

정보검색기술 기말프로젝트

KMA Search 프로젝트

- AI 기반 마켓데이터분석 플랫폼



컴퓨터정보통신대학원
디지털융합금융학과 1조
정낙현, 엄호천, 박진수, 김주원

Contents

1. 프로젝트 배경
2. Problem
3. KMA Search
4. 진행사항
5. 기대효과





프로젝트 배경

01

1-1. 주식투자



1-2. 주식 분석



2020 상반기 베스트 애널리스트

부문	성명	회사
반도체	김경민	하나금융투자
통신	김홍식	하나금융투자
통신 및 네트워크 장비 단말기	김복호	하나금융투자
전기전자·가전	김동원	KB증권
디스플레이	김동원	KB증권
인터넷·소프트웨어	황승택	하나금융투자
엔터테인먼트·레저	이기훈	하나금융투자
미디어·광고	이기훈	하나금융투자
유통	박종대	하나금융투자
운송	강성진	KB증권
증권·보험·기타금융	장효선	삼성증권
은행·신용카드	최정욱	하나금융투자
유틸리티	유재선	하나금융투자
자동차·타이어	김준성	메리츠증권
조선·중공업·기계	김현	메리츠증권
제약·바이오	선민정	하나금융투자
석유화학	윤재성	하나금융투자
음식료·담배	김정욱	메리츠증권
생활소비재(화장품 등)·교육	박종대	하나금융투자
섬유·패션	하누리	메리츠증권
철강·금속	박성봉	하나금융투자
건설·시멘트	채상욱	하나금융투자
지주회사	정대로	미래에셋대우
거시경제·금리	나중혁	하나금융투자
투자전략	이재만	하나금융투자
계량분석	이경수	하나금융투자
데일리 시황	김용구	하나금융투자
파생상품	김용구	하나금융투자
글로벌 투자전략(미국·선진국)	이재만	하나금융투자
글로벌 투자전략(중국·신흥국)	김경환	하나금융투자
원자재	황병진	NH투자증권
채권	윤여삼	메리츠증권
신용분석	김상훈	신한금융투자
글로벌 자산배분	김종원	현대차증권
글로벌 ETF	이창환	현대차증권
스몰캡	코스닥 벤처팀	하나금융투자





Problem

02

1-1. 문제점

1) 기존 문제점

내용 추가





KMA Search

03

3-1. KMA Search ?

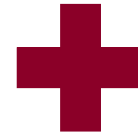
KMA Search
Korea university
MiaeassetDaewoo
AI analyst
Research



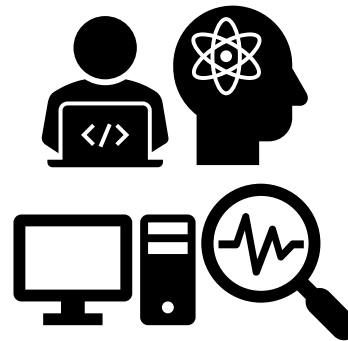
3-2. KMA Search ?



