

목차

1. 서론

1.1 주제 선정 이유

1.2 데이터 수집

1.3 데이터 소개

2. 분석

2.1 데이터 전처리

2.2 분산분석

2.3 주성분 분석

3. 결론

1. 서론

1.1 주제 선정 이유

2020년 1분기, 코로나로 인해 전 세계 증시가 큰 폭으로 하락하였다. 코스피 지수 역시 1400p 까지 하락한 모습을 보였으며, 이에 많은 개인투자자들이 주식시장에 대거 유입되었다. 코로나 확산이 장기화되면서 개인투자자들이 기관과 외국인에 맞서 주식을 대거 매입하면서 '동학개미운동'이라는 신조어까지 생겨났다. 이후 코스피가 사상 최고치를 기록하며 많은 사람들이 주식투자에 관심을 가지게 된 만큼 우리 조는 2020년 한 해 동안 유가증권시장에 상장된 종목들을 대상으로 성과를 분석하여 어떠한 특성을 가진 종목에 투자했었을 때, 성과가 좋았는지 분석하고자 하였다.

도표 1. 코스피 지수 추이



자료: 네이버금융

도표 2. 동학개미운동



자료: 네이버 지식백과

1.2 데이터 수집

코스피에 상장된 종목들의 데이터를 분석하기 위해 한국거래소(KRX)에 상장된 종목들의 데이터를 수집하고자 하였다. Python을 이용하여 크롤링하였으며, pykrx패키지를 사용하였다. 유가증권시장에 상장된 2566개 종목 중 시가총액 상위 2000여개 종목의 데이터를 수집하였으며, 이 중 표본 기간 중 거래정지 등의 이유로 수집이 어려운 종목들을 제외한 1939종목의 데이터를 수집하였다.

Python 데이터 크롤링 코드

```
from pykrx import stock
import time
import pandas as pd

stock_code = stock.get_market_ticker_list("20210521", market="ALL")
df = stock.get_market_cap_by_ticker("20210529")

#코스피 시가총액 상위종목 순서로 정렬
```

```

a=df.index.values
stock_code=a.tolist()

res = pd.DataFrame()
i = 1
for ticker in stock_code[:1000]:
    try:
        df = stock.get_market_ohlc_by_date(fromdate="20191201", todate="20210521", ticker=ticker)
        df1 = stock.get_market_cap_by_date(fromdate="20191201", todate="20210521", ticker=ticker)
        df3 = stock.get_market_fundamental_by_date(fromdate="20191201", todate="20210521",
        ticker=ticker)
        df = pd.concat([df, df1, df3], axis=1)
        df = df.assign(종목코드 = ticker, 종목명 = stock.get_market_ticker_name(ticker), brand = i)
        res = pd.concat([res, df], axis=0)
    except Exception:
        pass
    else:
        i = i+1
        time.sleep(1)
res = res.reset_index()
res

res.to_excel("stock_data.xlsx")

```

1.3 데이터 소개

데이터 분석을 위해 종속변수 5개와 설명변수 6개를 설정하였다. 종속변수로는 로그수익률, 일종변동성, 역사적변동성, 회전율, 거래량 지표를 사용하였으며, 설명변수로는 시가총액, 주가, 배당수익률, PER, PBR, 분기를 사용하였다.

로그수익률(x1) : 전일대비 주가 상승률(%)

일종변동성(x2) : Garman and Klass가 On the estimation of security price volatilities from historical data(1980)에서 산출한 1일 단위 변동성(%)

역사적변동성(x3) : 20일간 로그수익률을 사용하여 산출한 20일 역사적변동성(%)

회전율(x4) : 시가총액 대비 거래대금으로 높을수록 유동성이 높음(%)

거래량(x5) : 대표적인 유동성 지표

시가총액(x6) : 시가총액이 높을수록 우량주

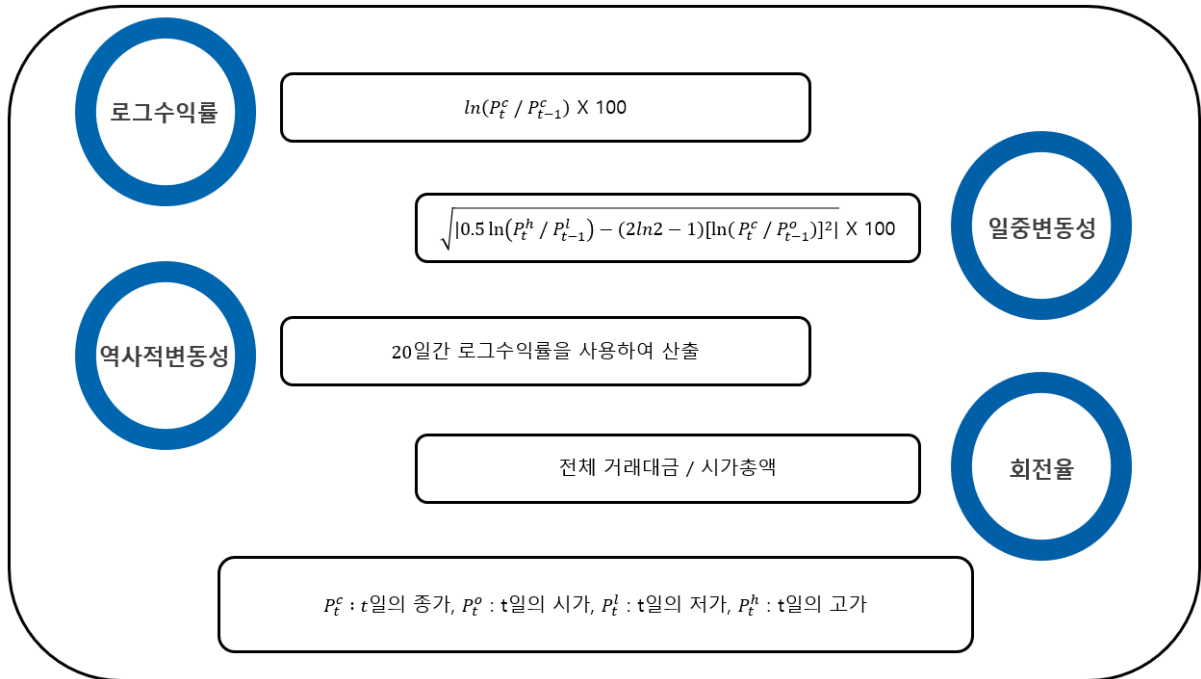
주가(x7) : 2018년 투자접근성을 높이고자 삼성전자 1주당 50주로 액면분할을 한 사례가 있음. 이에 착안하여 1주당 주가 수준 차이에 의한 투자성과 차이를 분석하고자 주가를 설명변수에 포함

배당수익률(x8) : 주가대비 배당액

PER(x9) : 일반적으로 PER이 높을수록 주가가 고평가 상태

PBR(x10) : 일반적으로 PBR이 높을수록 주가가 고평가 상태

분기(x11) : 1분기~4분기



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			로그수익률	일중변동성	역사적변동성	회전율	거래량	시가총액	주가	배당수익률	per	pbr	투자시기	
2	날짜	brand	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	
3	2020-01-02 00:00:00	1	-1.08109	5.060883	23.19961	0.218389	12993228	3295320	55200	2.57	8.54	1.56	1	
4	2020-01-03 00:00:00	1	0.542007	11.9741	22.55646	0.259628	15422255	3313229	55500	2.55	8.59	1.57	1	
5	2020-01-06 00:00:00	1	0	7.939681	22.01331	0.171271	10278951	3313229	55500	2.55	8.59	1.57	1	
6	2020-01-07 00:00:00	1	0.539085	12.69468	21.94247	0.168179	10009778	3331139	55800	2.54	8.64	1.58	1	
7	2020-01-08 00:00:00	1	1.776246	12.56297	21.92404	0.392511	23501171	3390836	56800	2.49	8.79	1.61	1	
8	2020-01-09 00:00:00	1	3.119837	15.13593	23.46881	0.401052	24102579	3498293	58600	2.42	9.07	1.66	1	
9	2020-01-10 00:00:00	1	1.524162	13.96759	23.65701	0.26615	16000170	3552021	59500	2.38	9.21	1.68	1	
10	2020-01-13 00:00:00	1	0.836825	11.92207	23.6599	0.189224	11359139	3581870	60000	2.36	9.29	1.7	1	
11	2020-01-14 00:00:00	1	0	12.57147	22.63455	0.285186	16906295	3581870	60000	2.36	9.29	1.7	1	
12	2020-01-15 00:00:00	1	-1.68071	5.21816	22.70753	0.240334	14300928	3522172	59000	2.4	9.13	1.67	1	
13	2020-01-16 00:00:00	1	2.840625	12.20545	24.26505	0.238019	14381774	3623658	60700	2.33	9.39	1.72	1	
14	2020-01-17 00:00:00	1	0.983614	15.58283	21.49588	0.269046	16025661	3659477	61300	2.31	9.49	1.73	1	
15	2020-01-20 00:00:00	1	1.778543	12.04807	21.623	0.209669	12528855	3725144	62400	2.27	9.66	1.77	1	
16	2020-01-21 00:00:00	1	-1.61554	7.486081	22.63541	0.187424	11142693	3665446	61400	2.31	9.5	1.74	1	
17	2020-01-22 00:00:00	1	1.455159	10.63017	22.83144	0.253788	15339565	3719175	62300	2.27	9.64	1.76	1	
18	2020-01-23 00:00:00	1	-2.43716	10.69938	24.68993	0.251353	14916555	3629628	60800	2.33	9.41	1.72	1	
19	2020-01-28 00:00:00	1	-3.34479	10.85416	27.79713	0.397228	23664541	3510232	58800	2.41	9.1	1.66	1	
20	2020-01-29 00:00:00	1	0.508907	10.88809	27.76745	0.276447	16446102	3528141	59100	2.4	9.15	1.67	1	
21	2020-01-30 00:00:00	1	-3.2677	2.030962	29.80096	0.352274	20821939	3414716	57200	2.48	8.85	1.62	1	
22	2020-01-31 00:00:00	1	-1.40847	11.49736	29.90054	0.336474	19749457	3366957	56400	2.51	8.73	1.6	1	
23	2020-02-03 00:00:00	1	1.408474	9.351743	29.95347	0.396532	23995260	3414716	57200	2.48	8.85	1.62	1	
24	2020-02-04 00:00:00	1	2.078719	17.86662	31.40656	0.361407	21800102	3516202	58000	2.4	9.12	1.67	1	

