

COCO '22

한국 주식시장과 경제지표 간의 상관관계 분석

서울 부동산 매매가_회귀분석을 중심으로

COCO 2팀 | 박태하 서주영 윤우성 윤이지

Start Session

주제 선정 이유

업무 중에도 증시 확인... 주식에 빠진 대한민국

조사대상: 성인남녀 639명

조사기간: 2년 1월 8일 ~ 11일 (4일간)

성인 3명 중 2명(67.7%), “주식 투자한다”

주식증시 확인횟수는? 일 평균 5.8회

1년간 매수금액 평균 1,257만원

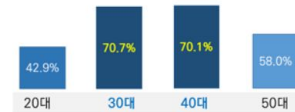
투자수익 평균 704만원



한편, 주식 입문시점은?



연령대별 투자 비율은?



Incrut x 알바바

인크루트가 조사한 설문에 따르면 성인 3명 중 2명은 주식 투자를 하는 것으로 나타났다. (사진=인크루트 홈페이지)

- 최근 2~3년간 대한민국에서 914만명이 주식 투자를 하고 있을 정도로 투자에 대한 관심이 급등
- 특히 20, 30대 사이에서 집 사는 것을 목표로 투자를 시작한 이들이 많음
- 부동산 매매가를 포함한 다양한 경제지표들이 주가에 얼마나 영향을 미치는지, 그리고 20, 30대가 주식시장에 얼마나 영향을 미치는지 분석해 보고자 이와 같은 주제를 선정하였음

주가와 상관관계 분석할 경제지표 선정 및 데이터 수집

대한민국의 주가지수인 코스피와 코스닥에 영향을 미치는 경제지표들을 다음과 같이 선정함

S&P 500

나스닥

세계 증시가 잘
반영되는 경제지표

출처: investing.com

한국 및 미국

기준금리

코스피와 코스닥의
직접적 원인이 되는
경제지표

출처: E-나라지표, FRED

비트코인

서울 부동산 매매가

최근 많은 관심을
받고 있는 투기자산

출처: investing.com,
부동산통계정보시스템

이외에도 어떤 연령이 코스피와 코스닥에 영향을 가장 많이 미치는지 알아보고자

연령별 보유주식수와 인구 수 데이터를 수집하였음

출처: KRX 한국거래소, 행정안전부

분석에 사용한 데이터프레임

		날짜	코스닥	코스피	SP500	나스닥	비트코인	한국기준금리	미국기준금리	매매가_서울				
구분 2016 Year	주주수 2016 KOSPI	0	2018-07-01	775.52	2295.26	2816.29	7671.79	8735000	1.50	1.91	573870	보유주식수 2020 율.9		
		1	2018-08-01	816.97	2322.88	2901.52	8109.54	7870000	1.50	1.91	578331			
		2	2018-09-01	822.27	2343.07	2913.98	8046.35	7417000	1.50	1.95	587393			
0	2016	2026.46	631.44	73302	240326	954088	1421435	1257549	649493	234449	56802			
1	2017	2467.00	798.42	78336	254906	939703	1408217	1305194	700180	255558	63587			
2	2018	2041.00	948.65	92766	316956	1041910	1533553	1427594	787420	282528	72928			
3	2019	2198.00	669.83	98612	381910	1072120	1576373	1510653	948584	363893	164336			
4	2020	2873.00	968.42	273710	1071086	1812487	2214905	1980018	1172221	401673	181128			
5	2021	2978.00	1033.98	656340	2042786	2856757	3165661	2754873	1555875	491060	213351			
8	합계	488744												
			12	2019-07-01	630.18	2024.55	2980.38	8175.42	11968000	1.50	2.40	643380	49672265343	100.0
			13	2019-08-01	610.55	1967.79	2926.46	7962.88	11614000	1.50	2.13	644713		
			14	2019-09-01	621.76	2063.05	2976.74	7999.33	9961000	1.50	2.04	646418		
			15	2019-10-01	658.52	2083.48	3037.56	8292.36	10698000	1.25	1.83	650238		

시각화 - 코스피/코스닥 Trend Line (2018.7~2022.6)



