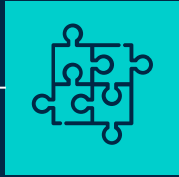




NETFLIX DATA ANALYSIS 2020

Presented By:
Ahmed Meshref
Paul Arah

TABLE OF CONTENTS



01

PROBLEM & SOLUTION

Overview of Netflix
and problem
proposed



02

OUR PROCESS

Data Analysis
processes taken to
solve the problem



03

SUMMARY

Visualizations and
findings summary

CONTEXT

The Rwanda Revenue Authority(RRA) is a financial services institution located in Kigali Rwanda. Their mandate is to collect tax revenue on behalf of the Republic of Rwanda. Due to the prevailing COVID-19 pandemic, its tax collection figures for the 2020-2021 fiscal year deteriorated by over 10% to a value of \$1.3 Billion. They are now planning in advance to improve their fiscal tax revenue for 2021-2022. A key goal of theirs is to invest in digital transformation. In line with their digital agenda, the institution is planning to establish an effective revenue collection platform to influence their decision making processes. They want to gain better insights from the vast amount of data they collect daily as well as from the media and other platforms. RRA published a Request for Proposal(RFP) which was submitted to tech companies that specialize in data analytics.

UNDERSTANDING THE PROBLEM

Due to the harsh economic realities of the COVID-19 pandemic, Netflix aims to understand its commercial performance in 2020 compared to the last five years. This is achieved by performing exploratory data analysis on Kaggle's updated Netflix data and exploring viable metrics like shows trends, countries with the most movie production, etc.

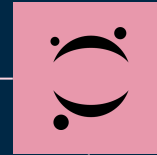
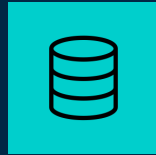
NETFLIX

Netflix is a streaming service that offers a wide variety of award-winning TV shows, movies, anime, documentaries, and more on thousands of internet-connected devices. (Netflix, 2021)



DATA AND TOOLS

DATASET
Latest Netflix
dataset from kaggle



JUPYTER
Interactive
computing tool

ANALYSIS TOOLS
Numpy & Pandas



VISUALIZATION
Matplotlib & Seaborn

ANALYSIS PROCESS

Understanding the
dataset

EXPLORING

01

02

CLEANING

Cleaning the dataset

Perform analysis &
Visualization

ANALYSIS

03

04

SUMMARY

Findings & Summary



01

UNDERSTAND THE DATA

DATASET COLUMNS



Type: Show Type
- Movie
- TV Show



Country: country
produced the movie



Release_year: the original
year when the show was
produced.



Title: Show Title



Date_added: the date
when the show was
added to Netflix (mm -
day - year)



Listed_in: all genres which
the movie relates to

Other self-explanatory columns: **cast**, **description**, **rating**, and **director**.

MORE INSIGHTS

	COUNT	Function	DESCRIPTION
--	-------	----------	-------------

NULL VALUES	~ 3000	<code>df.isnull().sum()</code>	Around 3000 missing values were found across multiple columns.
-------------	--------	--------------------------------	--

DEPLICATES	0	<code>df.duplicated().any()</code>	No duplicates were found
------------	---	------------------------------------	--------------------------

UNIQUE ROWS	7789	<code>df.nunique()</code>	All rows were unique
-------------	------	---------------------------	----------------------



02

DATA CLEANING

CLEANING

Function

DESCRIPTION

DROP NULL

`df.dropna()`

Drop null values to create more robust analysis

DATE FORMAT

`pd.to_datetime(df['date_added'])`

Change date format of the date_added column

CREATE YEAR_ADDED
COL

`df['date_added'].
dt.year`

Extract the year added into a new column for a show as it is required in the analysis.

CLEAN DATA HEAD

show_id	type		title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	year_added
1	s2	Movie	7:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	2016-12-23	2016	TV-MA	93 min	Dramas, International Movies	After a devastating earthquake hits Mexico Cit...	2016
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence ...	Singapore	2018-12-20	2011	R	78 min	Horror Movies, International Movies	When an army recruit is found dead, his fellow...	2018
3	s4	Movie	9	Shane Acker	Elijah Wood, John C. Reilly, Jennifer Connelly...	United States	2017-11-16	2009	PG-13	80 min	Action & Adventure, Independent Movies, Sci-Fi...	In a postapocalyptic world, rag-doll robots hi...	2017
4	s5	Movie	21	Robert Luketic	Jim Sturgess, Kevin Spacey, Kate Bosworth, Aar...	United States	2020-01-01	2008	PG-13	123 min	Dramas	A brilliant group of students become card-coun...	2020
5	s6	TV Show	46	Serdar Akar	Erdal Beşikçioğlu, Yasemin Allen, Melis Birkan...	Turkey	2017-07-01	2016	TV-MA	1 Season	International TV Shows, TV Dramas, TV Mysteries	A genetics professor experiments with a treatm...	2017

