

# 하둡 분산 파일 시스템 프로젝트

## All About Marvel Cinematic Universe

👤 박 경 진

📞 010-5148-1895

✉ safebreaking@naver.com



## Contents

- Schedule
- Motive
- HDFS
- Project

# Schedule



8월 13일

8월 20일

8월 27일

9월 3일

9월 5일

Java, Spring, 하둡 저장 및 시각화

화면 구성 HTML, CSS

수정 및 교정

테스트 및 보완

발표 준비



# Motive

각종 Analytics의 발달로 한결 쉬워진 데이터 분석의 시대

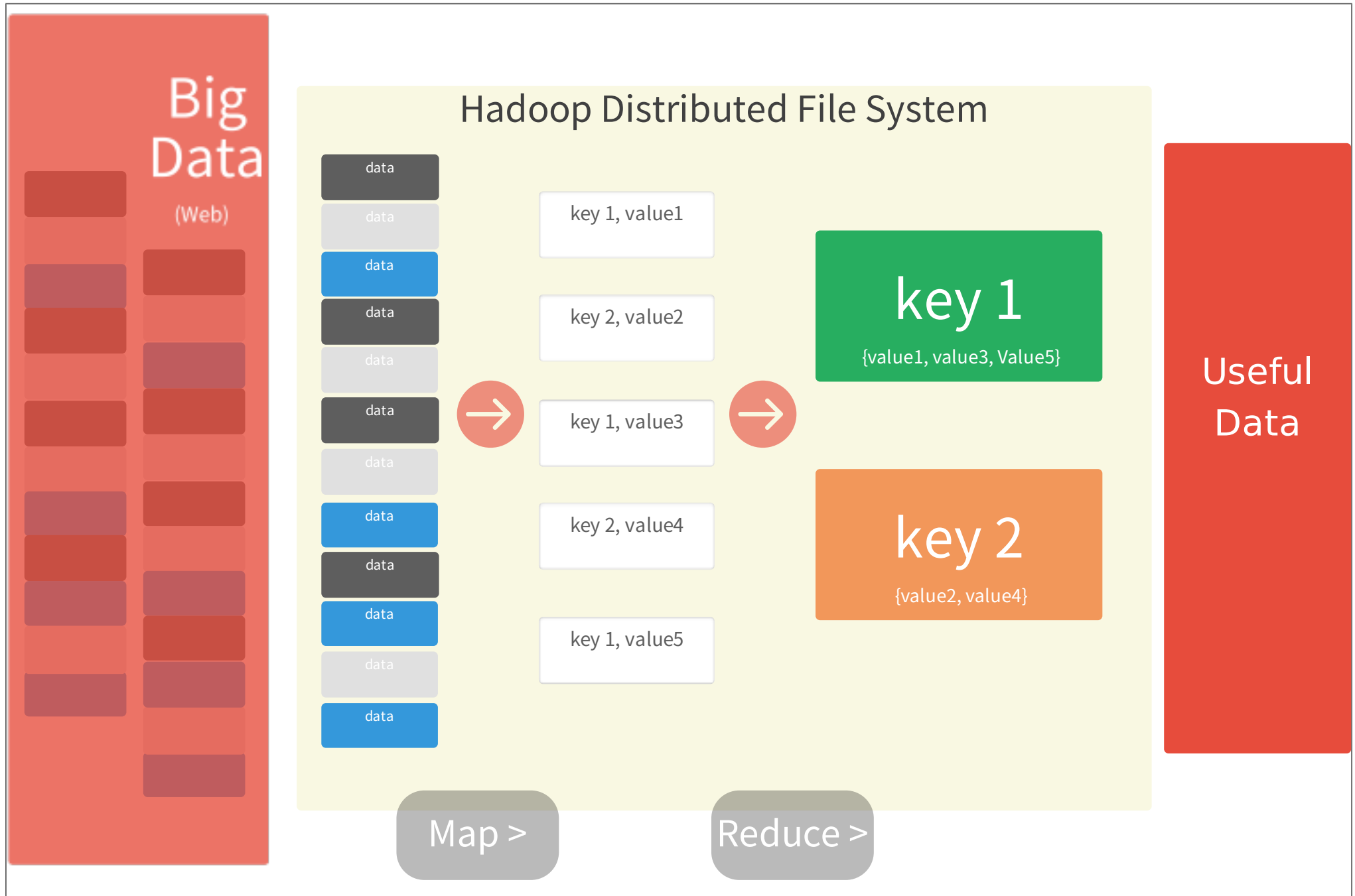
1. 직접 선별한 마블 캐릭터 Google 검색량 데이터를  
하둡 분산 파일 시스템으로 정제할 수 있는가?
2. 선별된 키워드의 검색량 변화를 데이터 전, 후 처리를 통해 시각화 할 수 있는가?
3. 스프링 프레임워크 안에서 마블(MCU) 소개 웹 페이지를 구현 할 수 있는가?



