

광주광역시 광산구 아파트 매매가격 예측

생각만하다 끝나조_(호남대)

윤성호, 이상현, 이수완, 정윤희, 정혜수(발표자)

목차

1. 주제 선정
2. 데이터 수집
3. 시각화
4. 데이터 전처리
5. 모델링(회귀)
6. 결론 및 아쉬운 점



1. 주제 선정

2030세대 저축을 하는 가장 큰 이유

내년 집값 어떻게 될까...전국 매매가 하락세, 수도권은 보합

김경민 기자 | 입력 : 2018.12.13 10:08:23

2019년 집값은 올해처럼 급등할까, 하락세로 돌아설까.

주요 연구기관이 내년 집값 전망치를 내놓으면서 실수요자 관심이 뜨겁다. 대체로 비관적인 예측이 많은 가운데 서울 집값은 강세, 약세 전망이 엇갈린다.

경조사비

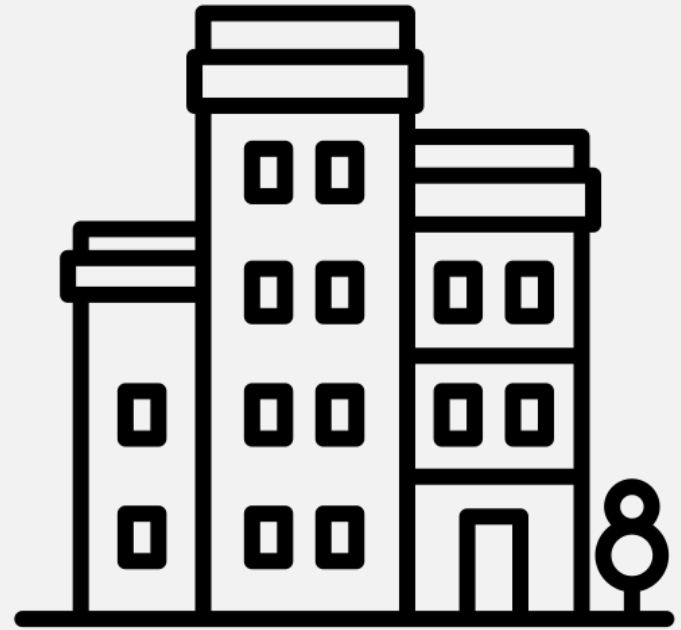
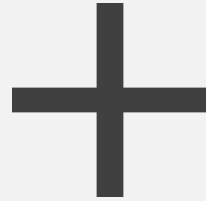
원문보기 | 경제>국제경제 | 경제>금융_재테크 | 경제>서비스_쇼핑 2019-07-17 하현옥(hyunock@joongang.co.kr)

가 가 ★스크랩 뉴스듣기 f t y

대부분 젊은 직장인이 돈을 모으려는 가장 큰 이유는 바로 '주택마련'입니다.