DATA ANALYSIS & VISUALIZATION

03

ANALYSIS QUESTIONS

Compare the number of Movies & Shows **added** to the website from year 2008-2021.

Compare number of movies & TV Shows **released** in each one of the last 5 years.

Show movies VS TV shows **trend** in the last 5 years at Netflix.

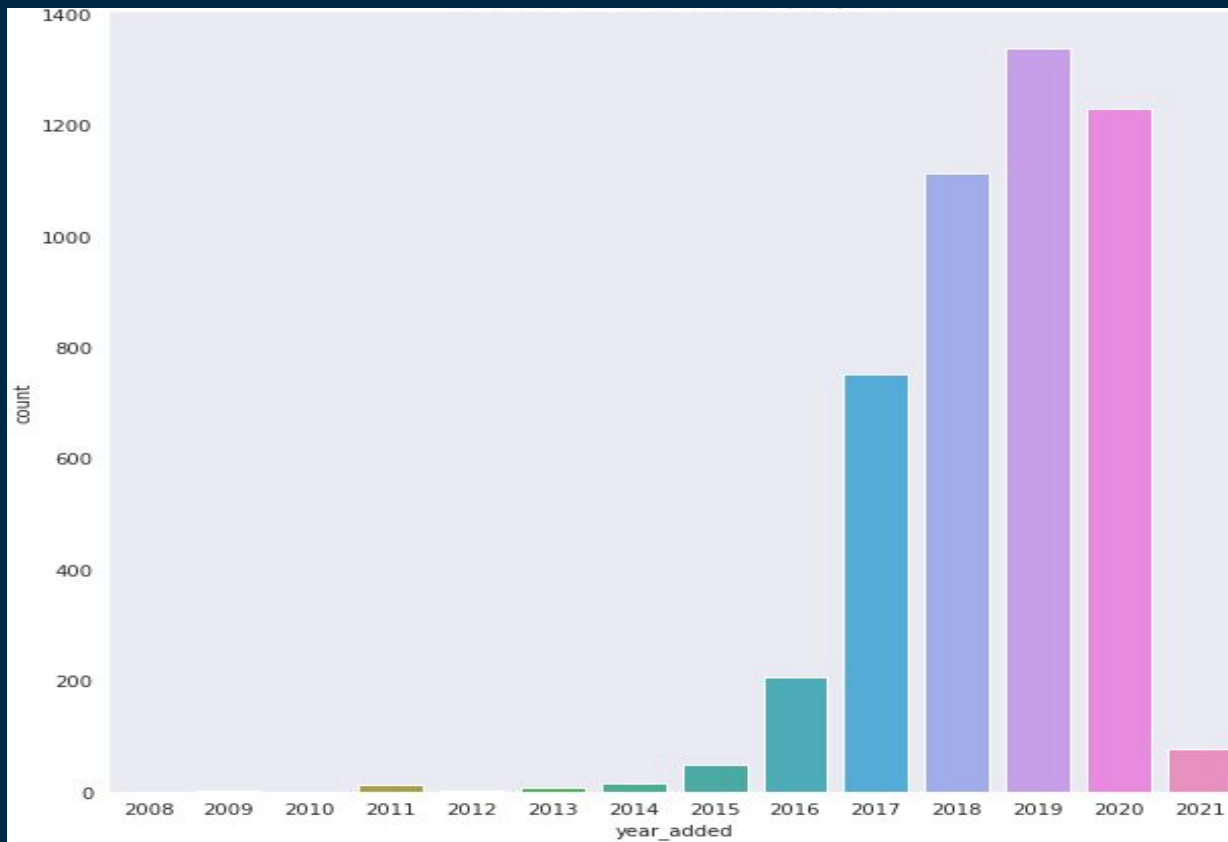
What is the number of movies VS TV shows in 2020?

Which are the top **5** countries that produces movies in the last 5 years?