2. 분석

2.1 데이터 전처리

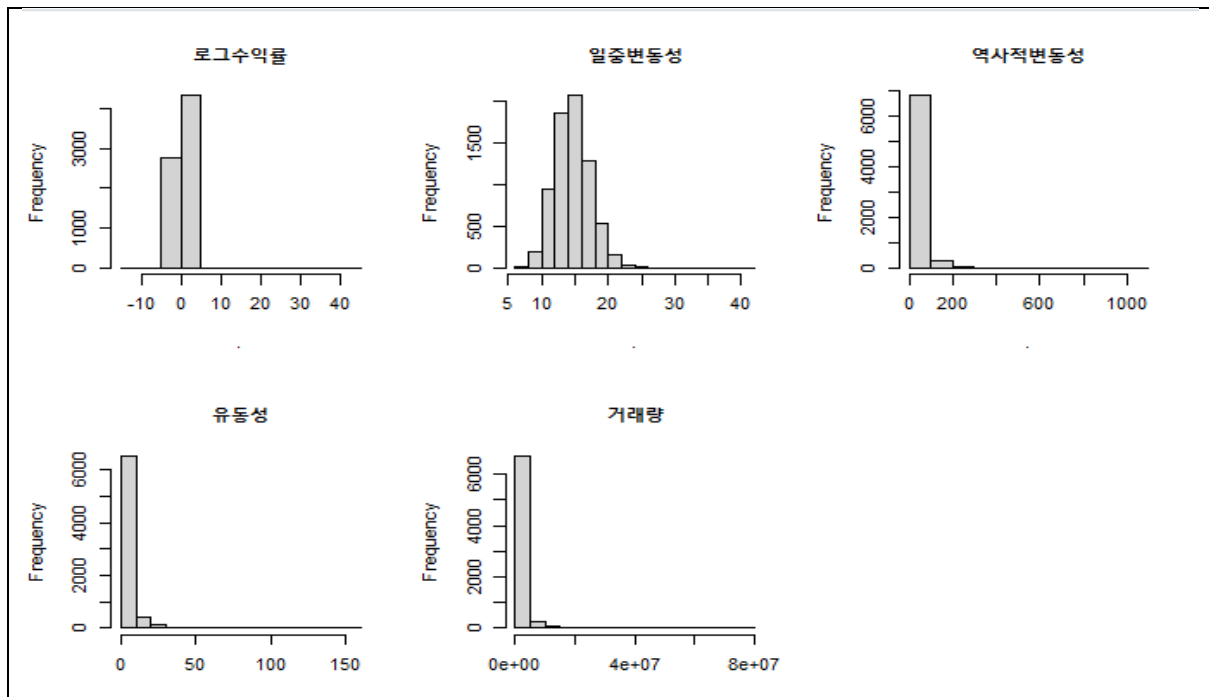
수집한 1939종목 중 우선주와 같이 PER, PBR과 같은 펀더멘탈을 산출할 수 없는 종목과 결측 데이터를 보유한 종목을 제외한 총 1827개 종목에 대한 분석을 실시하고자 하였다. 먼저 6개의 설명변수들이 연속형 변수들이므로 다음과 같은 기준으로 범주화 하였다.

구분	시가총액	주가	배당수익률	PER	PBR	분기
1	10조원 이상	10만원 이상	0%	0	1미만	1분기
2	1조원 이상	5만원 이상	0%대	0초과 10미만	1이상 2미만	2분기
3	5000억원 이상	3만원 이상	1%대	10이상 20미만	2이상 3미만	3분기
4	2000억원 이상	1만원 이상	2%대	20초과 40미만	3이상	4분기
5	1000억원 이상	5천원 이상	3% 이상	40이상 80미만		
6	1000억원 미만	5천원 이하		80이상		

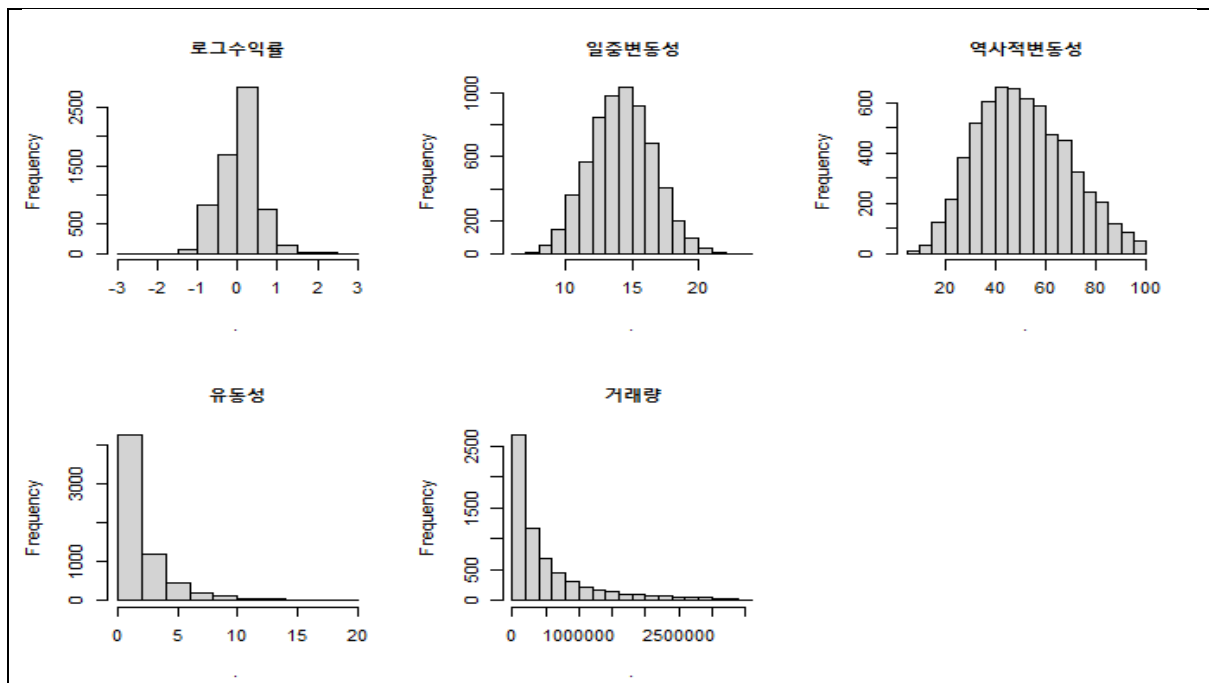
	brand	기업명	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11
1	1	삼성전자	-2.512819e-01	11.238987	33.21986	0.4154745	24784526.84	1	2	4	2	2	1
2	1	삼성전자	1.648068e-01	10.681619	37.66597	0.3614445	21570588.05	1	2	4	3	2	2
3	1	삼성전자	1.521471e-01	9.695329	28.14375	0.3318704	19770543.48	1	2	4	3	2	3
4	1	삼성전자	5.419079e-01	10.494756	24.85571	0.3713184	22196522.64	1	2	4	4	2	4
5	2	SK하이닉스	-1.966282e-01	12.580678	38.87025	0.6203346	4513227.37	1	2	3	2	2	1
6	2	SK하이닉스	3.504670e-02	11.107617	48.03786	0.6015350	4378000.44	1	2	3	4	2	2
7	2	SK하이닉스	-2.032849e-02	9.965695	27.38062	0.5316977	3865252.30	1	2	3	4	2	3
8	2	SK하이닉스	5.640921e-01	12.581989	35.74899	0.5246104	3824908.26	1	2	3	4	2	4
9	3	NAVER	-1.494077e-01	12.070824	32.01064	0.3654189	602649.02	1	1	2	4	4	1
10	3	NAVER	7.400823e-01	14.601287	50.11960	0.6286049	1033917.44	1	1	2	5	4	2
11	3	NAVER	1.637477e-01	12.624565	37.95314	0.6352146	1042693.05	1	1	2	5	4	3
12	3	NAVER	-2.226648e-02	10.553657	30.85239	0.4599352	755031.92	1	1	2	5	4	4
13	4	LG화학	-6.478394e-02	13.922033	52.04404	0.7608337	535761.02	1	1	3	3	2	1
14	4	LG화학	7.788746e-01	14.029245	60.69141	0.8481770	599999.90	1	1	2	5	2	2
15	4	LG화학	4.495032e-01	13.988257	52.91047	1.0532735	743672.17	1	1	2	6	3	3
16	4	LG화학	3.787921e-01	12.803410	45.18906	0.6567303	465672.64	1	1	2	6	4	4
17	5	삼성바이오로직스	1.729135e-01	13.207224	42.91263	0.3263935	216273.68	1	1	1	6	4	1
18	5	삼성바이오로직스	7.785556e-01	14.781913	56.06388	0.5630436	373209.77	1	1	1	6	4	2
19	5	삼성바이오로직스	-1.815179e-01	10.958374	34.02902	0.2816622	186120.05	1	1	1	6	4	3
20	5	삼성바이오로직스	2.949232e-01	11.533459	39.09447	0.2300666	152473.49	1	1	1	6	4	4
21	6	카카오	2.085691e-02	13.086836	33.84514	0.8601520	745821.85	1	3	2	6	3	1
22	6	카카오	8.893029e-01	14.734323	45.66822	1.4496436	1267973.05	1	3	2	6	4	2
23	6	카카오	4.834642e-01	13.631452	41.06622	1.3447136	1183983.64	1	2	2	1	4	3
24	6	카카오	1.087518e-01	10.794887	32.43657	0.6869152	606558.84	1	2	2	1	4	4
25	7	현대차	-4.941772e-01	12.116666	40.84478	0.7237621	1543928.74	1	1	5	3	1	1
26	7	현대차	1.584290e-01	12.203752	51.55885	0.8132853	1739071.00	1	2	5	3	1	2
27	7	현대차	9.416985e-01	15.062255	50.97332	1.3200897	2828334.81	1	1	4	3	1	3
28	7	현대차	1.195193e-01	11.233452	32.47110	0.9490005	2030623.31	1	1	4	3	1	4
29	8	삼성SDI	3.046497e-02	14.507782	51.32343	1.0125071	694390.18	1	1	2	4	2	1
30	8	삼성SDI	6.771462e-01	14.215718	64.47610	1.1099051	763604.84	1	1	2	5	2	2
31	8	삼성SDI	2.751758e-01	13.174819	41.32211	0.7844805	539813.91	1	1	2	5	3	3
32	8	삼성SDI	6.076203e-01	12.501826	37.02776	0.6166078	425320.31	1	1	2	6	3	4
33	9	셀트리온	3.794004e-01	13.863502	40.05892	0.9011255	1169187.81	1	1	1	6	4	1
34	9	셀트리온	4.751817e-01	15.235024	63.28134	1.2457896	1642046.11	1	1	1	6	4	2

