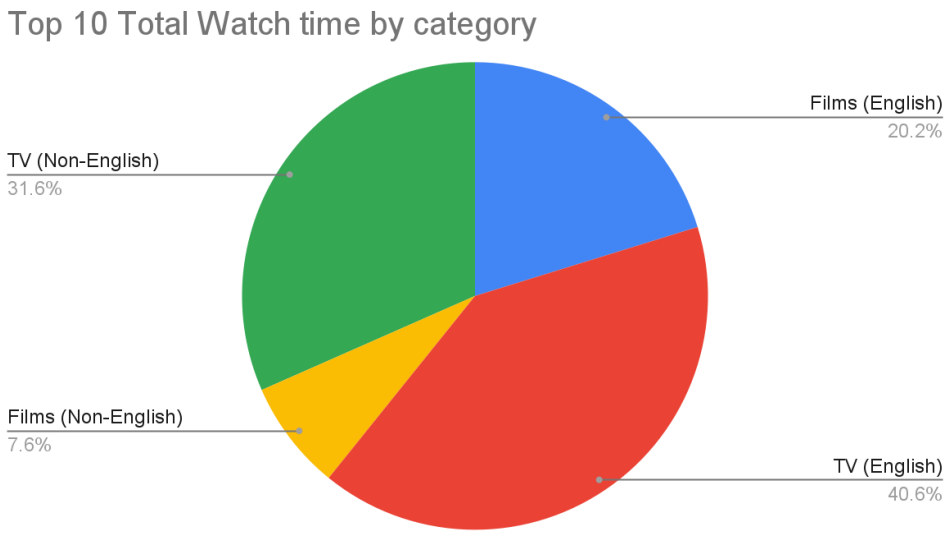
INTRODUCTION

From the big screen to your living room, the age of streaming has made the consumption of media easier than ever. As media continues to evolve, we have begun to see films ditch the big screen and opt for streaming platform releases. With tech giants like Netflix reporting 221.64 million paid subscribers worldwide (Statista), it is easy to see how streaming releases would still reach their target audiences. As with any successful company, knowing your audience is essential to success, but what does the demographic of Netflix, a huge international streaming service, really look like?

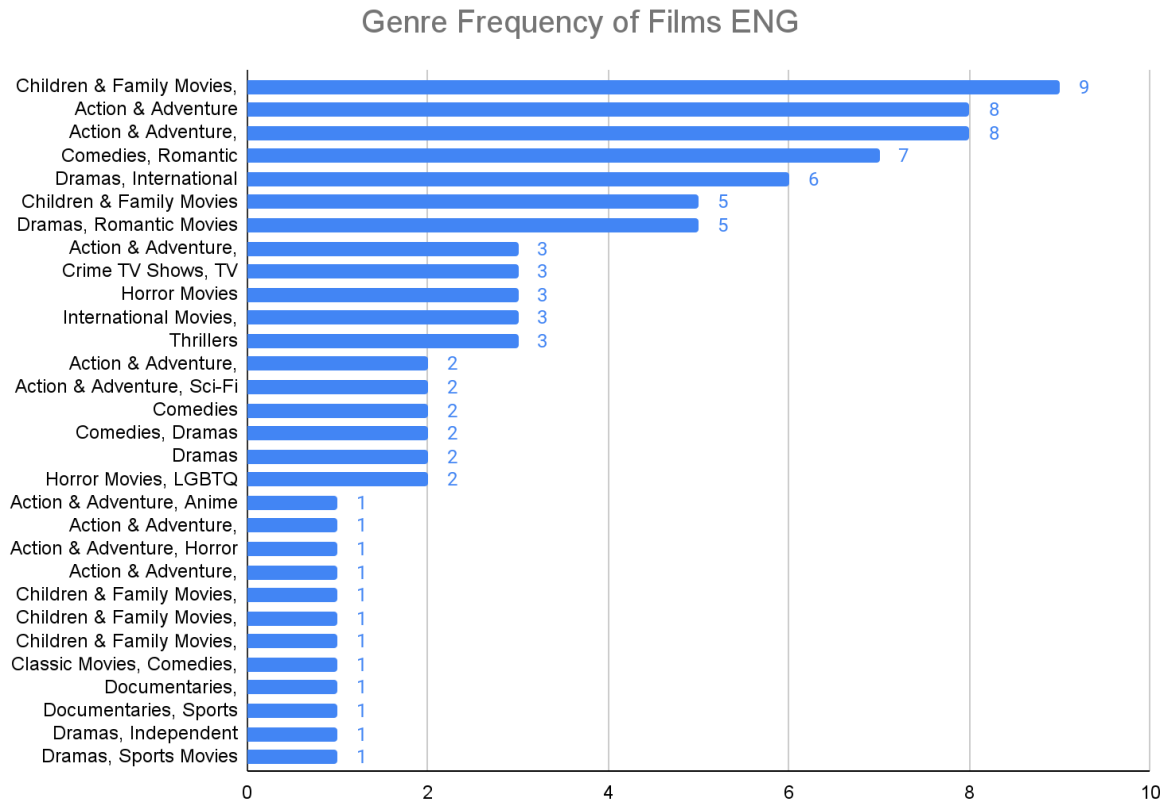
Well for starters, the United States and Canada make up roughly ⅓ of netflix's total audience with 74.58 million paid subscribers (Statista). With that being the case, we would assume that English media would compose roughly half of all consumption, especially, once we take into account, English speaking countries outside of North America. Therefore, if English media makes up roughly half of consumption, we would expect that it is consumed at the same rate as non-English media. However if we look at the data, we can see our initial hypothesis does not hold. English media watch time greatly outweighs its counterpart.

