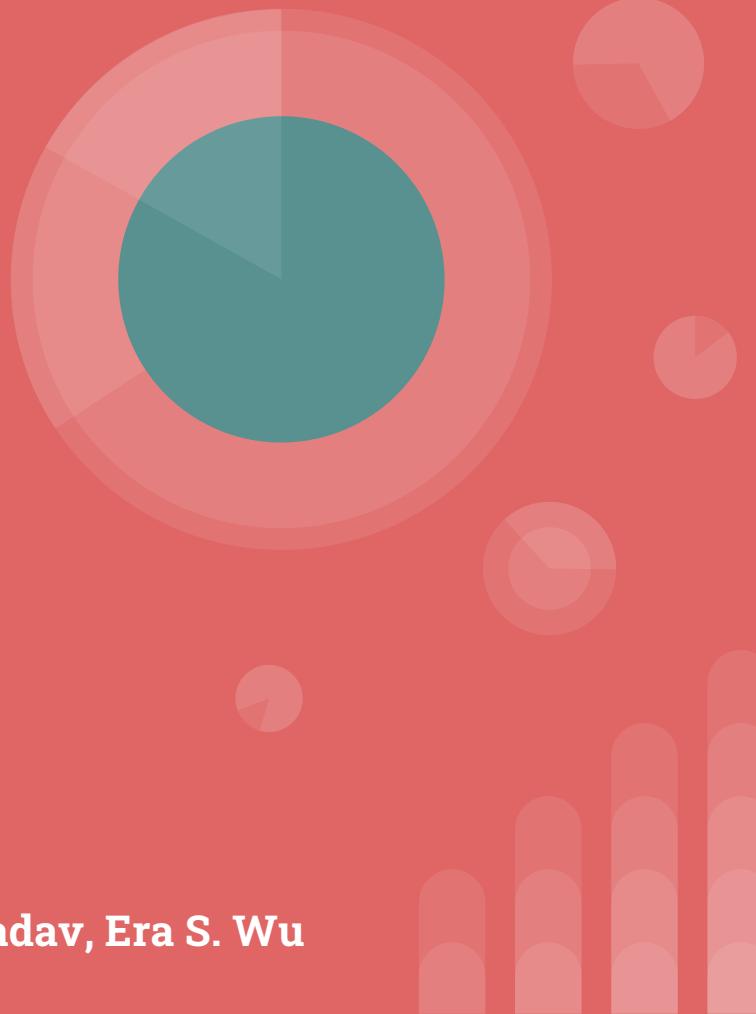


# A Story about Stories



Ansh Gupta, Bo Qin, Chinaza Nnam, Digvijay Yadav, Era S. Wu



# Introduction



# The dataset

**“Netflix Movies and TV Shows”** (<https://www.kaggle.com/datasets/shivamb/netflix-shows>)

- Show\_id = unique ID for every movie/TV show
- Type = identifier of movie/TV show
- Title
- Director
- Cast
- Country = country where the movie/TV show was produced
- Date\_added = date added to Netflix
- Release\_year = actual year of release
- Rating = TV rating (e.g., PG, TV-Y7)
- Duration = total duration, in minutes or # of seasons
- Listed\_in = genre (e.g., Docuseries, International Movies)



# The dataset

**“Netflix Movies and TV Shows”** (<https://www.kaggle.com/datasets/shivamb/netflix-shows>)

- Show\_id = unique ID for every movie/TV show
- Type = identifier of movie/TV show
- Title
- Director
- Cast
- Country = country where the movie/TV show was produced
- Date\_added = date added to Netflix
- Release\_year = actual year of release
- Rating = TV rating (e.g., PG, TV-Y7)
- Duration = total duration, in minutes or # of seasons
- Listed\_in = genre (e.g., Docuseries, International movies)





# The dataset

## Description

As her father nears the end of his life, filmmaker Kirsten Johnson stages his death in inventive and comical ways to help them both face the inevitable.

After crossing paths at a party, a Cape Town teen sets out to prove whether a private-school swimming star is her sister who was abducted at birth.

To protect his family from a powerful drug lord, skilled thief Mehdi and his expert team of robbers are pulled into a violent and deadly turf war.

Feuds, flirtations and toilet talk go down among the incarcerated women at the Orleans Justice Center in New Orleans on this gritty reality series.

In a city of coaching centers known to train India's finest collegiate minds, an earnest but unexceptional student and his friends navigate campus life.

The arrival of a charismatic young priest brings glorious miracles, ominous mysteries and renewed religious fervor to a dying town desperate to believe.

Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pegasi and Unicorns should be pals, and, hoof to heart, she's determined to prove it.

On a photo shoot in Ghana, an American model slips back in time, becomes enslaved on a plantation and bears witness to the agony of her ancestral past.

A talented batch of amateur bakers face off in a 10-week competition, whipping up their best dishes in the hopes of being named the U.K.'s best.

A woman adjusting to life after a loss contends with a feisty bird that's taken over her garden, and a husband who's struggling to find a way forward.

Sicily boasts a bold "Anti-Mafia" coalition. But what happens when those trying to bring down organized crime are accused of being criminals themselves?

Struggling to earn a living in Bangkok, a man joins an emergency rescue service and realizes he must unravel a citywide conspiracy.

After most of her family is murdered in a terrorist bombing, a young woman is unknowingly lured into joining the very group that killed them.

When the clever but socially-awkward Tetv™ joins a new school, she'll do anything to fit in. But the queen bee among her classmates has other ideas.

Cameras following Bengaluru police on the job offer a rare glimpse into the complex and challenging inner workings of four major crime investigations.

Students of color navigate the daily slights and slippery politics of life at an Ivy League college that's not nearly as "post-racial" as it thinks.

Declassified documents reveal the post-WWII life of Otto Skorzeny, a close Hitler ally who escaped to Spain and became an adviser to world presidents.

Strangers Diego and Isabel flee their home in Mexico and pretend to be a married couple to escape his drug-dealing enemies and her abusive husband.

After a deadly home invasion at a couple's new dream house, the traumatized wife searches for answers, and learns the real danger is just beginning.



# A story about stories?

- How do we analyze and learn more about these stories?
- Using data, we want to tell a story about the stories!



# Data Cleaning

# Initial Formatting

- Import and create original data sheet
- Spell check : Review > Spelling
- Filter unique rows/Remove duplicates: Data> Advanced > Unique Records only
- Data > DataTools > Remove Duplicates
- Both yielded same # of rows
- Finding missing values, Ensuring consistency in notation, Treating Blanks
- Examining date formats, Formerly in MM-DD-YYYY format. Went ahead to separate into respective columns

AutoSave OFF Home Insert Draw Page Layout Formulas Data Review View Tell me

Normal Page Break Preview Page Layout Custom Views Ruler Formula Bar Gridlines Headings Zoom 130% Zoom to Selection Zoom to 100% New Window Arrange All Unfreeze Panes Freeze Top Row Freeze First Column Split Hide Switch Windows View Macros Record Macro Use Relative References