이후, 위 데이터의 이상치를 제거하고 파워변환을 통한 전처리 과정을 진행하여 데이터 분석에 용이하게 데이터를 가공하였다.

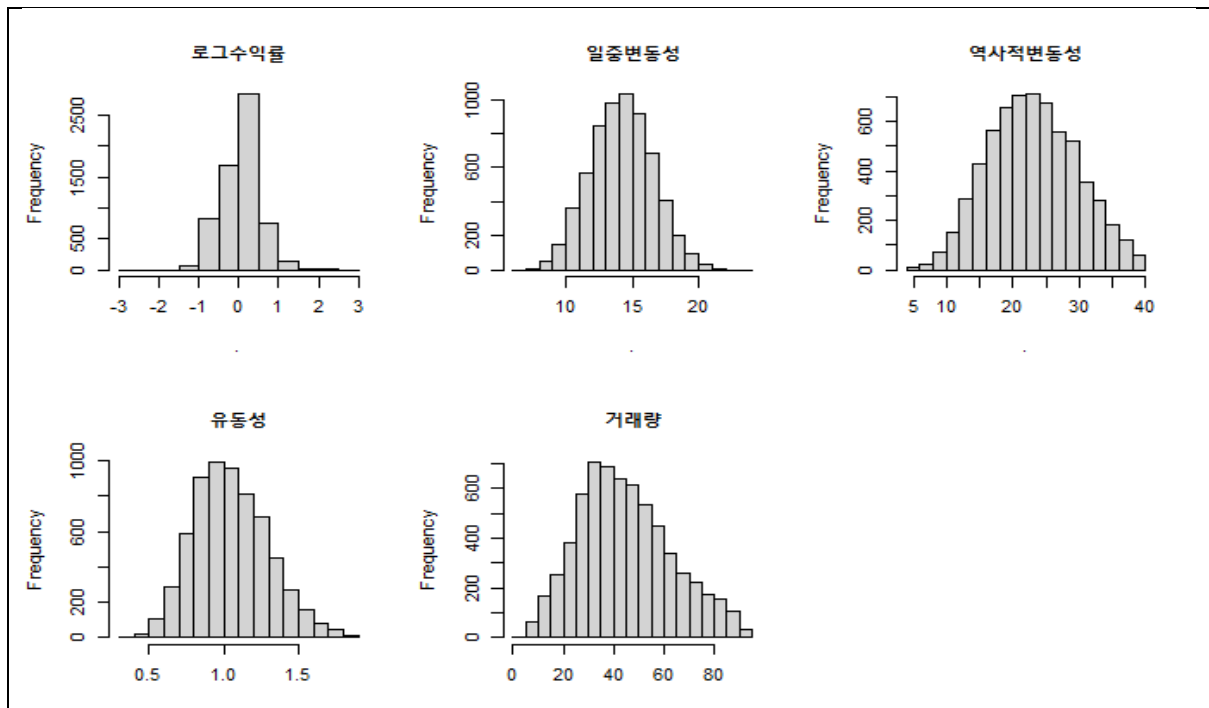
종속변수(x1~x5)의 히스토그램은 다음과 같다.(이상치 제거 X)



각 종속변수별 이상치를 제거한 후 히스토그램은 다음과 같다.



각 종속변수별 파워변환 후 히스토그램은 다음과 같다.



2.2 분산분석

1) 시가총액(x6)에 대한 분석

시가총액(x6) 수준에 따른 x1~x5의 평균이 동일한지 알아보기 위해 일원배치 다변량 분산분석을 실시하였다. 일원배치 다변량 분산분석 결과는 다음과 같다.

	Df	wilks	approx F	num Df	den Df	Pr(>F)
x6	5	0.77931	65.451	25	23568	< 2.2e-16 ***
Residuals	6348					

p-value가 유의수준 0.05보다 작아 "5개의 종속 변수의 평균 간 차이가 없다"는 귀무가설 H0가 기각된다. 이후 LSD.test를 진행하였으며, 결과는 다음과 같다.

로그수익률(x1)			일중변동성(x2)			역사적변동성(x3)		
\$groups			\$groups			\$groups		
	x1 groups			x2 groups			x3 groups	
1	0.11456890	a	6	14.74044	a	6	23.59369	a
4	0.07464685	a	5	14.33587	b	5	23.19710	ab
2	0.06738177	ab	4	14.06783	c	4	22.87938	bc
3	0.06384624	ab	3	13.69318	d	3	22.37046	cd
5	0.05893998	ab	2	13.23971	e	2	21.97982	d
6	0.03535422	ab	1	12.23891	f	1	19.31110	e

회전율(x4)	거래량(x5)	x6의 최대유의선형결합
<pre>\$groups x4 groups 6 1.0854780 a 5 1.0712972 a 4 1.0315139 b 3 0.9884376 c 2 0.9604911 c 1 0.8811553 d</pre>	<pre>\$groups x5 groups 1 56.26238 a 2 48.70286 b 5 44.85824 c 6 44.44186 cd 3 42.99060 de 4 42.68689 e</pre>	<pre>\$vectors [1,] [1,] -0.18029644 [2,] 0.11307030 [3,] -0.03449558 [4,] 0.97641604 [5,] -0.01140538</pre>

시가총액이 큰 우량기업일수록 일중변동성과 역사적변동성이 작게 나타났다. 시가총액이 낮은 기업일수록 회전율이 큰 것으로 나타났는데, 이는 거래대금을 시가총액으로 나눈 회전율의 특성상 시가총액이 낮을수록 회전율이 크게 나온 것으로 예상된다. 시가총액이 가장 큰 1그룹과 2그룹의 로그수익률 및 거래량이 비교적 크게 나타났다.

최대유의선형결합에서 가장 중요한 변수는 회전율(x4)로 나타났다.

2) 주가(x7)에 대한 분석

주가(x7) 수준에 따른 x1~x5의 평균이 동일한지 알아보기 위해 일원배치 다변량 분산분석을 실시하였다. 일원배치 다변량 분산분석 결과는 다음과 같다.

	Df	wilks	approx F	num	Df	den	Df	Pr(>F)
x7	5	0.72263	86.155	25	23568	<	2.2e-16	***
Residuals	6348							

p-value가 유의수준 0.05보다 작아 "5개의 종속 변수의 평균 간 차이가 없다"는 귀무가설 H0가 기각된다. 이후 LSD.test를 진행하였으며, 결과는 다음과 같다.

