

Goodie Movie

스프링 프레임워크 기반의
빅데이터 분석 및 시각화 처리 프로젝트

구디 아카데미
빅데이터 2기
이종화

목차

- 1) 소개 & 제작 목적
- 2) 개발 환경 & 사용 기술
- 3) 프로젝트 내용
- 4) 개선점
- 5) 시연
- 6) Q&A

1) 소개 & 제작 목적



GOODEE MOVIE

GOODIE MOVIE란?

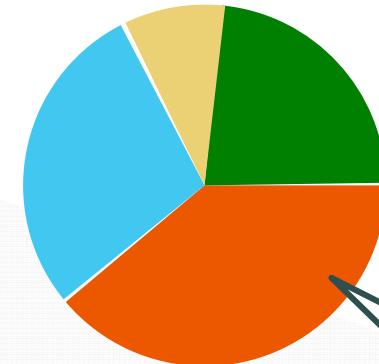
◆ 소개

- 1) 1874년부터 올해까지
45,000개에 달하는 영화
데이터의 시각화된 분석
통계를 제공
- 2) 가입을 통해 영화 리뷰를
쓰고 평점을 매기거나,
게시글을 통해 토론하는
커뮤니티 기능 제공

타겟 데이터
파일(CSV)

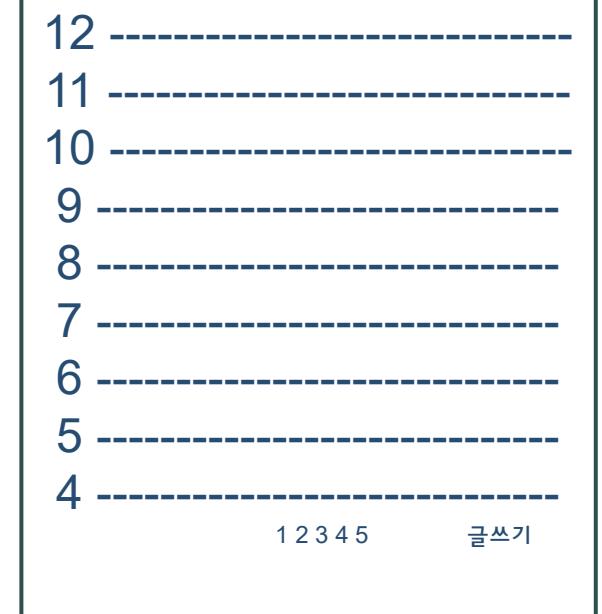


+



분석을 통한
차트화

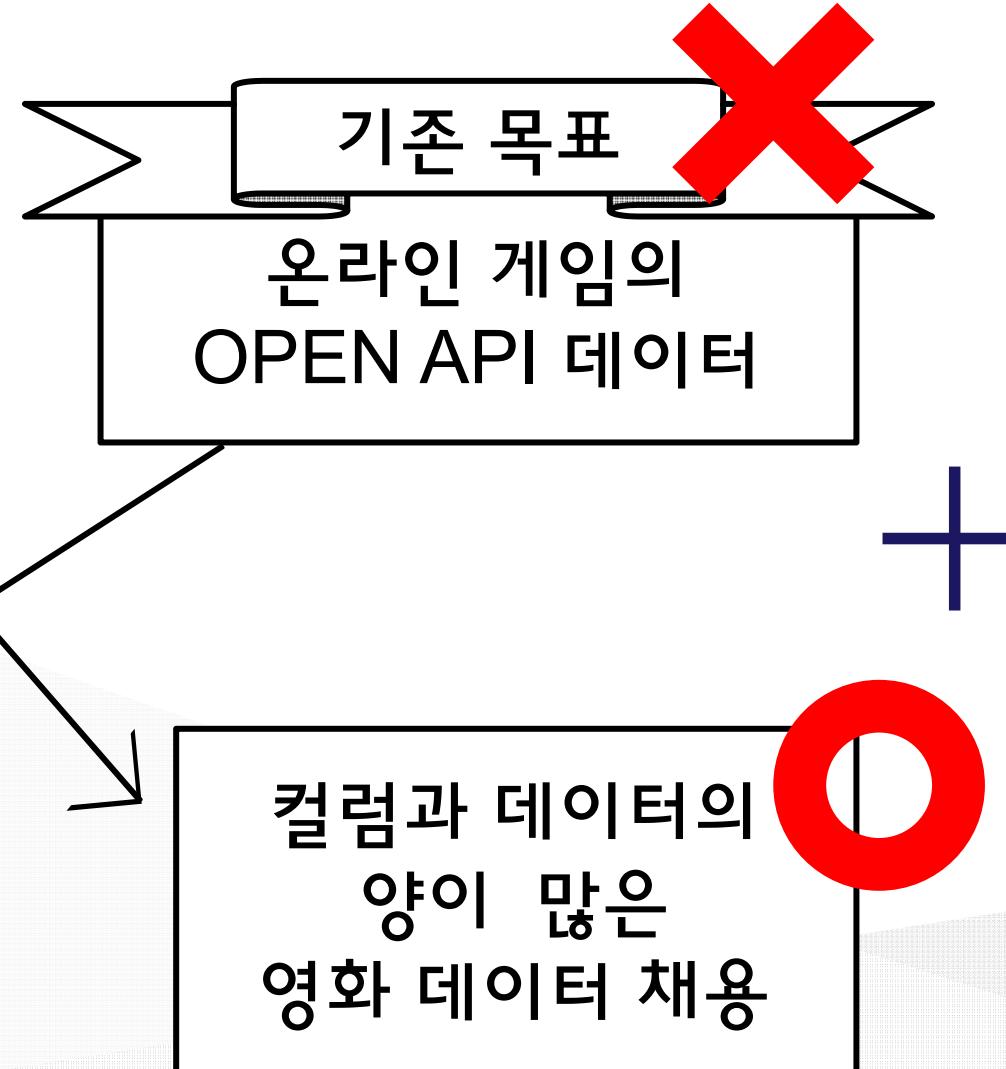
리뷰 게시판



제작목적

프로젝트 설계 조건 (요구사항)

- 스프링 프레임워크 기반으로 웹사이트를 구현
- 하둡 소프트웨어를 통해 1000개 이상의 빅데이터를 분석&시각화



부가적으로
게시판 기능
구현

참조 데이터 (CSV)