SPRING  
Framework



Data-Driven Documents





# Project

1. 위키피디아, 네이버지식인 데이터 크롤링
2. 구글 키워드 검색 데이터 ► 하둡 맵 리듀스 및 시각화
3. HTML, CSS, JS 를 사용하여 웹 사이트 구현

**MARVEL** STUDIOS



<b>Marvel Cinematic Universe</b>	
<b>Marvel Cinematic Universe</b>	
Marvel Cinematic Universe Infante from Marvel Studios: Assembling a Universe (2014)	
<b>Created by</b>	Marvel Studios
<b>Original work</b>	April 4th 2008
	<i>Print publications</i>
<b>Comics</b>	Marvel Cinematic Universe tie-in comics
	<i>Films and television</i>
<b>Film(s)</b>	Marvel Cinematic Universe films
<b>Short film(s)</b>	Marvel One-Shots
<b>Television series</b>	Television series
<b>Web series</b>	Wrecked Agents of S.H.I.E.L.D. Slingshot
	<i>Audio</i>
<b>Original music</b>	Music of the Marvel Cinematic Universe
	<i>Miscellaneous</i>
<b>Theme park attractions</b>	Marvel Land Marvel Experience Guardians of the Galaxy – Mission: Breakout!

```
[root@name ~]# hdfs dfs -cat /input/data/team1
2013-01-01,26,8,56,8,11
2013-01-02,39,12,60,6,6
2013-01-03,12,12,32,0,6
2013-01-04,24,9,27,0,0
2013-01-05,23,7,27,0,0
2013-01-06,13,10,23,7,0
2013-01-07,6,8,31,0,0
2013-01-08,39,22,20,0,0
2013-01-09,11,17,31,0,0
2013-01-10,11,14,28,0,11
2013-01-11,26,0,26,0,0
2013-01-12,33,23,43,7,7
2013-01-13,39,19,29,6,0
2013-01-14,22,8,30,0,5
2013-01-15,17,8,30,6,6
2013-01-16,11,11,19,0,6
2013-01-17,28,11,14,0,0
2013-01-18,26,17,17,6,0
2013-01-19,13,10,39,0,7
2013-01-20,35,6,16,0,0
2013-01-21,27,14,24,5,0
```

## 네이버지식인

## 구글트랜드

위키피디아

## 화면 구성

MARVEL STUDIOS

MARVEL STUDIOS

Nav 버튼



MARVEL STUDIOS

EL STUDIOS MARVEL STUDIOS MARVEL STUDIOS MARVEL STUDIOS MARVEL STUDIOS MARVEL STUDIOS MARVEL

Marvel Cinematic Univers 10th Anniversary Fan Page.

You may have better understanding of Marvel movies and characters.  
Thank you for visiting. We hope you have MARVELous experience here.



This site is optimized for learning and enjoying Marvel Studio.  
While using this site, you agree to have read and accepted our terms of use, cookie and privacy policy.  
Copyright 2018 by Kyoungjin Park. All Rights Reserved.



마블 스튜디오 10 주년

MCU는 무엇이고, 마블은 무엇인가?

MCU from Wikipedia

Marvel from 지식인

내 용 화 면

메 뉴



## 크롤링

MARVEL STUDIOS

MARVEL STUDIOS

MARVEL STUDIOS

## Marvel Cinematic Universe

The **Marvel Cinematic Universe (MCU)** is an American media franchise and **shared universe** that is centered on a **series of superhero films**, independently produced by **Marvel Studios** and based on characters that appear in **American comic books** published by **Marvel Comics**. The franchise has expanded to include comic books, short films, television series, and digital series. The shared universe, much like the original **Marvel Universe** in comic books, was established by **crossing over** common plot elements, settings, cast, and characters. **Phil Coulson**, portrayed by **Clark Gregg**, is an original character to the MCU and the only character to appear across all its different media.

