



1 / 37



치킨4조



모바일/PC게임
데이터
엔지니어링
프로젝트



게임 데이터 파이프라인 구축을 활용한 카테고리 및 리뷰 분석



목차

1. 프로젝트 개요

- 프로젝트 소개 및 목표
- 팀원 역할 소개

2. 파이프라인 구성

- 전체 데이터 파이프라인 소개

3. 모바일 게임

- 모바일 데이터 파이프라인
- 모바일 데이터 ETL 과정 및 분석

4. PC 게임

- 스팀 데이터 파이프라인
- 스팀 데이터 ETL 과정 및 분석
- 메타크리틱 데이터 파이프라인
- 메타크리틱 데이터 ETL 과정 및 분석

5. 게임 방송

- 트위치 데이터 파이프라인
- 트위치 데이터 ETL 과정 및 분석

6. 개발 후기

- 개발 후기 및 발전 방향 제시

01.

프로젝트 개요

프로젝트 소개 및 목표
팀원 역할 소개



1. 프로젝트 개요

01

프로젝트 소개

- 모바일과 PC 각 플랫폼 게임의 인기 카테고리 유형과 리뷰 분석을 통해 사용자들이 카테고리별로 어떤 평가를 하는지 파악하고자 한다.
- 전문가 평점과 일반 사용자 평점을 비교하여 어떠한 차이가 있는지 확인한다.
- 트위치 방송과 시청자수를 통해 사용자들이 어떠한 게임에 직, 간접적인 관심이 있는지 파악한다.



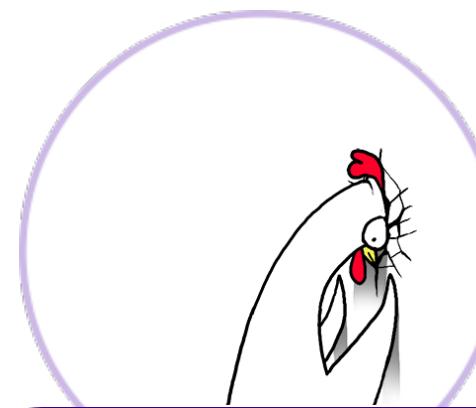
1. 프로젝트 개요

01

팀원 소개

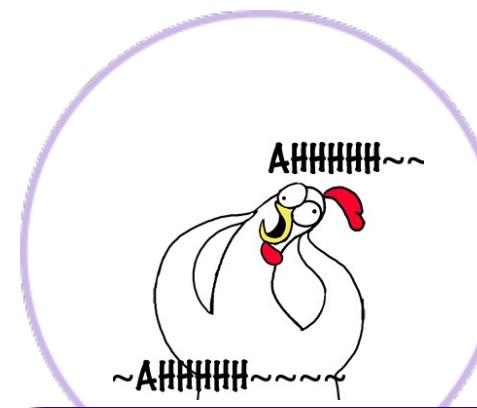


치킨4조



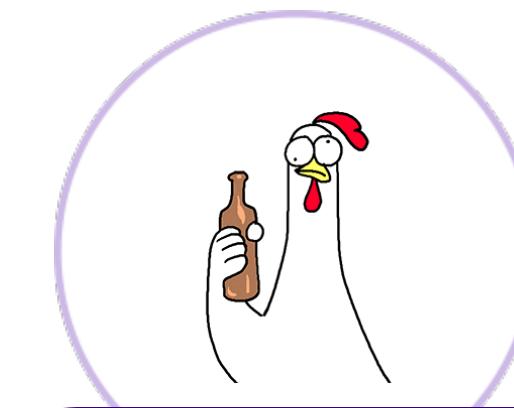
팀장 : 민윤지

모바일 게임 담당



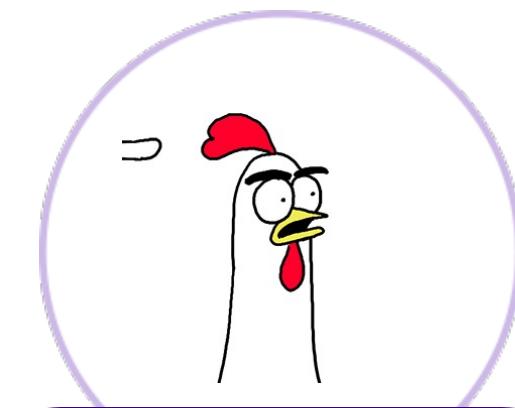
강규란

게임 방송 담당



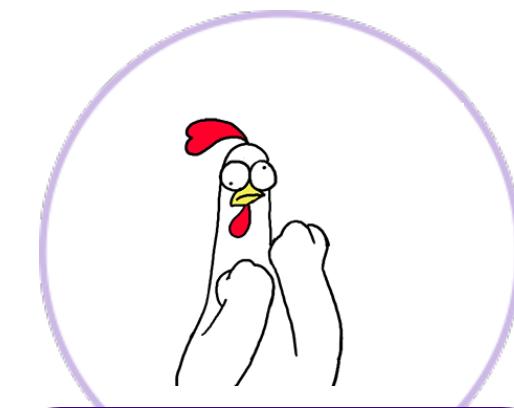
김희원

PC 게임 담당



박시현

모바일 게임 담당



추호연

PC 게임 담당

•팀원 가나다 순 표기

02.

파이프라인 구성

전체 데이터 파이프라인 소개

2. 파이프라인 구성

02

전체 데이터 파이프라인

데이터 소스



Selenium
BeautifulSoup

크롤링

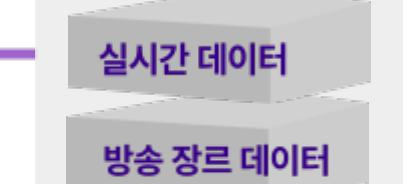
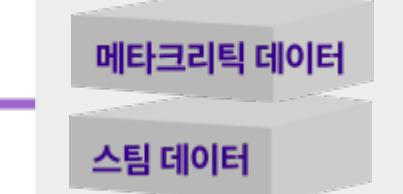
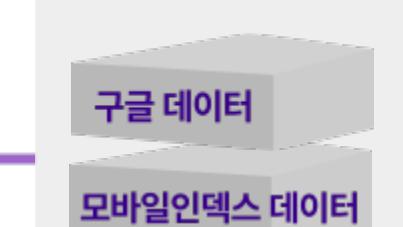
Scrapy
BeautifulSoup

크롤링

Scrapy Selenium

크롤링

데이터 웨어하우스

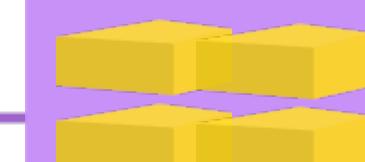
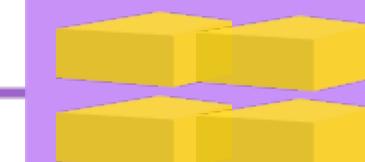
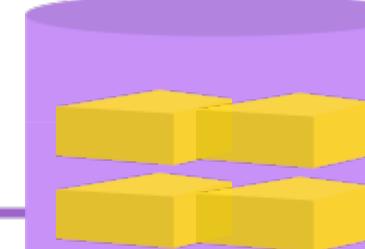


스파크 데이터 가공

스파크 데이터 가공

스파크 데이터 가공

데이터 마트



시각화



머신러닝

긍부정 리뷰
판단 머신러닝
구현

점수 기반 게임
추천 머신러닝
구현



03.

모바일 게임

모바일 데이터 파이프라인

모바일 데이터 ETL 과정 및 분석



3. 모바일 게임

03

모바일 게임

- 대상 : 구글 플레이 스토어에 등록된 게임
- 기간 : 2019년 1월 1일 ~ 2022년 6월 14일

3. 모바일 게임

03

모바일 인기 게임 카테고리



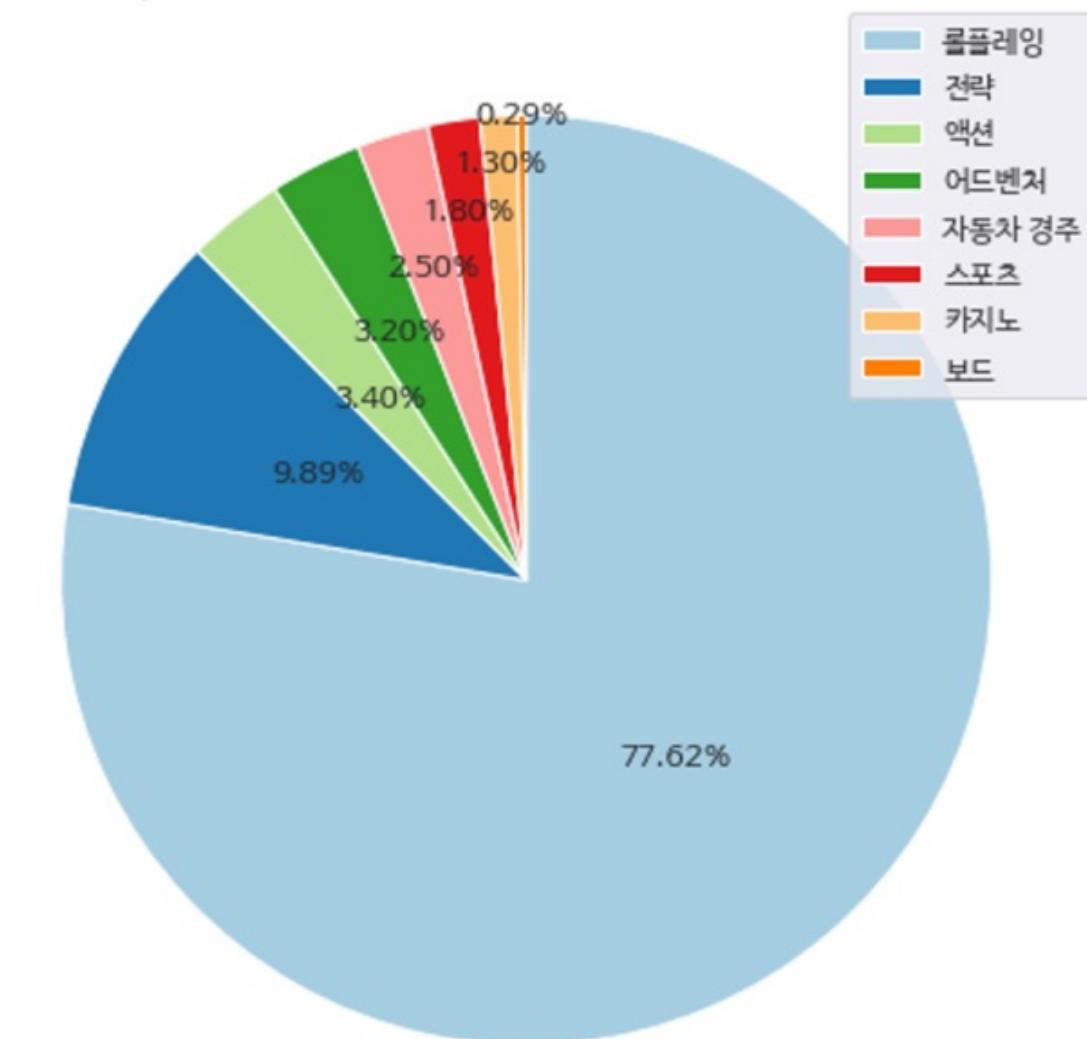
MOBILEINDEX
I N S I G H T

크롤리

1 ~ 45위 까지
순위와 게임

카테고리	카테고리별 랭크 횟수
0 룰플레잉	12594
1 전략	1599
2 액션	545
3 어드벤처	518
4 자동차 경주	402
5 스포츠	284
6 카지노	213
7 보드	53

top10에 가장 많이 랭크된 카테고리 비율



- 룰플레잉 : RPG라고도 부르며, 각자에게 할당된 캐릭터를 조작하고 가상의 상황에서 주어지는 시련을 극복하며 목표 달성을 목적으로 하는 게임

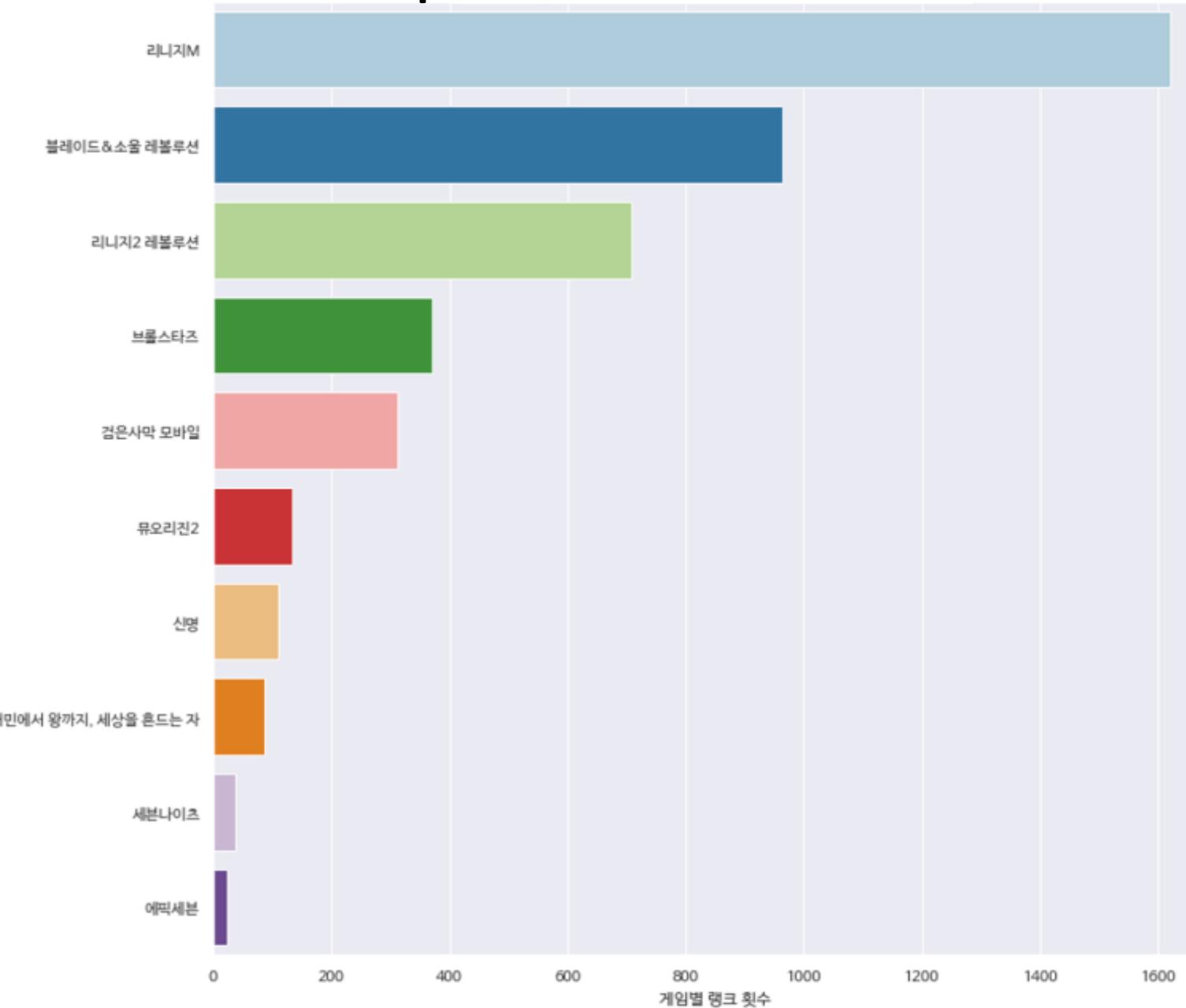
3. 모바일 게임

01

모바일 인기 게임

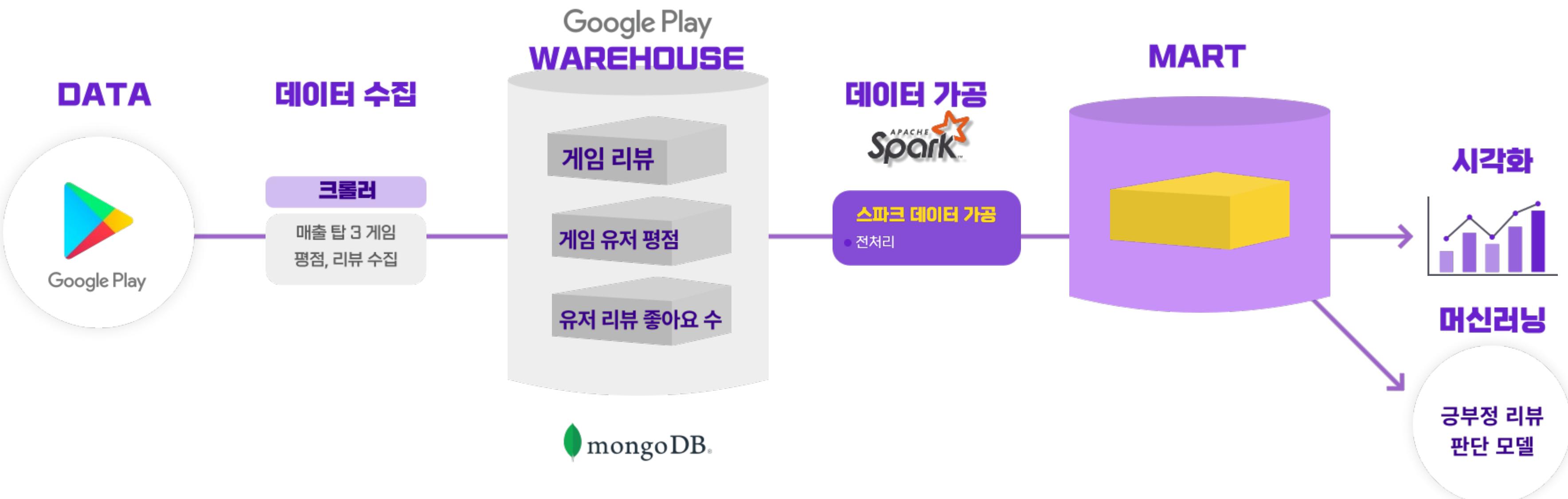
게임명	카테고리	게임별 랭크 횟수
리니지M	롤플레잉	1620
블레이드&소울 레볼루션	롤플레잉	963
검은사막 모바일	롤플레잉	311
리니지2 레볼루션	롤플레잉	708
뮤오리진2	롤플레잉	133
브롤스타즈	액션	370
왕이되는자 - 서민에서 왕까지, 세상을 흔드는 자	롤플레잉	88
에픽세븐	롤플레잉	23
세븐나이즈	롤플레잉	38
신명	롤플레잉	110

