

TD1

Data Visualization

Alexandre Cogordan



Table of contents

Table of contents	2
Exercise 0	3
Exercise 1	4
Age and gender demographics that have the most sales	4
Exercise 2	6
Question 2 - Comparative analysis of films and television shows	6
Question 3 - Month with the lowest amount of content	10
Question 4 - Do a movie audience analysis	11
Question 5 - Other analysis	16

Exercise 0

2.1 What type of data concerns the Country and the New Case respectively (Nominal, Ordinal, Interval and Ratio)?

The data is nominal.

2.2 What type of chart do you think we should use?

A bar chart.

2.3 Why do you choose this type of chart?

Because the data is nominal and therefore is separate but doesn't have a particular order.

2.4 After sorting, please indicate the three countries with the most New Cases.

France, Germany, Italy.

3.1 What default data type is the Year Week value?

Ordinal.

3.2 Could we use it directly? If not, what data type do we need?

No, we would need to use new cases or an iteration that uses new cases.

3.3 What type of chart do you think we should use?

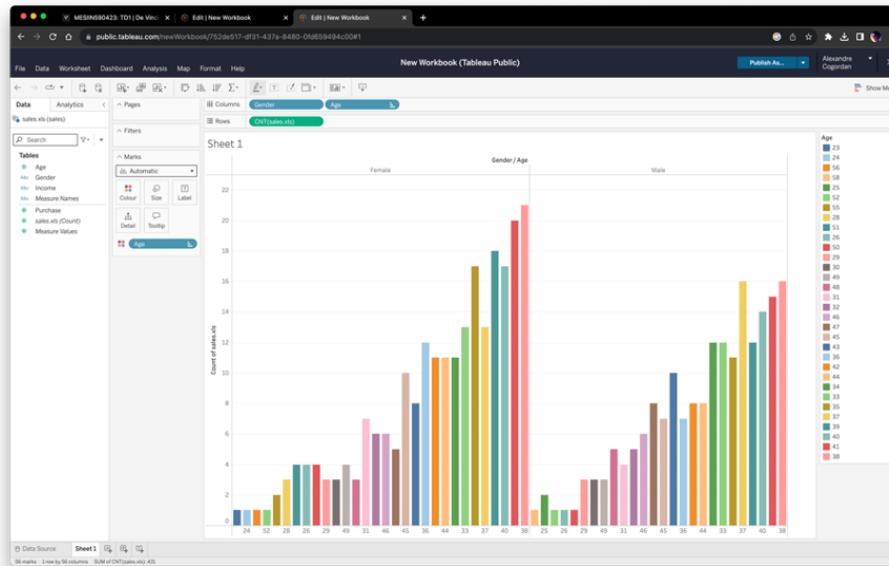
A line graph.

5. Which month has the most New Case in 2022?

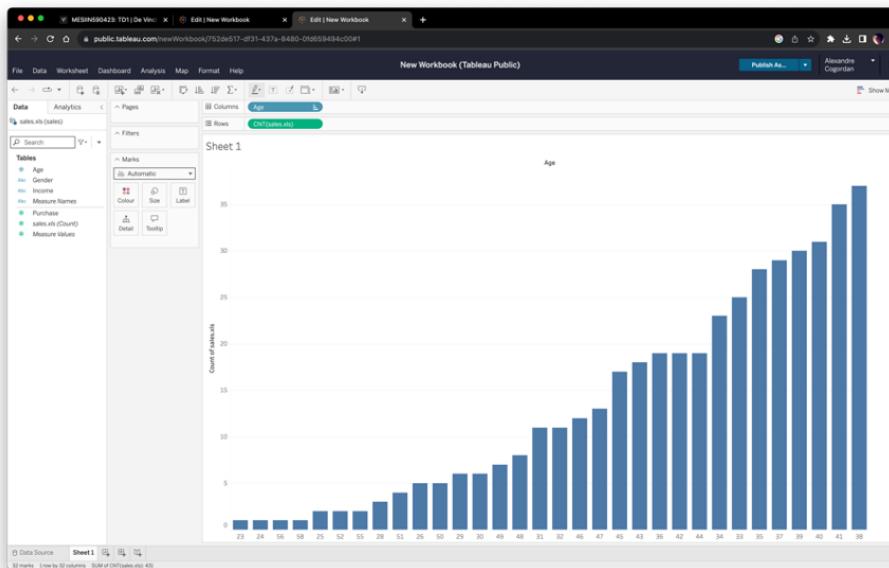
March.

Exercise 1

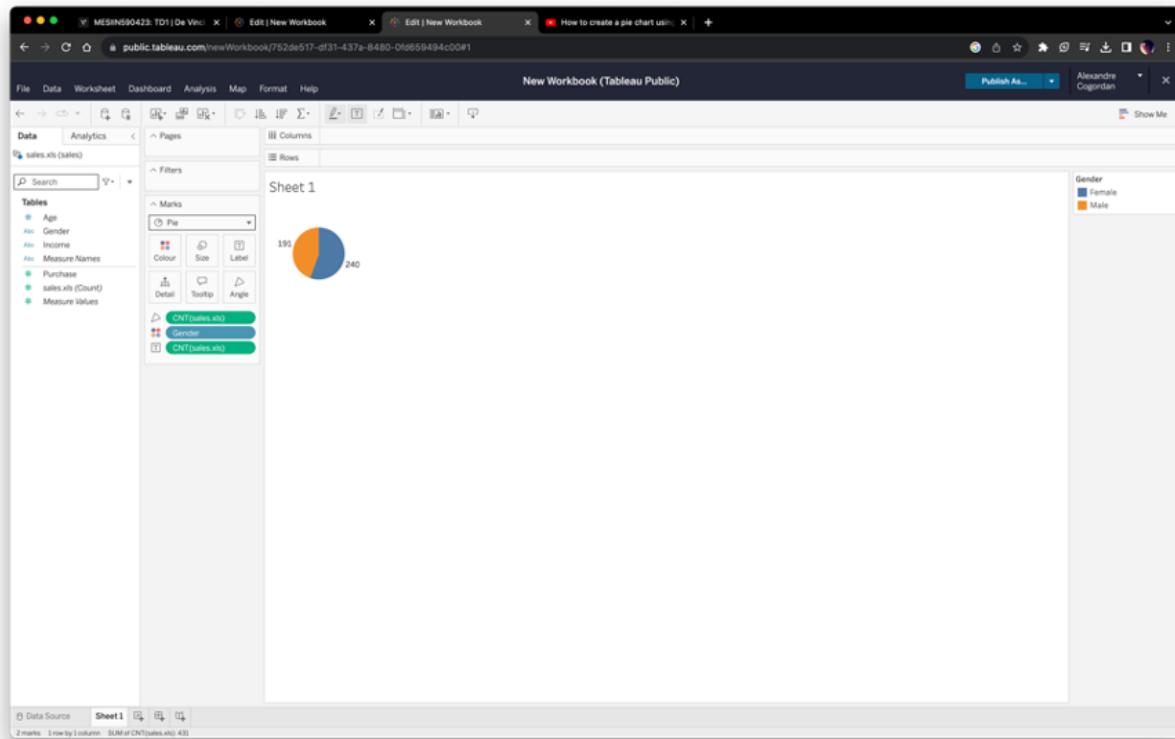
Age and gender demographics that have the most sales