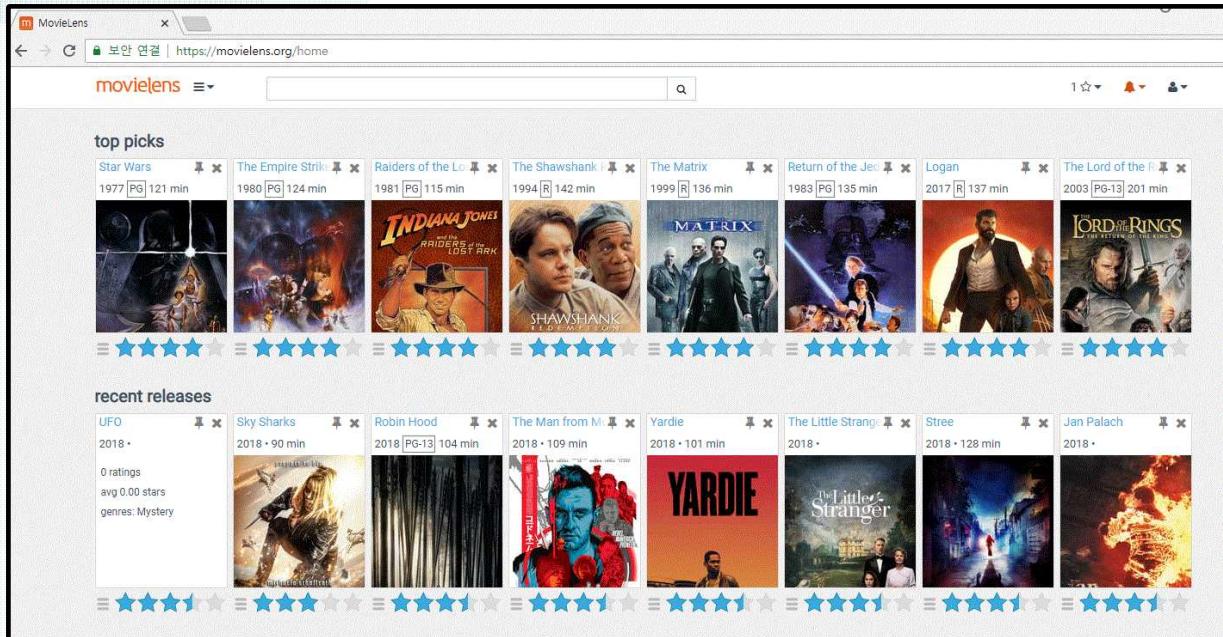
<데이터 출처>

<https://www.kaggle.com/rounakbanik/the-movies-dataset>

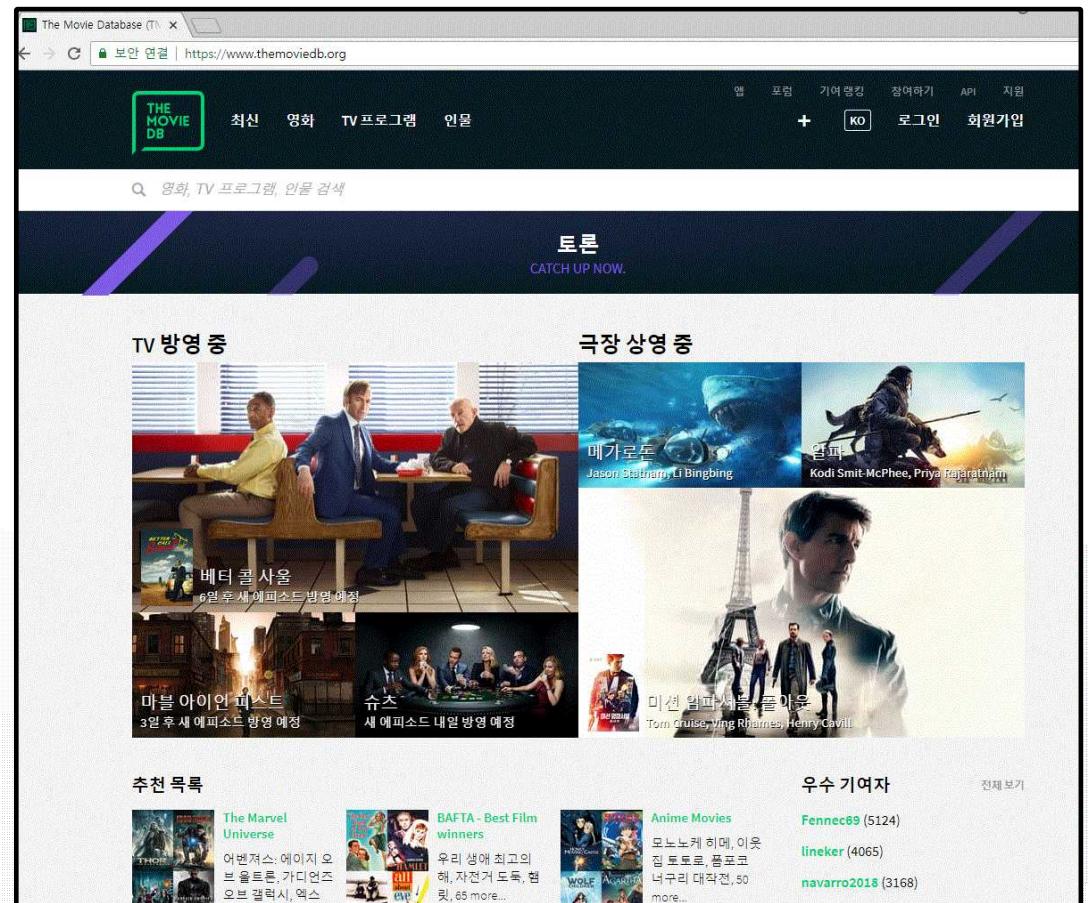
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	adult	belongs_to_budget	genres	homepage	id	imdb_id	original_lang	original_title		overview	popularity	poster_path	production_
2	FALSE	{"id": 1019, "300000000	[{"id": 16, "http://toysto	862	tt0114709	en		Toy Story		Led by Wo	21,94694	/rhLRbceoE	[{"name": "P
3	FALSE		650000000	[{"id": 12, "name": "Adv	8844	tt0113497	en	Jumanji		When sibling	17,01554	/vzmL6fP7a	[{"name": "I
4	FALSE	{"id": 11905, "0	[{"id": 10749, "name": "	15602	tt0113228	en	Grumpier Old Men		A family wed	11,7129	/6ksm1sjKM	[{"name": "W	
5	FALSE		160000000	[{"id": 35, "name": "Cor	31357	tt0114885	en	Waiting to Exhale		Cheated on	3,859495	/16XOMpE	[{"name": "C
6	FALSE	{"id": 9687, "0	[{"id": 35, "name": "Cor	11862	tt0113041	en	Father of the Bride Part	I	Just when G	8,387519	/e64sOl48t	[{"name": "S	
7	FALSE		600000000	[{"id": 28, "name": "Acti	949	tt0113277	en	Heat		Obsessive r	17,92493	/zMyfPUelu	[{"name": "R
8	FALSE		580000000	[{"id": 35, "name": "Cor	11860	tt0114319	en	Sabrina		An ugly duc	6,677277	/JQh15y5Yl	[{"name": "F
9	FALSE			[{"id": 28, "name": "Acti	45325	tt0112302	en	Tom and Huck		A mischievo	2,561161	/sGO5Qa5!	[{"name": "V
10	FALSE		350000000	[{"id": 28, "name": "Acti	9091	tt0114576	en	Sudden Death		International	5,23158	/eoWvKD6C	[{"name": "U
11	FALSE	{"id": 645, "580000000	[{"id": 12, "http://www.	710	tt0113189	en	GoldenEye		James Bond	14,68604	/5c0ovjT41	[{"name": "U	
12	FALSE		620000000	[{"id": 35, "name": "Cor	9087	tt0112346	en	The American President		Widowed U	6,318445	/lymPNGLZ	[{"name": "O
13	FALSE			[{"id": 35, "name": "Cor	12110	tt0112896	en	Dracula: Dead and Lovin	When a law	5,430331	/xve4cgfYlt	[{"name": "C	
14	FALSE	{"id": 11769, "0	[{"id": 10751, "name": "	21032	tt0112453	en	Balto		An outcast h	12,14073	/gV5PCAVC	[{"name": "U	
15	FALSE		440000000	[{"id": 36, "name": "Hist	10858	tt0113987	en	Nixon		An all-star c	5,092	/clCkmCEIX	[{"name": "B
16	FALSE		980000000	[{"id": 28, "name": "Acti	1408	tt0112760	en	Cutthroat Island		Morgan Ada	7,284477	/odM9973k	[{"name": "U
17	FALSE		520000000	[{"id": 18, "name": "Dra	524	tt0112641	en	Casino		The life of th	10,13739	/xo517bXE	[{"name": "U
18	FALSE		165000000	[{"id": 18, "name": "Dra	4584	tt0114388	en	Sense and Sensibility		Rich Mr. Da	10,67317	/IA9HTy84B	[{"name": "O
19	FALSE		4000000	[{"id": 80, "name": "Crin	5	tt0113101	en	Four Rooms		It's Ted the	9,026586	/eQs5hh9rx	[{"name": "U
20	FALSE	{"id": 3167, "300000000	[{"id": 80, "name": "Crin	9273	tt0112281	en	Ace Ventura: When Natu	Summoned	8,205448	/wRIGnJhEz	[{"name": "O		
21	FALSE		600000000	[{"id": 28, "name": "Acti	11517	tt0113845	en	Money Train		A vengeful M	7,337906	/jSozzzVOR	[{"name": "O
22	FALSE	{"id": 91698, "302500000	[{"id": 35, "name": "Cor	8012	tt0113161	en	Get Shorty		Chili Palmer	12,66961	//WtDUUgC	[{"name": "U	
23	FALSE			[{"id": 18, "name": "Dra	1710	tt0112722	en	Copycat		An agorapho	10,7018	/80czeJGS	[{"name": "B
24	FALSE		500000000	[{"id": 28, "name": "Acti	9691	tt0112401	en	Assassins		Assassin Rc	11,06594	/xAx5MP7D	[{"name": "S
25	FALSE			[{"id": 18, "name": "Dra	12665	tt0114168	en	Powder		Harassed by	12,13309	/1uRKsxOC	[{"name": "O
26	FALSE		3600000	[{"id": 18, "http://www.	451	tt0113627	en	Leaving Las Vegas		Ben Sander	10,33203	/37qHRJxn	[{"name": "U