<Top 10에 가장 많이 랭크된 게임>



3. 모바일 게임

구글플레이스토어 파이프라인



3. 모바일 게임

구글플레이스토어 데이터 전처리

- content : 리뷰
- score : 평점

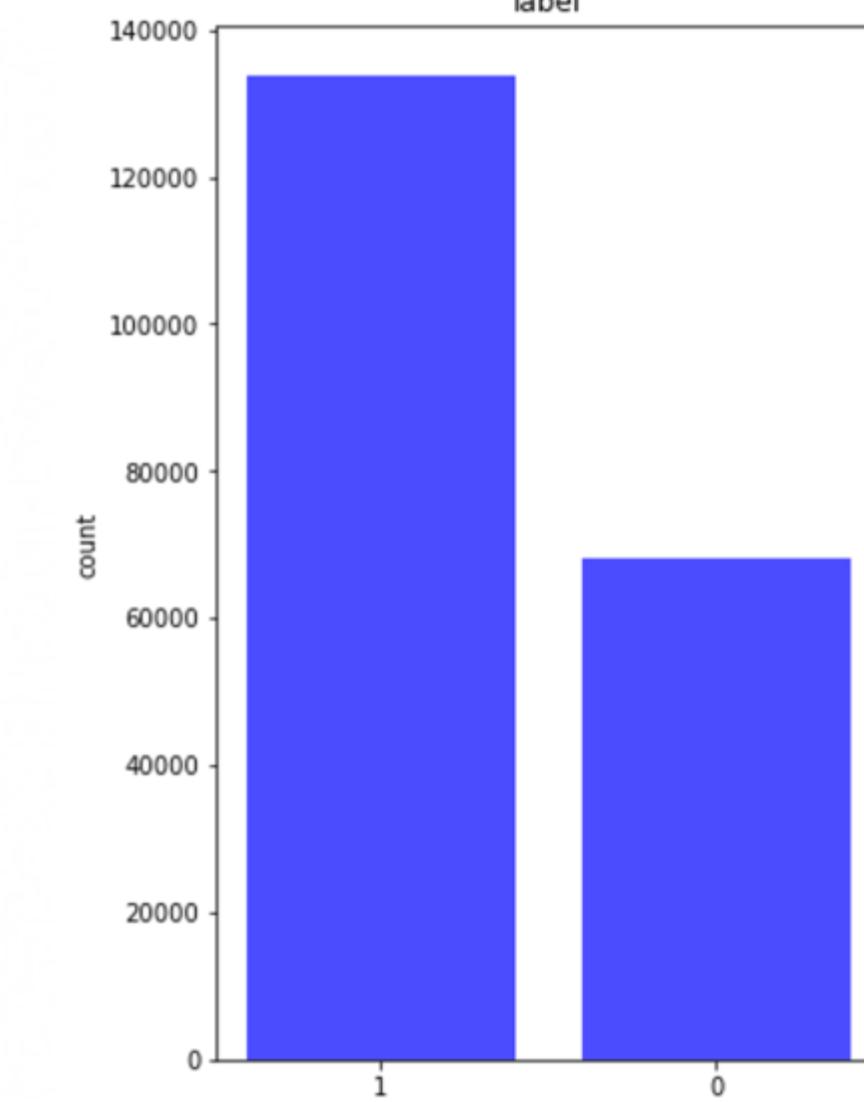
평점 4, 5 점이면 긍정 리뷰 (label = 1)

score	label	content
5	1	인류 역사상 최고의 게임
4	1	드뎌 —망 —드
4	1	감사.
5	1	제가잘못했습니다 계정락좀 풀어주...
5	1	굿
5	1	사냥터 분류좀 해주세요
5	1	Nc는 회사입니다.
4	1	5주년 쿠폰 일반 룬복구 쿠폰 ...
5	1	옛날보다좋아짐
5	1	1990년대 리니지 PC 유료게...
5	1	갓
5	1	재밌어요
5	1	캐릭접속시 자꾸팅김
5	1	4년을해도 그자리 —ㅜㅜ 과금러...
5	1	업데이트 다운로드 잘 됨.
5	1	업데이트를 하루 종일할래?
5	1	NC 화이팅!
4	1	피케이를 싫어하는 다수의유저가 ...
5	1	굿
5	1	재밌게 놀고 있다

평점 1, 2, 3 점이면 부정 리뷰 (label = 0)

score	label	content
1	0	노잼현질유도게임 핵과금유저 넘쳐...
1	0	nc벌럼 이거 아직 안망했어?
1	0	자동사냥하면 안돼나요 왜 자동 ...
1	0	에휴
1	0	도박장
1	0	페이투원
1	0	500명 대기자있어서 3번기달렸...
1	0	재미 읍다
1	0	한국산 게임의 수준을 보여주는 쓰레기
1	0	망겜 입니다 하지마세요 과금유도...
1	0	10억 있다면하세요 근데 10억...
1	0	아니 설치도 안했는데 결제가 1...
1	0	이걸 아직도 한다고요? 접은지 ...
1	0	엔씨게임은 거르는게 답
1	0	암게임 현금이 3천이상 있는사람...
1	0	그냥 현질 많이 하는 사람이 짱...
1	0	평점0점!!
2	0	이
1	0	아직도 돈 써가며 게임하는 사람...
1	0	걍 인간 쓰레기게임

label	count
1	133891
0	68219



3. 모바일 게임

리뷰감성분석 머신러닝 구현

- 판다스 데이터프레임으로 가져온 후 훈련데이터와 테스트 데이터를 3:1의 비율로 분리함

	content	score	label
노잼현질유도게임 핵과금유저 넘쳐나서 무과들은 게임 시작도 전에 끝나는게임 안 하는게 이득	1	0	
nc벌럼 이거 아직 안망했어?	1	0	
자동사냥하면 안돼나요 왜 자동 사냥 하면 와서 죽이나요 npc /리칸트 사진두 찍 어...	1	0	
에휴	1	0	
도박장	1	0	

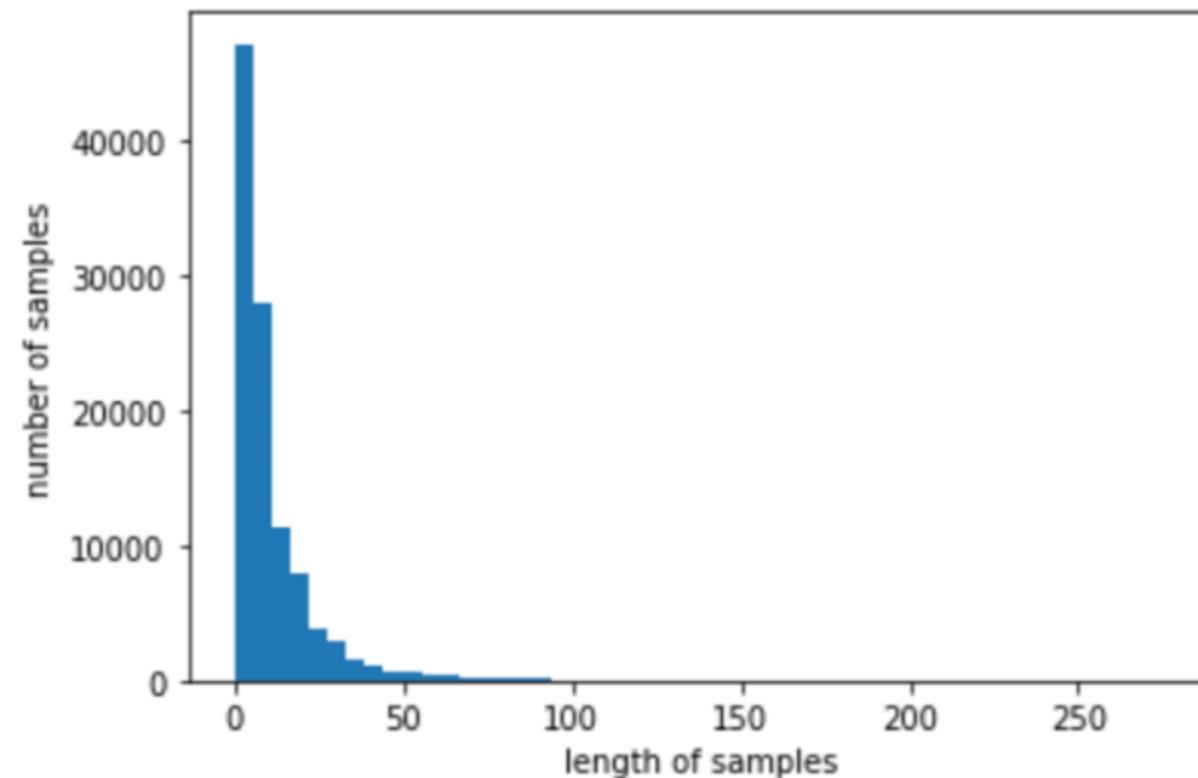
```
# 훈련데이터와 테스트데이터를 3:1의 비율로 분리한다
train_data, test_data = train_test_split(data, test_size = 0.25, random_state = 42)
print('훈련용 리뷰의 개수 :', len(train_data))
print('테스트용 리뷰의 개수 :', len(test_data))
```

훈련용 리뷰의 개수 : 106784
 테스트용 리뷰의 개수 : 35595

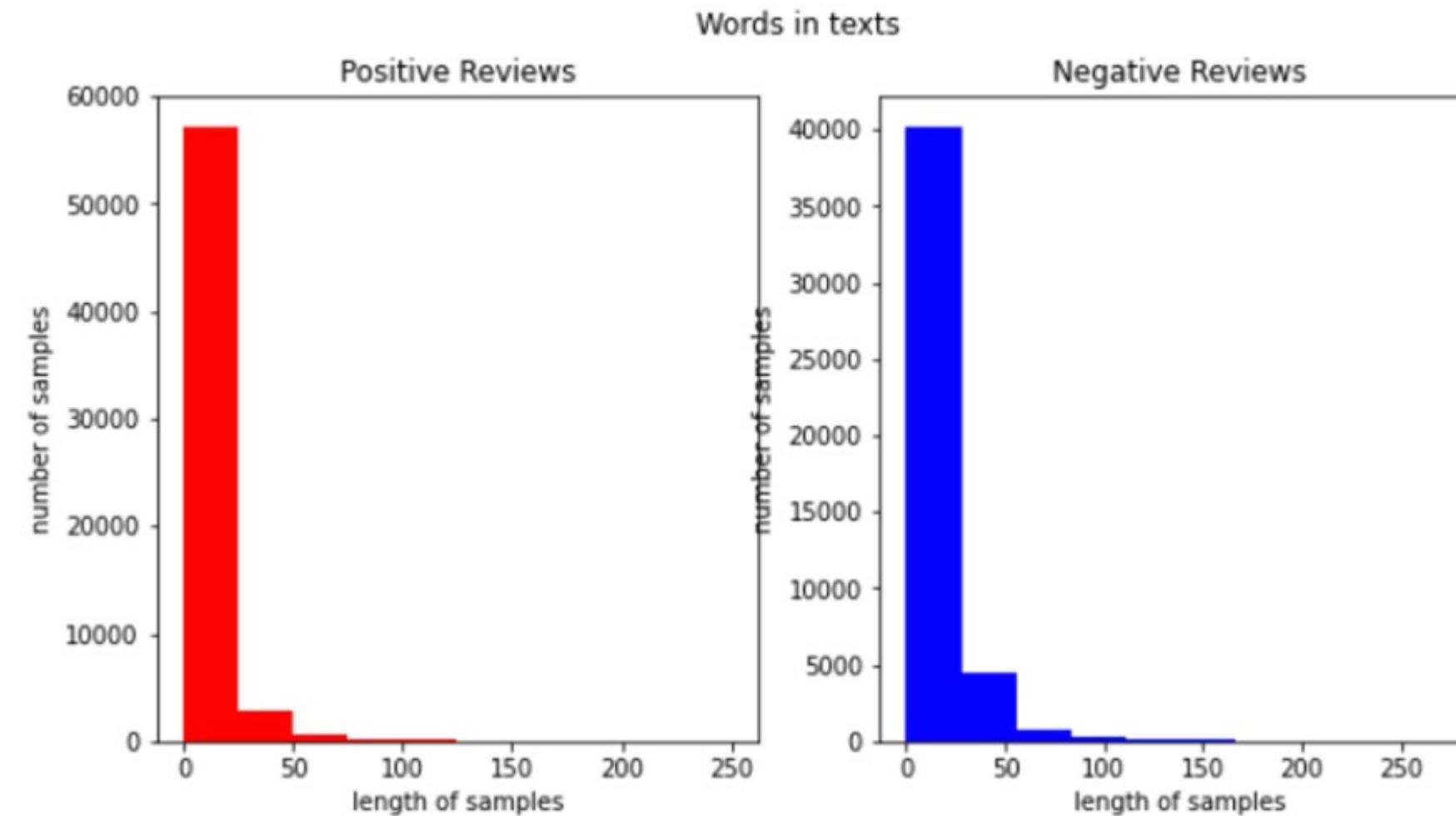
3. 모바일 게임

리뷰감성분석 머신러닝 구현

리뷰의 최대 길이 : 277
리뷰의 평균 길이 : 11.096278468684448



긍정 리뷰의 평균 길이 : 8.533998916238362
부정 리뷰의 평균 길이 : 14.4969597907813



긍정리뷰와 부정리뷰 평균 길이

- 부정 리뷰가 긍정 리뷰보다 더 길게 작성되는 경향이 있음.

3. 모바일 게임

리뷰감성분석 머신러닝 구현

분석의 정확도

```
28] loaded_model = load_model('review_model.h5')
    print("\n 테스트 정확도: %.4f" % (loaded_model.evaluate(X_test, y_test)[1]))  
  
1113/1113 [=====] - 41s 37ms/step - loss: 0.3869 - acc: 0.8288  
  
테스트 정확도: 0.8288
```

코랩에서 구현

https://colab.research.google.com/drive/1q26EFsyet1D_uPCqmU68001iW66UvzJ#scrollTo=24tsE8IlgioL

04.

PC 게임

스팀 데이터 파이프라인

스팀 데이터 ETL 과정 및 분석

메타크리틱 데이터 파이프라인

메타크리틱 데이터 ETL 과정 및 분석



4. PC 게임

01

스팀 크롤링 데이터

	all_rev_desc	all_rev_sum	developers	final_price	franchise	genre	meta_score	original_price	publishers	recent_rev_desc	recent_rev_sum	release	title	user_def_tags
0	- 79% of the 3,044 user reviews for this game ...	Mostly Positive	Valve	₩ 5,500	NaN	Action	NaN	₩ 5,500	Valve	- 90% of the 20 user reviews in the last 30 da...	Very Positive	1 Nov, 2000	Ricochet	Action,FPS,Multiplayer,Classic,First-Person,Sci...
1	- 95% of the 11,961 user reviews for this game...	Overwhelmingly Positive	Gearbox Software	₩ 5,500	Half-Life	Action	NaN	₩ 5,500	Valve	- 92% of the 166 user reviews in the last 30 d...	Very Positive	1 Nov, 1999	Half-Life: Opposing Force	FPS,Action,Classic,Sci-fi,Singleplayer,Shooter...
2	- 96% of the 58,981 user reviews for this game...	Overwhelmingly Positive	Valve	₩ 10,500	Half-Life	Action	96.0	₩ 10,500	Valve	- 96% of the 688 user reviews in the last 30 d...	Overwhelmingly Positive	8 Nov, 1998	Half-Life	FPS,Sci-fi,1990's,Singleplayer,Action,Multipla...
3	- 87% of the 3,195 user reviews for this game ...	Very Positive	Valve	₩ 5,500	NaN	Action	79.0	₩ 5,500	Valve	- 80% of the 21 user reviews in the last 30 da...	Very Positive	1 May, 2003	Day of Defeat	FPS,World War II,Multiplayer,Shooter,Action,Wa...
4	- 79% of the 1,529 user reviews for this game ...	Mostly Positive	Valve	₩ 5,500	NaN	Action	NaN	₩ 5,500	Valve	- 84% of the 13 user reviews in the last 30 da...	Very Positive	1 Jun, 2001	Deathmatch Classic	Action,FPS,Classic,Multiplayer,Shooter,First-P...
...
82389	- 83% of the 549 user reviews for this game ar...	Very Positive	Drastic Games	₩ 31,000	NaN	Action,Adventure	NaN	₩ 31,000	Noodlecake	- 80% of the 366 user reviews in the last 30 d...	Very Positive	11 May, 2022	Soundfall	Rhythm,Looter Shooter,Dungeon Crawler,Multip...
82390	- 90% of the 141 user reviews for this game ar...	Very Positive	Vinci Games	Free	NaN	Action,Casual,Simulation,Sports,Early Access	NaN	Free	Vinci Games	- 88% of the 34 user reviews in the last 30 da...	Very Positive	19 Apr, 2022	Blacktop Hoops	Early Access,Action,Basketball,VR,Arcade,Casua...
82391	NaN	NaN	Lion's Shade	₩ 3,300	NaN	NaN	NaN	₩ 3,300	Lion's Shade	NaN	NaN	11 Jun, 2022	Cliff Empire Soundtrack	NaN
82392	NaN	NaN	Zoryth Games	₩ 2,970	NaN	Casual,Indie	NaN	₩ 3,300	Zoryth Games	NaN	NaN	9 Jun, 2022	S.I.N. Unit - Farmhouse Map DLC	Indie,Casual,3D,First-Person,Multiplayer,Explor...
82393	NaN	NaN	Pixel House Games	Free To Play	NaN	Casual,Indie,Strategy	NaN	Free To Play	INDIECN	NaN	1 user reviews	9 Jun, 2022	Big Bia - 外星异能卡组	Strategy,Indie,Casual

82394 rows x 14 columns

4. PC 게임

01

스팀 데이터 전처리

일반인평가를 1~9등급으로 나눔(9등급이 높은 평가)