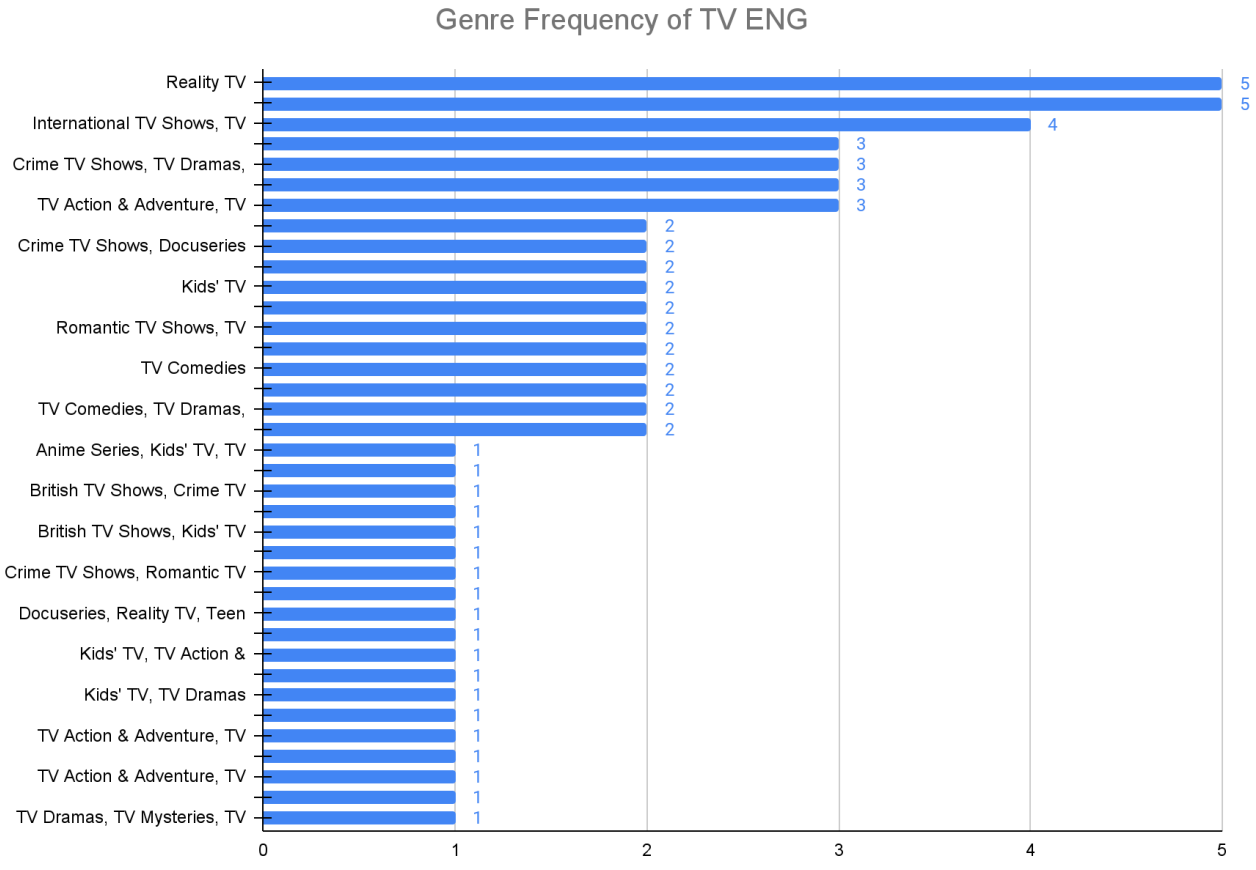


By looking at the Top 10 of Netflix as a metric for what is being watched most, we can see media popularity over the course of a 42 week period. As this pie chart demonstrates 60.8% of all popular media consumed over the 42 week period was English media, but why is this the case? And would attempting to grow the popularity of