



EDA: Netflix

Completed by:
Allan Gil and
Andres Jimenez-
Bonilla

Introduction

**How do we get organize data on shows and
most importanly how do we get to know
what people think of the shows thots what
we are going to find out today**

these are the columns we have

"title"	"rating"
"ratingLevel"	"ratingDescription"
"release_year"	"user_rating_score"
	"user_rating_size"

Who can watch

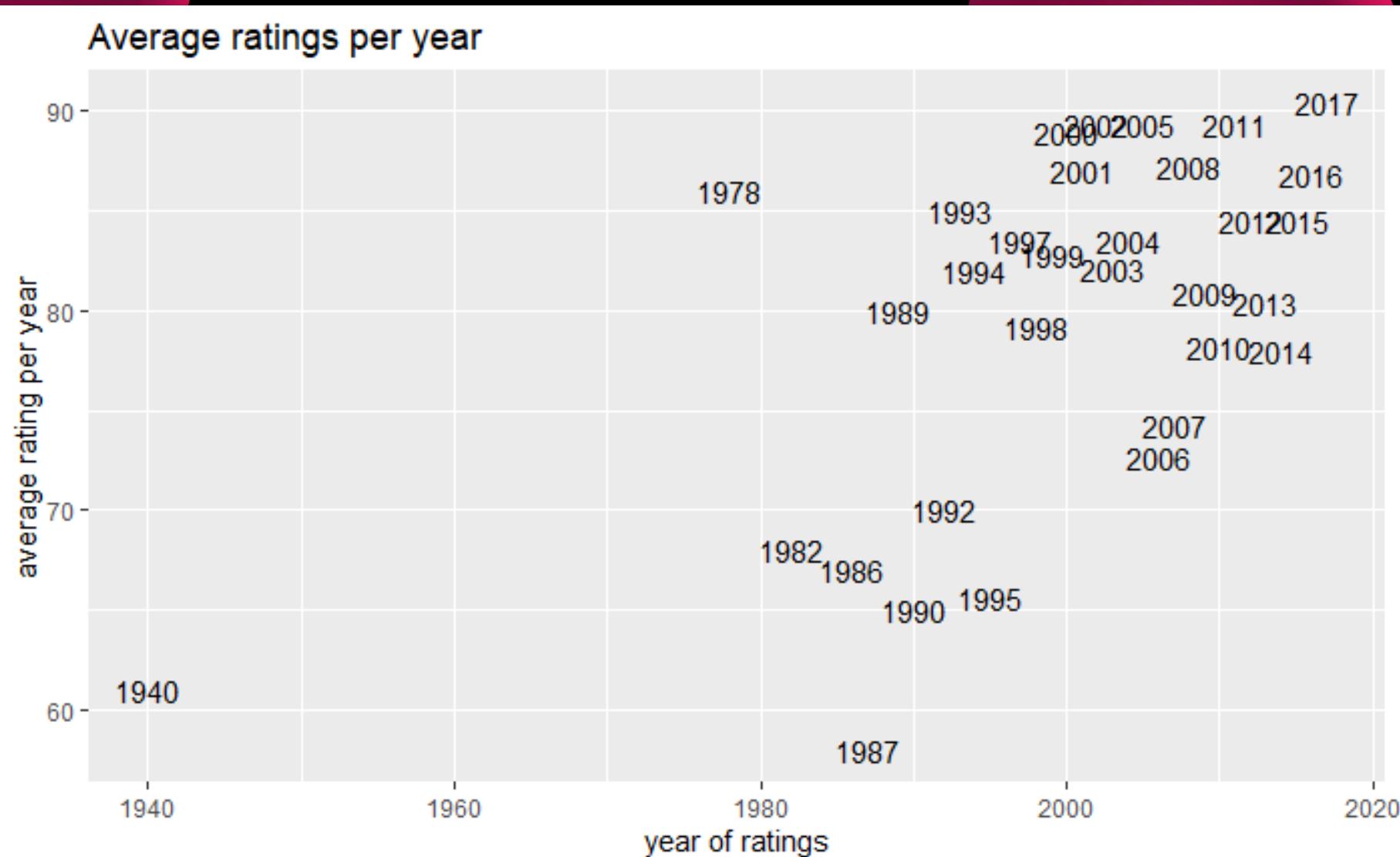
We wanted to find out what Netflix series young kids could watch so we posed the question; what are the newest Netflix series most suitable for children under 13?

```
{r}  
netflixShows %>%  
  select(title, rating, ratingLevel, ratingDescription, release_year) %>%  
  filter(ratingDescription <=70) %>%  
  arrange(desc(release_year))%>%  
  summarize(title, rating, ratingLevel)  
..
```

A tibble: 569 x 3

title	rating	ratingLevel
<chr>	<chr>	<chr>
One Day at a Time	TV-PG	Parental guidance suggested. May not be suitable for all children.
Dinotrux	TV-Y7-FV	Suitable for children ages 7 and older. Content contains fantasy violence.
Edgar Rice Burroughs' Tarzan and Jane	TV-Y7-FV	Suitable for children ages 7 and older. Content contains fantasy violence.
Dinotrux	TV-Y7-FV	Suitable for children ages 7 and older. Content contains fantasy violence.
Edgar Rice Burroughs' Tarzan and Jane	TV-Y7-FV	Suitable for children ages 7 and older. Content contains fantasy violence.
Voltron: Legendary Defender	TV-Y7-FV	Suitable for children ages 7 and older. Content contains fantasy violence.
Buddy Thunderstruck	TV-PG	Parental guidance suggested. May not be suitable for all children.
Dragons: Race to the Edge	TV-Y7-FV	Suitable for children ages 7 and older. Content contains fantasy violence.
Dawn of the Croods	TV-Y7	Suitable for children ages 7 and older
Dinotrux	TV-Y7-FV	Suitable for children ages 7 and older. Content contains fantasy violence.

what year has the the most movies with the highest average rating



What Netflix series have no age rating

we wanted o know what Netflix series have no ratings so we formulated this code

```
{r}  
netflixShows %>%  
  select(release_year, rating, title) %>%  
  filter(rating == "NR")  
  ...
```

A tibble: 14 x 3

	release_year <int>	rating <chr>	title <chr>
	2016	NR	The 100
	2013	NR	Nymphomaniac: Volume 1
	2013	NR	Hunter X Hunter (2011)
	2014	NR	Sex, Love and Therapy
	2016	NR	Stronger Than the World
	2016	NR	Back and Forth
	2016	NR	The 100
	2016	NR	Almost Adults
	2014	NR	Honeytrap
	2016	NR	Like Cotton Twines

only for adults

Lastly we wanted to discover which content is just for adults to separate the content

title	rating	ratingDescription
Breaking Bad	TV-MA	110
The Walking Dead	TV-MA	110
Black Mirror	TV-MA	110
The 100	NR	124
<ed><f1> Pai, <ed><f1> - Look at This	TV-MA	110
Nymphomaniac: Volume 1	NR	124
Hunter X Hunter (2011)	NR	124
Marvel's Luke Cage	TV-MA	110
Marvel's Iron Fist	TV-MA	110
Narcos	TV-MA	110

Conclusion

After our data exploration, we were able to gain some insight into how release year, and content rating, affect the review scores. This project has helped me and my team adapt to an online environment and github has helped with this process. In this project we further learned important R skills.