Ascending order of sales by age bar graph separated by gender



Ascending order of sales by age

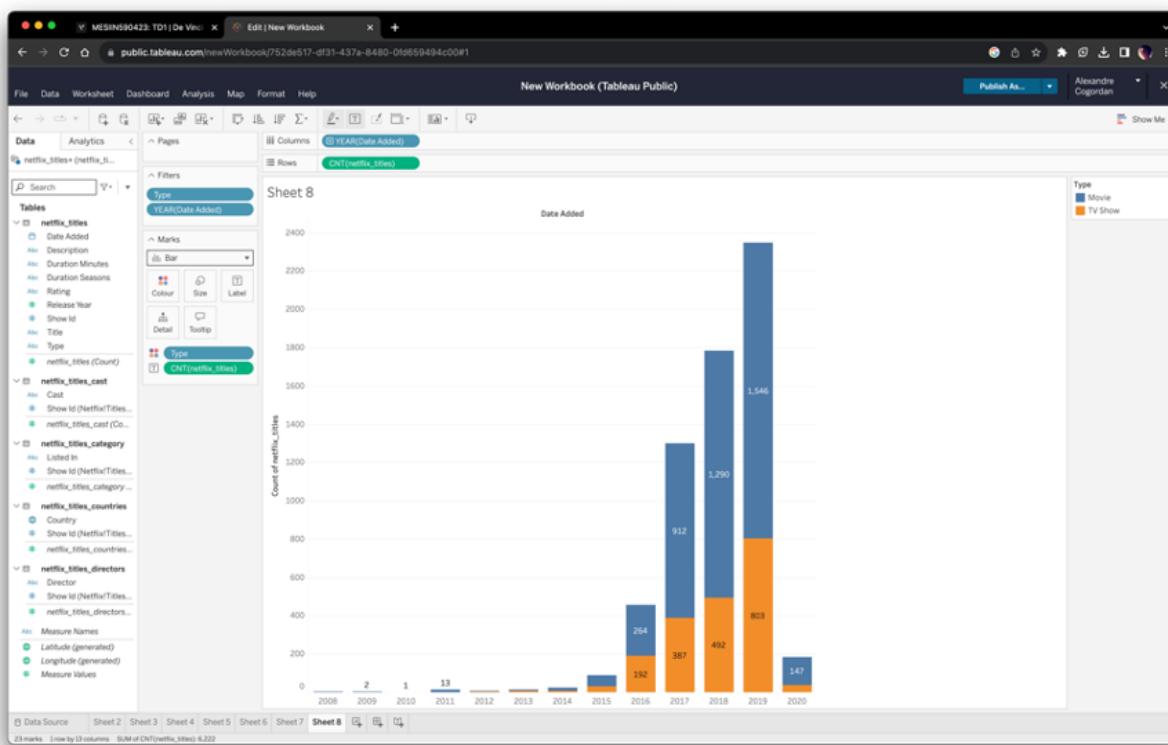


Total number of sales by gender

Exercise 2

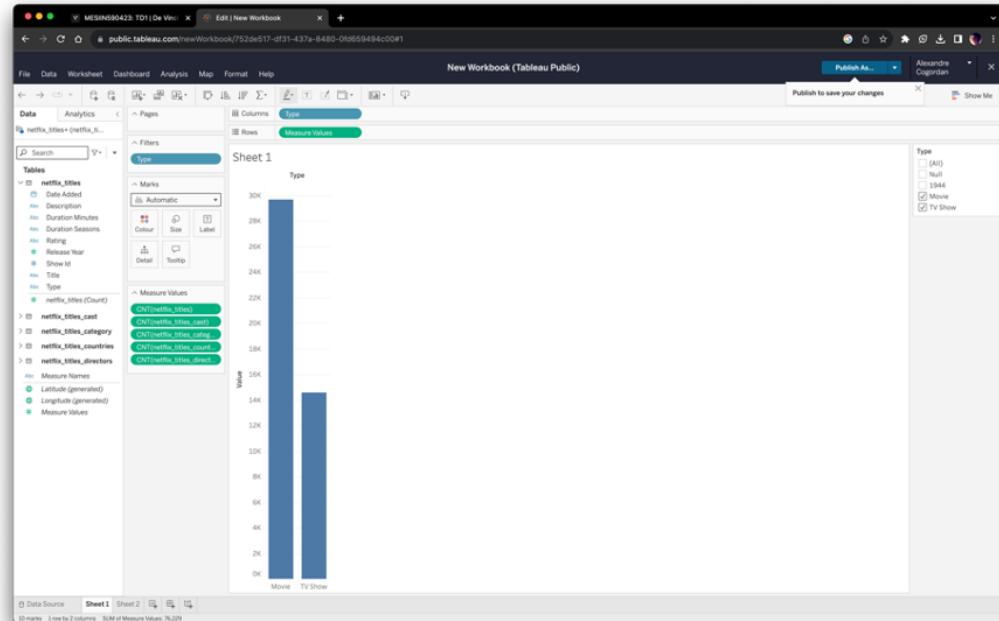
Question 2 - Comparative analysis of films and television shows