- 코스피와 코스닥은 **2020년 7월**부터 상승했다가 **2021년 7월** 이후로 하락하는 추세를 보임

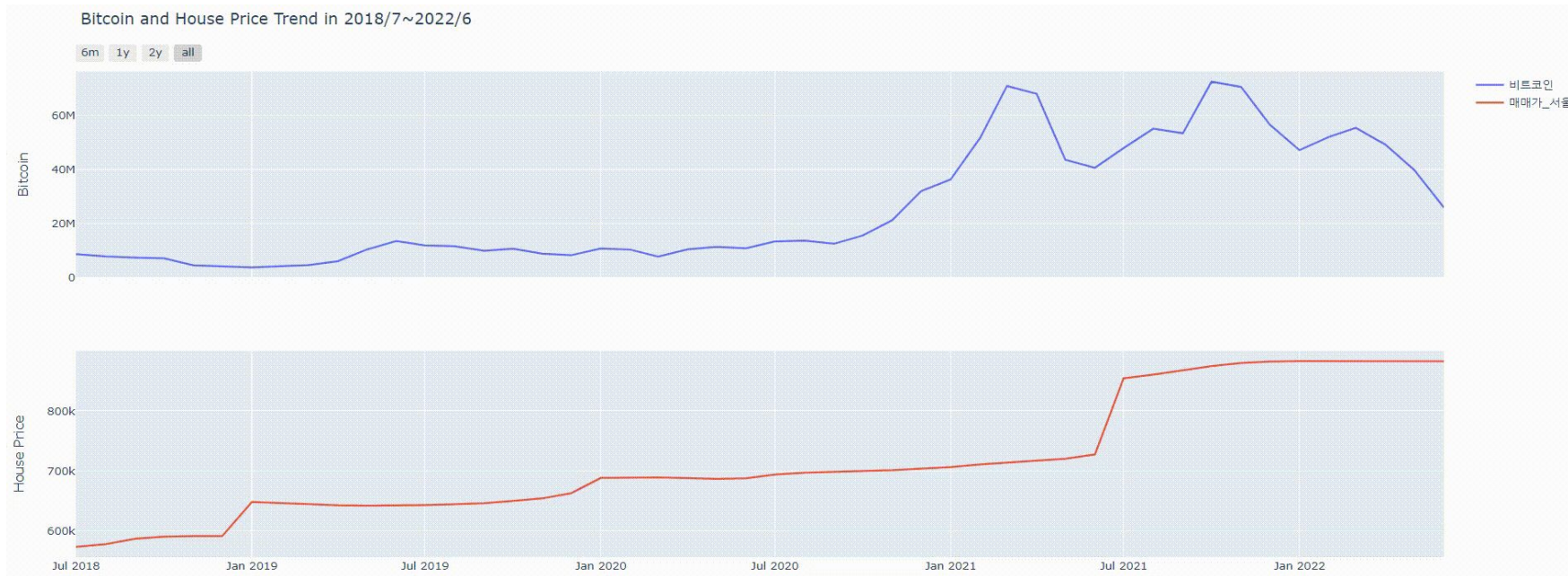
시각화 - S&P500/나스닥 Trend Line (2018.7~2022.6)