UI구현과 설계 참조 사이트

◆ <https://movielens.org/home>



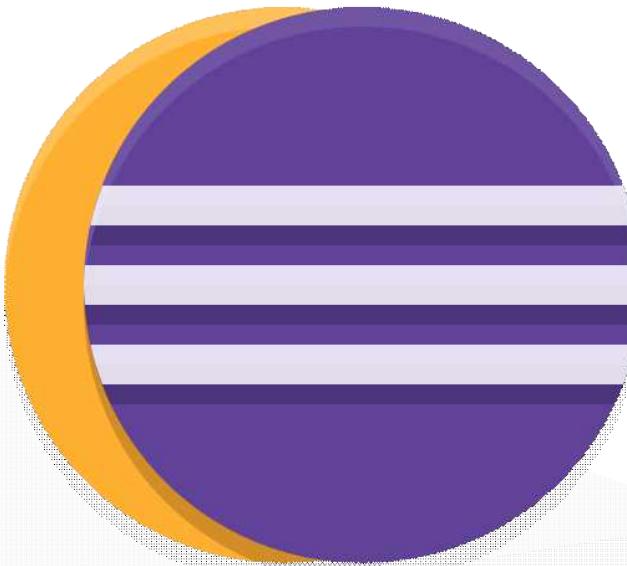
◆ <https://www.themoviedb.org/>



개발 스케줄

일	월	화	수	목	금	토
8월	5	6	7	8	9 설계 기획	10 개발 시작
	12	13 UI구현	14	15	16 X	17 X
	19	20	21	22 기능 설계	23	24
	26	27 X	28	29	30	31 TEST
	2	3 TEST	4 포폴준비	5 발표	6	7
9월	8					

2) 개발 환경 & 사용 기술



개발 환경 & 사용기술

언어

Java & JavaScript

데이터베이스

MariaDB

프로그램

Spring & Hadoop & Tomcat

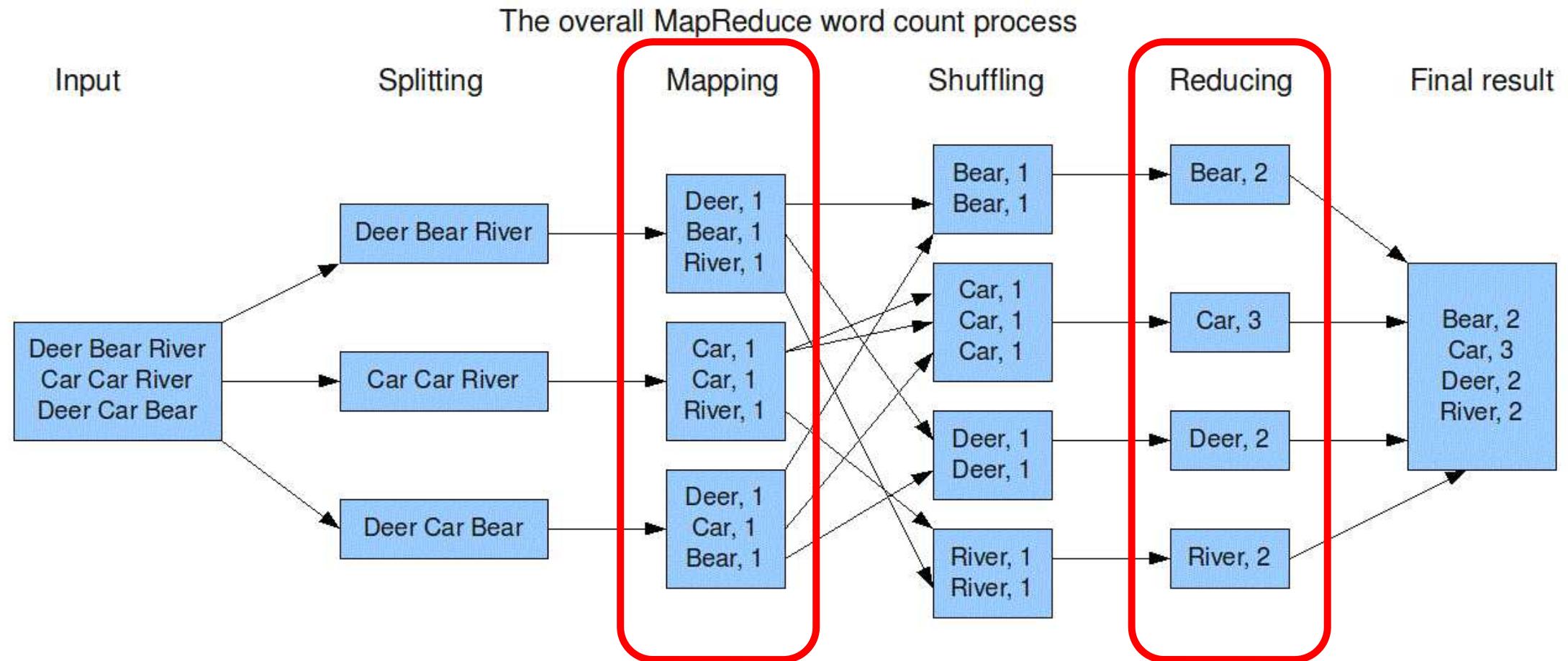
UI

HTML(+JSP) & CSS
jQuery(JavaScript)
Google Chart

하둡이란 무엇인가?

