

# 트위터, 대선 테마주와 연관성이 있을까?



2조

최승우 김준용 문진희 박민정

# Content

- Introduction
- Choose the Election-oriented Stocks
- Simple Correlation Analysis
- Principal Component Analysis
  - Index vs. Twit. Regression
- Volatility Measure
  - Volat. vs. Twit. Regression
- Volume vs. Twit. Regression
- Conclusion



# Process Distribution

## **Jin-hee, Moon**

- BE in Industrial Engineering
- Correlation Analysis for Election-oriented Stocks of Each President Candidate (SAS EG, SAS 9.2)
- Plot Time-series Data for Presentation

## **Seung-woo, Choi**

- BE in Computer Engineering
- Principal Component Analysis for the stocks; especially plot the Eigenvalue graphs (XLMiner)
- Collect Entire Presentation Data

## **Jun-yong Kim**

- BBA in Business Administration
- Regression Analysis between Twitter information and the profile of stocks (SAS 9.2)
- Compose PPT and Make a Presentation

## **Min-jeong, Park**

- BBA in Technology Management
- Principal Component Analysis for the stocks; especially compute each principal component for the candidate (WEKA)
- Design and Outline PPT



# Introduction

## Our Claim

대선과 금융 시장은  
과연 얼마나 밀접한가?



# 문재인 후보 테마주

기업	사유
우리들제약	이상호 우리들병원 병원장이 노무현 전 대통령의 주치의
서희건설	이봉관 서희건설 회장이 경희대 동문
유성티엔에스	경희대 동문인 이봉관 서희건설 회장이 최대주주
바른손	문재인 후보의 전 법무법인이 법률고문 맡았었음
위노바	노무현 전 대통령 주치의였던 이상호의 아들 이승렬씨가 사장
에이엔피	친노계 인사 송철호 변호사가 비상근 이사

# 박근혜 후보 테마주

기업	사유
EG	박근혜 동생인 박지만씨가 이 회사 회장
비트컴퓨터	조현정 대표이사 회장이 한나라당(현 새누리당) 비대위 위원
대유에이텍	박근혜 조카사위가 이 회사 최대주주인 박영우 회장
아가방컴퍼니	저출산 복지 정책 관련
보령메디앙스	저출산 복지 정책 관련

# 안철수 후보 테마주

기업	사유
미래산업	안원장과 정문술 회장 친분
안랩	안철수 원장 설립사 최대주주 지분 28.56%
솔고바이오	사외이사가 안원장과의 친분
링네트	이주석 대표, 스탠퍼드 동문 일자리 창출 스마트워크 관련주
우성사료	최대주주가 신경민 의원과 사돈지간, 신경민 의원 청춘 콘서트 멘토
오픈베이스	정진섭 회장이 스탠퍼드대 출신 안랩과 사업 연관성
노루페인트	안경수 회장이 서울대, 스탠퍼드 대학원 출신

# 분석 방향



주가  
Stock Price



거래량  
Trading Volume



변동성  
Volatility

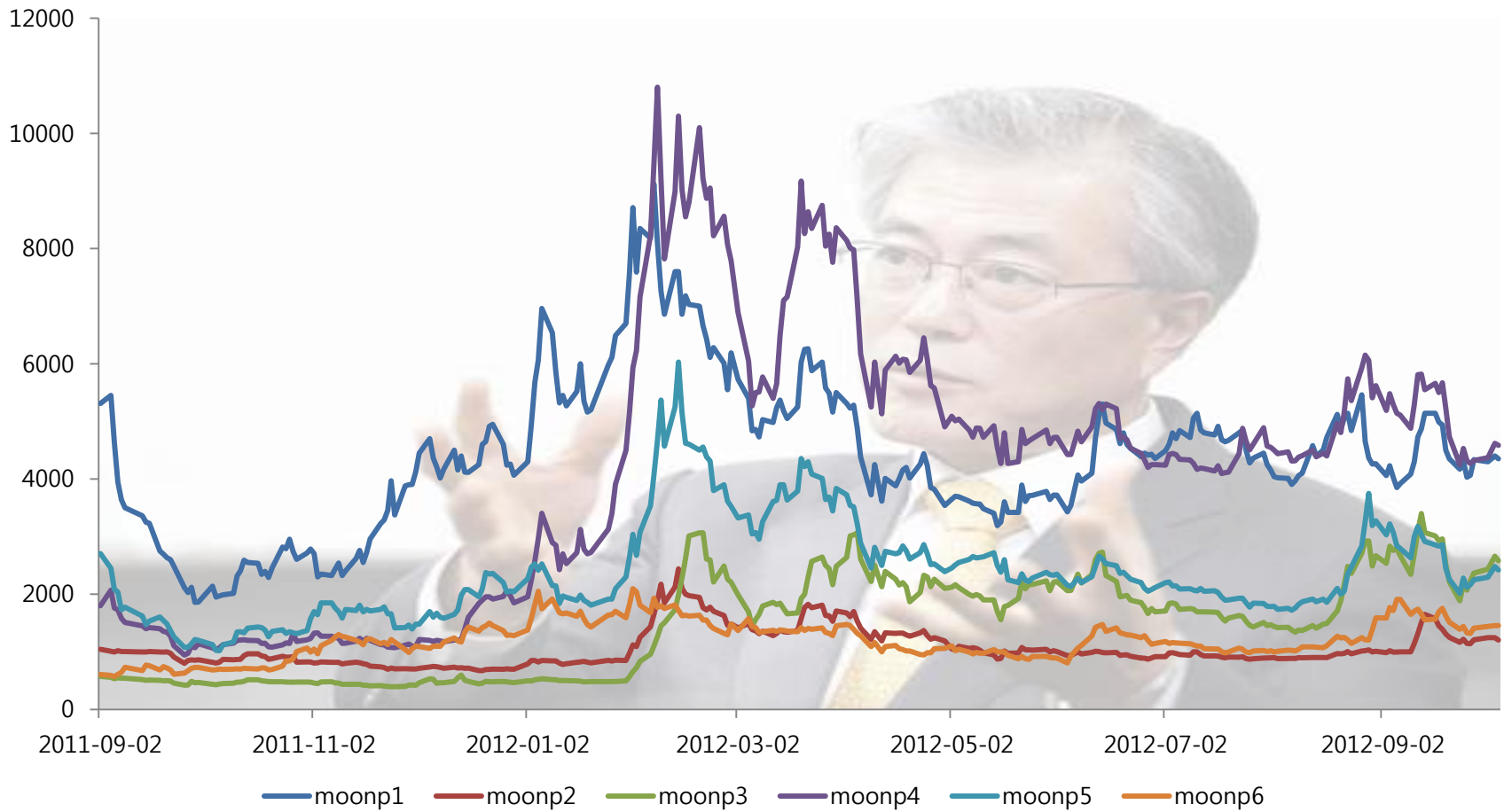


# Chapter 1 Stock Prices!

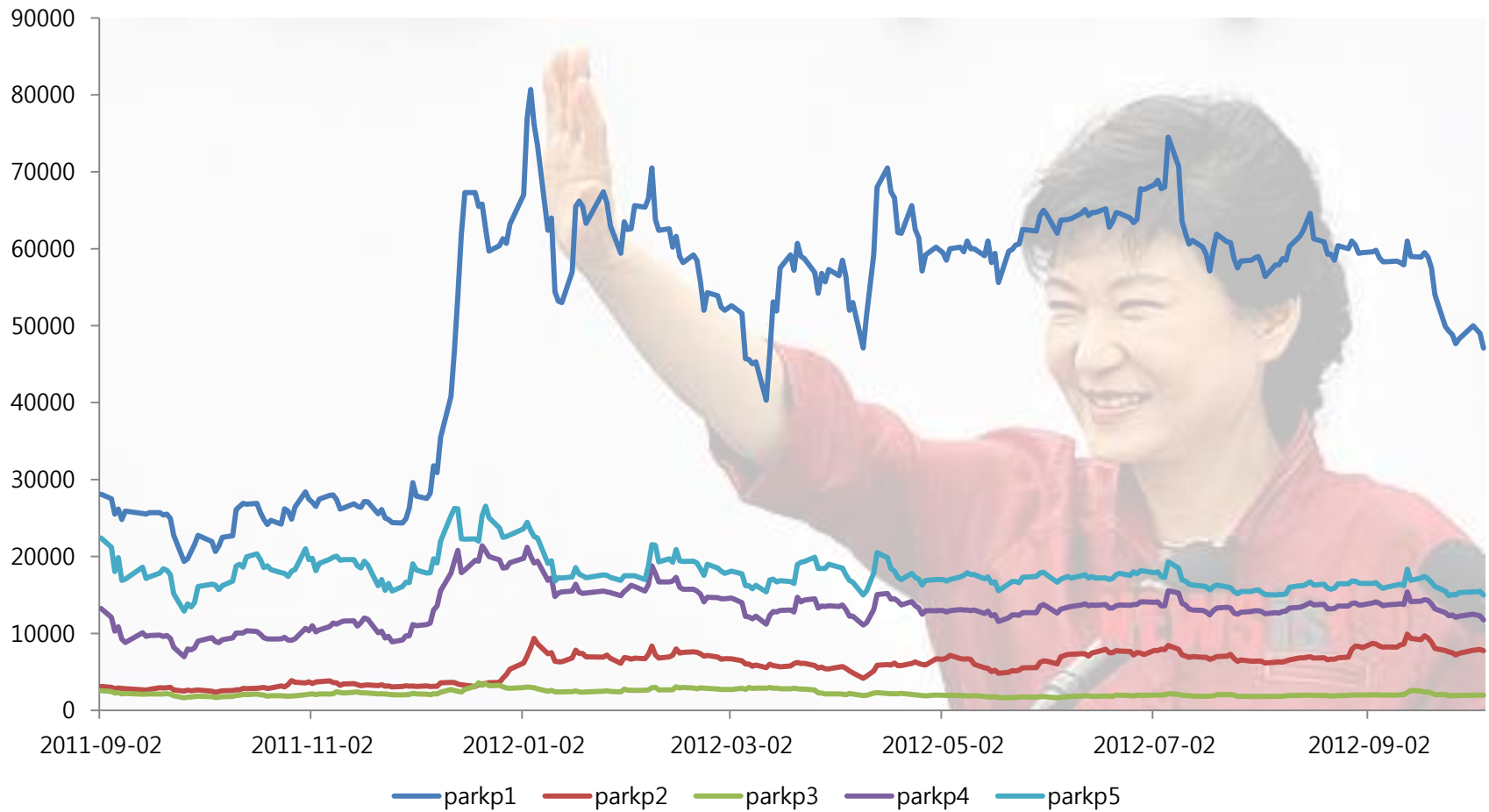


Does TWITTER MENTION matter in Stock Pricing?