05

- S&P 500과 나스닥은 2019년 12월부터 상승하다가 2022년 12월부터 하락하는 추세 보임

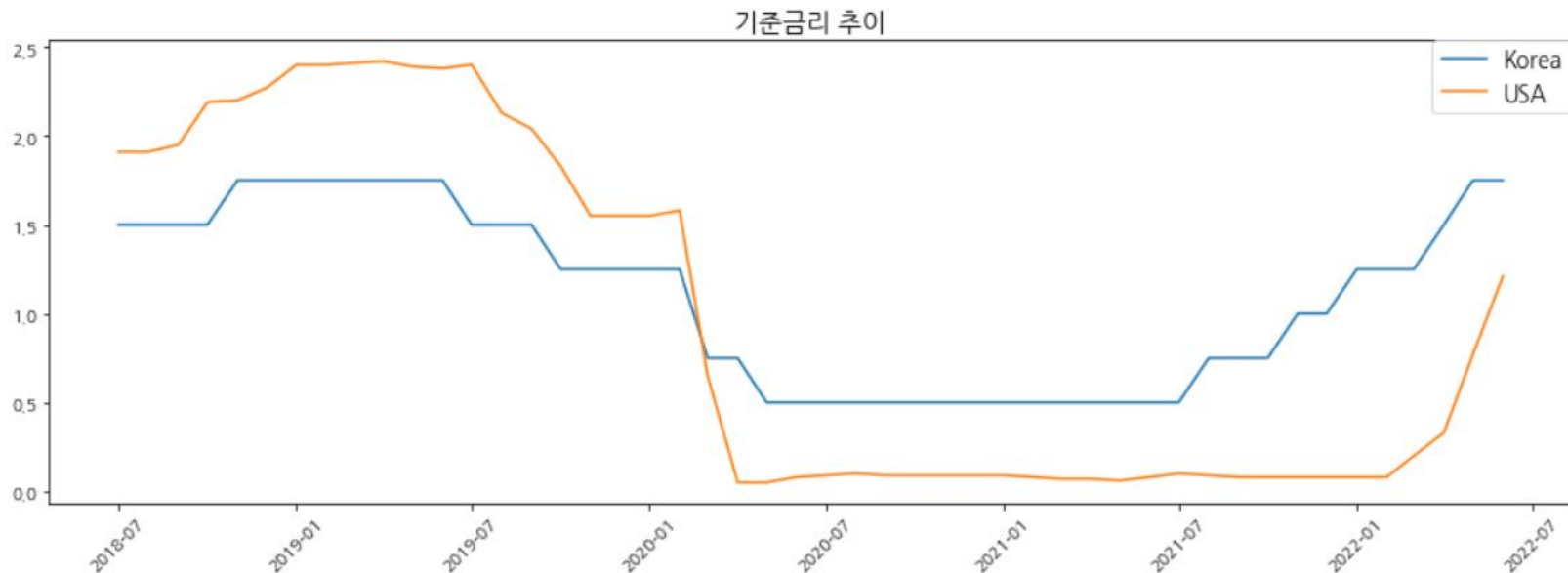
시각화 - 비트코인 / 서울부동산매매가 Trend Line (2018.7~2022.6)



06

- 비트코인 매매가는 **2020년 9월부터** 상승하다가 **2022년 5월부터** 하락하는 추세 보임
- 서울 부동산 매매가는 지난 **5년간** 계속 상승세였으며, **2021년 6월에** 특히 급등함

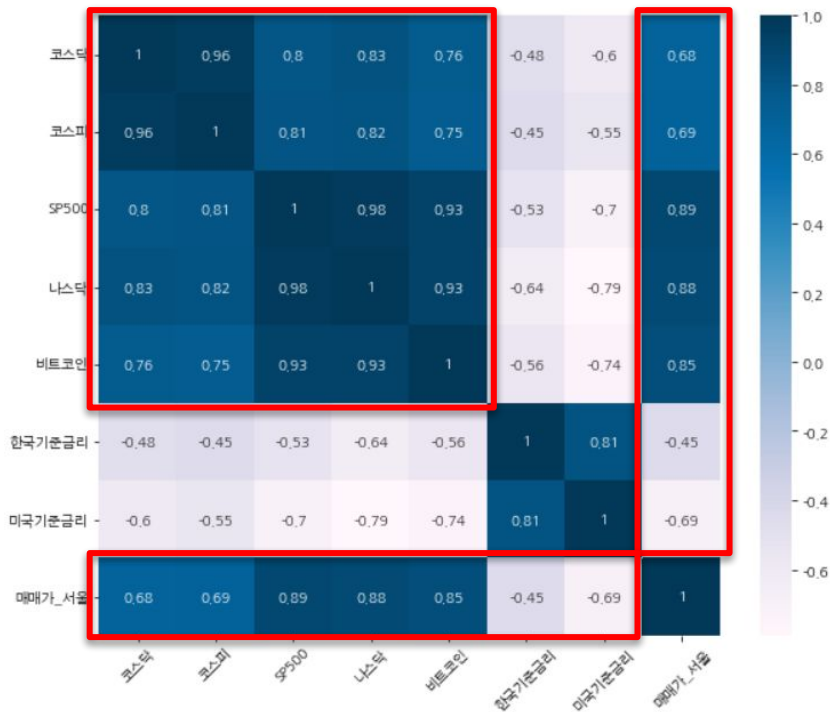
시각화 - 한국 및 미국 기준금리 Trend Line (2018.7~2022.6)



