

A close-up photograph of a vintage video camera, likely a Sony CCD camera, with the word "NETFLIX" overlaid in large, bold, red capital letters. The camera's viewfinder and control panel are visible, showing various settings like "P3.2K", "F40", "NF89", "Z31", "AUDIO", "CH1 SELECT", "CH2 SELECT", "INT(L)", "INPUT 1", "INPUT 2", "CH1", "CH2", "AUTO", "MANU", "FUNCTION", "SHTR/F.RATE", "DIAL SEL", "MODE CHK", "DISP/ AUTO MANU". The background is dark and out of focus.

NETFLIX

Netflix Movies and TV shows

Team 6: Zengyi Xu, Siyuan Huang, Hang Yin, Yi Lin, Yubei Tang, Harshitha Tummalapalli

Netflix Movies and TV shows



Scenario & Original Data



Normalization Plan



ETL



Analytical Procedures

Scenario

Everything is a Recommendation



Over 75% of what people watch comes from our recommendations

The problem:

people don't always want to watch five-star videos: a member might give "The Godfather" five stars and "Friends" three stars but still be more interested in watching "Friends."

Requirements:

Design a database system

- new information can easily be updated and properly and safely stored
- create different authorization levels for departments to allow them to have quick access to only the information they need

Applications:

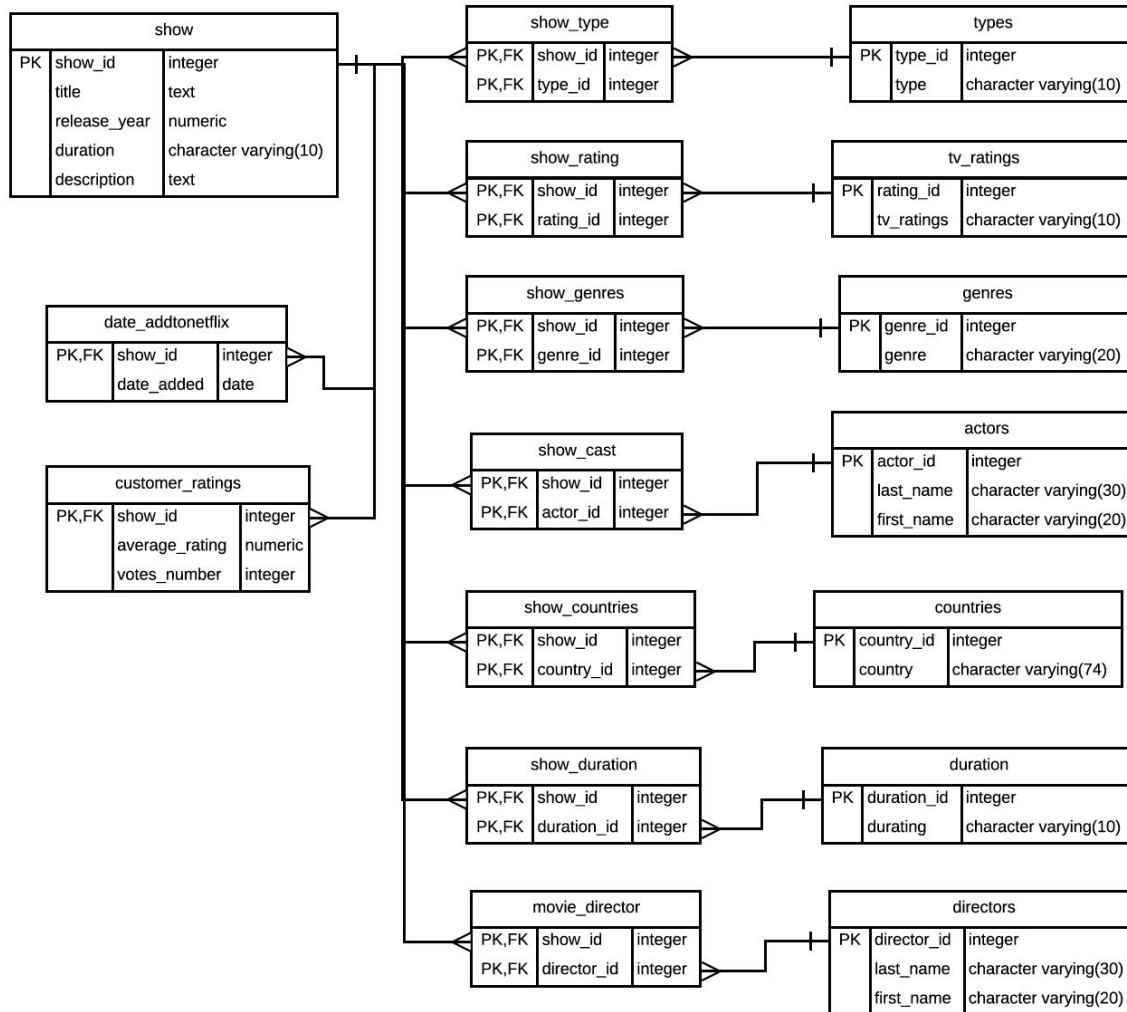
- Determine the trend in different regions or genres
- Provide better personalized recommendation
- Understand its strengths and weaknesses and be able to improve its content quality and management system