로그수익률(x1)	일중변동성(x2)	역사적변동성(x3)
<pre>\$groups x1 groups 2 0.10227287 a 3 0.10049958 a 4 0.08615391 a 1 0.07686117 a 5 0.06238144 a 6 0.00931552 b</pre>	<pre>\$groups x2 groups 6 14.58565 a 4 14.52628 a 5 14.20123 b 3 13.65713 c 1 12.83051 d 2 12.73548 d</pre>	<pre>\$groups x3 groups 4 23.99266 a 6 23.32884 b 5 23.09083 b 3 22.26798 c 1 20.45106 d 2 19.69876 d</pre>

회전율(x4)	거래량(x5)	x7의 최대유의선형결합
<pre>\$groups x4 groups 5 1.0768265 a 4 1.0703136 a 6 1.0639306 a 3 1.0072099 b 1 0.9096206 c 2 0.8988685 c</pre>	<pre>\$groups x5 groups 6 52.24758 a 5 43.94492 b 4 42.09444 c 3 38.76280 d 1 33.22280 e 2 31.87858 e</pre>	<pre>\$vectors [,1] [1,] 0.23845583 [2,] -0.16473245 [3,] 0.04316339 [4,] 0.95589427 [5,] -0.02012721</pre>

주가에 대한 분석에서 눈에 띄는 변수는 변동성(x2, x3)과 유동성(x4, x5)이다. 주가가 높은 그룹(1~3)의 변동성이 비교적 낮은 모습을 보이며, 반대로 주가가 낮은 그룹(4~6)은 유동성이 큰 모습을 보였다. 이는 1주당 주식가격이 낮을수록 접근성이 높아져 거래가 활발히 이뤄지며, 이에 따라 주가가 낮은 그룹의 변동성이 높은 그룹보다 비교적 높게 나타난 것이라 예상된다. 실제로 삼성전자는 2018년 4월 액면분할을 통해 투자 접근성을 높여 삼성전자 주식거래의 유동성을 높이고자 한 사례가 있다.

최대유의선형결합에서 가장 중요한 변수는 회전율(x4)로 나타났다.

3) 배당수익률(x8)에 대한 분석

배당수익률(x8) 수준에 따른 x1~x5의 평균이 동일한지 알아보기 위해 일원배치 다변량 분산분석을 실시하였다. 일원배치 다변량 분산분석 결과는 다음과 같다.

	Df	wilks	approx F	num Df	den Df	Pr(>F)
x8	4	0.71024	114.35	20	21045	< 2.2e-16 ***
Residuals	6349					

p-value가 유의수준 0.05보다 작아 “5개의 종속 변수의 평균 간 차이가 없다”는 귀무가설 H0가 기각된다. 이후 LSD.test를 진행하였으며, 결과는 다음과 같다.

로그수익률(x1)	일중변동성(x2)	역사적변동성(x3)
<pre>\$groups x1 groups 2 0.14176379 a 3 0.07208532 b 1 0.04838188 b 4 0.04331353 b 5 -0.02239840 c</pre>	<pre>\$groups x2 groups 1 15.20769 a 2 14.61835 b 3 14.01915 c 4 13.50636 d 5 12.21516 e</pre>	<pre>\$groups x3 groups 1 25.38520 a 2 24.32507 b 3 22.42428 c 4 20.92000 d 5 18.19081 e</pre>

회전율(x4)	거래량(x5)	x8의 최대유의선형결합
\$groups	\$groups	\$vectors
x4 groups	x5 groups	[,1]
1 1.1278736 a	1 51.69766 a	[1,] 0.414711694
2 1.1276697 a	2 45.54917 b	[2,] -0.284150901
3 1.0306302 b	3 41.24162 c	[3,] 0.032734600
4 0.9617249 c	4 38.54914 d	[4,] -0.863794636
5 0.8503706 d	5 35.23091 e	[5,] -0.007729746

배당수익률이 가장 높은 5그룹(3%이상)의 경우 로그수익률이 음수가 나오며 가장 저조한 수익률을 보였다. 배당수익률이 낮은 그룹일수록 변동성과 유동성이 높게 나왔다. 일반적으로 고배당주의 경우 기업 영업이익의 큰 비중을 주주에게 환원하는 특징이 있다. 저배당주의 경우 주주에게 환원하는 배당금은 적지만 그만큼 투자자의 활동을 통해 기업실적 향상을 도모하는 성장형 기업이 많이 분포한다. 따라서, 성장형 기업이 많이 분포한 저배당주 그룹의 변동성 및 유동성이 높게 나온 것으로 보인다.

최대유의선형결합에서 가장 중요한 변수는 회전율(x4)과 로그수익률(x1)으로 나타났다.

4) PER(x9)에 대한 분석

PER(x9) 수준에 따른 x1~x5의 평균이 동일한지 알아보기 위해 일원배치 다변량 분산분석을 실시하였다. 일원배치 다변량 분산분석 결과는 다음과 같다.

	Df	wilks	approx F	num Df	den Df	Pr(>F)
x9	5	0.81286	54.074	25	23568	< 2.2e-16 ***
Residuals	6348					

p-value가 유의수준 0.05보다 작아 "5개의 종속 변수의 평균 간 차이가 없다"는 귀무가설 H0가 기각된다. 이후 LSD.test를 진행하였으며, 결과는 다음과 같다.

로그수익률(x1)	일중변동성(x2)	역사적변동성(x3)
\$groups	\$groups	\$groups
x1 groups	x2 groups	x3 groups
6 0.13065042 a	1 15.18648 a	6 25.15395 a
5 0.10806123 ab	6 14.86545 b	1 25.00084 a
4 0.09955675 ab	5 14.75056 b	5 24.60000 a
3 0.07066809 b	4 14.40619 c	4 23.31157 b
1 0.02275330 c	3 13.87348 d	3 22.16945 c
2 0.01270684 c	2 13.15257 e	2 20.58447 d

회전율(x4)	거래량(x5)	x9의 최대유의선형결합
<pre>\$groups x4 groups 6 1.1609763 a 5 1.1209037 b 1 1.1089935 bc 4 1.0919292 c 3 1.0223550 d 2 0.9274704 e</pre>	<pre>\$groups x5 groups 6 52.32000 a 1 50.76795 a 5 47.90738 b 4 44.64551 c 3 41.61434 d 2 38.26579 e</pre>	<pre>\$vectors [,1] [1,] -0.58793790 [2,] 0.40645896 [3,] -0.07512620 [4,] 0.69512632 [5,] 0.01660097</pre>

PER이 낮은 기업일수록 변동성과 유동성이 작게 나타났다. 그룹1은 PER이 0인 그룹으로 PER이 가장 낮지만 PER=시가총액 / 순이익 이므로 그룹1은 저평가 상태가 아닌 순이익 0 이하인 적자 기업으로 분류된다. 분석결과 PER이 높은 그룹(4~6)의 수익률이 비교적 높게 나타났으며, 변동성과 유동성도 높은 것으로 나타났다.

최대유의선형결합에서 가장 중요한 변수는 회전율(x4)과 로그수익률(x1)으로 나타났다.

5) PBR(x10)에 대한 분석

PBR(x10) 수준에 따른 x1~x5의 평균이 동일한지 알아보기 위해 일원배치 다변량 분산분석을 실시하였다. 일원배치 다변량 분산분석 결과는 다음과 같다.

	Df	wilks	approx F	num Df	den Df	Pr(>F)
x10	3	0.82412	84.777	15	17519	< 2.2e-16 ***
Residuals	6350					

p-value가 유의수준 0.05보다 작아 "5개의 종속 변수의 평균 간 차이가 없다"는 귀무가설 H0가 기각된다. 이후 LSD.test를 진행하였으며, 결과는 다음과 같다.

로그수익률(x1)	일중변동성(x2)	역사적변동성(x3)
<pre>\$groups x1 groups 4 0.13045699 a 3 0.06823245 b 2 0.06432323 b 1 0.01965514 c</pre>	<pre>\$groups x2 groups 4 15.45603 a 3 14.77477 b 2 14.52001 c 1 13.39306 d</pre>	<pre>\$groups x3 groups 4 26.30785 a 3 24.51650 b 2 23.76400 c 1 20.82827 d</pre>
회전율(x4)	거래량(x5)	x10의 최대유의선형결합
<pre>\$groups x4 groups 4 1.1705532 a 3 1.1360672 b 2 1.0981273 c 1 0.9421122 d</pre>	<pre>\$groups x5 groups 4 49.55667 a 3 48.77536 a 2 47.04514 b 1 40.07296 c</pre>	<pre>\$vectors [,1] [1,] -0.084044021 [2,] 0.040676286 [3,] -0.003190146 [4,] 0.995618016 [5,] -0.004078178</pre>

PBR이 높은 기업일수록 모든 종속변수(x1~x5)가 높게 나타났다. PER과 PBR이 낮을수록 주가가 저평가 상태에 있음을 의미한다. PER과 PBR이 높은 업종은 IT, 바이오, 제약과 같은 고부가 가치 기업이다. 이들 기업은 미래에 기대가치가 높기에 현재실적대비 주가가 고평가 된다는 특징이 있다. 작년 한해동안 코로나로 인해 큰 주가하락이 있었지만 코로나 회복에 대한 기대감과 더불어 많은 자금이 증시로 유입되면서 증시가 사상 최고가를 기록하기도 하였다. 작년 한해동안 성과를 분석하였을 때, PER과 PBR이 높은 주식의 수익률 및 유동성이 높게 나타났지만 그만큼 변동성도 크게 나타난 모습이다.

최대유의선형결합에서 가장 중요한 변수는 회전율(x4)로 나타났다.