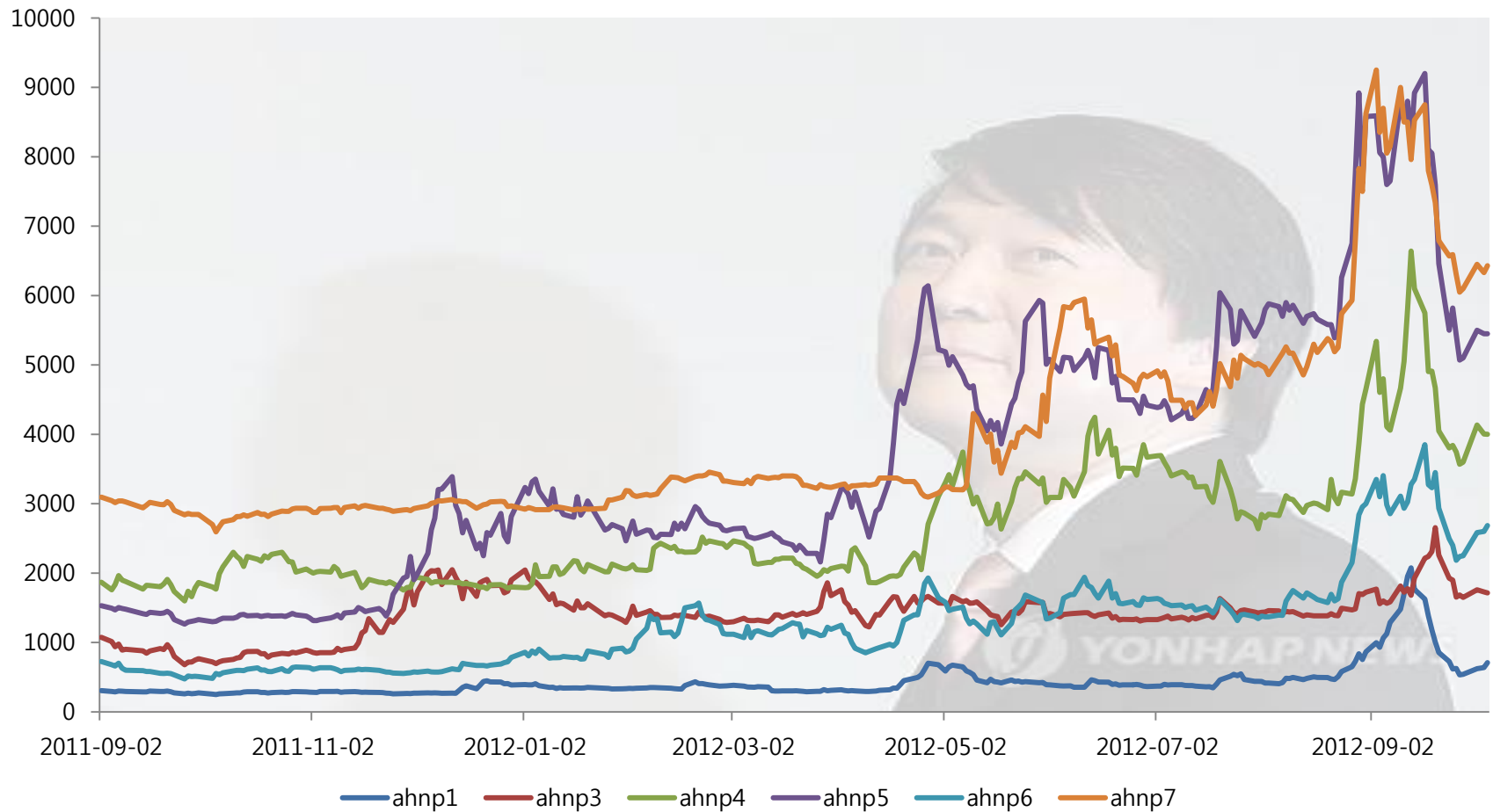
# 테아주 트렌드 - 문재인



# 테아주 트렌드 - 박근혜



# 테마주 트렌드 - 안철수





# 테마주 트렌드 - 안철수

$$\text{Corr}(x_t, y_t) = \frac{\text{Cov}(x_t, y_t)}{\sqrt{\text{Var}(x_t)} \sqrt{\text{Var}(y_t)}}$$

$$\text{Var}(x_t) = (T-1)^{-1} \sum_{t=1}^T (x_t - \bar{x})^2$$

$$\text{Cov}(x_t, y_t) = (T-1)^{-1} \sum_{t=1}^T (x_t - \bar{x})(y_t - \bar{y})$$

# 상관관계 - 문재인

피어슨 상관 계수, N = 271  
H0: Rho=0 가정하에서 Prob > |r|

	moonr1	moonr2	moonr3	moonr4	moonr5	moonr6
moonr1 유성티앤에스	1.00000	0.51357 <.0001	0.39682 <.0001	0.56363 <.0001	0.51073 <.0001	0.35194 <.0001
moonr2 서희건설	0.51357 <.0001	1.00000	0.56932 <.0001	0.57516 <.0001	0.64737 <.0001	0.27245 <.0001
moonr3 우리들제약	0.39682 <.0001	0.56932 <.0001	1.00000	0.52980 <.0001	0.58801 <.0001	0.23880 <.0001
moonr4 바른손	0.56363 <.0001	0.57516 <.0001	0.52980 <.0001	1.00000	0.63754 <.0001	0.29760 <.0001
moonr5 위노바	0.51073 <.0001	0.64737 <.0001	0.58801 <.0001	0.63754 <.0001	1.00000	0.28094 <.0001
moonr6 에미엔피	0.35194 <.0001	0.27245 <.0001	0.23880 <.0001	0.29760 <.0001	0.28094 <.0001	1.00000

# 상관관계 - 박근혜

피어슨 상관 계수, N = 271  
H0: Rho=0 가정하에서 Prob > |r|

	parkr1	parkr2	parkr3	parkr4	parkr5
parkr1 EG	1.00000	0.50621 <.0001	0.44940 <.0001	0.70845 <.0001	0.65497 <.0001
parkr2 비트컴퓨터	0.50621 <.0001	1.00000	0.45017 <.0001	0.45918 <.0001	0.48248 <.0001
parkr3 대유에이텍	0.44940 <.0001	0.45017 <.0001	1.00000	0.50467 <.0001	0.52220 <.0001
parkr4 아가방컴퍼니	0.70845 <.0001	0.45918 <.0001	0.50467 <.0001	1.00000	0.91170 <.0001
parkr5 보령메디앙스	0.65497 <.0001	0.48248 <.0001	0.52220 <.0001	0.91170 <.0001	1.00000

# 상관관계 - 안철수

피어슨 상관 계수, N = 271  
H0: Rho=0 가정하에서 Prob > |r|

	ahnr1	ahnr2	ahnr3	ahnr4	ahnr5	ahnr6	ahnr7
ahnr1 미래산업	1.00000	0.13364 0.0278	0.19480 0.0013	0.41066 <.0001	0.34879 <.0001	0.37675 <.0001	0.31264 <.0001
ahnr2 안랩	0.13364 0.0278	1.00000	0.55283 <.0001	0.17341 0.0042	0.53628 <.0001	0.24252 <.0001	0.15904 0.0087
ahnr3 슬고바이오	0.19480 0.0013	0.55283 <.0001	1.00000	0.24779 <.0001	0.63323 <.0001	0.37267 <.0001	0.30223 <.0001
ahnr4 링네트	0.41066 <.0001	0.17341 0.0042	0.24779 <.0001	1.00000	0.25787 <.0001	0.37269 <.0001	0.33637 <.0001
ahnr5 우성사료	0.34879 <.0001	0.53628 <.0001	0.63323 <.0001	0.25787 <.0001	1.00000	0.39917 <.0001	0.37155 <.0001
ahnr6 오픈베이스	0.37675 <.0001	0.24252 <.0001	0.37267 <.0001	0.37269 <.0001	0.39917 <.0001	1.00000	0.34220 <.0001
ahnr7 노루페인트	0.31264 <.0001	0.15904 0.0087	0.30223 <.0001	0.33637 <.0001	0.37155 <.0001	0.34220 <.0001	1.00000



