

# 네티플릭스 콘텐츠 데이터 분석



공조팀 : 정성범, 장재준, 김꿈새

...



# 목차 Contents

• • •

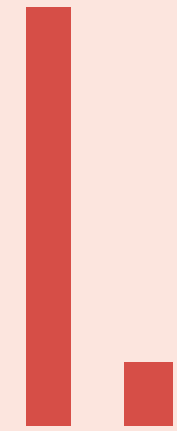
I. 넷플릭스 데이터 소개

III. 넷플릭스 데이터 시각화 및 분석

• • •

II. 넷플릭스 데이터 전처리

IV. 결론 및 Q&A



# 넷플릭스 데이터 소개



Dataset

netflix\_title

by Rohit Thomare

10 months ago • 1 MB • 1

netflix\_title

- 데이터 출처 : kaggle
- <https://www.kaggle.com/datasets/rohithomare17/netflix-title>

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lon
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train l

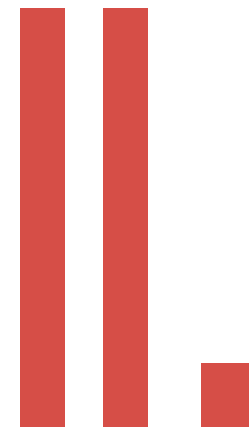
## 칼럼 소개

- show\_id : 콘텐츠 해당 아이디 번호
- type : tv쇼 또는 영화 중 콘텐츠가 속한 범주
- title : 콘텐츠 이름
- director : 콘텐츠 제작자 또는 감독
- cast : 영화와 tv쇼의 캐스팅 배우 목록
- country : 콘텐츠 제작 국가
- date\_added : 콘텐츠 업데이트 된 날짜
- release\_year : 콘텐츠 개봉 연도 또는 방영 시작 연도
- rating : 콘텐츠 관람 또는 시청 등급
- duration : 콘텐츠 상영 또는 방영 시간
- listed\_in : 콘텐츠 해당 장르 목록
- description : 콘텐츠 상세 설명