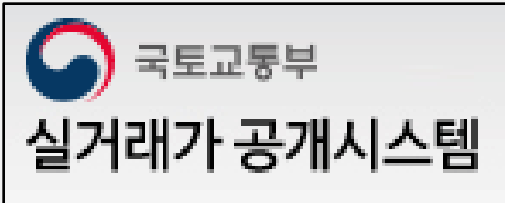
1. 주제 선정

-4-



광주광역시 광산구 일대 아파트매매 가격예측

2. 데이터 수집



24000개의 광주광역시 광산구 아파트 매매가격 데이터 수집



아파트 가격을 결정하는 데 영향을 미치는 요인을 컬럼으로 추가

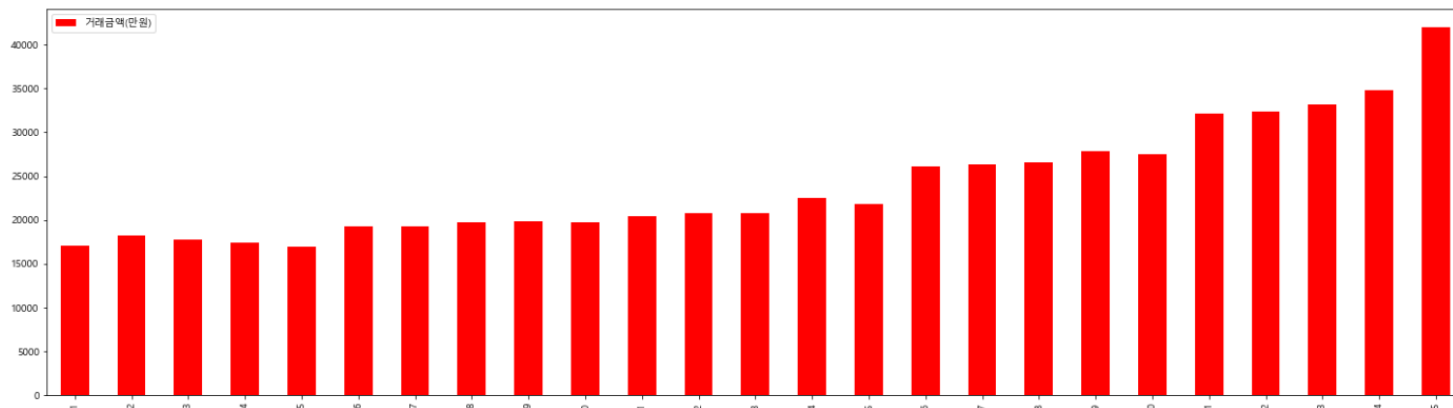
2. 데이터 수집

시군구	번지	본번	부번	단지명	전용면적(m²)	계약년월	계약일	거래금액(만원)	층	건축년도	도로명
광주광역시 광산구 도산동	1138.2	1138		2 대덕1	59.31	201907	20	7650	8	1990	송도로
광주광역시 광산구 도산동	790	790		0 대주피오레1차	84.98	201907	1	17900	11	2000	송도로182번길
광주광역시 광산구 도산동	790	790		0 대주피오레1차	84.98	201907	18	17500	3	2000	송도로182번길
광주광역시 광산구 도산동	1138.1	1138		1 무등파크맨션	59.88	201907	9	9000	7	1994	도산로
광주광역시 광산구 도산동	56.5	56		5 보경해피드림	84.9546	201907	10	18800	13	2006	송도로212번길
광주광역시 광산구 도산동	1304.1	1304		1 삼라미주	31.395	201907	24	5000	3	1997	남동길
광주광역시 광산구 도산동	1136.1	1136		1 호반1	59.95	201907	12	10150	5	1995	도산로9번길
광주광역시 광산구 도산동	782	782		0 호반2차	59.95	201907	4	13400	11	1998	송도로162번길
광주광역시 광산구 도산동	782	782		0 호반2차	59.95	201907	11	12200	7	1998	송도로162번길
광주광역시 광산구 도산동	782	782		0 호반2차	59.95	201907	18	12300	1	1998	송도로162번길
광주광역시 광산구 도산동	782	782		0 호반2차	59.95	201908	5	13000	7	1998	송도로162번길
광주광역시 광산구 도산동	840.1	840		1 호반청암빌라트	49.65	201907	5	7000	6	1993	도산로12번길
광주광역시 광산구 도천동	627.1	627		1 중흥1차	59.94	201907	6	9000	10	1995	비아로62번길
광주광역시 광산구 도천동	627.1	627		1 중흥1차	59.94	201907	6	9800	12	1995	비아로62번길
광주광역시 광산구 도천동	627.1	627		1 중흥1차	59.94	201907	16	8900	9	1995	비아로62번길
광주광역시 광산구 도천동	628.1	628		1 중흥2차	59.76	201907	15	8900	14	1995	비아로62번길
광주광역시 광산구 도천동	628.1	628		1 중흥2차	59.76	201907	22	8800	12	1995	비아로62번길
광주광역시 광산구 도천동	628.1	628		1 중흥2차	59.76	201907	22	8100	10	1995	비아로62번길
광주광역시 광산구 비아동	152.5	152		5 호반	39.68	201907	2	7500	16	1998	비아로
광주광역시 광산구 비아동	152.5	152		5 호반	37.83	201907	6	7600	4	1998	비아로
광주광역시 광산구 산월동	882.1	882		1 부영1차	84.9763	201907	2	19300	1	2003	월계로
광주광역시 광산구 산월동	882.1	882		1 부영1차	84.9763	201907	12	20900	9	2003	월계로
광주광역시 광산구 산월동	882.1	882		1 부영1차	84.9763	201907	22	21000	13	2003	월계로
광주광역시 광산구 산월동	882.1	882		1 부영1차	84.9763	201907	25	21000	6	2003	월계로
광주광역시 광산구 산월동	882.1	882		1 부영1차	84.9763	201907	26	20500	4	2003	월계로
광주광역시 광산구 산월동	882.1	882		1 부영1차	84.9763	201907	30	19100	2	2003	월계로
광주광역시 광산구 산월동	882.1	882		1 부영1차	84.9763	201907	30	22000	13	2003	월계로
광주광역시 광산구 산월동	885.1	885		1 부영2차	84.7886	201907	3	19500	1	2004	첨단중앙로68번길
광주광역시 광산구 산월동	885.1	885		1 부영2차	84.7886	201907	4	19500	3	2004	첨단중앙로68번길
광주광역시 광산구 산월동	885.1	885		1 부영2차	84.7886	201907	8	21400	12	2004	첨단중앙로68번길
광주광역시 광산구 산월동	885.1	885		1 부영2차	84.7886	201907	16	20500	8	2004	첨단중앙로68번길
광주광역시 광산구 산월동	885.1	885		1 부영2차	84.7886	201907	16	20500	1	2004	첨단중앙로68번길

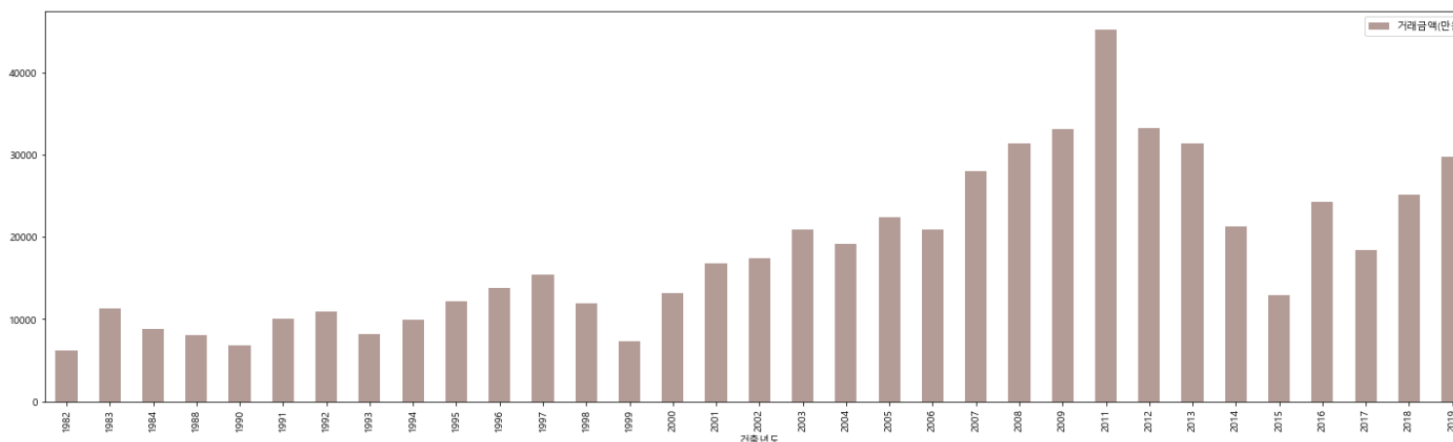
3. 시각화

- 가격대비 층별 시각화
- X축 : 층
- Y축 : 매매가격

```
In [127]: apart.groupby('층').mean()[['거래금액(만원)']].plot(kind='bar', figsize=(25, 7),
                                         stacked=True, color=['r']);
```