The first film released in the MCU was *Iron Man* (2008), which began the first phase of films culminating in the crossover film *Marvel's The Avengers* (2012). Phase Two began with *Iron Man 3* (2013), and concluded with *Ant-Man* (2015). The films are currently in Phase Three, which began with the release of *Captain America: Civil War* (2016) and is set to conclude with an untitled *Avengers* film (2019). Marvel Television expanded the universe further, first to network television with *Marvel's Agents of S.H.I.E.L.D.* on ABC in the 2013–14 television season, followed by online streaming with *Marvel's Daredevil* on Netflix in 2015 and *Marvel's Runaways* on Hulu in 2017, and then to cable television with *Marvel's Cloak & Dagger* in 2018 on Freeform. Marvel Television has also produced the digital series *Marvel's Agents of S.H.I.E.L.D.: Slingshot*, which is a supplement to *Agents of S.H.I.E.L.D.* Soundtrack albums have been released for all of the films, along with many of the television series, as well as the release of compilation albums containing existing music heard in the films. The MCU also includes **tie-in comics** published by Marvel Comics, while Marvel Studios has also produced a series of **direct-to-video short films** and a **viral marketing** campaign for its films and the universe with the faux news program *WHIH Newsfront*.

The franchise has been commercially successful as a multimedia shared universe, though some critics have found that some of its films and television series have suffered in service of the wider universe. It has inspired other film and television studios with comic book character adaptation rights to attempt to create similar shared universes. The MCU has also been the focus of other media, outside of the shared universe, including attractions at various **Walt Disney Parks and Resorts**, an attraction at *Discovery Times Square*, a *Queensland Gallery of Modern Art* exhibit, two television specials, guidebooks for each film, multiple tie-in video games, and commercials.



## 마블 스튜디오 10주년

MCU는 무엇이고, 마블은 무엇인가?

MCU from Wikipedia

Marvel from 지식인

MARVEL STUDIOS

## 마블 지식인 성지글

마블 코믹스 속 세계에는 수많은 위험이 도사리고 있습니다. 현실에 비해 압도적으로 인간의 돌연변이 출산율이 높을때다, 출산되는 돌연변이들의 태반이 각종 가공할 위험의 초능력이라는 것이 바로 위험거리였죠. 이러한 돌연변이들은 대다수가 외관상의 구별이 어려우려다 제어할 방법이 돌연변이 본인의 의지가 아닌 이상은 없다시피 하기에 마블 코믹스의 범죄들은 하늘을 찌를 수밖에 없었고, 높은 범죄율의 원인이 되는 만큼 돌연변이들은 사회적으로 천대를 받는 것이 일반적입니다. 그에 맞춰 돌연변이의 통제가 필요하다고 느끼는 돌연변이들의 뜻이 맞아 탄생한 집단이 바로 '엑스맨'입니다.

엑스맨(X-Men)은 여러 의미를 가진 고유명사입니다. 돌연변이를 발전시키는 X 유전자를 보유한 사람이라는 의미와, 자비에 영재 학교의 설립자인 찰스 자비에 교수의 성에서 따온 XX(Xavier)에서 따온, 자비에의 부하 내지는 사람들이라는 의미를 가지죠. 찰스 자비에의 전 세계에서 가장 강력한 텔레파시 능력을 가진 인물로, 자신의 능력을 통제하지 못해 고통받거나 올바른 인도를 받지 못해 악의 길을 걷게 되는 아이들을 받아들여 올바른 사상과 능력 통제를 가르쳐 왔던 돌연변이가 되도록 훈련하는 것이 바로 자비에와 그가 설립한 자비에 영재 학교의 목표입니다.

자비에 영재 학교의 교수진과 몇몇 학생들은 능력이 능력인만큼 학업 뿐만이 아니라 다양한 평화 유지 업무, 돌연변이 인권 운동에도 관여합니다. 대개 돌연변이와 관련된 범죄나 테러에 관여하는 편이며, 가끔 범우주적인 능력을 가진 돌연변이들이 등장하기도 하기 때문에 우주에까지 활동 범위를 넓히는 경우도 간혹 있습니다.

마블 코믹스의 위험은 그뿐만이 아닙니다. 유독 현실에 비해 다양한 사상을 가진 초월적인 수준의 천재들이 많이 나오는 탓에, 각종 첨단 장비나 생체적인 개조를 통한 범죄들도 많이 일어납니다. 돌연변이 범죄들만으로도 사회가 휘청대는데 거기다 각종 악당들의 범죄까지 추가되는 것이죠. 지구에 넘쳐나는 악당들 때문에 굴이 아픈, 평화를 원하는 이들이 모여 탄생한 집단이 바로 '어벤저스'입니다.