95 - 99% : Overwhelmingly Positive	9
94 - 80% : Very Positive	8
80 - 99% + few reviews: Positive	7
70 - 79% : Mostly Positive	6
40 - 69% : Mixed	5
20? - 39% : Mostly Negative	4
0 - 39% + few reviews: Negative	3
0 - 19% : Very Negative	2
0 - 19% + many reviews: Overwhelmingly Negative	1



title	user_def_tags	grade	reviews	release_year
Ricochet	Action,FPS,Multiplayer,Classic,First-Person,Sci-Fi	8.0	20.0	2000.0
Half-Life: Opposing Force	FPS,Action,Classic,Sci-Fi,Singleplayer,Shooter	8.0	166.0	1999.0
Half-Life	FPS,Sci-Fi,1990's,Singleplayer,Action,Multiplayer	9.0	688.0	1998.0
Day of Defeat	FPS,World War II,Multiplayer,Shooter,Action,War	8.0	21.0	2003.0
Deathmatch Classic	Action,FPS,Classic,Multiplayer,Shooter,First-Person	8.0	13.0	2001.0
...
Soundfall	Rhythm,Looter Shooter,Dungeon Crawler,Multiplayer	8.0	366.0	2022.0
Blacktop Hoops	Early Access,Action,Basketball,VR,Arcade,Casual	8.0	34.0	2022.0
Cliff Empire Soundtrack	...	NaN	NaN	2022.0
S.I.N. Unit - Farmhouse Map DLC	Indie,Casual,3D,First-Person,Multiplayer,Exploration	NaN	NaN	2022.0
Big Bia - 外星异能卡组	Strategy,Indie,Casual	NaN	NaN	2022.0



4. PC 게임

일반인 평가 vs. 전문가평가 상관관계

_c0	developers	final_price	franchise	genre	meta_score	original_price	publishers	title	user_def_tags	grade	reviews	release_year
0	Valve	5500.0	null	Action	null	5500.0	Valve	Ricochet	Action,FPS,Multip...	8.0	20.0	2000.0
1	Gearbox Software	5500.0	Half-Life	Action	null	5500.0	Valve	Half-Life: Oppositi... FPS,Action,Classi...	8.0	166.0	1999.0	
2	Valve	10500.0	Half-Life	Action	96.0	10500.0	Valve	Half-Life FPS,Sci-Fi,1990's...	9.0	688.0	1998.0	
3	Valve	5500.0	null	Action	98.0	5500.0	Valve	Day of Defeat FPS,World War II,...	8.0	21.0	2003.0	
4	Valve	5500.0	null	Action	null	5500.0	Valve	Deathmatch Classic Action,FPS,Classi...	8.0	13.0	2001.0	
5	Valve	5500.0	null	Action	null	5500.0	Valve	Team Fortress Cla... Action,FPS,Multip...	8.0	46.0	1999.0	
6	Gearbox Software	5500.0	Half-Life	Action	71.0	5500.0	Valve	Half-Life: Blue S... FPS,Action,Sci-Fi...	8.0	139.0	2001.0	
7	Valve	10500.0	null	Action	88.0	10500.0	Valve	Counter-Strike: S... Shooter,Action,FP...	9.0	708.0	2004.0	
8	Valve	10500.0	null	Action	80.0	10500.0	Valve	Day of Defeat: So... FPS,World War II,...	8.0	96.0	2010.0	
9	Valve	5500.0	null	Action	null	5500.0	Valve	Half-Life 2: Deat... Action,FPS,Multip...	8.0	63.0	2004.0	
10	Valve	10500.0	Half-Life	Action	96.0	10500.0	Valve	Half-Life 2 FPS,Action,Sci-Fi...	9.0	1044.0	2004.0	
11	Valve	10500.0	null	Action	90.0	10500.0	Valve	Portal Puzzle,Puzzle Pla...	9.0	1015.0	2007.0	
12	Valve	8500.0	Half-Life	Action	90.0	8500.0	Valve	Half-Life 2: Epis... FPS,Action,Sci-Fi...	9.0	268.0	2007.0	
13	Valve	0.0	null	Action,Free to Play	92.0	0.0	Valve	Team Fortress 2 Free to Play,Hero...	8.0	7696.0	2007.0	
14	Valve	10500.0	null	Action	89.0	10500.0	Valve	Left 4 Dead Zombies,Co-op,Mul...	9.0	321.0	2008.0	
15	Valve	10500.0	null	Action	90.0	10500.0	Valve	Portal Puzzle,Puzzle Pla...	9.0	1015.0	2007.0	
16	Valve	0.0	Dota Action,Free to Pl...	Action	90.0	0.0	Valve	Dota 2 Free to Play,MOBA...	6.0	21410.0	2013.0	
17	Valve	0.0	null	Action	77.0	0.0	Valve	Alien Swarm Free to Play,Co-o...	8.0	65.0	2010.0	
18	Valve	2200.0	null	Action,Adventure	95.0	2200.0	Valve	Portal 2 Platformer,Puzzle...	9.0	2232.0	2011.0	
19	Valve,Hidden Path...	0.0	null	Action,Free to Play	83.0	0.0	Valve	Counter-Strike: G... FPS,Shooter,Multip...	8.0	59726.0	2012.0	

meta.
grade
널값
제거

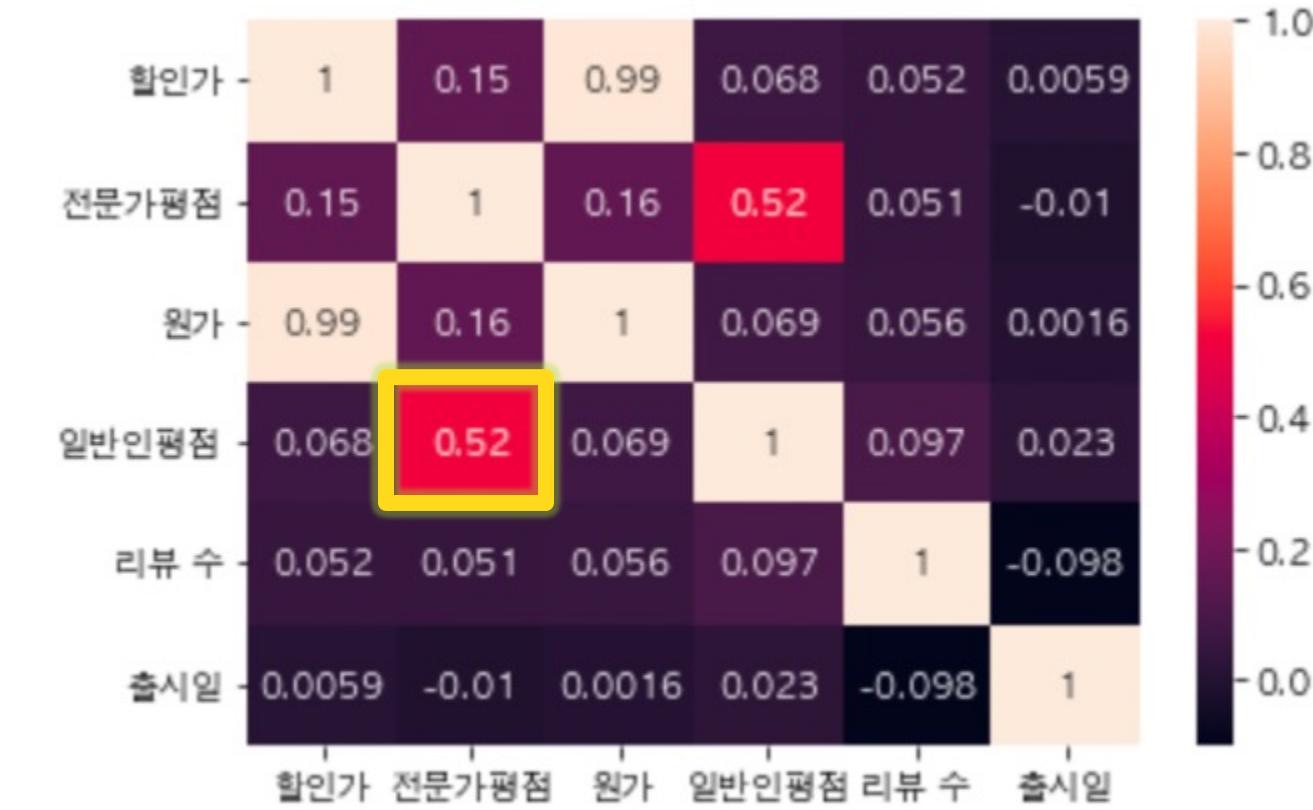
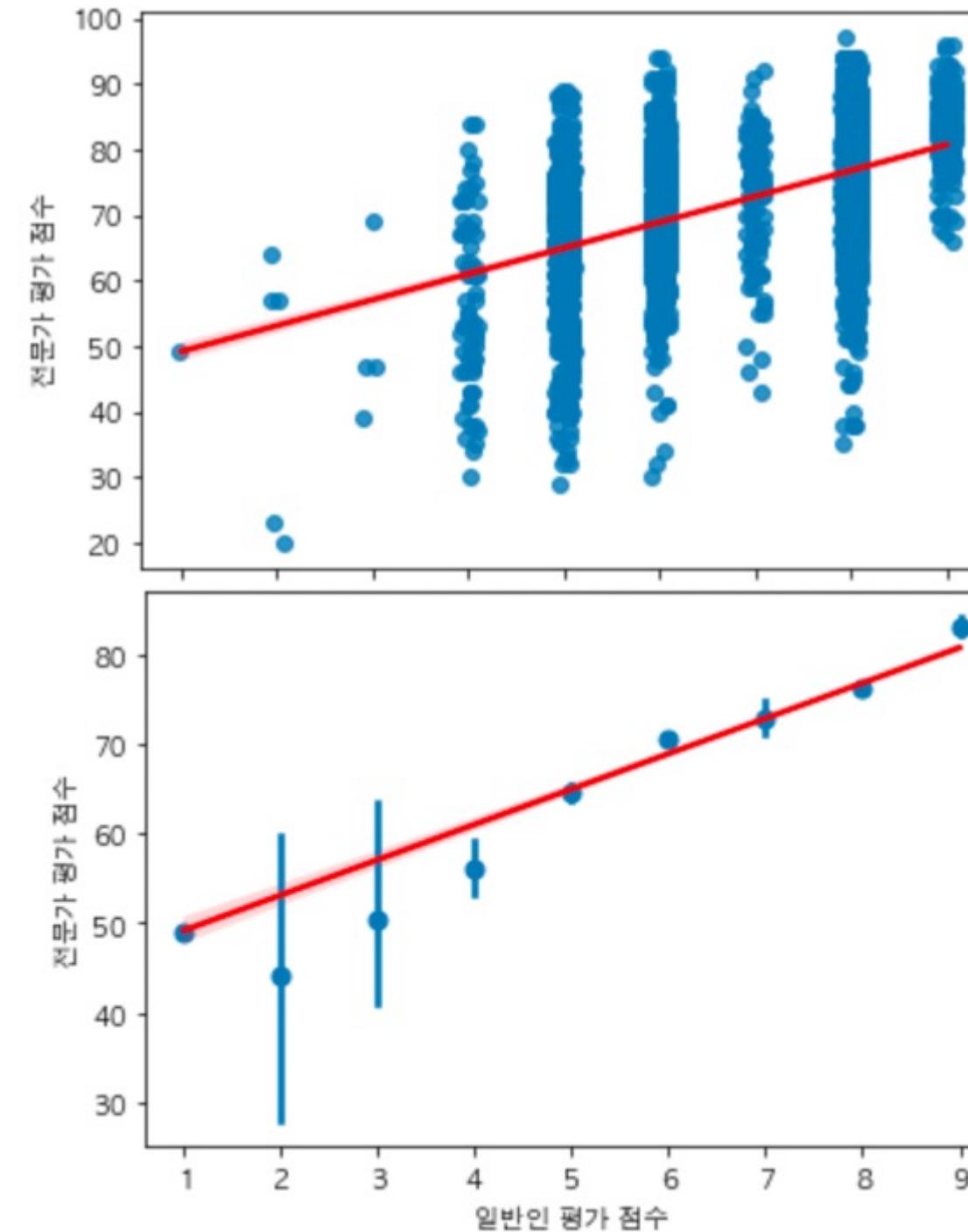
_c0	developers	final_price	franchise	genre	meta_score	original_price	publishers	title	user_def_tags	grade	reviews	release_year
2	Valve	10500.0	Half-Life	Action	96.0	10500.0	Valve	Half-Life FPS,Sci-Fi,1990's...	9.0	688.0	1998.0	
3	Valve	5500.0	null	Action	79.0	5500.0	Valve	Day of Defeat FPS,World War II,...	8.0	21.0	2003.0	
6	Gearbox Software	5500.0	Half-Life	Action	71.0	5500.0	Valve	Half-Life: Blue S... FPS,Action,Sci-Fi...	8.0	139.0	2001.0	
7	Valve	10500.0	null	Action	88.0	10500.0	Valve	Counter-Strike: S... Shooter,Action,FP...	9.0	708.0	2004.0	
8	Valve	10500.0	null	Action	80.0	10500.0	Valve	Day of Defeat: So... FPS,World War II,...	8.0	96.0	2010.0	
10	Valve	10500.0	Half-Life	Action	96.0	10500.0	Valve	Half-Life 2 FPS,Action,Sci-Fi...	9.0	1044.0	2004.0	
11	Valve	10500.0	null	Action	90.0	10500.0	Valve	Portal Puzzle,Puzzle Pla...	9.0	1015.0	2007.0	
12	Valve	8500.0	Half-Life	Action	90.0	8500.0	Valve	Half-Life 2: Epis... FPS,Action,Sci-Fi...	9.0	268.0	2007.0	
13	Valve	0.0	null	Action,Free to Play	92.0	0.0	Valve	Team Fortress 2 Free to Play,Hero...	8.0	7696.0	2007.0	
14	Valve	10500.0	null	Action	89.0	10500.0	Valve	Left 4 Dead Zombies,Co-op,Mul...	9.0	321.0	2008.0	
15	Valve	10500.0	null	Action	90.0	10500.0	Valve	Portal Puzzle,Puzzle Pla...	9.0	1015.0	2007.0	
16	Valve	0.0	Dota Action,Free to Pl...	Action	90.0	0.0	Valve	Dota 2 Free to Play,MOBA...	6.0	21410.0	2013.0	
17	Valve	0.0	null	Action	77.0	0.0	Valve	Alien Swarm Free to Play,Co-o...	8.0	65.0	2010.0	
18	Valve	2200.0	null	Action,Adventure	95.0	2200.0	Valve	Portal 2 Platformer,Puzzle...	9.0	2232.0	2011.0	
19	Valve,Hidden Path...	0.0	null	Action,Free to Play	83.0	0.0	Valve	Counter-Strike: G... FPS,Shooter,Multip...	8.0	59726.0	2012.0	
21	Tripwire Interactive	5500.0	Tripwire Interactive	Action	81.0	5500.0	Tripwire Interactive	Red Orchestra: Os... World War II,Acti...	8.0	null	2006.0	
22	Tripwire Interactive	21000.0	Tripwire Interactive	Action	72.0	21000.0	Tripwire Interactive	Killing Floor FPS,Zombies,Co-op...	8.0	61.0	2009.0	
24	Ritual Entertainment	10500.0	null	Action	75.0	10500.0	Ritual Entertainment	SIN Episodes: Em... Action,FPS,Cyber...	8.0	520.0	2006.0	
25	Introversion Soft...	10500.0	null	Indie,Strategy	75.0	10500.0	Introversion Soft...	Uplink Hacking,Indie,Strateg...	8.0	1521.0	2006.0	
26	Introversion Soft...	10500.0	null	Indie,Strategy	84.0	10500.0	Introversion Soft...	Darwinia Retro,Atmospheric...	6.0	624.0	2005.0	

Group by
grade

grade	mean(meta_score)	count(meta_score)
0.0	null	2
1.0	49.0	1
2.0	44.2	5
3.0	50.5	4
4.0	56.18571428571428	70
5.0	64.59724137931035	725
6.0	70.59476309226933	802
7.0	72.86666666666666	105
8.0	76.22237697307335	2154
9.0	83.0592105263158	152