# PCA(Principle Component Analysis)

For a given matrix  $A$

$$Ax = \lambda x$$

$$\det(A - \lambda I) = 0$$

Eigenvalue of  $A$  is  $\lambda_{1,2, \dots}$  and Eigenvector of  $A$  is  $x$

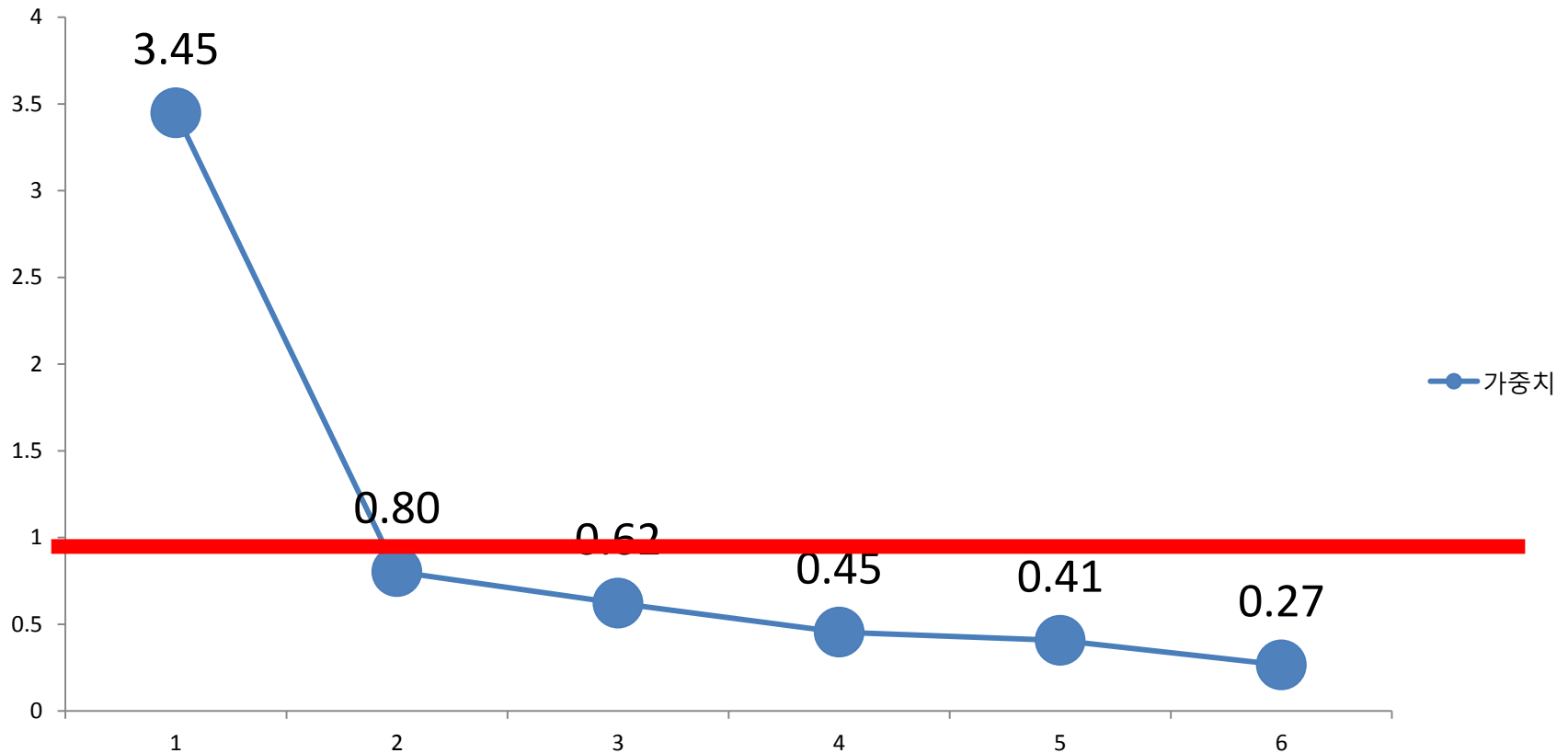
$$P_{\text{문재인}} = 0.40S_1 + 0.33S_2 + 0.45S_3 + 0.48S_4 + 0.47S_5 + 0.23S_6$$

$$P_{\text{박근혜}} = 0.44S_1 + 0.43S_2 + 0.34S_3 + 0.50S_4 + 0.49S_5$$

$$P_{\text{안철수}} = 0.33S_1 + 0.40S_2 + 0.41S_3 + 0.29S_4 + 0.48S_5 + 0.43S_6 + 0.17S_7$$

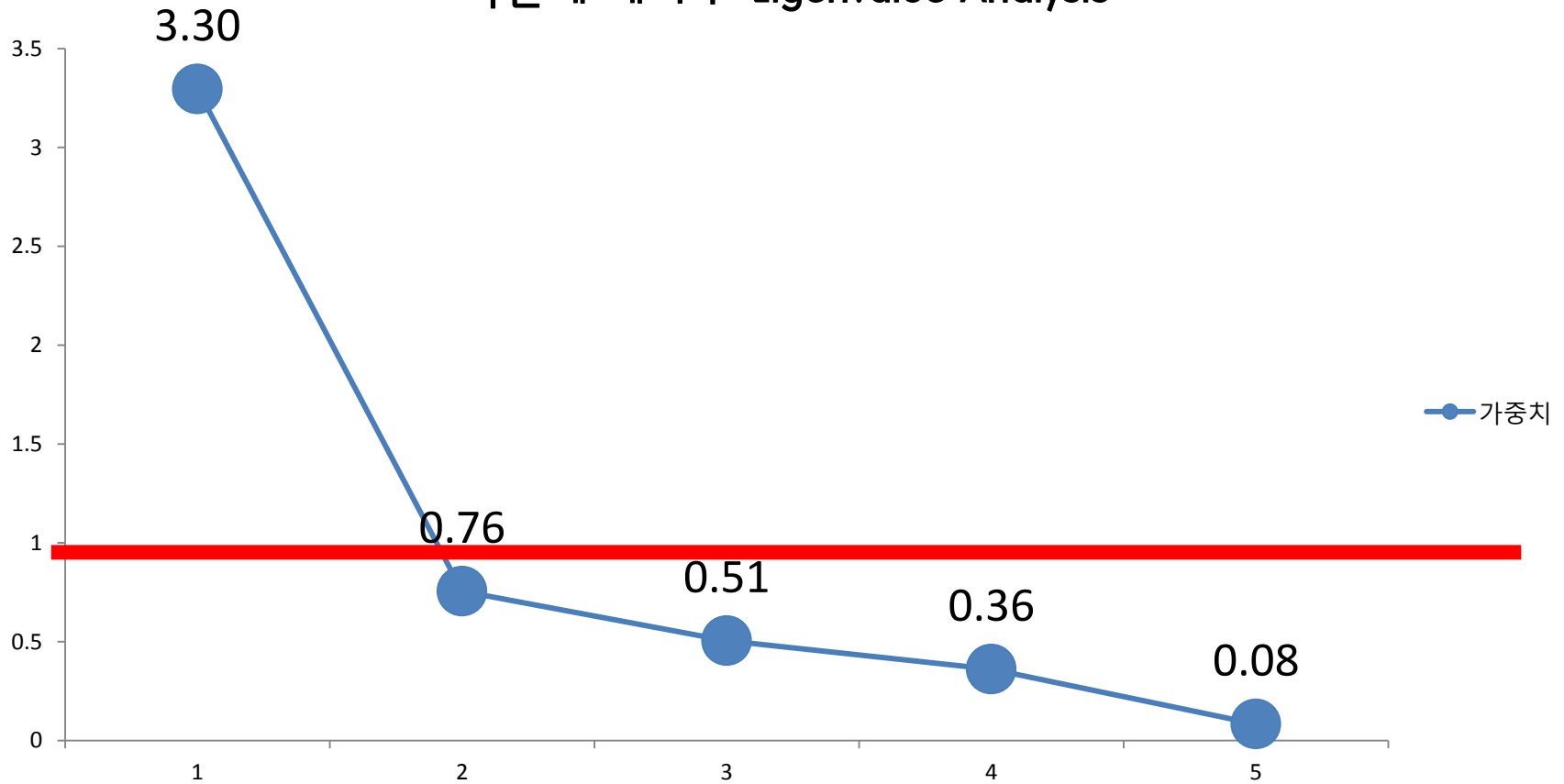
# PCA(Principle Component Analysis)