Sample of original data

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description		
81145628	Movie	Norm of the North: King Siz	Richard Finn, Tim Maltbv	Alan Marriott, Andrew	United State	9-Sep-19	2019	TV-PG	90 min	Children & F	Before planning an awesome wed		
80117401	Movie	Jandino: Whatever it Takes		Jandino Asporaat	United King	9-Sep-16	2016	TV-MA	94 min	Stand-Up C	Jandino Asporaat riffs on the chal		
70234439	TV Show	Transformers Prime		Peter Cullen, Sumalee	United State	8-Sep-18	2013	TV-Y7-FV	1 Season	Kids' TV	With the help of three human allie		
80058654	TV Show	Transformers: Robots in Disguise		Will Friedle, Darren Ci	United State	8-Sep-18	2016	TV-Y7	1 Season	Kids' TV	When a prison ship crash unleash		
80125979	Movie	#realityhigh	Fernando Lebrija	Nesta Cooper, Kate W	United State	8-Sep-17	2017	TV-14	99 min	Comedies	When nerdy high schooler Dani fi		
80163890	TV Show	Apaches		Alberto Ammann, Elo	Spain	8-Sep-17	2016	TV-MA	1 Season	Crime TV Sh	A young journalist is forced into a		
70304989	Movie	Automata	Gabe Ib谩帽ez	Antonio Banderas, Dy	Bulgaria, Un	8-Sep-17	2014	R	110 min	International	In a dystopian future, an insuranc		
80164077	Movie	Fabrizio Copano: Solo piens	Rodrigo Toro, Francisco Schultz	Fabrizio Copano	Chile	8-Sep-17	2017	TV-MA	60 min	Stand-Up C	Fabrizio Copano takes audience p		
80117902	TV Show	Fire Chasers			United State	8-Sep-17	2017	TV-MA	1 Season	Docuseries,	As California's 2016 fire season rag		
70304990	Movie	Good People	Henrik Ruben Genz	James Franco, Kate H	United State	8-Sep-17	2014	R	90 min	Action & Ac	A struggling couple can't believe t		
80169755	Movie	Joaqu铆n Reyes: Una y no n	Jos茅 Miguel Contreras	Joaqu铆n Reyes		8-Sep-17	2017	TV-MA	78 min	Stand-Up C	Comedian and celebrity imperson.		
70299204	Movie	Kidnapping Mr. Heineken	Daniel Alfredson	Jim Sturgess, Sam Wc	Netherlands	8-Sep-17	2015	R	95 min	Action & Ac	When beer magnate Alfred "Fredc		
80182480	Movie	Krish Trish and Baltiboy		Damandeep Singh Baggan, Smita		8-Sep-17	2009	TV-Y7	58 min	Children & F	A team of minstrels, including a r		
80182483	Movie	Krish Trish and Baltiboy: Bat	Munjal Shroff, Tilak Shetty	Damandeep Singh Baggan, Smita		8-Sep-17	2013	TV-Y7	62 min	Children & F	An artisan is cheated of his payme		
80182596	Movie	Krish Trish and Baltiboy: Bes	Munjal Shroff, Tilak Shetty	Damandeep Singh Baggan, Smita		8-Sep-17	2016	TV-Y	65 min	Children & F	A cat, monkey and donkey team u		