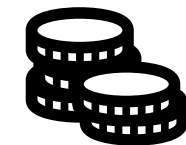
SNS



주가, 재무제표, 뉴스



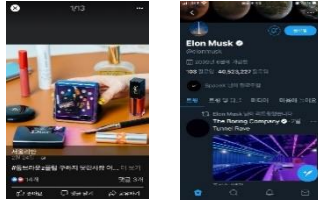
KMA Search



마케팅
투자 추천
리서치

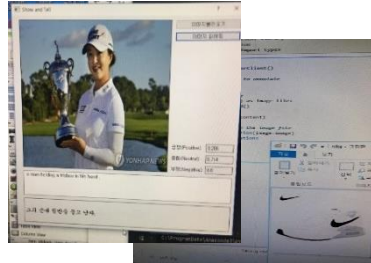


3-3. KMA Search 구성도



SNS, 사진

이미지분석

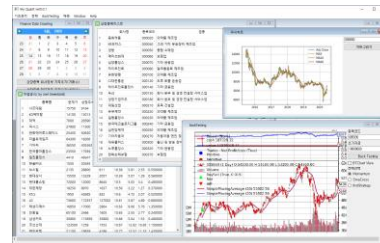


삭제예정

뉴스,기사,검색어



텍스트 분석

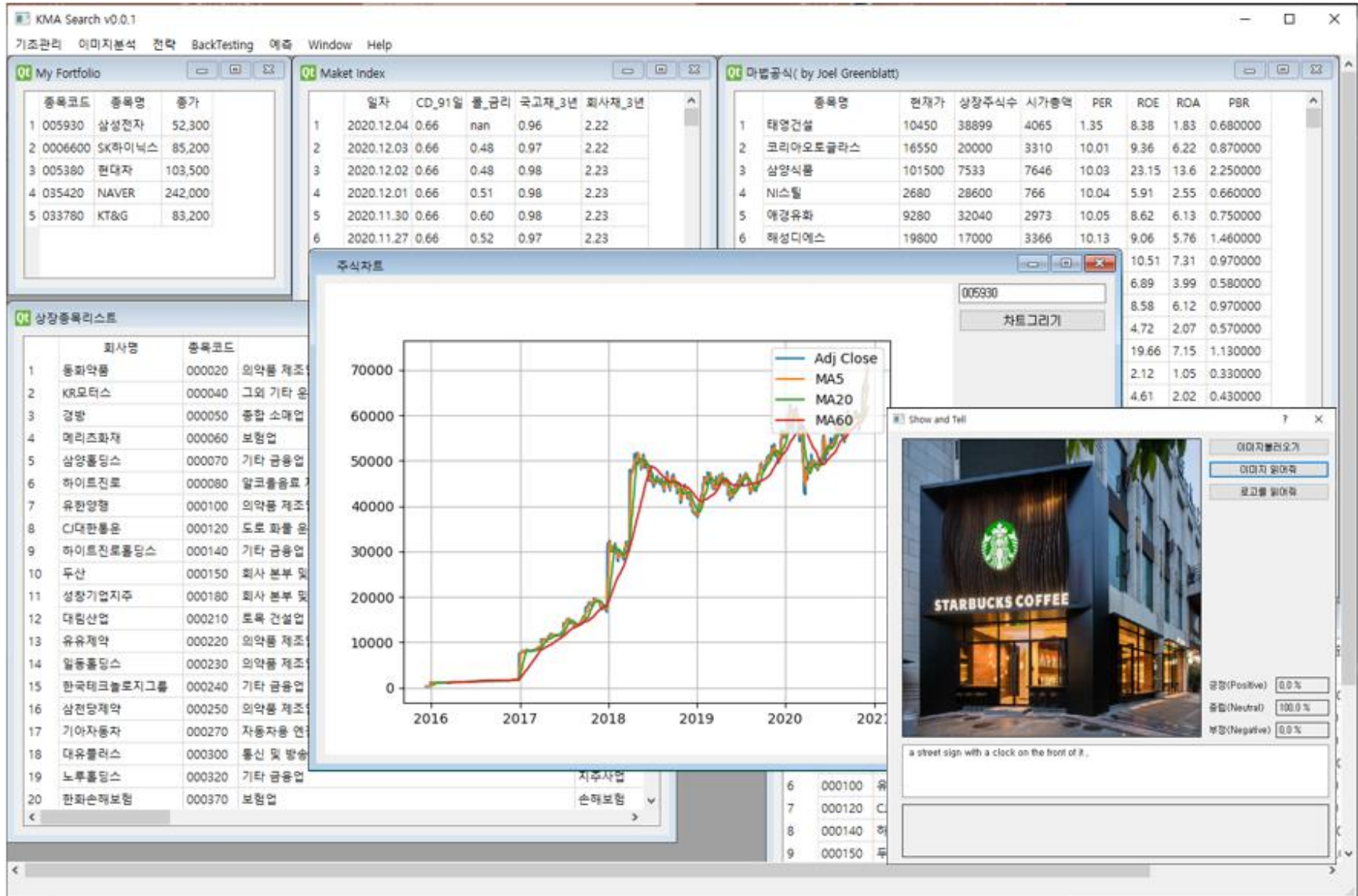


종목추천

종목명	30일 증가	T스코어
한국토지신탁 034830	1,955원 +0.26%	81
동부건설 005960	13,700원 +2.24%	81
국순당 043650	5,600원 +4.67%	80



