

EDA

영화 흥행에 미치는 요인 분석

11기 김지원 송예원
12기 김경환 김지은



0

목차

1 주제 소개

2 데이터 수집 및 전처리

3 데이터 시각화

4 Quantile Regression

5 sns 분석

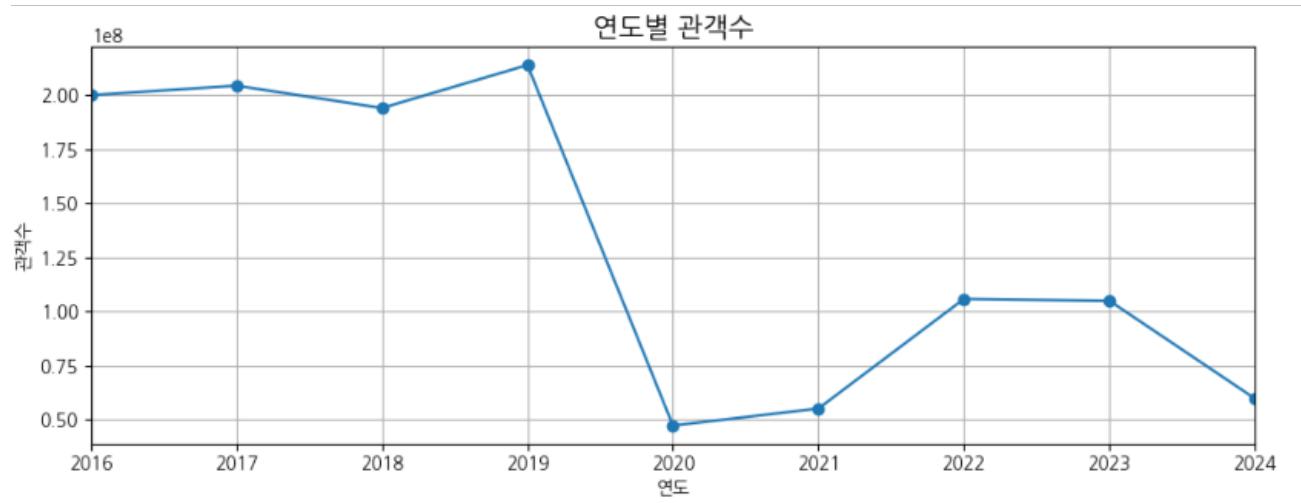
6 결론 및 한계점

주제 소개

극장 영화 산업의 위기



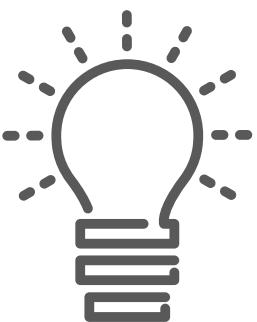
천만 아님 안 봐... 관객수 양극화



올해 관객수 1, 2위 1000만 돌파
3~9위까지는 100만명대 머물러
OTT 인기에 극장 찾는 발길 줄어

2

데이터 수집 및 전처리



데이터 수집

[1] KOBIS 박스오피스 데이터

- 관객수 100명 이상의 데이터 수집

순위		영화명	개봉일	매출액	매출액	점유율	관객수	스크린수	상영횟수	대표국적	국적	배급사
0	1	부산행	2016-07-20	93178283048		0.056	11565479	1788	151315	한국	한국	(주)넥스트엔터테인먼트월드(NEW)
1	2	검사외전	2016-02-03	77320403264		0.046	9707581	1812	153532	한국	한국	(주)쇼박스
2	3	캡틴 아메리카: 시빌 워	2016-04-27	72672111827		0.044	8677249	1991	161696	미국	미국	월트디즈니컴퍼니코리아 유한책임회사, 월트디즈니컴퍼니코리아(주)
3	4	밀정	2016-09-07	61269783831		0.037	7500420	1444	138161	한국	한국, 미국	워너브러더스 코리아(주)
4	5	터널	2016-08-10	57529484417		0.034	7120508	1105	113201	한국	한국	(주)쇼박스
...
114	396	킹 VS. 머신킹	2024-02-15	781000		0.000	163	10	55	미국	미국	(주)제이앤씨미디어그룹
115	401	서치: 데스게임	2024-01-18	960000		0.000	160	5	20	러시아	러시아	주식회사 원더스튜디오
116	515	아마존: 더 비기닝	2024-01-11	606000		0.000	101	5	11	미국	미국	(주)누리픽쳐스
117	522	도토리	2024-06-27	500000		0.000	100	1	1	한국	한국	원코리아미디컴
118	522	뺑소니: 최고로 최악의 날	2024-06-16	667300		0.000	100	6	60	일본	일본	(주)미디어캐슬

데이터 수집

[2] 네이버 영화 관람평 크롤링

- 수집 변수: 아이디, 별점, 관람평, 작성 날짜, 공감 수, 비공감 수

• 공감순 • 최신순

스포일러 포함

별점	작성자	내용	작성일	공감수	비공감수
★★★★★ 10	fnrd****	"어른이 된다는게 이런건가봐 기쁨이 줄어드는거"	2024.06.12. 12:07	8,147	285
★★★★★ 10	smgg****	결국 시작은 '잘 해내야한다'는 불안감. 그냥 좀 못합시다, 뭐 어때요	2024.06.12. 13:07	4,975	152
★★★★★ 10	perc****	어른이 되고 나니 '불안이'의 마음을 너무 잘 알겠어서, 예상치 못한 장면들에서 눈물이 났다. 불안이라는 감정과 친해질 수 있는 영화가 될 듯	2024.06.12. 19:08	4,019	74



	keyword	id	rating	comment	date	up	down
1	인사이드 아웃 2	fnrd****	별점(10점 만점 중) 10	"어른이 된다는게 이런건가봐 기쁨이 줄어드는거"	2024.06.12.	7,773	274
2	인사이드 아웃 2	smgg****	별점(10점 만점 중) 10	결국 시작은 '잘 해내야한다'는 불안감. 그냥 좀 못합시다, 뭐 어	2024.06.12.	4,738	142
3	인사이드 아웃 2	perc****	별점(10점 만점 중) 10	어른이 되고 나니 '불안이'의 마음을 너무 잘 알겠어서, 예상치	2024.06.12.	3,804	71
4	인사이드 아웃 2	cber****	별점(10점 만점 중) 10	빙봉은 없었다. 하지만 영화가 끝났을땐 그 안에 내가 있었다.	2024.06.12.	3,700	219
5	인사이드 아웃 2	blue****	별점(10점 만점 중) 10	<엘리멘탈>로 사랑의 감동을 주더니 1년만에 우리 모두가 가	2024.06.12.	2,174	69
6	인사이드 아웃 2	kenn****	별점(10점 만점 중) 9	내 감정들과 내 자신을 돌아볼 수 있었던 시간	2024.06.12.	1,930	50

데이터 수집

[3] 네이버 영화 정보 크롤링

- 수집 변수: 등급, 장르, 러닝타임

개봉	2016.04.27.
등급	12세 이상 관람가
장르	액션, SF, 스릴러
국가	미국
러닝타임	147분
배급	월트 디즈니 컴퍼니 코리아



	keyword	year	grade	genre	running_time
0	부산행	2016	15세 이상 관람가	액션, 스릴러	118분
1	검사외전	2016	15세 이상 관람가	범죄, 코미디	126분
2	캡틴 아메리카: 시빌 워	2016	12세 이상 관람가	액션, SF, 스릴러	147분
3	밀정	2016	15세 이상 관람가	액션	140분
4	터널	2016	12세 이상 관람가	드라마	126분

4

데이터 수집

[4] 인스타그램 영화관 계정 크롤링 (cgv, 메가박스, 롯데시네마)

- 수집 변수: 게시 날짜, 게시글, 좋아요 수



	account	post_num	post_date	text	num_likes
0	megaboxon	1	22시간 전	megaboxon\n💡SUIT-UP💡\n	\n📍 한 예매...
1	megaboxon	2	1일 전	megaboxon\n카카오뱅크 저금통 with 메가박스💰\n	\n지금 ...
2	megaboxon	3	1일 전	megaboxon\n슈트-업 한 악당 처리반 어떻게 참는데…👉\n피규어, 빔 키...	463
3	megaboxon	4	3일 전	megaboxon\n한 세기를 대표하는 시대의 명작\n다시 만나고 싶은 영화가 있나...	707
4	megaboxon	5	4일 전	megaboxon\n미니언의 최애 바나나 우유👉\n우유 이름 말하기 한호흡 챌린지!...	3458
...
1750	megaboxon	1750	2014년 7월 2일	megaboxon\n#메가박스 #MEGABOX #cinema #design #col...	32
1751	megaboxon	1751	2014년 7월 1일	megaboxon\n이 야밤에 간절히 생각나는 너\n\n#메가박스 #S포테이토 #M...	58
1752	megaboxon	1752	2014년 7월 1일	megaboxon\n오늘같이 선선한 저녁엔 붕붕이를 끌고 자동차극장에 가서 영화 한...	32
1753	megaboxon	1753	2014년 7월 1일	megaboxon\n당신의 꿈을 개봉합니다. #메가박스\n525주	56

데이터 전처리

[1] KOBIS 박스오피스 데이터

- rating_mean에서의 많은 결측치

1	영화명	개봉일	매출액	관객수	스크린수	상영횟수	대표국적	배급사	rating_mean	grade	genre	running_time
2	부산행	2016-07-20	93178283048	11565479	1788	151315	한국	(주)넥스트엔터	2.9222973	15세 이상 관람가		118분
3	검사외전	2016-02-03	77320403264	9707581	1812	153532	한국	(주)쇼박스	4.07142857	15세 이상 관람가		126분
4	캡틴 아메리카: 카이로의 복수	2016-04-27	72672111827	8677249	1991	161696	미국	월트디즈니컴퍼니	5.83549784	12세 이상 관람가		147분
5	밀정	2016-09-07	61269783831	7500420	1444	138161	한국	워너브라더스 코	9.02028986	15세 이상 관람가		140분
6	터널	2016-08-10	57529484417	7120508	1105	113201	한국	(주)쇼박스	6.51902174	12세 이상 관람가		126분
7	인천상륙작전	2016-07-27	55101410303	7049643	1049	97019	한국	(주)씨제이이엔티	9.31379668	12세 이상 관람가		110분
8	럭키	2016-10-13	56444461256	6975290	1234	155977	한국	(주)쇼박스	7.87967581	15세 이상 관람가		112분
9	곡성	2016-05-12	55863520382	6879908	1485	130322	한국	이십세기폭스코	2.03868472	15세 이상 관람가		156분
10	덕혜옹주	2016-08-03	44393828109	5599229	964	97312	한국	롯데쇼핑(주)롯데	6.53773585	12세 이상 관람가		127분
11	마스터	2016-12-21	39784641065	4935320	1501	70786	한국	(주)씨제이이엔티	4.16574279	15세 이상 관람가		143분
12	주토피아	2016-02-17	37096057742	4706158	735	131005	미국	월트디즈니컴퍼니	9.68032787	전체 관람가	애니메이션	108분
13	판도라	2016-12-07	33955398164	4313864	1184	90758	한국	(주)넥스트엔터테인먼트	8.40547945	12세 이상 관람가		136분
14	아가씨	2016-06-01	35211229991	4287839	1171	90365	한국	(주)씨제이이엔티	2.73387097	청소년 관람불가		144분
15	쿵푸팬더3	2016-01-28	31340798156	3984812	1365	97404	미국	(주)씨제이이엔티	7.39557229	전체 관람가	애니메이션	95분
16	데드풀	2016-02-17	27595618651	3317196	1036	82521	미국	월트디즈니컴퍼니	3.65869565	청소년 관람불가		128분
17	나우 유 씨	2016-07-13	25116690113	3100113	1200	56981	미국	롯데쇼핑(주)롯데	3.45028011	12세 이상 관람가		129분

rating_mean 결측치 : 1054
 grade 결측치 : 696
 genre 결측치 : 668
 running_time 결측치 : 633

데이터 전처리

[1] KOBIS 박스오피스 데이터

- 여전히 존재하는 결측치는 분포를 통해 결정

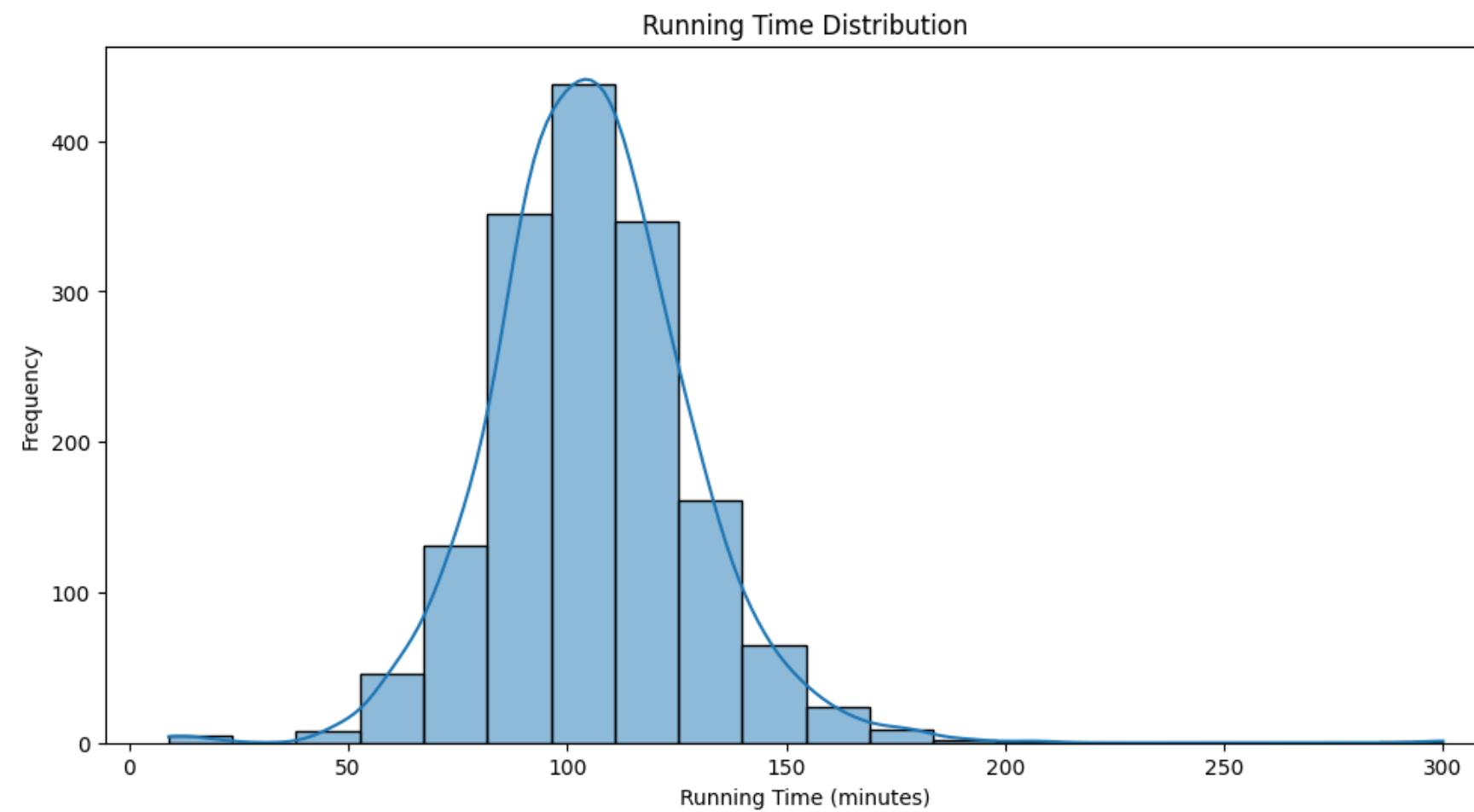
rating_mean 결측치 개수 : 0개
 grade 결측치 개수 : 60개
 genre 결측치 개수 : 53개
 running_time 결측치 개수 : 44개