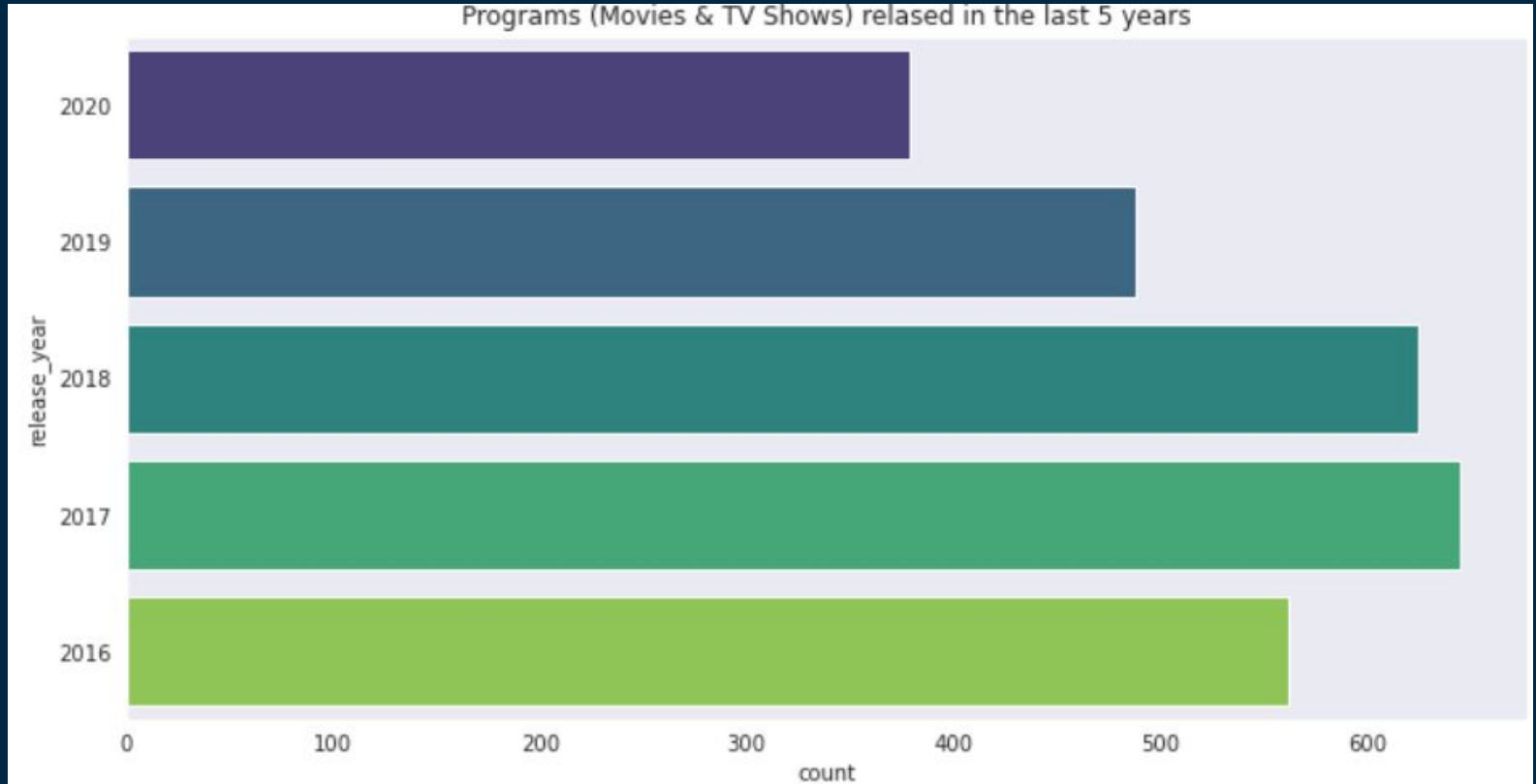
Analyze the **US contribution** to movies and TV shows in the last 5 years.