To start off our comparative analysis, we thought it would be good to introduce the number of movies and series added each year. This would allow us to not only a perspective of the evolution in content addition on the website, but also potential favouritism for types of content. From what we can see, the number of content added increases each year. We also realise that there are far more movies added than series, but this seems normal considering series tend to stay for a longer period of time and newer series are less prominent than new movies. Nevertheless, the ratio is not totally uneven which would seem to emphasise the importance of keeping and adding both types of content on the platform.

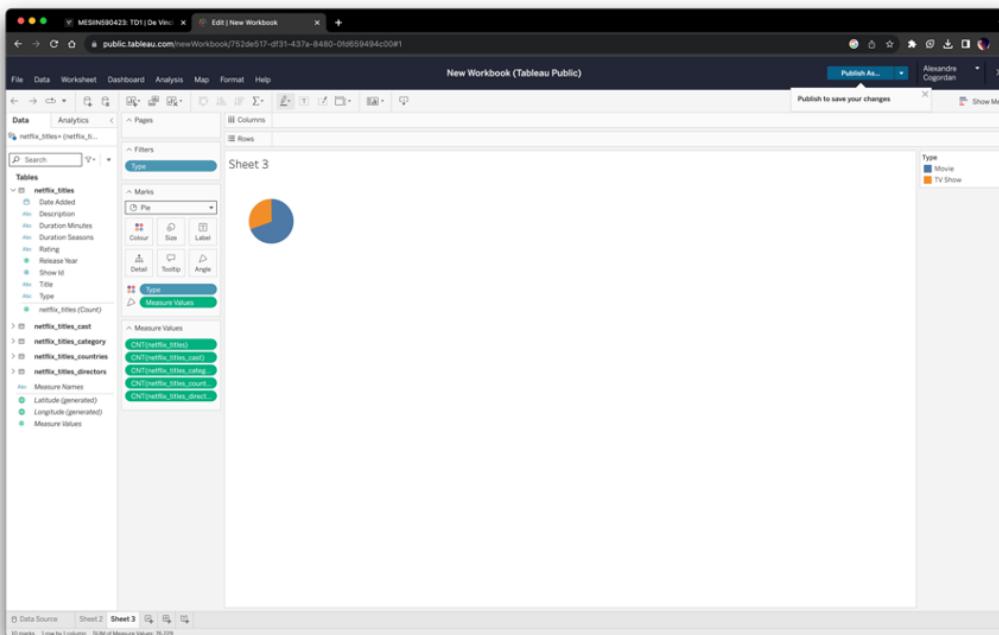


Movies and series added on Netflix each year

To further compare the difference in size between the two types, we also compared their total counts against each other in both a bar graph and pie chart:

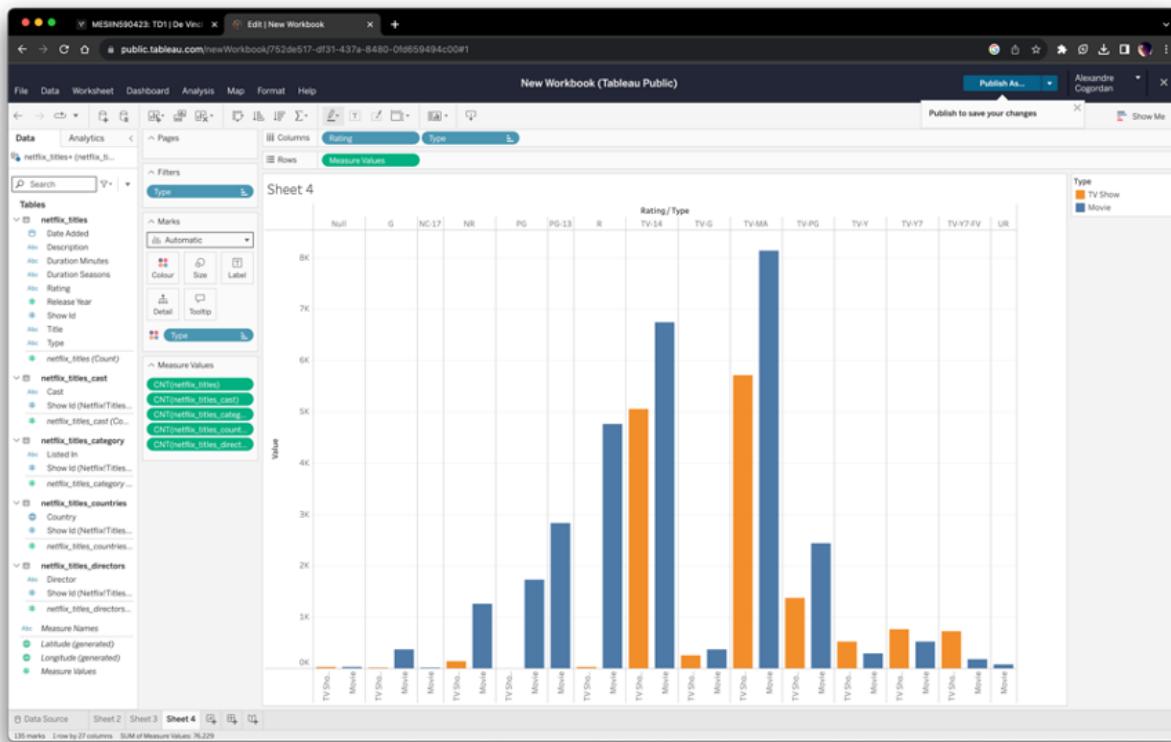


Bar graph for the count of movies and series



Pie chart for the count of movies and series

Thereafter, we thought about how movies and series could differ in audience ratings:



Bar graph of movies and TV Series by audience ratings

Seeing this graph allows us to see if the usual audiences change by types of content. To give a general examination of our results, it seems that the TV series present on Netflix are more focused towards children than movies (where they thrive in TV-Y and TV-7). On the other hand, movies are comparatively more adapted to older audiences as they win over the R and TV-MA ratings.