80182482	Movie	Krish Trish and Baltiboy: Coi	Tilak Shetty	Damandeep Singh Baggan, Smita		8-Sep-17	2012	TV-Y7	61 min	Children & F	In three comic-strip-style tales, a l		
80182597	Movie	Krish Trish and Baltiboy: Ov	Tilak Shetty	Rishi Gambhir, Smita Malhotra, De		8-Sep-17	2017	TV-Y7	65 min	Children & F	A cat, monkey and donkey learn t		
80182481	Movie	Krish Trish and Baltiboy: Part II		Damandeep Singh Baggan, Smita		8-Sep-17	2010	TV-Y7	58 min	Children & F	Animal minstrels narrate stories at		
80182621	Movie	Krish Trish and Baltiboy: The	Munjal Shroff, Tilak Shetty	Damandeep Singh Baggan, Smita		8-Sep-17	2013	TV-Y7	60 min	Children & F	The consequences of trickery are e		
80057969	Movie	Love	Gaspar No茅	Karl Glusman, Klara K	France, Belg	8-Sep-17	2015	NR	135 min	Cult Movies	A man in an unsatisfying marriage		
80060297	Movie	Manhattan Romance	Tom O'Brien	Tom O'Brien, Katherin	United State	8-Sep-17	2014	TV-14	98 min	Comedies, li	A filmmaker working on a docume		
80046728	Movie	Moonwalkers	Antoine Bardou-Jacquet	Ron Perlman, Rupert	France, Belg	8-Sep-17	2015	R	96 min	Action & Ac	A brain-addled war vet, a failing t		
80046727	Movie	Rolling Papers	Mitch Dickman		United State	8-Sep-17	2015	TV-MA	79 min	Documentar	As the newspaper industry takes a		
70304988	Movie	Stonehearst Asylum	Brad Anderson	Kate Beckinsale, Jim S	United State	8-Sep-17	2014	PG-13	113 min	Horror Movi	In 1899, a young doctor arrives at		
80057700	Movie	The Runner	Austin Stark	Nicolas Cage, Sarah P	United State	8-Sep-17	2015	R	90 min	Dramas, Ind	A New Orleans politician finds his		
80045922	Movie	6 Years	Hannah Fidell	Taissa Farmiga, Ben R	United State	8-Sep-15	2015	NR	80 min	Dramas, Ind	As a volatile young couple who h		
80244601	TV Show	Castle of Stars		Chaivanol Pupart, Jintanutda Lum		7-Sep-18	2015	TV-14	1 Season	International	As four couples with different lifes		
80203094	Movie	City of Joy	Madeleine Gavin		United State	7-Sep-18	2018	TV-MA	77 min	Documentar	Women who've been sexually bru		
80190843	TV Show	First and Last				7-Sep-18	2018	TV-MA	1 Season	Docuseries	Take an intimate look at the emot		
70241607	Movie	Laddaland	Sopon Sukdapisit	Saharat Sangkapreech	Thailand	7-Sep-18	2011	TV-MA	112 min	Horror Movi	When a family moves into an upsi		
80988892	Movie	Next Gen	Kevin R. Adams, Joe Ksander	John Krasinski, Charly	China, Cana	7-Sep-18	2018	TV-PG	106 min	Children & F	When lonely Mai forms an unlikel		
80239639	Movie	Sierra Burgess Is A Loser	Ian Samuels	Shannon Purser, Krist	United State	7-Sep-18	2018	PG-13	106 min	Comedies, R	A wrong-number text sparks a virt		