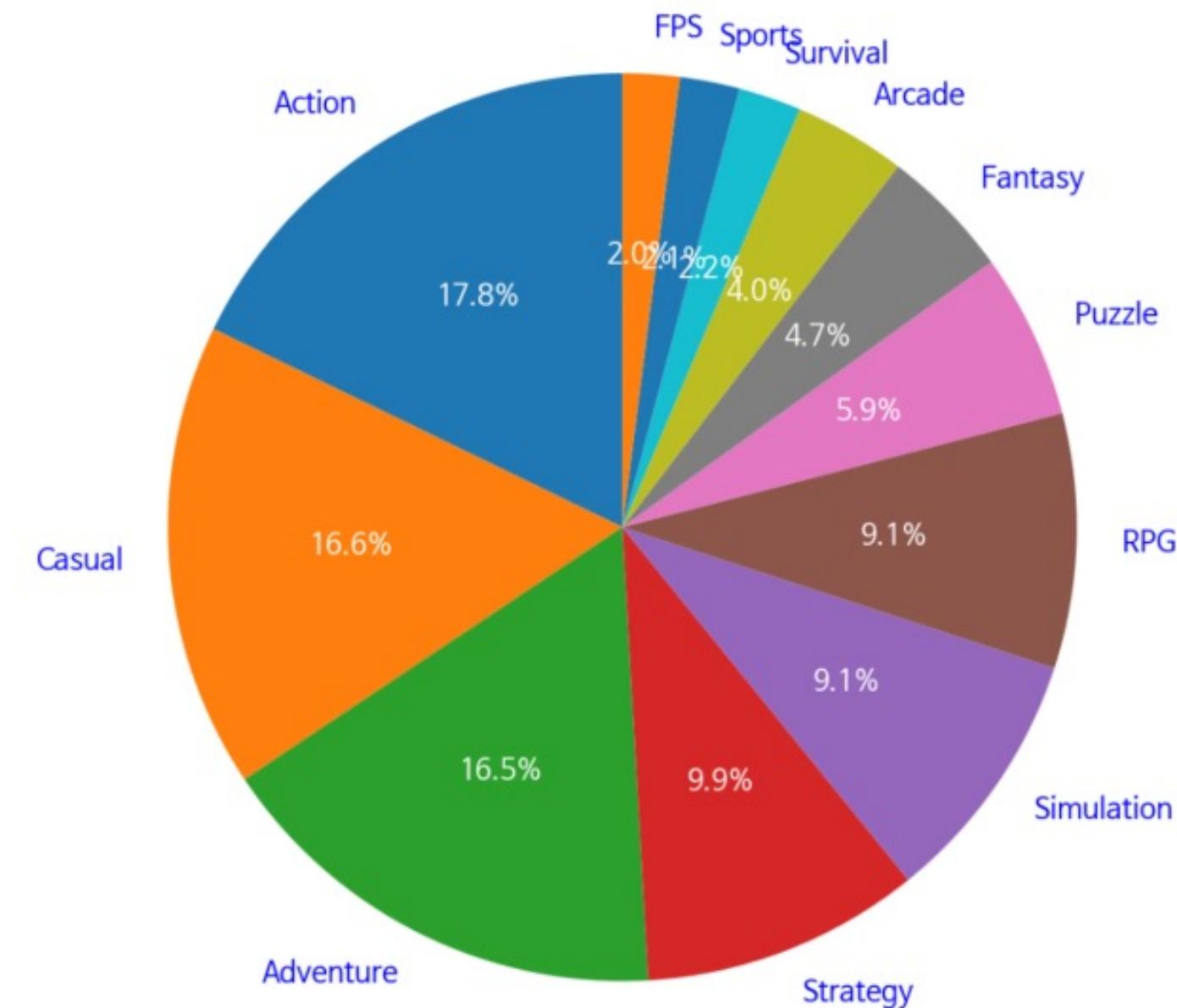
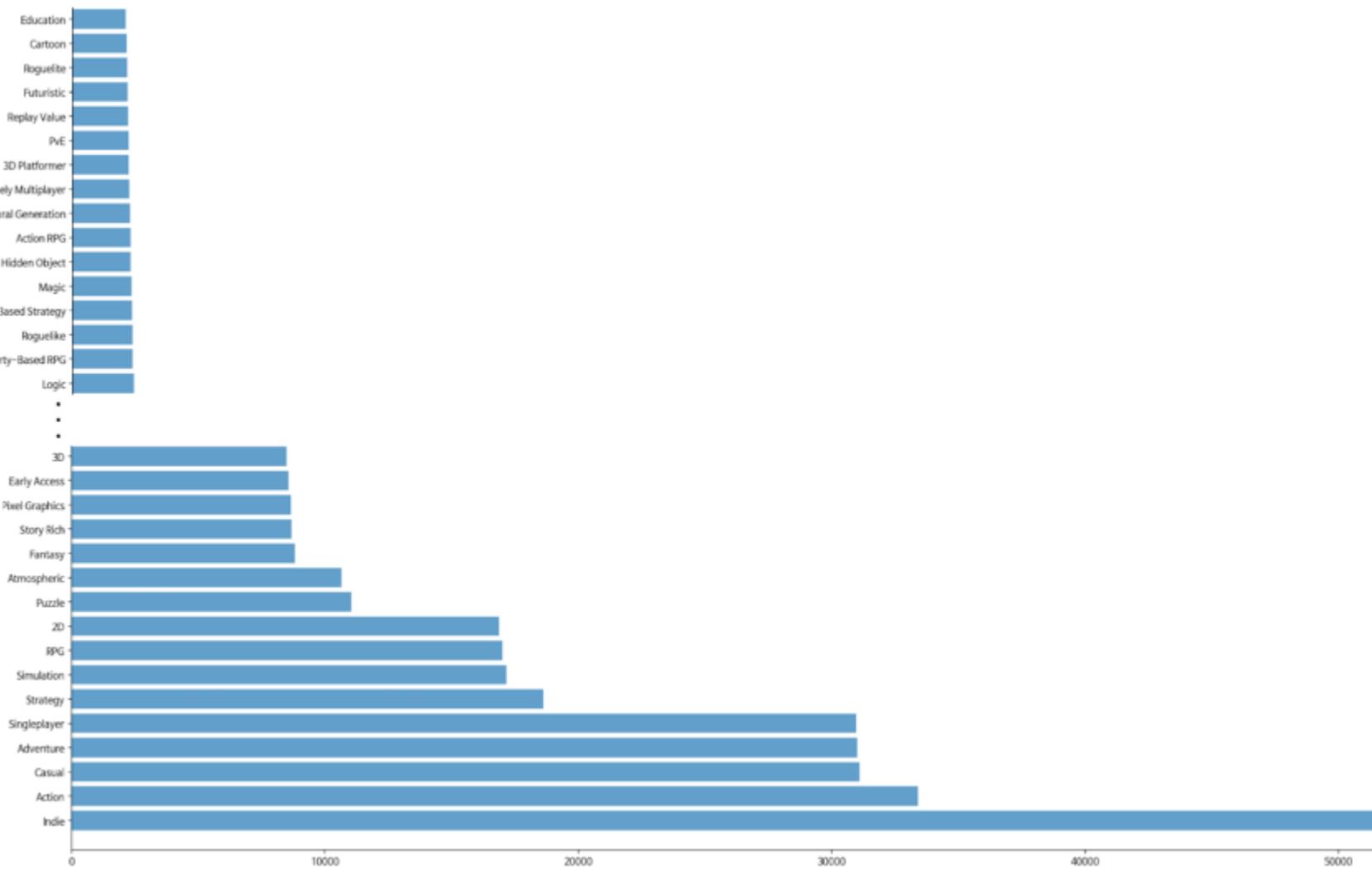
4. PC 게임

일반인 평가 vs. 전문가평가 상관관계



4. PC 게임

전체 게임 카테고리 분석 (TOP 100)



카테고리 12가지 장르로 압축



시대별 카테고리 분석

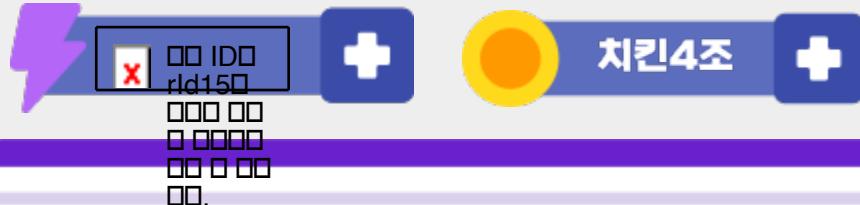
_c0	developers	final_price	franchise	genre	meta_score	original_price	publishers	title	user_def_tags	grade	reviews	release_year
0	Valve	5500.0	null	Action	null	5500.0	Valve	Ricochet	Action,FPS,Multip...	8.0	20.0	2000.0
1	Gearbox Software	5500.0	Half-Life	Action	null	5500.0	Valve	Half-Life: Oppos... FPS,Action,Classi...	8.0	166.0	1999.0	
2	Valve	10500.0	Half-Life	Action	96.0	10500.0	Valve	Half-Life FPS,Sci-Fi,1990's...	9.0	688.0	1998.0	
3	Valve	5500.0	null	Action	79.0	5500.0	Valve	Day of Defeat FPS,World War II,...	8.0	21.0	2003.0	
4	Valve	5500.0	null	Action	null	5500.0	Valve	Deathmatch Classic Action,FPS,Classi...	8.0	13.0	2001.0	
5	Valve	5500.0	null	Action	null	5500.0	Valve	Team Fortress Cl... Action,FPS,Multip...	8.0	46.0	1999.0	
6	Gearbox Software	5500.0	Half-Life	Action	71.0	5500.0	Valve	Half-Life: Blue S... FPS,Action,Sci-Fi...	8.0	139.0	2001.0	
7	Valve	10500.0	null	Action	88.0	10500.0	Valve	Counter-Strike: S... Shooter,Action,FP...	9.0	708.0	2004.0	
8	Valve	10500.0	null	Action	80.0	10500.0	Valve	Day of Defeat: So... FPS,World War II,...	8.0	96.0	2010.0	
9	Valve	5500.0	null	Action	null	5500.0	Valve	Half-Life 2: Deat... Action,FPS,Multip...	8.0	63.0	2004.0	
10	Valve	10500.0	Half-Life	Action	96.0	10500.0	Valve	Half-Life 2 FPS,Action,Sci-Fi...	9.0	1044.0	2004.0	
11	Valve	10500.0	null	Action	90.0	10500.0	Valve	Portal Puzzle,Puzzle Pla...	9.0	1015.0	2007.0	
12	Valve	8500.0	Half-Life	Action	90.0	8500.0	Valve	Half-Life 2: Epis... FPS,Action,Sci-Fi...	9.0	268.0	2007.0	
13	Valve	0.0	null	Action,Free to Play	92.0	0.0	Valve	Team Fortress 2 Free to Play,Hero...	8.0	7606.0	2007.0	
14	Valve	10500.0	null	Action	89.0	10500.0	Valve	Left 4 Dead Zombies,Co-op,Mul...	9.0	321.0	2008.0	
15	Valve	10500.0	null	Action	90.0	10500.0	Valve	Portal Puzzle,Puzzle Pla...	9.0	1015.0	2007.0	
16	Valve	0.0	Dota Action,Free to Pl...	Action	90.0	0.0	Valve	Dota 2 Free to Play,MOBA,...	6.0	21410.0	2013.0	
17	Valve	0.0	null	Action	77.0	0.0	Valve	Alien Swarm Free to Play,Co-o...	8.0	65.0	2010.0	
18	Valve	2200.0	null	Action,Adventure	95.0	0.0	Valve	Portal 2 Platformer,Puzzle...	9.0	2232.0	2011.0	
19	Valve,Hidden Path...	0.0	null	Action,Free to Play	83.0	0.0	Valve	Counter-Strike: G... FPS,Shooter,Multip...	8.0	59726.0	2012.0	

Year
구간
설정

	user_def_tags	release_year	
Adventure,Point &...	2021.0	Tactical,Realisti...	2018.0
Free to Play,Indi...	2020.0	Action,Parkour,In...	2017.0
Sports,Free to Pl...	2020.0	Surreal,Nudity,Ad...	2016.0
Free to Play,Indi...	2020.0	Casual,Simulation...	2016.0
Adventure,RPG,Ind...	2020.0	Puzzle Platformer...	2016.0
Action,Indie,Stra...	2020.0	Pixel Graphics,Ad...	2016.0
Vehicular Combat,...	2020.0	Strategy,Board Ga...	2015.0
Action,Gore,FPS,C...	2020.0	Free to Play,Stra...	2018.0
Puzzle,Sokoban,Di...	2020.0	Multiplayer,Racin...	2016.0
Multiplayer,Physi...	2021.0	Action,Mythology,...	2015.0
Multiplayer,FPS,A...	2021.0	Puzzle,Exploratio...	2016.0
Adventure,Indie,C...	2020.0	Open World Surviv...	2016.0
Indie,Action,Poli...	2020.0	Free to Play,Mult...	2015.0
RPG,Strategy,West...	2020.0	Massively Multipl...	2017.0
RPG,Adventure,Ind...	2020.0	Action,Indie,RPG,...	2016.0
Action,Indie,Adve...	2020.0	Early Access,Acti...	2015.0
Action,Adventure,...	2021.0	Utilities,Audio P...	2017.0
Space,Sandbox,MMO...	2020.0	Action RPG,Dark F...	2016.0
Action,RPG,Action...	2021.0	Space,Simulation,...	2015.0
Indie,Simulation	2022.0	Music,Education,S...	2016.0

4. PC 게임

	1990-1994	1995-1999	2000-2004	2005-2009	2010-2014	2015-2019	2020-2022
0	Retro	Classic	Singleplayer	Singleplayer	Indie	Indie	Indie
1	Adventure	Action	Action	Action	Action	Action	Singleplayer
2	Classic	Singleplayer	Classic	Adventure	Singleplayer	Adventure	Casual
3	Action	Adventure	Strategy	Strategy	Adventure	Casual	Action
4	1990's	1990's	Adventure	Casual	Casual	Strategy	Adventure
5	Singleplayer	Retro	Atmospheric	Indie	Strategy	Singleplayer	2D
6	Simulation	Sci-fi	Multiplayer	Multiplayer	RPG	Simulation	Strategy
7	Pixel Graphics	Strategy	Simulation	Simulation	Simulation	RPG	RPG
8	2D	First-Person	Sci-fi	Atmospheric	2D	2D	3D
9	Strategy	Shooter	Great Soundtrack	Puzzle	Multiplayer	Puzzle	Simulation
10	Point & Click	FPS	Shooter	RPG	Great Soundtrack	Early Access	Colorful
11	Shooter	Point & Click	First-Person	Classic	Puzzle	Atmospheric	Atmospheric
12	RPG	Casual	RTS	Great Soundtrack	Atmospheric	VR	Puzzle
13	First-Person	Atmospheric	Casual	Shooter	Fantasy	Free to Play	Pixel Graphics
14	Old School	Great Soundtrack	Story Rich	Story Rich	Platformer	Great Soundtrack	Fantasy
15	Flight	Simulation	FPS	Sci-fi	Story Rich	Multiplayer	Cute
16	Sci-fi	RPG	RPG	First-Person	Co-op	Story Rich	Story Rich
17	Puzzle	Puzzle	Indie	FPS	Sci-fi	Fantasy	Exploration
18	FPS	Story Rich	Point & Click	Open World	Shooter	Pixel Graphics	Early Access
19	Funny	Multiplayer	Fantasy	Co-op	Female Protagonist	Difficult	Arcade
20	Platformer	Comedy	Retro	Fantasy	Difficult	Anime	First-Person
21	Story Rich	2D	Third Person	RTS	Funny	Violent	Multiplayer
22	Atmospheric	Cult Classic	Puzzle	Point & Click	Open World	Arcade	Action-Adventure
23	FMV	FMV	2D	Third Person	Arcade	Retro	Relaxing
24	Combat	Indie	Cult Classic	Female Protagonist	First-Person	Horror	Shooter
25	Comedy	Funny	Tactical	Funny	Point & Click	Platformer	Family Friendly
26	Casual	Horror	Old School	2D	Retro	Funny	Funny
27	Exploration	Fantasy	Funny	Historical	Turn-Based	Sports	Platformer
28	Historical	Family Friendly	Open World	Comedy	Comedy	Cute	Sci-fi
29	War	Gore	War	Horror	Horror	First-Person	2D Platformer



4. PC 게임

점수가 높거나 낮은 게임의 카테고리

_c0	developers	final_price	franchise	genre	meta_score	original_price	publishers	title	user_def_tags	grade	reviews	release_year
0	Valve	5500.01	null	Action	null	5500.01	Valve	Ricochet	Action,FPS,Multip...	8.0	20.01	2000.01
1	Gearbox Software	5500.01	Half-Life	Action	null	5500.01	Valve	Half-Life: Oppositi...	Action,FPS,Classi...	8.0	166.01	1999.01
2	Valve	10500.01	Half-Life	Action	96.01	10500.01	Valve	Half-Life FPS,Sci-Fi,1990's...	9.0	688.01	1998.01	
3	Valve	5500.01	null	Action	79.01	5500.01	Valve	Day of Defeat FPS,World War II,...	Action,FPS,Classi...	8.0	21.01	2003.01
4	Valve	5500.01	null	Action	null	5500.01	Valve	Deathmatch Classic Action,FPS,Classi...	Action,FPS,Multip...	8.0	13.01	2001.01
5	Valve	5500.01	null	Action	null	5500.01	Valve	Team Fortress Clas...	Action,FPS,Multip...	8.0	46.01	1999.01
6	Gearbox Software	5500.01	Half-Life	Action	71.01	5500.01	Valve	Half-Life: Blue Shift Action,Sci-Fi,...	Action,FPS,Classi...	8.0	139.01	2001.01
7	Valve	10500.01	null	Action	88.01	10500.01	Valve	Counter-Strike: S...	[Shooter,Action,FP...	9.0	708.01	2004.01
8	Valve	10500.01	null	Action	80.01	10500.01	Valve	Day of Defeat: So...	[FPS,World War II,...]	8.0	96.01	2010.01
9	Valve	5500.01	null	Action	null	5500.01	Valve	Half-Life 2: Death...	Action,FPS,Multip...	8.0	63.01	2004.01
10	Valve	10500.01	Half-Life	Action	96.01	10500.01	Valve	Half-Life 2 FPS,Action,Sci-Fi,...	Action,FPS,Classi...	9.0	1044.01	2004.01
11	Valve	10500.01	null	Action	90.01	10500.01	Valve	Portal Puzzle,Puzzl...	Puzzle,Action,Sci-Fi...	9.0	1015.01	2007.01
12	Valve	8500.01	Half-Life	Action	90.01	8500.01	Valve	Half-Life 2: Epis...	[FPS,Action,Sci-Fi,...]	9.0	268.01	2007.01
13	Valve	0.01	null	Action,Free to Play	92.01	0.01	Valve	Team Fortress 2 Free to Play,Hero...	Hero,Action,FP...	8.0	7696.01	2007.01
14	Valve	10500.01	null	Action	89.01	10500.01	Valve	Left 4 Dead Zombies,Co-op,Mul...	Zombie,Co-op,Multipl...	9.0	321.01	2008.01
15	Valve	10500.01	null	Action	90.01	10500.01	Valve	Portal Puzzle,Puzzl...	Puzzle,Action,Sci-Fi...	9.0	1015.01	2007.01
16	Valve	0.01	Data Action,Free to Play ...	Action	90.01	0.01	Valve	DotA 2 Free to Play,MOBA...	MOBA,Action,Sci-Fi...	6.0	21410.01	2013.01
17	Valve	0.01	null	Action	77.01	0.01	Valve	Alien Swarm Free to Play,Co-o...	Co-op,Action,Sci-Fi...	8.0	65.01	2010.01
18	Valve	2200.01	null	Action,Adventure	95.01	0.01	Valve	Portal 2 Platformer,Puzzl...	Platformer,Action,Sci-Fi...	9.0	2232.01	2011.01
19	Valve,Hidden Path,...	0.01	null	Action,Free to Play	83.01	0.01	Valve	Counter-Strike: Global Offensive FPS,Shooter,Mult...	Shooter,Action,Sci-Fi...	8.0	59726.01	2012.01

user_def_tags grade	user_def_tags grade
FPS,Sci-fi,1990's... 9.0	Racing 3.0
Shooter,Action,FP... 9.0	Indie,Adventure,P... 2.0
FPS,Action,Sci-fi... 9.0	Indie,Adventure,P... 2.0
Puzzle,Puzzle Pla... 9.0	FPS,Sci-fi,Flight... 3.0
FPS,Action,Sci-fi... 9.0	FPS,Sci-fi,Flight... 3.0
Zombies,Co-op,Mul... 9.0	Action,Indie 3.0
Puzzle,Puzzle Pla... 9.0	Simulation,Strate... 2.0
Platformer,Puzzl... 9.0	Psychological Hor... 2.0
FPS,Classic,Actio... 9.0	Simulation,Strate... 2.0
RPG,Vampire,Cult ... 9.0	Space,Simulation,... 1.0
Racing,Combat Rac... 9.0	Strategy,Action,M... 2.0
Tower Defense,Zom... 9.0	Simulation,Space,... 1.0
Sandbox,Multiplay... 9.0	Action,Indie,Earl... 2.0
Zombies,Co-op,Mul... 9.0	Strategy,4X,Indie... 2.0
Action,FPS,Tactic... 9.0	Action,RPG,Indie 3.0
Action,Multiplaye... 9.0	Simulation,Buildi... 2.0
Turn-Based Strate... 9.0	Simulation,Buildi... 2.0
Warhammer 40K,Str... 9.0	Strategy,Blood,Go... 2.0
Warhammer 40K,Str... 9.0	Casual,Farming Si... 2.0
Great Soundtrack,... 9.0	Simulation,Manage... 2.0

	good game	bad game
0	Great Soundtrack	VR
1	Funny	Management
2	Anime	Realistic
3	Cute	Building
4	Pixel Graphics	Sports
5	Visual Novel	Hidden Object
6	Free to Play	PvP
7	Multiple Endings	Arcade
8	Retro	Racing
9	Choices Matter	Economy
10	Moddable	Zombies
11	Mystery	Driving
12	Dark	Action-Adventure
13	Short	Space
14	Physics	Tactical
15	Post-apocalyptic	Massively Multiplayer
16	Dating Sim	Survival Horror
17	Memes	Online Co-Op
18	Romance	Futuristic
19	Classic	Automobile Sim