6) 분기(x11)에 대한 분석

PER(x9) 수준에 따른 x1~x5의 평균이 동일한지 알아보기 위해 일원배치 다변량 분산분석을 실시하였다. 일원배치 다변량 분산분석 결과는 다음과 같다.

	df	wilks	approx F	num Df	den Df	Pr(>F)
x11	1	0.54329	1067.3	5	6348	< 2.2e-16 ***
Residuals	6352					

p-value가 유의수준 0.05보다 작아 "5개의 종속 변수의 평균 간 차이가 없다"는 귀무가설 H0가 기각된다. 이후 LSD.test를 진행하였으며, 결과는 다음과 같다.

로그수익률(x1)	일중변동성(x2)	역사적변동성(x3)
\$groups x1 groups 2 0.3830962 a 4 0.1727287 b 3 0.1664154 b 1 -0.4879392 c	\$groups x2 groups 2 15.23454 a 1 14.41385 b 3 14.07007 c 4 13.19715 d	\$groups x3 groups 2 26.90159 a 1 23.25580 b 3 22.39553 c 4 19.62436 d
회전율(x5)	거래량(x6)	x11의 최대유의선형결합
\$groups x4 groups 3 1.0834952 a 2 1.0622083 b 4 1.0583912 b 1 0.9848693 c	\$groups x5 groups 3 46.59544 a 2 45.53277 ab 4 45.22930 b 1 41.24904 c	\$vectors [,1] [1,] 0.4369764484 [2,] -0.0925471131 [3,] -0.0172703910 [4,] 0.8945323787 [5,] -0.0004151778

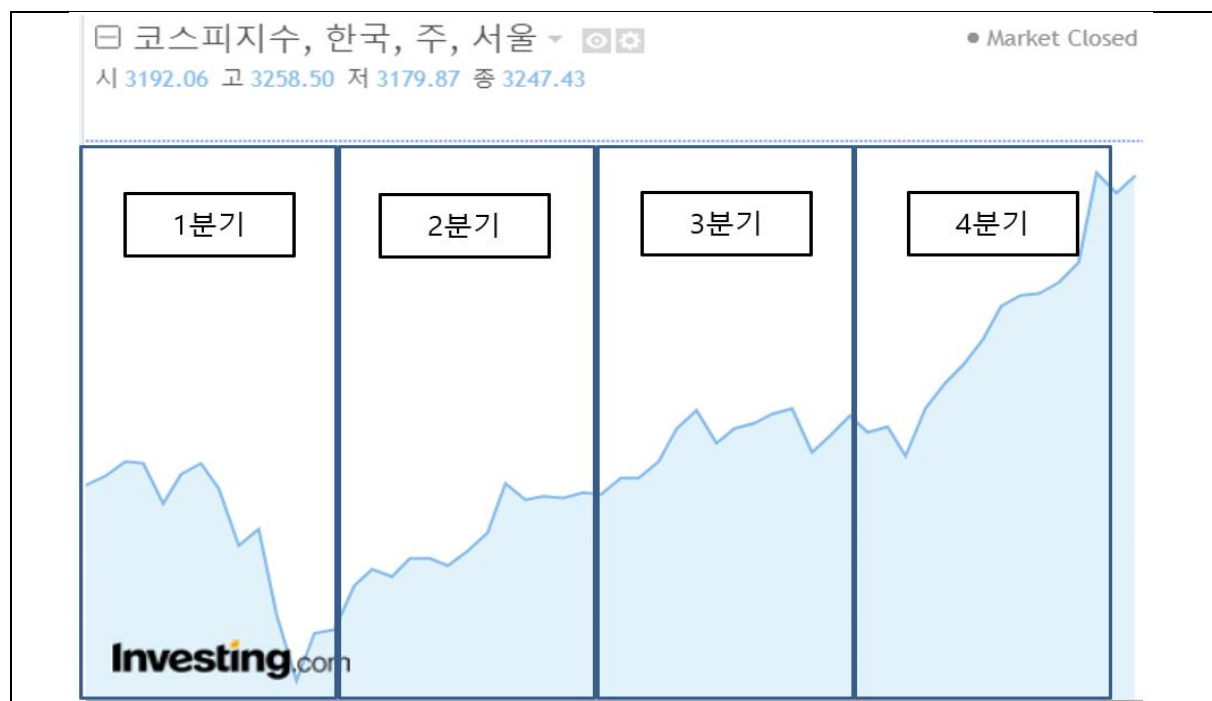
분석결과 1분기의 수익률, 유동성이 다른 분기에 비해 가장 낮은 것을 볼 수 있다. 2020년 1분기는 코로나로 인해 주가가 하락한 시기이다. 따라서 증시가 침체되어 수익률과 유동성이 가장

낮은 시기로 보인다. 수익률과 변동성이 가장 큰 시기는 2분기로 코스피가 연중 최저점을 찍은 1분기를 지나 주가가 이전 수준으로 회복하는 기간이었다. 유동성이 가장 큰 시기는 3분기로 나타났으며, 변동성이 가장 작은 시기는 4분기였다. 4분기는 코스피 지수가 연중 가장 큰 폭으로 상승한 시기였다. 하지만 변동성이 가장 작은 시기로 나타났는데, 이는 코스피 지수가 시가총액 가중 평균 방식이기에 상위 우량주에 편향된 성격을 띠기 때문으로 예상된다.

실제로 4분기 데이터를 대상으로 로그수익률을 종속변수, 시가총액을 설명변수로 한 일원배치 일변량 분산분석 결과 시가총액이 높은 1그룹과 2그룹의 로그수익률이 가장 크게 나타났다.

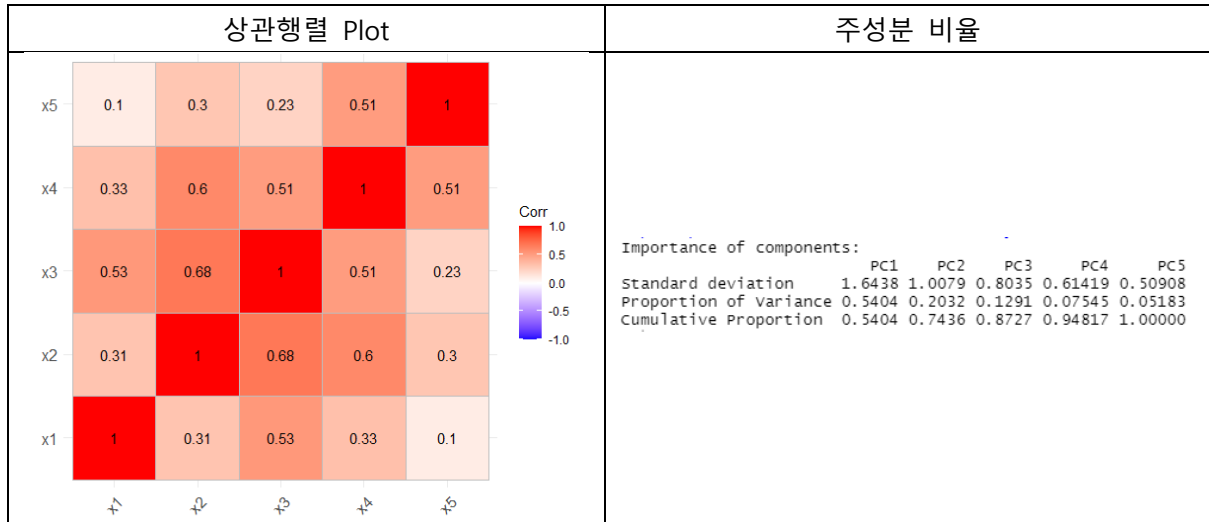
최대유의선형결합에서 가장 중요한 변수는 회전율(x4)과 로그수익률(x1)으로 나타났다.

일원배치 일변량 분산분석(4분기 데이터 : x1~x6)									
	Df	wilks	approx F	num Df	den Df	Pr(>F)	\$groups		
x6	5	0.72513	21.518	25	5952.7	< 2.2e-16 ***	1	0.2808200	a
							2	0.2239869	a
							6	0.1832725	ab
							4	0.1783269	ab
							5	0.1468392	b
							3	0.1393324	b
Residuals	1606								