non-English media prove fruitful or just a waste of time and money?

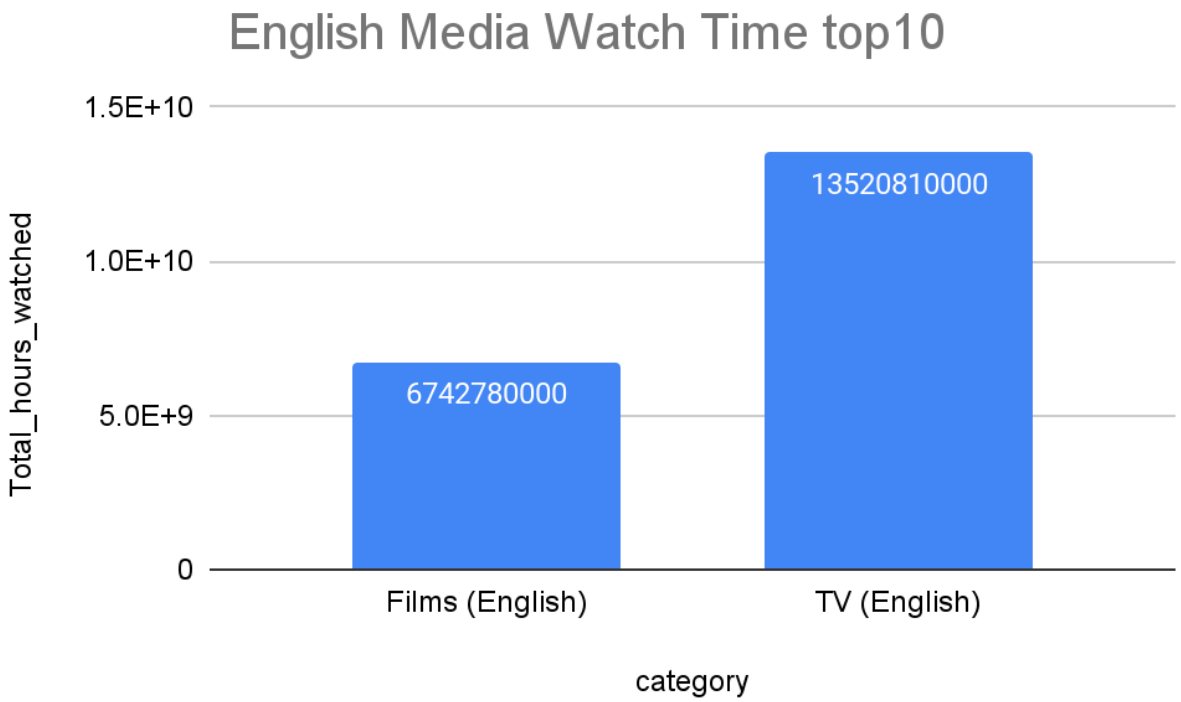
ANALYSIS

To determine if non-English media can become as popular as English media, we should first look at some of the attributes of popular English media. Once again utilizing our top 10 data we can gain an understanding of which genres of film and tv are in demand.