4. PC 게임

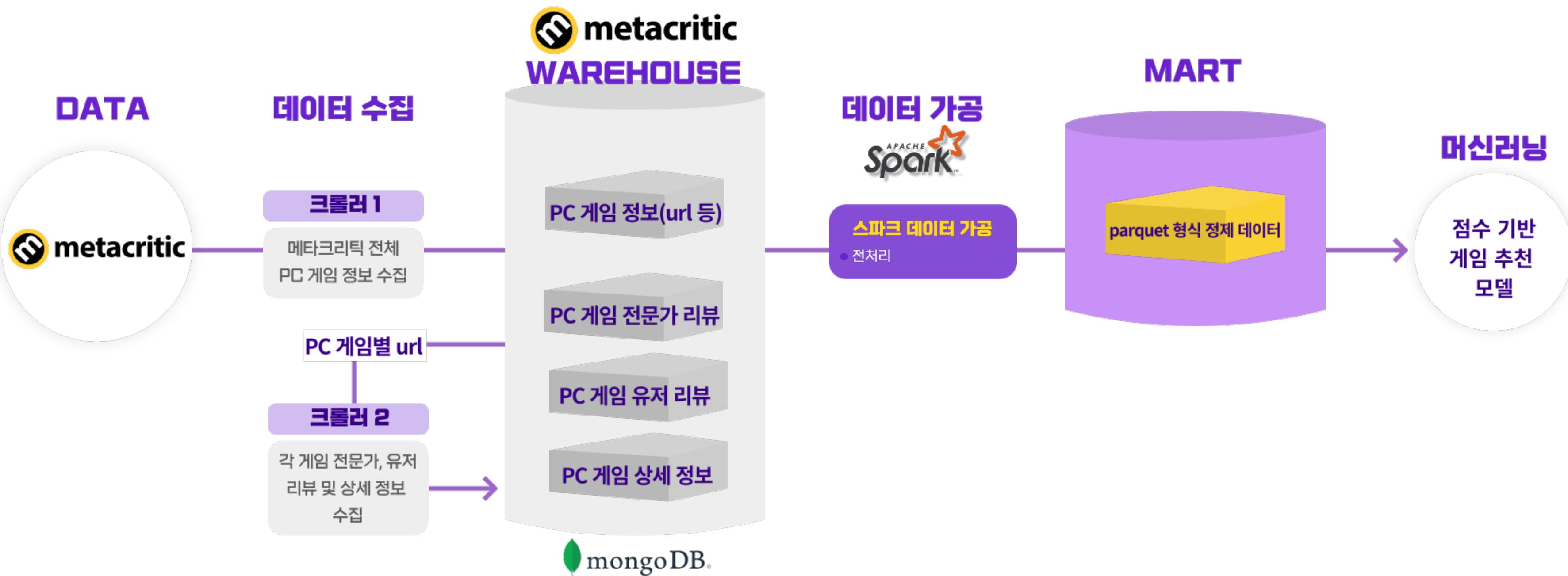
01

메타크리틱 게임 정보

- 대상 : 메타크리틱 pc 게임 중 메타크리틱 점수가 존재하는 게임
- 기간 : 1998년 1월 1일 ~ 2022년 6월 1일

4. PC 게임

메타크리틱 파이프라인



4. PC 게임

메타크리틱 1차 크롤링

The screenshot shows the Metacritic website's PC Games page. It lists several upcoming titles, each with a thumbnail, title, platform (PC), and release date. A purple arrow points from this screenshot to the '크롤링' section below.

Game	Platform	Release Date
Chesskoban Cyber	PC	June 19, 2022
The Asmodian Princesses and the Witch in the Forest	PC	June 18, 2022
Fly By Knight	PC	June 18, 2022
Flowers: Le volume sur hiver	PC	June 18, 2022
Green Diver (Dnovel)	PC	June 18, 2022
DEEP 8	PC	June 17, 2022

크롤링

raw_game_info_row : 71291			
	title	href	metascore
	Police Shootout	police-shootout	tbd 2022-06-01 00:00:00
	Failed Ninja Academy	failed-ninja-academy	tbd 2022-06-01 00:00:00
	SILT	silt	76 2022-06-01 00:00:00
	Deathrun TV	deathrun-tv	tbd 2022-06
	Hack and Slime	hack-and-slime	tbd 2022-06
	Dark Dimension	dark-dimension	tbd 2022-06
	Fishing Paradiso	fishing-paradiso	tbd 2022-06-01 00:00:00
	Exile Origin	exile-origin	tbd 2022-06-01 00:00:00
	PolyShot (TEKNOSOFT)	polyshot-teknosoft	tbd 2022-06-01 00:00:00
	Bushi	bushi	tbd 2022-06-01 00:00:00

only showing top 10 rows

멀티프로세싱 사용하여
병렬처리

metascore 점수가
있는(!=tbd) 게임의 href 추출

href
+-----+-----+
nra-varmint-hunter
rise-to-hell-retribution
implants-weapons-of-mass-destruction
arkania-blade-of-destiny-revised
generals
ai-in-the-dark-illumination
leisure-suit-larry-box-office-bust
citadels
infestation-survivor-stories-the-war-z
winter-sports-2012-feel-the-spirit
+-----+-----+

```
if __name__=='__main__':
    with Pool(processes=multiprocessing.cpu_count()) as pool:
        data = pool.map(get_game_info_and_reviews, fail_page)
```

4. PC 게임

메타크리틱 2차 크롤링

GRAND THEFT AUTO V PC

Rockstar Games | Release Date: Apr 13, 2015
Also On: PlayStation 3, PlayStation 4, PlayStation 5, Xbox 360, Xbox One, Xbox Series X

Summary **Critic Reviews** **User Reviews** **Details & Credits** **Trailers & Videos**

96 Metascore Universal acclaim based on 57 Critic Reviews

Critic score distribution: Positive: 56 Mixed: 1 Negative: 0

Buy Now **Buy on amazon.com**

Critic score Publication By date

DarkStation Jul 2, 2015
If you hadn't had a chance to play it yet at all though, and your computer can handle it, the PC version does definitely feel like the definitive version.
[All this publication's reviews](#) | [Read full review](#)

CD-Action Jun 29, 2015
Rockstar promised that the PC edition of GTA V would be the best way to experience this breathtaking game, and that's exactly what it is. [06/2015, p.48]
[All this publication's reviews](#)

GameOver.gr Jun 2, 2015
Beautiful, witty, rude and captivating Rockstar's GTA V PC release is a must play title. Entangling open world, a great variety in missions, tons of online content and a port that respects the game itself and PC gamers alike.
[All this publication's reviews](#) | [Read full review](#)

LEVEL (Czech Republic) May 12, 2015
Even more free, prettier, smoother and in every aspect a more comfortable Los Santos than we all remember. Old song for the extra money you say? C'mon, for game like this you can never pay too much. [Issue#252]
[All this publication's reviews](#)

ImpulseGamer May 11, 2015
Ultimately you will get everything you expected from GTA 5 for PC; a tremendous game with many months of playability and the potential to only get better.
[All this publication's reviews](#) | [Read full review](#)

Gamer.no May 4, 2015
Quotation forthcoming.
[All this publication's reviews](#) | [Read full review](#)

유저 리뷰

raw_user_review_row : 448448						
title	name grade	review	date	total_ups	total_thumbs	
Aragami 2	Kilio	1 Meh story, weak a...	2022-03-29T00:00:00	0	0	
Aragami 2	Oshiri_Maeda	7 Nice game for cou...	2022-01-26T00:00:00	0	0	
Aragami 2	3t3rnal	4 I have to say tha...	2022-01-09T00:00:00	2	2	
Aragami 2	Furuya65	3 das spiel ist ein...	2021-10-18T00:00:00	1	1	
Aragami 2	gamezn_ru	8 Каждое ваше дейст...	2021-10-03T00:00:00	0	1	
Aragami 2	Nobilis1984	3 Bigger but not be...	2021-09-19T00:00:00	2	4	
Aragami 2	Joe_Friday	0 A technical mess ...	2021-09-18T00:00:00	3	4	
Aragami 2	Ocesone	3 El juego está muy...	2021-09-18T00:00:00	2	2	
Iwakura_Kiwami	chin7771	Al Game is 10/10 but...	2022-06-06T00:00:00	0	0	
raw_critic_review_row : 110728						
title	source	link	grade	review	date	
Aragami 2	GameGrin https://www.gamegrin...	90 Overall, Aragami ...	2021-09-15T00:00:00			
Aragami 2	GamingBolt https://gamingbolt...	80 With improved tra...	2021-09-16T00:00:00			
Aragami 2	Multiplayer.it https://multiplay...	80 Aragami 2 is a gr...	2021-09-15T00:00:00			
Aragami 2	Merlin'in Kazanı ... http://www.merlin...	78 Aragami 2, in add...	2021-09-15T00:00:00			
Aragami 2	Everyeye.it https://www.every...	75 With Aragami 2, t...	2021-09-19T00:00:00			
Aragami 2	Wccftech https://wccftech....	75 Aragami 2 is anot...	2021-09-15T00:00:00			
Aragami 2	TheSixthAxis https://www.thesix...	70 Aragami 2 is a hu...	2021-09-29T00:00:00			
Aragami 2	Areajugones https://areajugon...	70 Even if I know th...	2021-09-21T00:00:00			
Aragami 2	Gamer Escape https://gamerescap...	70 For all my compla...	2021-09-20T00:00:00			
Aragami 2	Jeuxvideo.com https://www.jeuxv...	70 With its simple a...	2021-09-15T00:00:00			
raw_game_detail_row : 6917						
title	critic_score	user_score	genres	developer	summary	
Aragami 2	65	5.3 ['Action Adventure'] Aragami 2 is a th...				
Synth Riders	83	8.0 ['Action', 'Rhyth...'] ['Retrowave VR,Kl... Get your groove o...				
Total War: Three ...	78	tbd ['Strategy', 'Rea...'] [('Creative Assemb... Determine the fat...				
Necronator: Dead ...	66	tbd ['Strategy', 'Rea...'] [('Toge Productions') Necronator: Dead ...				
SILT	73	7.2 ['Action Adventure'] ['Spiral Circus'] Silt is a puzzle...				
Yakuza Kiwami	80	7.9 ['Action Adventure'] ['Ryu ga Gotoku S... In 2005, a legend...				
Journey For Elysium	63	tbd ['Adventure', '3D...'] ['Mantis'] Journey for Elys...				
Severed Steel	74	6.8 ['Action', 'First...'] [('Greylock Studio') Severed Steel is ...				
Blind Drive	87	6.6 ['Adventure', 'Ge...'] [('Lo-Fi People') Blind Drive is an...				
Unexplored 2: The...	69	tbd ['Role-Playing', ...] ['Big Sugar Films... Unexplored 2: The...				

전문가
리뷰

기타 정보

DB 저장

4. PC 게임

데이터 전처리

raw_critic_review_row : 110728					
title	source	link	grade	review	date
Aragami 2	GameGrin https://www.gamegrin.com/reviews/aragami-2/	90	Overall, Aragami ...	2021-09-15T00:00:00	
Aragami 2	GamingBolt https://gamingbolt.com/reviews/aragami-2-review	80	With improved tra...	2021-09-16T00:00:00	
Aragami 2	Multiplayer.it https://multiplayer.it/reviews/aragami-2/	80	Aragami 2 is a gr...	2021-09-15T00:00:00	
Aragami 2	Merlin'in Kazanı http://www.merlinin-kazani.com/reviews/aragami-2/	78	Aragami 2, in add...	2021-09-15T00:00:00	
Aragami 2	Everyeye.it https://www.everyeye.it/reviews/aragami-2/	75	With Aragami 2, t...	2021-09-19T00:00:00	
Aragami 2	Wccftech https://wccftech.com/reviews/aragami-2-review	75	Aragami 2 is anot...	2021-09-15T00:00:00	
Aragami 2	TheSixthAxis https://www.thesixthaxis.com/reviews/aragami-2-review	70	Aragami 2 is a hu...	2021-09-29T00:00:00	
Aragami 2	Areajugones https://areajugones.com/reviews/aragami-2/	70	Even if I know th...	2021-09-21T00:00:00	
Aragami 2	Gamer Escape https://gamerescapes.com/reviews/aragami-2/	70	For all my compla...	2021-09-20T00:00:00	
Aragami 2	Jeuxvideo.com https://www.jeuxvideo.com/reviews/aragami-2/	70	With its simple a...	2021-09-15T00:00:00	

raw_user_review_row : 448448						
title	name	grade	review	date	total_ups	total_thumbs
Aragami 2	Kilio	1	Meh story, weak a...	2022-03-29T00:00:00	0	0
Aragami 2	Oshiri_Maeda	7	Nice game for cou...	2022-01-26T00:00:00	0	0
Aragami 2	3t3rnal	4	I have to say tha...	2022-01-09T00:00:00	2	2
Aragami 2	Furuya65	3	das spiel ist ein...	2021-10-18T00:00:00	1	1
Aragami 2	gamezn_ru	8	Каждое ваше действ...	2021-10-03T00:00:00	0	1
Aragami 2	Nobilis1984	3	Bigger but not be...	2021-09-19T00:00:00	2	4
Aragami 2	Joe_Friday	0	A technical mess ...	2021-09-18T00:00:00	3	4
Aragami 2	Ocesone	3	El juego está muy...	2021-09-18T00:00:00	2	2
Yakuza Kiwami	shin777	0	Game is 10/10, bu...	2022-06-06T00:00:00	0	0
Yakuza Kiwami	d3uzz	9	Yakuza Kiwami - o...	2022-05-28T00:00:00	0	0

문자열 게임제목과 유저명
수치형으로 변경

title	gameId	name	userId
Tales of Arise	0	Hackerman42069	0
Coromon	1	henriyouu	1
Masquerada: Songs...	2	Bazoslav	2
Ember	3	Umadbro	3
Planet Zoo	4	Astralniy	
Dungeons 2	5	Kerber	
Human: Fall Flat	6	ArmyRetGuy	6
Californium	7	Hurrikhan	7
Volume	8	Squallish	8
The Elder Scrolls...	9	Metacritic376	9

Parquet 형식으로
정제된 데이터 저장

userId	gameId	rating
1061	0	70.0
156004	0	40.0
159061	0	40.0
81073	0	100.0
19036	0	60.0
159761	0	80.0
159769	0	100.0
158155	0	80.0
105	0	90.0
81779	0	90.0

4. PC 게임

메타크리틱 게임 추천 모델

테스터 데이터 예측

userId	gameId	rating	prediction
613	3749	70.0	57.55703
847	3749	70.0	74.57218
847	5803	70.0	62.239864
1618	2366	70.0	73.240005
5110	5803	100.0	78.55769

RMSE, MAE 값

RMSE: 9.057132696335273

MAE: 4.174443947454175

userId	recommendations
31	[{0, 0.0}, {10, 0...}
34	[{4025, 283.88574...}
53	[{4546, 535.662}, ...]
65	[{3682, 121.82384...}
78	[{4025, 345.741}, ...]

gameId	recommendations
26	[{17954, 4301.002...}
27	[{54743, 12489.11...}
28	[{118278, 8311.11...}
31	[{118278, 6721.86...}
34	[{118278, 7005.82...}]

유저 별 추천, 게임 별 추천,
게임 타이틀 데이터셋 병합

gameId	userId	ex_gameId	title
3360	17190	3682	I Love You, Colon...
3360	17190	2779	Grimmwood: They C...
3360	17190	181	Hell is Other Demons
1371	17190	3682	I Love You, Colon...
1371	17190	2779	Grimmwood: They C...
1371	17190	181	Hell is Other Demons
1371	17190	3682	I Love You, Colon...
6453	17190	2779	Grimmwood: They C...
6453	17190	181	Hell is Other Demons
1915	17190	3682	I Love You, Colon...