- 기준금리는 미국과 한국 둘 다 **2019년 7월**부터 감소하는 추세이며, 미국의 감소 폭이 더 큼
- 한국 기준금리는 **2021년 7월**부터, 미국 기준금리는 **2021년 1월**부터 다시 상승하는 추세

코스피/코스닥과 타 경제지표들 간의 상관분석

Spearman Correlation of Features



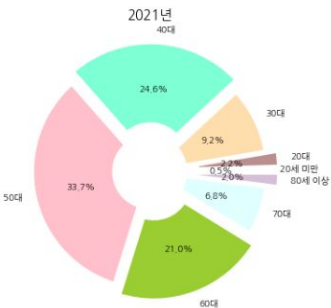
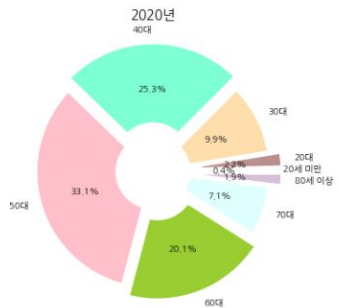
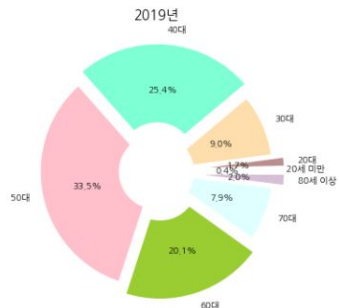
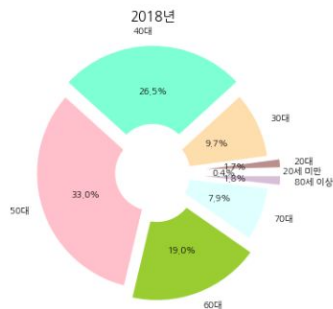
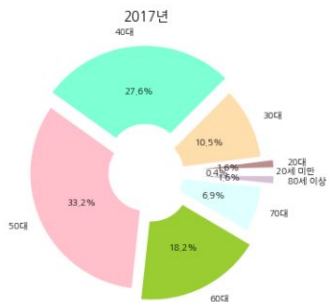
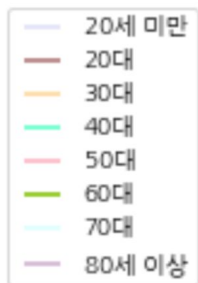
<spearman 사용근거>

- **pearson**은 충분한 데이터가 있어야 하고, 데이터가 정규분포를 따라야함.
- **df_3, 4**의 데이터가 **30**개 미만이기 때문에 상관기법을 통일하기 위해 **spearman** 사용

<상관분석 결과>

- 코스피, 코스닥, **S&P500**, 나스닥, 비트코인은 서로 간의 상관관계수가 모두 **0.7**이상으로, 매우 강한 상관성을 보임.
- 흥미로운 점은 서울부동산의 매매가가 한국투자지표보다 미국투자지표 및 비트코인과 강한 상관성을 보이는 것

데이터시각화 : 연령별 보유주식수 비율 (2017-2021)