3-4. KMA Search



3-5. SNS, 뉴스 이미지 → TEXT 가져오기

김세영, LPGA 펠리컨 챔피언십 우승...상금왕 선두

동료 2020-11-23 09:32 수정 2020-11-23 09:50

f t + -

스포츠 많이 보는 기사

1. "핵 주먹" 마이크 타이슨, 기회가 없었다
2. SK서 워터닉 집합, 한화 아글스로 간다
3. 양현종-이대호-최형우 위 경인 등-KBO FA전후 팀 단 확정
4. 박세리의 매질내 '명예의 전당' 입성
5. 타이슨 대 존스 '세계 헤비급 레전드' 맞대결

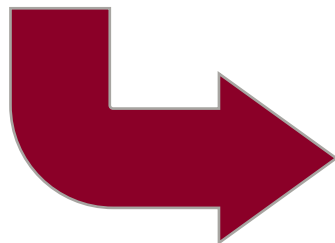
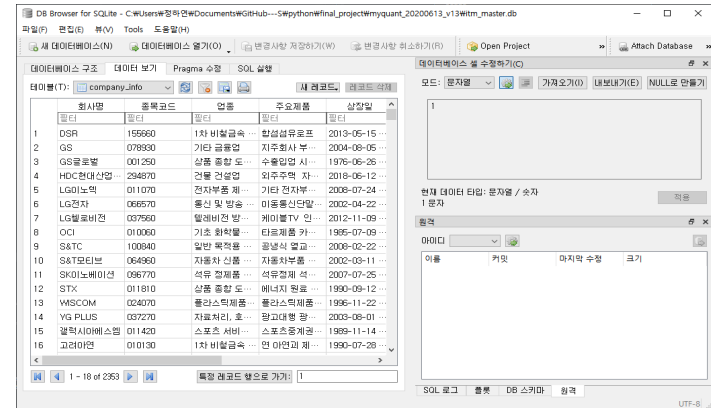
한겨레와 친구하기

1 like 2020-11-23 09:32

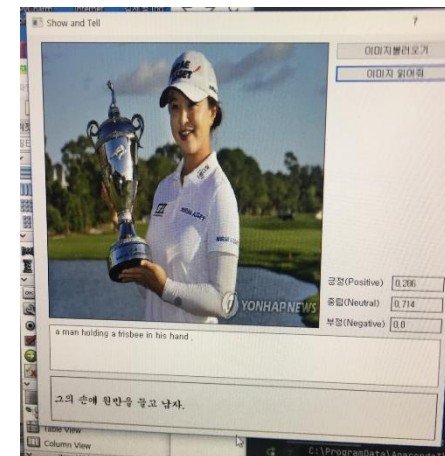


김세영이 23일(한국시간) 미국여자프로골프(LPGA) 투어 펠리컨 챔피언십에서 우승을 차지한 뒤 우승 트로피를 들어 올리고 있다. 발매/AP 연합뉴스

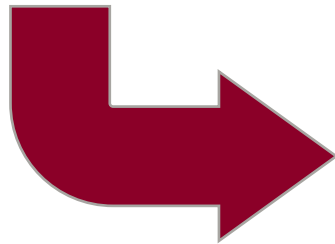
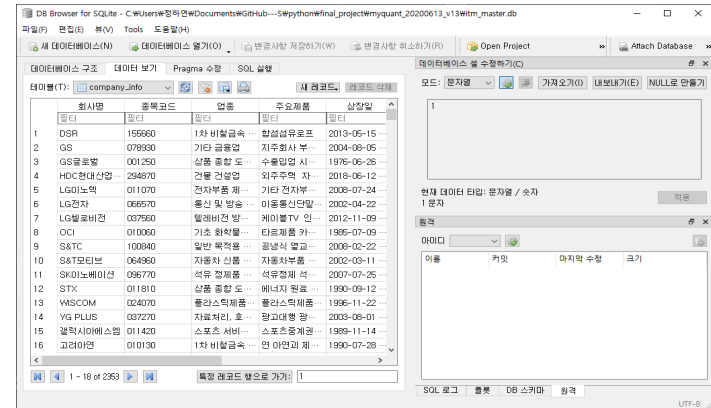
김세영(27)이 미국여자프로골프(LPGA) 투어 펠리컨 챔피언십(총상금 150만 달러) 우승을 차지