Group by gameId
collect_set()

gameId	recommended
0	[Death's Gambit: Afterlife', '99Vidas', 'Shroud of the Avatar: Forsaken Virtues', 'Nancy Drew: The Secret of Shadow Ranch', 'Crownstakes', 'Aeterna Noctis', 'Super Seducer: How to Talk to Girls']
1	[Super Mega Baseball 3', 'RACE: The WTCC Game', 'Emergency 2: The Ultimate Fight for Life', '8BitBoy', 'AnaCapri: The Dream', 'LEGO City Undercover', 'Age of Wushu', 'Splitgate', 'Traffic Giant']
2	[Star Wars: The Old Republic - Rise of the Hutt Cartel', '2064: Read Only Memories', 'Medieval Kingdom Wars', 'Darkstar: The Interactive Movie', '911 Fire Rescue', 'Traffic Giant', 'Time Machine VR']
3	[Nottu Love 2: Devotion', 'Space Rangers HD: A War Apart', 'Rakuen', 'BONWORKS', 'Parkasaurus', 'Ghostbusters', 'Rusty Lake Hotel']
4	[Jimmy and the Pulsating Mass', 'Death's Gambit: Afterlife', 'Tiny Metal: Full Metal Rumble', 'Dreamkards', 'Decisive Battles of WWII Vol 2: Kursk Pocket', '99Vidas', 'Return to Castle Wolfenstein: Enemy Territory', 'Gray Dawn']
5	[Super Mega Baseball 3', 'RACE: The WTCC Game', 'Traffic Giant']
6	[Super Mega Baseball 3', 'Shadow of Destiny', 'RACE: The WTCC Game', 'Battle Princess Madelyn', 'Traffic Giant', 'WASTED: A Post-Apocalyptic Pub Crawler']
7	[Winds of Change', 'Death's Gambit: Afterlife', 'Star Drift Evolution', 'LISA', 'Midnight Nowhere', 'Dungeon Fighter Online', 'Traffic Giant', 'Angels with Scaly Wings']
8	[Killer7', 'RACE: The WTCC Game', 'Death's Gambit: Afterlife', 'SpinTires', 'Traffic Giant', 'Angels with Scaly Wings', 'Gray Dawn', 'Project Wingman']
9	[Super Mega Baseball 3', 'Star Wars: Battlefront II', 'RACE: The WTCC Game', 'Death's Gambit: Afterlife', 'Ultimate Fishing Simulator', 'Karma Incarnation 1', '99Vidas', 'What Happened', 'Traffic Giant']

05.

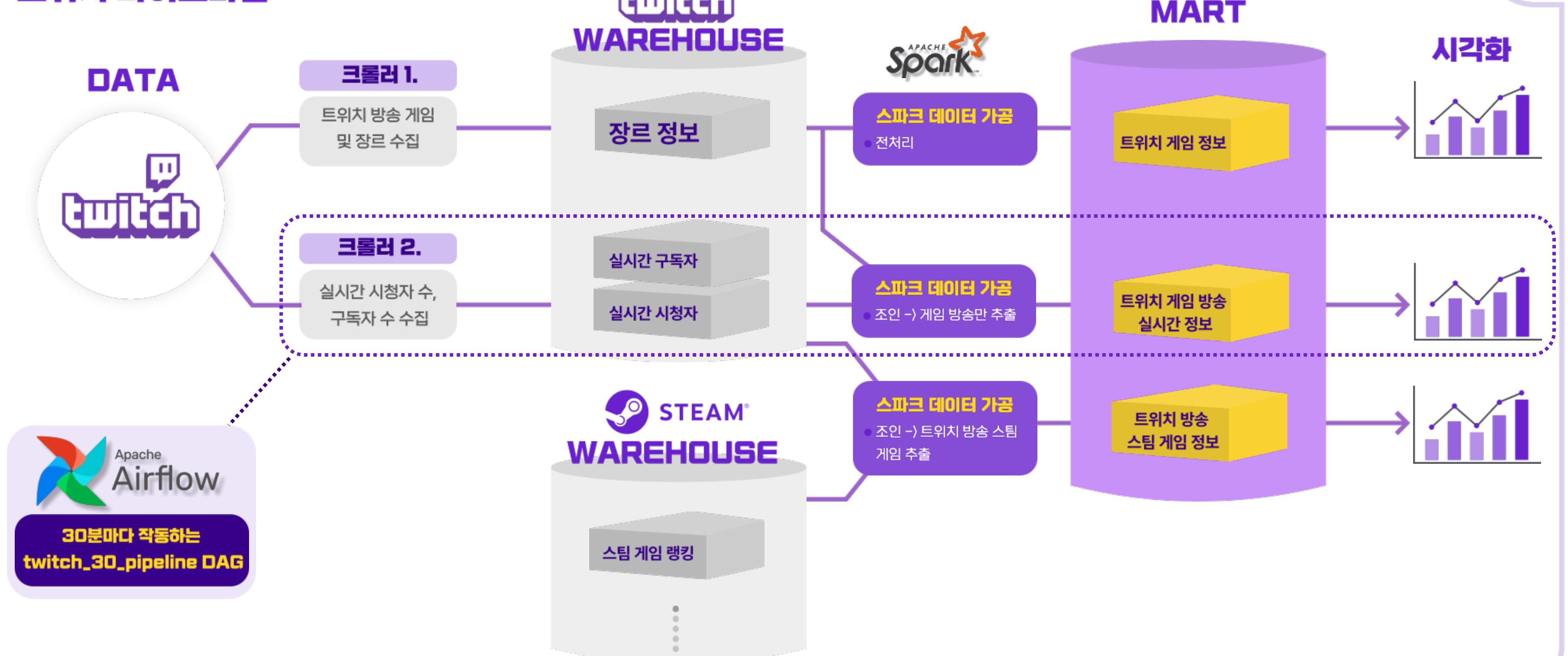
게임 방송

트위치 데이터 파이프라인

트위치 데이터 ETL 과정 및 분석

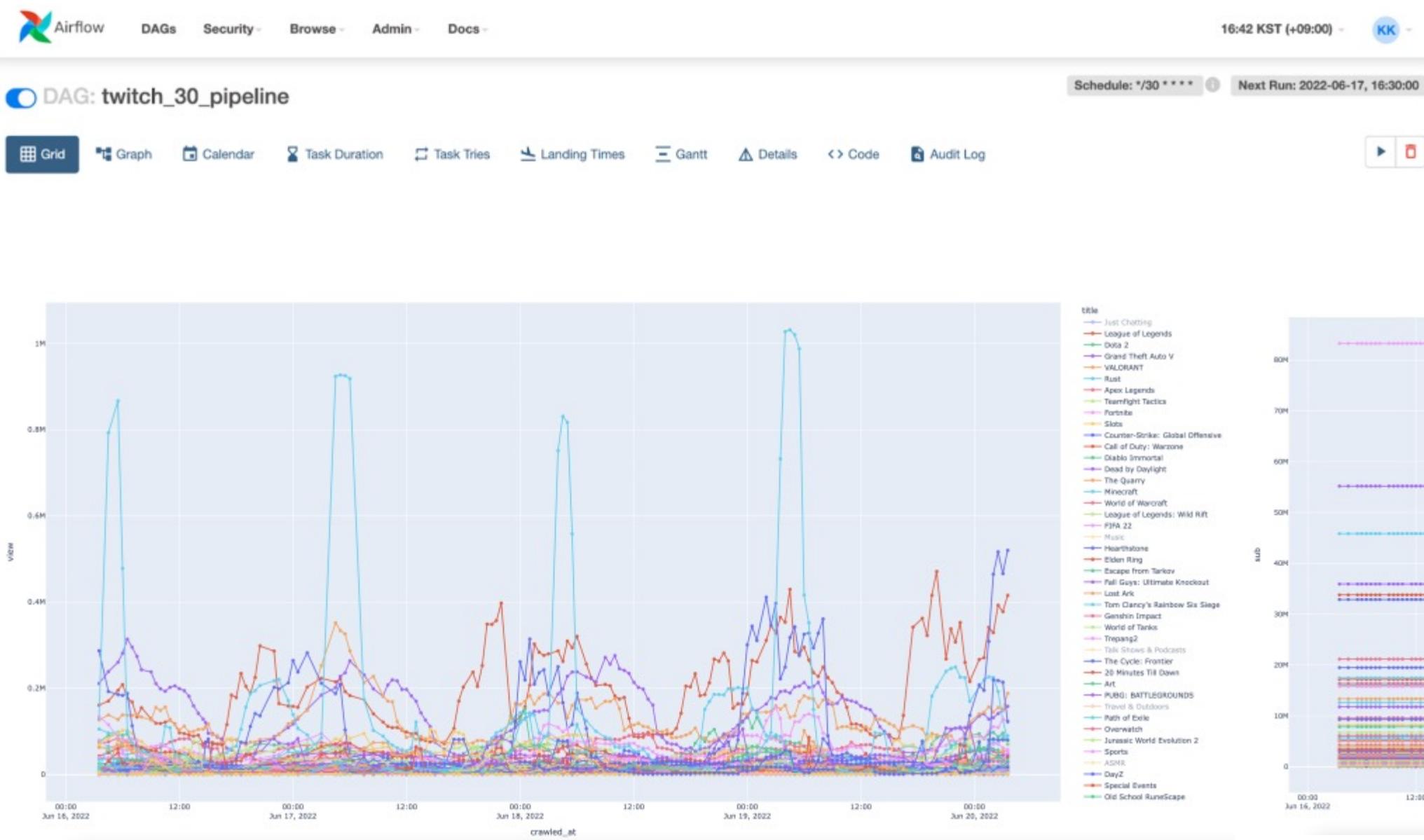
5. 게임 방송

트위치 파이프라인



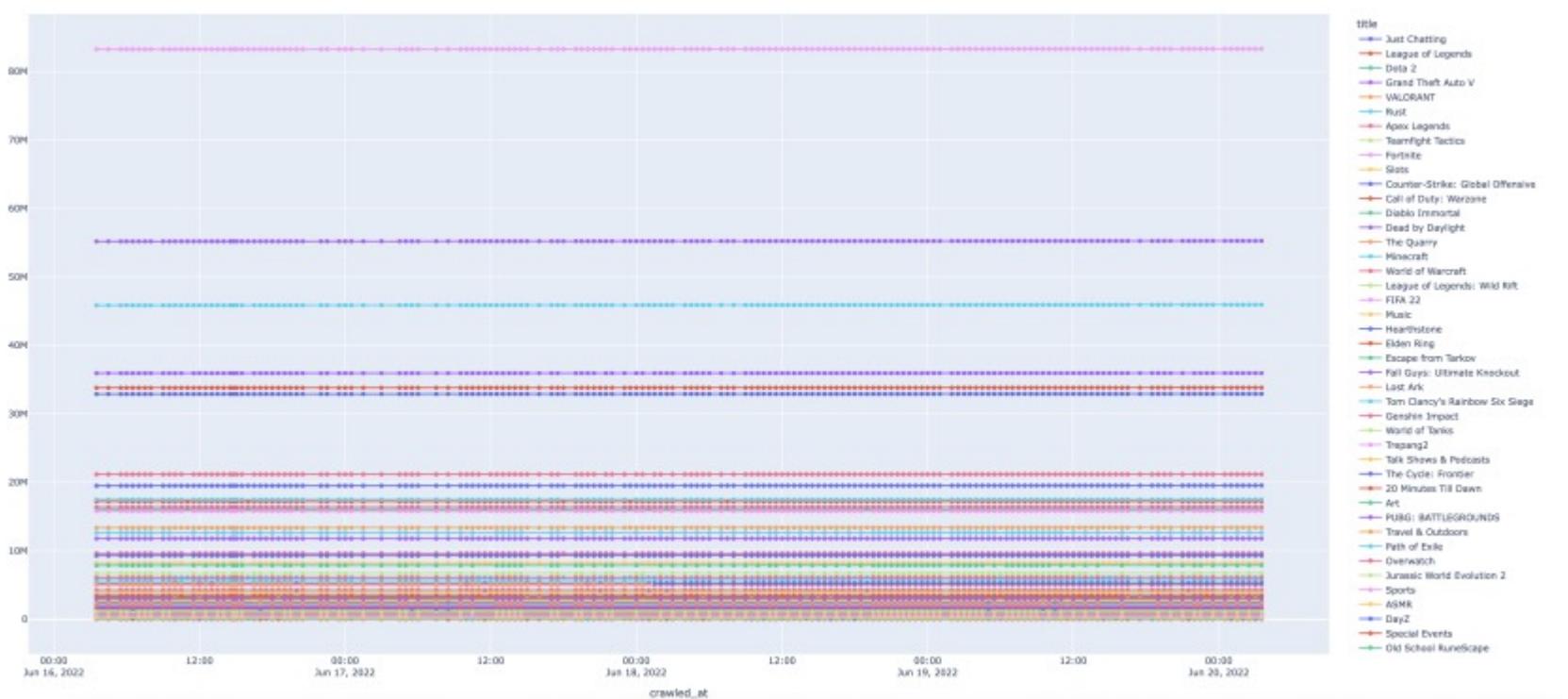
5. 게임 방송

에어플로우 트위치 DAG



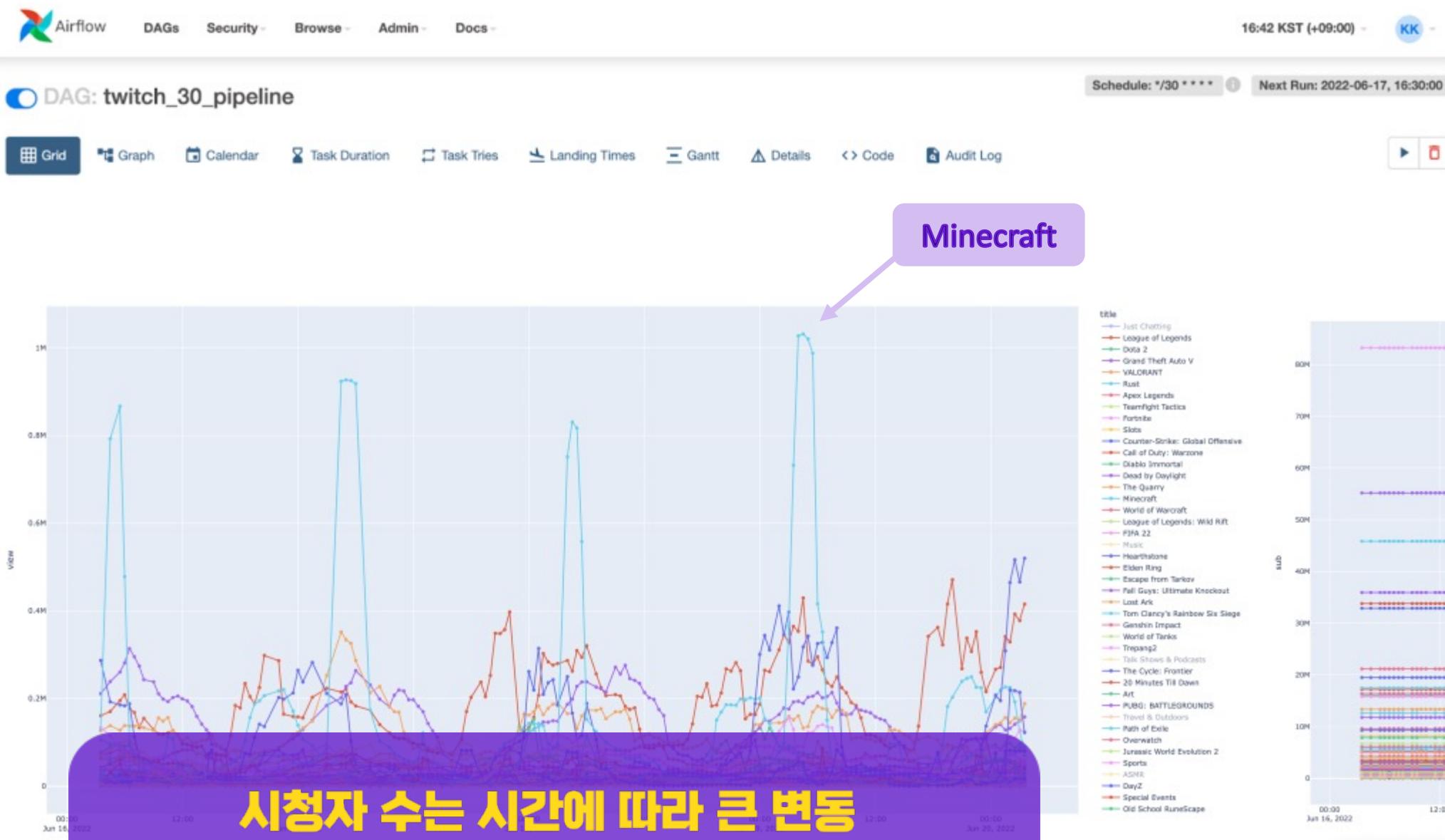
30분마다 데이터 크롤링 ~ 시각화

- 게임별 시청자 수 및 구독자 수 크롤링부터 시각화까지



5. 게임 방송

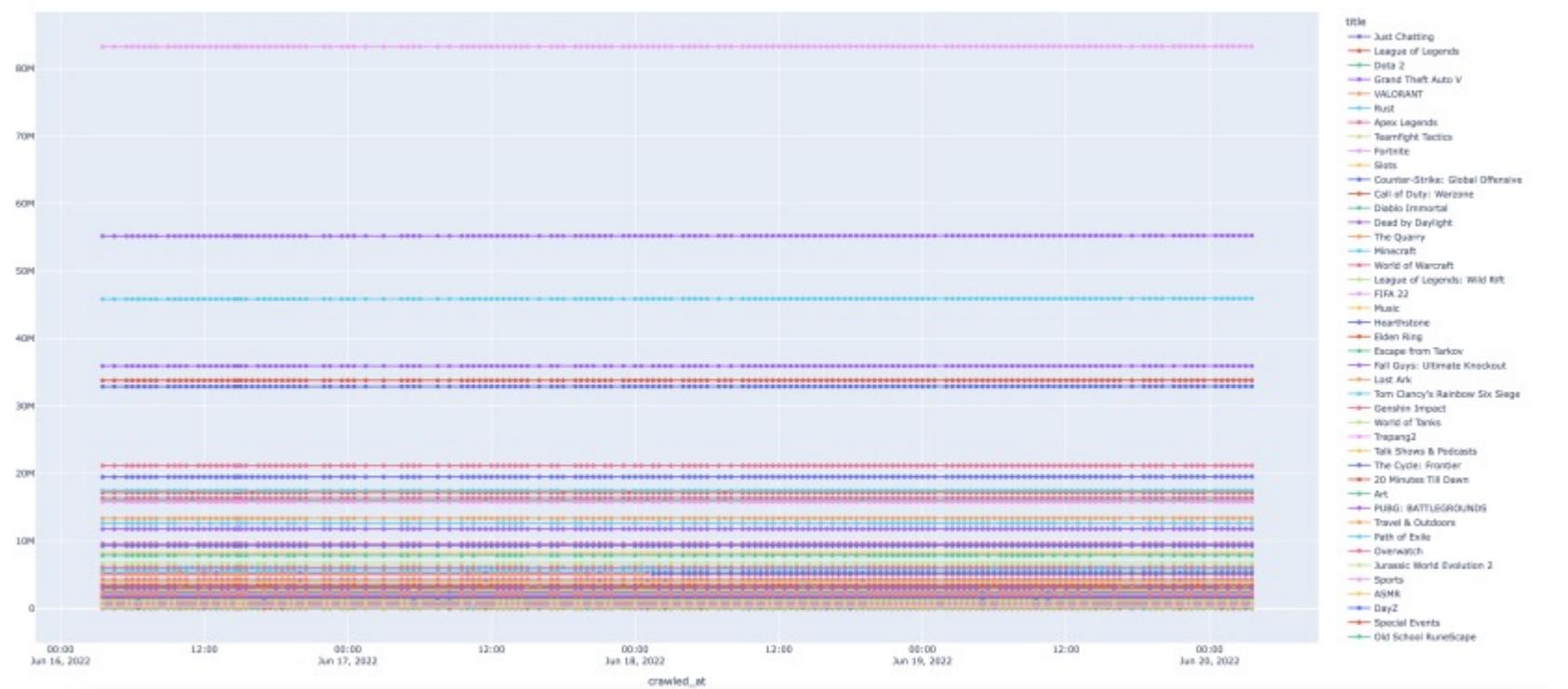
에어플로우 트위치 DAG



특히 Minecraft는 가장 큰 변동을 보이고, 가장 높은 피크 값을 가지지만 평균 시청 점유율은 3위
→ 전반적 인기보다 특정 트위치 채널에서 방송할 때 엄청난 양의 시청자 유입 유추 가능