1	영화명	개봉일	매출액	관객수	스크린수	상영횟수	대표국적	배급사	rating_mean	grade	genre	running_time
2	부산행	2016-07-20	93178283048	11565479	1788	151315	한국	(주)넥스트엔터	2.922297297	15세 이상 관람 액션, 스릴러	118분	
3	검사외전	2016-02-03	77320403264	9707581	1812	153532	한국	(주)쇼박스	4.071428571	15세 이상 관람 범죄, 코미디	126분	
4	캡틴 아메리카: 퍼시픽 전선	2016-04-27	72672111827	8677249	1991	161696	미국	월트디즈니컴퍼니	5.835497835	12세 이상 관람 액션, SF, 스릴러	147분	
5	밀정	2016-09-07	61269783831	7500420	1444	138161	한국	워너브라더스	9.020289855	15세 이상 관람 액션	140분	
6	터널	2016-08-10	57529484417	7120508	1105	113201	한국	(주)쇼박스	6.519021739	12세 이상 관람 드라마	126분	
7	인천상륙작전	2016-07-27	55101410303	7049643	1049	97019	한국	(주)씨제이이언	9.31379668	12세 이상 관람 전쟁, 드라마	110분	
8	럭키	2016-10-13	56444461256	6975290	1234	155977	한국	(주)쇼박스	7.87967581	15세 이상 관람 코미디	112분	
9	곡성	2016-05-12	55863520382	6879908	1485	130322	한국	이십세기폭스코리아	2.03868472	15세 이상 관람 미스터리, 스릴러	156분	
10	덕혜옹주	2016-08-03	44393828109	5599229	964	97312	한국	롯데쇼핑(주)롯데	6.537735849	12세 이상 관람 서사, 드라마	127분	
11	닥터 스트레인저	2016-10-20	47401000396	5446239	1503	102056	미국	월트디즈니컴퍼니	NA	12세 이상 관람 액션, 판타지, 스릴러	120분	
12	마스터	2016-12-21	39784641065	4935320	1501	70786	한국	(주)씨제이이언	4.165742794	15세 이상 관람 액션, 범죄	143분	
13	주토피아	2016-02-17	37096057742	4706158	735	131005	미국	월트디즈니컴퍼니	9.680327869	전체 관람가 애니메이션, 액션	108분	
14	신비한 동물사	2016-11-16	39300515858	4662534	1431	115076	미국	워너브라더스	NA	12세 이상 관람 모험, 가족, 판타지	142분	

데이터 전처리

[1] KOBIS 박스오피스 데이터

- running_time의 결측치는 중앙값으로 대체



```
from scipy import stats
running_time = df_cleaned_2['running_time']

# Shapiro-Wilk 검정
shapiro_test = stats.shapiro(running_time)

print('Shapiro-Wilk 검정 결과')
print('검정 통계량:', shapiro_test.statistic)
print('p-값:', shapiro_test.pvalue)

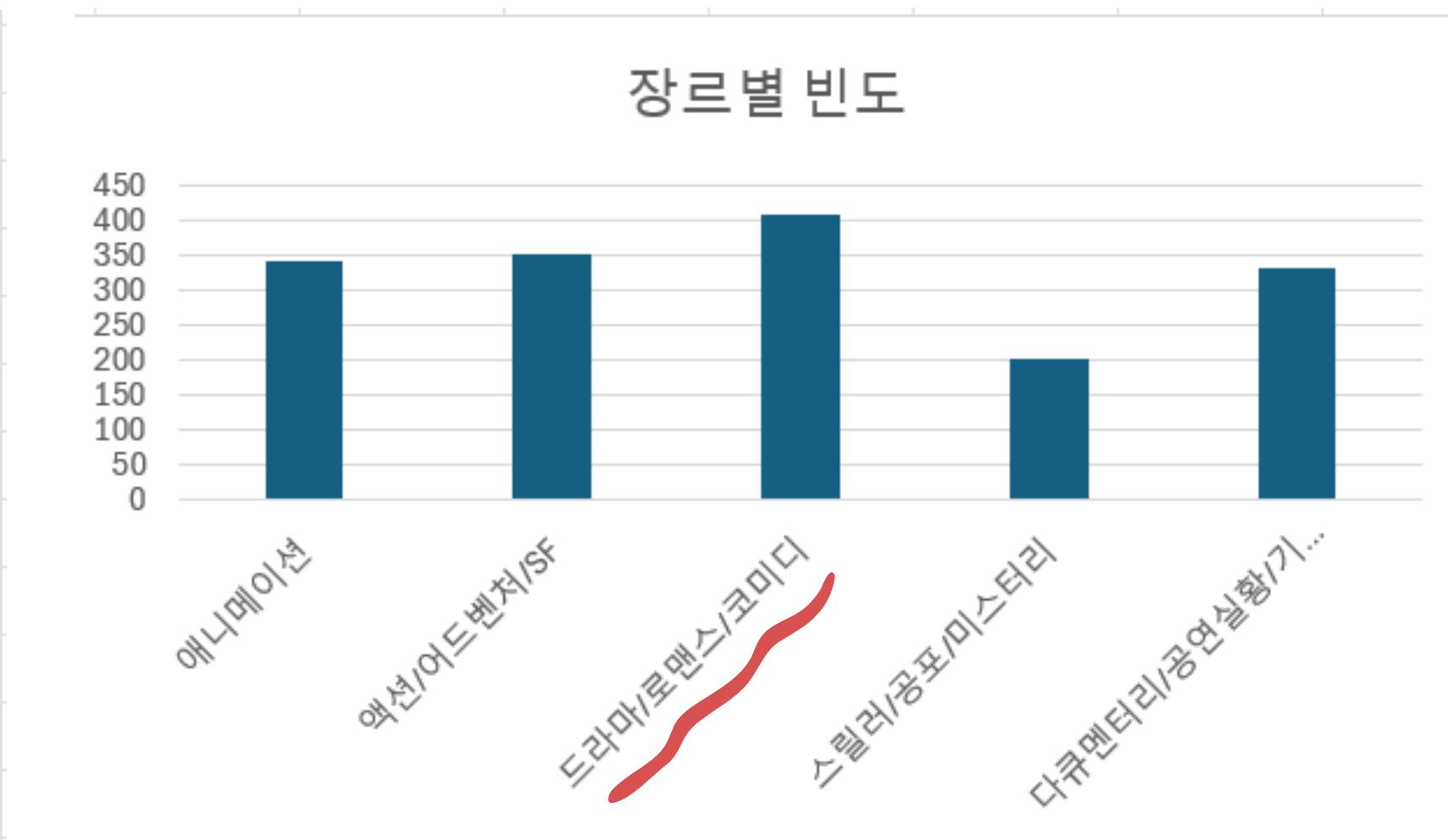
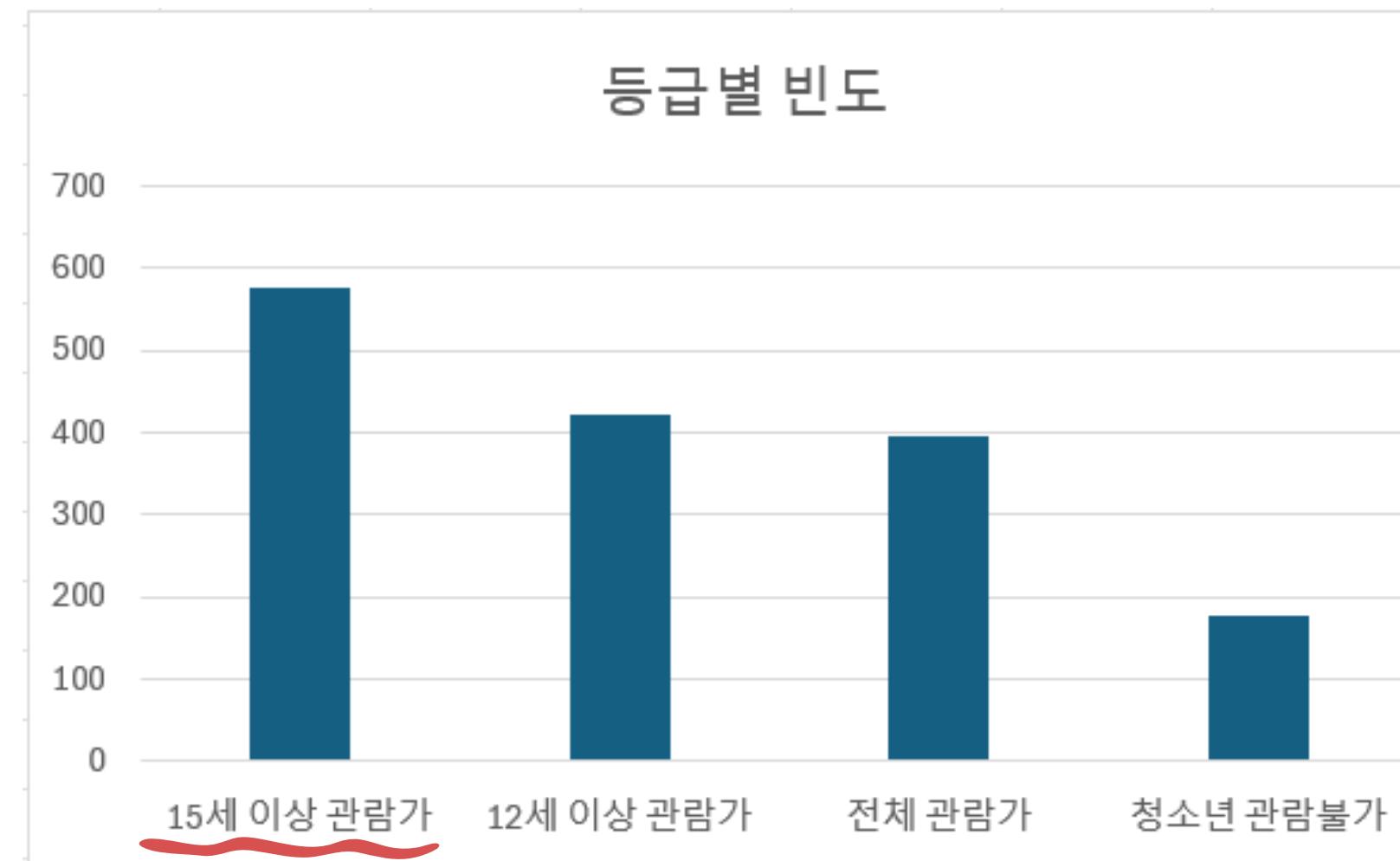
if shapiro_test.pvalue>0.05:
    print("정규분포를 따른다.")
else:
    print("running_time은 정규분포를 따르지 않는다.")

Shapiro-Wilk 검정 결과
검정 통계량: 0.9706499621483644
p-값: 1.7836088316064375e-17
running_time은 정규분포를 따르지 않는다.
```

데이터 전처리

[1] KOBIS 박스오피스 데이터

- 범주형 변수인 grade와 genre는 최빈값으로 대체



데이터 전처리

[1] KOBIS 박스오피스 데이터

- 2016~2024년 KOFIC 한국 영화산업 결산 - 배급사별 순위 참고
- 1~11로 순위 매김 (11위는 10위에 오르지 못한 기타 배급사들)

〈표 19〉 2023년 전체영화 배급사별 매출액 및 관객 점유율 순위 (매출액 기준)

(단위: 편 | 원 | 명 | %)

순위	배급사	상영 편수	매출액	매출액 점유율	관객 수	관객 점유율	주요 배급작
1	플러스엠	7.5	210,659,399,586	16.7%	21,562,879	17.2%	서울의 봄 / 범죄도시3 (공동배급)
2	디즈니	16	206,508,196,791	16.4%	19,211,268	15.4%	엘리멘탈 / 아바타: 물의 길 / 가디언즈 오브 갤럭시: Volume 3
3	롯데	16.5	145,200,785,933	11.5%	14,870,377	11.9%	마션 임파서블: 데드 레코닝 PART ONE / 콘크리트 유토피아
4	NEW	17.5	120,665,374,134	9.6%	12,127,702	9.7%	더 퍼스트 슬램덩크 / 밀수
5	UPI	19	99,665,143,968	7.9%	9,828,767	7.9%	오펜하이머 / 슈퍼 마리오 브라더스
6	CJ ENM	14	79,491,517,560	6.3%	8,095,294	6.5%	천박사 퇴마 연구소: 설경의 비밀 / 영웅
7	쇼박스	5	72,153,015,633	5.7%	7,149,187	5.7%	스즈메의 문단속
8	에이비오	0.5	52,343,868,886	4.1%	5,341,407	4.3%	범죄도시3 (공동배급)
9	마인드마크	4	37,153,885,887	2.9%	3,821,255	3.1%	30일
10	워너	11	27,122,452,561	2.2%	2,640,283	2.1%	아쿠아맨과 로스트 킹덤
기타		2,435	210,446,610,806	16.7%	20,487,510	16.4%	
합계		2,546	1,261,410,251,744	100.0%	125,135,927	100.0%	

배급사	rating_mean	grade	genre	running_time	배급사 순위
(주)넥스트엔터테인먼트월드(NEW)	2.922297297	15세 이상 관람가	액션/어드벤처/SF	118분	6
(주)쇼박스	4.071428571	15세 이상 관람가	스릴러/공포/미스터리	126분	2
월트디즈니컴퍼니코리아 유한책임회사	5.835497835	12세 이상 관람가	액션/어드벤처/SF	147분	3
워너브러더스 코리아(주)	9.020289855	15세 이상 관람가	액션/어드벤처/SF	140분	4
(주)쇼박스	6.519021739	12세 이상 관람가	드라마/로맨스/코미디	126분	2
(주)씨제이이엔엠	9.31379668	12세 이상 관람가	다큐멘터리/공연실황/기타	110분	1
(주)쇼박스	7.87967581	15세 이상 관람가	드라마/로맨스/코미디	112분	2

데이터 전처리

[2] 인스타그램 데이터

- ‘real_tags’ 컬럼 생성: 게시글에서 #tag 들만 추출해 새로운 컬럼 생성
- ‘num_tags’ 컬럼: #tag 개수

	account	post_num	post_date	text	num_likes	real_tags	year	month	year_month	num_tags
24	cgv_korea	3200	2015-07-24	['cgv_korea', '미니언이 몰려온다...科技进步', '지금 CGV에서 절찬리...']	268	[#미니언즈, #minions, #미니언즈콤보, #미니언즈피규어, #미니언, #CG...]	2015	7	2015-07	7
40	cgv_korea	3184	2016-04-05	['cgv_korea', '봄 바람 휘날리며~▶▶▶ 흘날리는 벚꽃 잎이~🌸', '까향!...']	555	[#봄꽃영화제, #벚꽃축제커플조심, #뒤에들이사귀는듯, #파코니는괜찮나, #설레지마...]	2016	4	2016-04	21
46	cgv_korea	3178	2016-04-20	['cgv_korea', '♥파코니 공고♥앵그리버드 사생팬 모집', 'CGV에 #앵...']	426	[#앵그리버드, #CGV, #CulturePlayNews, #앵그리버드, #Angri...]	2016	4	2016-04	10
65	cgv_korea	3159	2016-05-20	['cgv_korea', '프렌들.. 어제 #세익스피어톡 왜 안왔어...?😊', '...']	378	[#세익스피어톡, #올리비아핫세, #로미오와줄리엣, #이상용평론가, #스크린문학전,...]	2016	5	2016-05	17
75	cgv_korea	3149	2016-06-09	['cgv_korea', '열심히 일한 당신, 떠나라 IMAX로~! 🎉', '파코...']	430	[#아이맥스여권, #정글북, #워크래프트, #CGV, #CulturePlayNews...]	2016	6	2016-06	15