- ❖ **위키에서는...**
- ◆ **아파치 하둡(Apache Hadoop)**은 대량의 자료를 처리할 수 있는 큰 컴퓨터 클러스터에서 동작하는 분산 응용 프로그램을 지원하는 프리웨어 자바 소프트웨어 프레임워크
- ◆ 분산처리 시스템인 구글 파일 시스템을 대체할 수 있는 **하둡 분산 파일 시스템(HDFS: Hadoop Distributed File System)**과 **맵리듀스**를 구현한 것

맵리듀스란?



간략도

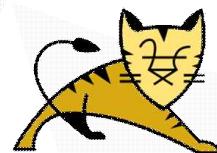


GitHub

저장소



spring
by Pivotal™



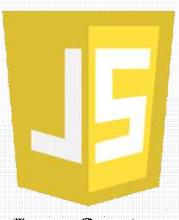
Apache
Tomcat



Java



AJAX



JavaScript



Google chart



hadoop

분석



프로젝트 상세 내용 확인

<저장소 GitHub>

<https://github.com/jhcorea9n9q/Project>



GitHub

3) 프로젝트 내용



프로젝트 주요내용 가이드

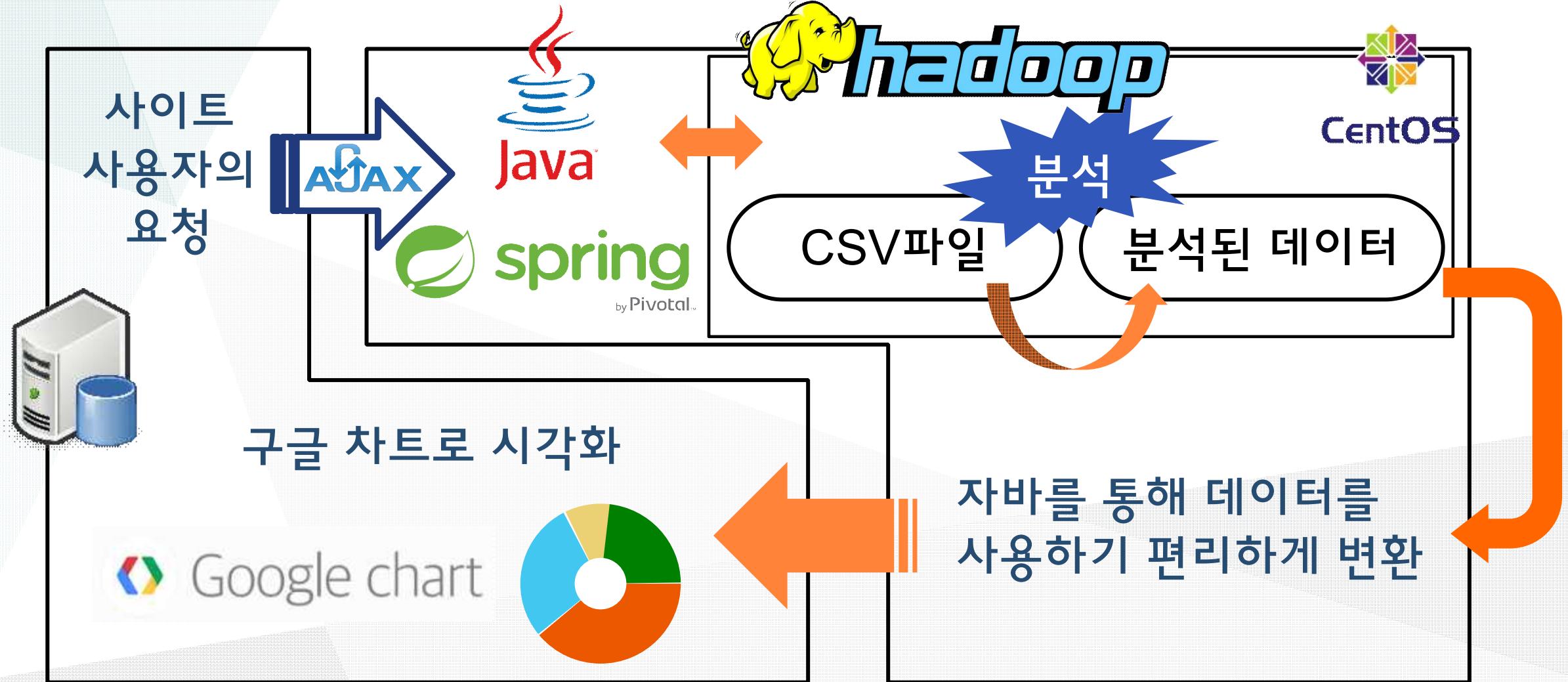
(1) 빅데이터 분석

(2) 유저 관리 기능

(3) 피드백과 게시판 구현

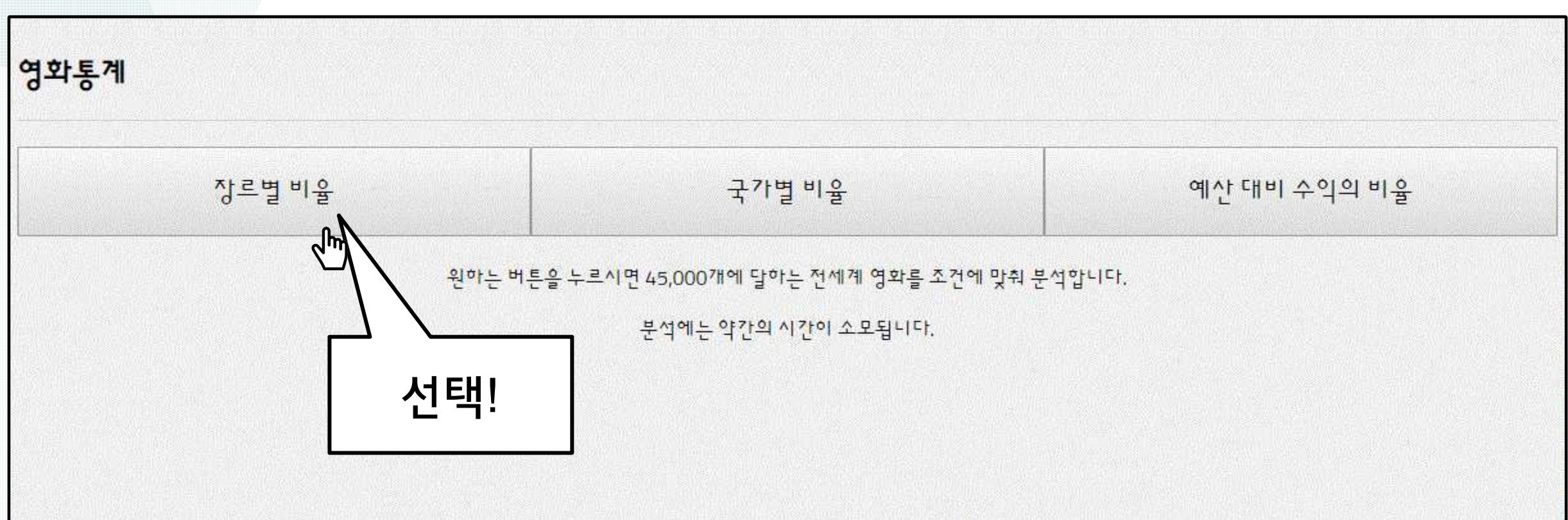
(1) 빅데이터 분석

◆ 빅데이터 분석 플랫폼인 HADOOP을 통한 차트 시각화가 목표



(1) 빅데이터 분석

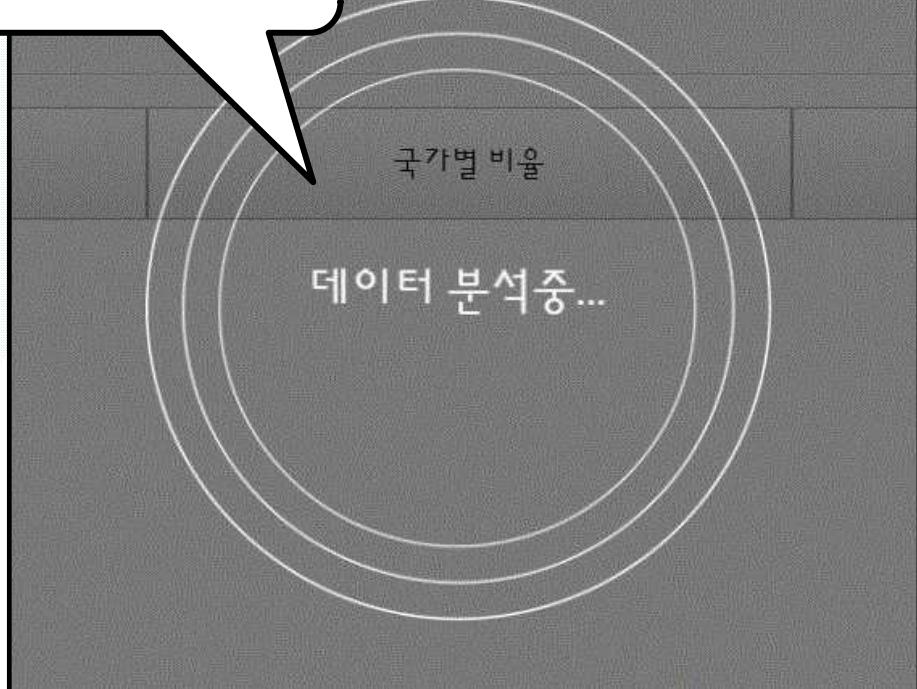
① GOODIE MOVIE 영화통계 페이지에서 원하는 내용을 선택



(1) 빅데이터 분석

② 약간의 대기 시간과 함께 HADOOP이 데이터를 분석,
분석된 데이터는 HADOOP 내 result 폴더에 실시간으로 저장됨

웹 화면

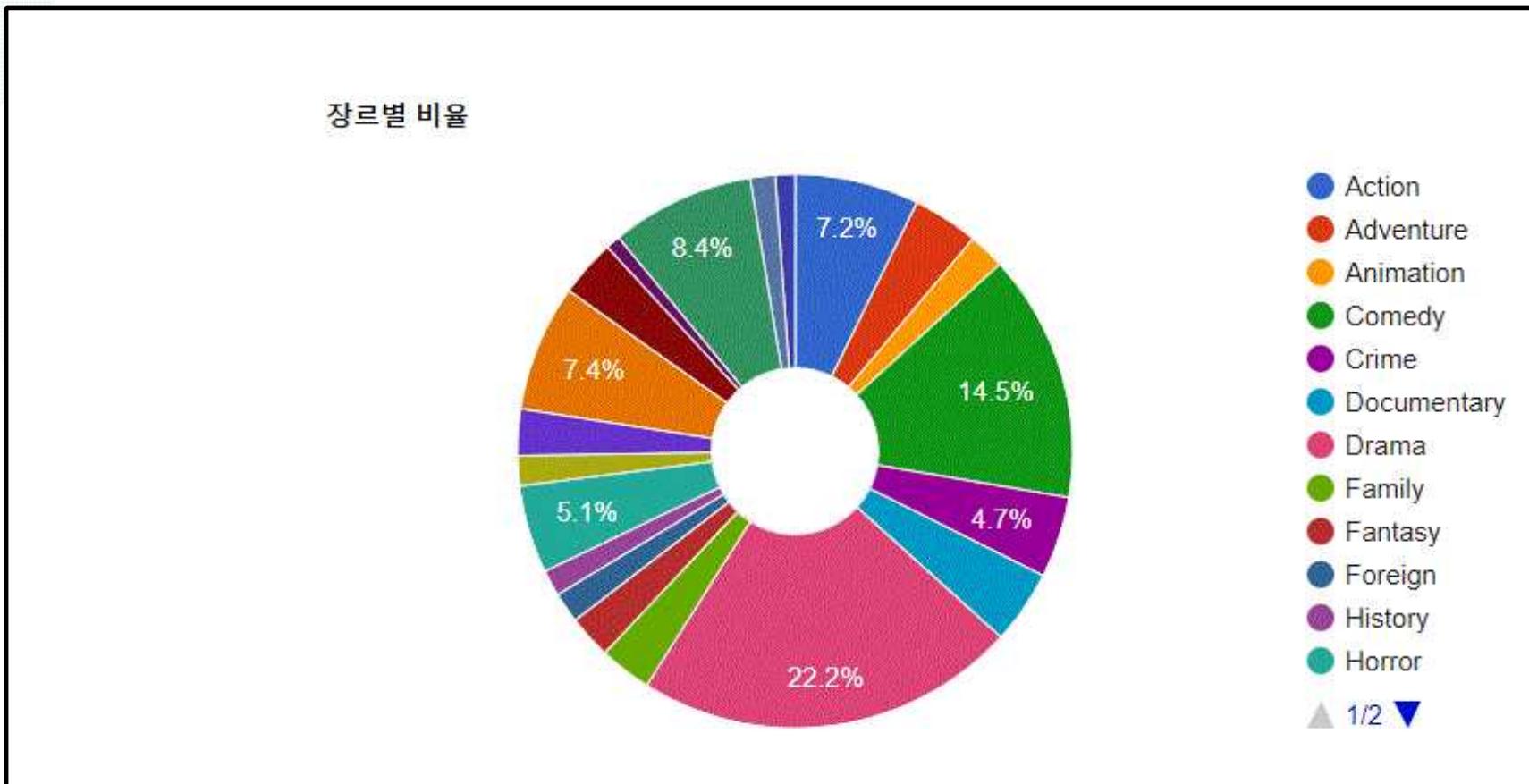


서버측

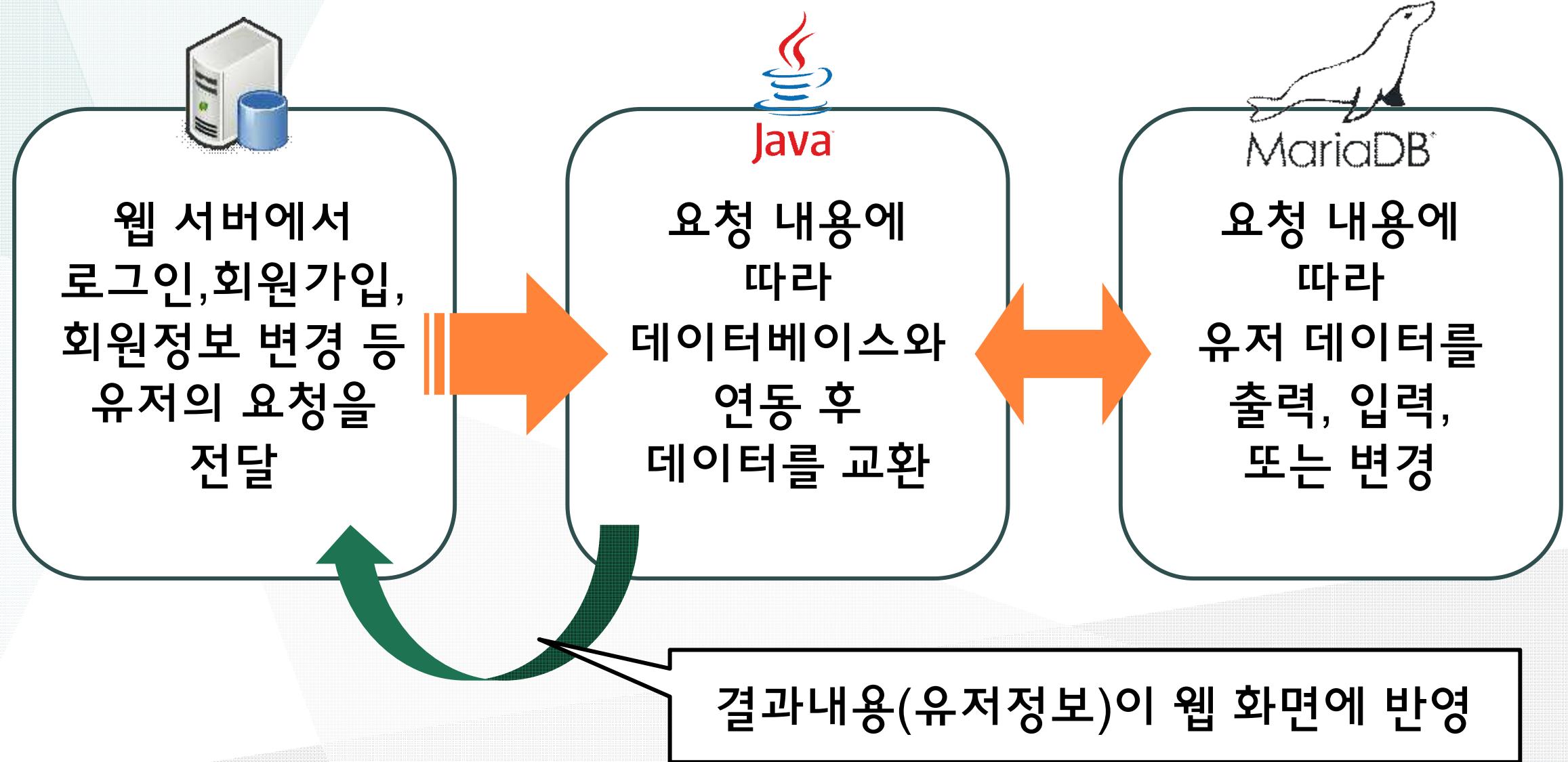
```
drwxr-xr-x - goodee supergroup          0 2018-09-04 09:00 /C/20180904_094624
-rw-r--r-- 3 goodee supergroup          0 2018-09-04 09:00 /C/20180904_094624/_SUCCESS