C16	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	
1	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description						
2	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	United States	25-Sep-21	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmmaker Kirsten Johnson stages his final days.							
3	s2	TV Show	Blood & Water		Ama Qamat	South Africa	24-Sep-21	2021	TV-MA	2 Seasons	International	After crossing paths at a party, a Cape Town teen sets out to prove whether she can stop her father's drug empire.						
4	s3	TV Show	Ganglands	Julien Leclerc	Sami Bouajila, Tracy Goto	Canada	24-Sep-21	2021	TV-MA	1 Season	Crime TV Shows	To protect his family from a powerful drug lord, skilled thief Mehdi and his crew must infiltrate a gangland prison.						
5	s4	TV Show	Jailbirds New Orleans			United States	24-Sep-21	2021	TV-MA	1 Season	Documentaries, I Feuds	In a city of coaching centers known to train India's finest collegiate minds, TV Dramas, The arrival of a charismatic young priest brings glorious miracles, ominous r						
6	s5	TV Show	Kota Factory		Mayur More	India	24-Sep-21	2021	TV-MA	2 Seasons	Documentaries, I Feuds, flirtations and toilet talk go down among the incarcerated women at International	TV Dramas, Children & F Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pegasi and						
7	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford,	United States	24-Sep-21	2021	TV-MA	1 Season	Documentaries, I Feuds	TV Dramas, Children & F Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pegasi and						
8	s7	Movie	My Little Pony: A New Generation	Robert Culpepper	Vanessa Hudgens, Kimiko	United States	24-Sep-21	2021	PG	91 min	Documentaries, I Feuds	TV Dramas, Children & F Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pegasi and						
9	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanabi	United States	24-Sep-21	1993	TV-MA	125 min	Documentaries, I Feuds	TV Dramas, Children & F Equestria's divided. But a bright-eyed hero believes Earth Ponies, Pegasi and						
10	s9	TV Show	The Great British Baking Show	Andy Devons	Mel Giedroyc	United Kingdom	24-Sep-21	2021	TV-14	9 Seasons	Documentaries, I Feuds	British TV Shows	A talented batch of amateur bakers face off in a 10-week competition, whipping up delicious treats and learning valuable life lessons along the way.					
11	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy	United States	24-Sep-21	2021	PG-13	104 min	Comedies	Documentaries, D A woman adjusting to life after a loss contends with a feisty bird that's taken						
12	s11	TV Show	Vendetta: Truth, Lies and The Mafia			United States	24-Sep-21	2021	TV-MA	1 Season	Documentaries, D	TV Shows	Sicily boasts a bold "Anti-Mafia" coalition. But what happens when those try to					
13	s12	TV Show	Bangkok Breaking	Kongkiat Kor Sukollawat	Kanarat, Sushant Singh	Thailand	23-Sep-21	2021	TV-MA	1 Season	Documentaries, D	Struggling to earn a living in Bangkok, a man joins an emergency rescue service.						
14	s13	Movie	Je Suis Karl	Christian Schell	Luna Wedler	Germany, Czech Republic	23-Sep-21	2021	TV-MA	127 min	Documentaries, D	After most of her family is murdered in a terrorist bombing, a young woman must find strength and justice in the face of tragedy.						
15	s14	Movie	Confessions of an Invisible Girl	Bruno Garofoli	Klara Castanho, Lucca Pic	Portugal	22-Sep-21	2021	TV-PG	91 min	Documentaries, D	Children & F When the clever but socially-awkward Tetv™ joins a new school, she'll do all						
16	s15	TV Show	Crime Stories: India Detectives			United States	22-Sep-21	2021	TV-MA	1 Season	Documentaries, D	British TV Shows	Cameras following Bengaluru police on the job offer a rare glimpse into the					
17	s16	TV Show	Dear White People		Logan Browning	United States	22-Sep-21	2021	TV-MA	4 Seasons	Documentaries, D	TV Comedies	Students of color navigate the daily slights and slippery politics of life at an elite college.					
18	s17	Movie	Europe's Most Dangerous Man: Otto Skorzeny in Spain	Pedro de la Rosa	Diego Garcíá, Pablo Azorín	Spain	22-Sep-21	2020	TV-MA	67 min	Documentaries, D	Declassified documents reveal the post-WWII life of Otto Skorzeny, a close ally of Hitler.						
19	s18	TV Show	Falsa Identidad	Luis Ernesto	Mexico	Mexico	22-Sep-21	2020	TV-MA	2 Seasons	Documentaries, D	Strangers Diego and Isabel flee their home in Mexico and pretend to be a married couple.						
20	s19	Movie	Intrusion	Adam Salter	Freida Pinto, Logan Marshall-Green	United States	22-Sep-21	2021	TV-14	94 min	Thrillers	Thrillers	After a deadly home invasion at a couple's new dream house, the trauma					
58	s57	Movie	Naruto Shippuden the Movie: Blood Prison	Masahiko Matsuoka	Junko Takeuchi	Japan	15-Sep-21	2011	TV-14	102 min	Action & Adventure	Mistakenly accused of an attack on the Fourth Raikage, ninja Naruto is imprisoned.						
59	s58	Movie	Naruto Shippuden the Movie: Bonds	Hajime Kamio	Junko Takeuchi	Japan	15-Sep-21	2008	TV-PG	93 min	Action & Adventure	When strange ninjas ambush the village of Konohagakure, it's up to adolescent ninjas to defend their home.						
60	s59	Movie	Naruto Shippuden the Movie: The Will of Fire	Masahiko Matsuoka	Junko Takeuchi	Japan	15-Sep-21	2009	TV-PG	96 min	Action & Adventure	When four out of five ninja villages are destroyed, the leader of the one spared must find a way to restore balance to the world.						
61	s60	Movie	Naruto Shippuden: The Movie	Hajime Kamio	Junko Takeuchi	Japan	15-Sep-21	2007	TV-PG	95 min	Action & Adventure	The adventures of adolescent ninja Naruto Uzumaki continue as he's tasked with a dangerous mission.						
62	s61	Movie	Naruto Shippuden: The Movie: The Lost Tower	Masahiko Matsuoka	Junko Takeuchi	Japan	15-Sep-21	2010	TV-14	85 min	Action & Adventure	When Naruto is sent to recover a missing nin, the rogue manages to send him on a dangerous mission.						
63	s62	Movie	Naruto the Movie 2: Legend of the Stone of Gelel	Hirotsugu Kaizuka	Junko Takeuchi	Japan	15-Sep-21	2005	TV-PG	97 min	Action & Adventure	While on a mission to return a missing pet, Naruto and two fellow ninjas are faced with a powerful enemy.						
64	s63	Movie	Naruto the Movie 3: Guardians of the Crescent Moon	Toshiyuki Tsukamoto	Junko Takeuchi	Japan	15-Sep-21	2006	TV-PG	95 min	Action & Adventure	Exuberant ninja Naruto teams up with his pals Sakura and Kakashi to escort the last remaining member of his clan back to safety.						
65	s64	Movie	Naruto the Movie: Ninja Clash in the Land of Snow	Tensai Okazaki	Junko Takeuchi	Japan	15-Sep-21	2004	TV-PG	83 min	Action & Adventure	Naruto, Sasuke, and Sakura learn they'll be protecting an actress from being captured by a group of bandits.						
66	s65	Movie	Nightbooks	David Yarrow	Winslow Fegley, Lidya Jewell	United States	15-Sep-21	2021	TV-PG	103 min	Children & Family	Scary story fan Alex must tell a spine-tingling tale every night, or stay trapped in a dark room.						
67	s66	TV Show	Numberblocks	Beth Chalme	United Kingdom	United Kingdom	15-Sep-21	2021	TV-Y	6 Seasons	Kids' TV	In a place called Numberland, math adds up to tons of fun when a group of colorful characters help children learn basic arithmetic.						
68	s67	TV Show	Raja Kasu Aur Anya Kahanianyan			India	15-Sep-21	2014	TV-G	1 Season	Documentaries, I	I Explore the history and flavors of regional Indian cuisine, from traditional Kitchens to modern street food.						
69	s68	TV Show	Saved by the Bell	Mark-Paul Gosselaar	United States	United States	15-Sep-21	1994	TV-PG	9 Seasons	Kids' TV	From middle school to college, best friends Zack, Kelly, Slater, Jessie, Screech, and Mr. T team up to solve problems and have fun.						
70	s69	Movie	Schumacher	Hanns Brügel	Michael Schumacher	Germany	15-Sep-21	2021	TV-14	113 min	Documentaries, I	Through exclusive interviews and archival footage, this documentary traces the life and career of Formula 1 legend Michael Schumacher.						
71	s70	TV Show	Stories by Rabindranath Tagore			India	15-Sep-21	2015	TV-PG	1 Season	Documentaries, I	The writings of Nobel Prize winner Rabindranath Tagore come to life in this international series.						
72	s71	TV Show	Too Hot To Handle: Latino	Itatvá Cantoral		United States	15-Sep-21	2021	TV-MA	1 Season	Documentaries, I	On this reality show, singles from Latin America and Spain are challenged to find love.						
73	s72	Movie	A StoryBot Space Adventure	David A. Vann	Evan Spiridellis, Erin Fitzgerald	United States	14-Sep-21	2021	TV-Y	13 min	Children & Family	Join the StoryBots and the space travelers of the historic Inspiration4 mission.						
74	s73	TV Show	Jack Whitehall: Travels with My Father	Jack Whitehall	United Kingdom	United Kingdom	14-Sep-21	2021	TV-MA	5 Seasons	Documentaries, I	British TV Show	Comedic Jack Whitehall invites his stuffy father, Michael, to travel with him around the world.					
75	s74	Movie	King of Boys	Kemi Adetiba	Sola Sobowa	Nigeria	14-Sep-21	2018	TV-MA	182 min	Documentaries, I	When a powerful businesswoman, King of Boys, has political ambitions, she threatens to expose her son's secret past.						
76	s75	TV Show	The World's Most Amazing Vacation Rentals	David Cimino	David Cimino	United States	14-Sep-21	2021	TV-PG	2 Seasons	Documentaries, I	With an eye for every budget, three travelers visit vacation rentals around the world.						
77	s76	Movie	Woman With a Gun	Carrie-Anne Moss	Christina Ricci	United States	14-Sep-21	2021	TV-C	102 min	Documentaries, I	Children & Family	A woman with a gun must make choices to save her family.					