Naturally, we'd thought otherwise considering we are students with R and TV-MA rated series suggested to us, but we've come to realise that there are a great number of kids' shows that explain this phenomenon.

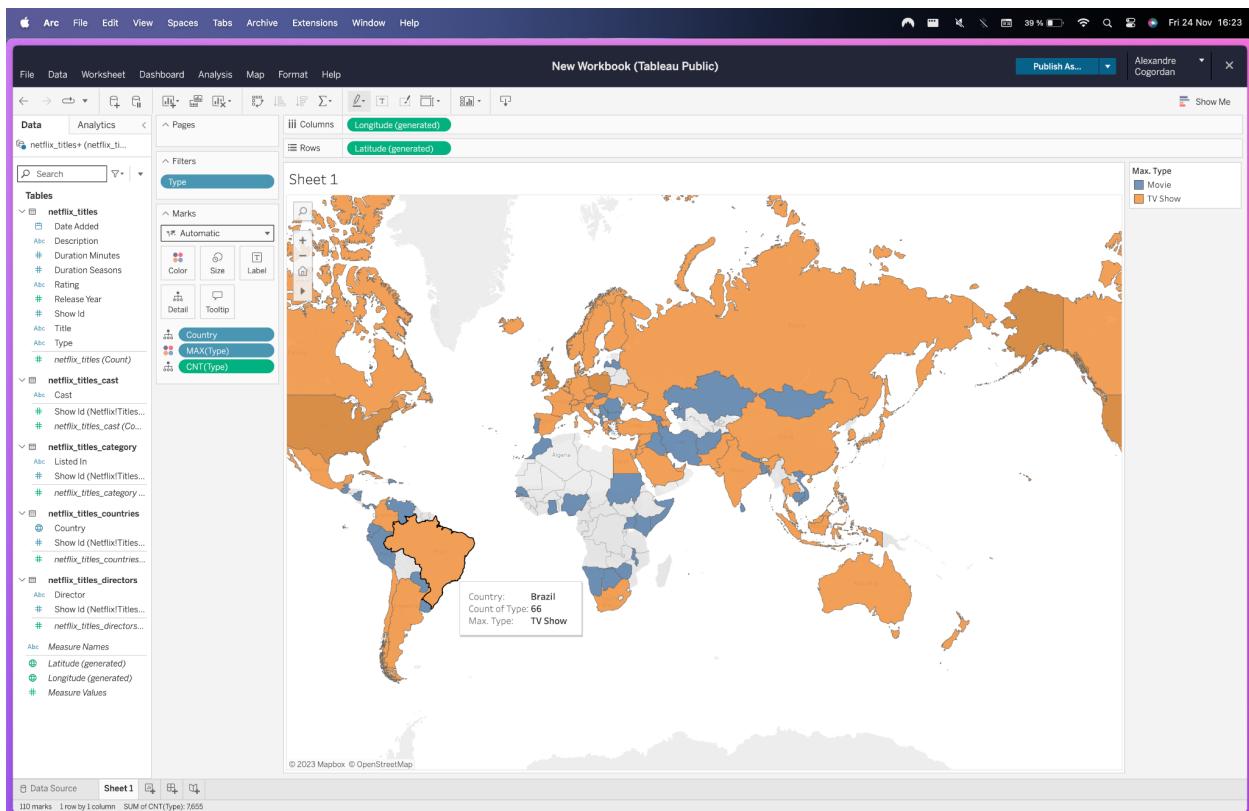


TV Series



Movies

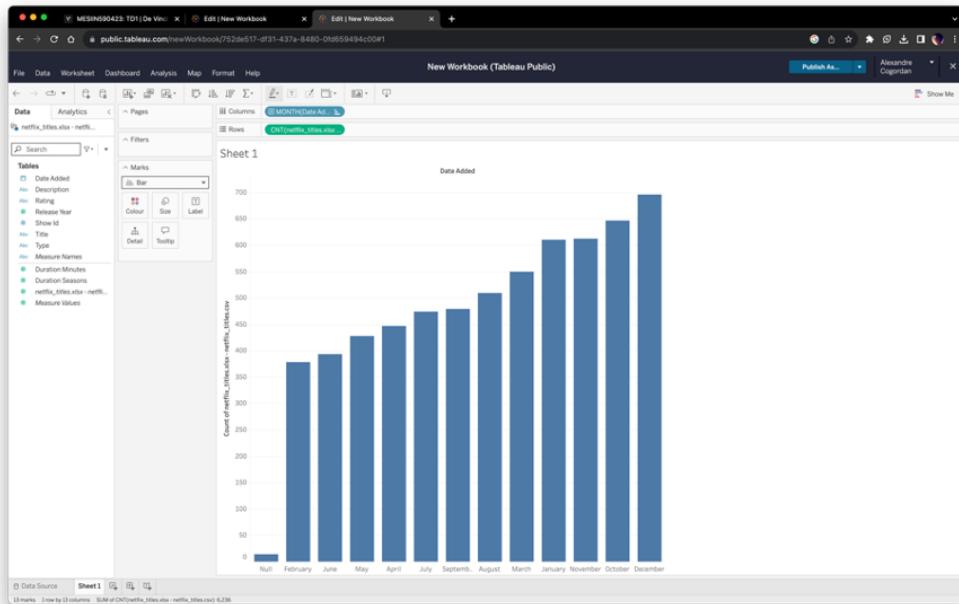
Most popular target audience by type of content



