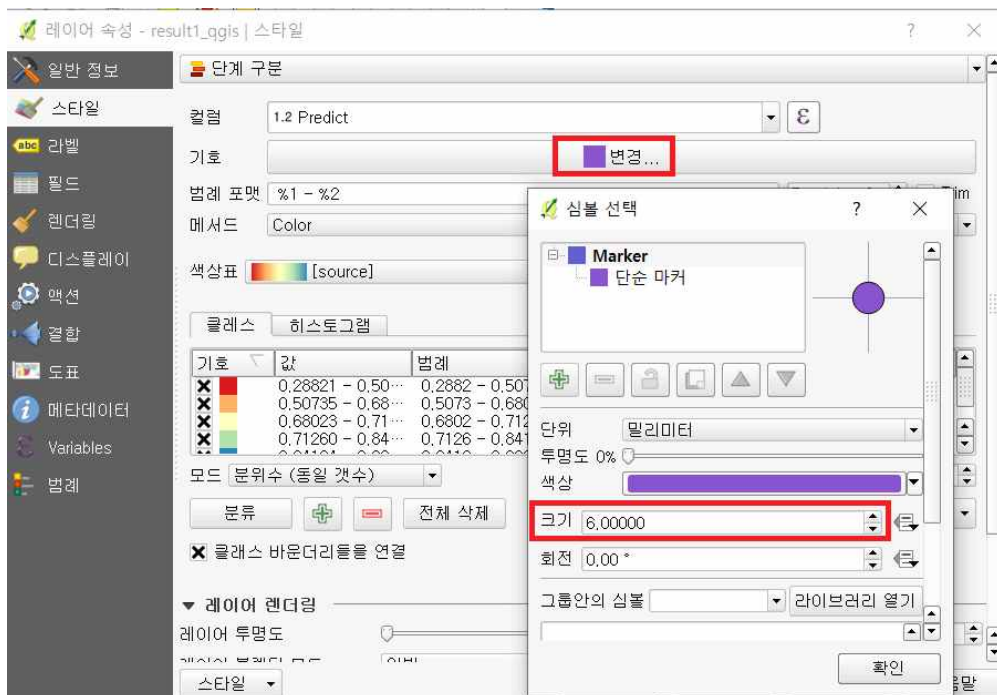
③result1_qgis 레이어를 우클릭 하여 속성을 선택한다.



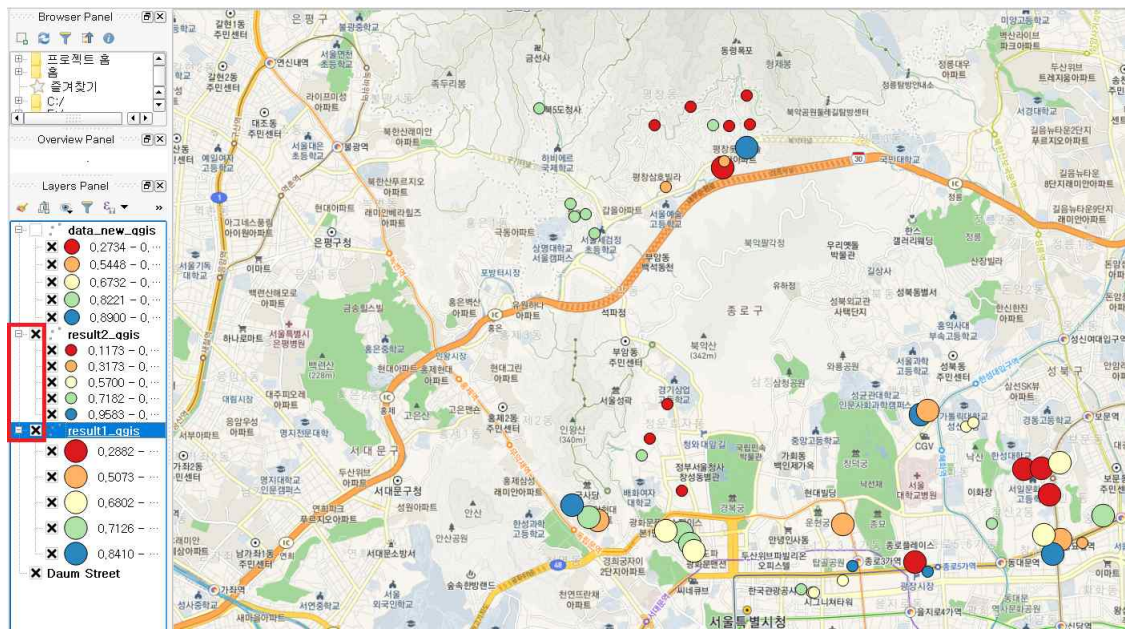
④스타일에서 옵션을 위와 같이 적용한다.

③, ④ 과정을 result2_qgis, data_new_qgis 레이어에 대해서도 반복해 준다.

4-3. 두 모델 비교

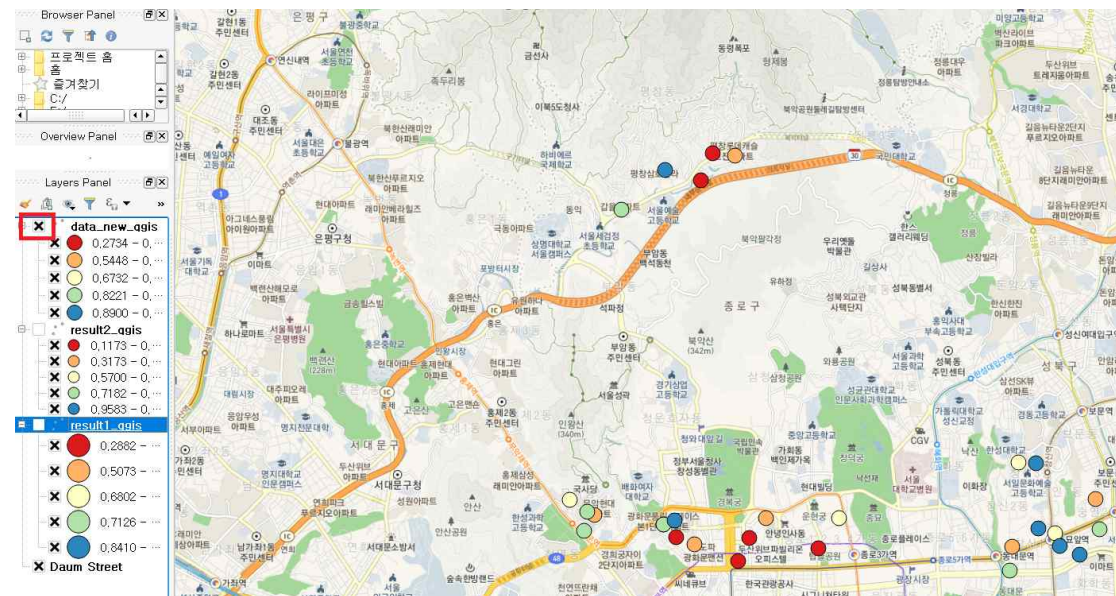


result1_qgis의 속성에서 기호 변경에서 크기를 6으로 지정한다.



result1_qgis 와 result2_qgis 레이어를 선택하여 비교 한다.

4-4. 아파트모델 결과도출용 데이터 적용 시각화



data_new_qgis 레이어를 선택한다.