표본 수 : 8807개

기간 : 1925년 ~ 2021년

넷플릭스에 등록되어 있는 각종 영화, 드라마 등 각 콘텐츠별 제목, 감독, 배우 등의 정보



# 넷플릭스 데이터 전처리

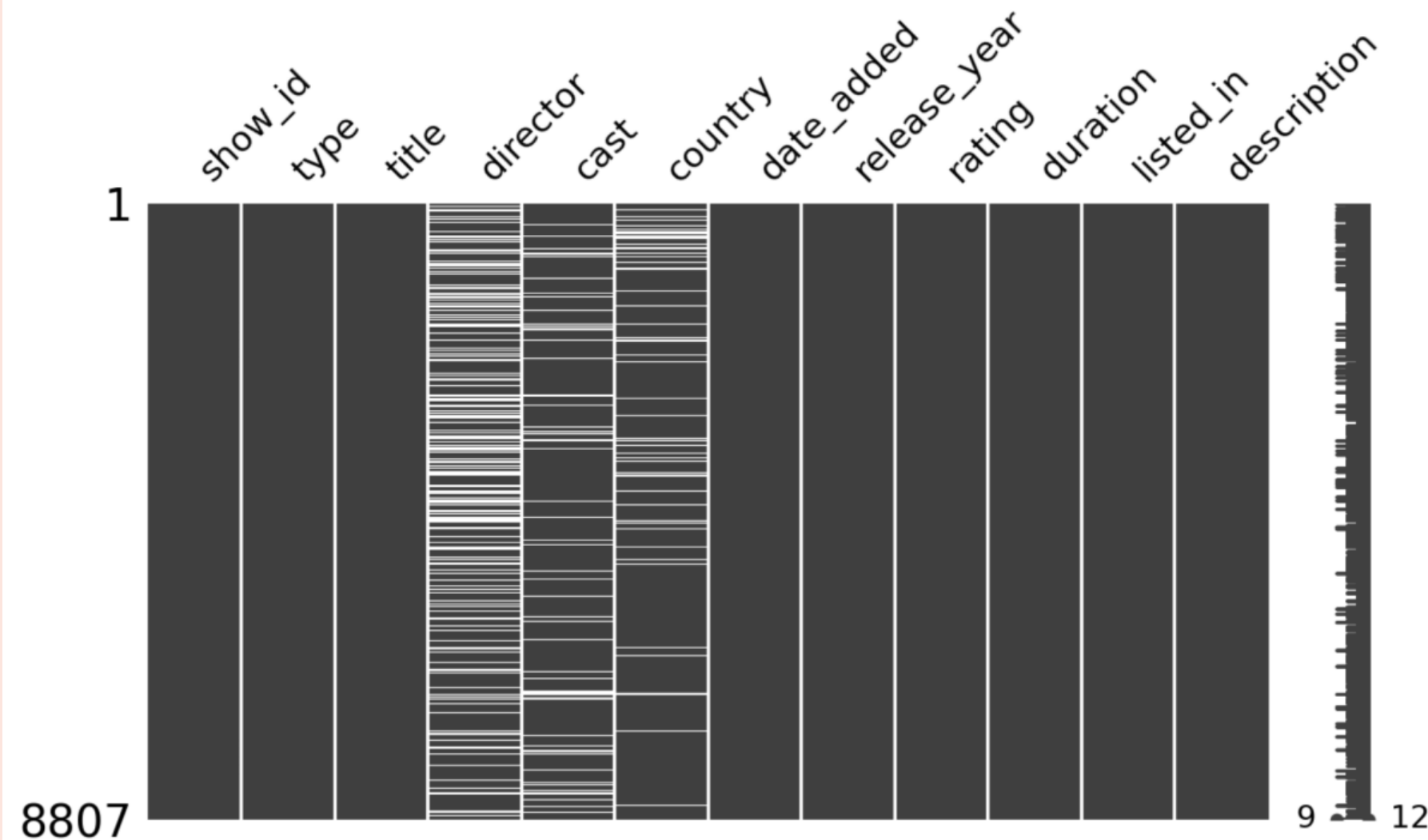
- 02 데이터 전처리
- director, cast, country 칼럼에서 결측치 발생함을 확인
- country, rating, duration 칼럼을 활용해 파생변수 생성

파생변수 생성

```
# 전처리 이전
df.isna().sum()
```

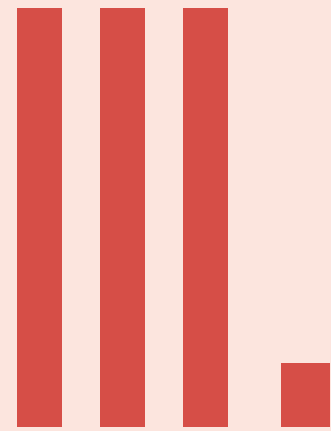
```
show_id      0
type         0
title        0
director    2634
cast        825
country     831
date_added   10
release_year  0
rating       4
duration     3
listed_in    0
description  0
dtype: int64
```

```
# 결측치 확인
sns.heatmap(df, figsize=(10, 5));
```



```
# 전처리 이후
df.isna().sum()
```

```
type         0
title        0
director     0
cast         0
country      0
date_added   0
release_year  0
rating       0
listed_in    0
month_added  0
month_name_added 0
year_added   0
country_1st  0
target_user  0
genre        0
dtype: int64
```