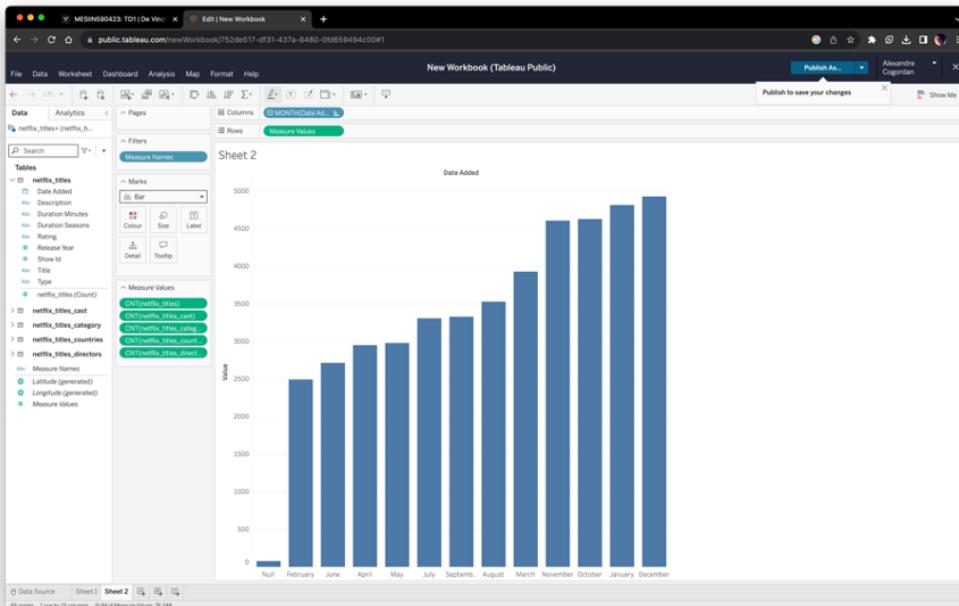
Favourite type of show by country

We've finished by looking using our geographic data. We've managed to count the number of movies and tv shows by country and find out which one was the most popular amongst them. As an example, we discovered that the Brazilian population had more tv shows available on its Netflix platform, this may be explained by the popularity of tv shows carried by its population. In most cases, we concluded that tv shows are more available than movies in most countries' Netflix platforms.