# Column Formatting

- Trimming spaces: =TRIM(D2:D8800)
- Splitting columns
  - **listed\_in**: to present the various genres of shows and movies captured in the data
  - **Cast**: split into separate columns. Selected first three to represent top three leading cast actors

# Natural Language Processing



# Natural Language Processing

- Word clouds
- Sentiment analysis
- Genre prediction

# NLP: Word Clouds

# Word Cloud 1: the universal theme?



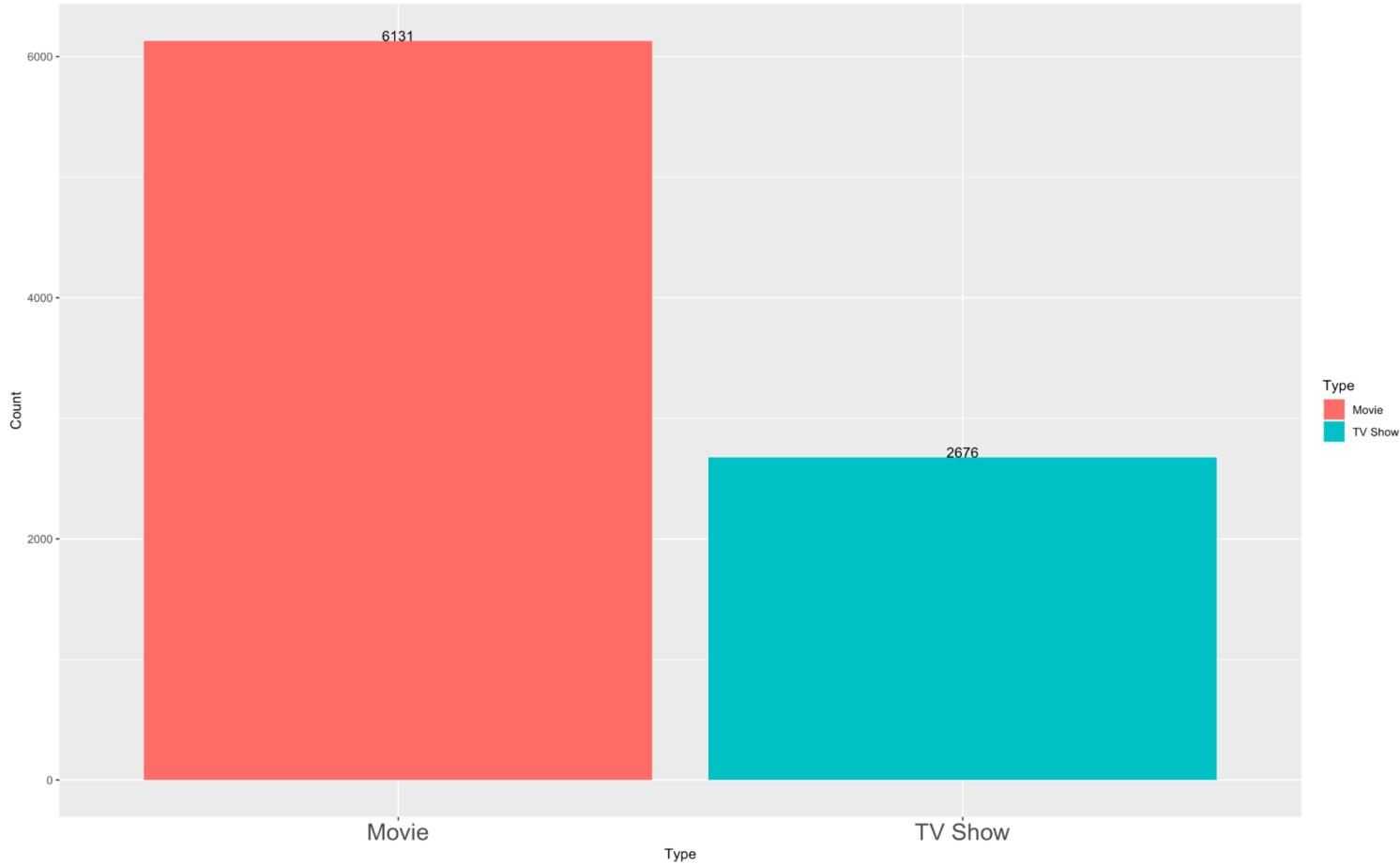
crew husband  
dreams forces reality  
search brother mission  
job romance student history  
game sets crime special drama business  
comic drug story team daughter couple  
social time city takes teen set falls music  
film son mother true town career  
power wife series world past live  
people day leads star school life love fight living  
stand based lives begins  
meets boy girl family home friend run  
battle war save women dark  
future comedy friends father returns adventure  
personal cop evil documentary college  
village stories murder secret struggles  
relationship mysterious journey events police  
american marriage comedian agent  
childhood learn

based on all entries



META, META.

# Word Cloud 2: types



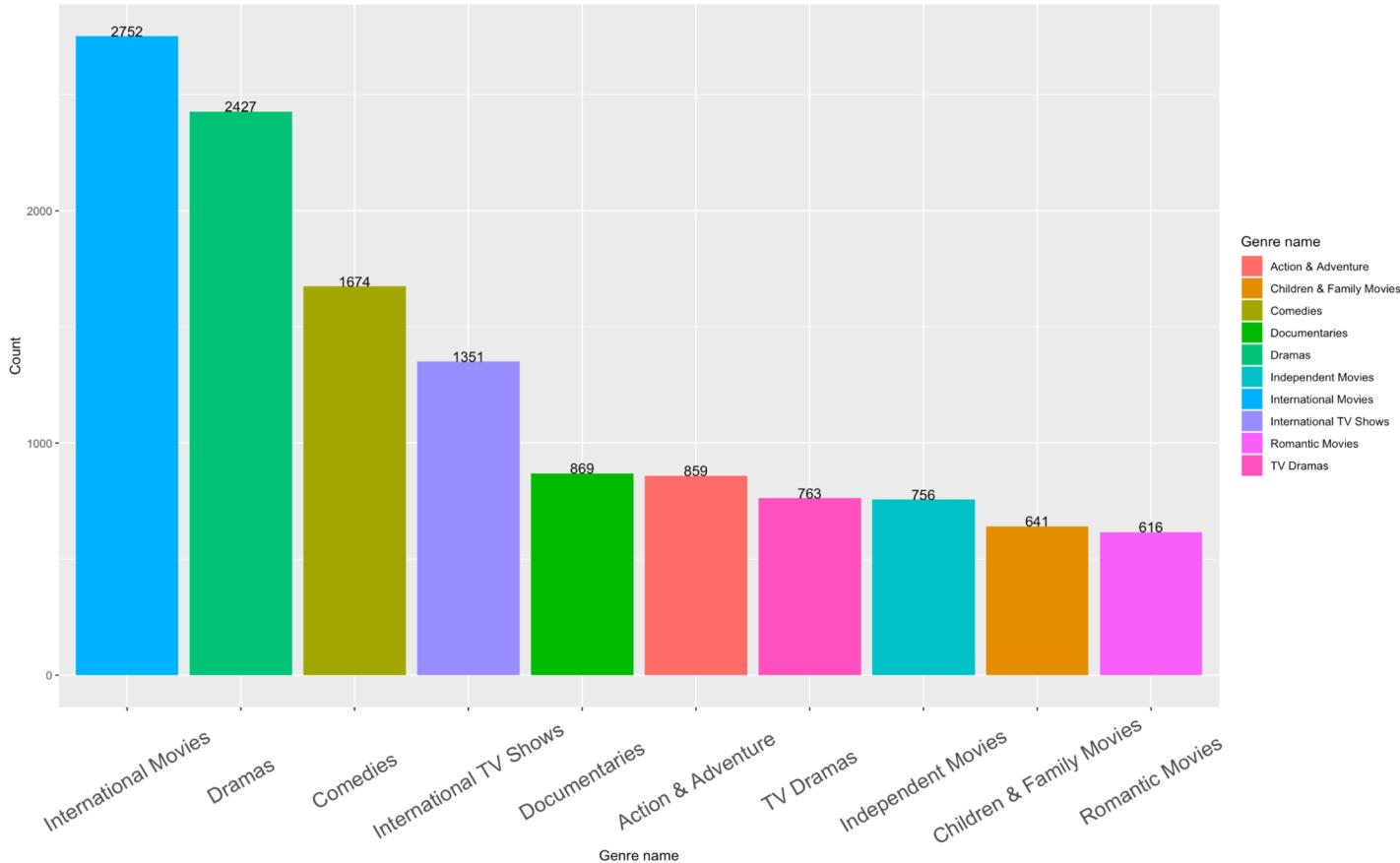
# Word Cloud 2: types

history revenge  
forced run relationship parents  
village comedian returns meets  
night star past team couple struggles  
brother mother stand falls career  
women romance takes war story series  
college school set game  
drug music father marriage  
real film girl town fight  
live teen save friend york  
stories mission day job  
husband wife home  
based city lives  
journey sets son crime  
comic wife begins  
police true boy student  
agent evil documentary murder  
stage secret time killer  
events battle death comedy dreams  
rescue mysterious american child  
search childhood drama learns  
christmas girlfriend struggling