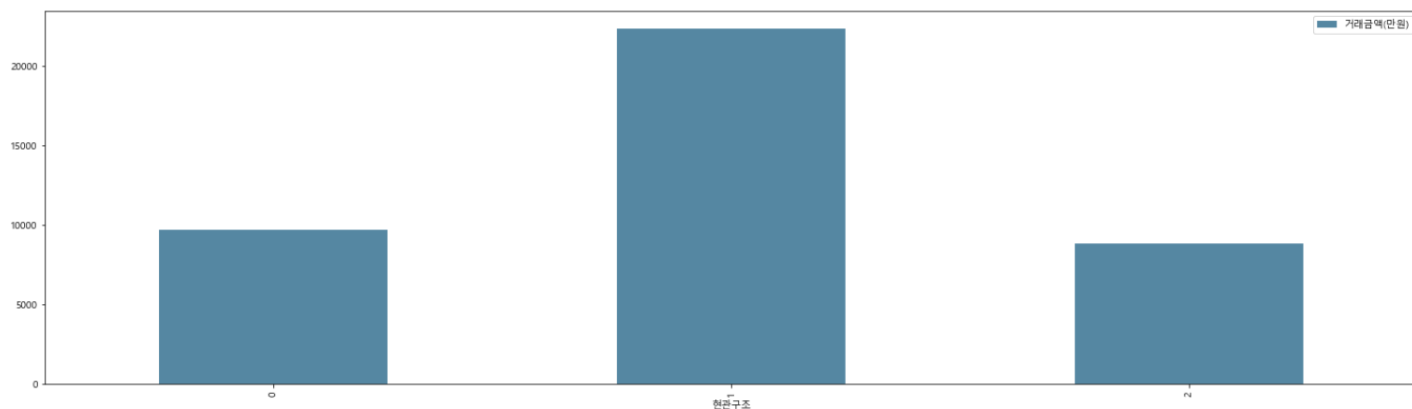
```
In [129]: apart.groupby('건축년도').mean()[['거래금액(만원)']].plot(kind='bar', figsize=(25, 7),
                                         stacked=True, color=['#B39B96']);
```



- 가격대비 건축년도 시각화
- X축 : 건축년도
- Y축 : 매매가격

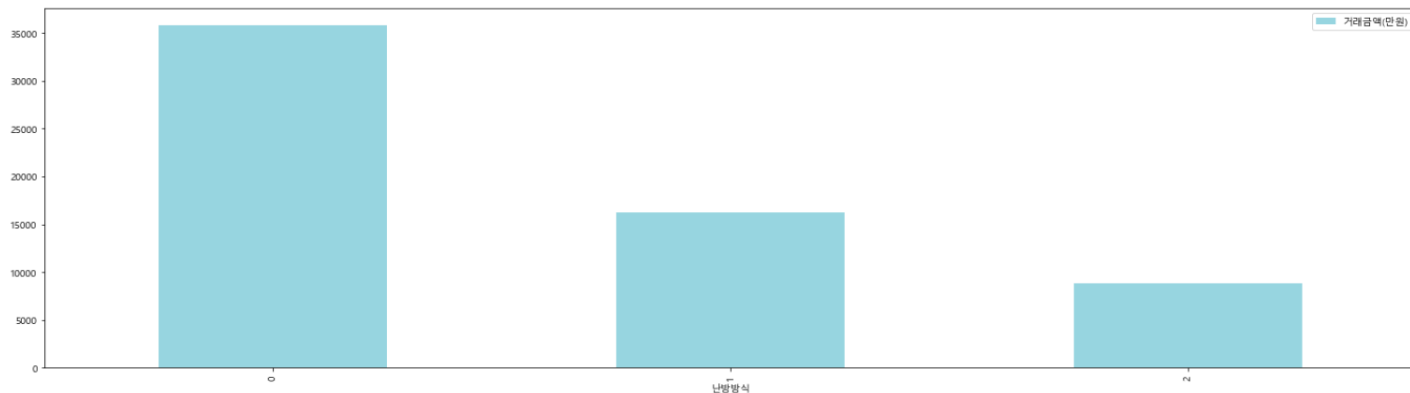
3. 시각화

```
In [115]: apart.groupby('현관구조').mean()[['거래금액(만원)']].plot(kind='bar', figsize=(25, 7),
                                             stacked=True, color=['#5587A2']);
```



- 현관구조 시각화
- X축 : 현관구조 방식
- Y축 : 매매가격

```
In [122]: apart.groupby('난방방식').mean()[['거래금액(만원)']].plot(kind='bar', figsize=(25, 7),
                                             stacked=True, color=['#97D5E0']);
```



- 난방방식 시각화
- X축 : 난방방식
- Y축 : 매매가격

4-1. 데이터 전처리

시군구	번지	본번	부번	단지명	전용면적(m²)	계약년월	계약일	거래금액(만원)	층	건축년도	도로명
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	59.31	201809	22	8,900	5	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	39.15	201811	1	6,350	10	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	39.15	201811	19	6,250	9	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	39.15	201812	14	4,700	6	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	59.31	201812	21	8,300	4	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	59.31	201902	19	8,800	8	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	59.31	201904	4	8,000	10	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	59.31	201904	6	8,550	8	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	39.15	201905	5	5,850	2	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	31.59	201906	7	5,500	10	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	39.15	201907	3	6,000	10	1990	송도로
광주광역시 광산구 도산동	1138-2	1138	0002	대덕1	59.31	201907	20	7,650	8	1990	송도로
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201808	8	17,450	11	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201808	20	16,300	1	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201809	1	18,600	1	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201809	5	18,500	11	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201809	19	18,500	6	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201809	29	18,250	4	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201810	4	17,200	2	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201810	6	18,000	6	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201810	13	18,600	8	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201810	15	18,400	6	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201810	27	17,600	4	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201812	19	17,400	4	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201812	31	18,700	12	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201901	12	18,300	5	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201902	16	18,100	6	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201902	23	17,600	4	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201903	13	17,650	6	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201903	16	17,950	7	2000	송도로182번길
광주광역시 광산구 도산동	790	0790	0000	대주피오레1차	84.98	201904	1	16,000	7	2000	송도로182번길