Genre Frequency of Films ENG
Children & Family Movies, 9
Action & Adventure 8
Action & Adventure, 8
Comedies, Romantic 7
Dramas, International 6
Children & Family Movies 5
Dramas, Romantic Movies 5
Action & Adventure, 3
Crime TV Shows, TV 3
Horror Movies 3
International Movies, 3
Thrillers 3
Action & Adventure, 2
Action & Adventure, Sci-Fi 2
Comedies 2
Comedies, Dramas 2
Dramas 2
Horror Movies, LGBTQ 2
Action & Adventure, Anime 1
Action & Adventure, 1
Action & Adventure, Horror 1
Action & Adventure, 1
Children & Family Movies, 1
Children & Family Movies, 1
Children & Family Movies, 1
Classic Movies, Comedies, 1
Documentaries, 1
Documentaries, Sports 1
Dramas, Independent 1
Dramas, Sports Movies 1
0
2
4
6
8
10

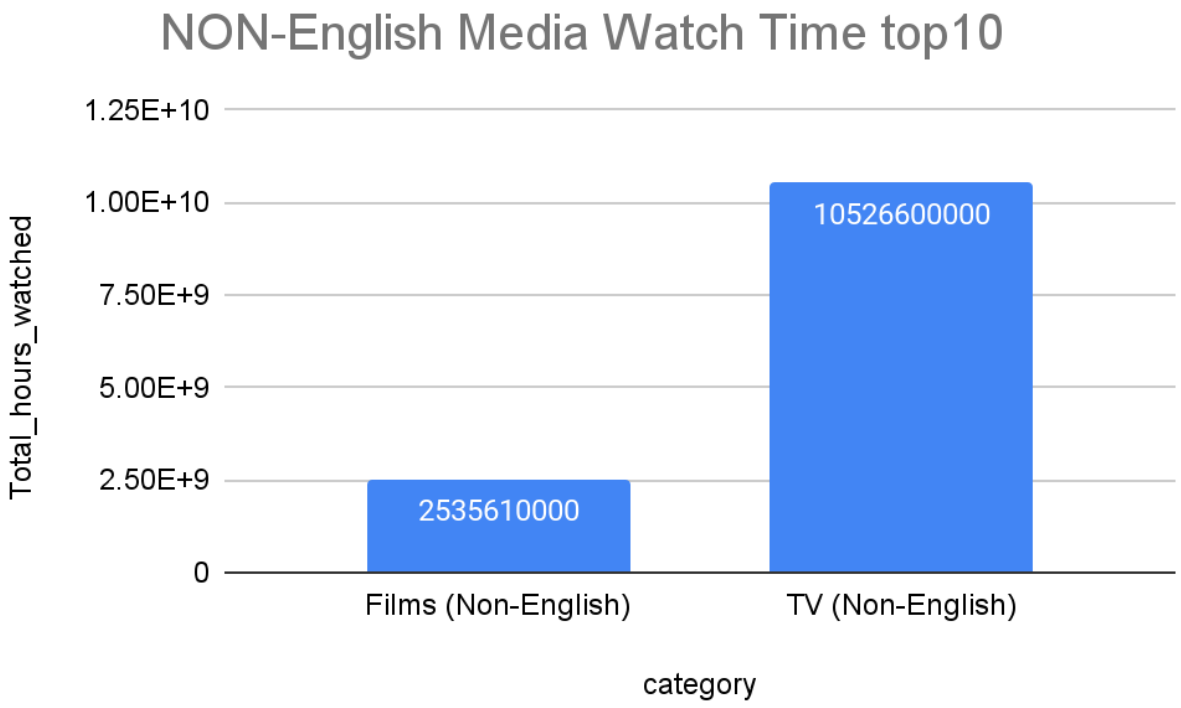

Genre Frequency of TV ENG
Reality TV 5
5
International TV Shows, TV 4
3
Crime TV Shows, TV Dramas, 3
3
TV Action & Adventure, TV 3
2
Crime TV Shows, Docuseries 2
2
Kids' TV 2
2
Romantic TV Shows, TV 2
2
TV Comedies 2
2
TV Comedies, TV Dramas, 2
2
Anime Series, Kids' TV, TV 1
1
British TV Shows, Crime TV 1
1
British TV Shows, Kids' TV 1
1
Crime TV Shows, Romantic TV 1
1
Docuseries, Reality TV, Teen 1
1
Kids' TV, TV Action & 1
1
Kids' TV, TV Dramas 1
1
TV Action & Adventure, TV 1
1
TV Action & Adventure, TV 1
1
TV Dramas, TV Mysteries, TV 1
0
1
2
3
4
5