30분마다 데이터 크롤링 ~ 시각화

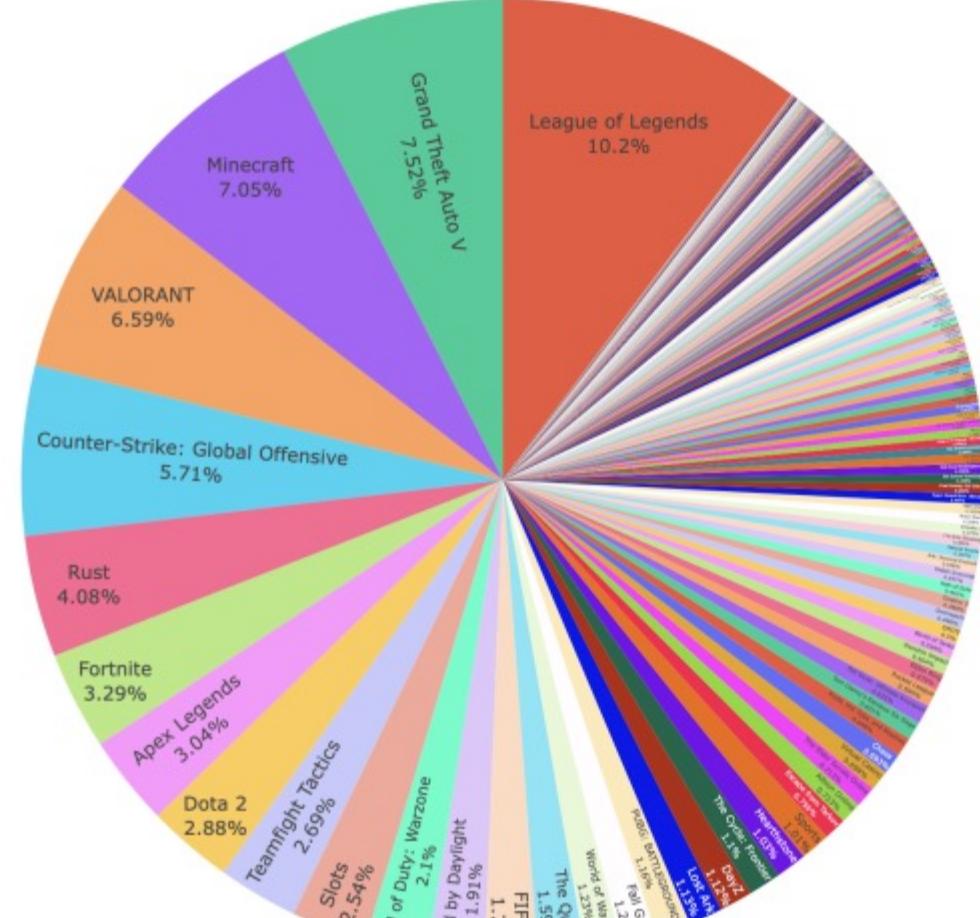
- 게임별 시청자 수 및 구독자 수 크롤링부터 시각화까지



5. 게임 방송

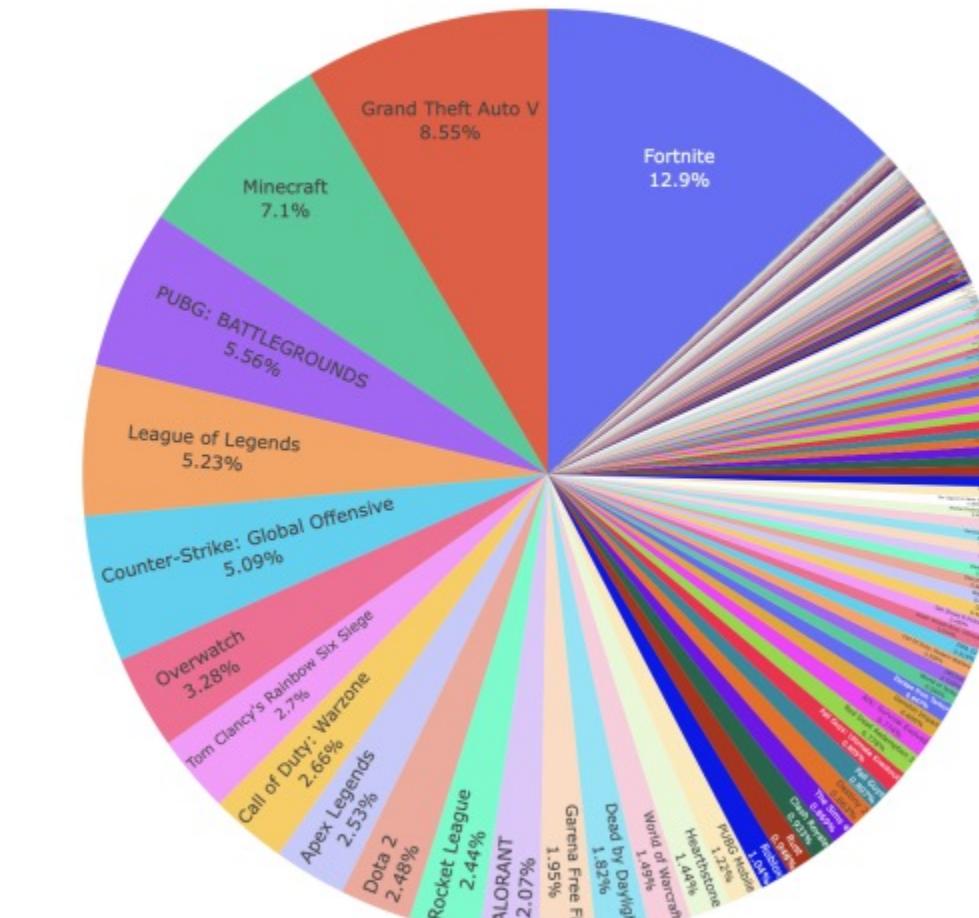
에어플로우 트위치 DAG : 게임별 평균 시청자 수 vs. 평균 구독자 수

Current Twitch Streaming Game Average Viewers



- Just Chatting
- League of Legends
- Grand Theft Auto V
- Minecraft
- VALORANT
- Counter-Strike: Global Offensive
- Rust
- Fortnite
- Apex Legends
- Dota 2
- Teamfight Tactics
- Slots
- Call of Duty: Warzone
- Dead by Daylight
- FIFA 22
- The Quarry
- Music
- World of Warcraft
- Fall Guys
- PUBG: BATTLEGROUNDS
- Lost Ark
- DayZ
- The Cycle: Frontier
- Hearthstone
- Sports
- ASMR
- Escape from Tarkov
- Albion Online
- The Elder Scrolls Online
- Virtual Casino
- Chess
- Pools, Hot Tubs, and Beaches
- Tom Clancy's Rainbow Six Siege
- Fall Guys: Ultimate Knockout
- Rocket League
- Art
- Elden Ring
- Genshin Impact
- World of Tanks
- SMITE
- Overwatch
- Destiny 2
- Path of Exile
- Diablo Immortal

Current Twitch Streaming Game Average Subscribers

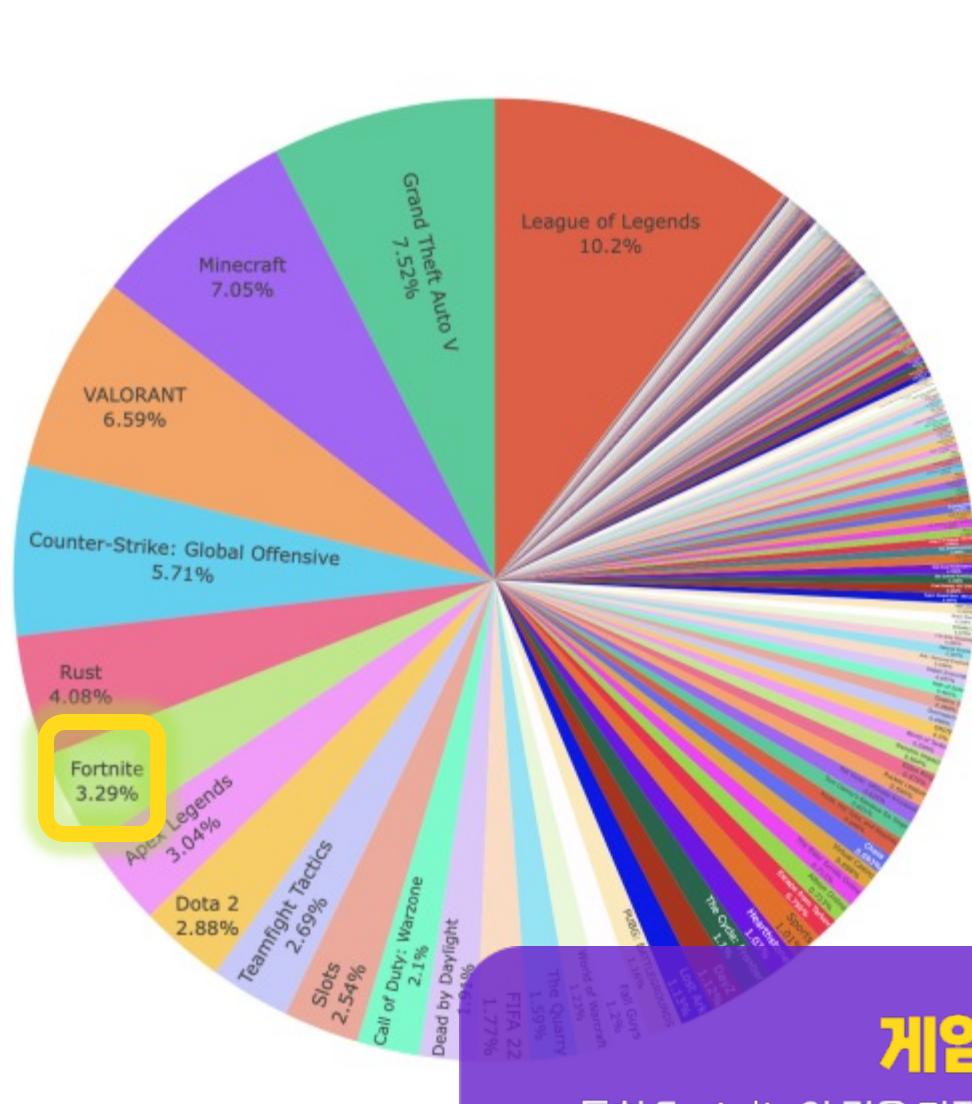


- Fortnite
- Grand Theft Auto V
- Minecraft
- PUBG: BATTLEGROUNDS
- League of Legends
- Counter-Strike: Global Offensive
- Overwatch
- Just Chatting
- Tom Clancy's Rainbow Six Siege
- Call of Duty: Warzone
- Apex Legends
- Dota 2
- Rocket League
- VALORANT
- Garena Free Fire
- Dead by Daylight
- World of Warcraft
- Hearthstone
- Music
- PUBG Mobile
- Roblox
- Rust
- Clash Royale
- The Sims 4
- Destiny 2
- Fall Guys
- Fall Guys: Ultimate Knockout
- Red Dead Redemption 2
- Ark: Survival Evolved
- Genshin Impact
- Escape from Tarkov
- ASMR
- World of Tanks
- VRChat
- Call Of Duty: Modern Warfare
- FIFA 22
- FIFA 22
- Super Smash Bros. Ultimate
- Art
- Talk Shows & Podcasts
- DayZ
- Warframe
- Sea of Thieves
- Dark Souls III
- Brawl Stars

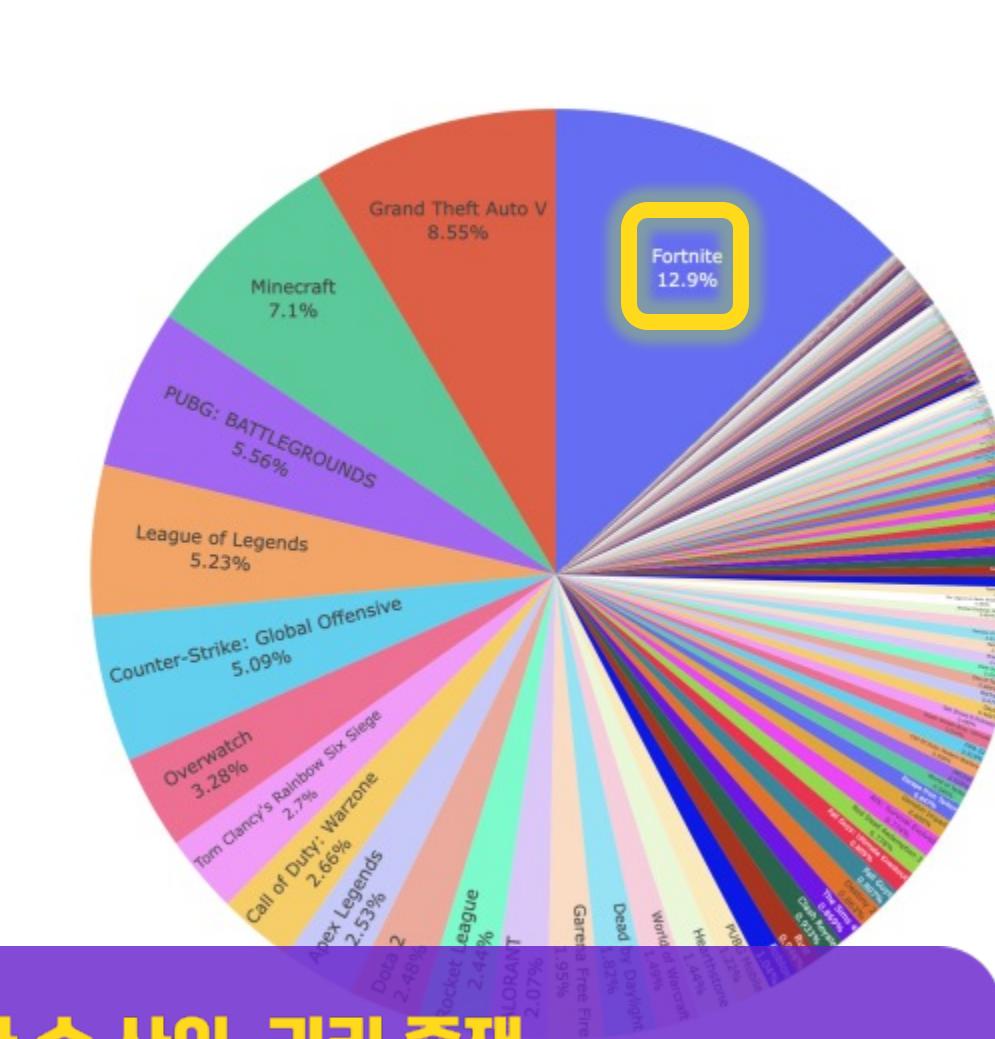
5. 게임 방송

에어플로우 트위치 DAG : 게임별 평균 시청자 수 vs. 평균 구독자 수

Current Twitch Streaming Game Average Viewers



Current Twitch Streaming Game Average Subscribers



게임 시청자 수와 구독자 수 사이 괴리 존재

- 특히 Fortnite의 경우 가장 많은 구독자 수를 보유하고 있으나, 시청 점유율에서는 3.29%에 그침 : 구독자 수는 시간에 따라 큰 변동이 없는 현상을 보이므로 과거 유입된 구독자가 많으나 현재 시청 유행에서 뒤쳐지는 것으로 유추할 수 있음