-rw-r--r-- 3 goodee supergroup          2031 2018-09-04 09:00 /C/20180904_094624/part-r-00000
drwxr-xr-x - goodee supergroup          0 2018-09-04 14:22 /G/20180903_052716
drwxr-xr-x - goodee supergroup          0 2018-09-03 17:27 /result
/G/20180903_052716/_SUCCESS            0 2018-09-03 17:27 /result
-rw-r--r-- 3 goodee supergroup          265 2018-09-03 17:27 /result
/G/20180903_052716/part-r-00000        0 2018-09-04 14:22 /result
drwxr-xr-x - goodee supergroup          0 2018-09-04 14:22 /result
/G/20180904_022208/_SUCCESS            0 2018-09-04 14:22 /result
-rw-r--r-- 3 goodee supergroup          265 2018-09-04 14:22 /result
/G/20180904_022208/part-r-00000        0 2018-09-03 17:27 /result
drwxr-xr-x - goodee supergroup          0 2018-09-03 17:27 /result
/Y/20180903_052733/_SUCCESS            0 2018-09-03 17:27 /result
-rw-r--r-- 3 goodee supergroup          0 2018-09-03 17:27 /result
/Y/20180903_052733/part-r-00000        91 2018-09-03 17:27 /result
[root@project ~]#
```

(1) 빅데이터 분석

③ 받아온 데이터가 웹 화면에서 차트로 완성됨



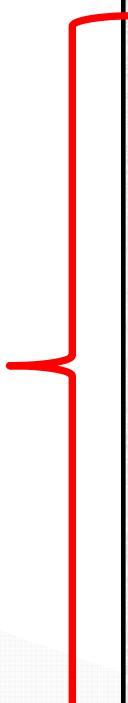
(2) 유저 관리 기능



(2) 유저 관리 기능

➤ 회원가입

① 사이트 요구조건에 따라
유저가 회원데이터를 작성



이메일

닉네임(15자 이하)

비밀번호(4~15자)

비밀번호 재확인

GOODEE MOVIE로부터 각종 알림 메일 수신 동의(선택사항)

완료!

취소