문재인 테마주 Eigenvalue Analysis



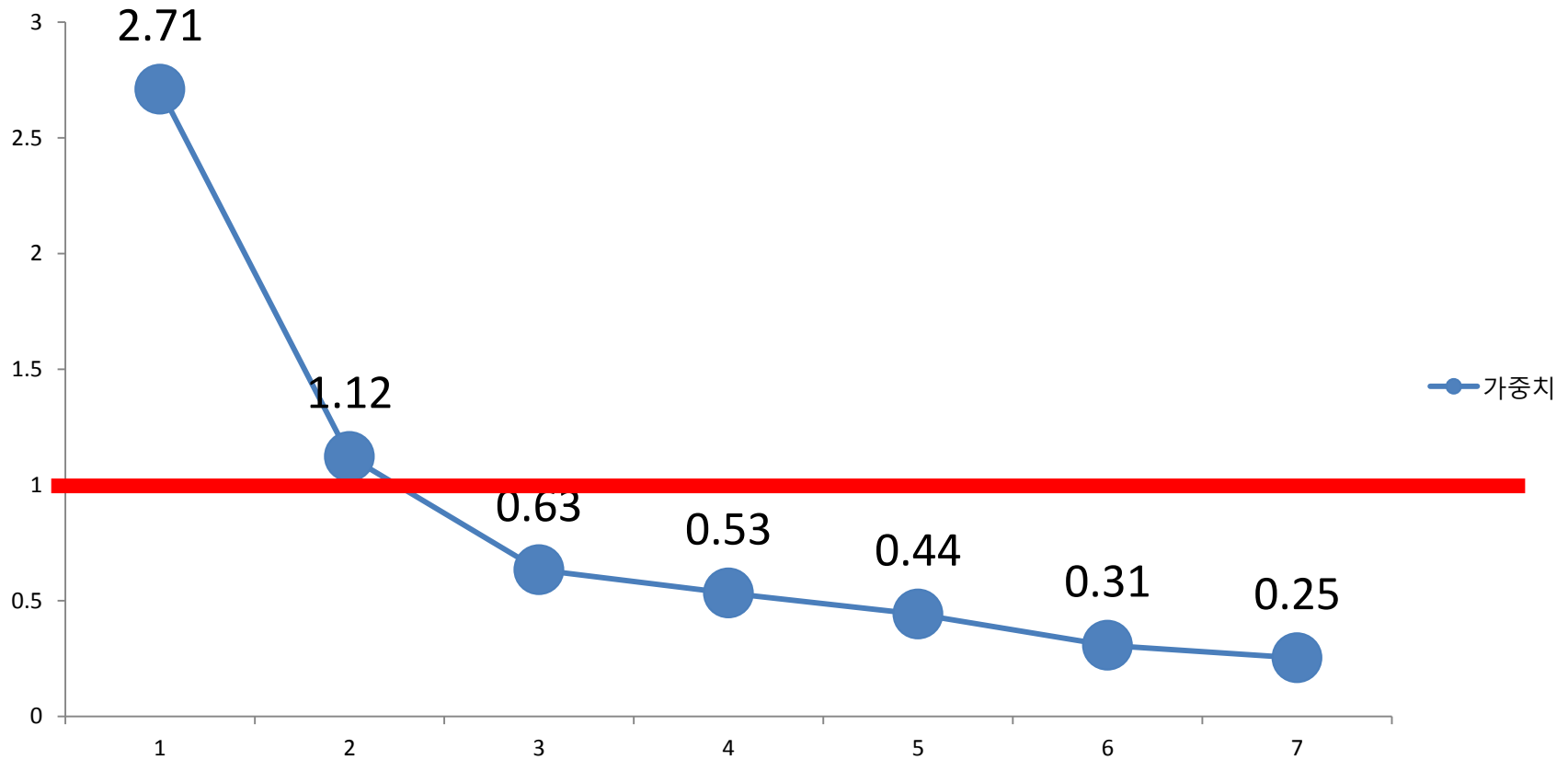
# PCA(Principle Component Analysis)

박근혜 테마주 Eigenvalue Analysis



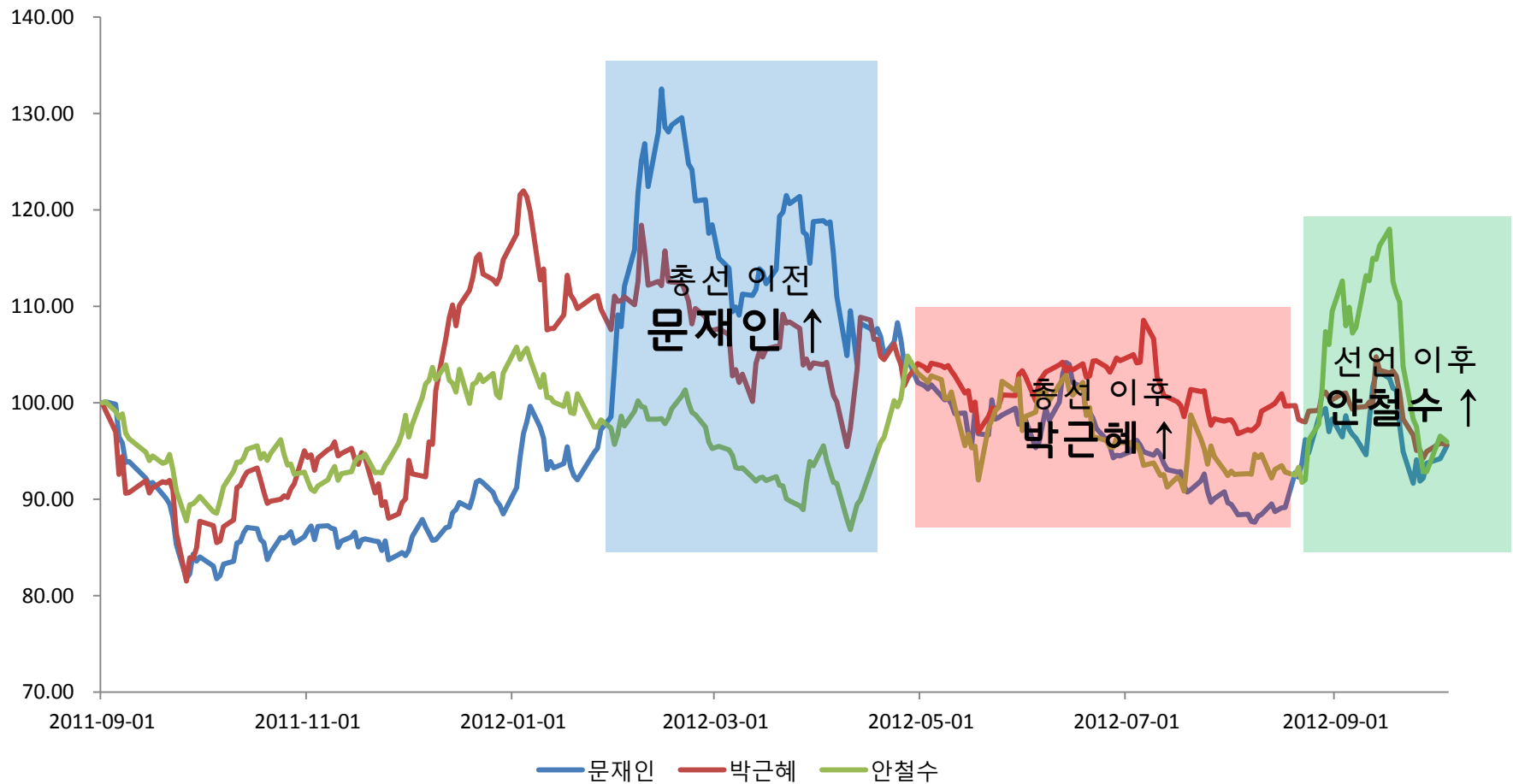
# PCA(Principle Component Analysis)

안철수 테마주 Eigenvalue Analysis





# Result – 주성분 분석



# Regression Model

$$P = f(TWEET) \rightarrow P_{i,t} = \alpha + \beta TWEET_{i,t} + u_{i,t}$$

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	Intercept	1	0.07255	0.12464	0.58	0.5610
moon	문재인트위터	1	-0.11332	0.07103	-1.60	0.1118

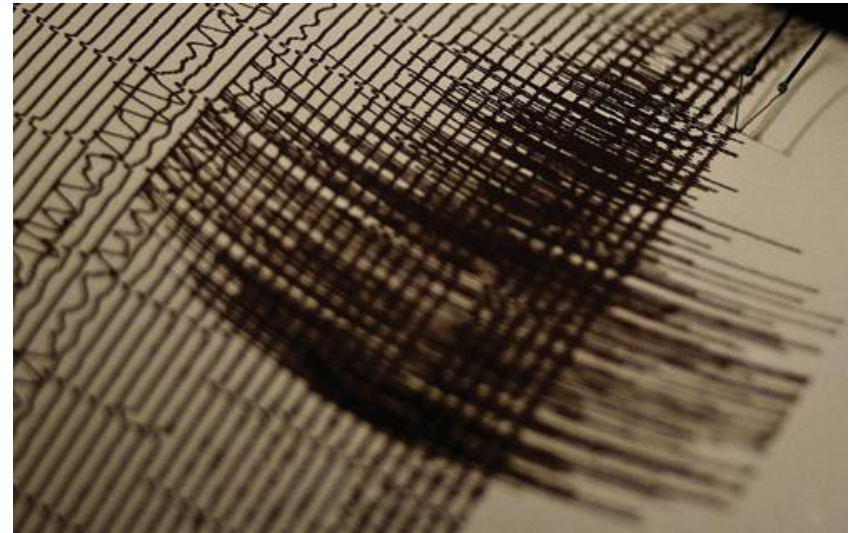
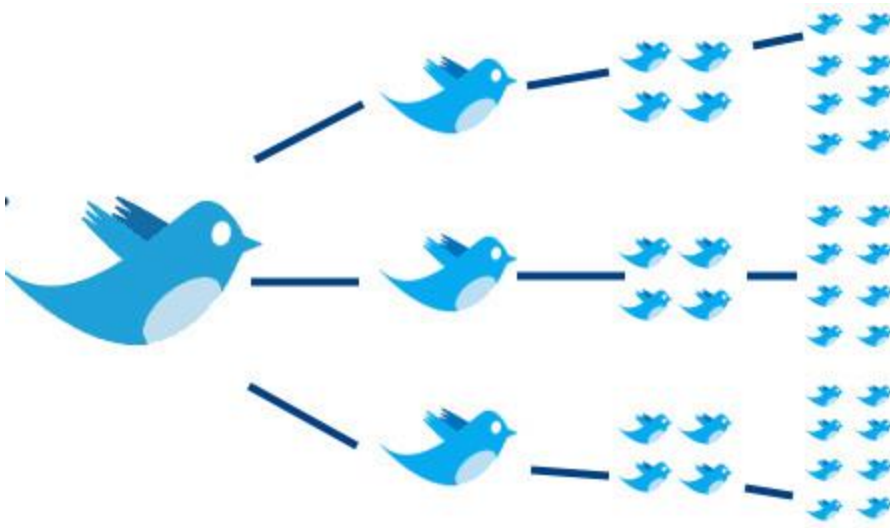
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	Intercept	1	0.04329	0.14292	0.30	0.7622
park	박근혜트위터	1	-0.02837	0.04325	-0.66	0.5124

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	Intercept	1	0.08882	0.11895	0.75	0.4558
ahn	안철수트위터	1	-0.07062	0.03701	-1.91	0.0574

**안철수 테마주 지수만 트위터에 영향을 받는다!**

(그러나 그 영향력이 극히 미미하다. 약 0.07%,  $R^2=0.0134$ )

# Chapter 2 Volatility!



과연 트위터 멘션이 정말 변동성을 유발하는 요인일까?