근본적으로 어벤저스와 엑스맨은 평화 유지가 목적인 만큼 서로의 활동 범위가 겹치기 마련입니다. 물론 엑스맨은 돌연변이와 관련된 갈등이 주요 목표라면 어벤저스는 지구의 모든 위험이 되는 존재와의 갈등에 관여하는 편이라서 활동 범위 자체는 어벤저스가 더 넓다고 할 수 있지만, 어찌 되었든 활동 범위가 겹치는 탓에 두 집단은 멤버들 간의 소통과 교류가 잦고, 그만큼 갈등도 많이 겪습니다. 어벤저스가 감당하지 못할 위험이 있다면 같이 힘을 합치는 경우도 있고, 개인적인 문제로 엑스맨을 떠나 어벤저스로 활동하거나 그 반대의 길을 걷는 이들도 많죠.

여기에 한술 더 떠서 지구에서의 문제를 주로 다루는 게 어벤저스와 엑스맨이라면, 범우주적인 영역을 아우르는 평화 수호 집단은 '가디언즈 오브 갤럭시'라고 할 수 있습니다. 당연히 우주 전체가 활동 범위이니 지구에서도 가끔 돌리고 서로 교류도 하는 편이며, 멤버의 이동도 간혹 있습니다. 이렇게 방대한 세계의 이야기를 다룬 것이 바로 마블 코믹스의 주된 주제로, 스케일이 차원이 다른 만큼 이야기의 크기도 말로 표현하지 못할 정도로 깊고 복잡합니다.

원래대로라면 이 세 집단(언급되지 않았지만 원작에는 한참 더 많습니다)에 소속된 수많은 멤버들의 영상화 저작권은 전부 마블 코믹스의 소유였습니다. 하지만 마블 코믹스가 1990년대에 재정난을 겪게 되고, 이런 매력적인 이야기를 가진 캐릭터들의 영상화 판권을 당시 할리우드의 메이저 스튜디오에 넘겨 파산 위기를 모면하곤 했던 과거가 있습니다.

대표적인 마블 코믹스의 강판 캐릭터인 스파이더맨은 소니/콜롬비아 픽처스에 넘어가(스파이더맨)상부작으로 제작되어 허여로 영화의 가능성을 열게 되었습니다. 원래



## 마블 스튜디오 10주년

MCU는 무엇이고, 마블은 무엇인가?

MCU from Wikipedia

Marvel from 지식인

# Hadoop MapReduce

**MARVEL** STUDIOS

```
[root@name ~]# hdfs dfs -cat /input/data/team1
2013-01-01,26,8,56,8,11
2013-01-02,39,12,60,6,6
2013-01-03,12,12,32,0,6
2013-01-04,24,9,27,0,0
2013-01-05,23,7,27,0,0
2013-01-06,13,10,23,7,0
2013-01-07,6,8,31,0,0
2013-01-08,39,22,20,0,0
2013-01-09,11,17,31,0,0
2013-01-10,11,14,28,0,11
2013-01-11,26,0,26,0,0
2013-01-12,33,23,43,7,7
2013-01-13,39,19,29,6,0
2013-01-14,22,8,30,0,5
2013-01-15,17,8,30,6,6
2013-01-16,11,11,19,0,6
2013-01-17,28,11,14,0,0
2013-01-18,26,17,17,6,0
2013-01-19,13,10,39,0,7
2013-01-20,35,6,16,0,0
2013-01-21,27,14,24,5,0
```



```
[root@name ~]# hdfs dfs -cat /result/20180903152705/part-r-00000
201301 Hulk 343
201302 Hulk 327
201303 Hulk 429
201304 Hulk 52
201305 Hulk 79
201306 Hulk 64
201307 Hulk 152
201308 Hulk 137
201309 Hulk 138
201310 Hulk 159
201311 Hulk 143
201312 Hulk 139
201401 Hulk 84
201402 Hulk 92
201403 Hulk 112
201404 Hulk 407
201405 Hulk 339
201406 Hulk 309
201407 Hulk 224
201408 Hulk 255
201409 Hulk 219
201410 Hulk 175
```

**MARVEL** STUDIOS

Hadoop Distributed File System데이터 분석중

**MARVEL** STUDIOS

Hulk

데이터 분석 완료

**MARVEL** STUDIOS

MCU 페이즈 2 캐릭터  
2013년 ~ 2018년 7월  
Google key word 검색 데이터

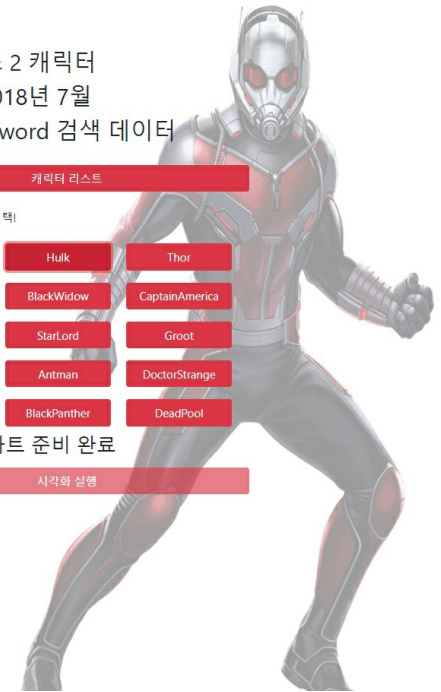
캐릭터 리스트

분석 하려는 캐릭터를 선택!