데이터 전처리

[2] 인스타그램 데이터

- KOBIS 박스 오피스 데이터에서 영화 이름만 추출
- ‘is_movie’ 컬럼: 영화 제목이 #tag에 포함되었을 경우 is_movie = 1, 아닐 경우 is_movie = 0
- ‘movie_name’ 컬럼: 영화 제목이 포함되었을 경우 해당하는 영화 제목

['부산행', '검사외전', '캡틴아메리카:시빌워', '밀정', '터널', '인천상륙작전', '럭키', '곡성', '덕혜옹주',

	account	post_num	post_date	text	num_likes	real_tags	year	month	year_month	num_tags	is_movie	movie_name
24	cgv_korea	3200	2015-07-24	['cgv_korea', '미니언이 몰려온다...科技进步', '지금 CGV에서 절찬 리...', ...]	268	[#미니언즈, #minions, #미니언즈콤보, #미니언즈피규어, #미니언, #CG...]	2015	7	2015-07	7	1	미니언즈
40	cgv_korea	3184	2016-04-05	['cgv_korea', '봄 바람 휘날리며~▶ 흘날리는 벚꽃 잎이~🌸', '꺄 핳!', ...]	555	[#봄꽃영화제, #벚꽃축제커플조심, #뒤에돌이사귀는듯, #파코니는괜자 나, #설레지마...]	2016	4	2016-04	21	1	주토피아
46	cgv_korea	3178	2016-04-20	['cgv_korea', '♥파코니 공고♥앵그리버드 사생팬 모집', 'CGV에 #앵...', ...]	426	[#앵그리버드, #CGV, #CulturePlayNews, #앵그리버드, #Angr...]	2016	4	2016-04	10	1	앵그리버드더무 비
65	cgv_korea	3159	2016-05-20	['cgv_korea', '프렌들.. 어제 #세익스피어톡 왜 안왔어...?😊', '...']	378	[#세익스피어톡, #올리비아핫세, #로미오와줄리엣, #이상용평론가, # 스크린문화학전,...]	2016	5	2016-05	17	1	로미오와줄리엣
75	cgv_korea	3149	2016-06-09	['cgv_korea', '열심히 일한 당신, 떠나라 IMAX로~ kino, '파코...']	430	[#아이맥스여권, #정글북, #워크래프트, #CGV, #CulturePlayNews...]	2016	6	2016-06	15	1	정글북