# 넷플릭스 데이터 시각화 및 분석

콘텐츠 유형, 국가, 소비자 연령, 장르, 계절 5가지 기준으로

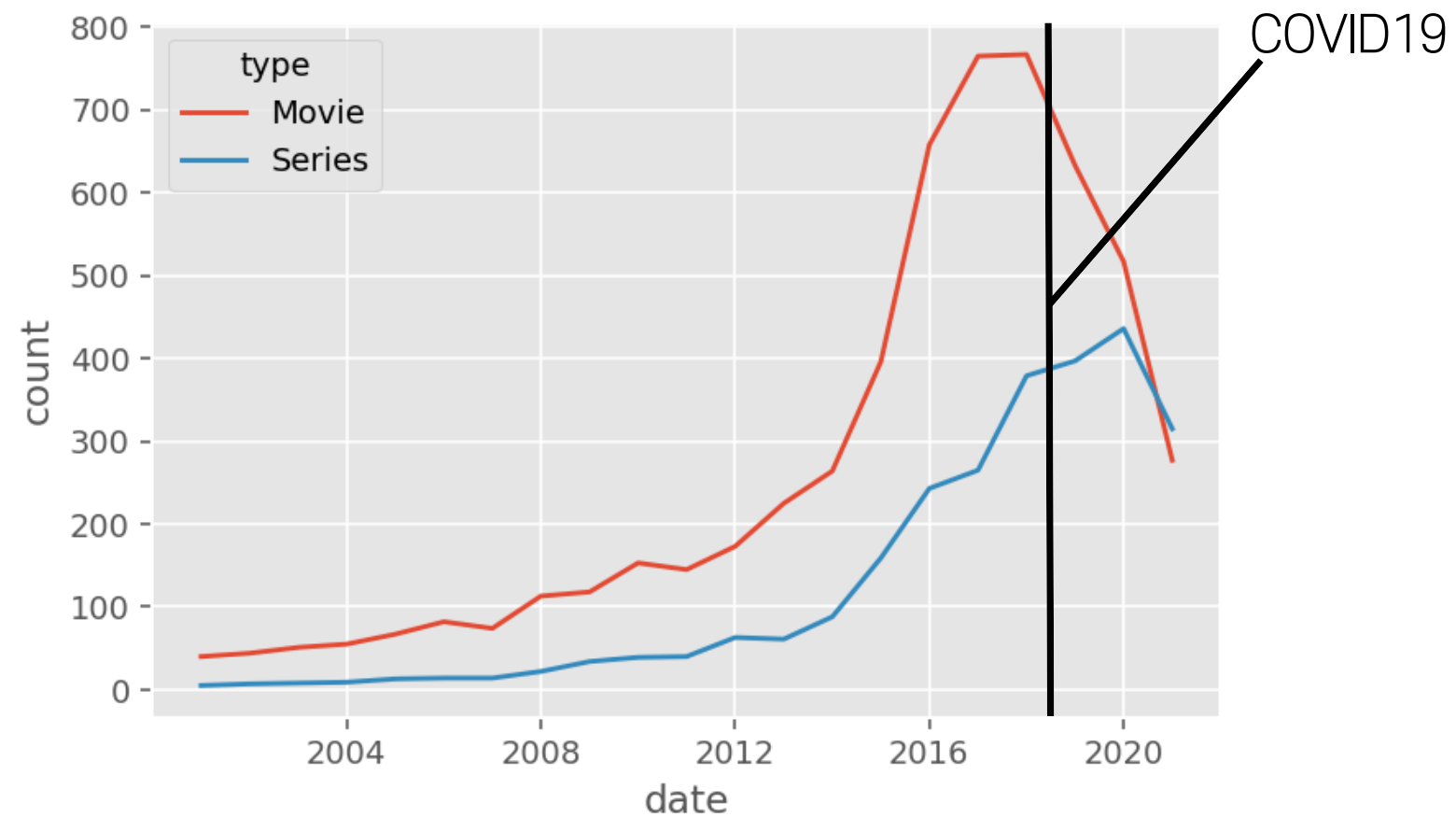
칼럼별 데이터 추출 및 분석 작업 진행

# III. 콘텐츠 유형별, 국가별 분석

01