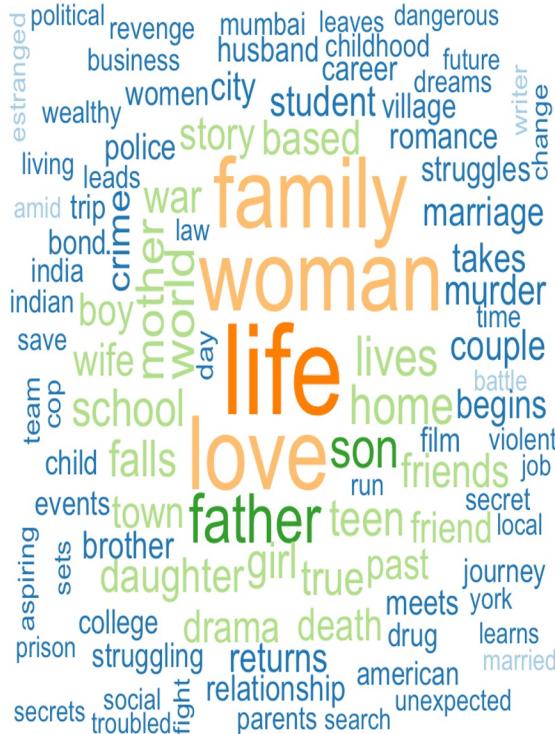
## Movie

TV Show

# Word Cloud 3: genres



# Word Cloud 3: genres - dramas and comedies



## Dramas



## Comedies

# Word Cloud 3: genres - internationals

money struggles  
personal drama romance married  
secrets nightstar mumbai events  
winsave secret relationship job  
dark setteen village daughter wealthy  
forced brother falls friends marriage sets  
trip police documentary school  
heart death town  
child wife murder singer  
killer boy past friend revenge  
time leads city son mission  
officer meets takes cop based  
team returns story girl drug  
women film war inday  
childhood india home true career  
mysterious mother fight crime  
parents student begins battle marry  
business change husband girlfriend  
change dreams gang political  
unexpected struggling

## International movies

challenges  
marriage history navigate  
business college reality families aspiring  
living people secret wealthy century  
husband father son war real  
career paths sets drama women  
journey wife boy girl school personal  
secrets friend team lives friendship joins  
cop chance past battle death star  
teen heart murder leads true future  
top home live students  
city dark world police time story  
share explore falls mysterious save  
explore falls detective student crime  
mother struggles daughter drug takes based single  
powerful power meets events begins fight london  
revenge dreams returns town day  
determined

## International shows

# Word Cloud 3: genres - documentaries and action & adventure

## documentary

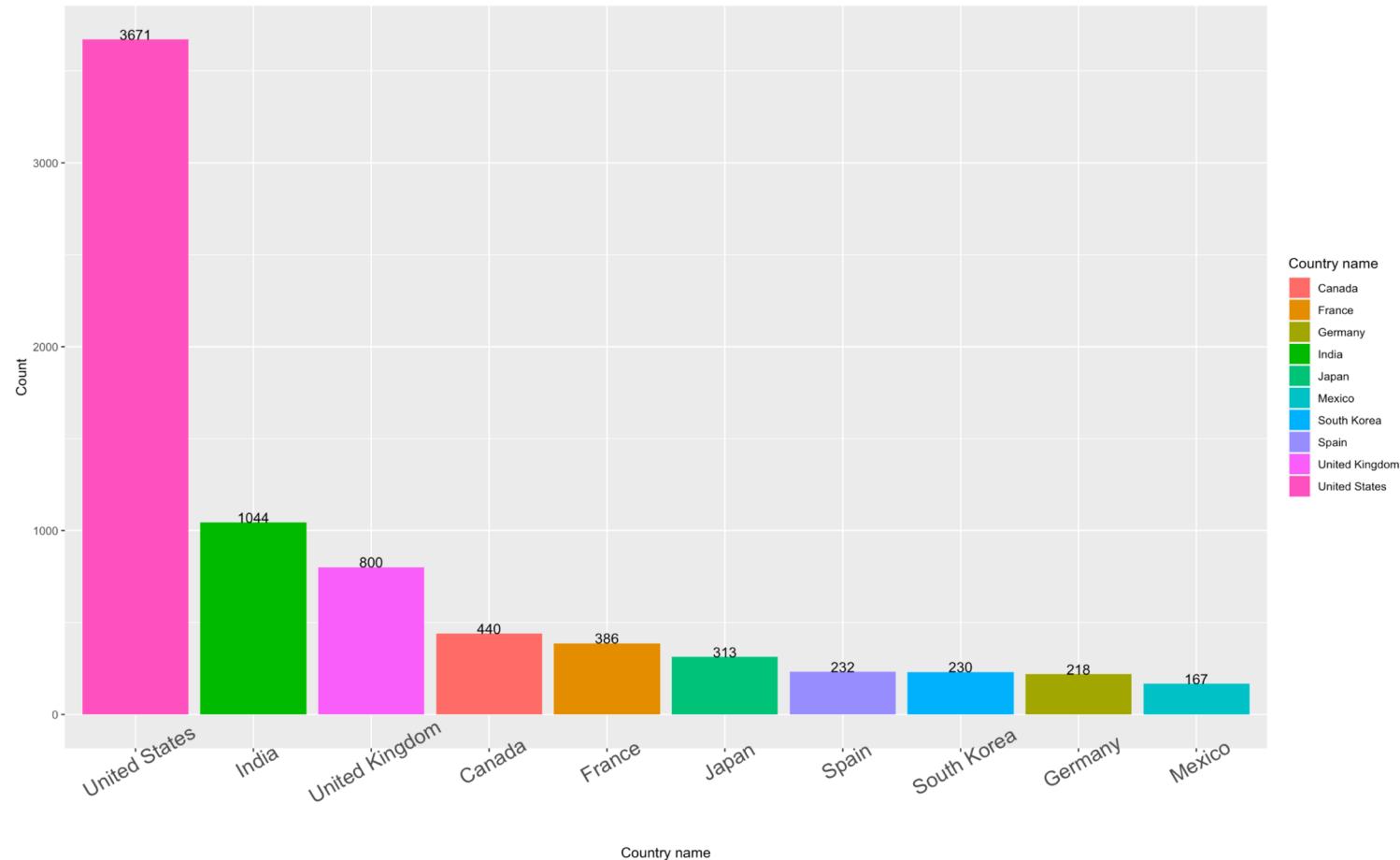
A word cloud centered around the word "documentary" in large orange font. The words are color-coded by category: life (orange), film (yellow), history (green), music (blue), and other genres (grey). The words include: reveals, u.s, journey, life, history, rise, details, john, power, rock, explores, film, music, follow, america, world's, experts, share, art, career, star, american, black, community, icon, athletes, examines, director, filmmaker, culture, battle, sport, decades, personal, musical, death, captures, woman, movement, controversial, extraordinary, friends.

Documentaries

A word cloud centered around the words "action" and "adventure" in large orange font. The words are color-coded by category: war (orange), life (yellow), mission (orange), secret (green), world (orange), rescue (red), mission (orange), powerful (green), forces (green), drug (red), city (green), band (green), son (green), discover (pink), past (pink), takes (pink), cia (pink), town (pink), teen (pink), police (orange), car (orange), officer (green), top (green), true (green), death (green), time (green), arts (green), teams (green), leader (green), gang (green), fight (red), save (red), criminal (green), team (green), attack (green), lord (green), ancient (green), assassin (green), returns (green), terrorist (red), father (red), earth (green), human (green), u.s (green), ruthless (green), kidnapped (green), dangerous (green), stop (green), special (green), soldiers (green), army (green), sets (green), brothers (green), hunter (green), prison (green), enemy (green), leads (green), planet (green), people (green).

Action & Adventure

# Word Cloud 4: countries of production



# Word Cloud 4: countries - similarities

romance forces drama  
ankees career mother  
power son death  
agent special  
woman day secret crew  
killer fight girl save fr  
wild father past city  
past set home  
stories real home  
drug teen true  
history time  
humantown  
wifestar  
journey stand  
mysterious lives  
u.s murde  
battle comedia  
struggles friend  
discov

legendary special events di  
stand comic an  
future father fri  
animals save  
romance school  
world's dark team  
stories agent time  
past team  
set home  
murder live ii  
game london  
epic son death  
document  
daughter far  
discovers based  
forced footage britain

relationship fight  
journey leads liv  
political village  
learns couple  
couple past  
day heart  
aspiring marry  
officer loves  
loves city  
career town  
true film  
town story  
friend win  
star sets  
sets home  
based boy  
seeks stories  
friends

korea's brother families  
king national gains begin  
class students married agent  
future live mysterious learn  
political murder crime sets game  
returns college secret dreams sets writer  
island police past woman  
singer lee friends love  
change north teen family life  
trip home animal  
joins star pop korea  
company city world day  
magical school evil  
husband law detective top  
law save reality lives  
daughter chance student art  
ability

U.S.