# Volatility Model = GARCH(1,1)?

$$y_t = \mu + \varepsilon_t \quad \varepsilon_t \sim N(0, h_t)$$

$$h_t := E(\sigma_t^2)$$

Variable	DF	Estimate	Standard Error	t Value	Approx Pr >  t
Intercept	1	0.001586	0.002891	0.55	0.5833
ARCH0	1	0.000000	0.000179	3.59	0.0003
ARCH1	1	0.3877	0.1337	2.90	0.0037
GARCH1	1	0.4653	0.1297	3.59	0.0003

$$= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

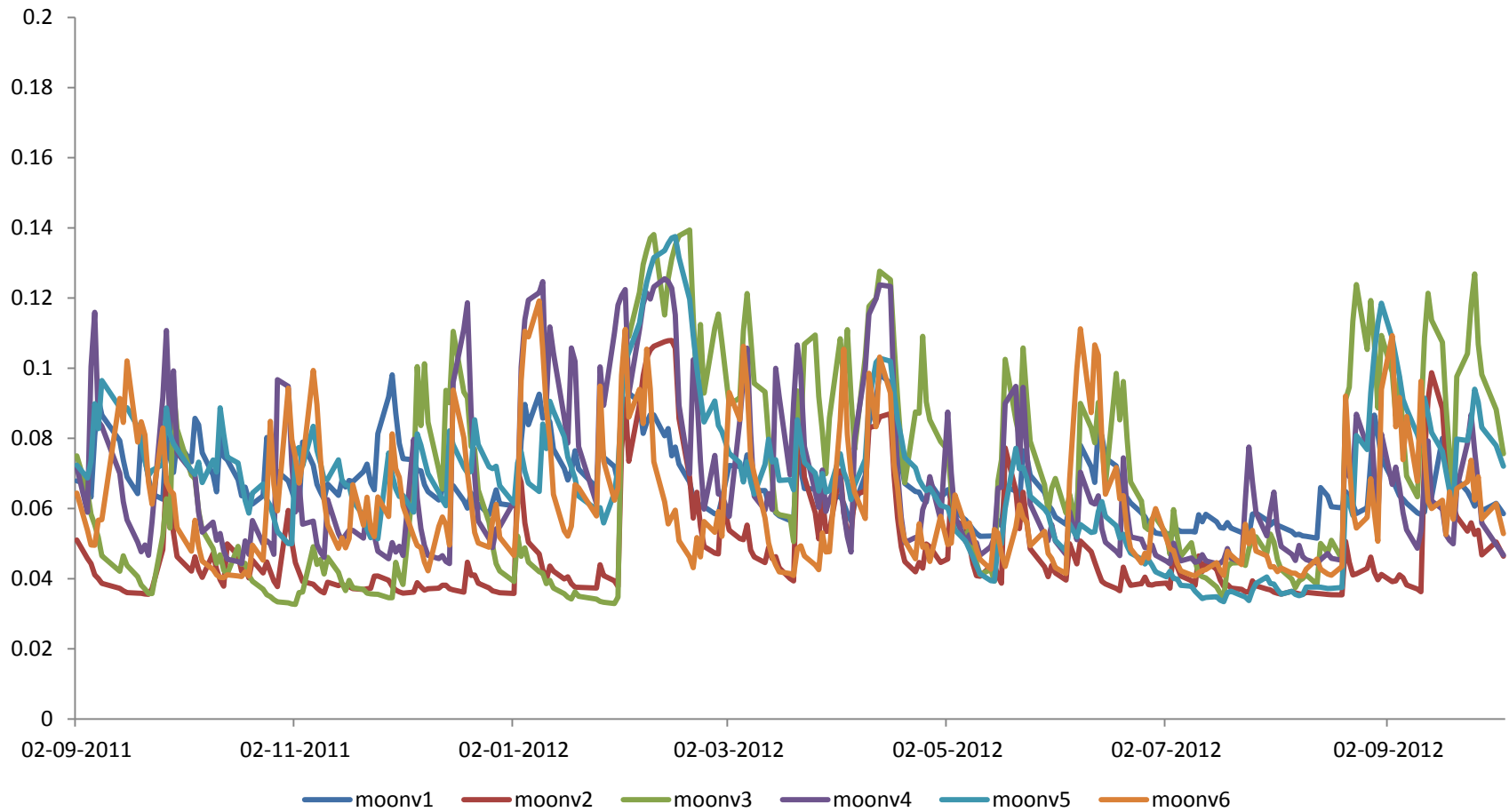
Tim Bollerslev, 1986, "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*

1. 표본 내 시간 가변 분산 구조 설명 가능 (본 연구에서 사용)
2. 추정된 모형을 바탕으로 미래 변동성 예측 (본 연구에선 하지 않음)