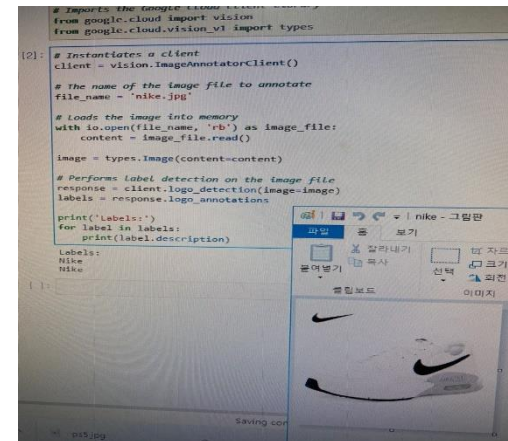
이미지 캡셔닝
긍정, 부정



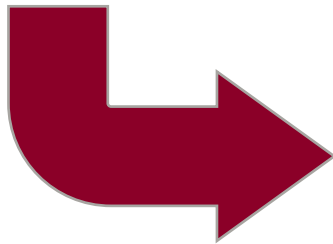
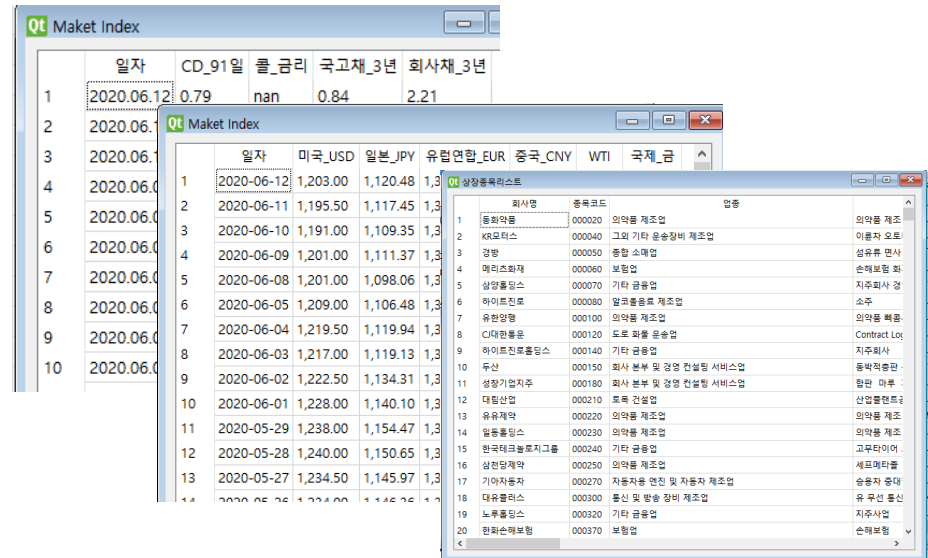
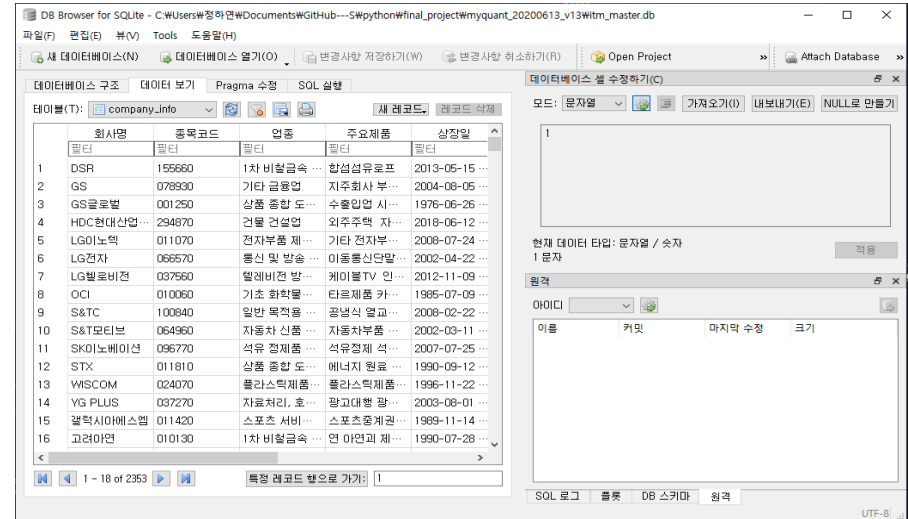
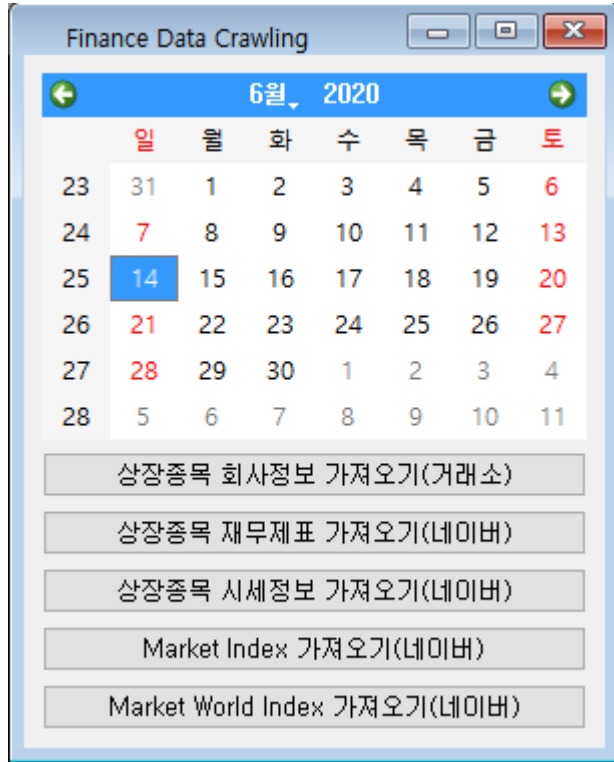
3-6. 이미지 → 로고 데이터 가져오기