U.K.

India

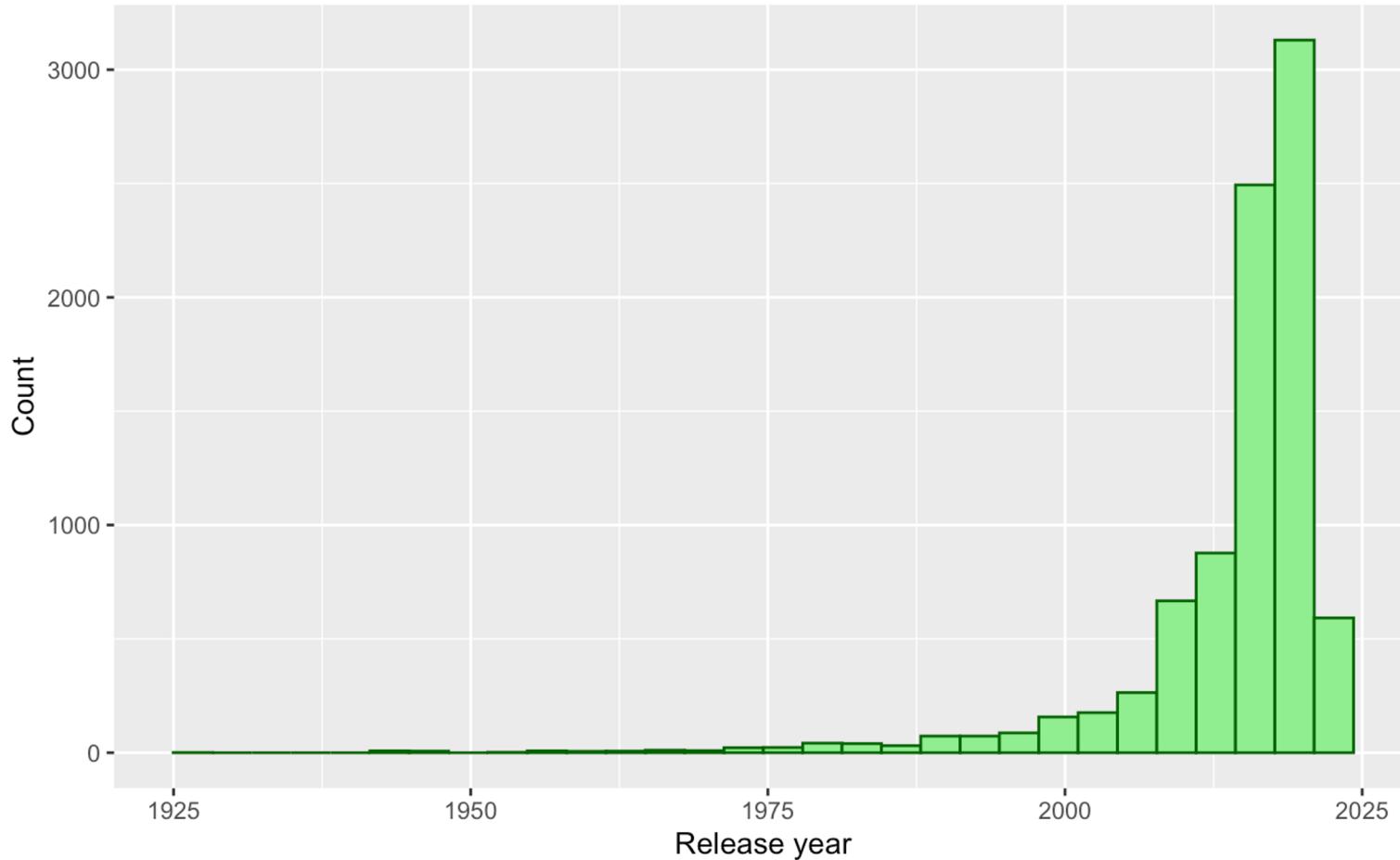
South  
Korea

## Word Cloud 4: countries - differences

## Mexico

# Japan

# Word Cloud 5: release year



# Word Cloud 5: release year ≈ zeitgeist?

nazi  
american  
era  
:=film  
war  
documentary  
collection  
combat

1925~1950

professor  
comedy  
world leads  
hired woman  
love war girl  
officer  
daughter life falls police  
wife family friends  
danger  
forbidden island

1951~1975

# Word Cloud 5: release year ≈ zeitgeist?

1976~2000

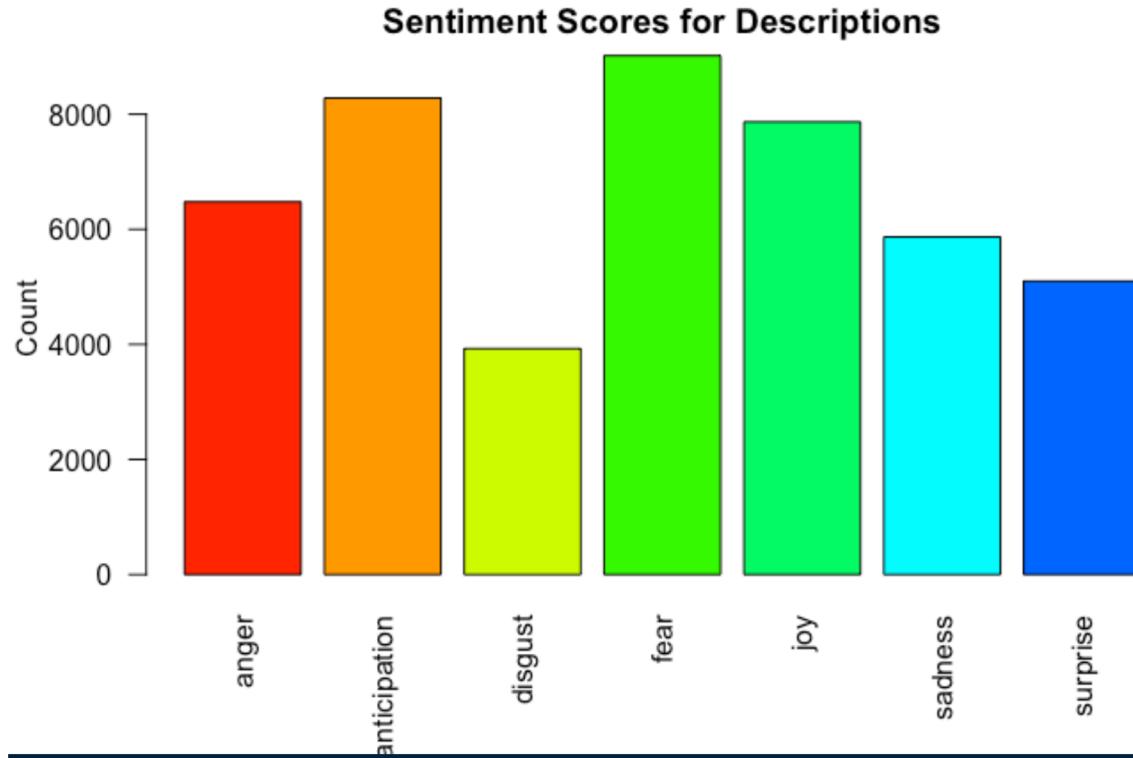
2000~2021



# NLP: Sentiment Analysis



# Sentiment Scores





# Interpretation



- **Fear and Sensationalism** is selling!!!
- **Tumultuous times** - recession, tech boom, pandemic
- **Change in tastes of viewers** - psychological thrillers  
(Jaws, Shutter Island rewatched), Stranger Things doing well