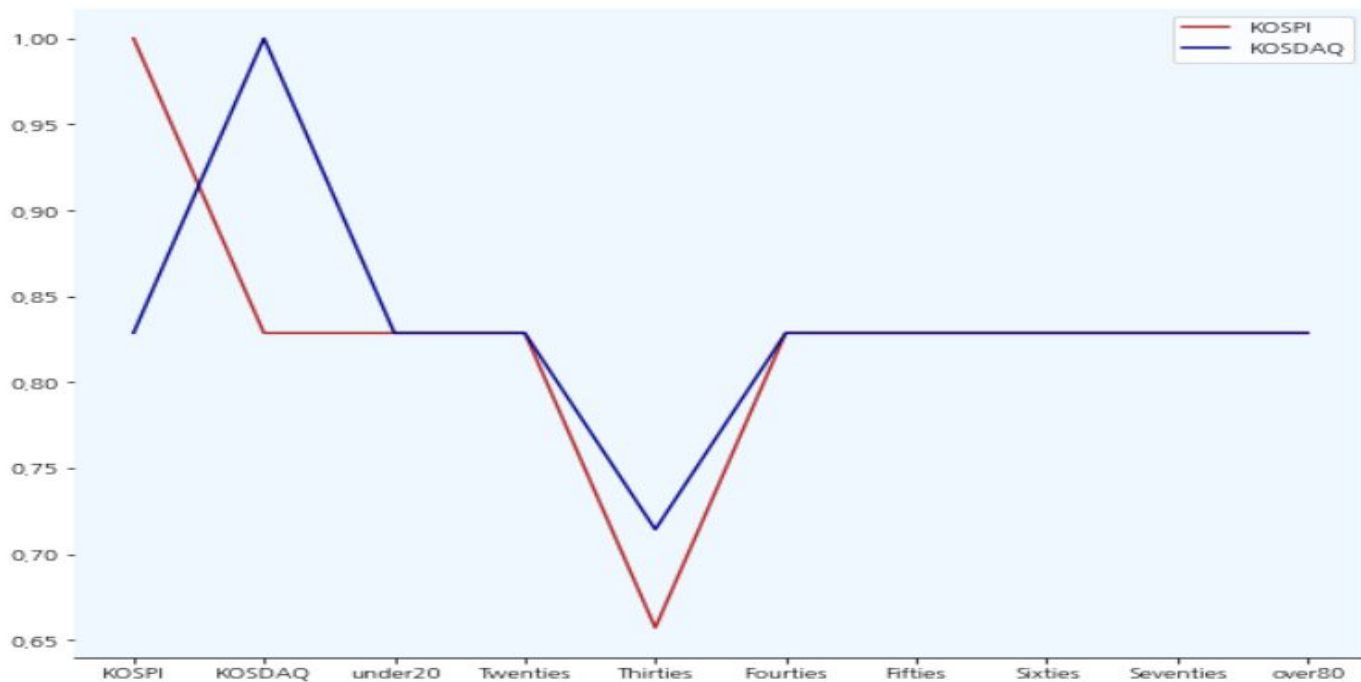
- 최근 **20, 30대** 주식 투기가 과열되고 있다는 기사를 접하여, 실제로 **20~30대**의 보유주식수에 유의미한 차이가 있는지 데이터를 통해 확인하고자 함