데이터 삭제

4-1. 데이터 전처리

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	시군구	단지명	번지	전용면적(㎡)	계약년월	계약일	최고층	층	건축년도	세대수	초등학교	200m 이내	건설회사	거래금액(만원)	
2	광주광역시	대덕1	1138.2	59.31	201902	19	10	8	1990	150	4	1	대덕건설	8800	
3	광주광역시	대덕1	1138.2	59.31	201904	2	10	10	1990	150	4	1	대덕건설	8000	
4	광주광역시	대덕1	1138.2	59.31	201904	6	10	8	1990	150	4	1	대덕건설	8550	
5	광주광역시	대덕1	1138.2	39.15	201905	3	10	2	1990	150	4	1	대덕건설	5850	
6	광주광역시	대덕1	1138.2	31.59	201906	3	10	10	1990	150	4	1	대덕건설	5500	
7	광주광역시	대덕1	1138.2	59.31	201801	3	10	3	1990	150	4	1	대덕건설	5000	
8	광주광역시	대덕1	1138.2	59.31	201802	3	10	5	1990	150	4	1	대덕건설	9000	
9	광주광역시	대덕1	1138.2	39.15	201805	3	10	5	1990	150	4	1	대덕건설	6600	
10	광주광역시	대덕1	1138.2	39.15	201806	3	10	9	1990	150	4	1	대덕건설	6480	
11	광주광역시	대덕1	1138.2	39.15	201807	3	10	1	1990	150	4	1	대덕건설	6500	
12	광주광역시	대덕1	1138.2	59.31	201807	2	10	3	1990	150	4	1	대덕건설	7800	
13	광주광역시	대덕1	1138.2	59.31	201809	2	10	5	1990	150	4	1	대덕건설	8900	
14	광주광역시	대덕1	1138.2	39.15	201811	1	10	10	1990	150	4	1	대덕건설	6350	
15	광주광역시	대덕1	1138.2	39.15	201811	1	10	9	1990	150	4	1	대덕건설	6250	
16	광주광역시	대덕1	1138.2	39.15	201812	1	10	6	1990	150	4	1	대덕건설	4700	
17	광주광역시	대덕1	1138.2	59.31	201812	2	10	4	1990	150	4	1	대덕건설	8300	
18	광주광역시	대덕1	1138.2	31.59	201704	2	10	8	1990	150	4	1	대덕건설	5800	
19	광주광역시	대덕1	1138.2	39.15	201704	2	10	8	1990	150	4	1	대덕건설	6500	
20	광주광역시	대덕1	1138.2	39.15	201706	2	10	8	1990	150	4	1	대덕건설	6900	
21	광주광역시	대덕1	1138.2	59.31	201706	2	10	3	1990	150	4	1	대덕건설	8000	
22	광주광역시	대덕1	1138.2	59.31	201707	1	10	10	1990	150	4	1	대덕건설	7200	
23	광주광역시	대덕1	1138.2	39.15	201708	3	10	7	1990	150	4	1	대덕건설	6500	
24	광주광역시	대덕1	1138.2	39.15	201709	2	10	10	1990	150	4	1	대덕건설	6500	
25	광주광역시	대덕1	1138.2	39.15	201710	1	10	7	1990	150	4	1	대덕건설	6800	
26	광주광역시	대덕1	1138.2	39.15	201711	1	10	1	1990	150	4	1	대덕건설	6800	
27	광주광역시	대덕1	1138.2	31.59	201712	3	10	7	1990	150	4	1	대덕건설	5700	
28	광주광역시	대덕1	1138.2	39.15	201712	2	10	6	1990	150	4	1	대덕건설	6800	
29	광주광역시	대덕1	1138.2	39.15	201601	1	10	6	1990	150	4	1	대덕건설	5000	
30	광주광역시	대덕1	1138.2	59.31	201602	3	10	1	1990	150	4	1	대덕건설	5000	
31	광주광역시	대덕1	1138.2	59.31	201605	3	10	8	1990	150	4	1	대덕건설	8100	
32	광주광역시	대덕1	1138.2	59.31	201606	2	10	9	1990	150	4	1	대덕건설	8000	
33	광주광역시	대덕1	1138.2	31.59	201607	3	10	9	1990	150	4	1	대덕건설	5950	

데이터 추가

4-1. 데이터 전처리

-11-

```
In [9]: apt_dum = pd.get_dummies(apt)
apt_test = pd.get_dummies(apt_test)
```

```
In [11]: print(apt_dum.shape)
print(apt_test.shape)

(23877, 152)
(305, 110)
```

```
In [12]: li = []
for i in apt_dum:
    if i not in apt_test:
        li.append(i)
    print(i)
```

```
In [13]: for i in li:
apt_test[i] = 0
```

```
In [14]: apt_test.shape
```

```
Out[14]: (305, 152)
```

```
In [16]: apt_a = apt_test.loc[:, "건설회사_0":]
apt_b = apt_test.loc[:, "시군구_광주광역시 광산구 흑석동"]
```

```
In [17]: apt_a = apt_a.sort_index(axis = 1)
apt_a.head()
```

```
In [18]: data = pd.concat([apt_b, apt_a], axis=1)
```

- 컬럼수를 맞춰주기 위해서 결측치를 0으로 채워 줌
- 새로 추가된 결측치 데이터를 이름 순으로 정렬해 데이터 통합

Random Forest

```
: from sklearn.ensemble import RandomForestRegressor
: from sklearn.model_selection import GridSearchCV

: rf_model = RandomForestRegressor(random_state=0)
: rf_model.fit(X_train_st, price)

: ...

: print("train score : ", rf_model.score(X_train_st, price))
train score : 0.9970699741208182

: pre_rf = rf_model.predict(X_test_st)

: mse_st = metrics.mean_squared_error(sub, y_pre)
: mse_st
: 9723545.173819937

: rmse = np.sqrt(mse_st)
: rmse
: 3118.2599593074237

: r2_score(sub, y_pre)
: 0.8989326902840442
```

```
: rmse = np.sqrt(mse_st)
: rmse
: 3118.2599593074237

: r2_score(sub, y_pre)
: 0.8989326902840442
```

- Random Forest를 사용 하였지만, Rmse(평균제곱오차)가 높아서 매매가격에 영향을 주는 요인을 찾기로 함.