Question 3 - Month with the lowest amount of content



Graph made using the csv file format

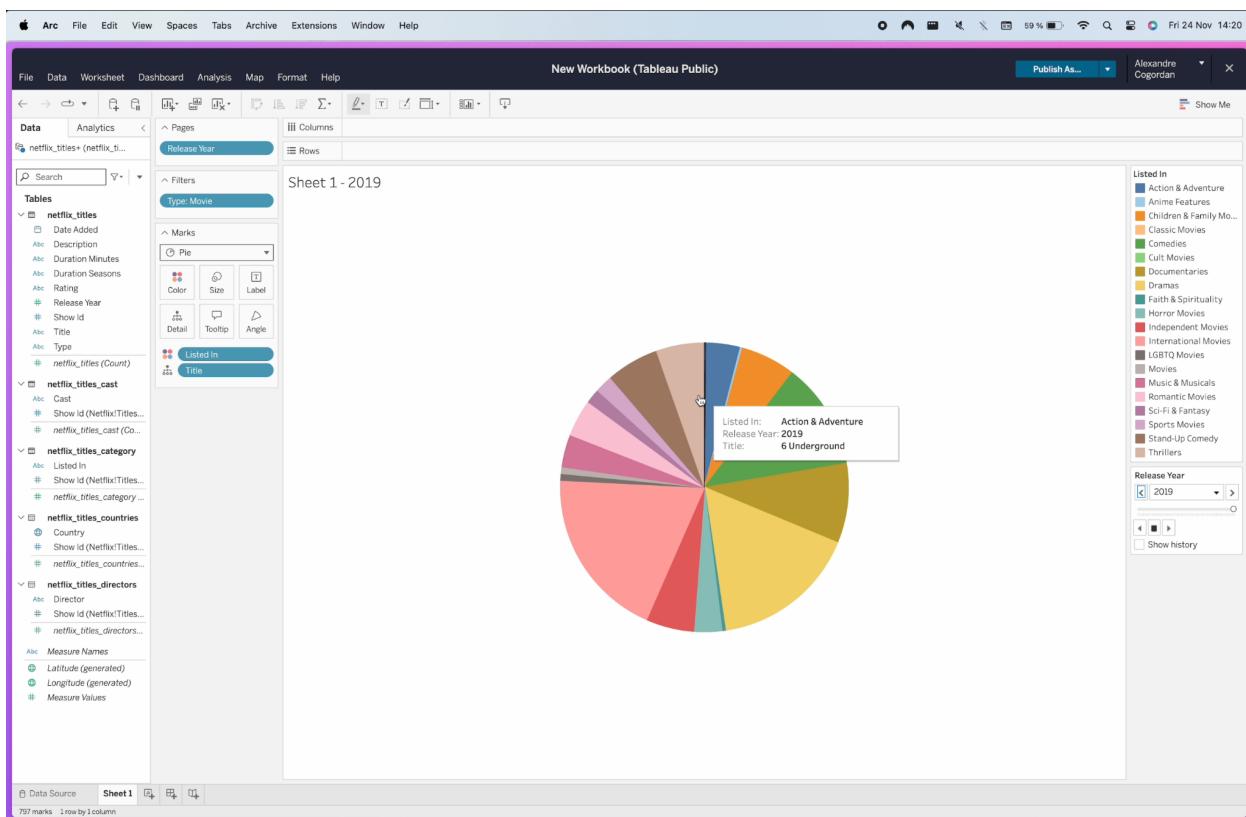


Graph made using the xlsx file format

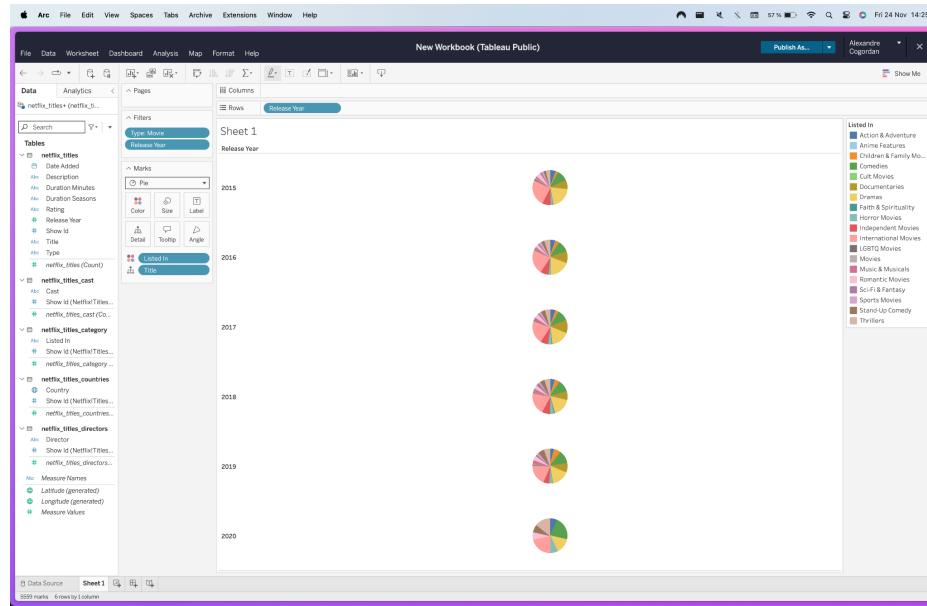
We've used both sources and their results to check our analysis. Judging from our results, February is the month in which new content should be broadcasted (as it is the month with the lowest content).

Question 4 - Do a movie audience analysis

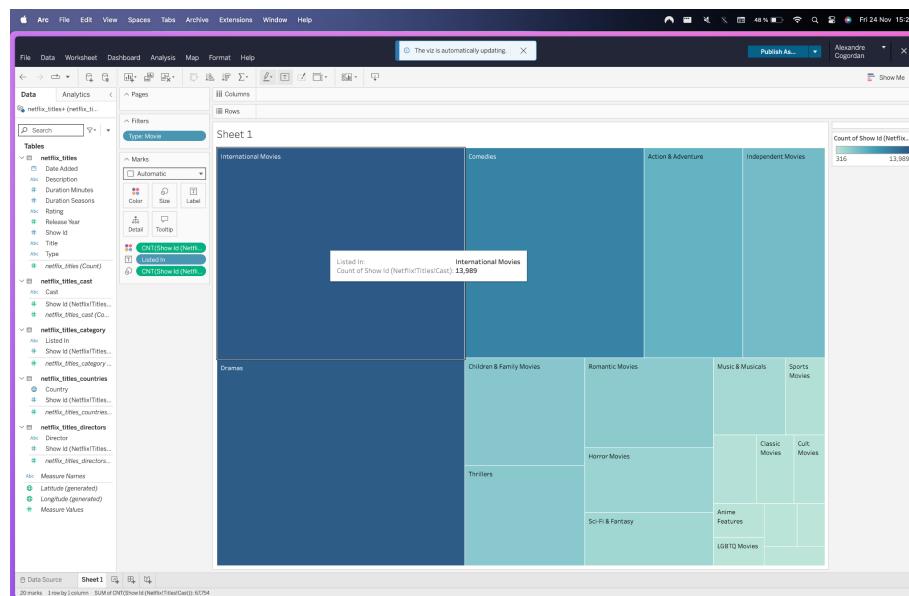
We've started our movie analysis by looking at the repartition of genres each year for movies. This allowed us to see if trends in genres changed as time passed. For example, we've come to understand that the number of international movies rose each year. However, we believe that this isn't mostly explained by trends but rather by the increased availability for Netflix to have more foreign movie licences.



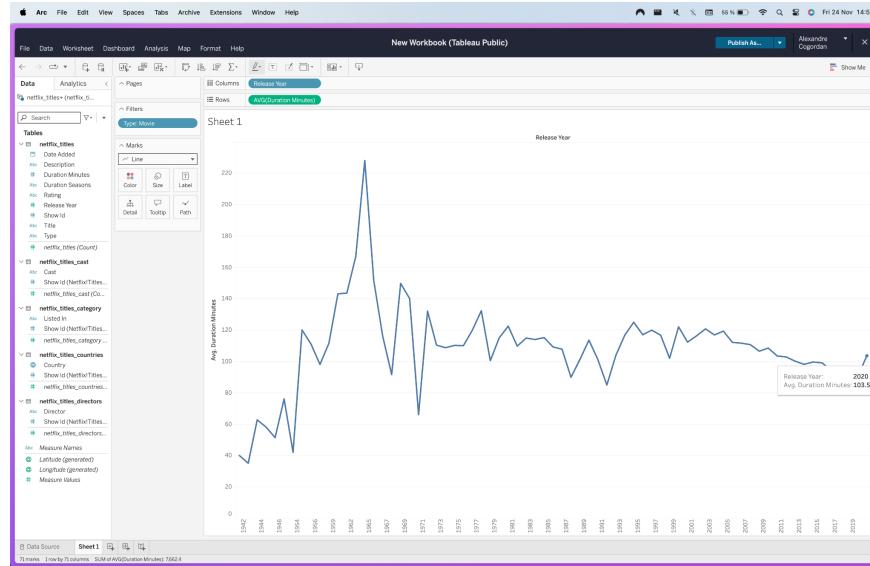
Movie repartition by genre per year - [Demonstration: https://youtu.be/3k5eifYzUoQ](https://youtu.be/3k5eifYzUoQ)