- 대체로 **50대, 40대, 60대** 순으로 보유주식수 비율이 높음. **40대**는 소폭 감소하고, **60대**는 소폭 증가

- 코스피 코스닥 가격과 **20~30대**의 보유주식수와의 상관성을 찾고자 분석했으나 유의미한 차이 없었음

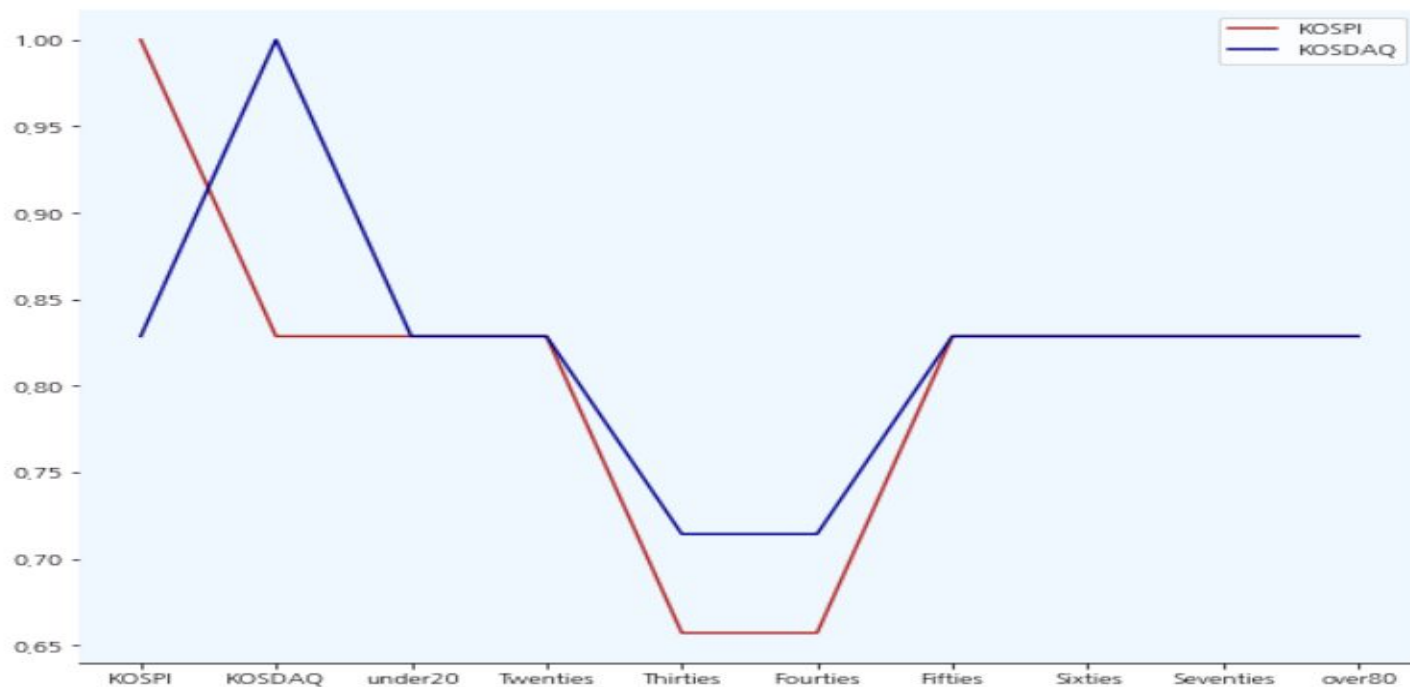
코스피/코스닥과 연령별 보유주식수 간의 상관관계

한국주식시장에 대한 연령별 보유주식수 상관관계



코스피/코스닥과 연령별 주주수 상관관계

한국주식시장에 대한 연령별 주주수 상관관계



코스피와 경제지표 간 회귀분석

OLS 회귀분석 결과

OLS Regression Results

Dep. Variable:

코스피

R-squared:

0.912

Model:

OLS

Adj. R-squared:

0.899

Method:

Least Squares

F-statistic:

70.74

Date:

Thu, 18 Aug 2022

Prob (F-statistic):

4.77e-20

Time:

00:44:45

Log Likelihood:

300.33

No. Observations:

48

AIC:

615.9

Df Residuals:

41

BIC:

629.0

Df Model:

6

Covariance Type:

nonrobust

	coef	std err	t	P> t	[0.025	0.975]
Intercept	1516.3488	332.218	4.564	0.000	845.421	2187.277
SP500	0.0630	0.267	0.236	0.815	-0.477	0.603
나스닥	0.2005	0.070	2.845	0.007	0.058	0.343
비트코인	3.774e-06	2.55e-06	1.483	0.146	-1.37e-06	8.91e-06
한국기준금리	131.4640	156.511	0.840	0.406	-184.615	447.543
미국기준금리	76.8455	82.007	0.937	0.354	-88.771	242.462
매매가_서울	-0.0024	0.001	-4.031	0.000	-0.004	-0.001

Omnibus:

0.068

Durbin-Watson:

0.789

Prob(Omnibus):

0.967

Jarque-Bera (JB):

0.020

Skew:

-0.005

Prob(JB):

0.990

Kurtosis:

2.900

Cond. No.

5.93e+08

Prob가 4.77e-20으로, 0.05보다 작아 통계적으로 유의미한 회귀모형임. P값이 0.05이상인 변수는 '나스닥'과 '매매가_서울'임.