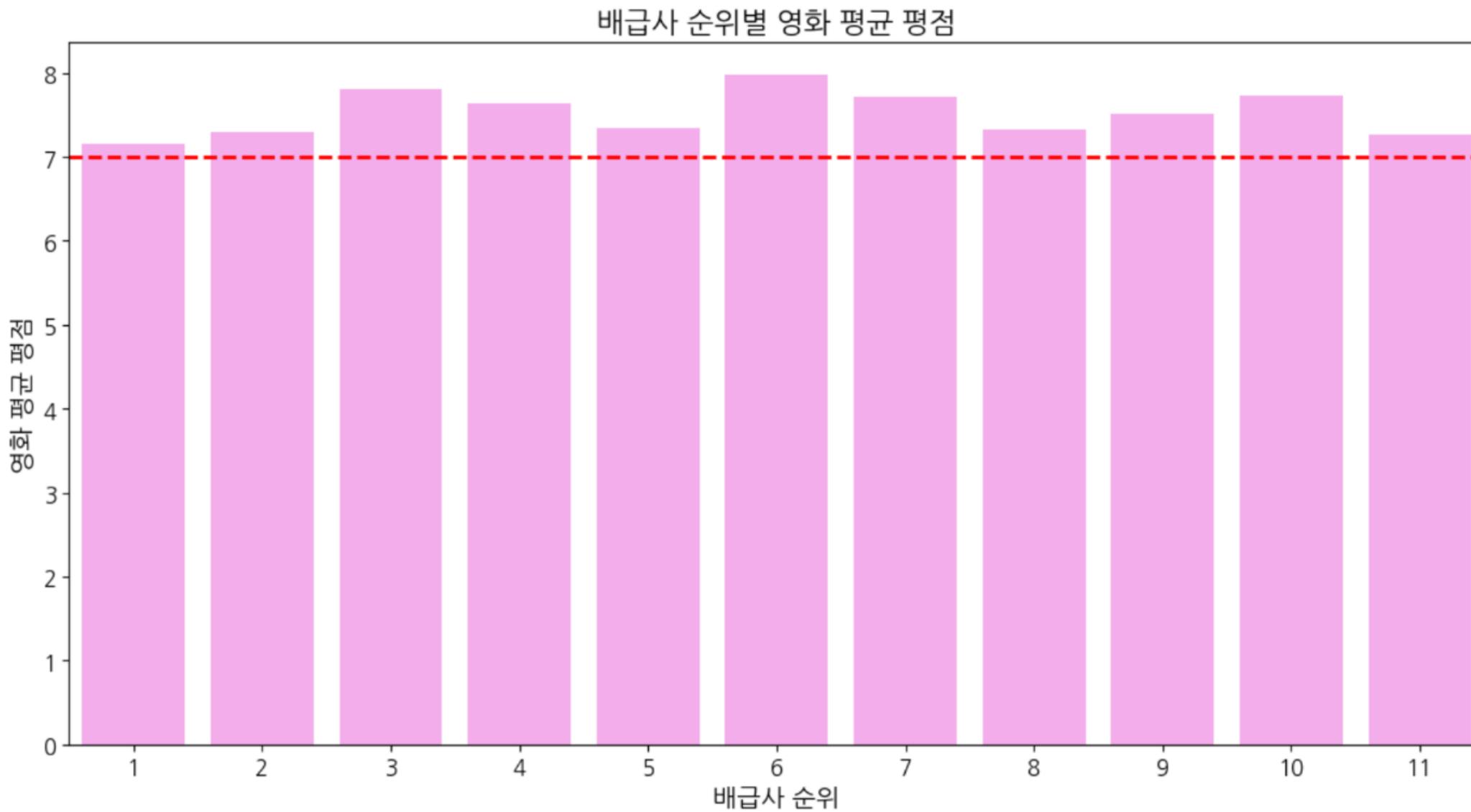
3

데이터 시각화



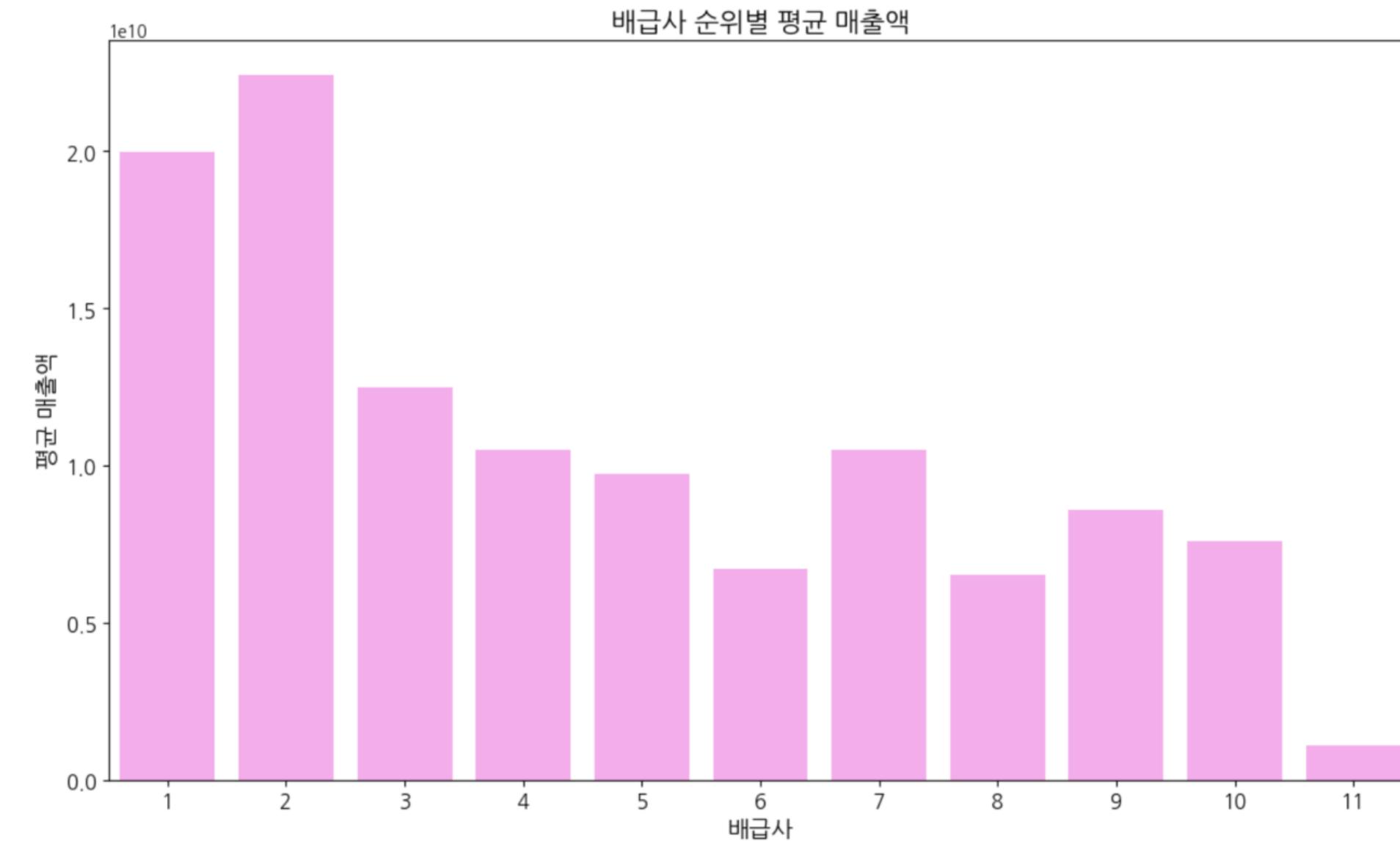
배급사 순위별 영화 평균 평점

- 배급사 평균 평점: 모두 7점 이상
- 배급사 순위 vs 영화 평균 평점:
양의 상관관계 X

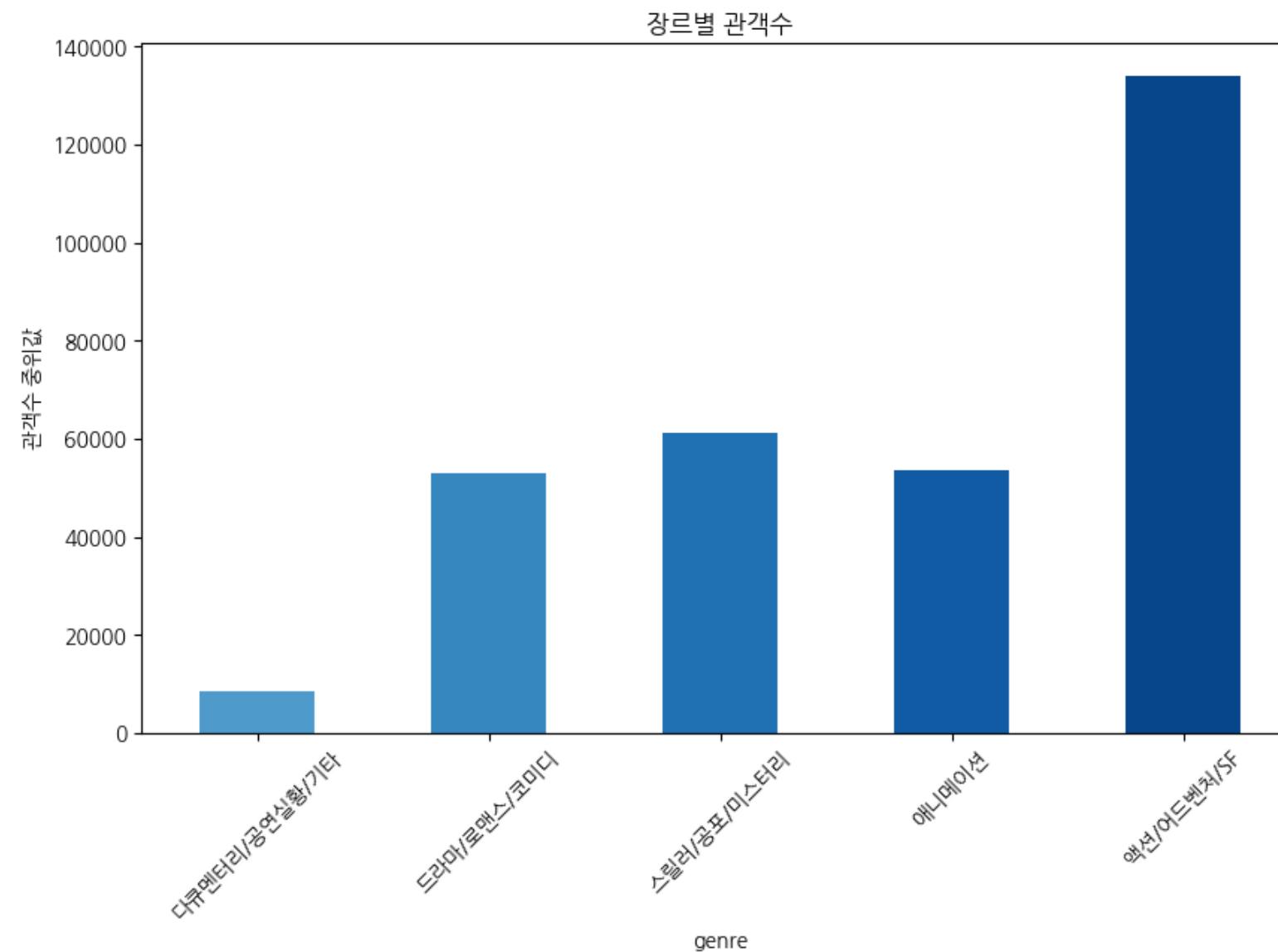


배급사 순위별 평균 매출액

- 1,2순위 배급사의 높은 평균 매출액
- 11위 배급사의 현저히 낮은 평균 매출액



장르별 관객수 분포



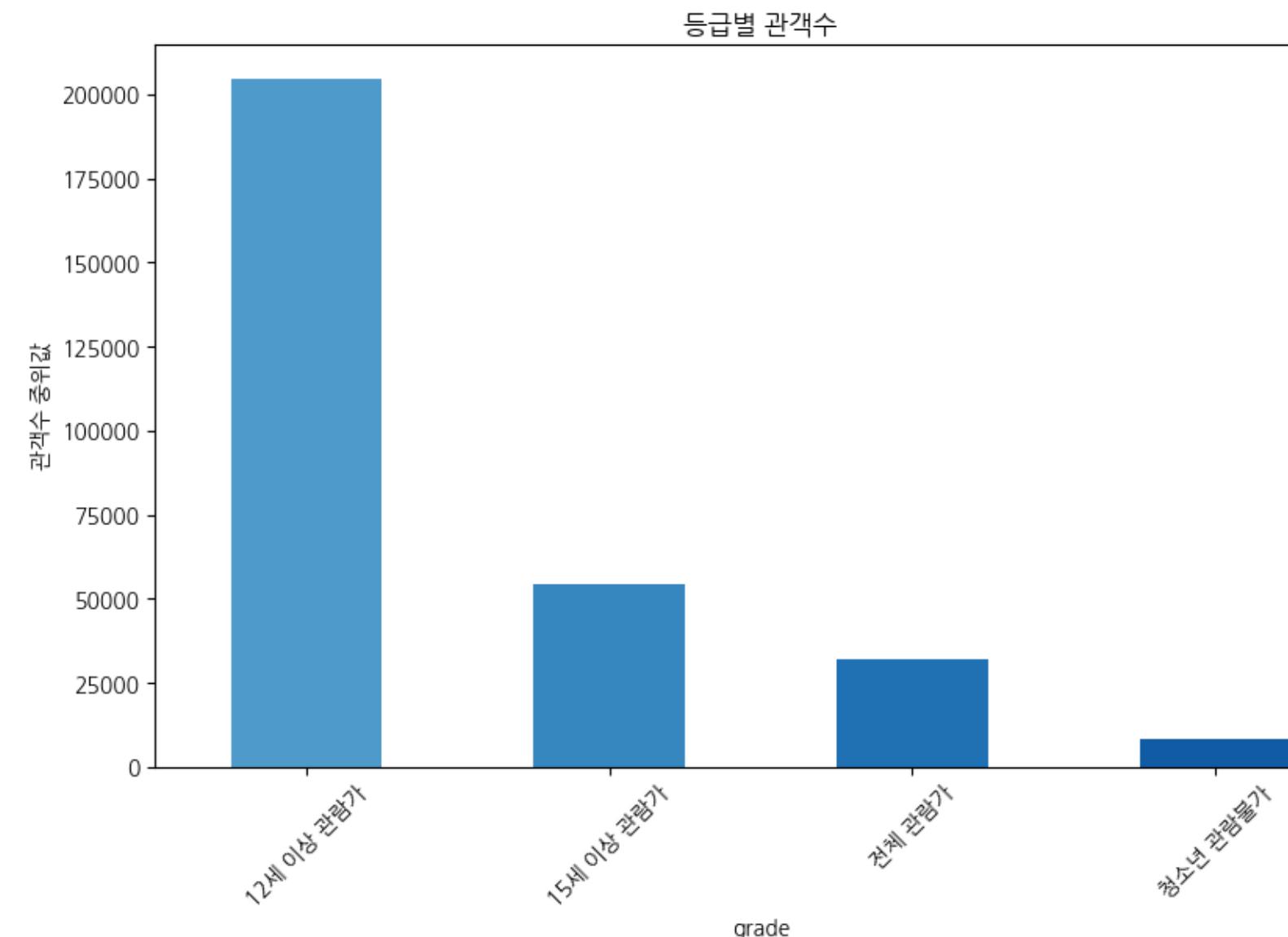
Kruskal-Wallis Test

통계량=66.713, p-값=0.000

장르별 관객수의 중위값에 차이가 있다.

- ✓ 애니메이션 관객수 중위값 < 액션/어드벤처/SF 관객수 중위값
- ✓ 다큐멘터리/공연실황/기타 < 다른 모든 장르
- ✓ 애니메이션 vs 드라마/로맨스/코미디 유의미 X
- ✓ 애니메이션 vs 드라마/로맨스/코미디 유의미 X

등급별 관객수 분포



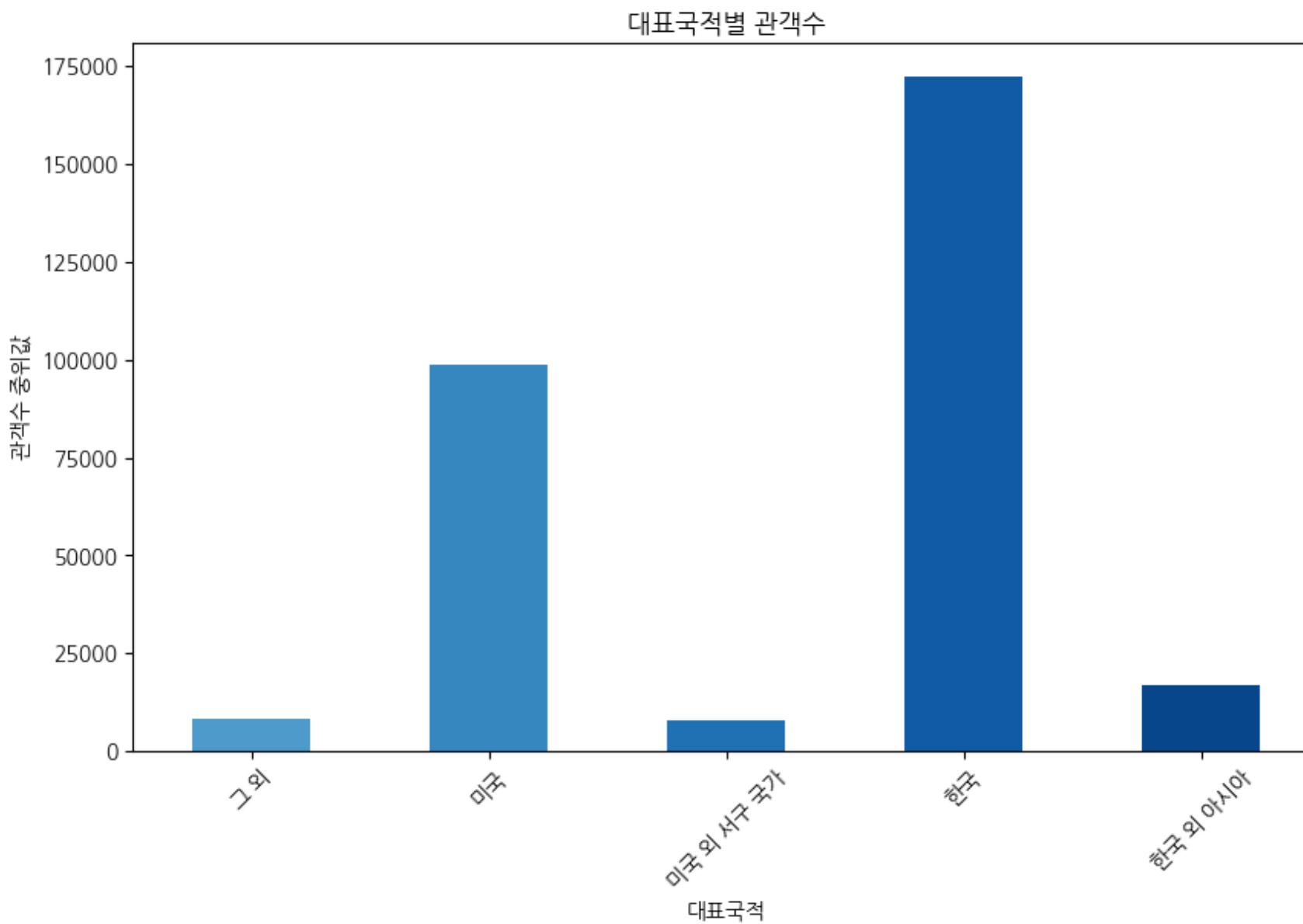
Kruskal-Wallis Test

통계량=81.442, p-값=0.000

등급별 관객수의 중위값에 차이가 있다.

- ✓ 청불 관객수 중위값 < 전체 관람가 관객수 중위값
- ✓ 다른 연령대의 영화 < 12세 이상 관람가 영화
- ✓ 캡틴 아메리카: 시빌 워 & 서울의 봄

대표국적별 관객수 분포



Kruskal-Wallis Test

통계량=177.541 , p-값=0.000

대표국적별 관객수의 중위값에 차이가 있다.

- ✓ 미국 관객수 중위값 < 한국 관객수 중위값
- ✓ 미국 외 서구 국가 < 한국 외 아시아
- ✓ 그 외 vs 미국 외 서구 국가 유의미 X