This image shows a user registration form for a website named 'GOODEE MOVIE'. The form consists of several input fields: '이메일' (Email), '닉네임(15자 이하)' (Nickname, 15 characters or less), '비밀번호(4~15자)' (Password, 4~15 characters), and '비밀번호 재확인' (Password Confirmation). Below the password fields is a checkbox for accepting email notifications from the site. At the bottom of the form are two buttons: a green '완료!' (Complete!) button and a red '취소' (Cancel) button. A red bracket on the left side of the image groups the four main input fields together.

(2) 유저 관리 기능

➤ 회원가입

gudi.ptime.org:10800 내용:
이미 가입되어 있는 아이디입니다.

회원가입

이메일
zhini0902@naver.com

닉네임(15자 이하)
구디맨

비밀번호(4~15자)
...

비밀번호 재확인
..

GOODEE MOVIE로부터 각종 알림 메일 수신 동의(선택사항)

② 조건 불일치 시
해당 내용에 대해 경고

“이미 가입되어 있는
아이디입니다” 등이
출력됨

(2) 유저 관리 기능

➤ 회원가입

③ 무사히 완료 시
회원 데이터가
저장되며,
관리자는
해당 유저
체크가 가능

사이트 관리		회원관리				
회원관리		회원번호	메일 주소	닉네임	리뷰 수	가입일
피드백 확인		2	goodee@gd.movies	구디맨	1	2018-08-24
		3	safebreaking@naver.com	kyoungjinPark	0	2018-08-24
		4	dd@dd	dddd	0	2018-08-24
		5	uuuu@uuuu	GUDIE	0	2018-08-28
		6	tis@tis	tis	0	2018-08-28
		7	port@port	포트바꿈!	0	2018-08-28
		8	nonono@nonono	nono	0	2018-08-28
		10	ameli@meli	ameli	0	2018-08-28

[1] [2] [3]

(2) 유저 관리 기능

▶ 로그인

① 회원가입한 내용을 입력.