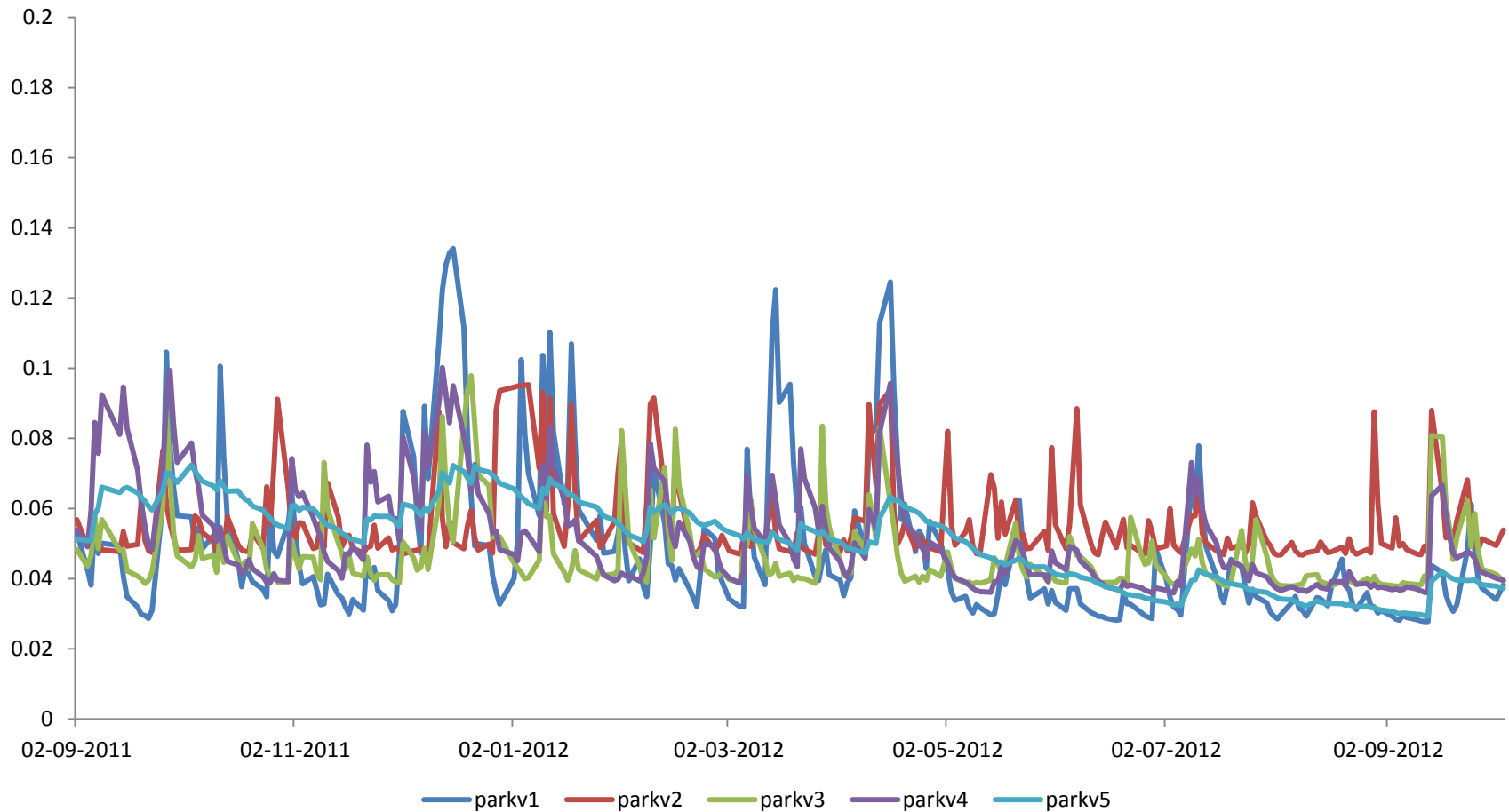




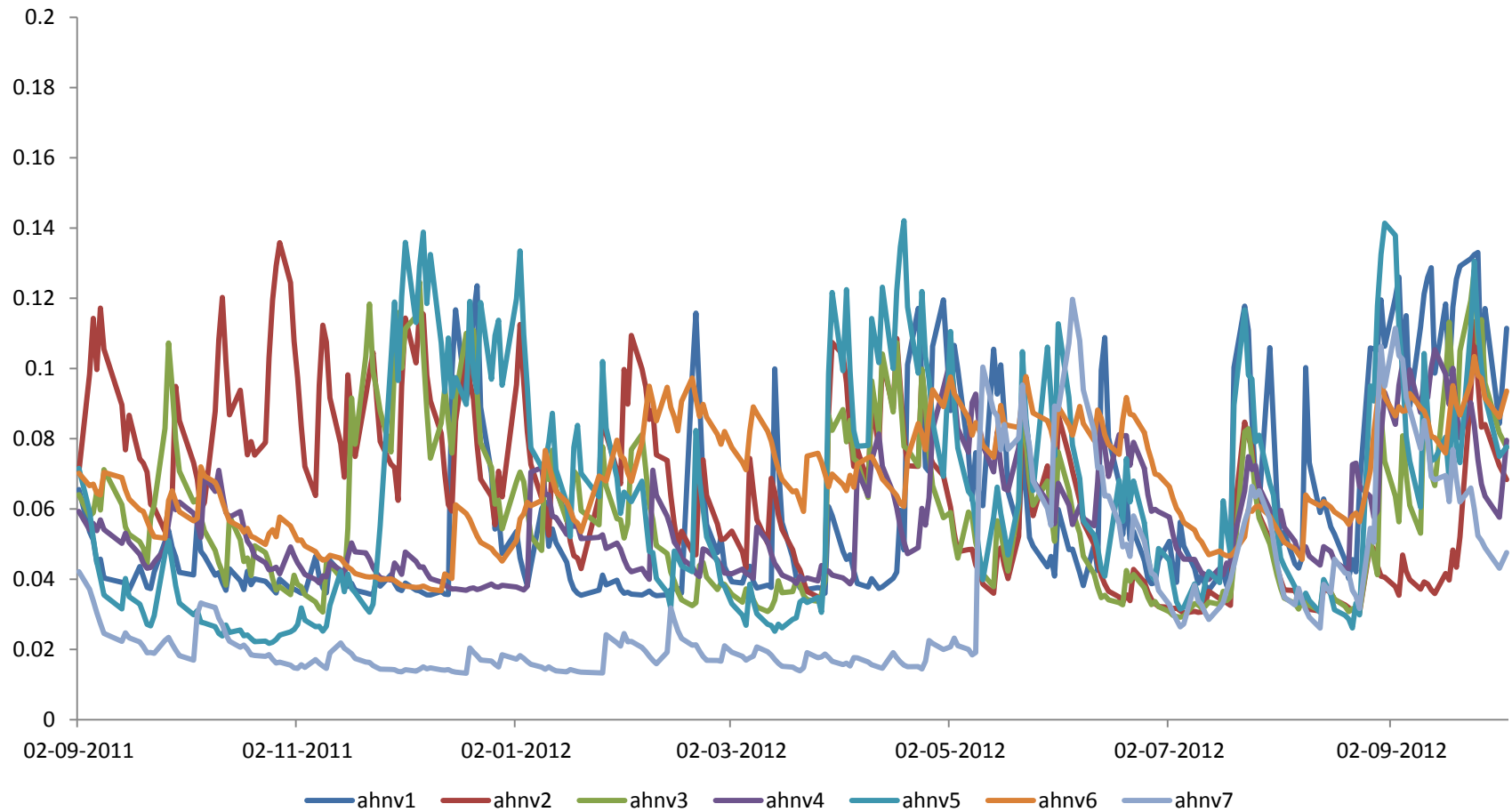
# 문재인 테마주 변동성



# 박근혜 테마주 변동성



# 안철수 테마주 변동성



# Analysis Methodology

$$VOLAT_{\text{유성티엔에스},t} = \alpha + \beta TWEET_{\text{문재인},t} + u_t$$

$$VOLAT_{\text{서희건설},t} = \alpha + \beta TWEET_{\text{문재인},t} + u_t$$

$$VOLAT_{\text{우리들제약},t} = \alpha + \beta TWEET_{\text{문재인},t} + u_t$$

...

$$VOLAT_{EG,t} = \alpha + \beta TWEET_{\text{박근혜},t} + u_t$$

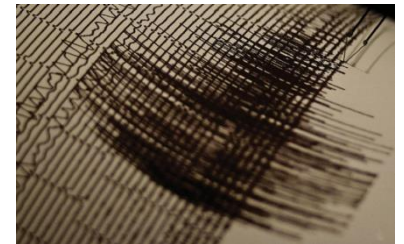
$$VOLAT_{\text{비트컴퓨터},t} = \alpha + \beta TWEET_{\text{박근혜},t} + u_t$$

$$VOLAT_{\text{대유에이텍},t} = \alpha + \beta TWEET_{\text{박근혜},t} + u_t$$

...

$$VOLAT_{\text{노루페인트},t} = \alpha + \beta TWEET_{\text{안철수},t} + u_t$$

종속변수?





# Analysis Methodology

$$VOLAT_{미래산업,t} = \alpha + \beta TWEET_{안철수,t} + u_t$$

$$VOLAT_{미래산업,t} = \alpha + \beta TWEET_{안철수,t} + u_t$$

$$H_0 : \beta_1 = \beta_2 \quad H_1 : \beta_1 \neq \beta_2$$

총선 이전과 총선 이후에 베타 계수는 다르게 나타나는가?  
구조 변화 분석 (Structural Change Analysis)



# Result

후보	테마주	총선 이전	유의수준	총선 이후	유의수준	전체	유의수준
문재인	유성티엔에스	0.543	0.846	0.695	0.279	-0.709	0.262
	서희건설	19.393	0.000	2.272	0.010	2.273	0.022
	우리들제약	39.696	0.000	5.014	0.003	6.893	0.000
	바른손	31.363	0.000	0.182	0.869	-0.837	0.554
	위노바	5.818	0.281	5.498	0.000	2.117	0.111
	에이엔피	15.418	0.007	2.092	0.070	1.416	0.217
박근혜	EG	2.283	0.446	-0.719	0.332	-2.644	0.002
	비트컴퓨터	1.831	0.269	0.127	0.782	-0.215	0.635
	대유에이텍	2.325	0.132	0.701	0.070	-0.215	0.612
	아가방컴퍼니	-2.186	0.246	-0.040	0.936	-2.310	0.000
	보령메디앙스	-3.745	0.000	-1.522	0.000	-4.263	0.000
안철수	미래산업	-3.236	0.349	7.040	0.000	8.600	0.000
	안랩	10.297	0.015	1.278	0.099	-1.383	0.104
	솔고바이오	-0.882	0.846	5.202	0.000	3.864	0.000
	링네트	2.234	0.198	2.586	0.000	4.331	0.000
	우성사료	-1.047	0.872	3.711	0.001	4.423	0.000
	오픈베이스	0.651	0.824	1.820	0.002	2.858	0.000
	노루페인트	4.040	0.000	2.744	0.005	5.987	0.000

# Result

후보	테마주	총선 이전	유의수준	총선 이후	유의수준	전체	유의수준
문재인	유성티엔에스	2.703	0.961	3.519	0.471	-17.062	0.082
	서희건설	1103.580	0.000	170.269	0.000	183.817	0.000
	우리들제약	1983.810	0.000	157.425	0.009	283.534	0.000
	바른손	511.892	0.000	2.012	0.858	-4.175	0.839
	위노바	107.232	0.074	38.069	0.000	27.912	0.014
	에이엔피	114.594	0.140	135.714	0.000	138.944	0.000
박근혜	EG	8.561	0.180	-0.872	0.399	-3.323	0.039
	비트컴퓨터	2.360	0.885	21.361	0.000	18.009	0.000
	대유에이텍	75.216	0.044	29.428	0.001	9.055	0.372
	아가방컴퍼니	-82.316	0.064	8.954	0.049	-45.103	0.000
	보령메디앙스	-17.694	0.066	0.502	0.608	-11.343	0.000
안철수	미래산업	-229.161	0.127	1317.747	0.000	1603.972	0.000
	안랩	51.168	0.004	2.884	0.064	-4.738	0.103
	솔고바이오	88.466	0.694	249.327	0.000	188.997	0.000
	링네트	11.804	0.086	-1.441	0.823	15.902	0.001
	우성사료	0.685	0.984	28.136	0.011	50.931	0.000
	오픈베이스	-9.640	0.809	50.433	0.011	104.168	0.000
	노루페인트	-0.184	0.541	9.721	0.023	19.245	0.000