4

Regression



regression

1

**What is
Quantile
regression**

2

**Why Quantile
regression**

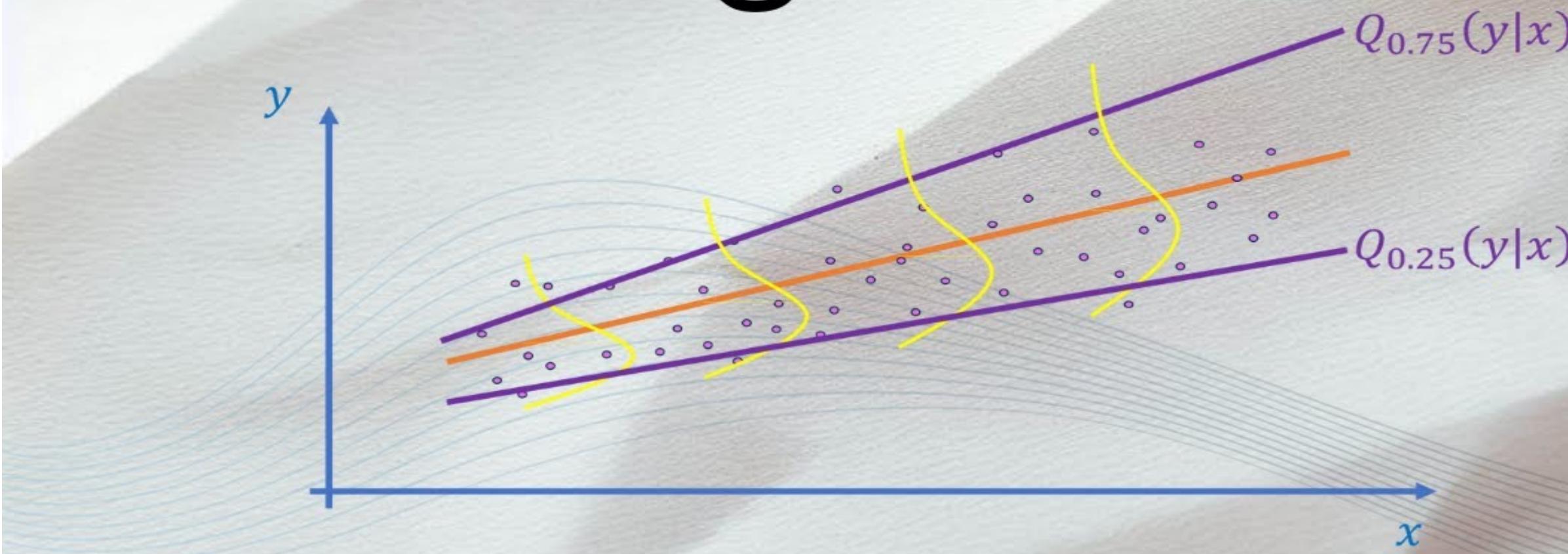
3

Result

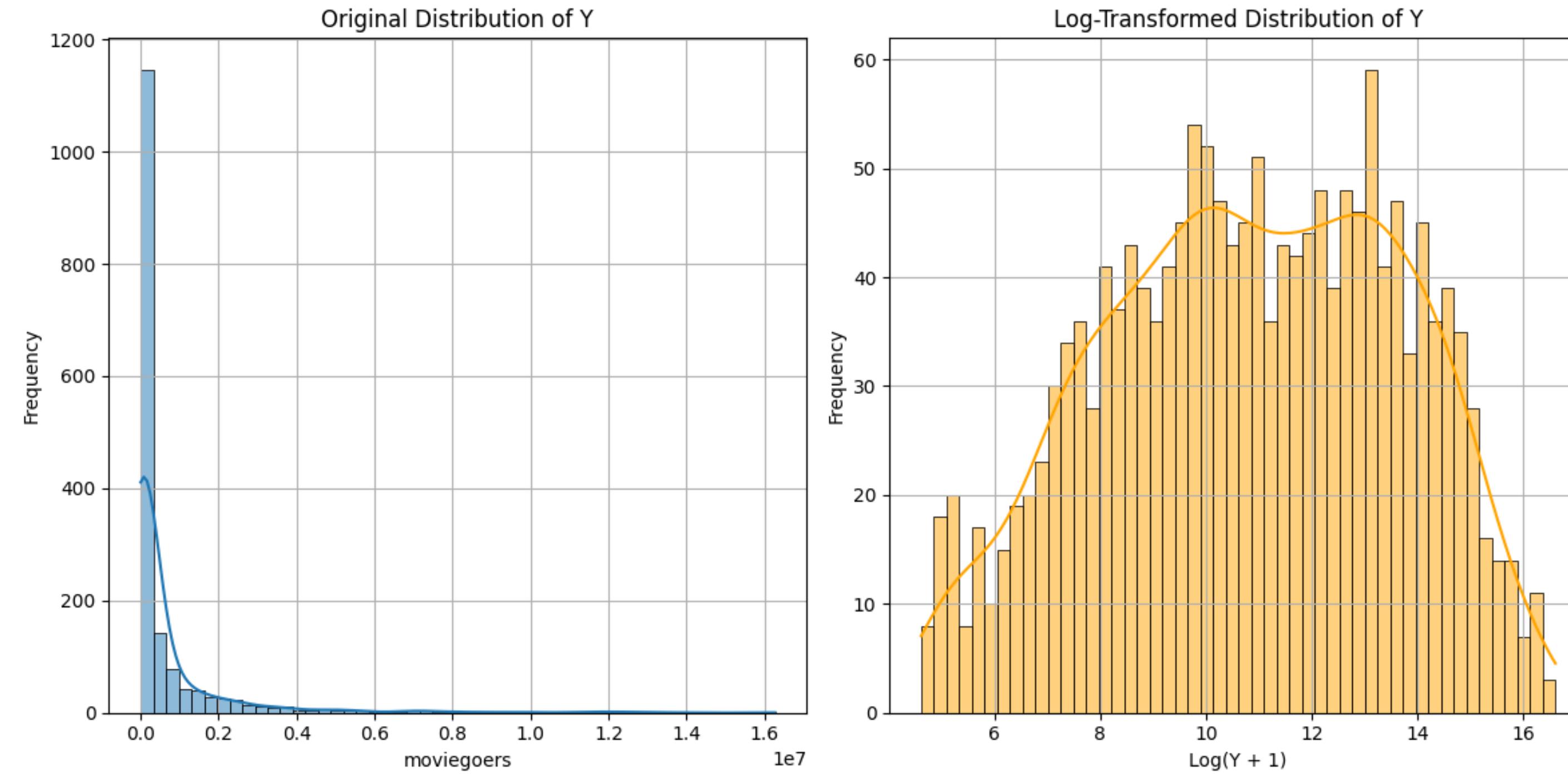
- running_time
- rating_mean

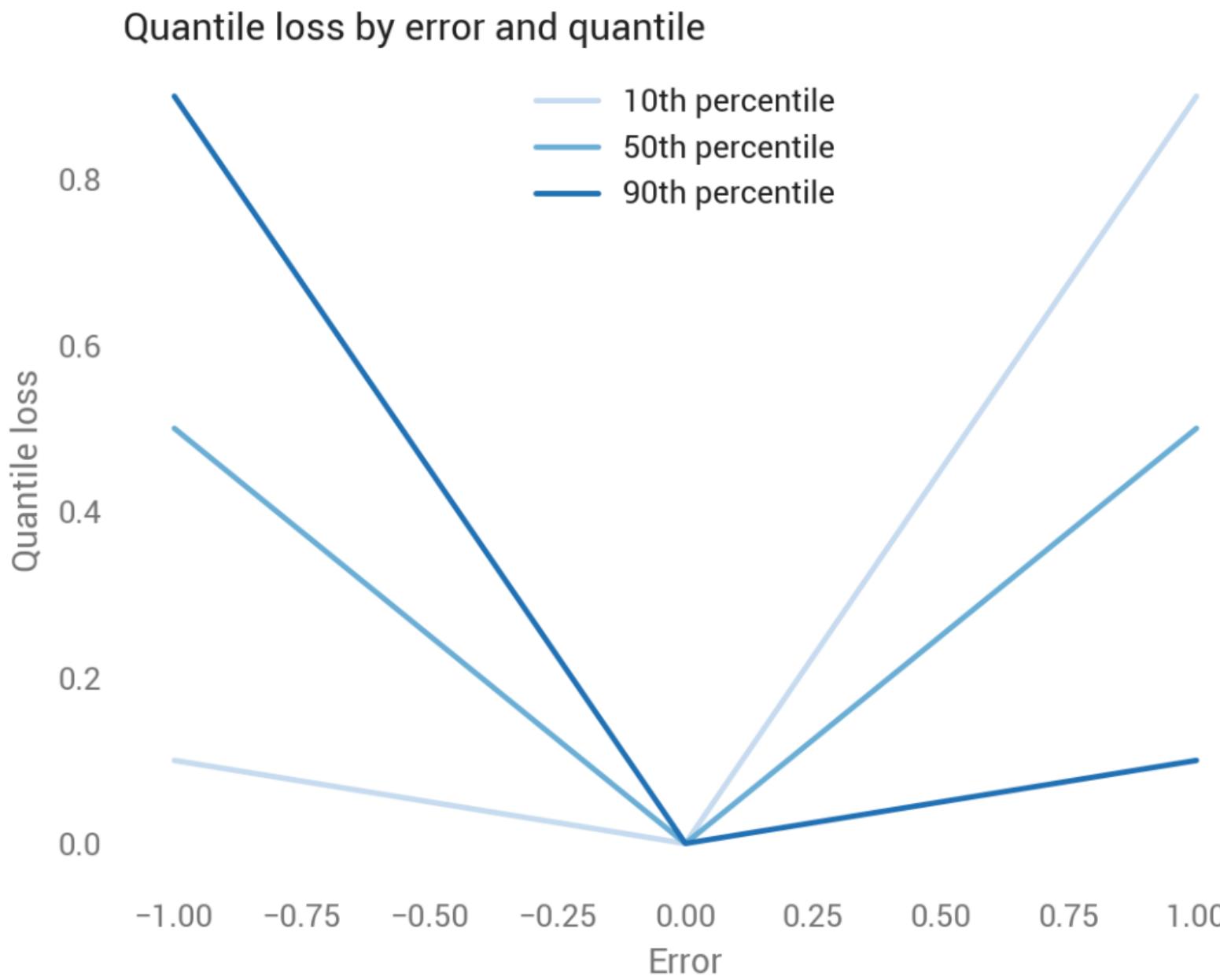
4

Linear vs. Quantile Regression



4

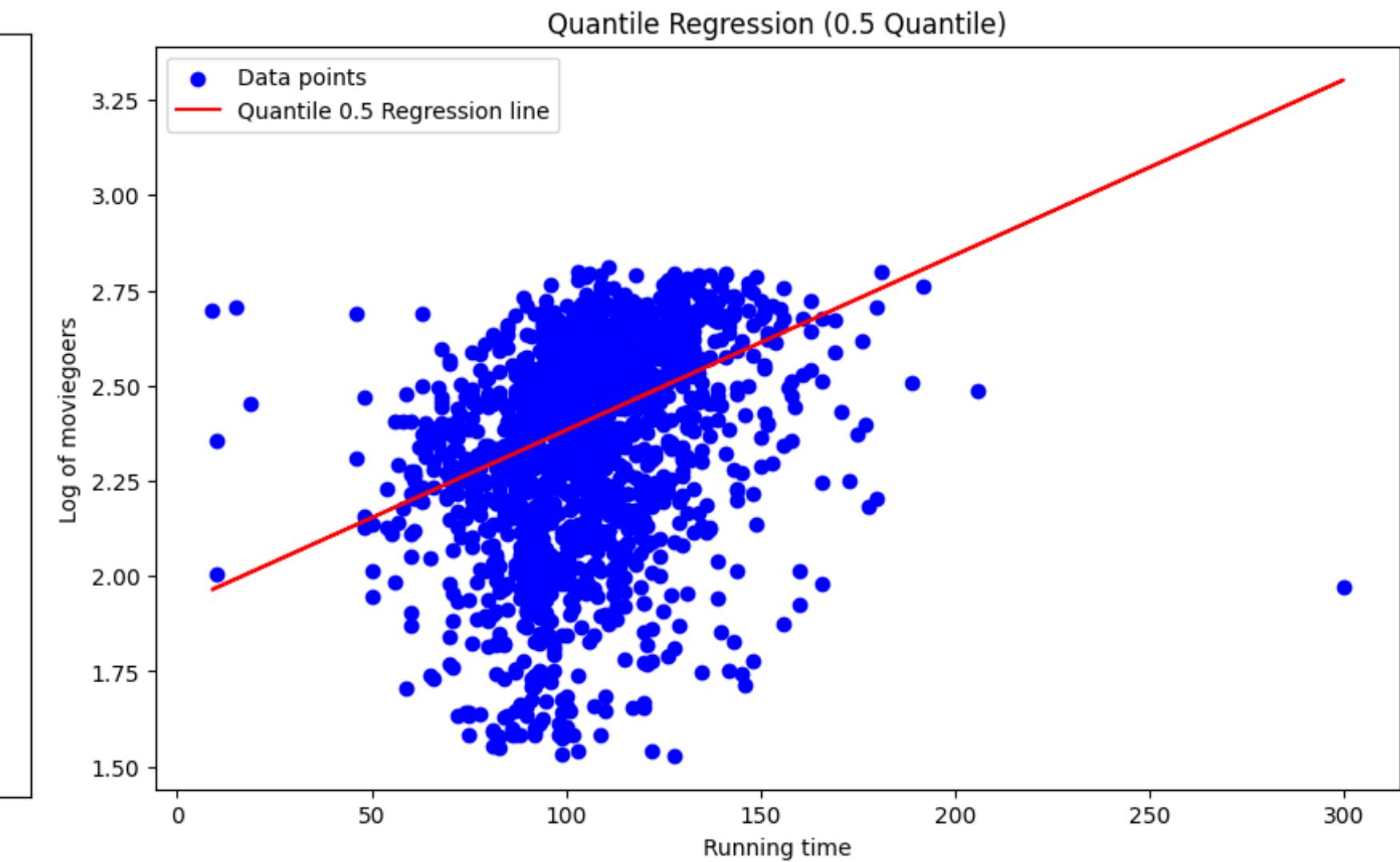
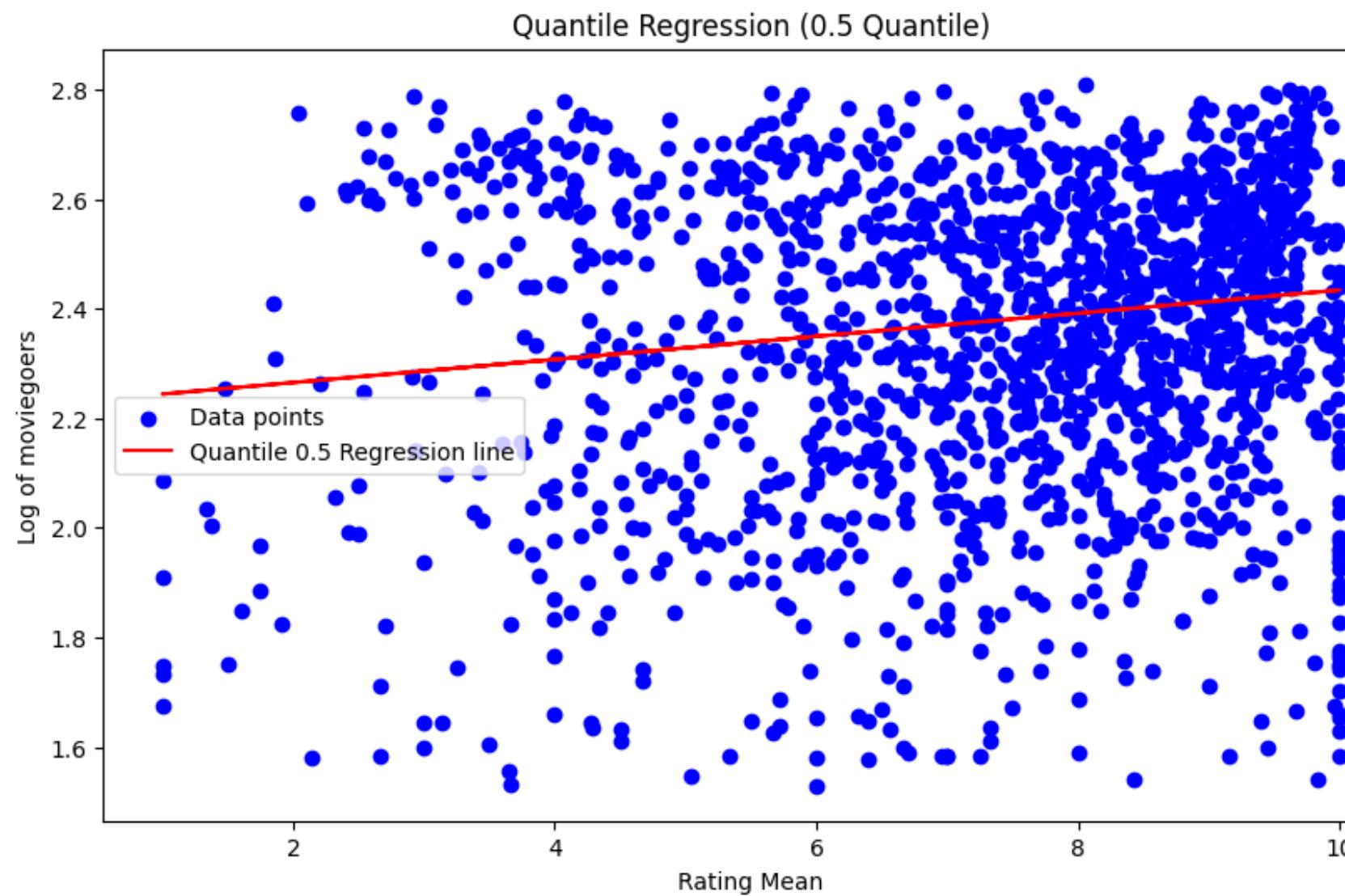




$$\ell(y, \hat{y}) = \begin{cases} \alpha \cdot (y - \hat{y}) & , \hat{y} \leq y \\ (1 - \alpha) \cdot (\hat{y} - y) & , \hat{y} > y \end{cases}$$

- quadratic \rightarrow linear
- different weight

result



둘 모두 양의 상관관계를 가지고 있다

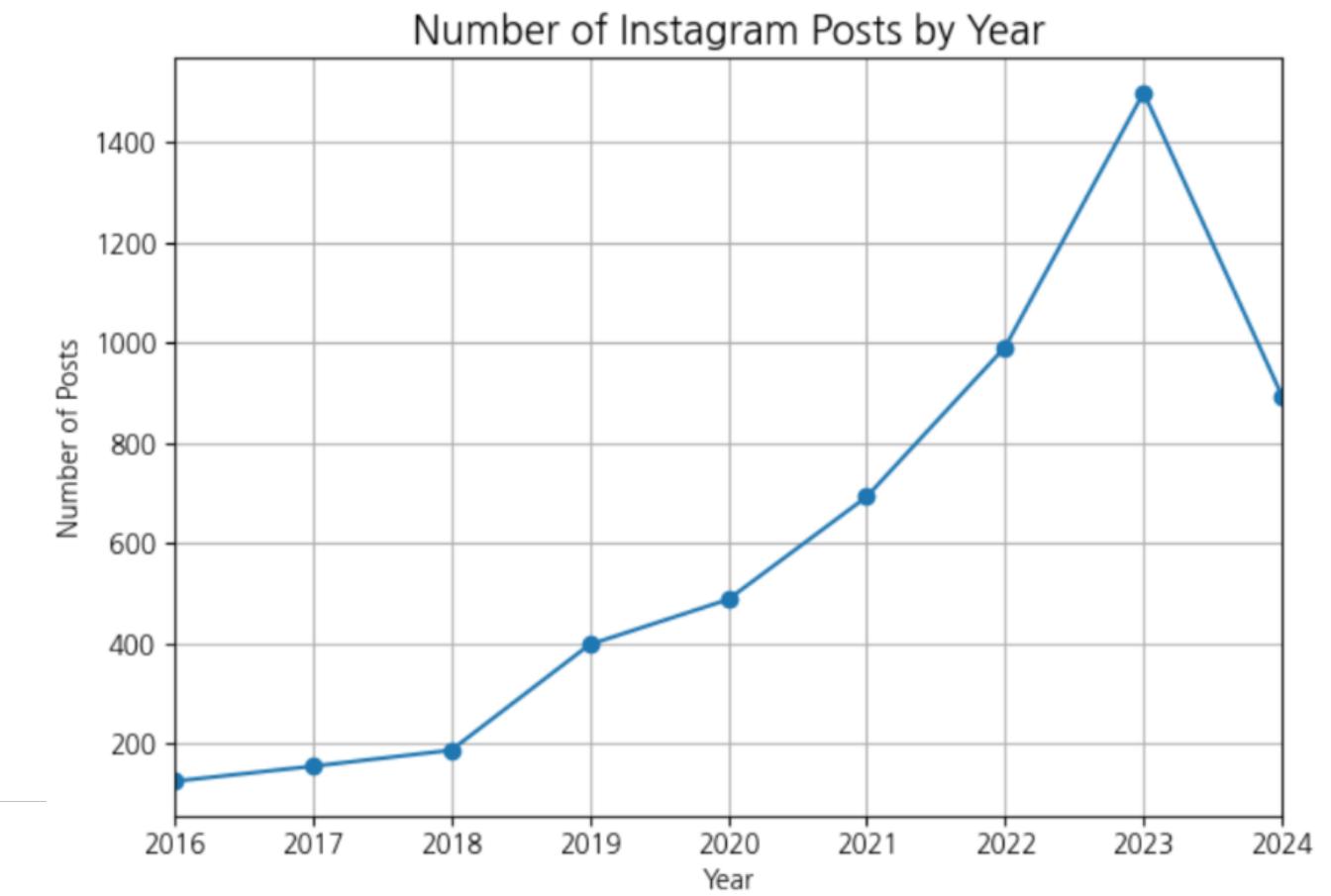
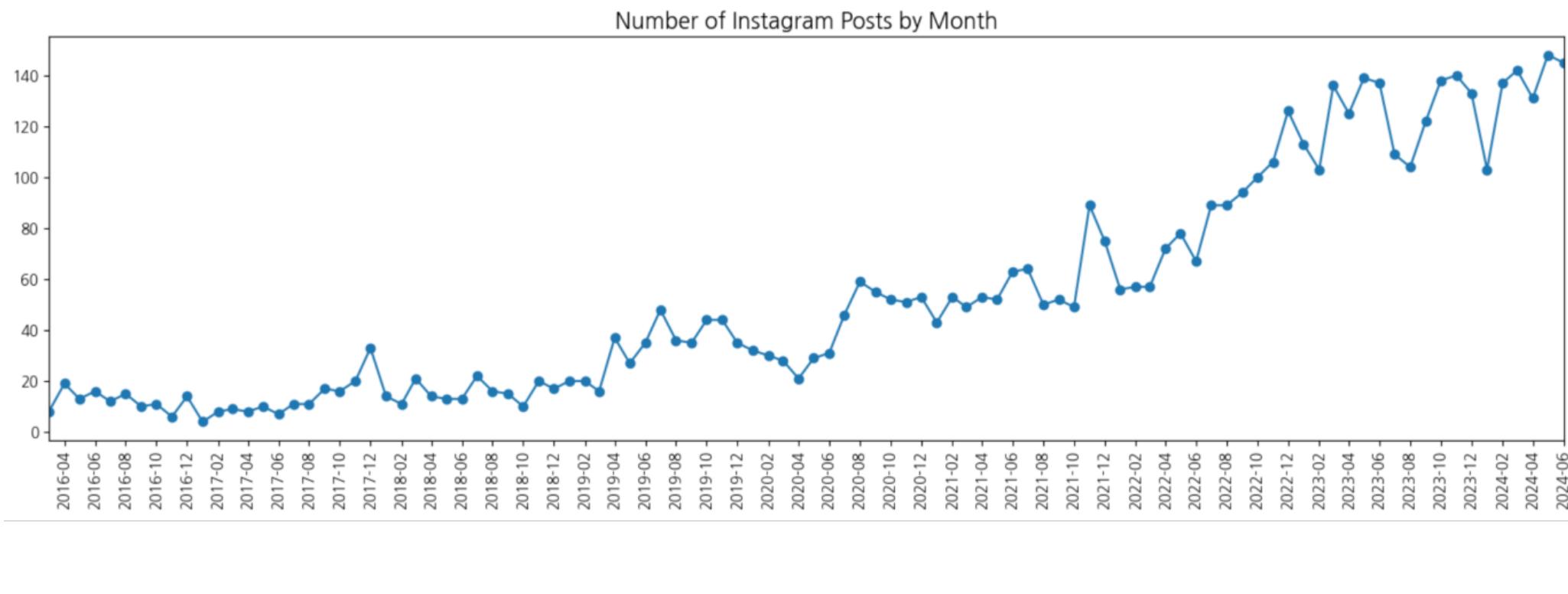
5