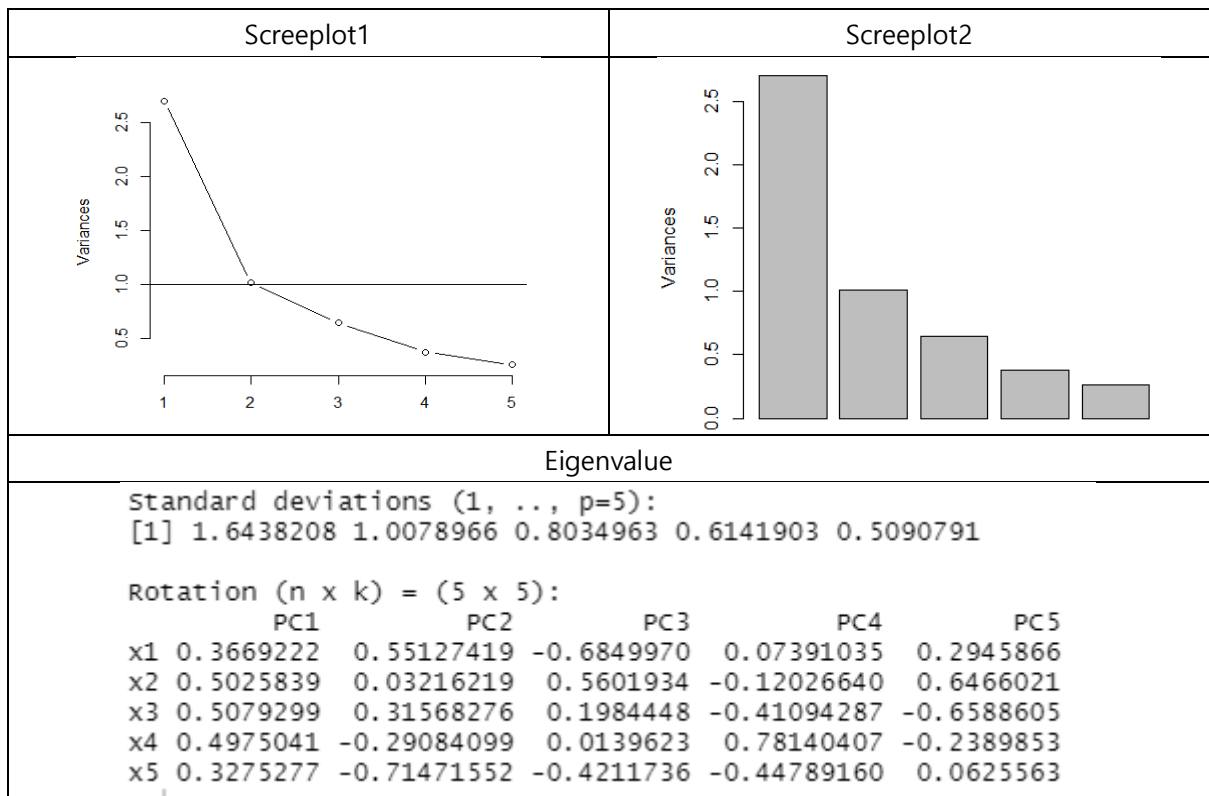


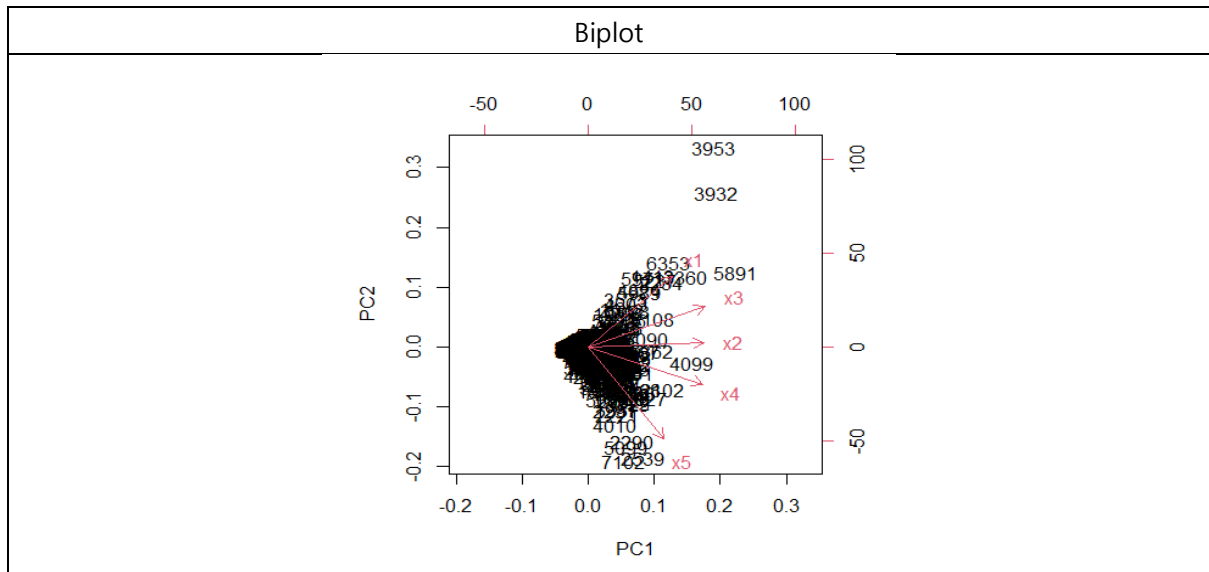
2.3 주성분 분석

독립변수들 간 상관관계를 확인하기 위해 상관행렬 plot을 그려보았으며, 주성분 분석 시행결과 분산비율은 다음과 같았다.



주성분 분석 시행 후 분산비율을 확인하였으며, 분산 10이 넘는 PC1과 PC2 두개의 주성분변수를 결정하였다. PC1과 PC2 두개의 변수로 약 74.36% 설명이 가능하다.





biplot확인 결과 PC1에서는 x2(일중변동성), x3(역사적변동성), x4(회전율)이 상대적으로 큰 영향을 미치고 이는 PC1이 변동성과 유동성에 대한 지표라고 볼 수 있다. PC2에서는 x5(거래량) 상대적으로 가장 큰 영향을 미치는 변수이고 거래량에 대한 지표라고 볼 수 있다.

기존변수와 주성분 점수는 다음과 같다.

기존변수와 주성분점수										
brand	기업명	x1	x2	x3	x4	x5	PC1	PC2		
1	1	삼성전자	-2.512819e-01	11.238987	33.21986	0.4154745	24784526.84	0.9474152926	-5.266858560	
2	1	삼성전자	1.648068e-01	10.681619	37.66597	0.3614445	21570588.05	0.7480686347	-4.324603423	
3	1	삼성전자	1.521471e-01	9.695329	28.14375	0.3318704	19770543.48	0.2529024568	-4.046595743	
4	1	삼성전자	5.419079e-01	10.494756	24.85571	0.3713184	22196522.64	0.7345229810	-4.360828251	
5	2	SK하이닉스	-1.966282e-01	12.580678	38.87025	0.6203346	4513227.37	-0.6444659909	-0.917932382	
6	2	SK하이닉스	3.504670e-02	11.107617	48.03786	0.6015350	4378000.44	-0.7142897234	-0.696820260	
7	2	SK하이닉스	-2.032849e-02	9.965695	27.38062	0.5316977	3865252.30	-1.2894448545	-0.810605101	
8	2	SK하이닉스	5.640921e-01	12.581989	35.74899	0.5246104	3824908.26	-0.4798849273	-0.372555178	
9	3	NAVER	-1.494077e-01	12.070824	32.01064	0.3654189	602649.02	-1.2133871628	-0.123770699	
10	3	NAVER	7.400823e-01	14.601287	50.11960	0.6286049	1033917.44	-0.1032654491	0.456912721	
11	3	NAVER	1.637477e-01	12.624565	37.95314	0.6352146	1042693.05	-0.8490356358	0.004339325	
12	3	NAVER	-2.226648e-02	10.553657	30.85239	0.4599352	755031.92	-1.4387773047	-0.117201954	
13	4	LG화학	-6.478394e-02	13.922033	52.04404	0.7608337	535761.02	-0.5368641592	0.116384754	
14	4	LG화학	7.788746e-01	14.029245	60.69141	0.8481770	599999.90	-0.0700235159	0.645533425	
15	4	LG화학	4.495032e-01	13.988257	52.91047	1.0532735	743672.17	-0.2793589885	0.354202389	
16	4	LG화학	3.787921e-01	12.803410	45.18906	0.6567303	465672.64	-0.6887471173	0.309777066	
17	5	삼성바이오로직스	1.729135e-01	13.207224	42.91263	0.3263935	216273.68	-0.7725076203	0.247313736	
18	5	삼성바이오로직스	7.785556e-01	14.781913	56.06388	0.5630436	373209.77	-0.0410277984	0.674286966	
19	5	삼성바이오로직스	-1.815179e-01	10.958374	34.02902	0.2816622	186120.05	-1.4474863002	-0.045704829	
20	5	삼성바이오로직스	2.949232e-01	11.533459	39.09447	0.2300666	152473.49	-1.1014564610	0.280001970	
21	6	카카오	2.085691e-02	13.086836	33.84514	0.8601520	745821.85	-0.8862577526	-0.053410726	
22	6	카카오	8.893029e-01	14.734323	45.66822	1.4496436	1267973.05	-0.0003137229	0.416448436	
23	6	카카오	4.834642e-01	13.631452	41.06622	1.3447136	1183983.64	-0.4338392561	0.159701245	
24	6	카카오	1.087518e-01	10.794887	32.43657	0.6869152	606558.84	-1.3205234821	-0.006581617	
25	7	현대자	-4.941772e-01	12.116666	40.84478	0.7237621	1543928.74	-1.0902189707	-0.452087221	
26	7	현대자	1.584290e-01	12.203752	51.55885	0.8132853	1739071.00	-0.6560347640	-0.039120953	
27	7	현대자	9.416985e-01	15.062255	50.97332	1.3200897	2828334.81	0.2942023955	0.173559904	
28	7	현대자	1.195193e-01	11.233452	32.47110	0.9490005	2030623.31	-1.0782011838	-0.306397659	
29	8	삼성SDI	3.046497e-02	14.507782	51.32343	1.0125071	694390.18	-0.3696040219	0.125300201	