IronMan	Hulk	Thor
Hawkeye	BlackWidow	CaptainAmerica
WinterSoldier	StarLord	Groot
Thanos	Antman	DoctorStrange
Spiderman	BlackPanther	DeadPool

Hulk데이터 차트 준비 완료

시각화 실행



## 시각화

MARVEL STUDIOS

MARVEL STUDIOS

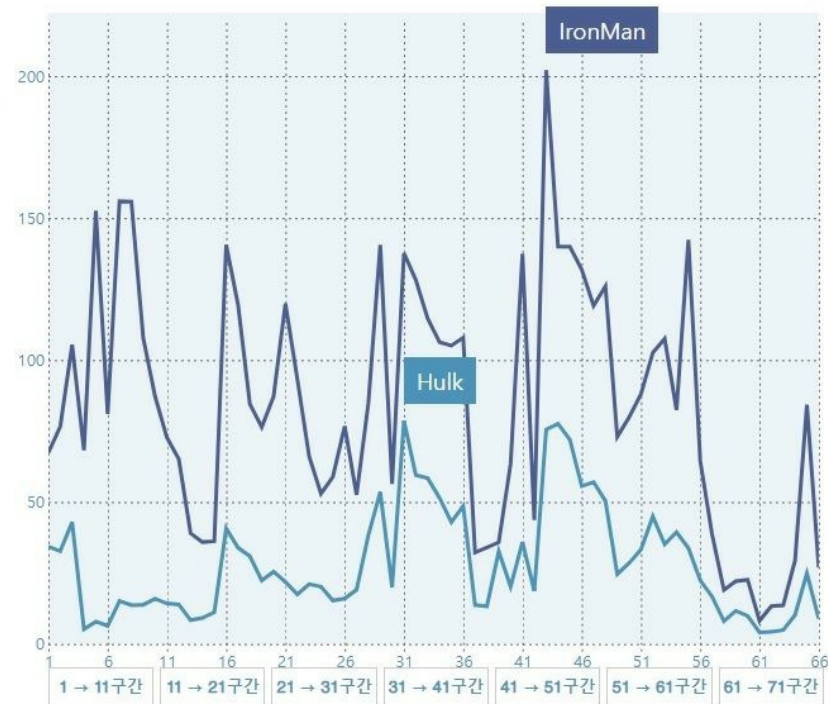
# MARVEL STUDIOS

## Marvel STATS :

캐릭터 이름 Keyword  
Google 검색량

IronMan

Hulk



## 기 타 화 면

MARVEL STUDIOS

MARVEL STUDIOS

MARVEL STUDIOS

MARVEL  
S T U D I O S

## PHASE 1

IRON MAN

May 2, 2008

June 13, 2008

THE INCREDIBLE HULK

IRON MAN 2

May 7, 2010

MARVEL STUDIOS

마블 영화를 즐기기 위한 Tip  
MCU 개봉 순서

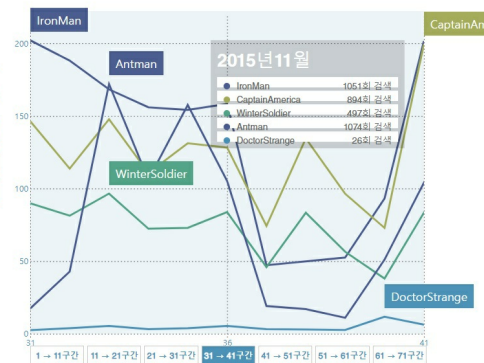
실영



## Marvel STATS :

캐릭터 이름 Keyword  
Google 검색량

IronMan Thor  
CaptainAmerica Hulk  
Hawkeye BlackWidow Spiderman  
WinterSoldier StarLord Groot  
Thanos Antman DoctorStrange  
BlackPanther Deadpool

핵심 MCU 캐릭터  
관심도 변화 분석

MCU 캐릭터 시작과 실영





# 프로젝트시연

[Youtube 영상 재생](#)

# Q & A



Fin.