4-2. 데이터 전처리

시군구	단지명	번지	전용면적(㎡)	계약년월	계약일	최고층	층	건축년도	주차수	세대수	현관구조	난방방식	초등학교	200m 이내	200m이내	200m 이내	1km 이내	건설회사	거래금액(만원)
광주광역시	대덕1	1138.2	59.31	201902	19	10	8	1990	0.5	150	0	1	4		1	1	0	대덕건설	8800
광주광역시	대덕1	1138.2	59.31	201904	4	10	10	1990	0.5	150	0	1	4		1	1	0	대덕건설	8000
광주광역시	대덕1	1138.2	59.31	201904	6	10	8	1990	0.5	150	0	1	4		1	1	0	대덕건설	8550
광주광역시	대덕1	1138.2	39.15	201905	5	10	2	1990	0.5	150	0	1	4		1	1	0	대덕건설	5850
광주광역시	대덕1	1138.2	31.59	201906	7	10	10	1990	0.5	150	0	1	4		1	1	0	대덕건설	5500
광주광역시	대덕1	1138.2	59.31	201801	8	10	3	1990	0.5	150	0	1	4		1	1	0	대덕건설	5000
광주광역시	대덕1	1138.2	59.31	201802	5	10	5	1990	0.5	150	0	1	4		1	1	0	대덕건설	9000
광주광역시	대덕1	1138.2	39.15	201805	31	10	5	1990	0.5	150	0	1	4		1	1	0	대덕건설	6600
광주광역시	대덕1	1138.2	39.15	201806	7	10	9	1990	0.5	150	0	1	4		1	1	0	대덕건설	6480
광주광역시	대덕1	1138.2	39.15	201807	9	10	1	1990	0.5	150	0	1	4		1	1	0	대덕건설	6500
광주광역시	대덕1	1138.2	59.31	201807	26	10	3	1990	0.5	150	0	1	4		1	1	0	대덕건설	7800
광주광역시	대덕1	1138.2	59.31	201809	22	10	5	1990	0.5	150	0	1	4		1	1	0	대덕건설	8900
광주광역시	대덕1	1138.2	39.15	201811	1	10	10	1990	0.5	150	0	1	4		1	1	0	대덕건설	6350
광주광역시	대덕1	1138.2	39.15	201811	19	10	9	1990	0.5	150	0	1	4		1	1	0	대덕건설	6250
광주광역시	대덕1	1138.2	39.15	201812	14	10	6	1990	0.5	150	0	1	4		1	1	0	대덕건설	4700
광주광역시	대덕1	1138.2	59.31	201812	21	10	4	1990	0.5	150	0	1	4		1	1	0	대덕건설	8300
광주광역시	대덕1	1138.2	31.59	201704	27	10	8	1990	0.5	150	0	1	4		1	1	0	대덕건설	5800
광주광역시	대덕1	1138.2	39.15	201704	28	10	8	1990	0.5	150	0	1	4		1	1	0	대덕건설	6500
광주광역시	대덕1	1138.2	39.15	201706	6	10	8	1990	0.5	150	0	1	4		1	1	0	대덕건설	6900
광주광역시	대덕1	1138.2	59.31	201706	27	10	3	1990	0.5	150	0	1	4		1	1	0	대덕건설	8000
광주광역시	대덕1	1138.2	59.31	201707	12	10	10	1990	0.5	150	0	1	4		1	1	0	대덕건설	7200
광주광역시	대덕1	1138.2	39.15	201708	3	10	7	1990	0.5	150	0	1	4		1	1	0	대덕건설	6500
광주광역시	대덕1	1138.2	39.15	201709	21	10	10	1990	0.5	150	0	1	4		1	1	0	대덕건설	6500
광주광역시	대덕1	1138.2	39.15	201710	14	10	7	1990	0.5	150	0	1	4		1	1	0	대덕건설	6800
광주광역시	대덕1	1138.2	39.15	201711	15	10	1	1990	0.5	150	0	1	4		1	1	0	대덕건설	6800
광주광역시	대덕1	1138.2	31.59	201712	6	10	7	1990	0.5	150	0	1	4		1	1	0	대덕건설	5700
광주광역시	대덕1	1138.2	39.15	201712	26	10	6	1990	0.5	150	0	1	4		1	1	0	대덕건설	6800
광주광역시	대덕1	1138.2	39.15	201601	15	10	6	1990	0.5	150	0	1	4		1	1	0	대덕건설	5000
광주광역시	대덕1	1138.2	59.31	201602	1	10	1	1990	0.5	150	0	1	4		1	1	0	대덕건설	5000
광주광역시	대덕1	1138.2	59.31	201605	2	10	8	1990	0.5	150	0	1	4		1	1	0	대덕건설	8100
광주광역시	대덕1	1138.2	59.31	201606	22	10	9	1990	0.5	150	0	1	4		1	1	0	대덕건설	8000
광주광역시	대덕1	1138.2	31.59	201607	1	10	9	1990	0.5	150	0	1	4		1	1	0	대덕건설	5950