We obtained a dataset of Netflix Movies & TV shows from Kaggle and supplemented this with a dataset from IMDb that provides information on customer ratings and votes.

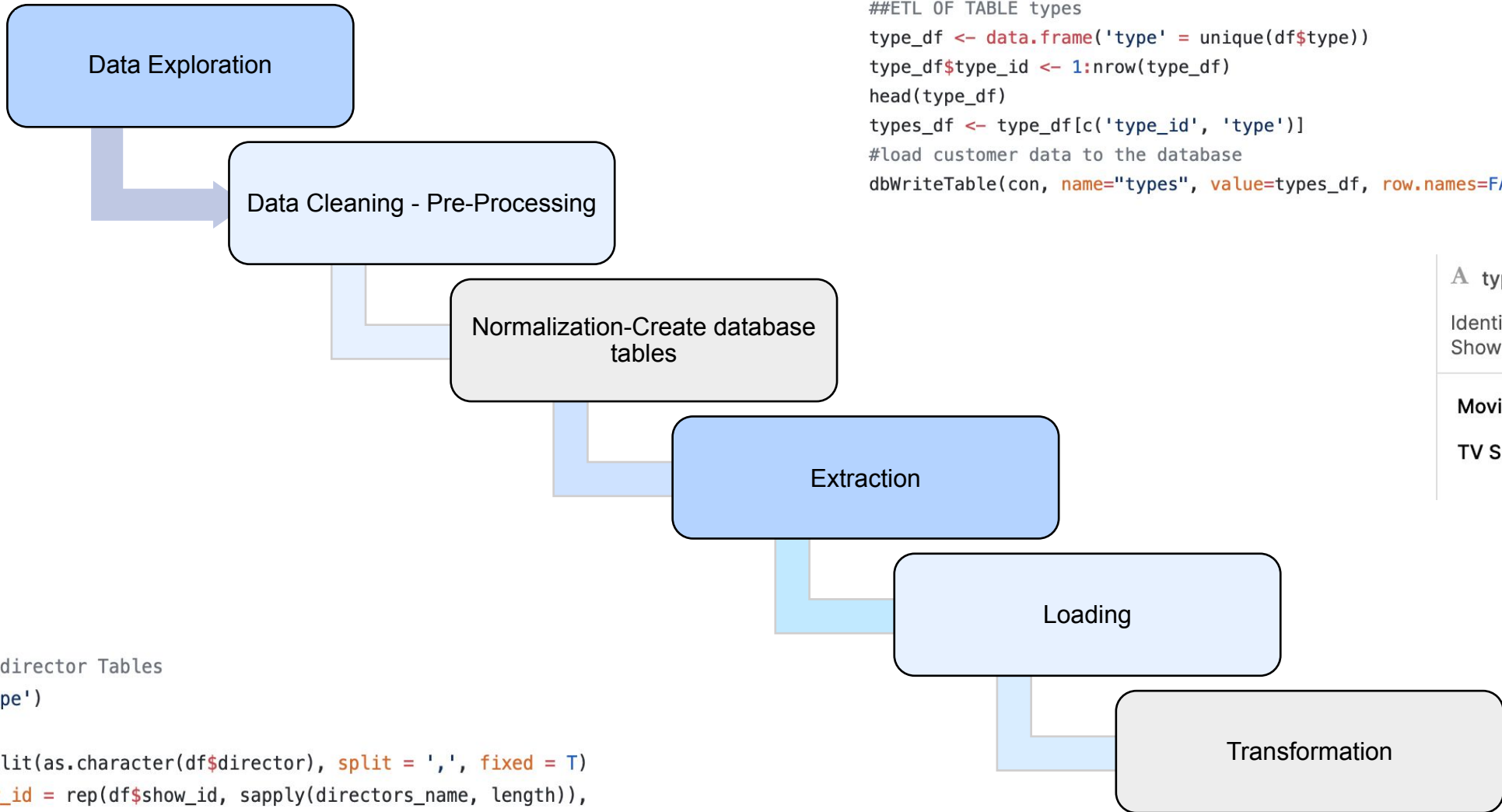
Normalization Plan



17 Tables

- Show
- date_addtoNetflix
- Customer_ratings
- Show_type
- Show_rating
- Show_genres
- Show_cast
- Show_countries
- Show_duration
- Movie_director
- Types
- TV_ratings
- Genres
- Actors
- Countries
- Duration
- Directors

ETL Process



```
##ETL OF TABLE types
type_df <- data.frame('type' = unique(df$type))
type_df$type_id <- 1:nrow(type_df)
head(type_df)
types_df <- type_df[c('type_id', 'type')]
#load customer data to the database
dbWriteTable(con, name="types", value=types_df, row.names=FALSE, append=TRUE)
```

A type	
Identifier - A Movie or TV Show	
Movie	68%
TV Show	32%

```
# directors and movie_director Tables
library('splitstackshape')
library(dplyr)
directors_name = strsplit(as.character(df$director), split = ',', fixed = T)
test = data.frame(show_id = rep(df$show_id, sapply(directors_name, length)),
                  name = unlist(directors_name))
```

Analytical Procedures

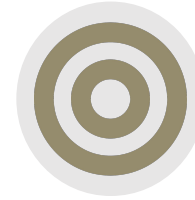


Customer Needs

- Optimize Content Management System and Recommendation System
- Improve future purchasing decision of TV/Movie

How

- We build 17 tables and each table contains one kind of specific information, such as country or type. Analysts can use those relations to find the relations and extract and analyze the data much easier.
- For the “C” level officers, it’s unnecessary to give the permission of accessing the database. We will set a reporting mechanism that does the summarization and visualization