<VIF(분산팽창지수)>

SP500: 80.923

나스닥: 102.362

비트코인: 7.838

한국기준금리: 15.009

미국기준금리: 16.089

매매가_서울: 9.078

VIF(분산팽창지수)

결과 '매매가_서울'과 '비트코인'만

다중공선성이 없음

회귀모델에서 '매매가_서울'만이 유의미한 독립변수임.

코스닥과 경제지표 간 회귀분석

OLS 회귀분석 결과

OLS Regression Results

Dep. Variable:	코스닥	R-squared:	0.899
Model:	OLS	Adj. R-squared:	0.884
Method:	Least Squares	F-statistic:	60.87
Date:	Thu, 18 Aug 2022	Prob (F-statistic):	7.55e-19
Time:	00:44:45	Log Likelihood:	250.26
No. Observations:	48	AIC:	514.5
Df Residuals:	41	BIC:	527.6
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	556.6830	115.581	4.816	0.000	323.263	790.103
SP500	-0.1590	0.093	-1.710	0.095	-0.347	0.029
나스닥	0.1200	0.025	4.896	0.000	0.071	0.170
비트코인	-5.553e-07	8.86e-07	-0.62	0.534	-2.34e-06	1.23e-06
한국기준금리	185.2992	54.451	3.403	0.001	75.333	295.265
미국기준금리	-33.1584	28.531	-1.162	0.252	-90.777	24.460
매매가_서울	-0.0009	0.000	-4.241	0.000	-0.001	-0.000

Omnibus:	0.315	Durbin-Watson:	1.226
Prob(Omnibus):	0.854	Jarque-Bera (JB):	0.479
Skew:	0.149	Prob(JB):	0.787
Kurtosis:	2.612	Cond. No.	5.93e+08

Prob가 **7.55e-19**으로, **0.05**보다 작아 통계적으로 유의미한 회귀모형임. P값이 **0.05**이상인 변수는 ‘나스닥’과 ‘매매가_서울’, ‘한국기준금리’임.

<VIF(분산팽창지수)>

SP500: 80.923

나스닥: 102.362

비트코인: 7.838

한국기준금리: 15.009

미국기준금리: 16.089

매매가_서울: 9.078

VIF(분산팽창지수)

결과 ‘매매가_서울’과 ‘비트코인’만

다중공선성이 없음

회귀모델에서 ‘매매가_서울’만이 유의미한 독립변수임.

코스닥과 경제지표 간 회귀분석

OLS 회귀분석 + VIF를 모두 고려한 결과

코스피와 코스닥 모두 매매가_서울만이 통계적으로 유의미한 독립변수로 나왔다.

<코스피>

매매가_서울	-0.0024	0.001	-4.031	0.000	-0.004	-0.001
--------	---------	-------	--------	-------	--------	--------

<코스닥>

매매가_서울	-0.0009	0.000	-4.241	0.000	-0.001	-0.000
--------	---------	-------	--------	-------	--------	--------

그렇다면 코스피와 코스닥 또한 매매가_서울에 영향을 미칠까??



서울 부동산 매매가와 코스피 / 코스닥과의 회귀분석

OLS 회귀분석 결과

OLS Regression Results

Dep. Variable:	매매가_서울	R-squared:	0.419
Model:	OLS	Adj. R-squared:	0.394
Method:	Least Squares	F-statistic:	16.26
Date:	Thu, 18 Aug 2022	Prob (F-statistic):	4.86e-06
Time:	00:44:45	Log-Likelihood:	-607.34
No. Observations:	48	AIC:	1221.
Df Residuals:	45	BIC:	1226.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	3.442e+05	6.64e+04	5.183	0.000	2.1e+05	4.78e+05
코스피	29.3754	90.022	0.328	0.746	-151.939	210.690
코스닥	372.2607	276.964	1.342	0.186	-185.573	930.095

Omnibus:	1.349	Durbin-Watson:	0.176
Prob(Omnibus):	0.509	Jarque-Bera (JB):	1.349
Skew:	0.359	Prob(JB):	0.509
Kurtosis:	2.602	Cond. No.	1.54e+04

하지만 매매가_서울을 반응변수로 코스피와 코스닥을 확인한 결과 이 모형의 **prob**는 **4.86e-06**으로 모형엔 적합하나, 코스피와 코스닥은 매매가 서울에 영향을 끼치는 변수가 아니었음.