기업로고 분석
긍정, 부정



3-6. 마켓데이터 가져오기



3-5. AI 리서치 보고서 생성

DB Browser for SQLite - C:\Users\정하연\Documents\GitHub\---\Swpthon\final_project\myquant_20200613_v13\witm_master.db

파일(F) 편집(E) 뷰(V) Tools 도움말(H)

새 데이터베이스(N) 데이터베이스 열기(O) 변경 사항 저장하기(W) 변경 사항 취소하기(R) Open Project Attach Database

데이터베이스 구조 데이터 보기 Pragma 설정 SQL 실행

데이터베이스: company_info

새 레코드... 레코드 삭제

필터	회사명	종목코드	업종	주요제품	상장일
1	DSR	155660	1차 비철금속	합성섬유로프	2013-05-15
2	GS	078930	기타 금융업	지주회사 부	2004-08-05
3	GS글로벌	001250	상품 종합 도	수출입업 시	1976-06-26
4	HDC현대산업	294870	건물 건설업	외주주력 자	2018-06-12
5	LG이노텍	011070	전자부품 제	기타 전자부	2008-07-24
6	LG전자	066570	통신 및 방송	이동통신단말	2002-04-22
7	LG화학	037560	알레비전 발	케이블TV 인	2012-11-09
8	OCI	010060	기초 화학물	타르제품 카	1995-07-09
9	S&T	100840	일반 목적용	공냉식 열교환	2008-02-22
10	S&T모티브	064960	자동차 부품	자동차 부품	2002-03-11
11	SK이노베이션	096770	석유 정제품	석유정제 석	2007-07-25
12	STX	011810	상품 종합 도	에너지 원료	1990-09-12
13	WISCOM	024070	플라스틱제품	플라스틱제품	1996-11-22
14	YG PLUS	037270	자료처리, 호	광고대행 할	2003-08-01
15	갤럭시이메스	011420	소프트 서비	소프트웨어권	1989-11-14
16	고려이연	010130	1차 비철금속	연 아연과 제	1990-07-28