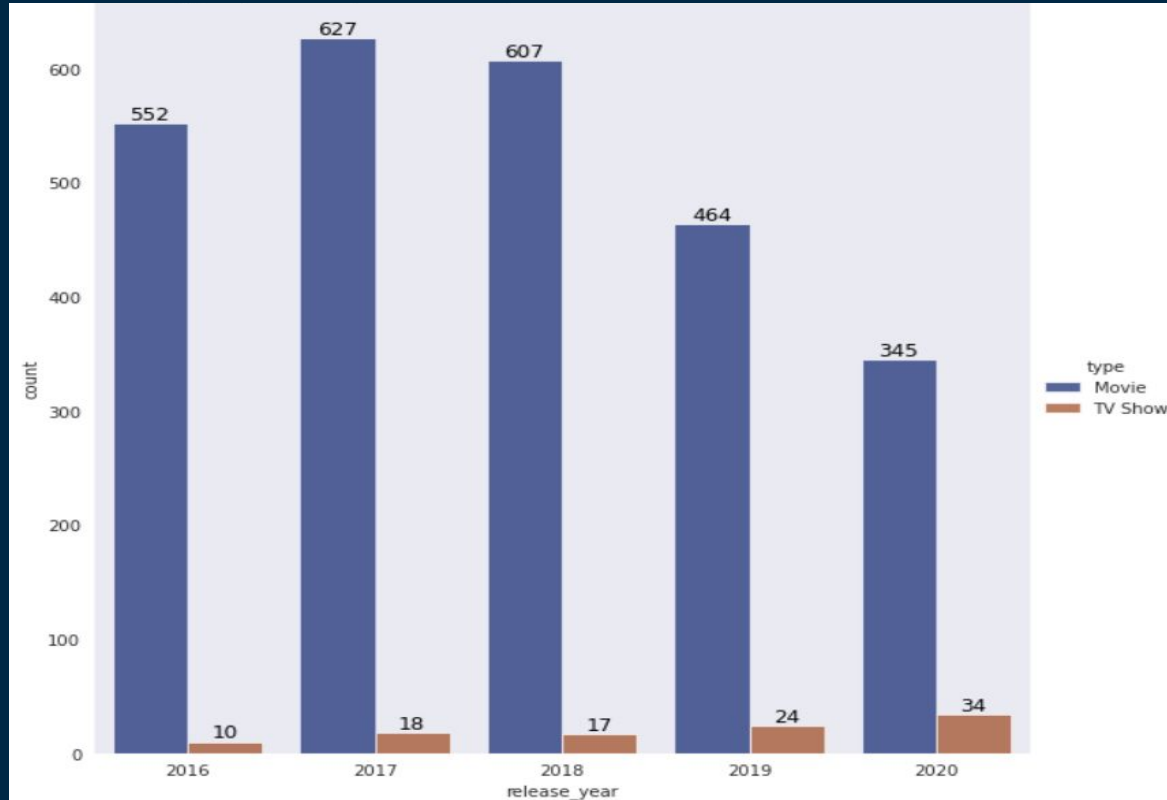
MOVIES & SHOWS ADDED 2008-2021



MOVIES & SHOWS RELEASED 2016-2020

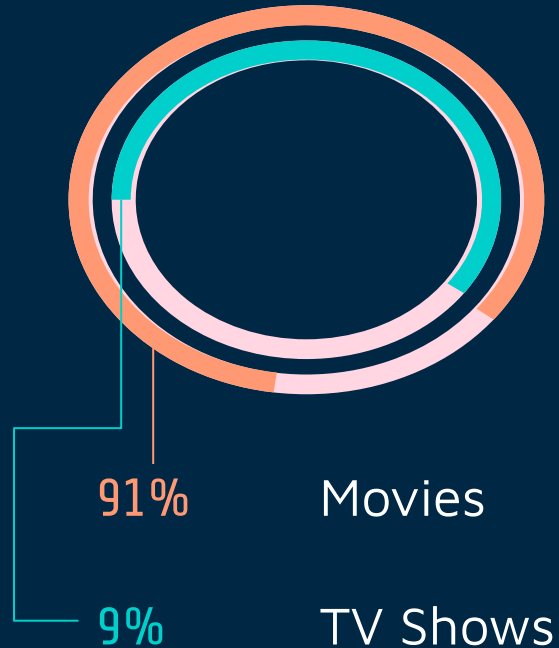


SHOWS VS MOVIES TREND

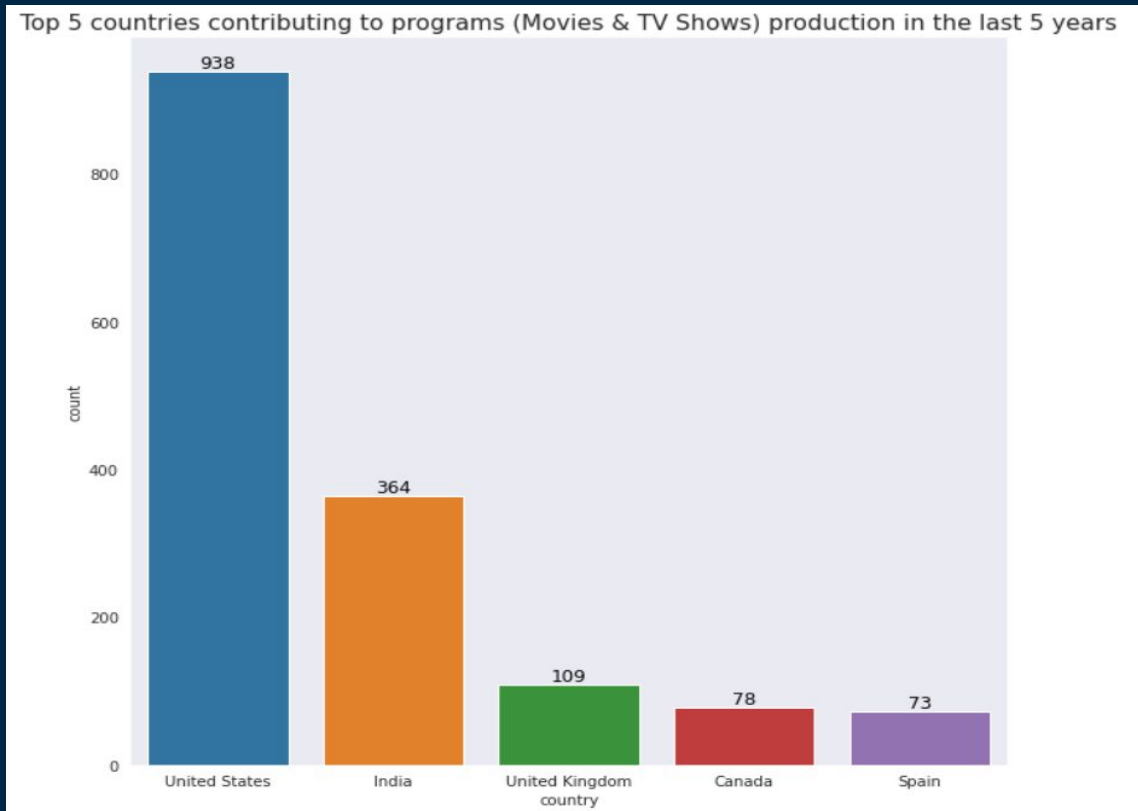


SHOWS VS MOVIES 2020

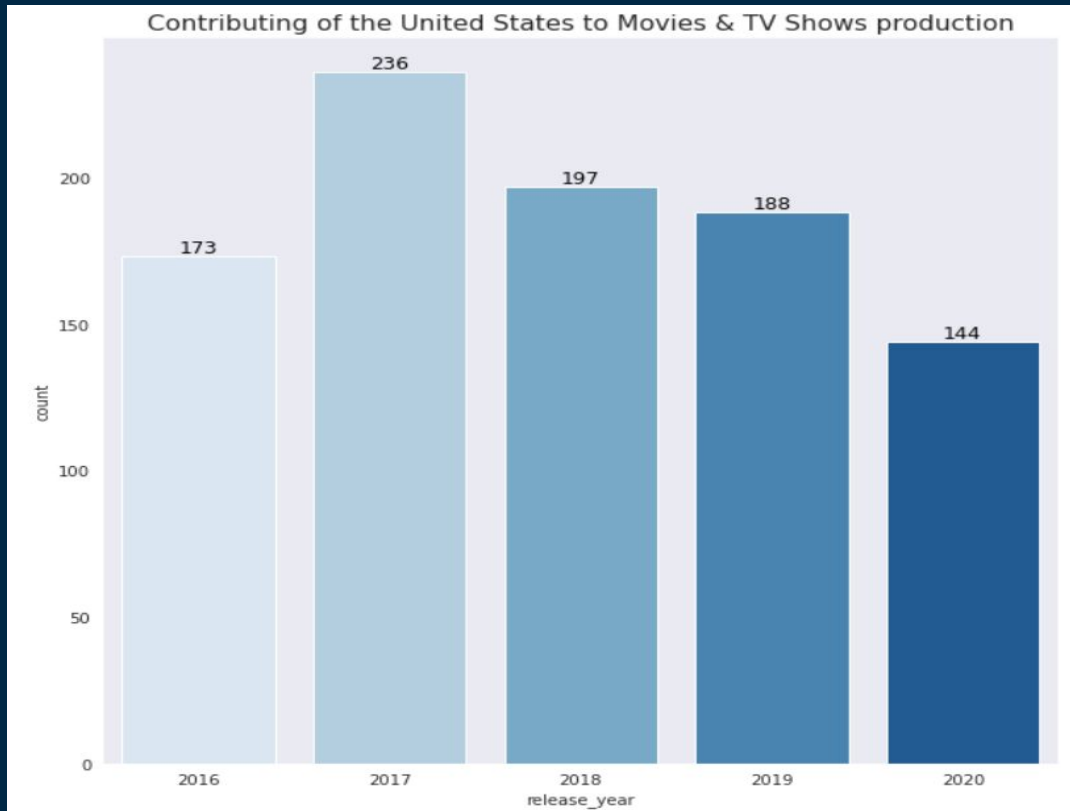
SHOWS TYPES



TOP CONTRIBUTORS 2016-2020



US CONTRIBUTION 2016-2020



KEY FINDINGS

04

BASED ON A SURVEY

85.00%

Our understanding of your future is
strongly anchored and the
future looks bright.

Presentations are
comprehensive and

TODAY'S HIGHLIGHTS

Final Introduction
Presentations
Agreement on Survey Process
Survey on Survey and Assessment
Presentations to Regulators
Survey Process
Regulatory and Assessment
Final Words

Discussion
Outline

Presentations by Regulators (10:00)

1229

Shows were added to the netflix in 2020 compared to 1337 in 2019. Shows added in 2020 are 8% less than shows added in 2019.