콘텐츠 유형별 분석

→ Movie & Series 시계열 추이

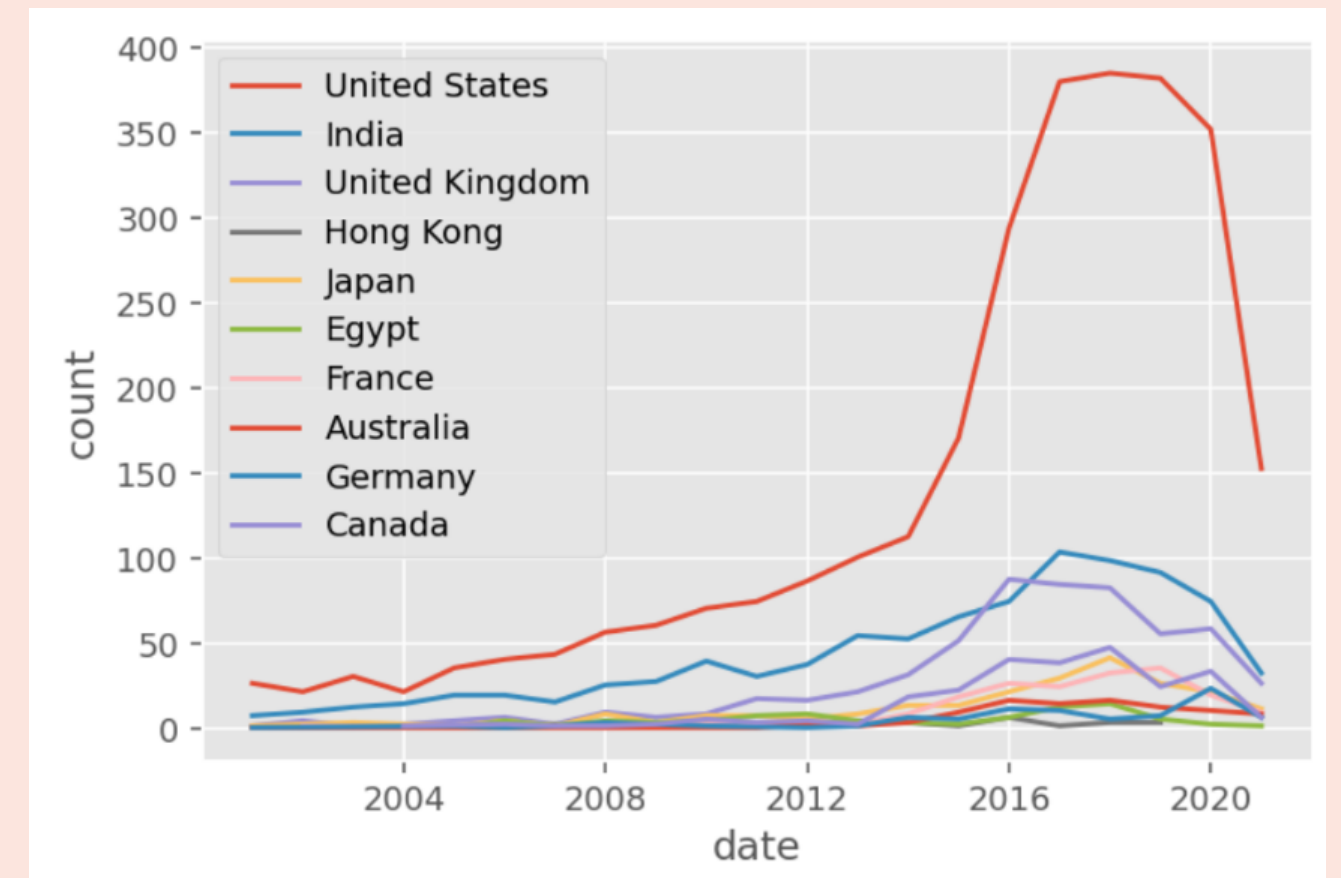


- 2015년 이후 영화 및 시리즈물 콘텐츠 모두 증가
- 코로나의 영향으로 영화 산업 축소