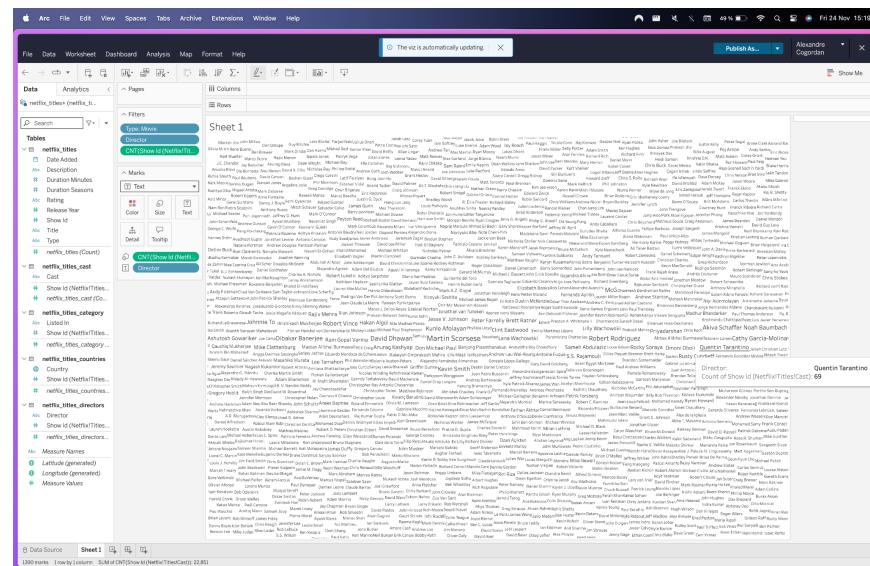
Whilst the previous comment was focused on the earlier years of Netflix, it seems that the number of international movies has reached its maturity state, but this allows us to see that most popular genres had around the same ratio each year.



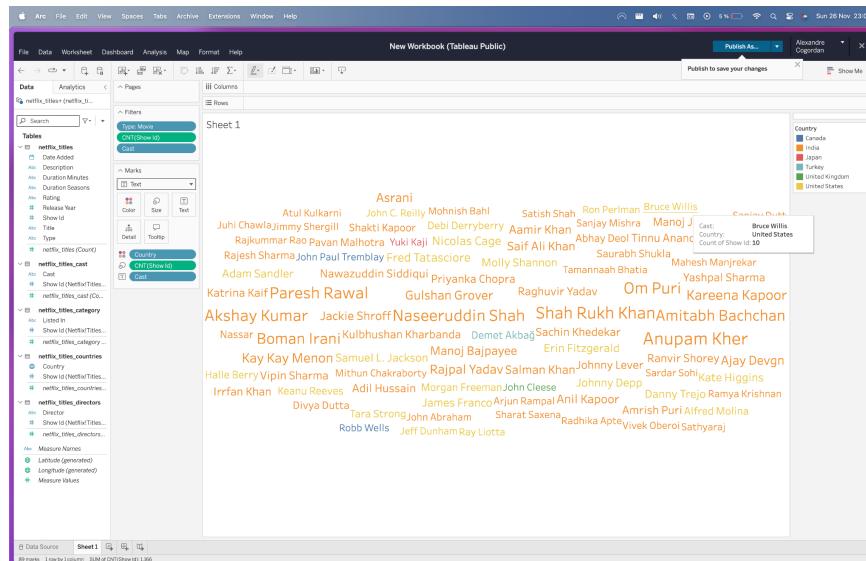
This is another representation of that idea, except that it isn't based on years. In that matter, regardless of the year, the number of international TV shows dominates most genres. However, if we didn't account where the movie was created, the most popular genre would be TV dramas.



Unlike what we'd believe, we discovered that the duration of movies had a peak in the 1960 with peaks of 225 minutes. It seems that the durations are now more reasonable with an average duration of 103.5 minutes in 2020.

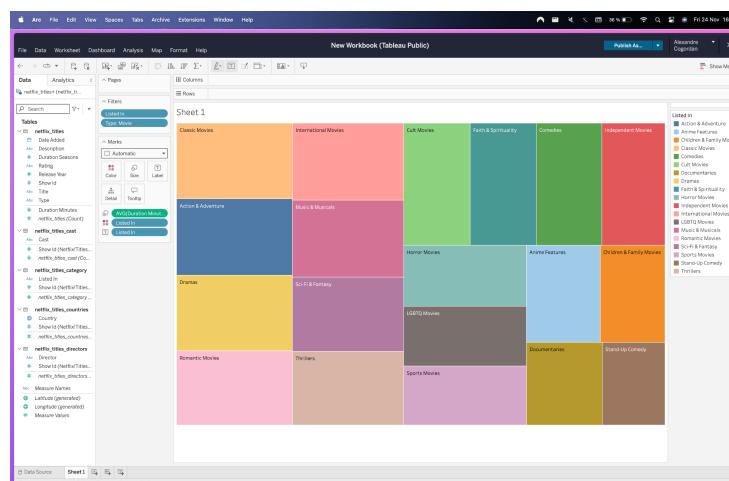


We've analysed the number of shows per directors for movies with the count as the size variable. This way, the directors with the most movies have their texts made bigger to outline themselves. In this case, we decided to keep american directors to avoid the analysis taking too much time but also to find out directors that are more known in the occident. Otherwise, it seemed that Indian directors dominated the counts of movies produced. This may be explained by the bollywood industry being so different to bollywood in terms of movies produced and how they're made.