만일 5회 이상 실패 시
다음 로그인 성공 때 경고

로그인

아직 GOODEE MOVIE에 가입하지 않으셨다면,
먼저 [이곳](#)에서 회원가입을 해주시기 바랍니다.

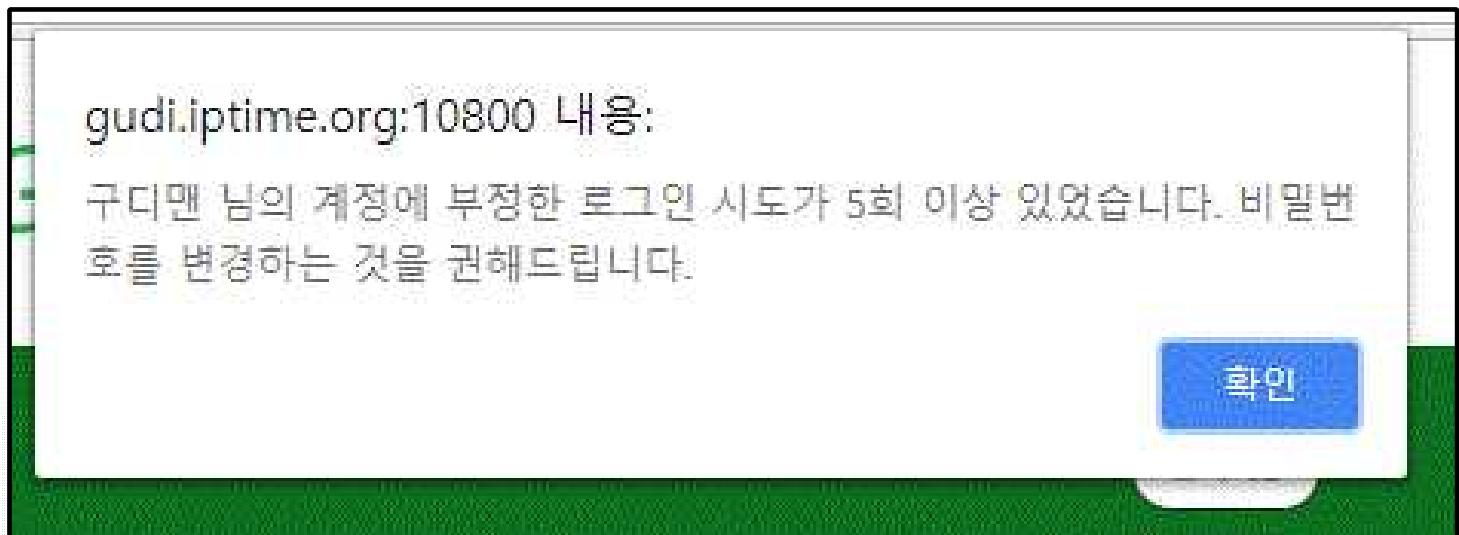
이메일 또는 닉네임

구디맨

비밀번호

....

로그인



(2) 유저 관리 기능

▶ 로그인

② 마이 페이지를 통해
자신의 정보를
확인 또는 변경 가능
(탈퇴 포함)