# NLP: Genre Prediction



# Proliferation

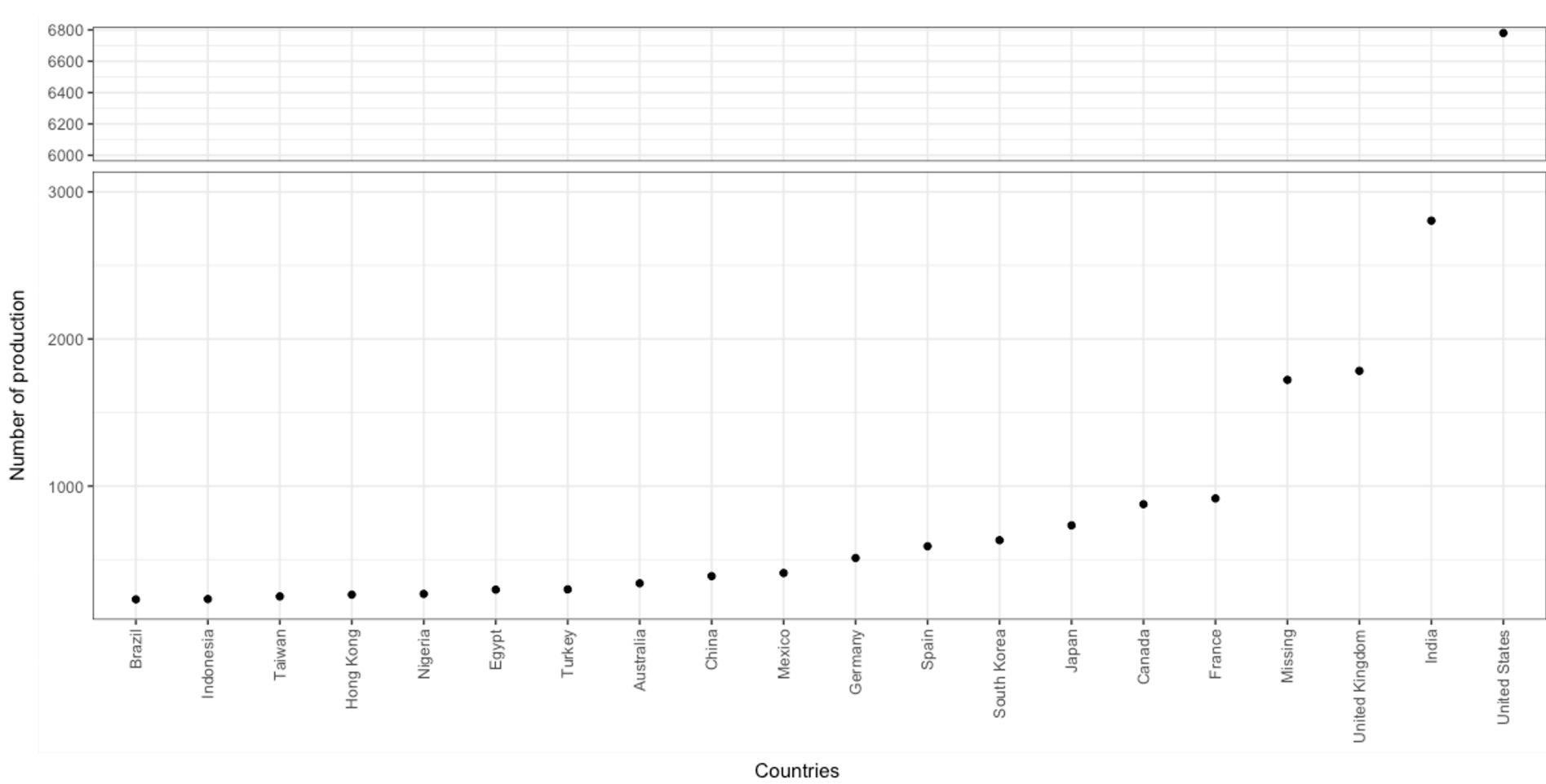
(Which country has highest number of production/genre types?)

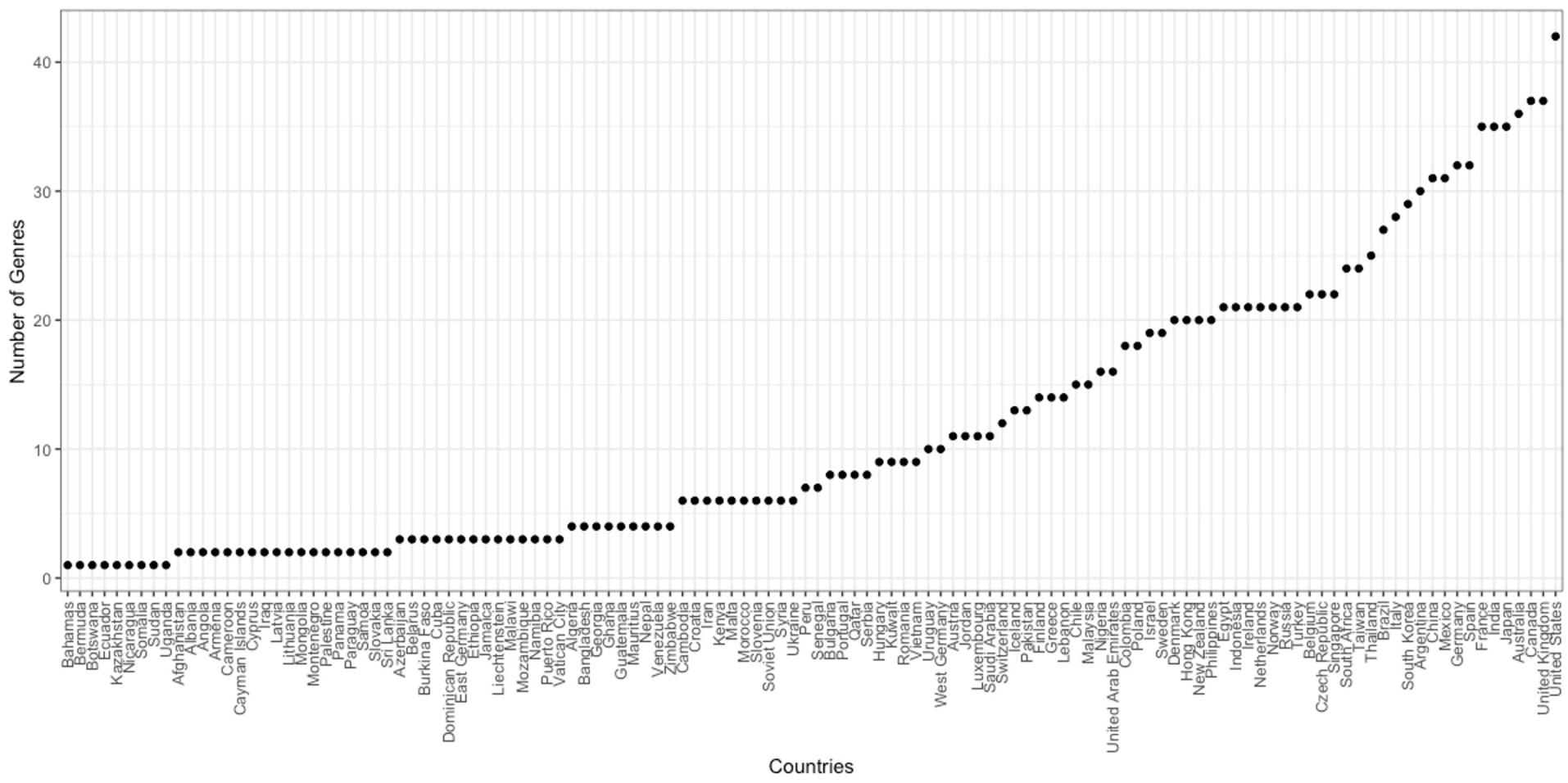
Each movie is associated with multiple countries and multiple genres

**Remove missing values** (only countries are missing)

-> Convert **Wide format** to **Long format** using python (Explode!)