데이터 추가

4-2. 데이터 전처리

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	시군구	단지명	번지	전용면적(㎡)	계약년월	계약일	최고층	층	건축년도	주차수	세대수	현관구조	난방방식	초등학교	200m 이내	200m이내	200m 이내	1km 이내	건설회사
2	광주광역시	대덕1	1138.2	59.31	201902	1	10	8	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
3	광주광역시	대덕1	1138.2	59.31	201904	1	10	10	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
4	광주광역시	대덕1	1138.2	59.31	201904	1	10	8	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
5	광주광역시	대덕1	1138.2	39.15	201905	1	10	2	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
6	광주광역시	대덕1	1138.2	31.59	201906	1	10	10	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
7	광주광역시	대덕1	1138.2	59.31	201801	1	10	3	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
8	광주광역시	대덕1	1138.2	59.31	201802	1	10	5	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
9	광주광역시	대덕1	1138.2	39.15	201805	3	10	5	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
10	광주광역시	대덕1	1138.2	39.15	201806	1	10	9	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
11	광주광역시	대덕1	1138.2	39.15	201807	1	10	1	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
12	광주광역시	대덕1	1138.2	59.31	201807	2	10	3	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
13	광주광역시	대덕1	1138.2	59.31	201809	2	10	5	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
14	광주광역시	대덕1	1138.2	39.15	201811	1	10	10	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
15	광주광역시	대덕1	1138.2	39.15	201811	1	10	9	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
16	광주광역시	대덕1	1138.2	39.15	201812	1	10	6	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
17	광주광역시	대덕1	1138.2	59.31	201812	2	10	4	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
18	광주광역시	대덕1	1138.2	31.59	201704	2	10	8	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
19	광주광역시	대덕1	1138.2	39.15	201704	2	10	8	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
20	광주광역시	대덕1	1138.2	39.15	201706	1	10	8	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
21	광주광역시	대덕1	1138.2	59.31	201706	2	10	3	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
22	광주광역시	대덕1	1138.2	59.31	201707	1	10	10	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
23	광주광역시	대덕1	1138.2	39.15	201708	1	10	7	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
24	광주광역시	대덕1	1138.2	39.15	201709	2	10	10	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
25	광주광역시	대덕1	1138.2	39.15	201710	1	10	7	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
26	광주광역시	대덕1	1138.2	39.15	201711	1	10	1	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
27	광주광역시	대덕1	1138.2	31.59	201712	1	10	7	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
28	광주광역시	대덕1	1138.2	39.15	201712	2	10	6	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
29	광주광역시	대덕1	1138.2	39.15	201601	1	10	6	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
30	광주광역시	대덕1	1138.2	59.31	201602	1	10	1	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
31	광주광역시	대덕1	1138.2	59.31	201605	1	10	8	1990	0.5	150	0	1	4	1	1	1	0	대덕건설
32	광주광역시	대덕1	1138.2	59.31	201606	2	10	9	1990	0.5	150	0	1	4	1	1	1	0	대덕건설