# 상관관계 – 안철수(총선 이후)

피어슨 상관 계수, N = 120  
H0: Rho=0 가정하에서 Prob > |r|

	ahnr1	ahnr2	ahnr3	ahnr4	ahnr5	ahnr6	ahnr7
ahnr1 미래산업	1.00000	0.32807 0.0003	0.20264 0.0264	0.47620 <.0001	0.47470 <.0001	0.45365 <.0001	0.32532 0.0003
ahnr2 안랩	0.32807 0.0003	1.00000	0.76388 <.0001	0.40606 <.0001	0.70026 <.0001	0.41516 <.0001	0.37455 <.0001
ahnr3 솔고바이오	0.20264 0.0264	0.76388 <.0001	1.00000	0.35000 <.0001	0.69264 <.0001	0.53197 <.0001	0.42341 <.0001
ahnr4 링네트	0.47620 <.0001	0.40606 <.0001	0.35000 <.0001	1.00000	0.34816 <.0001	0.55775 <.0001	0.37227 <.0001
ahnr5 우성사료	0.47470 <.0001	0.70026 <.0001	0.69264 <.0001	0.34816 <.0001	1.00000	0.59430 <.0001	0.50246 <.0001
ahnr6 오픈베이스	0.45365 <.0001	0.41516 <.0001	0.53197 <.0001	0.55775 <.0001	0.59430 <.0001	1.00000	0.44216 <.0001
ahnr7 노루페인트	0.32532 0.0003	0.37455 <.0001	0.42341 <.0001	0.37227 <.0001	0.50246 <.0001	0.44216 <.0001	1.00000

## [특징주] 박근혜 테마주, 상승장에서도 '비실'

머니투데이 이현수 기자 | 입력 : 2012.09.07 09:30

공유:  +1 { 0 }  Tweet { 0 }  좋아요 { 0 } | 소셜댓글: 0

'박근혜 관련주'로 분류된 정치테마주들이 코스피가 2% 넘게 오르는 상승장에서도 하락세를 보이고 있다. 안철수 교수 측의 폭로에 따른 것으로 풀이된다.

7일 오전 9시 28분 현재 EG (47,100원 ▼1900 -3.9%)는 전일 대비 0.34% 떨어진 5만8100원에 거래되고 있다. 보령메디앙스 (15,000원 ▼450 -2.9%)와 비트컴퓨터 (7,750원 ▼170 -2.1%)는 각각 0.32%, 0.49% 내리고 있다.



전날 안철수 서울대 융합과학기술대학원장 측 금태섭 변호사는 긴급 기자회견을 열고 "박근혜 후보 측 인사가 전화를 걸어와 뇌물과 여자 문제를 폭로하겠다고 협박하면서 안 원장의 대선 불출마를 종용했다"고 주장했다.

한편 이날 안랩 (81,600원 ▲300 0.4%)은 2.76% 오른 12만2800원에 거래되고 있다.

[경제]

게재 일자 : 2012년 09월 11일(火)

## 안철수 테마주, 1756% 뚫은 종목도 있다

이상급등 대선 테마주 '허실' ...2분기 영업이익은 -4000억



유력 대선 주자 및 정책들과 관련된 이른바 '테마주'들이 실적 부진에도 불구하고 지난 8개월여 동안 평균 226%나 급등한 것으로 나타났다.

일부 종목의 경우 최대 1756%까지 급등해 한국거래소가 '투자 주의' 경고를 내렸는데도 주가는 요지부동이다. 이에 따라 금융감독 당국은 테마주 급등 과정에 [시세조종](#) 세력 개입 여부 등 불공정거래 혐의를 정밀 조사하고 있는 것으로 알려졌다.

11일 거래소와 [금융감독](#)원에 따르면 지난해 말부터 9월6일까지 대표적인 정치인 테마주 16개 종목의 주가 등락을 비교 분석한 결과, 8개월여 동안 평균 226.17%나 급등한 것으로 조사됐다. 같은 기간 코스피지수와 코스닥지수는 각각 3.04%, 1.31% 상승하는데 그쳤다.

### 주요 대선주자 관련 테마주 주가 등락률 현황 (단위:%) ※2011년 말 종가와 9월6일 종가 비교

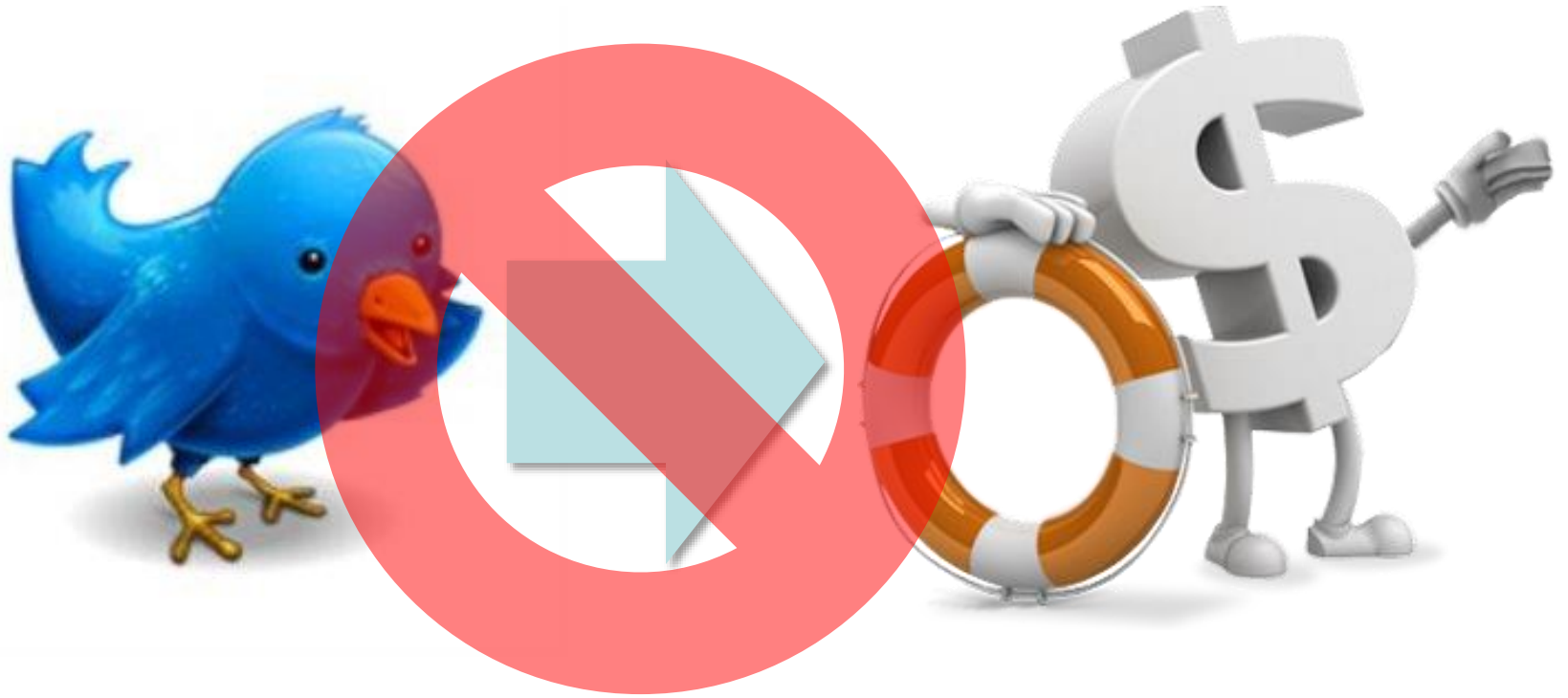


자료: 한국거래소



# Conclusion

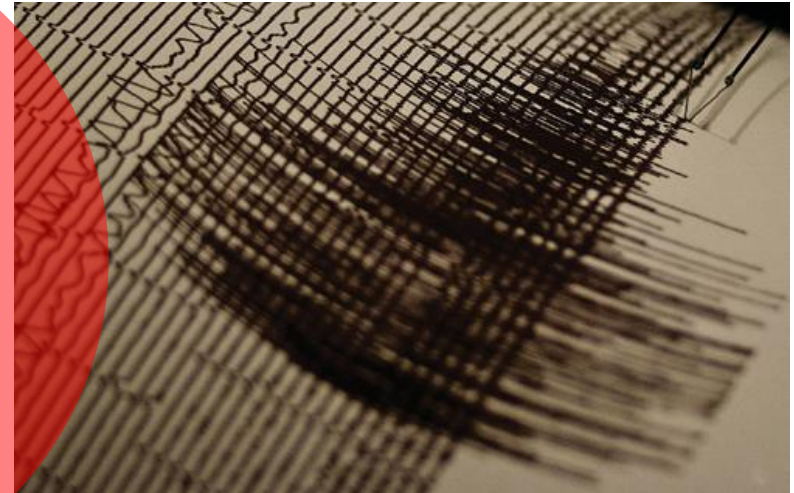
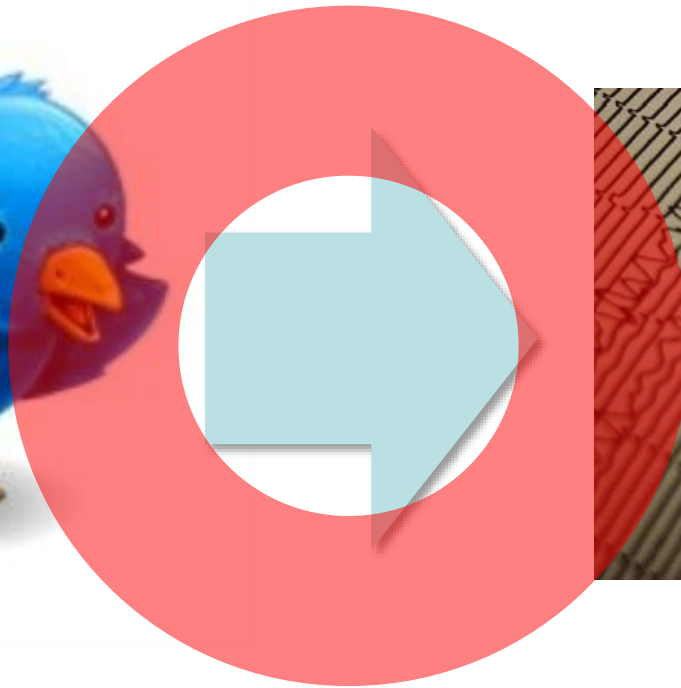
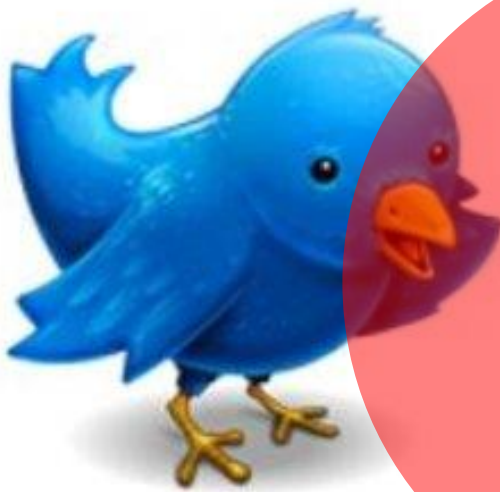
## 1. 트위터와 주가?