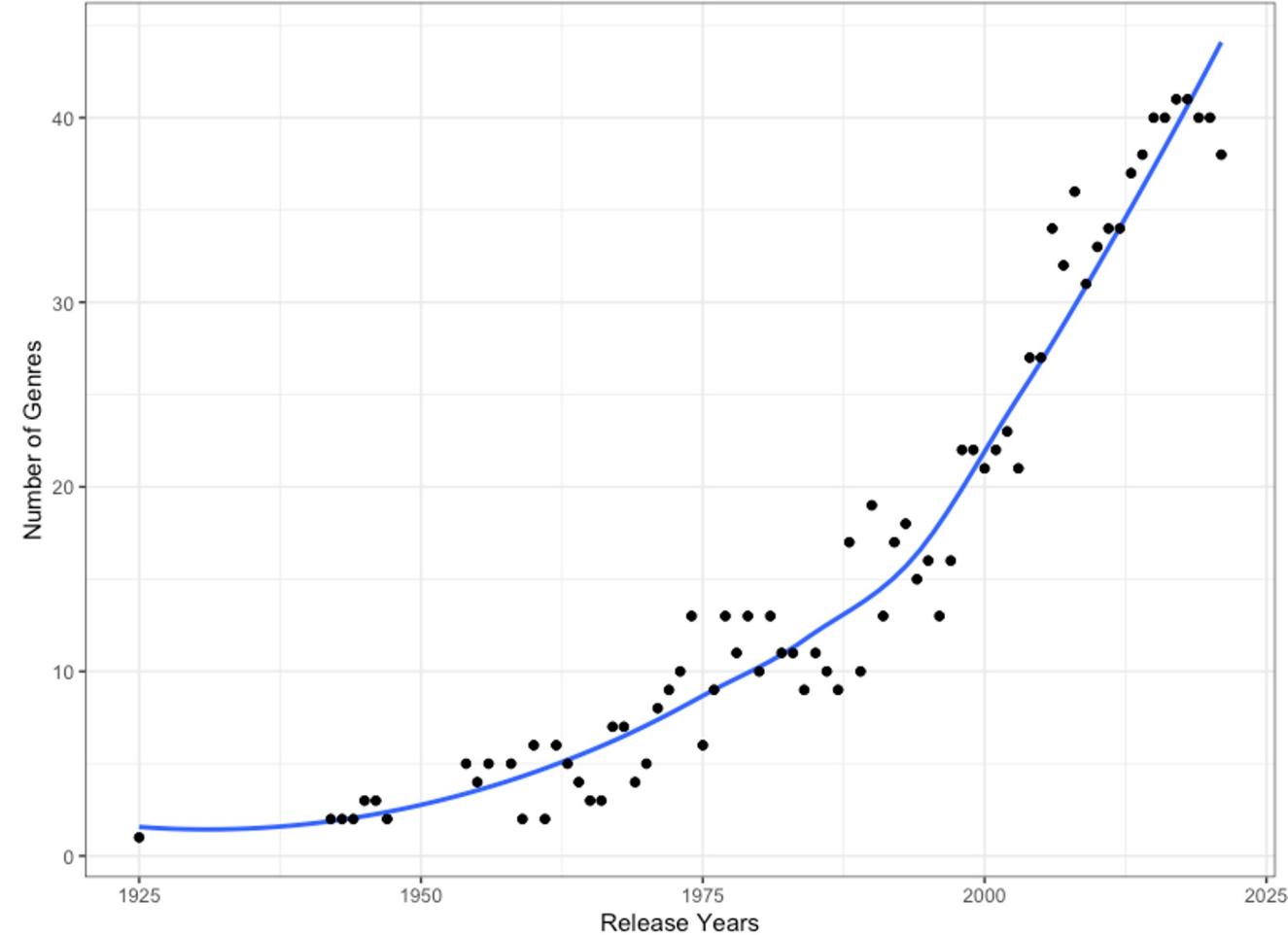
Missing	Not Missing	Total
1722	7085	8807







## Number of genres by production years (as included in this dataset).





# Genre Prediction Model

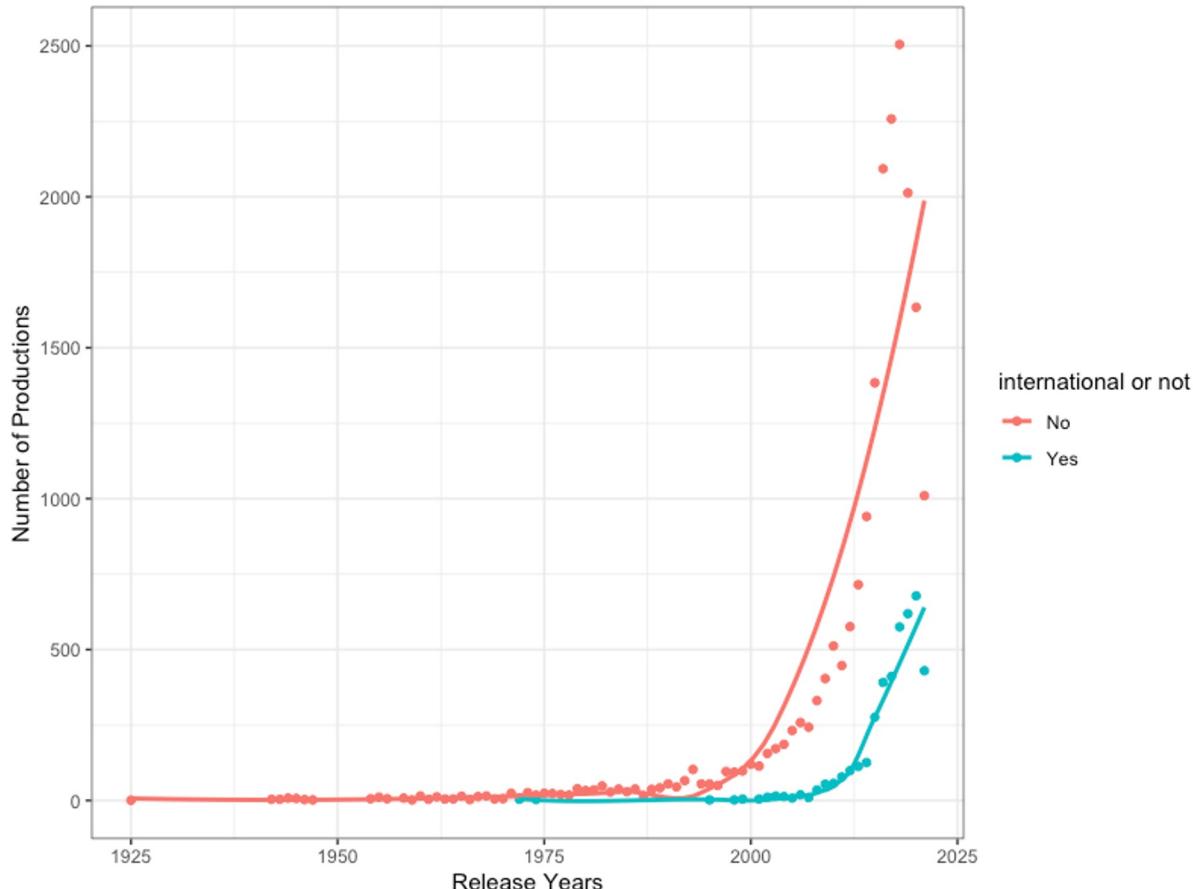
Things we considered for the model:

- 1, Limited data
- 2, Repetitive Genre types (e.g., Comedies, TV Comedies)
- 3, Imbalanced outcomes

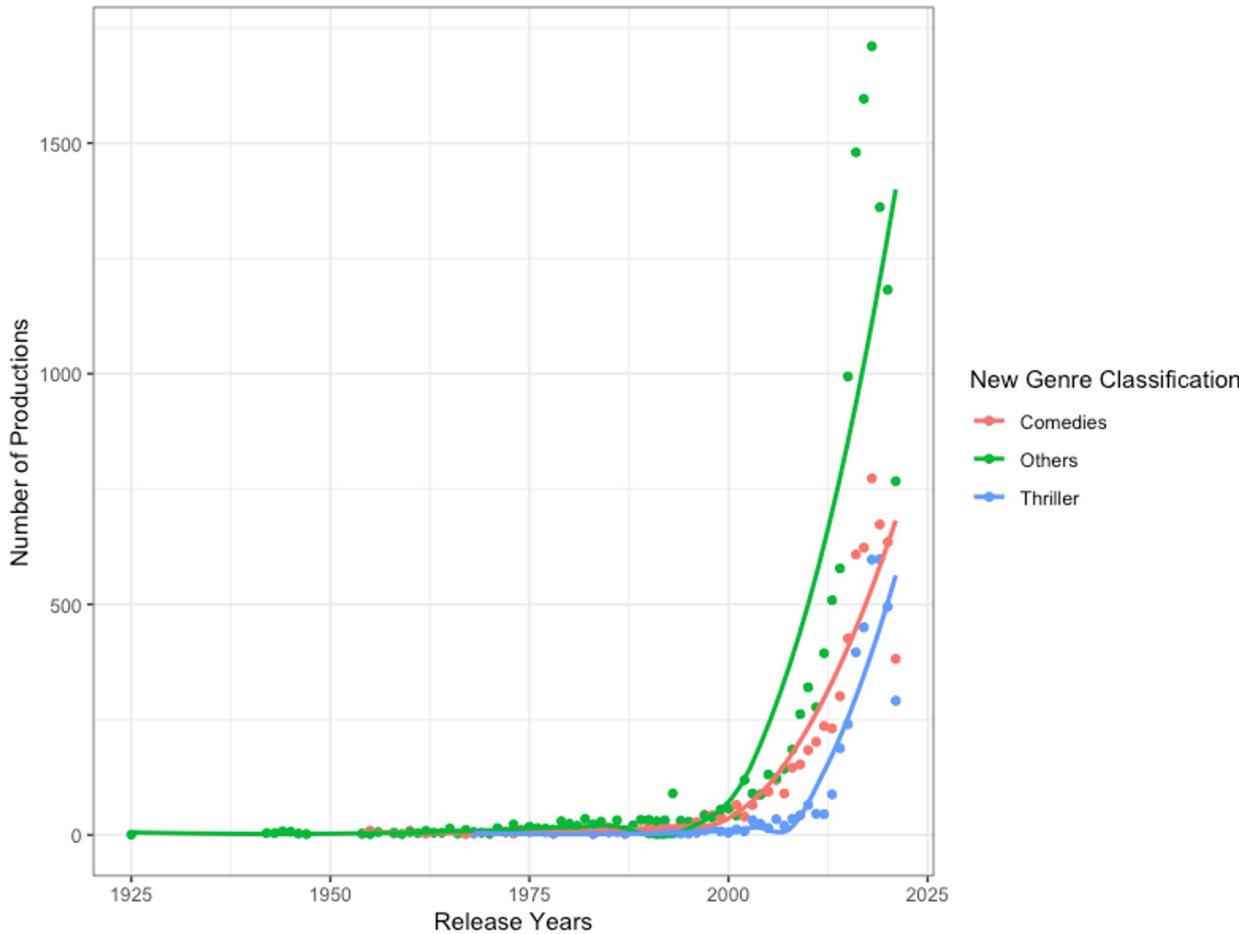
**Created new variables to account for these problems:**