인스타그램을 활용한 SNS 분석



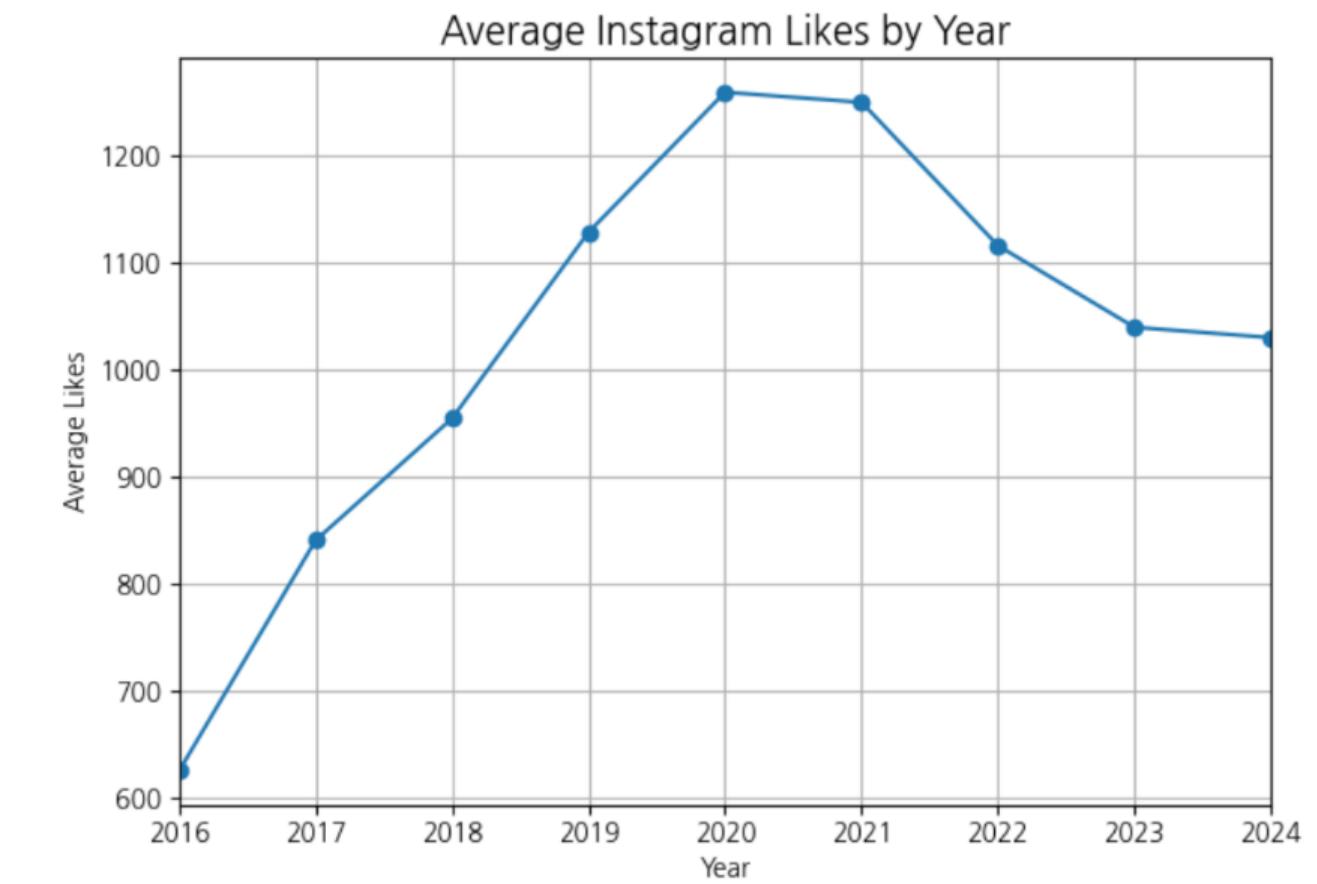
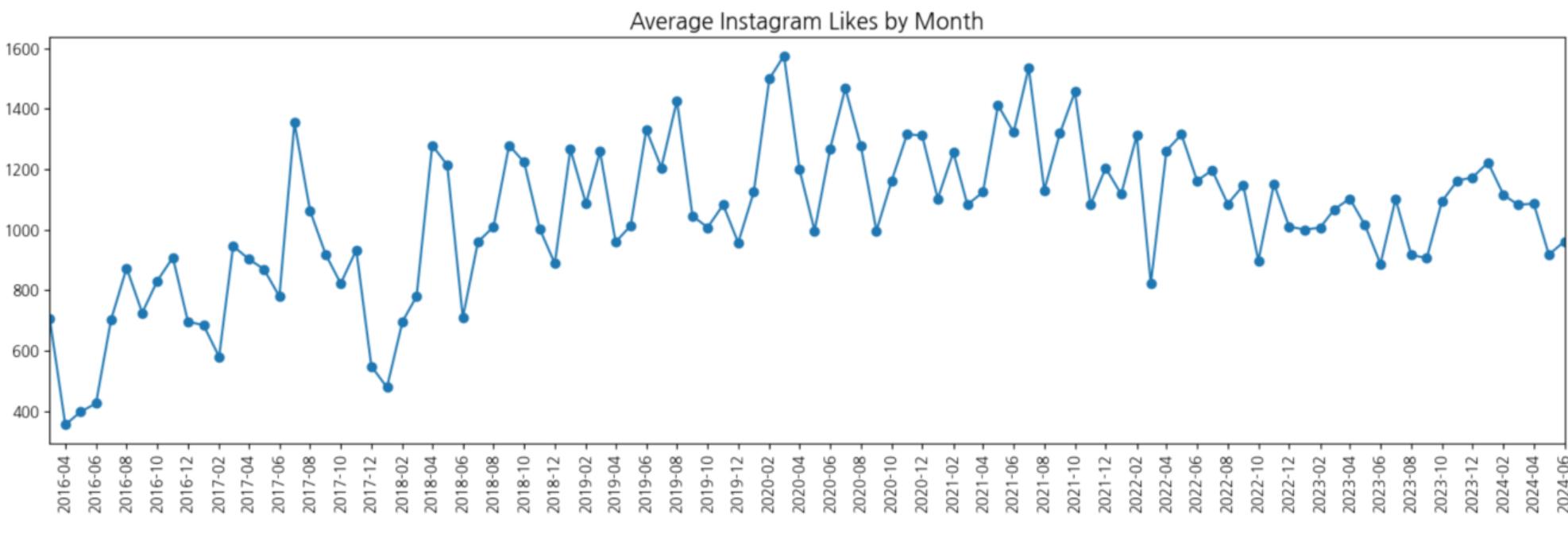
연도별 인스타그램 게시물 수 추이

- 게시물 수의 꾸준한 증가 → 홍보활동의 중요성



연도별 인스타그램 평균 좋아요 수 추이

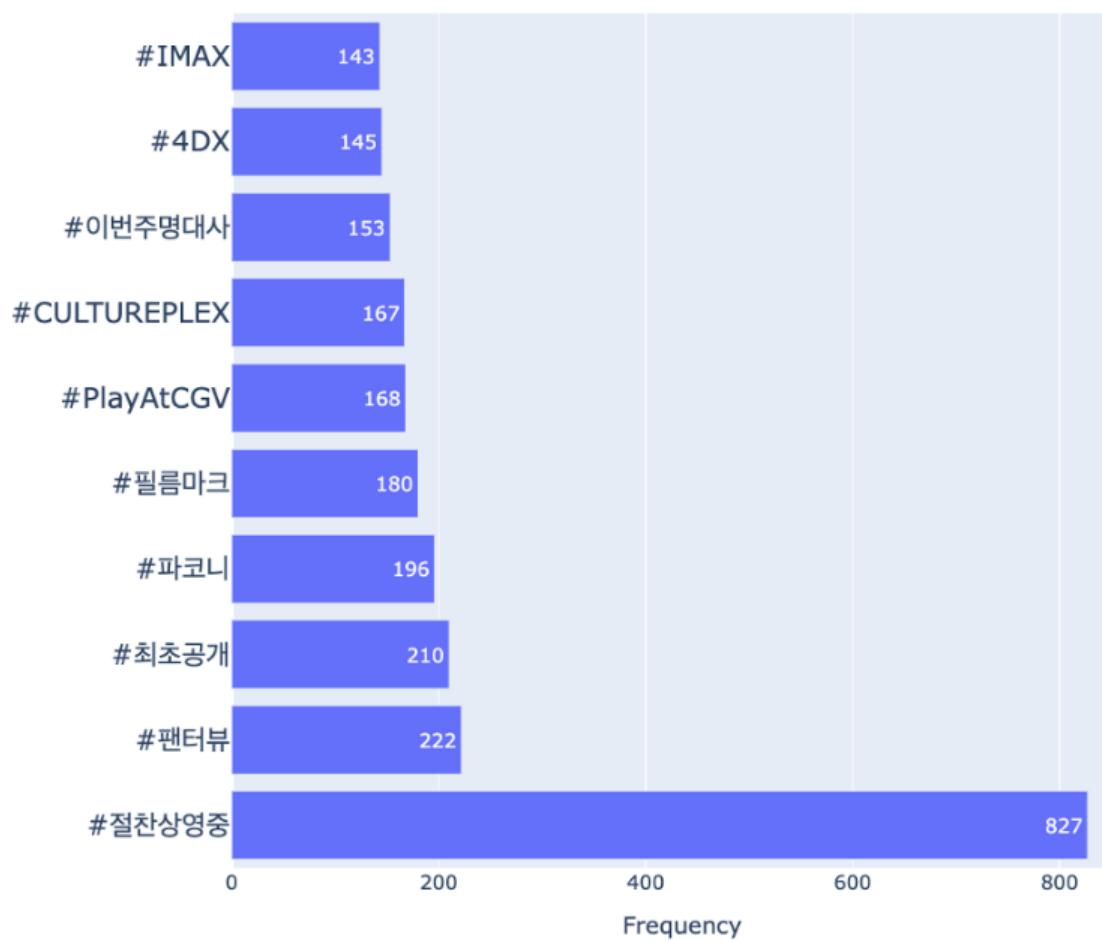
- 연도별로 올라오는 게시물의 수는 증가하고 있음에도 불구하고
- 평균 좋아요 수는 2020년까지는 가파르게 상승하다, 2020년 부터 천천히 하락하고 있다.