1 - 18 of 2353

특정 레코드 번호로 가기: 1

데이터베이스 설정하기(C)

모드: 문자열 가져오기(O) 내보내기(E) NULL로 만들기

현재 데이터 타입: 문자열 / 숫자

원격

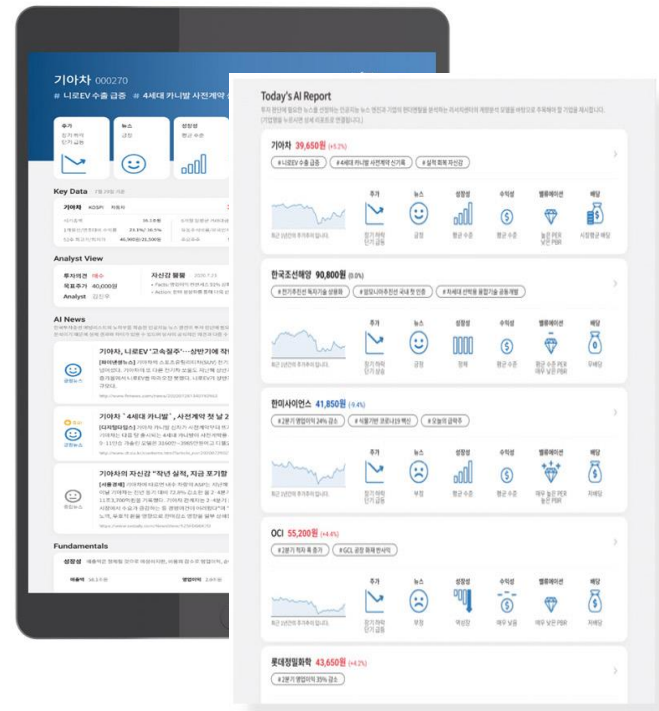
아이디

이름 커밋 마지막 수정 크기

SQL 로그 플롯 DB 스키마 원격

UTF-8

BERT, GPT3



3-5. AI 추천종목

DB Browser for SQLite - C:\Users\정하연\Documents\GitHub\---\Swpthon\final_project\myquant_20200613_v13\itm_master.db

파일(F) 편집(E) 뷰(V) Tools 도움말(H)

새 데이터베이스(N) 데이터베이스 열기(O) 변경 사항 저장하기(W) 변경 사항 취소하기(R) Open Project Attach Database

데이터베이스 구조 데이터 보기 Pragma 수정 SQL 실행

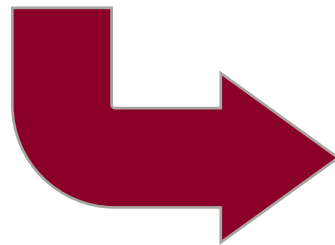
데이터베이스(T): company_info

회사명	종목코드	업종	주요제품	상장일
1 DSR	155660	1차 비철금속	합성섬유로프	2013-05-15
2 GS	078930	기타 금공업	지주회사 부	2004-08-05
3 GS글로벌	001250	상표 종합 도	수출입업 시	1976-06-26
4 HDC현대산업	294870	건물 건설업	외주제작 자	2018-06-12
5 LG이노텍	011070	전자부품 제	기타 전자부	2008-07-24
6 LG전자	066570	통신 및 방송	이동통신단말	2002-04-22
7 LG필로비전	037560	텔레비전 발	케이블TV 인	2012-11-09
8 OCI	010060	기초 화학물	타르질소 카	1985-07-09
9 S&T	100840	일반 목적용	공냉식 열교환	2008-02-22
10 S&T모티브	064960	자동차 부품	자동차부품	2002-03-11
11 SK이노베이션	096770	석유 정제품	석유정제 석	2007-07-25
12 STX	011810	상표 종합 도	에너지 원료	1990-09-12
13 WISCOM	024070	플라스틱제품	플라스틱제품	1996-11-22
14 YG PLUS	037270	자료처리, 호	광고대행 광	2003-08-01
15 엘렉시아베스	011420	스포츠 서비	스포츠중계권	1989-11-14
16 고려아연	010130	1차 비철금속	연 연연과 제	1990-07-28