트위터 멘션이 직접적으로 주가에 미치는 영향은 거의 없음!  
(멘션 ↑ → 주가 ↑ or ↓? Not Make Sense)

# Conclusion

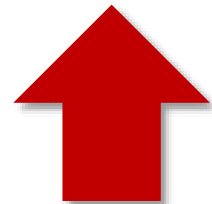
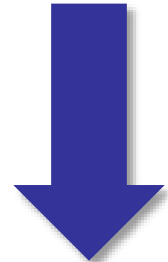
## 2. 트위터와 변동성?



**트위터 멘션이 많다면 주가의 변동성은 커지는 경향이 있다!**

31개 Case 중 26개(약 84%, 멘션 ↑ → 거래 집중 ↑ Make Sense!)

# Conclusion



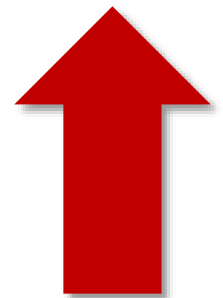
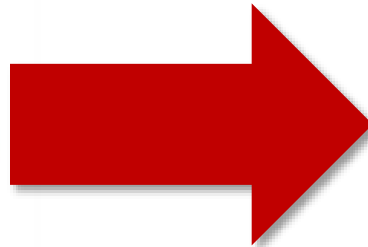
트위터 맨션 활성화가 대선 주자별 테마주에 얼마만큼 영향을 미치는가?  
(안철수, 문재인 ↑, 박근혜 ↓)

# Conclusion

총선 전 7종목 중 2종목 (약 26%)



**총선 전 7종목 일체 (100%)**

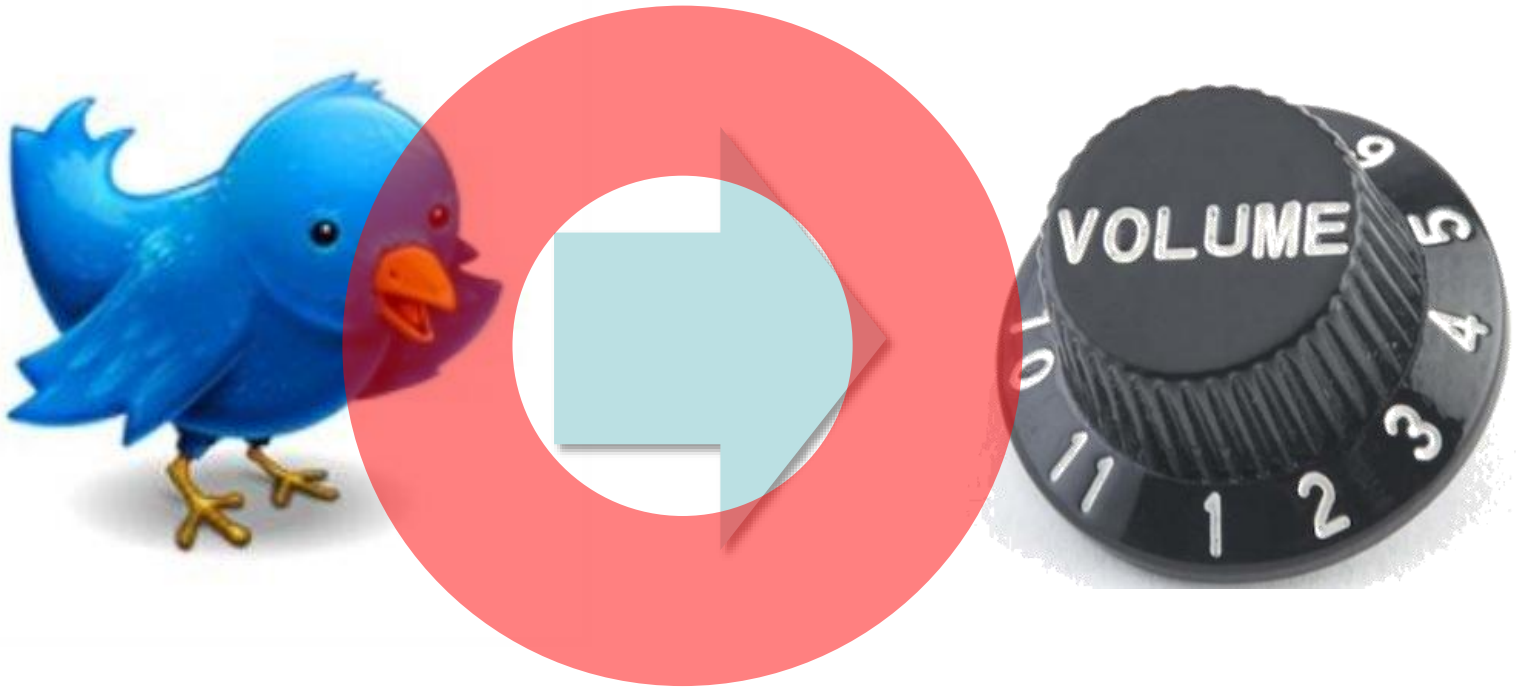


**테마주 주인이 인터넷 공간에서 더 Sensational한 인물이라면!?**  
(온라인 공간에서 안, 문, 박의 Man Power는?)



# Conclusion

## 3. 트위터와 거래량?



**트위터 멘션이 많다면 거래가 활발해지는 경향이 있다!**  
37개 Case 중 31개(약 84%, Make Sense!)



# Worthwhile Claims

- 전통적인 Financial Pricing Model의 Logic을 따르지 않음
  - 주요 변수를 간과한 모형 → 낮은  $R^2$  값 (Acceptable?)
  - 단, 멘션 수는 기업 가격 결정에 중요한 비체계적 요인일 수 있음
- Twitter의 대표성 문제
  - Naver? Google? Facebook? Youtube?
  - Twit이 2~30대의 전유물은 아닌지...
- 이 같은 시계열은 정형화된 언급만을 고려함
  - 오타(박그네), 정당 언급(민주통합당), 기타(Chul-soo Ahn, 사진 등)
  - Twit 멘션이 긍정적이었는지 부정적이었는지 전혀 고려하지 않음
- 테마주 선정의 문제; 무엇을 테마주로 볼 것인가?
  - 테마주가 아닌데 상관성은 높은 경우 (Spurious Regression 등)

# Reference

## 키워드 추출

<http://news.mt.co.kr/mtview.php?no=2012091810313705009>

<http://www.munhwa.com/news/view.html?no=2012091101071824034002>

[http://news.khan.co.kr/khnews/khan\\_art\\_view.html?artid=201209271624211&code=920100](http://news.khan.co.kr/khnews/khan_art_view.html?artid=201209271624211&code=920100)

<http://news.mt.co.kr/mtview.php?no=2012091810313705009>

<http://www.munhwa.com/news/view.html?no=2012091101071824034002>

<http://news.mt.co.kr/mtview.php?no=2012091810313705009>

<http://www.munhwa.com/news/view.html?no=2012091101071824034002>

<http://news.hankooki.com/lpage/economy/201209/h2012091921384284010.htm>

[http://www.fnnews.com/view?ra=Sent0301m\\_View&corp=fnnews&arcid=201209210100183330011100&cDateYear=2012&cDateMonth=09&cDateDay=21](http://www.fnnews.com/view?ra=Sent0301m_View&corp=fnnews&arcid=201209210100183330011100&cDateYear=2012&cDateMonth=09&cDateDay=21)

<http://www.mt.co.kr/view/mtview.php?type=1&no=2012090709304722736&outlink=1>

## 논문

[Bollerslev, Tim, 1986, "Generalized Autoregressive Conditional Heteroskedasticity", \*Journal of Econometrics\* 31, 307-327](#)

## 기사

<http://www.munhwa.com/news/view.html?no=2012091101071824034002>

THANK  
YOU!