Train Data 컬럼 20개, 컬럼 별 데이터 24,000개

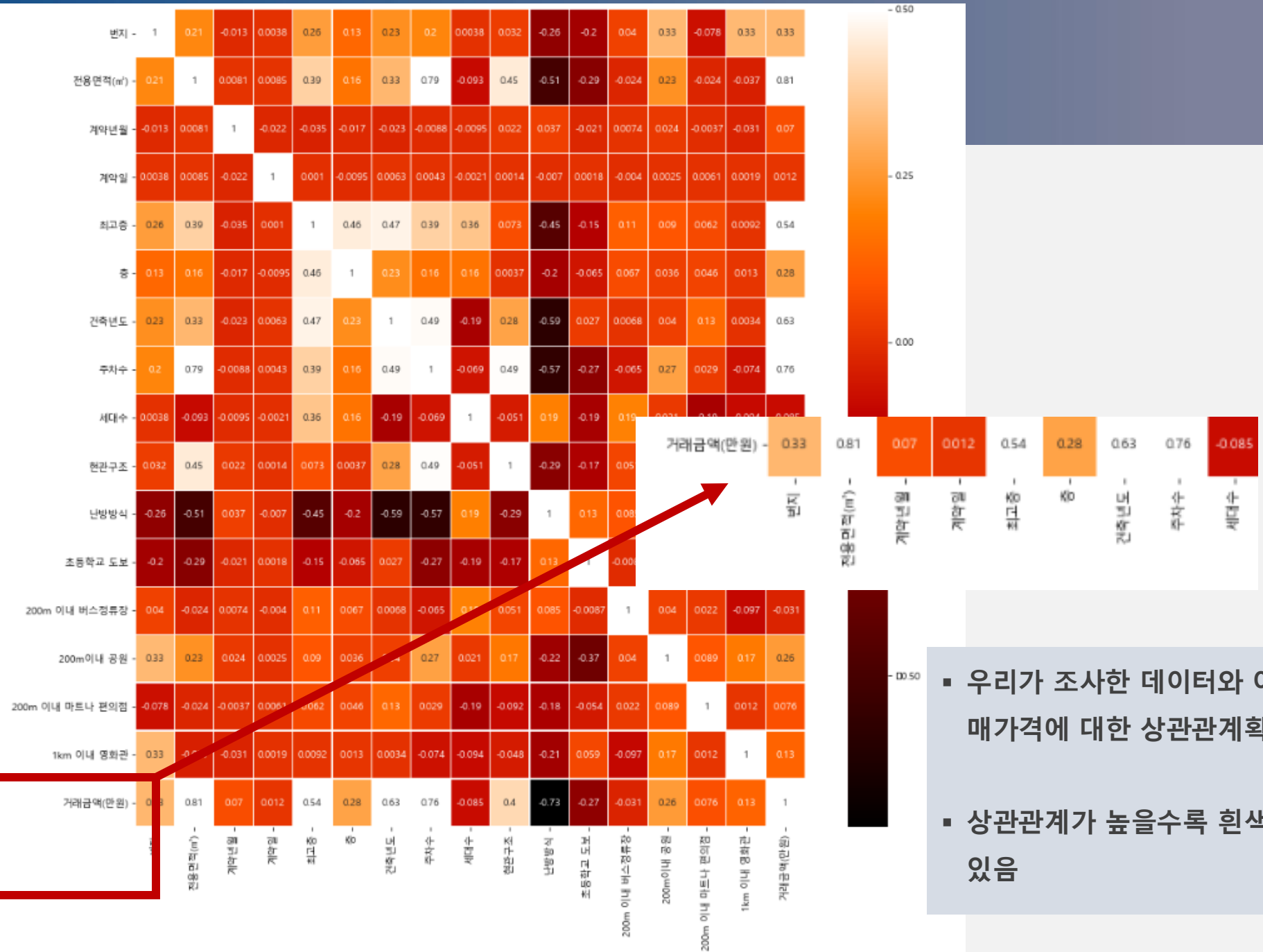
4-2. 데이터 전처리

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	시군구	단지명	번지	전용면적(㎡)	계약년월	계약일	최고층	층	건축년도	주차수	세대수	현관구조	난방방식	초등학교	200m 이내	200m이내	200m 이내	1km 이내	건설회사	거래금액(만원)	
2	광주광역시	대덕1	1138.2	59.31	201907	20	10	8	1990	0.5	150	0	1	4	1	1	1	0	대덕건설	7650	
3	광주광역시	대주피오르	790	84.98	201907	1	13	11	2000	1.01	414	1	1	8	0	0	1	0	대주건설	17900	
4	광주광역시	대주피오르	790	84.98	201907	18	13	3	2000	1.01	414	1	1	8	0	0	1	0	대주건설	17500	
5	광주광역시	무등파크민	1138.1	59.88	201907	9	10	7	1994	0.46	169	0	1	5	1	1	1	0	무등건설	9000	
6	광주광역시	보경해피드	56.5	84.9546	201907	10	15	13	2006	1	115	1	1	4	0	0	0	0	성화종합건	18800	
7	광주광역시	삼라미주	1304.1	31.395	201907	24	12	3	1997	0.43	228	0	1	3	1	1	1	0	삼라건설	5000	
8	광주광역시	호반1	1136.1	59.95	201907	12	13	5	1995	0.7	338	0	1	6	1	1	1	0	호반건설	10150	
9	광주광역시	호반2차	782	59.95	201907	4	13	11	1998	0.76	462	1	1	6	1	0	1	0	호반산업기	13400	
10	광주광역시	호반2차	782	59.95	201907	11	13	7	1998	0.76	462	1	1	6	1	0	1	0	호반산업기	12200	
11	광주광역시	호반2차	782	59.95	201907	18	13	1	1998	0.76	462	1	1	6	1	0	1	0	호반산업기	12300	
12	광주광역시	호반2차	782	59.95	201908	5	13	7	1998	0.76	462	1	1	6	1	0	1	0	호반산업기	13000	
13	광주광역시	호반청암빌	840.1	49.65	201907	5	9	6	1993	0.6	133	0	1	8	1	0	0	0	호반청암건	7000	
14	광주광역시	중흥1차	627.1	59.94	201907	6	15	10	1995	0.72	1056	0	1	15	1	1	0	0	중흥건설	9000	
15	광주광역시	중흥1차	627.1	59.94	201907	6	15	12	1995	0.72	1056	0	1	15	1	1	0	0	중흥건설	9800	
16	광주광역시	중흥1차	627.1	59.94	201907	16	15	9	1995	0.72	1056	0	1	15	1	1	0	0	중흥건설	8900	
17	광주광역시	중흥2차	628.1	59.76	201907	15	15	14	1995	0.7	1056	0	1	13	1	0	0	0	중흥건설	8900	
18	광주광역시	중흥2차	628.1	59.76	201907	22	15	12	1995	0.7	1056	0	1	13	1	0	0	0	중흥건설	8800	
19	광주광역시	중흥2차	628.1	59.76	201907	22	15	10	1995	0.7	1056	0	1	13	1	0	0	0	중흥건설	8100	
20	광주광역시	호반	152.5	39.68	201907	2	20	16	1998	0.71	951	1	1	10	1	0	0	0	호반건설	7500	
21	광주광역시	호반	152.5	37.83	201907	6	20	4	1998	0.71	951	1	1	10	1	0	0	0	호반건설	7600	
22	광주광역시	부영1차	882.1	84.9763	201907	2	15	1	2003	1.14	1198	1	1	7	1	1	0	0	부영	19300	
23	광주광역시	부영1차	882.1	84.9763	201907	12	15	9	2003	1.14	1198	1	1	7	1	1	0	0	부영	20900	
24	광주광역시	부영1차	882.1	84.9763	201907	22	15	13	2003	1.14	1198	1	1	7	1	1	0	0	부영	21000	
25	광주광역시	부영1차	882.1	84.9763	201907	25	15	6	2003	1.14	1198	1	1	7	1	1	0	0	부영	21000	
26	광주광역시	부영1차	882.1	84.9763	201907	26	15	4	2003	1.14	1198	1	1	7	1	1	0	0	부영	20500	
27	광주광역시	부영1차	882.1	84.9763	201907	30	15	2	2003	1.14	1198	1	1	7	1	1	0	0	부영	19100	
28	광주광역시	부영1차	882.1	84.9763	201907	30	15	13	2003	1.14	1198	1	1	7	1	1	0	0	부영	22000	
29	광주광역시	부영2차	885.1	84.7886	201907	3	15	1	2004	1.15	600	1	1	4	1	1	0	0	부영	19500	
30	광주광역시	부영2차	885.1	84.7886	201907	4	15	3	2004	1.15	600	1	1	4	1	1	0	0	부영	19500	
31	광주광역시	부영2차	885.1	84.7886	201907	8	15	2	2004	1.15	600	1	1	4	1	1	0	0	부영	21400	
32	광주광역시	부영2차	885.1	84.7886	201907	16	15	8	2004	1.15	600	1	1	4	1	1	0	0	부영	20500	
33	광주광역시	부영2차	885.1	84.7886	201907	16	15	1	2004	1.15	600	1	1	4	1	1	0	0	부영	20500	

Test Data 컬럼 20개, 컬럼 별 데이터 305개

RandomForestRegressor	RandomForestRegressor
In [107]: <code>from sklearn.ensemble import RandomForestRegressor</code> <code>from sklearn.model_selection import GridSearchCV</code>	In [107]: <code>from sklearn.ensemble import RandomForestRegressor</code> <code>from sklearn.model_selection import GridSearchCV</code>
In [108]: <code>rf_model = RandomForestRegressor()</code> <code>rf_model.fit(X_train_transformed, price)</code>	In [108]: <code>rf_model = RandomForestRegressor(random_state=0)</code> <code>rf_model.fit(X_train_transformed, price)</code>
In [109]: <code>print("train score : ", rf_model.score(X_train_transformed, price))</code>	In [109]: <code>print("train score : ", rf_model.score(X_train_transformed, price))</code>
train score : 0.99698938	train score : 0.9969893895567251
In [110]: <code>pre_rf = rf_model.predict(X_test_transformed)</code>	In [110]: <code>pre_rf = rf_model.predict(X_test_transformed)</code>
In [111]: <code>mse_rf = metrics.mean_squared_error(sub, pre_rf)</code> <code>mse_rf</code>	In [111]: <code>mse_rf = metrics.mean_squared_error(sub, pre_rf)</code> <code>mse_rf</code>
Out[111]: 969572.2843096539	Out[111]: 969572.2843096539
In [112]: <code>rmse = np.sqrt(mse_rf)</code> <code>rmse</code>	In [112]: <code>rmse = np.sqrt(mse_rf)</code> <code>rmse</code>
Out[112]: 984.6686164947342	Out[112]: 984.6686164947342
In [113]: <code>r2_score(sub, pre_rf)</code>	In [113]: <code>r2_score(sub, pre_rf)</code>
Out[113]: 0.98992218778248	Out[113]: 0.98992218778248