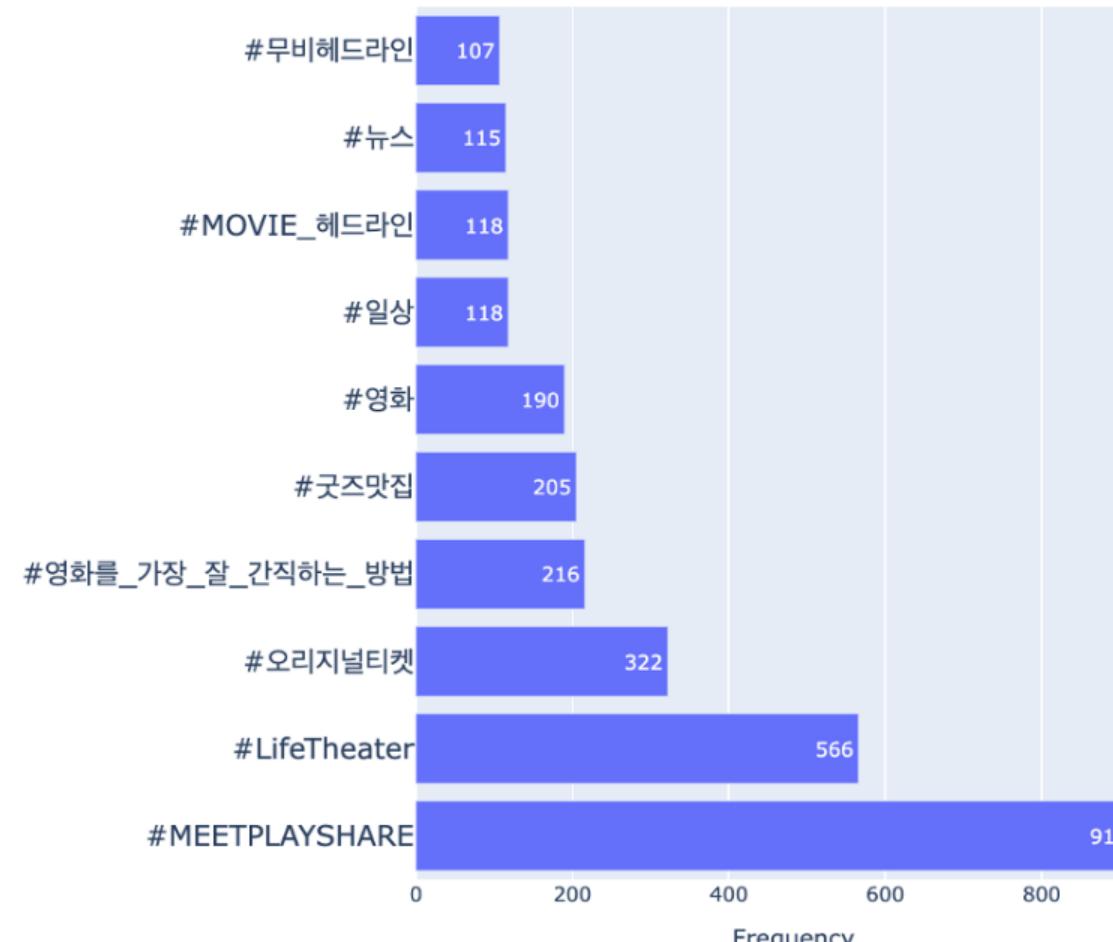
많이 쓰인 해시태그

- 공식 캐릭터, 브랜드 홍보 콘텐츠, 이벤트, 영화 홍보 콘텐츠

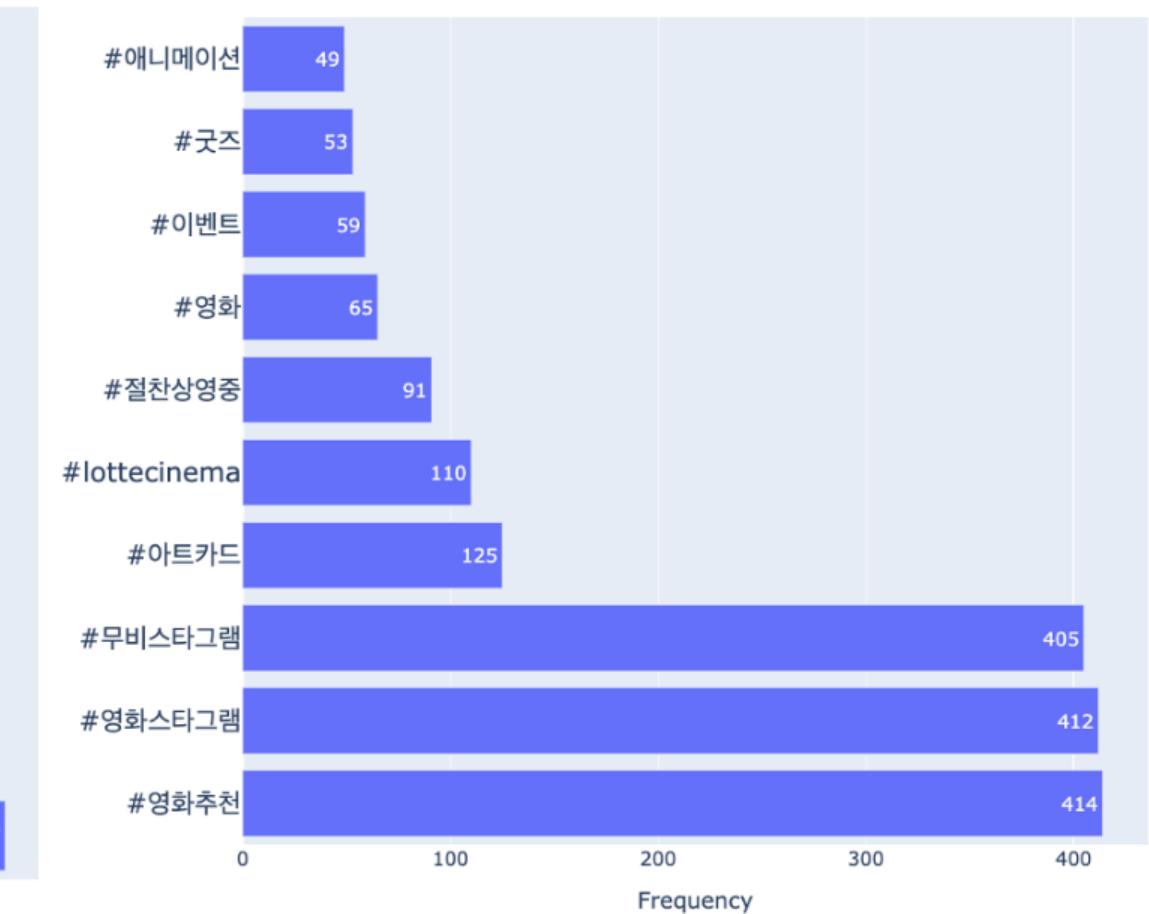
Top 10 Hashtags by Frequency @cgv_korea



Top 10 Hashtags by Frequency @megaboxon

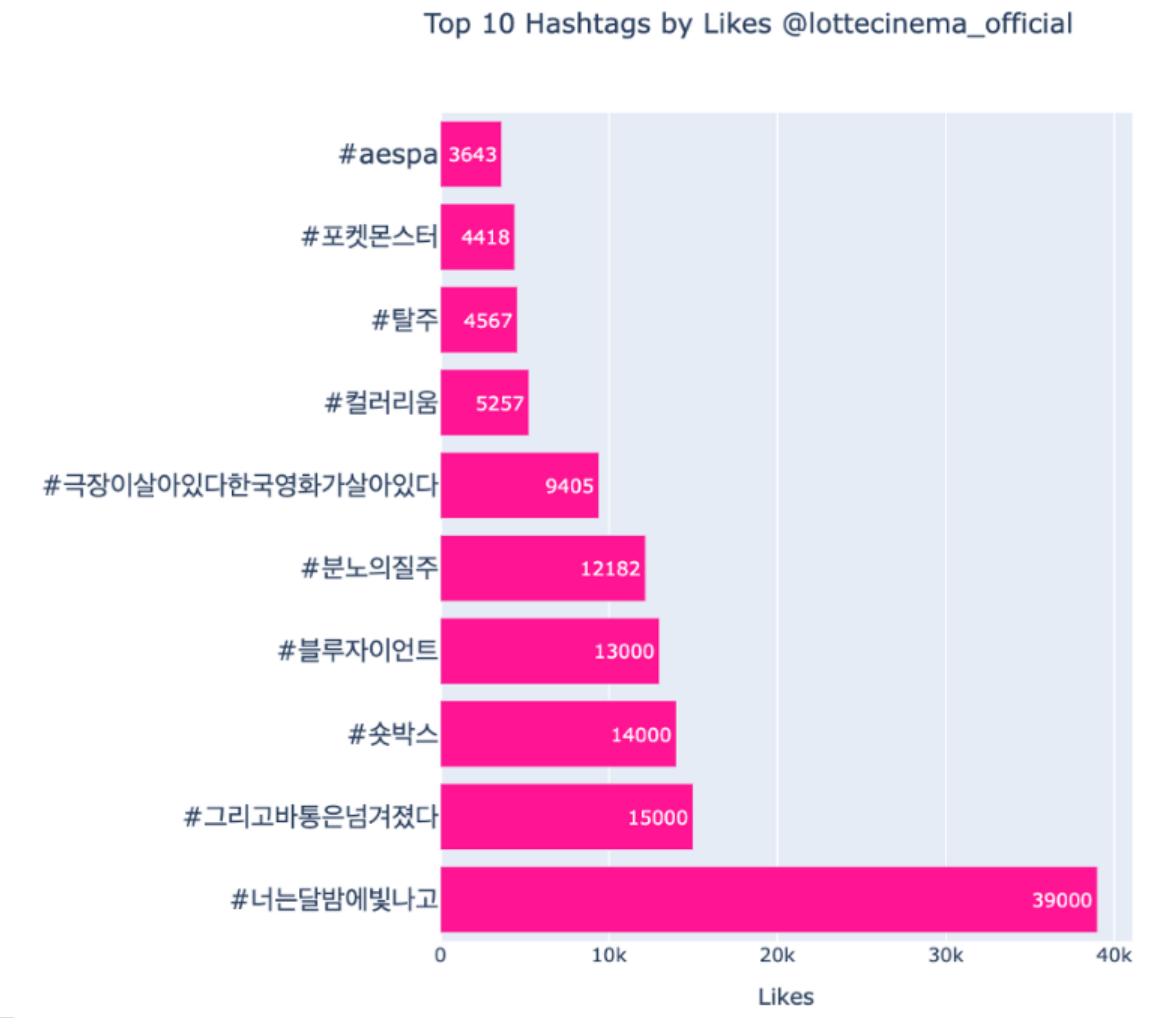
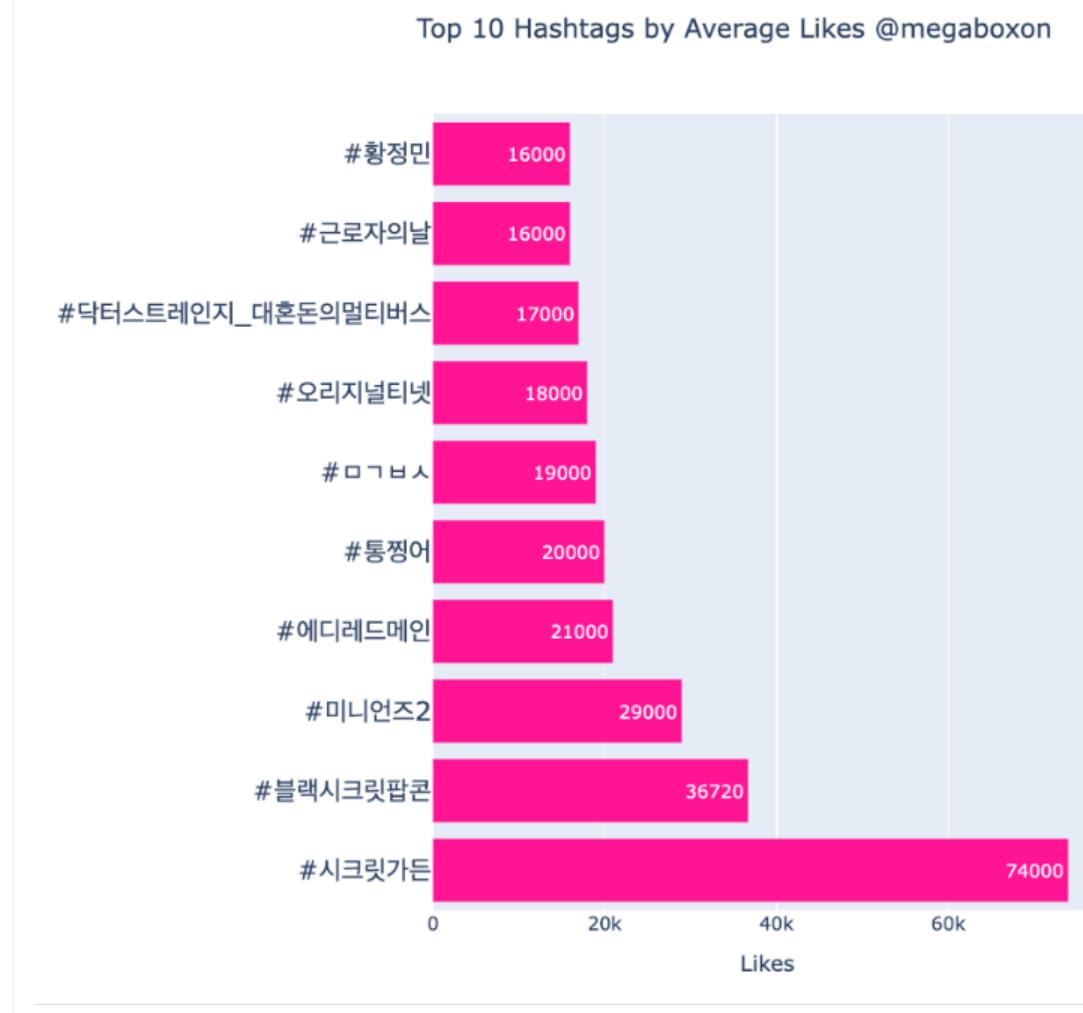
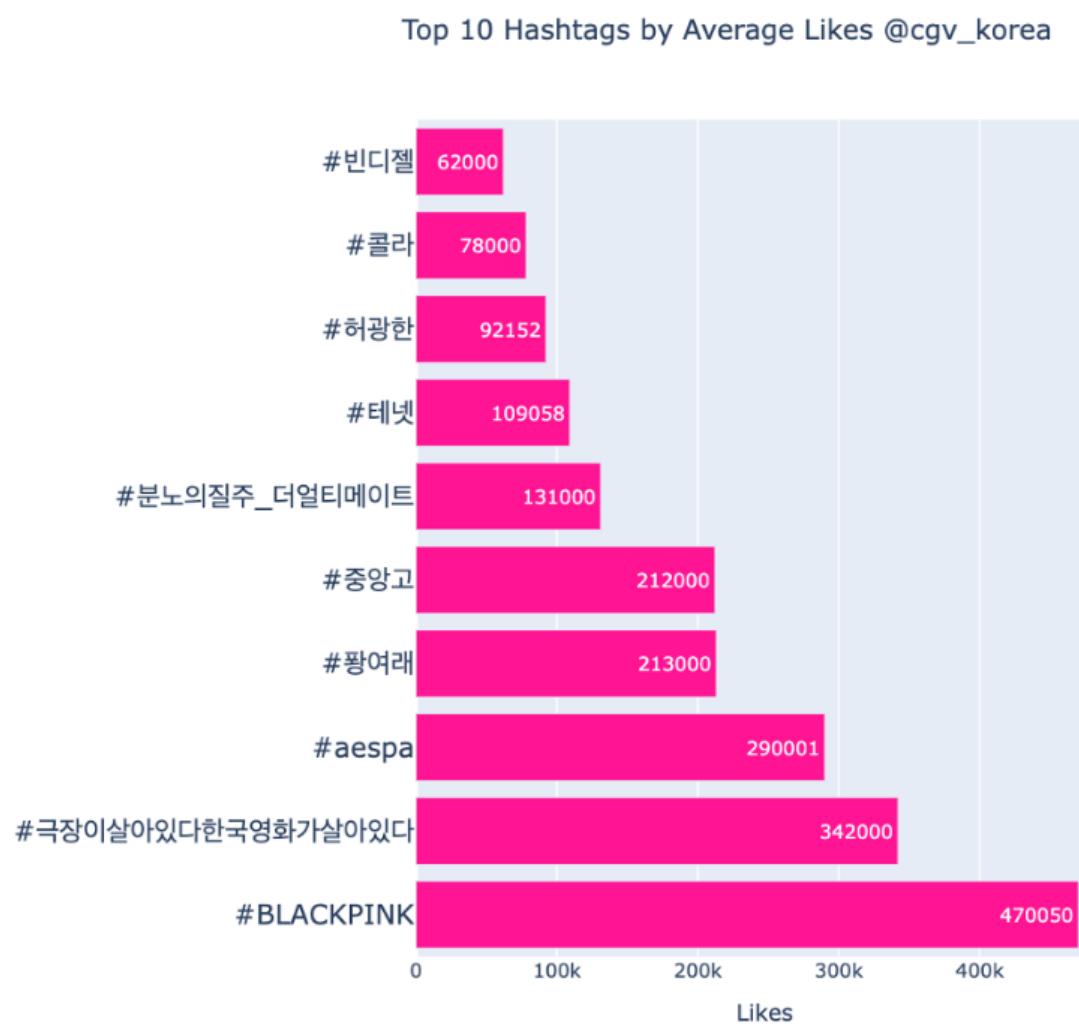