1 - 18 of 2353

특정 레코드 번호로 가기: 1

SQL 로그 실행 DB 스키마 보기



퀀트 모형

KMA Search v0.1

기초관리 이미지분석 전략 BackTesting 예측 Window Help

마블공식 (by Joel Greenblatt)

종목명	현재가	상장주식수	시가총액	PER	ROE	ROA	PBR
1 태영건설	10450	38899	4065	1.35	8.38	1.83	0.680000
2 코리아오토글라스	16550	20000	3310	10.01	9.36	6.22	0.870000
3 삼양식품	101500	7533	7646	10.03	23.15	13.6	2.250000
4 NIS	2680	28600	766	10.04	5.91	2.55	0.660000
5 애경유화	9280	32040	2673	10.05	8.62	6.13	0.750000
6 해성디에스	1980						
7 한국콜마물류	2505						
8 국도화학	5610						
9 한농화학	8060						
10 KTcs	2270						
11 화승전선	7240						
12 기온전선	2215						
13 KT	2440						
14 대성에너지	5480						
15 현대글로비스	1860						
16 LG	7410						
17 현대글로비스	7900						
18 휴비스	8220						
19 LF	1505						
20 두산인프라코어	8280						
21 꽃우유	2010						

가치투자 (by Benjamin Graham)

종목명	현재가	상장주식수	시가총액	PER	ROE	ROA	PBR
1 한화생명	2255	868530	19585	8.57	0.51	0.04	0.120000
2 경동엔비스트	29250	2365	692	13.27	2.26	1.47	0.160000
3 동양생명	3610	161359	5825	7.31	4.73	0.35	0.180000
4 세화증권	8090	10000	809	5.42	1.85	0.75	0.190000
5 태화증권	786000	1113	8751	8.44	4.34	3.78	0.200000
			1746	12.96	2.72	1.97	0.210000
			1374	7.19	4.37	3.78	0.210000
			19198	4.0	7.07	0.59	0.220000
			1100	4.83	5.76	0.46	0.220000
			1202	7.52	2.8	2.62	0.220000
			2101	5.82	6.37	3.31	0.230000
			1166	9.47	5.05	3.86	0.240000
			2514	13.85	5.96	2.87	0.250000
			12584	3.8	7.31	0.53	0.250000
			2323	4.95	8.78	5.72	0.250000
			2240	5.0	5.58	4.05	0.250000
			6992	6.44	4.98	0.3	0.250000
			3383	12.27	1.38	1.5	0.260000
			724	12.46	6.47	5.36	0.260000
			2229	13.18	3.39	2.66	0.260000
			476	14.0	3.32	3.12	0.260000

이기는투자 (by Peter Lynch)

종목명	현재가	상장주식수	시가총액	PER	ROE	ROA	PBR
1 세화	1930	57563	1111	4.5	34.66	6.12	1.300000
2 계룡건설	26150	8931	2335	2.6	16.94	4.0	0.450000
3 한신공영	15800	11571	1828	2.43	15.21	4.2	0.330000
4 태영건설	10450	38899	4065	1.35	8.38	1.83	0.680000
5 대일건설	30800	22053	6792	3.89	23.04	11.54	0.910000
6 메리츠금융지주	10000	134011	13401	2.88	15.26	1.38	0.400000
7 에이리츠	8130	3793	308	4.97	25.42	8.25	0.840000
8 동양고속	26350	2896	763	5.6	26.85	17.11	0.610000
9 사조대일	15800	9164	1448	6.54	30.57	9.72	0.330000
10 대성건설	9800	36213	3549	2.98	13.07	5.0	0.550000
11 법랑건설	5080	24832	1261	7.97	32.82	15.36	1.690000
12 다우기술	23350	44867	10476	3.75	14.62	1.84	0.660000
13 SK가스	109000	9230	10061	2.55	9.85	3.93	0.520000
14 시디즈	51700	2000	1034	7.32	27.34	11.24	1.970000
15 HDC현대산업개발	22100	65907	14566	5.48	20.3	8.92	0.550000
16 동남건설	42100	3610	1520	7.18	26.23	19.21	2.130000
17 동부건설	13400	22702	3042	4.37	15.3	7.45	0.690000
18 삼표식품	44550	4568	2035	5.36	18.57	12.43	1.070000
19 삼양식품	62300	3000	1869	4.12	14.22	12.64	0.500000
20 한국자산신용	4045	123978	5015	4.22	14.28	7.98	0.670000