Expected Insights

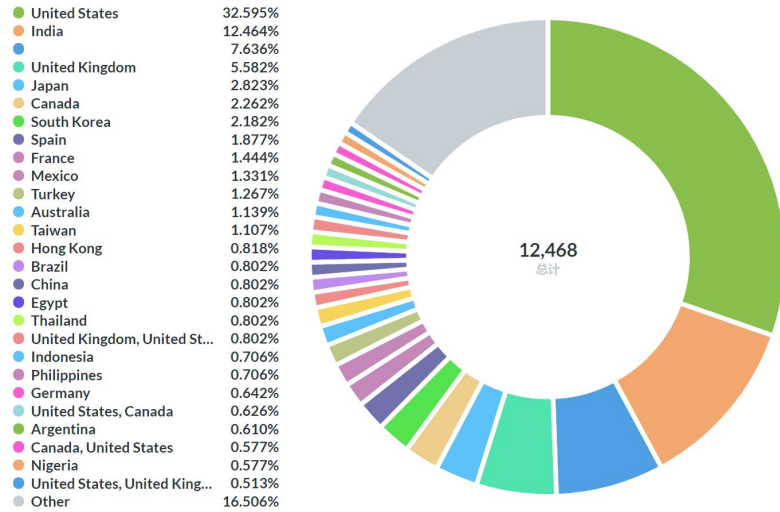
- Analyze the trends and provide an indication of the target audience and their preference
- Find the relationships between different variables
- Discover insights can also be used to improve Netflix marketing strategies



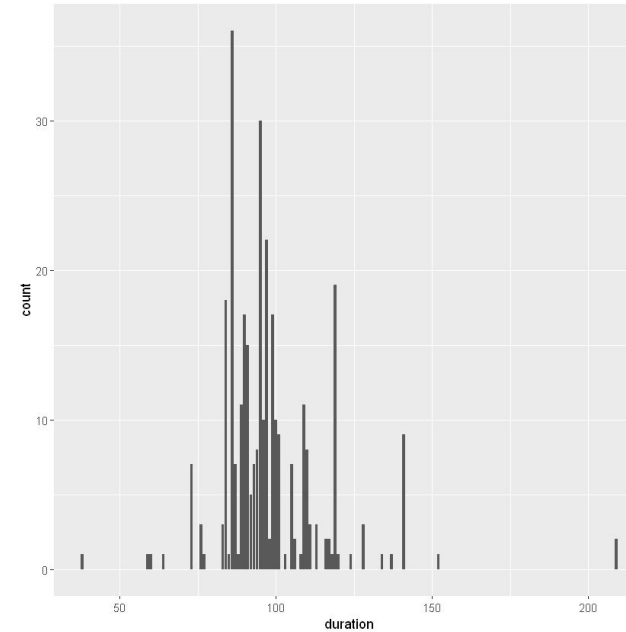
Interaction Demo

- Connect the Tableau with PostgreSQL on AWS: By typing the server name “cu-spring2020-group6.cggz75b61mlh.us-east-2.rds.amazonaws.com”
- Connect to the Metabase: <http://team06-metabase.herokuapp.com>

Analytical Applications



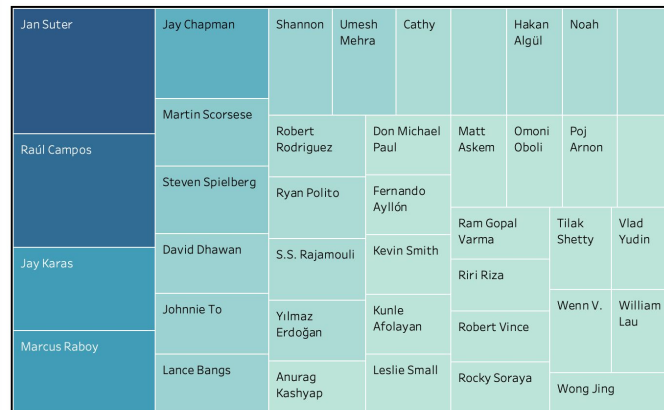
Movie distribution over countries



duration	count
86 min	36
95 min	30
97 min	22
119 min	19
84 min	18
90 min	17
99 min	17
91 min	15
109 min	11
89 min	11
100 min	10
96 min	10

Movie distribution over duration

Top 40 directors



Most popular directors

Analytical Applications

```
Call:
lm(formula = V2 ~ rating + genres, data = finaltable)
```

Residuals:

Min	1Q	Median	3Q	Max
-5.3916	-0.5638	0.0810	0.6917	3.6956

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.247524	0.211836	29.492	< 2e-16 ***
ratingNR	0.018665	0.201268	0.093	0.926114
ratingPG	-0.193578	0.199073	-0.972	0.330892
ratingPG-13	-0.070123	0.192540	-0.364	0.715725
ratingR	0.048406	0.189877	0.255	0.798781
ratingTV-14	-0.035363	0.188597	-0.188	0.851269
ratingTV-G	0.237382	0.251827	0.943	0.345904
ratingTV-MA	-0.004705	0.187607	-0.025	0.979993
ratingTV-PG	-0.020997	0.191324	-0.110	0.912614
ratingTV-Y	-0.586373	0.382364	-1.534	0.125195
ratingTV-Y7	-0.275689	0.260347	-1.059	0.289677
ratingTV-Y7-FV	-0.286005	0.279832	-1.022	0.306794
ratingUR	-0.372201	0.503995	-0.739	0.460240
genresAction	-0.467339	0.228742	-2.043	0.041088 *
genresAction, Adult, Crime	1.087840	1.134173	0.959	0.337524
genresAction, Adventure	0.449691	0.257538	1.746	0.080842 .
genresAction, Adventure, Animation	1.014207	0.155239	6.533	6.98e-11 ***
genresAction, Adventure, Biography	0.326583	0.660302	0.495	0.620903
genresAction, Adventure, Comedy	0.232362	0.182769	1.271	0.203657
genresAction, Adventure, Crime	0.790095	0.165921	4.762	1.96e-06 ***
genresAction, Adventure, Documentary	1.273474	1.134854	1.122	0.261845
genresAction, Adventure, Drama	0.934667	0.157373	5.939	3.03e-09 ***
genresAction, Adventure, Family	0.686939	0.287388	2.390	0.016866 *
genresAction, Adventure, Fantasy	0.664688	0.196418	3.384	0.000719 ***
genresAction, Adventure, History	1.857181	0.805372	2.306	0.021146 *
genresAction, Adventure, Horror	-0.765933	0.660186	-1.160	0.246024
genresAction, Adventure, Mystery	0.728726	0.355953	2.047	0.040677 *
genresAction, Adventure, Romance	0.222540	0.660312	0.337	0.736112
genresAction, Adventure, Sci-Fi	0.664664	0.218862	3.037	0.002401 **

	title	release_year	type	V2
3358	Pegasus	2019	Movie	9.1
5140	The Last Laugh	2019	Movie	9.1
1222	Domino	2019	Movie	9.0
1997	Homeland	2019	Movie	9.0
5138	The Last Laugh	2019	Movie	8.9
1548	For Love or Money	2019	Movie	8.8
1601	Fractured	2019	Movie	8.8
2276	Joy	2019	Movie	8.8
2450	Let's Dance	2019	Movie	8.8
5122	The King	2019	Movie	8.7

Top movies & TV shows in 2019

Non-personalized Recommendation

	title	release_year	type	V2
4827	The Gift	2019	TV Show	9.5
6261	What If?	2019	TV Show	9.5
293	Bad Blood	2019	TV Show	9.4
1837	Haunted	2019	TV Show	9.4
905	Chosen	2019	TV Show	9.3
6094	Undercover	2019	TV Show	9.2
5524	The Protector	2019	TV Show	9.1
6294	White Gold	2019	TV Show	9.1
5726	The Spy	2019	TV Show	9.0
187	Arrow	2019	TV Show	8.9

Linear Regression Model

INSIGHTS & CONCLUSION



Insight

- A user friendly Movie and Series profile database system that safely stores data in a structured format and provides departments with different authorization levels allowing them quick access to the information.
- Analysis of the relationships between different variables helps Netflix improve its recommendation system.



Interaction Authorization

- For Analysts, they can access the database and query or extract the data.
- For “C” level client, they will not be able to access the database directly in case they disorganized the data. Instead, there will be a report mechanism to help them summarize and visualize the report.



Recommendations for Investments

- Top 40 directors
- Genres: Drama, Adventure, Animation and Fantasy
- Movies with duration of 80 to 100 minutes
- TV shows with up to 3 seasons in most countries with exceptions in UK, USA & Japan where viewers are willing to watch up to 8 or 9 seasons.



Future Plan

- In the future, continue analyze relations between user's satisfaction rate and TV/Movies ratings to polish a personalized recommendation system
- Improve the content management system and their video sources.