계정

계정정보

활동내역

계정정보 변경

비밀번호 변경

알림 설정

계정 삭제

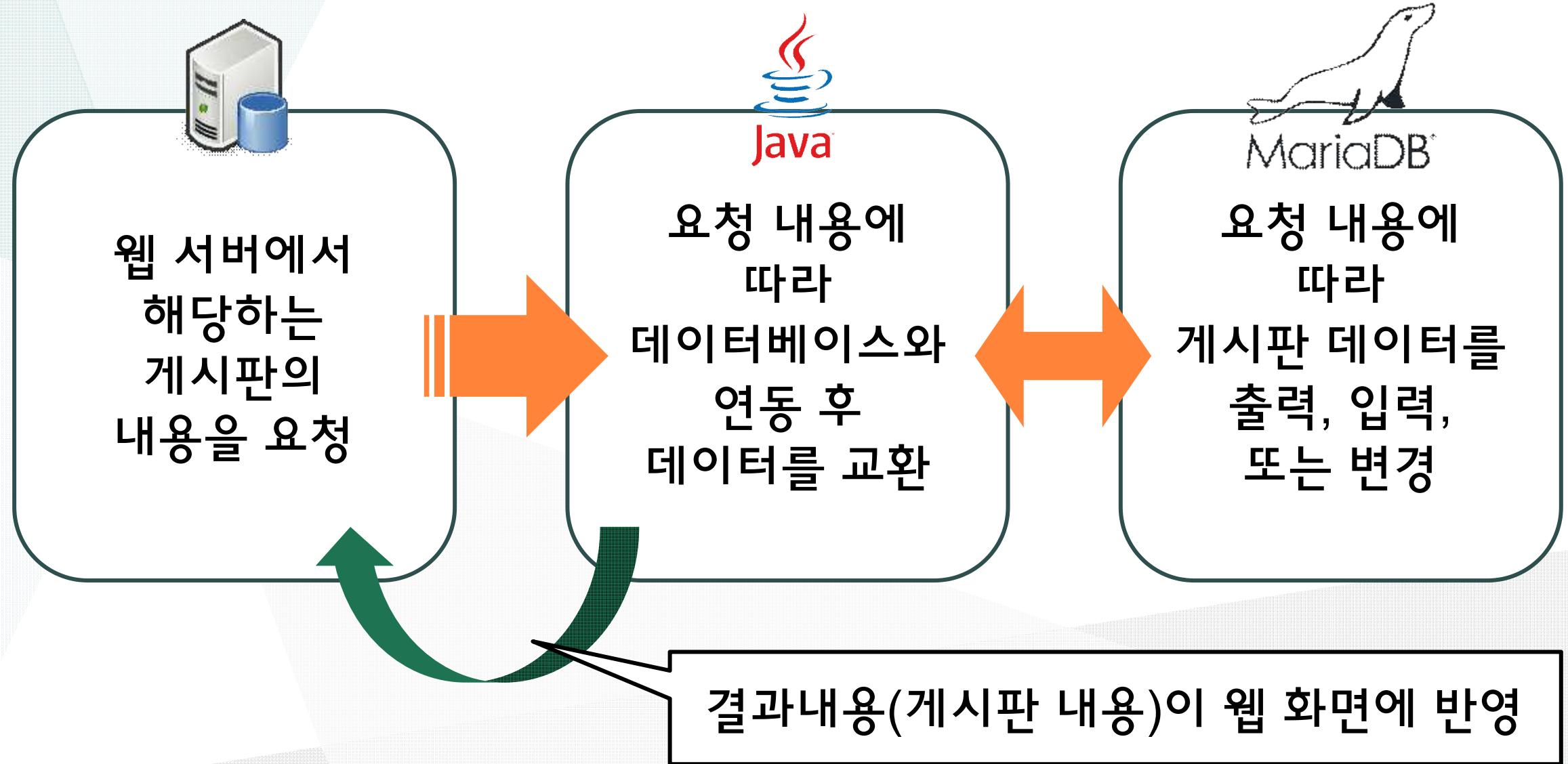
계정정보

이메일 : goodee@gd.movies

닉네임 : 구디맨

가입일 : 2018-08-24

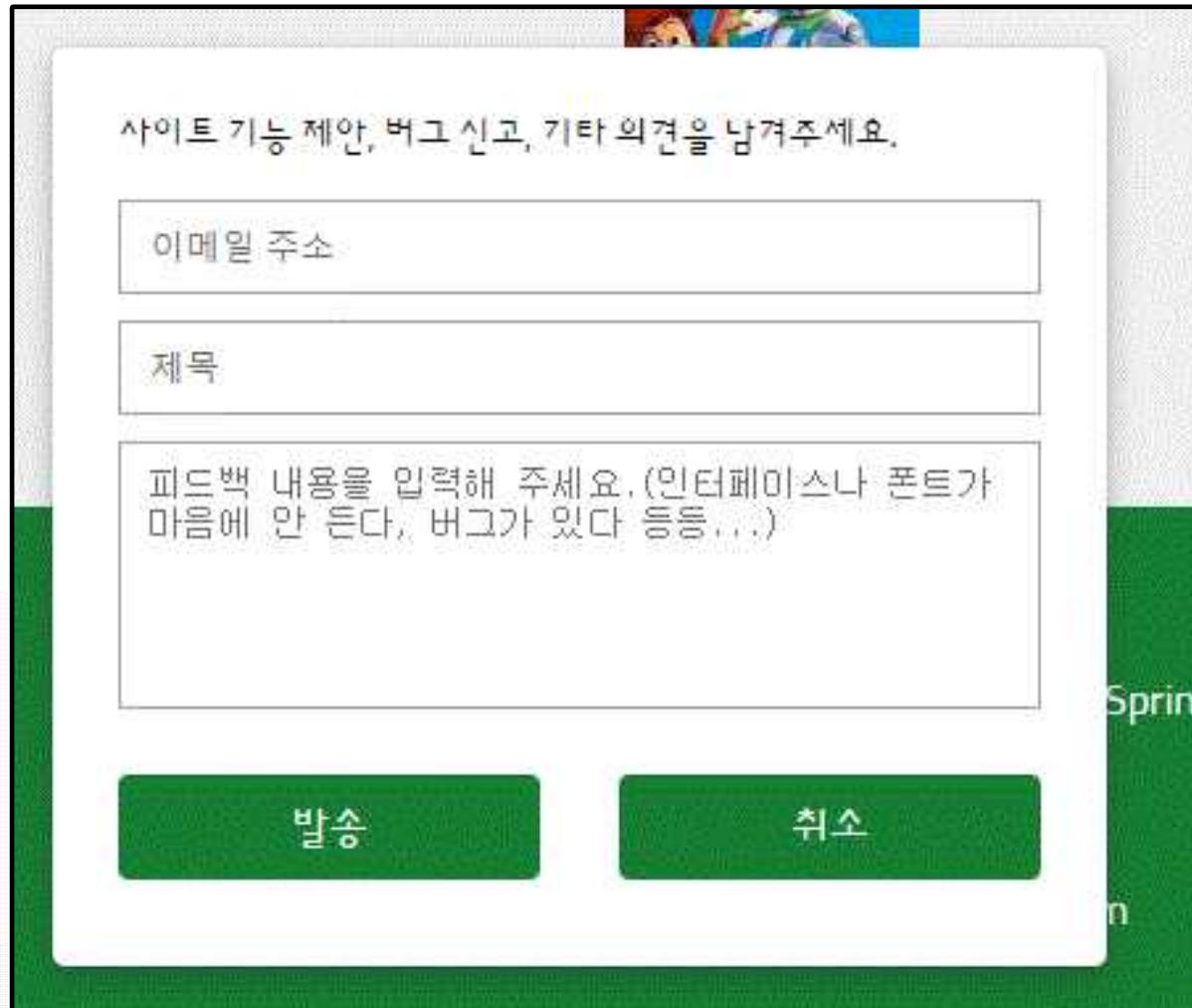
(3) 피드백과 게시판 구현



(3) 피드백과 게시판 구현

▶ 피드백

메인 화면에서
관리자에게
피드백이 전달 가능



(3) 피드백과 게시판 구현

▶ 게시판 구현

① 관리자를 위한
피드백 게시판과
유저 관리 게시판,
회원 전용
리뷰 게시판이 존재

사이트 관리		피드백 확인				
회원관리		번호	메일 주소	타이틀	발송일	체크
1	zhini0902@naver.com	구디무비 쪽 보고있습니다	2018-08-23	<input type="button" value="확인"/>		
2	hitohi@hi	UI가 깜에 만들어요.	2018-08-23	<input type="button" value="확인"/>		
3	zhini0902@naver.com	피드백 입력 성공?	2018-08-28	<input type="button" value="확인"/>		
4	zhini0902@naver.com	dddd	2018-08-28	<input type="button" value="확인"/>		
5	zhini0902@naver.com	123	2018-08-28	<input type="button" value="확인"/>		
6	zhini0902@naver.com	123	2018-08-28	<input type="button" value="확인"/>		
7	zhini0902@naver.com	23523adg	2018-08-28	<input type="button" value="확인"/>		
8	zhini0902@naver.com	88888	2018-08-28	<input type="button" value="확인"/>		

회원번호	메일 주소	닉네임	리뷰 수	가입일	탈퇴
2	goodee@gd.movies	구디맨	1	2018-08-24	<input type="button" value="탈퇴"/>

(3) 피드백과 게시판 구현

➤ 게시판 구현

② 페이지 클릭을 통해 제한된 데이터만 받아오게끔 설계함.

- ✓ 현재 페이지는 굵은 글씨로 표시
- ✓ 게시물의 양에 따라 페이지 수도 한정됨

[이전] [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [다음]

(3) 피드백과 게시판 구현

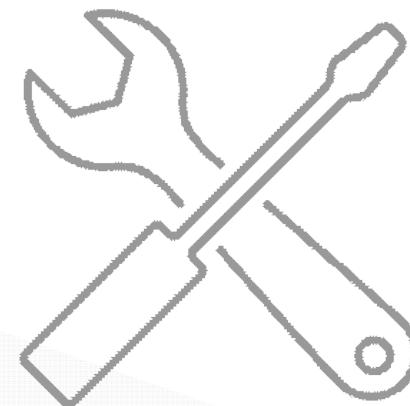
➤ 게시판 구현

③ 유저 게시판은 현재 미완성,

글쓰기/글 확인을 구현할 계획

번호	영화명	제목	닉네임	등록일	조회수	추천!
11	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
10	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
9	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
8	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
7	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
6	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
5	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
4	토이스토리	테스트입력	구디맨	09-03	0	<button>추천!</button>
2	토이스토리	토이스토리 괜찮죠.	dddd	08-31	0	<button>추천!</button>
1	토이스토리	토이스토리를 보고..	구디맨	08-31	2	<button>추천!</button>

4) 시행착오, 개선점



시행착오

- ✓ 개발 초기에 분석 대상 파일 탐색에 실패
- ✓ 원천 데이터 정제의 어려움 (“{'data':'data'}”)
- ✓ 하둡의 설정 관련 문제

개선점

- ✓ 계획과 달리 미완성된 기능이 존재 (유저게시판)
- ✓ 하둡을 통한 분석 내용의 디테일화 (연도별 또는 다른 차트의 사용)
- ✓ 게시글 추천이나 덧글 기능의 구현
- ✓ 코드의 간략화, 재사용성↑
- ✓ 모바일 전용 페이지의 구현

5) 시연



<http://gudi.iptime.org:10800>

6) Q & A



수고하셨습니다.