3. 시각화



- 우리가 조사한 데이터와 아파트 매매가격에 대한 상관관계확인
- 상관관계가 높을수록 흰색을 띄고 있음

4-3. 데이터 전처리

```
In [9]: combined = apt.append(apt_test, ignore_index=True)
        combined.head()
```

...

```
In [10]: combined.drop(["난방방식", "세대수"], axis=1, inplace=True)
```

```
In [14]: from sklearn.preprocessing import StandardScaler
```

```
In [15]: scaler = StandardScaler()
        scaler.fit(X_train)
        X_train_transformed = scaler.transform(X_train)
        X_test_transformed = scaler.transform(X_test)
```

- Train 데이터 뒤에 Test 데이터를 추가
- 음의 상관관계 "난방방식"과 "세대수" 컬럼을 삭제

RandomForestRegressor

```
In [48]: from sklearn.ensemble import  
from sklearn.model_selection
```

```
In [58]: rf_model = RandomForestRe  
rf_model.fit(X_train_tran
```

```
In [59]: print("train score : ",rf  
train score : 0.99688748
```

```
In [60]: pre_rf = rf_model.predict
```

```
In [61]: mse_rf = metrics.mean_sq  
mse_rf
```

```
Out [61]: 917727.4737704918
```

```
In [62]: rmse = np.sqrt(mse_rf)  
rmse
```

```
Out [62]: 957.980936016209
```

RandomForestRegressor

```
In [48]: from sklearn.ensemble import RandomForestRegressor  
from sklearn.model_selection import GridSearchCV
```

```
In [49]: rf_model = RandomForestRegressor(random_state=0, min_samples_leaf=1, n_estimators=100)  
rf_model.fit(X_train_transformed, price)
```

```
Out [49]: RandomForestRegressor(bootstrap=True, criterion='mse', max_depth=None,  
max_features='auto', max_leaf_nodes=None,  
min_impurity_decrease=0.0, min_impurity_split=None,  
min_samples_leaf=1, min_samples_split=2,  
min_weight_fraction_leaf=0.0, n_estimators=100, n_jobs=None,  
oob_score=False, random_state=0, verbose=0, warm_start=False)
```

```
In [50]: print("train score : ",rf_model.score(X_train_transformed, price))  
train score : 0.9977766620676803
```

```
In [51]: pre_rf = rf_model.predict(X_test_transformed)
```

```
In [52]: mse_rf = metrics.mean_squared_error(sub, pre_rf)  
mse_rf
```

```
Out [52]: 902591.4194440893
```

```
In [53]: rmse = np.sqrt(mse_rf)  
rmse
```

```
Out [53]: 950.0481142784765
```

5. 기대효과

결론

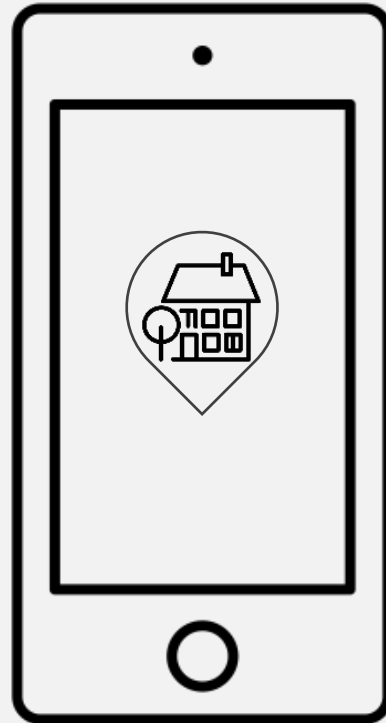


- 이처럼 아파트 가격에는 다양한 요인이 영향을 미치는 것을 알 수 있었고, 다양한 시각화 과정 속에서 상관 관계에 따라 변동 흐름을 보면서 더 나은 결과를 도출해 낼 수 있었습니다.
- 이번 과정을 통해서 협력하고, 함께 진행하면서 많은 것을 배우고 느꼈습니다.

기대효과



아파트 매매 시점 추천



모바일 APP개발



반응형 Web개발

5. 아쉬운 점

-22-

광주광역시 광산구 일대 판매추이

공

광주일보

거래량 줄고 상승세 주춤 ... 광주 부동산 시장 '위축'

원문보기 경제>부동산 2018-10-19

가 가 ★ 스크랩 뉴스듣기 f t

9·13 부동산 대책 발표 이후 거래 한 달 새 38.9%로 떨어져 매매 가격도 0.15% 상승에 그쳐 ... “더 지켜봐야” 신중론도



광주 부동산 시장이 위축되는 분위기다. 정부의 부동산 대책 발표 이후 거래량은 급격히 줄었고 급등세를 보였던 부동산 지표에서도 멈칫하는 분위기가 감지되고 있다.

1 2 3 4 5 6 7 8 9 10 11 12

판매월

[1] 네이버 블로그

(http://www.kwangju.co.kr/news_view.php?aid=1539874800643998005)

[2] 중앙일보 기사

[3] 국토 교통부

(<http://rtdown.molit.go.kr/>)

[4] 네이버 부동산

(<https://land.naver.com/>)

[5] 사랑방 부동산

(<http://home.sarangbang.com/>)

[6] 광주일보 기사

(http://www.kwangju.co.kr/news_view.php?aid=1539874800643998005)

[7] 매일경제 기사

(<https://www.mk.co.kr/news/culture/view/2018/12/777664/>)



윤성호 팀장

정보통신공학과

데이터 전처리 및 모델링



이상현 팀원

소프트웨어학과

데이터 추가 및 데이터 모델링



정혜수 팀원

소프트웨어학과

발표, 데이터 조사, PPT제작



정윤희 팀원

정보통신공학과

데이터 모델링 및 PPT제작



이수완 팀원

소프트웨어학과

데이터 수집 및 자료조사



Q&A



감사합니다.