02

국가별 분석

→ Top 10 국가 시계열 추이



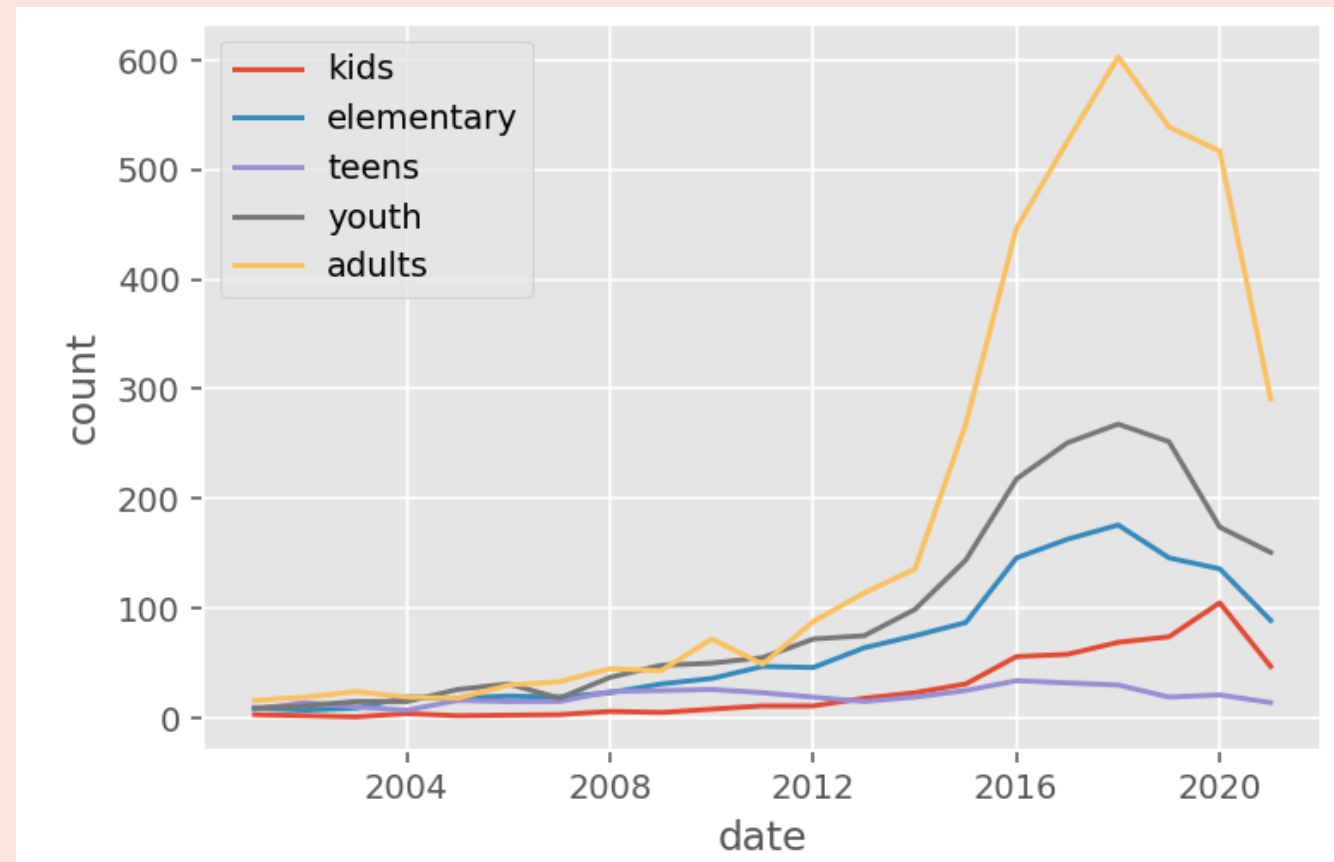
- 미국에 치중된 콘텐츠 수
  - 의외인 것은 인도 콘텐츠 비중이 높음
- Why? 높은 인구, 발리우드