진행사항

04

4-1. 프로젝트 인원



정낙현 (조장)
미래에셋대우
금융상품시스템팀



엄호천
미래에셋대우
전략팀
프로젝트 기획



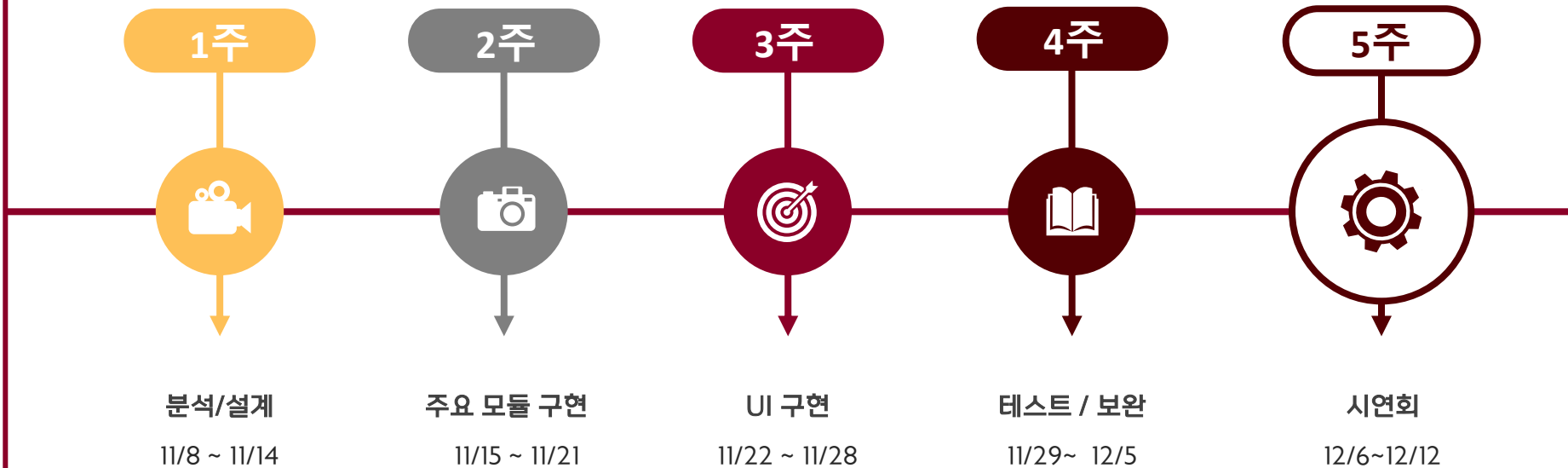
박진수
미래에셋대우 자금팀
프로그램 개발



김주원
미래에셋대우 디지털
사례조사 ,업무분석



4-3.프로젝트 진행 일정



4-2. 프로젝트 개발 사항

1. 프로젝트 계획
2. 테스트 코딩모형 및 투자 전략 모듈
3. 백테스팅 모듈



4-5. 진행사항

1) 뉴스 분석

- 뉴스 크롤링 , 뉴스요약

2)재무 분석

- 재무 데이터 크롤링 , 마켓데이터 수신

3) SNS 분석

- 실시간검색어 , 이미지 분석

4)마케팅 자료

- 추천종목

5) 리서치 자료

- 시황,매수,매도 의견





기대효과

05

5-1. 기대효과

- 1) 뉴스 분석
- 2) 재무 분석
- 3) SNS 분석
- 4) 마케팅

내용재정리





감사합니다