PC1 오름차순 정렬											
	brand	기업명	x1	x2	x3	x4	x5	PC1	PC2		
1	1534	동방아그로	0.041504603	6.988533	7.258589	0.07542872	10290.0656	-2.502005	-0.155776450		
2	925	진로발효	0.026919230	7.142507	7.820328	0.05510184	3645.9016	-2.473486	-0.154869151		
3	942	신라교역	-0.071935687	7.446860	8.592808	0.17278573	27689.3934	-2.432066	-0.209950947		
4	1271	신홍	-0.008542941	7.585793	6.661868	0.02853526	2751.4688	-2.423769	-0.178151886		
5	875	동일산업	0.156246196	7.194744	6.579791	0.19597980	4770.7377	-2.422452	-0.099669739		
6	456	퍼시스	0.030340759	7.513987	8.575081	0.02712107	3118.6719	-2.395697	-0.140637625		
7	925	진로발효	0.026085748	7.507959	8.919493	0.05180781	3424.7500	-2.391605	-0.141244737		
8	520	세아홀딩스	-0.685054076	7.882615	23.159517	0.08426881	3373.2742	-2.383746	-0.409839118		
9	1190	미창석유	0.108011933	7.571506	8.284566	0.03178156	554.1475	-2.360337	-0.098923143		
10	1389	인천도시가스	0.003281972	7.702972	11.667555	0.05341917	2338.2951	-2.325639	-0.127483678		
11	1389	인천도시가스	-0.322414035	8.306509	13.667206	0.03546096	1553.8548	-2.309023	-0.283307769		
12	821	일신방직	-0.713634964	8.766199	17.906943	0.11419451	2750.6774	-2.303927	-0.462503866		
13	1645	삼화왕관	0.023045475	8.179583	7.393980	0.08341101	1766.6066	-2.288752	-0.149463182		
14	1453	신영와코루	-0.563802381	8.442375	19.044411	0.19210525	1728.5645	-2.285874	-0.376229557		
15	1190	미창석유	0.082606618	7.969332	11.211377	0.03355349	584.5156	-2.255513	-0.082910870		
16	520	세아홀딩스	-0.028549635	8.075997	15.099836	0.01065663	426.1250	-2.224258	-0.108508232		
17	1526	삼양식품	0.078503734	8.120435	10.255160	0.35538035	35971.4590	-2.214554	-0.113722119		
18	1281	송일	0.022276794	8.297150	11.350781	0.13454447	8265.2295	-2.207277	-0.117605016		
19	390	빙그레	-0.053661192	8.049118	16.105320	0.23464072	23109.8197	-2.204904	-0.128882956		
20	456	퍼시스	0.059107226	8.485712	9.011434	0.01065055	1222.7869	-2.202173	-0.108263386		
21	405	신영증권	-0.268768373	8.533015	17.044036	0.07081882	6642.3871	-2.196912	-0.223932994		
22	576	농심홀딩스	0.064804654	8.059524	14.310637	0.12994388	6050.5410	-2.194039	-0.070174553		
23	1111	미원화학	0.129452920	7.752514	17.607390	0.05944403	1312.0820	-2.185911	-0.004796734		
24	41	KT	-0.496182030	8.125202	21.916931	0.33506418	875631.6452	-2.183351	-0.507401289		
25	557	LG헬로비전	0.002113919	8.103731	14.021884	0.33988143	262583.3607	-2.172324	-0.170490339		
26	1779	금비	-0.533311674	9.213045	17.822199	0.10912355	1062.4516	-2.156906	-0.357009570		
27	1047	노루홀딩스	-0.028857420	8.626700	12.501012	0.06360588	8477.4688	-2.155076	-0.129007047		
28	107	에스원	-0.047520552	8.243342	17.236554	0.18862817	71686.6230	-2.149974	-0.121440281		
29	719	KPX홀딩스	-0.506055343	9.208727	18.997165	0.03717583	1569.8710	-2.136573	-0.328480124		
30	884	제주은행	-0.829224457	9.502825	23.451752	0.09404880	30241.2097	-2.132636	-0.474641221		
31	1442	전일고속	-0.485653375	8.972765	22.413458	0.05017220	720.1935	-2.123324	-0.290441634		
32	1389	인천도시가스	0.063914841	8.560386	13.116877	0.10497936	4472.5000	-2.121390	-0.073767948		
33	1111	미원화학	0.130996529	8.330668	15.102925	0.07156896	1575.6562	-2.113580	-0.019640698		
34	389	메일유업	-0.247450493	8.725995	18.962922	0.32512267	25486.6094	-2.105227	-0.208451920		
35	1452	정상제이엘에스	0.058270389	8.684379	11.479702	0.32264836	50746.6393	-2.102527	-0.109302031		

PC2 오름차순 정렬											
	brand	기업명	x1	x2	x3	x4	x5	PC1	PC2		
1	1824	미래산업	-0.74573379	14.364130	74.32177	9.2823332	75506611	7.4863172	-16.211629		
2	647	에이비프라이오	1.34318274	20.525326	99.22914	31.4942354	75683875	11.4697128	-15.796732		
3	1311	서울식품	0.71427048	15.211040	59.10149	19.8597168	67270286	7.9903753	-14.266468		
4	584	에스맥	0.22491987	19.367090	108.74243	25.3701950	62787104	9.2618209	-13.361989		
5	1027	이화전기	0.46021238	14.967808	97.54100	7.7090372	54251551	5.6385387	-11.139325		
6	566	이아이디	0.44327285	18.013549	50.15681	8.2106664	48840209	6.1665530	-9.712718		
7	584	에스맥	0.40597675	16.335694	70.38861	16.7397706	43105372	5.6719519	-9.104361		
8	499	이트론	0.09681598	18.224854	98.34366	12.4067033	43063842	5.9588285	-8.806055		
9	185	박셀바이오	-13.76644086	41.650637	395.33562	66.6742106	4831850	9.7912232	-8.061334		
10	846	KEC	1.34764793	20.541679	120.81056	33.1472872	38818490	8.3527929	-7.927979		
11	531	조록범미디어	0.97591330	17.698629	66.70625	29.3064246	35064278	6.2751410	-7.678624		
12	1311	서울식품	-0.31250041	15.904541	63.28457	10.7382702	34512540	3.9360841	-7.494236		
13	979	빅텍	1.68200857	22.932760	116.02611	90.6547141	24621097	11.9126679	-7.361621		
14	400	신성이엔지	0.84458299	17.947696	96.13631	18.3828588	36147225	5.9482660	-7.227117		
15	851	오성첨단소재	0.96358470	19.801142	113.93572	48.4037060	30672899	8.3690927	-7.185266		
16	481	케이피엠텍	1.97653224	21.440057	118.64614	30.6310702	36063265	8.2607047	-6.893424		
17	1103	엔투텍	0.01819491	24.115205	163.36741	52.2247166	26819727	9.4280342	-6.591038		
18	1094	한솔홈데코	0.67043892	18.603687	88.47521	35.8245814	28053349	6.4608184	-6.467351		
19	531	조록범미디어	0.71534285	18.288628	82.70784	33.6757952	28285725	6.1953537	-6.448476		
20	312	우리기술투자	1.18759396	19.610382	103.45076	35.5834074	29401024	7.1589143	-6.309226		
21	469	SK증권	-0.30209524	11.363465	67.66067	6.1582560	29369313	2.3201394	-6.215590		
22	544	남선알마늄	0.36282274	17.345737	86.06468	25.5018401	27761113	5.2583790	-6.147839		
23	536	알레라	-1.46442115	39.712861	244.99812	103.7281280	13718470	15.5939838	-6.086220		
24	1329	웹크론	1.29563026	22.832688	138.58850	76.3587180	21083348	10.6254645	-5.992640		
25	1405	비케이타입	-1.78227061	22.644284	127.66938	19.4631383	24436272	5.2340485	-5.945265		
26	1311	서울식품	0.14075621	14.036427	54.72962	8.2122658	27427781	2.7638158	-5.734935		
27	481	케이피엠텍	-0.48246658	18.344391	122.87395	19.9065007	25970002	5.0418713	-5.656355		
28	594	아이에이	0.94827662	18.121667	100.17416	11.0907551	29619066	4.8841985	-5.430719		
29	1	삼성전자	-0.25128194	11.238987	33.21986	0.4154745	24784527	0.9474153	-5.266859		
30	531	조록범미디어	-0.20178761	17.568807	75.33764	26.0943159	21512586	4.3827478	-5.266121		
31	647	에이비프라이오	0.04077584	16.509007	83.32925	12.2972294	25050549	3.6680784	-5.195871		
32	891	sv인베스트먼트	0.33693279	19.356386	103.57099	38.6727631	20400160	6.1695561	-5.031357		
33	647	에이비프라이오	0.25754437	15.310952	67.77196	10.7936760	23982330	3.0911560	-4.932801		
34	856	첼온	0.69855562	20.589835	90.78333	33.7488049	21254357	6.0527045	-4.885001		
35	800	시노텍	0.76268880	18.287789	94.31908	29.9475979	21510656	5.4357607	-4.728005		
36	469	SK증권	0.88039585	14.880176	62.34181	5.3728667	25164373	2.8627029	-4.642769		
37	400	신성이엔지	0.68572404	18.301856	90.80842	12.7010422	23539771	4.2265564	-4.450868		
38	870	KNN	0.34091293	17.482693	100.56389	16.7822651	22068842	4.2585790	-4.441837		
39	462	에이프론헌제약	0.24725064	19.095519	103.64214	9.9811023	23378090	4.1647869	-4.416989		
40	1336	위지트	1.23407450	17.590796	80.57568	26.1623875	21472526	4.9939680	-4.415214		
41	1	삼성전자	0.54190787	10.494756	24.85571	0.3713184	22196523	0.7345230	-4.360828		

3. 결론

2020년도 투자성과를 분석한 결과를 요약하면 다음과 같다.