# Ⅲ. 소비자 연령별, 주요 콘텐츠 장르별 분석

## 03

### 소비자 연령별 분석

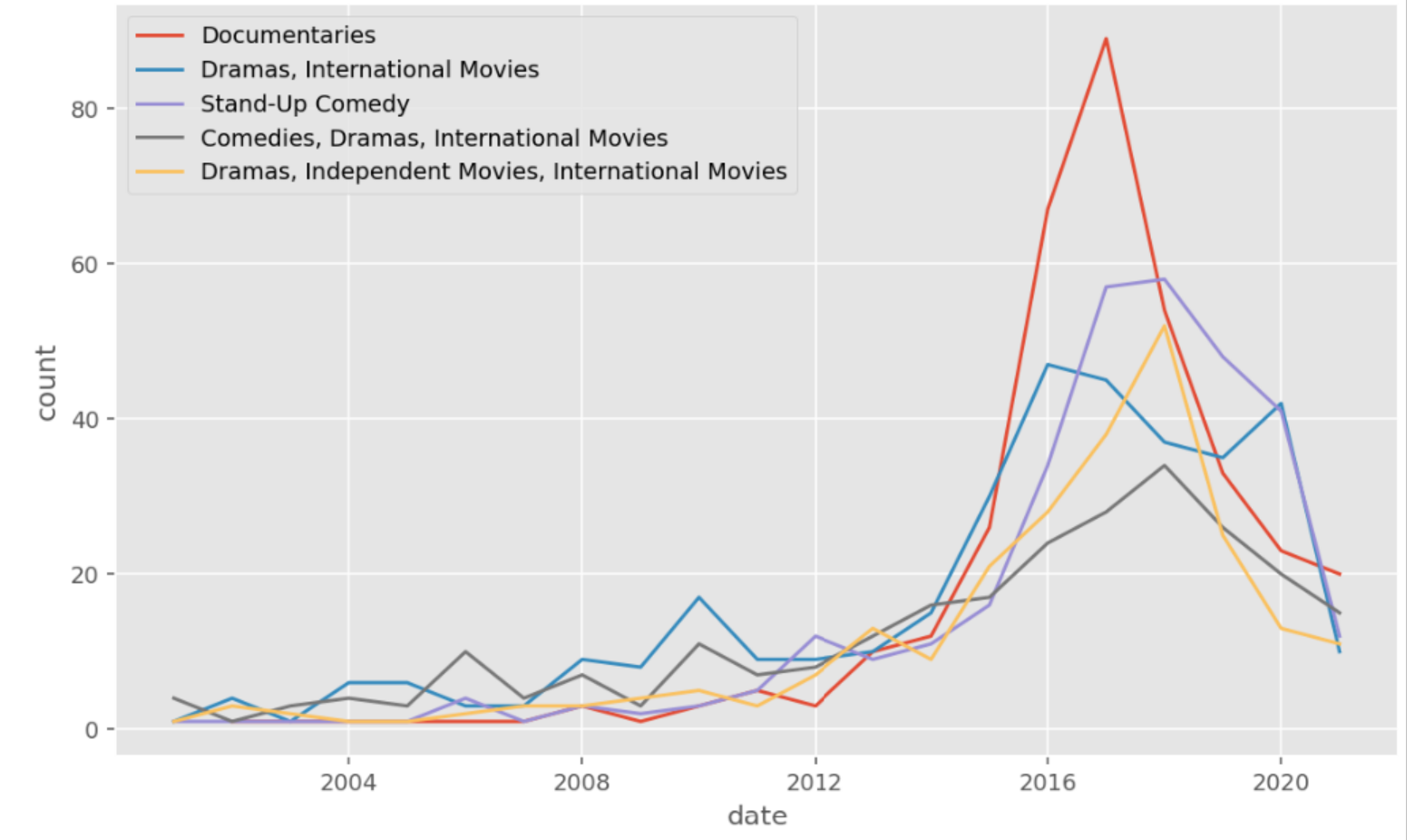


kids(전체관람가), elementary(7세 이상), teens(13세 이상)  
youth(16세 이상), adults(18세 이상, 청불포함)  
→ 소비자 수요를 반영해 20~30대 주요 고객층을 타겟

## 04

### 장르별 분석

→ Top 5 장르 시계열 추이



- 2015년부터 다큐멘터리 boom이 존재
- 최근 넷플릭스는 다큐멘터리 수급에 집중하지 않고 있음

### III. 계절별 콘텐츠 타이틀 워드클라우드



## # Killer



## # Christmas



## 04 결론

**01** 영화, 시리즈물 중 코로나19 이후  
시리즈물에 대한 선호도가 증가했다.

**02** 넷플릭스는 미국(자국) 콘텐츠에  
더 집중하고 있는 것으로 보인다.  
의외인 것은 인도 콘텐츠 비중이  
꽤 높다.

**03** 소비자 수요를 반영해 20~30대 주요  
고객층을 타겟으로 하고 있다.

**04** 2015년에 다큐멘터리 boom이  
존재했지만 2017년부터 넷플릭스는  
다큐멘터리 수급에 집중하지 않고 있다.

**05** 여름에는 스릴러, 겨울에는 Christmas  
관련 콘텐츠를 주로 제공하고 있다.

# Q & A

*Thank you* 발표를 들어주셔서

**감사합니다**

...