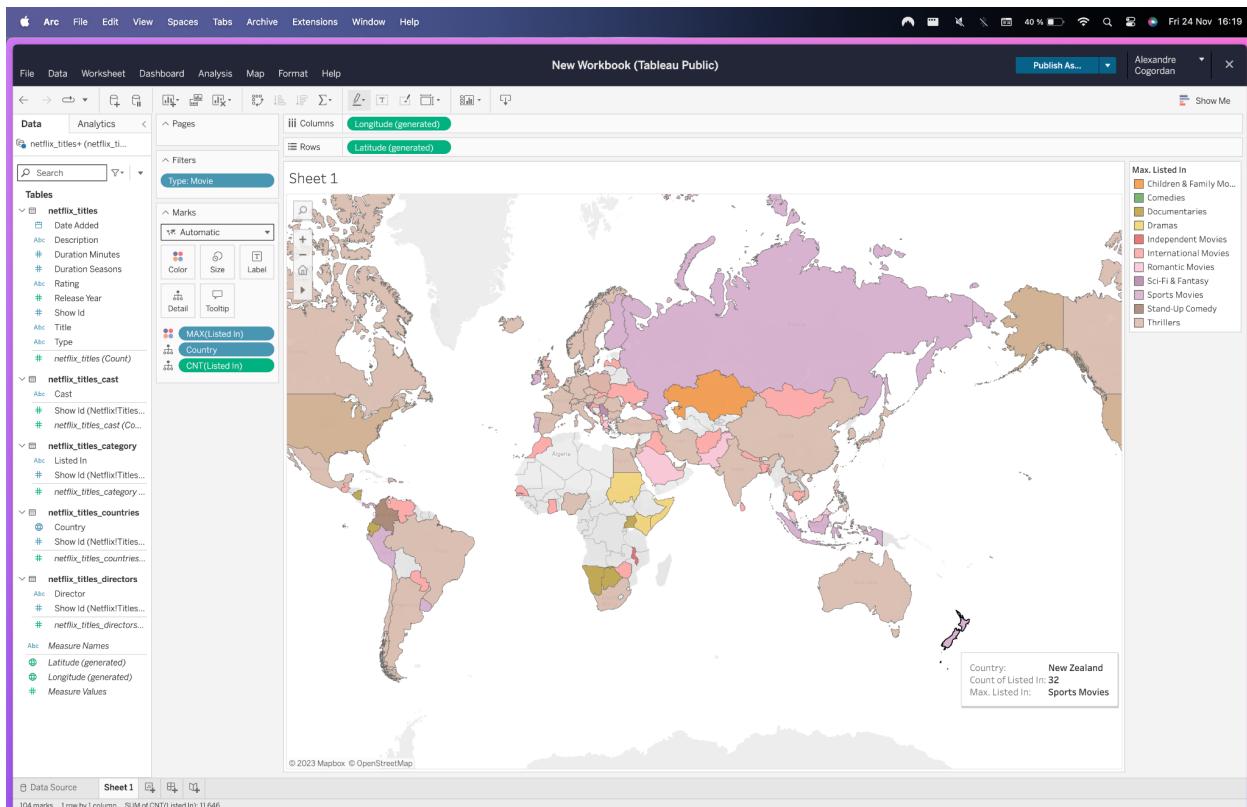
Number of movies per cast

The same thing was done for the cast. Yet again, Indian actors dominated the scene, but it appears that some American actors were favoured by the platform, hence probably beloved by netflix users. This is the case for Bruce Willis where he made 10 appearances on the platform. The colour lets us see the nationality of the cast to get a bigger overview of the type of cast we're dealing with.



Duration of movies by type

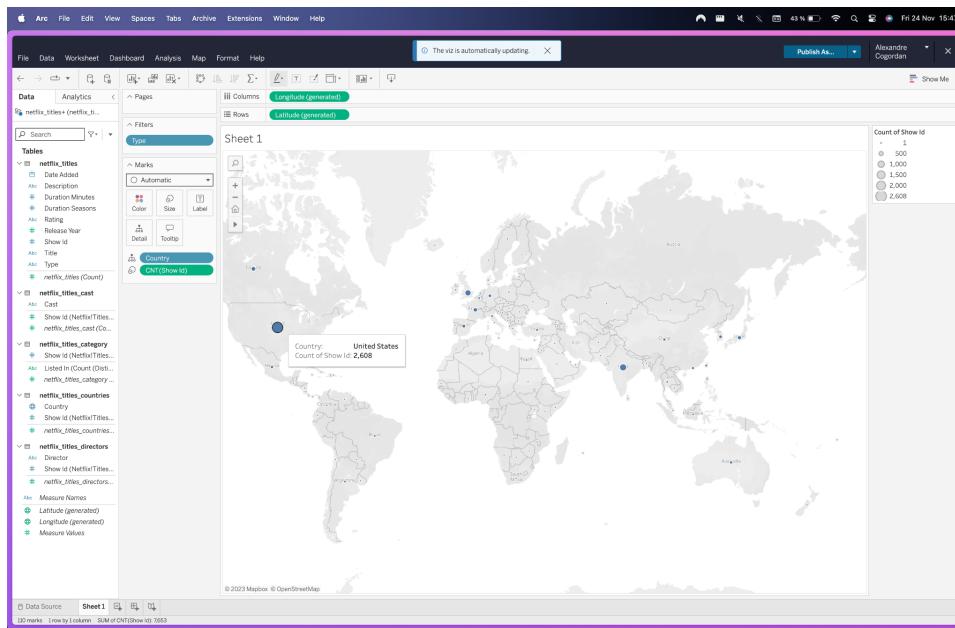
We continue with the duration of movies by type for movies. Surprisingly, the top genres had around the same durations which meant very little variation in duration.



Top movie categories by country

We finalise our analysis of the top movie categories by country. With the colours indicating the movie categories, we can visualise the trends by each country. We understood that Netflix platforms in the east had more sports related movies than the occident ones for example. What would this be caused by? Trends from the population, or rather the country's want to make its citizens healthier?

Question 5 - Other analysis



Amongst the other analysis, we managed to find out that most Netflix platforms had American movies. This is probably affected by the Hollywood influence on most countries. Finally, we've also carried out the same analysis made with the movies on the TV shows to get more insight into them also. Find below the mentioned analysis.