SHOW PRODUCTION TREND



645 shows produced



488 shows produced

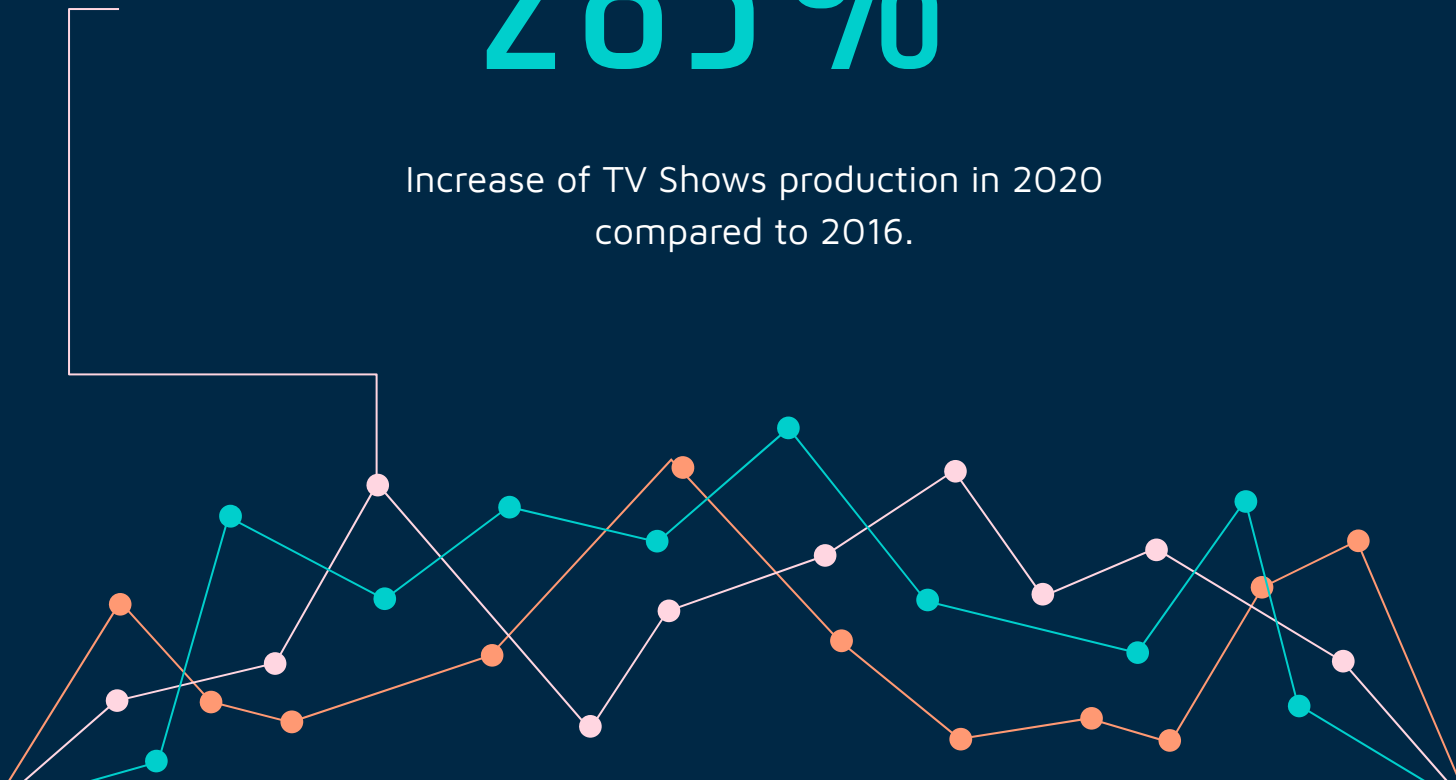


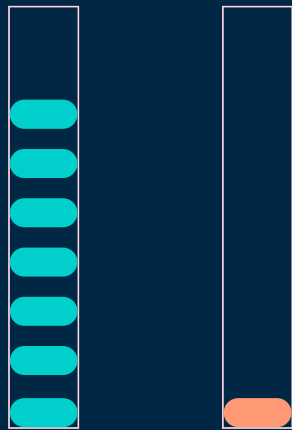
379 shows produced

2020 witnessed 41.24% reduction in the number of produced shows compared to 2018.

285%

Increase of TV Shows production in 2020
compared to 2016.





91%

Movies

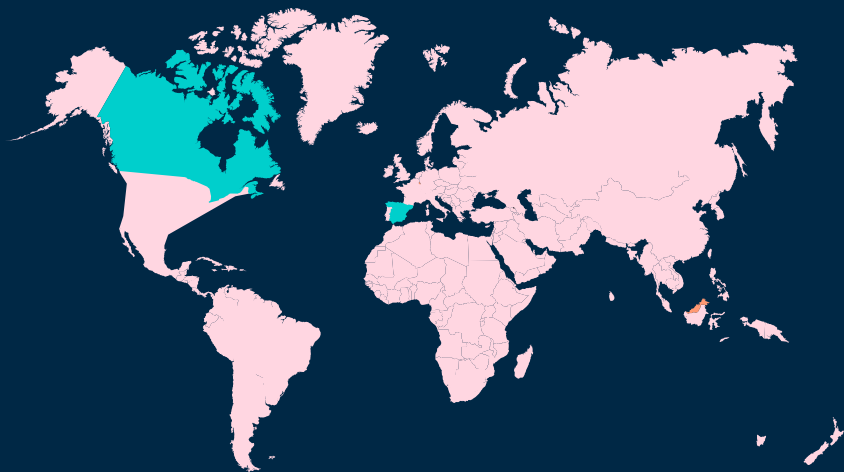
9%

TV Shows

Movies production is higher than TV Shows production by 82% in 2020.



TOP CONTRIBUTORS IN 2020



38%
US



14%
INDIA

US and INDIA contributed to more than
half of the shows released in 2020

SUMMARY

In conclusion, we explored the following fields in the dataset: The show types, countries, release year for movies. Title of different shows, dates, etc. Using the the right data cleaning and exploratory data analysis techniques, we were able to remove redundant values and derive insights from the data. Movies had a 91% prevalence over TV shows with 9%. 2019 was the year with highest number of movies added while 2017 was the year with the highest release of new movies. The United states is the highest contributor, followed by India. Netflix has seen a whooping 285% growth in TV shows since 2016 till date.

By netflix publicly available data that shows a growing revenue over the years, we can conclude that this increase of TV shows most likely has translated into an increase in revenue for netflix.



Do you have any questions?

a.meshref@alustudent.com

ALU

THANK YOU!



REFERENCES

1- "Unlimited movies, TV shows, and more.," *Netflix*. [Online]. Available: <https://www.netflix.com/eg-en/>. [Accessed: 07-Mar-2021].

2- S. Bansal, "Netflix Movies and TV Shows," *Kaggle*, 18-Jan-2021. [Online]. Available: <https://www.kaggle.com/shivamb/netflix-shows>. [Accessed: 07-Mar-2021].