- 1) 시가총액이 높은 우량주일수록 변동성이 낮았다.
- 2) 주가가 클수록 변동성이 낮았으며, 주가가 낮을수록 유동성이 높았다.
- 3) 배당수익률이 높을수록 변동성이 낮았으나, 수익률과 유동성도 저조했다.
- 4) PER과 PBR이 높은 그룹의 수익률과 유동성이 높게 나타났으나, 변동성 또한 크게 나타났다.
- 5) 2분기의 투자성과가 좋았으나, 우량주에 투자한 경우 4분기에 성과가 좋았다.

1번과 2번의 경우 일반적으로 주식시장에서 나타나는 현상이다. 2020년 코로나 확산으로 인해 주식시장이 큰 변동성을 보였음에도 1번과 2번과 같은 현상이 유지되었음을 확인하였다.

배당수익률이 높고 PER과 PBR이 낮을수록 주가가 저평가 상태라고 판단한다. 일반적으로 저평가 상태의 주식을 매입하는 것이 투자성과가 좋게 나타난다. 하지만 작년 주식시장에서는 저배당, 고PER, PBR 주식들이 강세를 보였다. 이는 작년 주식투자를 한 투자자들이 공격적인 투자를 한 결과로 예상된다. 실제로 동학개미운동으로 인한 개인투자자들의 유입에 힘입어 코스피지수는 역사상 최고치를 기록하였다.

본 분석결과는 2020년 코로나 사태라는 특수한 상황에서의 시장분석결과이므로 향후 미래에 투자지표로 활용되기는 어렵다고 판단된다.

R 분석코드

```
rm(list=ls())
stock<-read.csv("stock_data.csv",header=T,fileEncoding = "euc-kr")
stock %>% str
stock

for(i in 8:13){
  stock[,i]<-as.factor(stock[,i])
}

#데이터의 히스토그램 확인
par(mfrow=c(2,3),family="AppleGothic") #맥 한글깨짐 방지
stock[,3] %>% hist(main="로그수익률")
stock[,4] %>% hist(main="일중변동성")
stock[,5] %>% hist(main="역사적변동성")
stock[,6] %>% hist(main="유동성")
stock[,7] %>% hist(main="거래량")

#이상치 제거 + 히스토그램 확인
stock2<-filter(stock,x4<20 & x3<100 & x1>-3 & x1<3 & x5<3500000)

par(mfrow=c(2,3),family="AppleGothic") #맥 한글깨짐 방지
stock2[,3] %>% hist(main="로그수익률")
stock2[,4] %>% hist(main="일중변동성")
stock2[,5] %>% hist(main="역사적변동성")
stock2[,6] %>% hist(main="유동성")
stock2[,7] %>% hist(main="거래량")

#파워변환
stock2<-mutate(stock2,x3=x3^0.8,x4=x4^0.2,x5=x5^0.3)

par(mfrow=c(2,3),family="AppleGothic") #맥 한글깨짐 방지
stock2[,3] %>% hist(main="로그수익률") #x1 음수포함 변환x
stock2[,4] %>% hist(main="일중변동성")
stock2[,5] %>% hist(main="역사적변동성")
stock2[,6] %>% hist(main="유동성")
stock2[,7] %>% hist(main="거래량")

#x6 시가총액
mst1<-manova(as.matrix(stock2[,3:7])~x6,data=stock2)
mst1 %>% summary(test='Wilks')
N=dim(stock2)[1]
W=(N-1)*cov(mst1$residuals)
B=(N-1)*cov(mst1$fitted.values)
solve(W)%*%B %>% eigen

(x1<-LSD.test(aov(x1~x6,data=stock2),"x6",p.adj = 'none',group=T))
(x2<-LSD.test(aov(x2~x6,data=stock2),"x6",p.adj = 'none',group=T))
(x3<-LSD.test(aov(x3~x6,data=stock2),"x6",p.adj = 'none',group=T))
```

```
(x4<-LSD.test(aov(x4~x6,data=stock2),"x6",p.adj = 'none',group=T))
(x5<-LSD.test(aov(x5~x6,data=stock2),"x6",p.adj = 'none',group=T))
```

#x7 추가

```
mst2<-manova(as.matrix(stock2[,3:7])~x7,data=stock2)
mst2 %>% summary(test='Wilks')
N=dim(stock2)[1]
W=(N-1)*cov(mst2$residuals)
B=(N-1)*cov(mst2$fitted.values)
solve(W)%*%B %>% eigen
```

```
(x1<-LSD.test(aov(x1~x7,data=stock2),"x7",p.adj = 'none',group=T))
(x2<-LSD.test(aov(x2~x7,data=stock2),"x7",p.adj = 'none',group=T))
(x3<-LSD.test(aov(x3~x7,data=stock2),"x7",p.adj = 'none',group=T))
(x4<-LSD.test(aov(x4~x7,data=stock2),"x7",p.adj = 'none',group=T))
(x5<-LSD.test(aov(x5~x7,data=stock2),"x7",p.adj = 'none',group=T))
```

#x8 배당수익률

```
mst3<-manova(as.matrix(stock2[,3:7])~x8,data=stock2)
mst3 %>% summary(test='Wilks')
N=dim(stock2)[1]
W=(N-1)*cov(mst3$residuals)
B=(N-1)*cov(mst3$fitted.values)
solve(W)%*%B %>% eigen
```

```
(x1<-LSD.test(aov(x1~x8,data=stock2),"x8",p.adj = 'none',group=T))
(x2<-LSD.test(aov(x2~x8,data=stock2),"x8",p.adj = 'none',group=T))
(x3<-LSD.test(aov(x3~x8,data=stock2),"x8",p.adj = 'none',group=T))
(x4<-LSD.test(aov(x4~x8,data=stock2),"x8",p.adj = 'none',group=T))
(x5<-LSD.test(aov(x5~x8,data=stock2),"x8",p.adj = 'none',group=T))
```

#x9 per

```
mst4<-manova(as.matrix(stock2[,3:7])~x9,data=stock2)
mst4 %>% summary(test='Wilks')
N=dim(stock2)[1]
W=(N-1)*cov(mst4$residuals)
B=(N-1)*cov(mst4$fitted.values)
solve(W)%*%B %>% eigen
```

```
(x1<-LSD.test(aov(x1~x9,data=stock2),"x9",p.adj = 'none',group=T))
(x2<-LSD.test(aov(x2~x9,data=stock2),"x9",p.adj = 'none',group=T))
(x3<-LSD.test(aov(x3~x9,data=stock2),"x9",p.adj = 'none',group=T))
(x4<-LSD.test(aov(x4~x9,data=stock2),"x9",p.adj = 'none',group=T))
(x5<-LSD.test(aov(x5~x9,data=stock2),"x9",p.adj = 'none',group=T))
```

#x10 pbr

```
mst5<-manova(as.matrix(stock2[,3:7])~x10,data=stock2)
mst5 %>% summary(test='Wilks')
N=dim(stock2)[1]
```

```
W=(N-1)*cov(mst5$residuals)
B=(N-1)*cov(mst5$fitted.values)
solve(W)%*%B %>% eigen
```

```
(x1<-LSD.test(aov(x1~x10,data=stock2),"x10",p.adj = 'none',group=T))
(x2<-LSD.test(aov(x2~x10,data=stock2),"x10",p.adj = 'none',group=T))
(x3<-LSD.test(aov(x3~x10,data=stock2),"x10",p.adj = 'none',group=T))
(x4<-LSD.test(aov(x4~x10,data=stock2),"x10",p.adj = 'none',group=T))
(x5<-LSD.test(aov(x5~x10,data=stock2),"x10",p.adj = 'none',group=T))
```

#x11 투자시기

```
mst6<-manova(as.matrix(stock2[,3:7])~x11,data=stock2)
mst6 %>% summary(test='Wilks')
N=dim(stock2)[1]
W=(N-1)*cov(mst6$residuals)
B=(N-1)*cov(mst6$fitted.values)
solve(W)%*%B %>% eigen
```

```
(x1<-LSD.test(aov(x1~x11,data=stock2),"x11",p.adj = 'none',group=T))
(x2<-LSD.test(aov(x2~x11,data=stock2),"x11",p.adj = 'none',group=T))
(x3<-LSD.test(aov(x3~x11,data=stock2),"x11",p.adj = 'none',group=T))
(x4<-LSD.test(aov(x4~x11,data=stock2),"x11",p.adj = 'none',group=T))
(x5<-LSD.test(aov(x5~x11,data=stock2),"x11",p.adj = 'none',group=T))
```

#주성분 분석

#상관행렬 플랏

```
stock[,3:7] %>% cor %>% ggcorrplot(lab = T)
```

#분산 비율

```
prcomp(stock[,3:7],scale=T) %>% summary
```

#screeplot

```
screeplot(a,type='l',main="")
abline(h=1)
screeplot(a,main = "")
```

#eigenvalue와 주성분점수

```
a<-prcomp(stock[,3:7],scale=T)
a$x
```

#biplot

```
biplot(a,c(1,2))
```

#주성분점수

```
b<-cbind(stock[,1:7],a$x[,1:2])
arrange(b,PC1)
arrange(b,PC2)
```