Top 10 Hashtags by Frequency @lottecinema_official

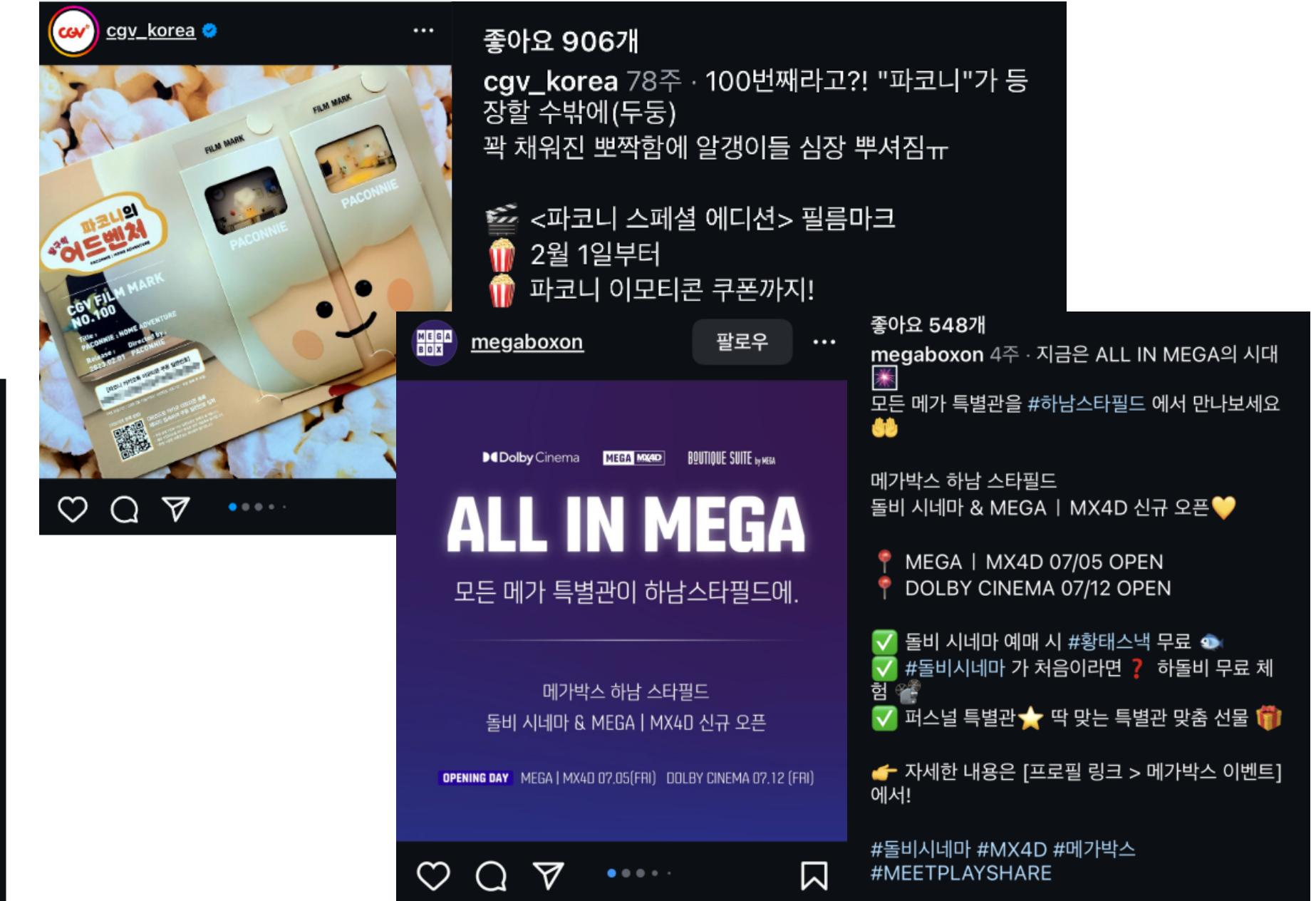
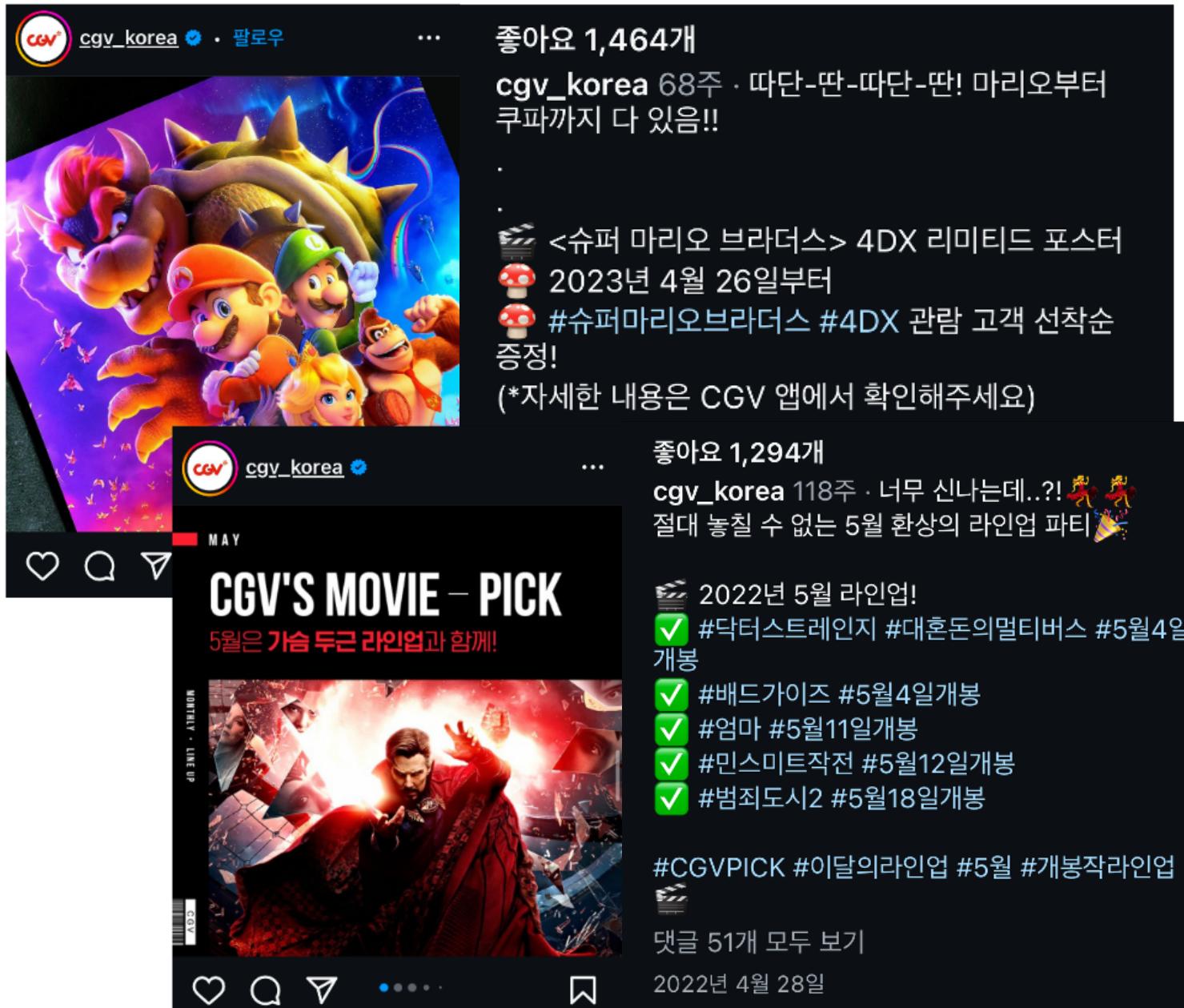


좋아요 수가 많은 해시태그

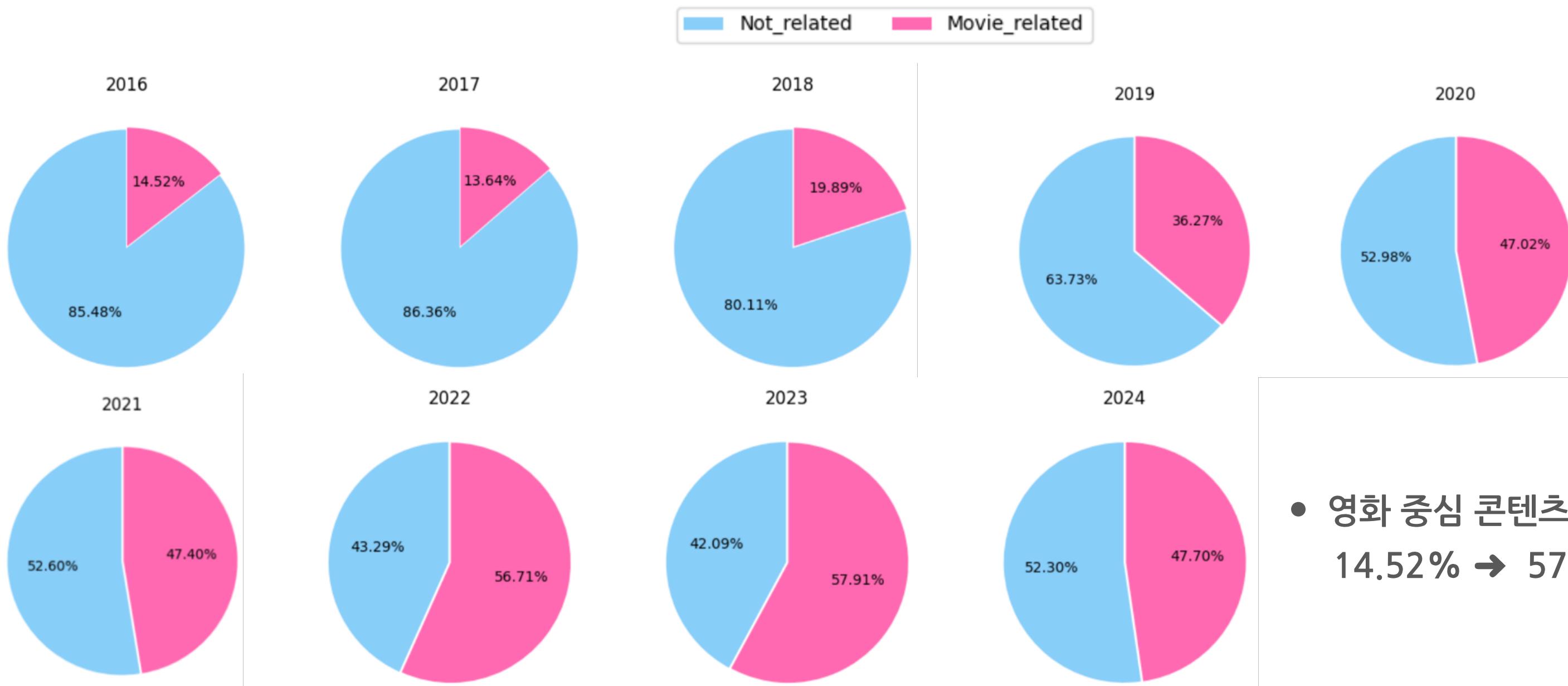
- 영화 제목, 영화 배우



영화 중심 콘텐츠 vs 브랜드 중심 콘텐츠



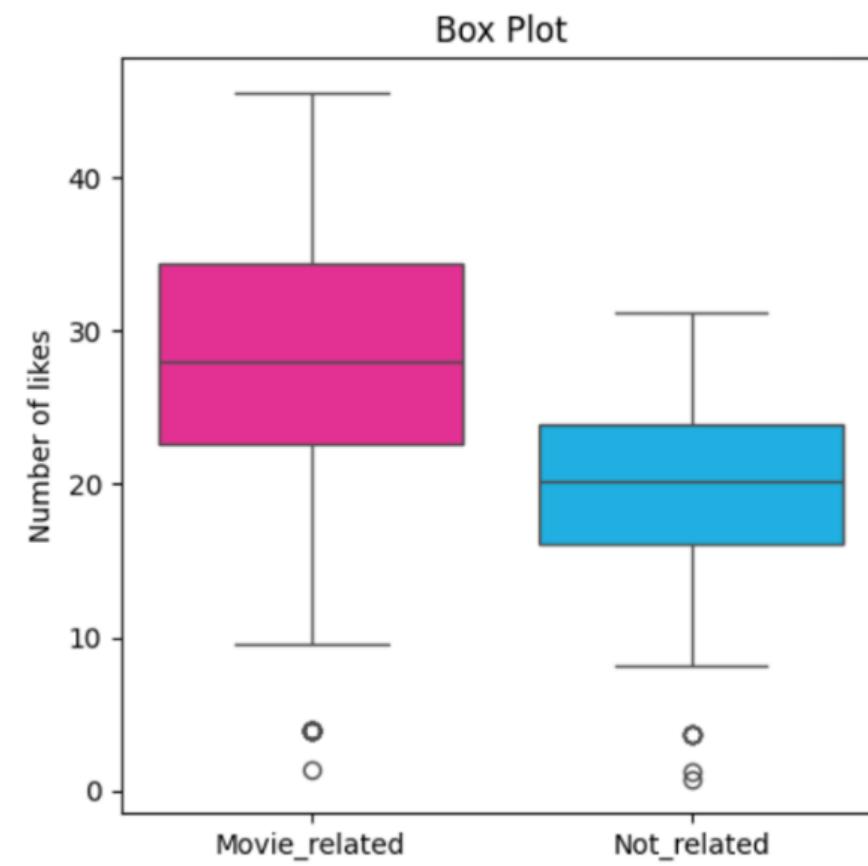
콘텐츠 비중 비교



- 영화 중심 콘텐츠의 비중 증가:
14.52% → 57.91%

좋아요 수 비교

U statistic: 4398680.0
P-value: 7.618172567543944e-116
Reject the null hypothesis – there is a significant difference between the two groups.
Movie-related posts have a higher median number of likes.



Mann-Whitney U test

통계량 = 4398680

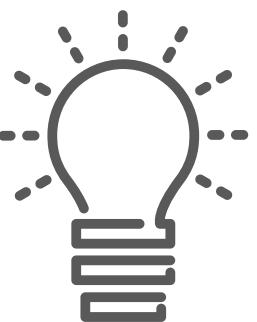
p-값 = 7.61817e-116

유의수준 0.05에서, 영화 중심 콘텐츠와 브랜드 중심 콘텐츠의 중위수에 통계적으로 유의한 차이가 있다.

영화 중심 콘텐츠의 좋아요 수가 더 많다고 할 수 있다.

6

결론 및 한계점



결론

영화의 장르, 등급, 국적별 관객 수에는 유의미한 차이가 있다.

→ 액션/어드벤쳐/SF 장르, 12세 이상 등급, 한국 영화가 더 선호되는 경향이 있음을 확인할 수 있다.

영화의 관객수(상위 75%)는 러닝타임과 별점과 양의 상관관계를 가지고 있다.

→ 영화의 흥행을 예측할 때 도움이 될 수 있다.

다만 러닝타임의 경우 데이터 내에 있는 구간 밖에서는 어떻게 될지 확실하지 않다.

2020년부터 전체적으로 감소하는 경향을 보이는 인스타그램 좋아요 수를 회복하기 위해

→ 영화 중심 콘텐츠의 양과 다양성 확대

→ 브랜드 중심 콘텐츠가 대중에게 보다 매력적으로 다가갈 수 있는 전략을 모색해야 한다.

한계점

- 크롤링으로 인한 많은 결측치 → 결측치를 다른 값으로 채우기 어려운 경우 데이터 자체를 사용하지 않았기에 데이터 수가 많지 않았다.
- 유의미한 변수의 부재 → 크롤링한 데이터들 중 일부는 현재까지 누적된 값이기에 예측에 활용할 수가 없었다.
- 대부분의 영화는 상업적으로 성공 X → skewed data



Thank You!

경청해주셔서 감사합니다.