5. 게임 방송

트위치 전체 게임 장르 vs. 평균 시청자 5000명 이상 게임 장르

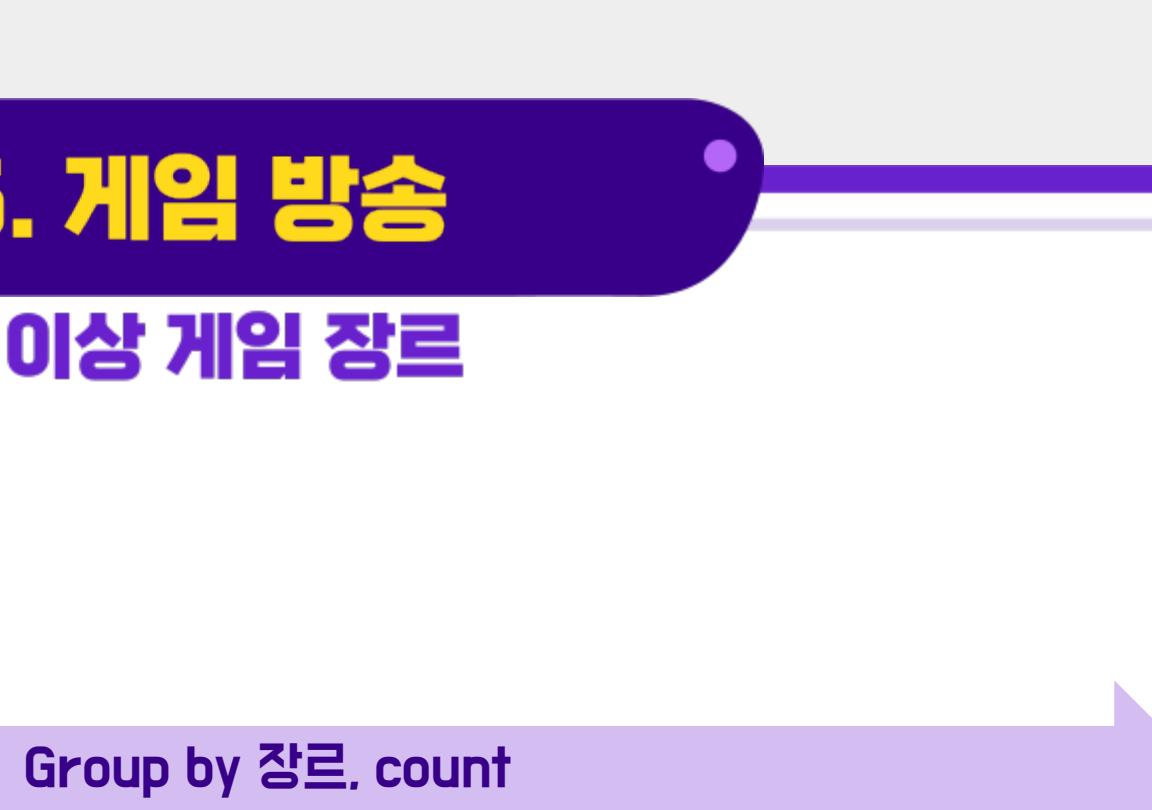
title	genres
Just Chatting	리얼라이프
League of Legends	MOBA
League of Legends	액션
Dota 2	MOBA
Dota 2	액션
Dota 2	전략
Grand Theft Auto V	어드벤처 게임
Grand Theft Auto V	슈팅 장르
VALORANT	어드벤처 게임
Rust	어드벤처 게임
Rust	슈팅 장르
Apex Legends	FPS
Apex Legends	슈팅 장르
Apex Legends	액션
Teamfight Tactics	전략
Teamfight Tactics	카드 및 보드 게임
Fortnite	슈팅 장르
Fortnite	액션
Fortnite	생존
Slots	겜블링 게임
Counter-Strike: G...	FPS
Counter-Strike: G...	슈팅 장르
...	
Fall Guys: Ultima...	플랫포머 게임

게임만 뽑기

title	genres
League of Legends	MOBA
Dota 2	MOBA
Dota 2	액션
Dota 2	전략
Grand Theft Auto V	어드벤처 게임
VALORANT	어드벤처 게임
Rust	어드벤처 게임
Apex Legends	FPS
Teamfight Tactics	카드 및 보드 게임
Fortnite	슈팅 장르
Fortnite	액션
Fortnite	생존
Slots	겜블링 게임
Counter-Strike: G...	FPS
Counter-Strike: G...	슈팅 장르
...	
Fall Guys: Ultima...	플랫포머 게임

Join Two Tables
on title

title	avg(view)	avg(sub)
Dota 2	44517.33870967742	1.60371645E7
Travel & Outdoors	2551.5	419549.224137931
Final Fantasy VII...	1184.5714285714287	291452.9285714286
Summoners War: Sk...	925.4883720930233	919731.0930232558
Super Mario 64	2568.785714285714	2393373.5
Fall Guys: Ultima...	10161.7627711864407	5202980.610169492
Metal Gear Solid ...	210.75	284765.475
Dynasty Warriors ...	2.6153846153846154	4960.527272727273
MapleStory	3401.814814815	923411.2037037037
Spelunky 2	193.48837209302326	117847.27906976744
Eternal Return	6083.368421052632	424492.70175438595
Dead by Daylight	34927.3870967742	1.177409841935484E7
DayZ	17665.80701754386	3119306.666666
Talk Shows & Podc...	6005.864406779661	3161589.84
Elden Ring	11094.338983050848	695412.309
Clash Royale	4459.0	6029040.96767445
Diablo II: Resurr...	1687.5581395348838	281110.1395348837
Legends of Runeterra	835.045454545455	619720.3636363636
Old School RuneScape	6222.578947368421	1137139.7719298245
Lineage 2	1578.408163265306	180125.5918367347



title	avg_view	avg_sub	genres
Dota 2	46053.46268656717	1.6037397074626865E7	전략
Dota 2	46053.46268656717	1.6037397074626865E7	액션
Dota 2	46053.46268656717	1.6037397074626865E7	MOBA
Final Fantasy VII...	1118.4	291463.516666666666	RPG
Final Fantasy VII...	1118.4	291463.516666666666	액션
Summoners War: Sk...	932.2391304347826	919736.2173913043	MMO
Summoners War: Sk...	932.2391304347826	919736.2173913043	액션
Summoners War: Sk...	932.2391304347826	919736.2173913043	전략
Super Mario 64	2498.066666666666	2393409.0888888887	FPS
Super Mario 64	2498.066666666666	2393409.0888888887	플랫포
Fall Guys: Ultima...	11294.78125	5203275.953125	플랫포
Fall Guys: Ultima...	11294.78125	5203275.953125	전략
Metal Gear Solid ...	199.27906976744185	284767.0465116279	전략
Metal Gear Solid ...	199.27906976744185	284767.0465116279	액션
Dynasty Warriors ...	2.5714285714285716	4960.5	액션
MapleStory	3365.2105263157896	923423.4561403509	어드벤처 게임
MapleStory	3365.2105263157896	923423.4561403509	액션
Spelunky 2	182.1304347826087	117048.34782608696	액션
Spelunky 2	182.1304347826087	117048.34782608696	플랫포
Eternal Return	5901.475409836065	424521.34426229505	생존

genres	genre_count
액션	786
RPG	568
어드벤처 게임	513
전략	333
시뮬레이션	328
슈팅 장르	323
인디 게임	240
FPS	177
플랫포 게임	154
퍼즐	138
공포	134
스포츠	90
운전/경주 게임	86
No Genre	83
격투게임	75
MMO	66
오픈 월드	64
RTS	46
모바일 게임	43
리듬 게임	43
생존	39
카드 및 보드 게임	32
비행 시뮬레이터	2
게임 오버레이	1

시청자 수가 많은 게임에는 FPS 장르가 많은 편

genres	top_view_genre_count
액션	231
FPS	16
슈팅 장르	13
MMO	12
전략	9
어드벤처 게임	7
시뮬레이션	6
오픈 월드	5
MOBA	5
플랫포	3
공포	3
카드 및 보드 게임	3
퍼즐	2
이케아드	2
생존	2
겜블링 게임	2
스포츠	2
운전/경주 게임	1
Point and Click	1
격투게임	1
플랫포 게임	1

5. 게임 방송

트위치로 "보는" 게임 vs. 실제로 "하는" 게임

game	rank
Counter-Strike: Global Offensive	1
Dota 2	2
Apex Legends	3
PUBG: BATTLEGROUNDS	4
Wallpaper Engine	5
Grand Theft Auto V	6
Destiny 2	7
ARK: Survival Evolved	8
Unturned	9
Lost Ark	10
Warframe	11
Team Fortress 2	12
ELDEN RING	13
V Rising	14
Tom Clancy's Rainbow Six Siege	15
Rust	16
NARAKA: BLADEPOINT	17
Rocket League	18
Euro Truck Simulator 2	19
Risk of Rain 2	20

game	rank	title	avg(view)	avg(sub)
Dota 2	2	Dota 2	44517.33870967742	1.60371645E7
Fall Guys: Ultimate Knockout	148	Fall Guys: Ultimate Knockout	10161.762711864407	5202980.610169492
METAL GEAR SOLID V: The Phantom Pain	181	Metal Gear Solid V: The Phantom Pain	210.75	284765.475
MapleStory	193	MapleStory	3401.814814814815	923411.2037037037
Dead by Daylight	43	Dead by Daylight	34927.3870967742	1.177409841935484E7
DayZ	29	DayZ	17665.80701754386	3119306.6666666665
ELDEN RING	13	Elden Ring	11094.338983050848	695412.3389830509
TEKKEN 7	133	Tekken 7	1092.933333333334	1821200.4444444445
Project Zomboid	22	Project Zomboid	1176.7906976744187	330742.79069767444
Rust	16	Rust	60179.12903225807	6120109.758064516
FINAL FANTASY XIV: A Realm Reborn	57	Final Fantasy XIV: A Realm Reborn	5821.035087719299	1637865.49122807
Phasmophobia	25	Phasmophobia	5699.294117647059	1538055.8823529412
Sea of Thieves	44	Sea of Thieves	4074.090909090909	2997318.8727272726
Paladins	123	Paladins	1685.5853658536585	2272063.7804878047
Street Fighter V	125	Street Fighter V	1245.2558139534883	2003347.1627906978
Hearts of Iron IV	55	Hearts of Iron IV	348.6444444444445	393942.1333333336
Tom Clancy's Rainbow Six Siege	15	Tom Clancy's Rainbow Six Siege	13600.949152542373	1.7473370949152544E7
DiRT Rally 2.0	228	DiRT Rally 2.0	419.64814814814815	110034.88888888889
VRChat	27	VRChat	4863.44	3414743.14
Red Dead Redemption 2	71	Red Dead Redemption 2	9789.018518518518	5024851.833333333
Assetto Corsa Competizione	132	Assetto Corsa Competizione	608.4594594594595	273088.3783783784
Resident Evil 2	161	Resident Evil 2	1656.6046511627908	416439.76744186046
...				
Mount & Blade II: Bannerlord	124	Mount & Blade II: Bannerlord	186.0612244897959	249134.4693877551

title	avg(view)	avg(sub)
Dota 2	44517.33870967742	1.60371645E7
Travel & Outdoors	2551.5	419549.224137931
Final Fantasy VII: Remastered	1184.5714285714287	291452.9285714286
Summoners War: Sky Arena	925.4883720930233	919731.0930232558
Super Mario 64	2568.785714285714	2393373.5
Fall Guys: Ultimate Knockout	10161.762711864407	5202980.610169492
Metal Gear Solid V: The Phantom Pain	210.75	284765.475
Dynasty Warriors 9	2.6153846153846154	4960.527272727273
MapleStory	3401.814814814815	923411.2037037037
Spelunky 2	193.48837209302326	117047.27906976744
Eternal Return	6083.368421052632	424492.70175438595
Dead by Daylight	34927.3870967742	1.177409841935484E7
DayZ	17665.80701754386	3119306.6666666665
Talk Shows & Podcasts	6005.864406779661	3161589.8474576273
Elden Ring	11094.338983050848	695412.3389830509
Clash Royale	4459.0	6029040.9069767445
Diablo II: Resurrected	1687.5581395348838	281110.1395348837
Legends of Runeterra	835.0454545454545	619720.3636363636
Old School RuneScape	6222.578947368421	1137139.7719298245
Lineage 2	1578.408163265306	180125.5918367347

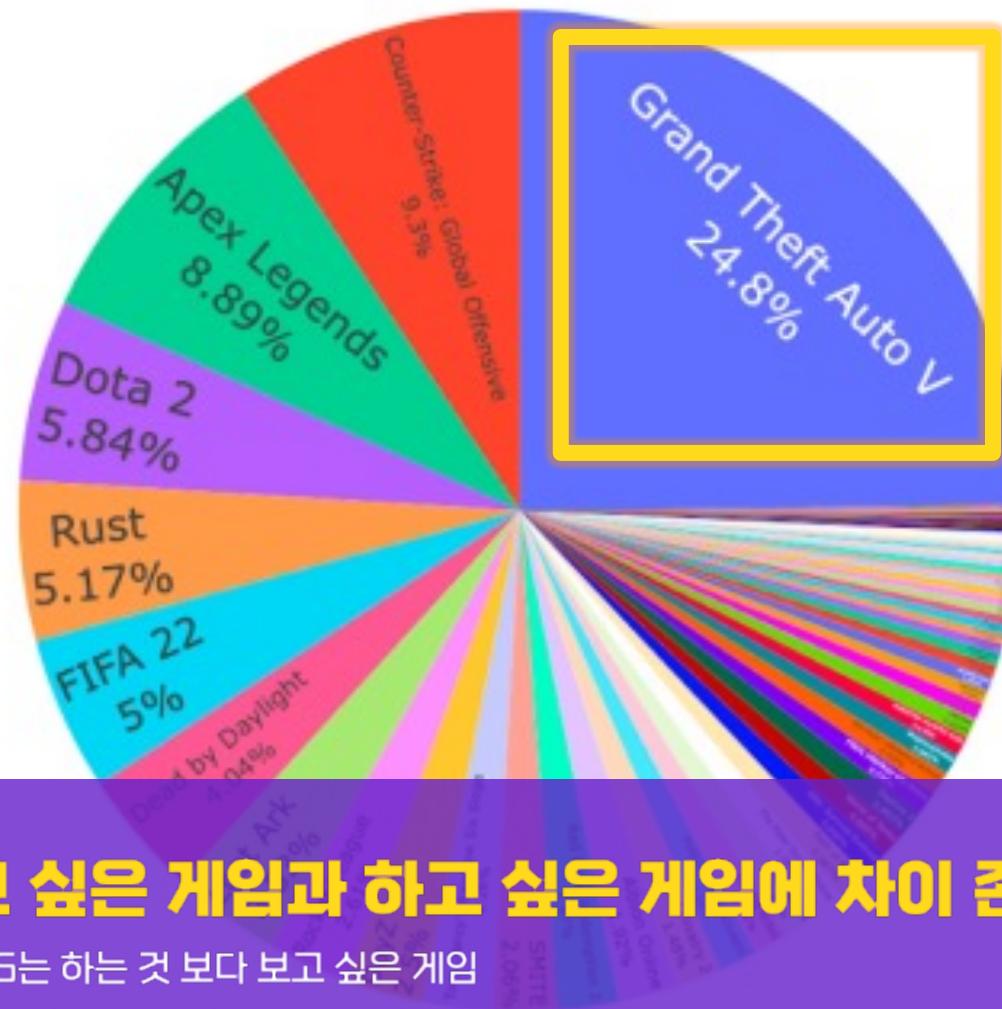
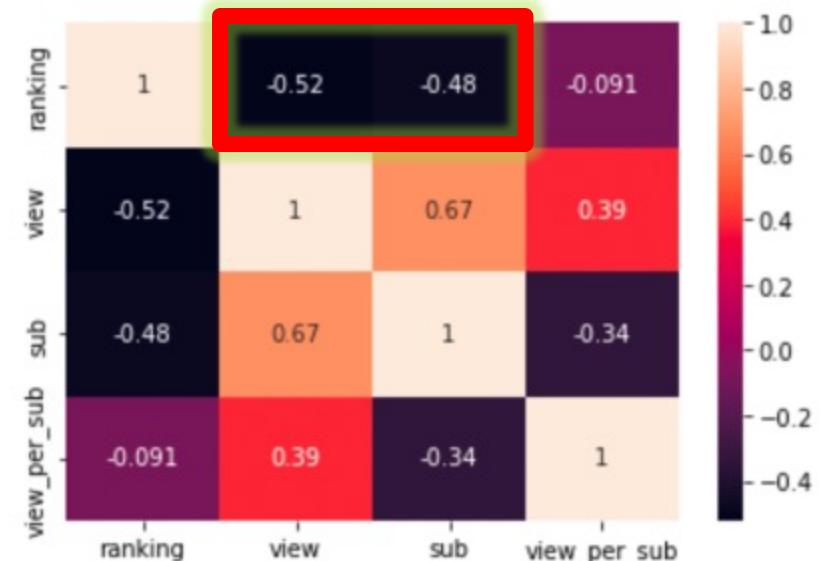
스팀 DB

트위치에서 방송되는 스팀 게임

트위치 DB

5. 게임 방송

트위치로 "보는" 게임 vs. 실제로 "하는" 게임



보고 싶은 게임과 하고 싶은 게임에 차이 존재

- GTA 5는 하는 것 보다 보고 싶은 게임
 - 배그는 보는 것보다 하고 싶은 게임

ranking	Game
1	Counter-Strike: Global Offensive
2	Dota 2
3	Apex Legends
4	PUBG: BATTLEGROUNDS
6	Grand Theft Auto V
7	Destiny 2
8	ARK: Survival Evolved
10	Lost Ark
11	Warframe
13	ELDEN RING
14	V Rising
15	Tom Clancy's Rainbow Six Siege
16	Rust
18	Rocket League
19	Euro Truck Simulator 2
21	War Thunder
22	Project Zomboid
24	MONSTER HUNTER RISE
25	Phasmophobia
27	VRChat
28	The Elder Scrolls Online
29	DayZ
30	FIFA 22
32	Stardew Valley
33	Vampire Survivors
39	7 Days to Die
43	Dead by Daylight
44	Sea of Thieves
47	F1 2021
48	The Sims 4
49	Black Desert
52	New World
55	Hearts of Iron IV
57	FINAL FANTASY XIV Online

06.

개발 후기

개발 후기 및 발전 방향 제시

6. 개발 후기

06

개발 후기 및 발전방향 제시

팀장 : 민윤지

스파크로 전처리하는 것이 만만치 않았다. 처음부터 공부하면서 하느라 시간도 오래 걸렸고, 대단한 결과물을 낸 것도 아니지만 스파크에 대해서 앞으로 어떤식으로 공부하면 될지 방향을 잡을 수 있었다

박시현

각보다 크롤러 만드는데 너무 많은 시간이 걸려 계획했던 만큼의 데이터를 뽑아오지 못해 아쉬웠다. 팀원들에게 많은 도움을 받아 이번 프로젝트를 마무리 할 수 있었다. 팀원들에게 너무 감사하다.

김희원

생각보다 시간이 짧아 에어플로우 활용이나 머신러닝 등 해보고 싶은 걸 전부 못해본 것 같아 아쉽습니다.

추호연

맡은 부분이 정적 페이지라 크롤링이 쉬울 줄 알았는데 몇만개의 페이지에 접근하다 보니 접속 경고 페이지나 오류 페이지가 떠서 데이터가 들어가지 않거나 샘플 페이지에서는 없었던 여러 예외 문제가 발생해서 생각보다 시간을 많이 소모했고, 그 외에도 수업 때는 없었던 문제들이 발생해서 그래서 프로젝트를 해보는거구나 하는 생각이 들었습니다.

강규란

함께 고생해준 팀원들에게 감사하다.
아직 하나 더 남았다니 믿을 수 없다.



37 / 37



치킨4조



감사합니다!

● 치킨 4조 : 데이터 엔지니어링 프로젝트

● 깃헙: <https://github.com/mitchell-pritchett/dags.git> (WIP)