<VIF(분산팽창지수)>

코스피: 11.813
코스닥: 11.813 ← 그 이유는 코스피와 코스닥 간에 다중공선성이 존재하기 때문.

서울 매매가 데이터를 세분화!!!

- 이론상 부동산시장은 주식시장 가격상승에 영향을 미치는 원인이지만, 본 회귀분석에서는 음의 회귀계수를 보여 원인을 분석함.
- 최근 5년간 서울지역에서 주택가격을 보면, 투기지구와 투기비과열지구의 가격 상승폭이 달라 데이터 자체가 희석되었을 수 있음. 그래서 서울지역을 좀 더 세분화해 회귀분석을 진행
- 서울지역을 강남, 강북, 동남, 서남, 동북, 도심, 서북(총 7개 지역)으로 세분화해 회귀분석을 진행함.

강남 부동산 매매가와 코스피 / 코스닥과의 회귀분석

OLS Regression Results

Dep. Variable:	코스피	R-squared:	0.910
Model:	OLS	Adj. R-squared:	0.897
Method:	Least Squares	F-statistic:	69.05
Date:	Thu, 25 Aug 2022	Prob (F-statistic):	7.45e-20
Time:	13:23:38	Log-Likelihood:	-301.46
No. Observations:	48	AIC:	616.9
Df Residuals:	41	BIC:	630.0
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	1867.0910	397.625	4.696	0.000	1064.070	2670.112
SP500	0.0228	0.268	0.085	0.933	-0.519	0.565
나스닥	0.1972	0.071	2.768	0.008	0.053	0.341
비트코인	3.928e-06	2.57e-06	1.528	0.134	-1.26e-06	9.12e-06
한국기준금리	11.0171	144.281	0.076	0.940	-280.365	302.399
미국기준금리	102.2224	79.806	1.282	0.207	-58.849	263.495
매매가_서북	-0.0031	0.001	-3.874	0.000	-0.005	-0.001

Omnibus:	0.264	Durbin-Watson:	0.731
Prob(Omnibus):	0.876	Jarque-Bera (JB):	0.156
Skew:	-0.134	Prob(JB):	0.925
Kurtosis:	2.919	Cond. No.	6.88e+08

OLS Regression Results

Dep. Variable:	코스피	R-squared:	0.910
Model:	OLS	Adj. R-squared:	0.896
Method:	Least Squares	F-statistic:	68.83
Date:	Thu, 25 Aug 2022	Prob (F-statistic):	7.91e-20
Time:	13:20:52	Log-Likelihood:	-301.53
No. Observations:	48	AIC:	617.1
Df Residuals:	41	BIC:	630.2
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	1740.9257	376.295	4.626	0.000	989.982	2500.869
SP500	0.0714	0.272	0.263	0.794	-0.477	0.620
나스닥	0.1971	0.071	2.761	0.009	0.053	0.341
비트코인	3.918e-06	2.57e-06	1.522	0.136	-1.28e-06	9.12e-06
한국기준금리	174.9272	166.831	1.049	0.301	-161.996	511.850
미국기준금리	40.4425	90.576	0.447	0.658	-142.479	223.364
매매가_동남	-0.0017	0.000	-3.853	0.000	-0.003	-0.001

Omnibus:	0.010	Durbin-Watson:	0.737
Prob(Omnibus):	0.995	Jarque-Bera (JB):	0.157
Skew:	-0.023	Prob(JB):	0.924
Kurtosis:	2.723	Cond. No.	6.67e+08

- 7개 지역 중 서북권이 회귀계수가 가장 낮고, 동남권이 회귀계수가 가장 높음.
- 주택시장의 가격변동이 주식시장에 비해 커서, 회귀계수가 음으로 나온것으로 보임.

Thank You