International	0	The shows are domestic production
	1	The shows are produced by other countries
Type	0	Movie
	1	TV
Genres	0	Other
	1	Thriller
	2	Comedy

# Check how good is our new variables

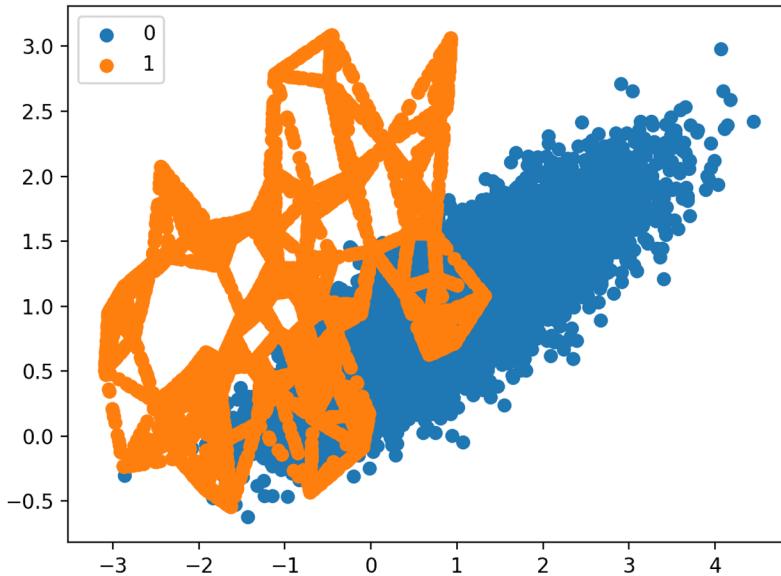
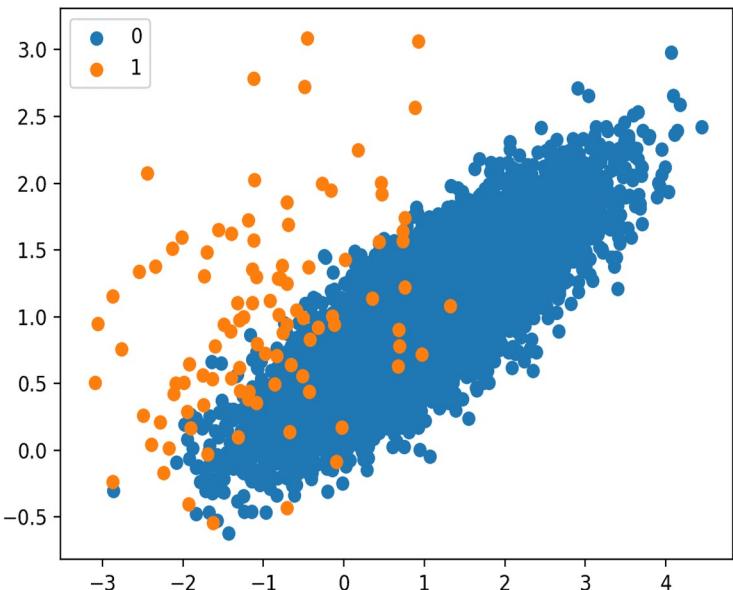


# Check how good is our new variables

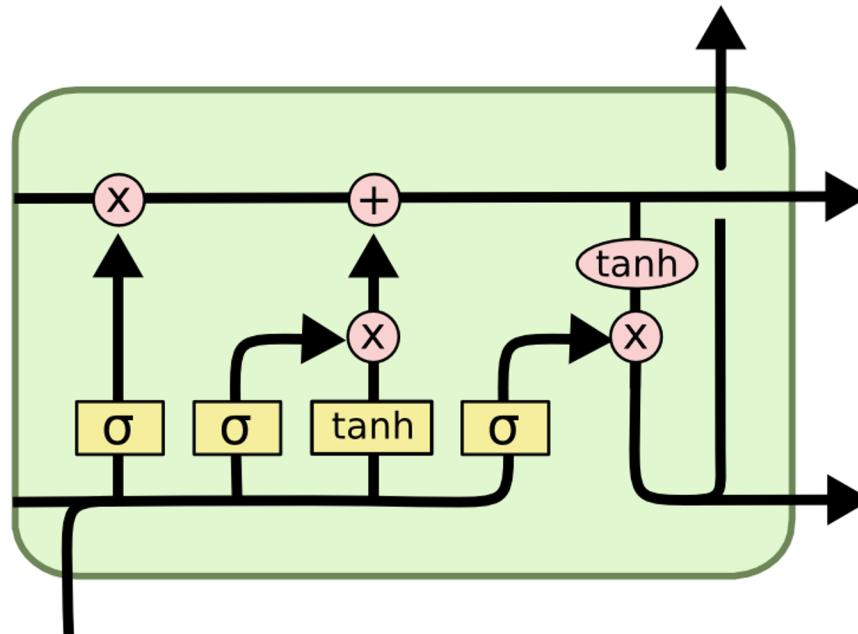




# SMOTE Sampling



# Long Short Term Memory Model (black magic)





# Model Performance

**8-2 train test split, drop out at 30% chance, SMOTE sampling to account for persistent imbalance.**

	Accuracy	Specificity	Sensitivity
International	89.9	87.7	90.2
Genre	67.7	67.8	67.6
Type	90.6	89.2	92.3

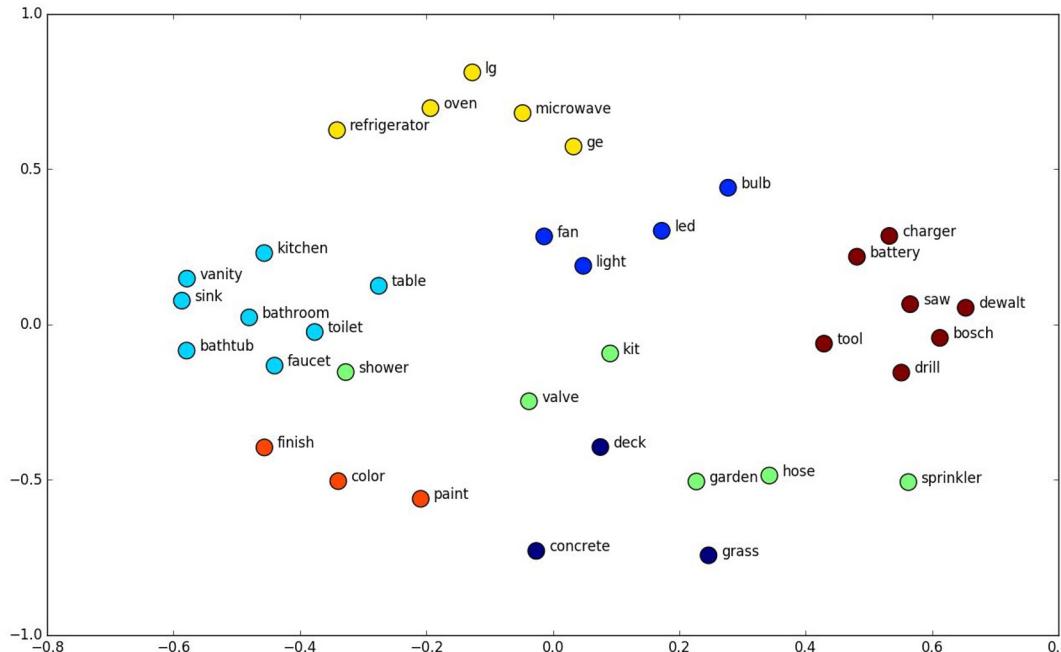
**Accuracy:** Total correct predictions/ Total predictions

**Specificity:** Total correct prediction of negative cases/ Total negative cases

**Sensitivity:** Total correct prediction of positive cases/ Total positive cases



# Word Embedding!



# Thank You!