As the data shows, for both tv and film, we can see certain genres stand out among the rest. When it comes to film, we can see categories having multiple tags, with 9 different genres including Action and Adventure. Similarly, we can see the popularity of tags including Comedy, Drama, Crime and Thriller. Once again with TV, we can see that similar tags are more frequent among the top 10 of English media consumed. From both of these tables, we can deduce that the tags that often produce the most popular media are that of Crime, Thriller, Comedy, Action & Adventure, and Drama. With film having a stand out in Children & Family Movies, and tv having a stand out in Reality TV. Even so, how does watch time reflect these interests? Below, we take a look at the raw amount of watch time for English Film vs English TV.

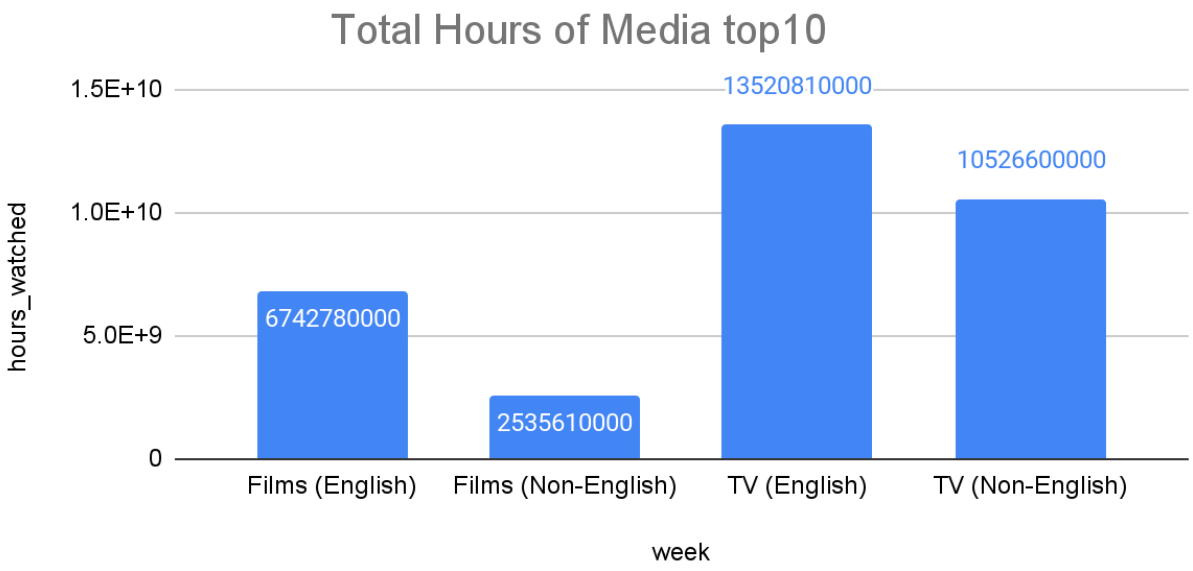


We can observe a staggering difference in total hours watched for Film and TV. Taking into account the amount of content a season can have as opposed to a movie makes sense. So the next question we should be asking is how many unique shows and movies are coming into the top 10? Well our data shows that 199 unique film titles appear in the top 10 over our 42 week period and only 106 unique show titles appear in our top 10 (see Appendix A), with 25 of those shows having more than one season (see Appendix B). Therefore English media consumption can best be categorized as consuming more tv than film, enjoying genres with action, drama, crime, and thriller, and preferring new titles as opposed to additional seasons of old titles. So how different could non-English media be?

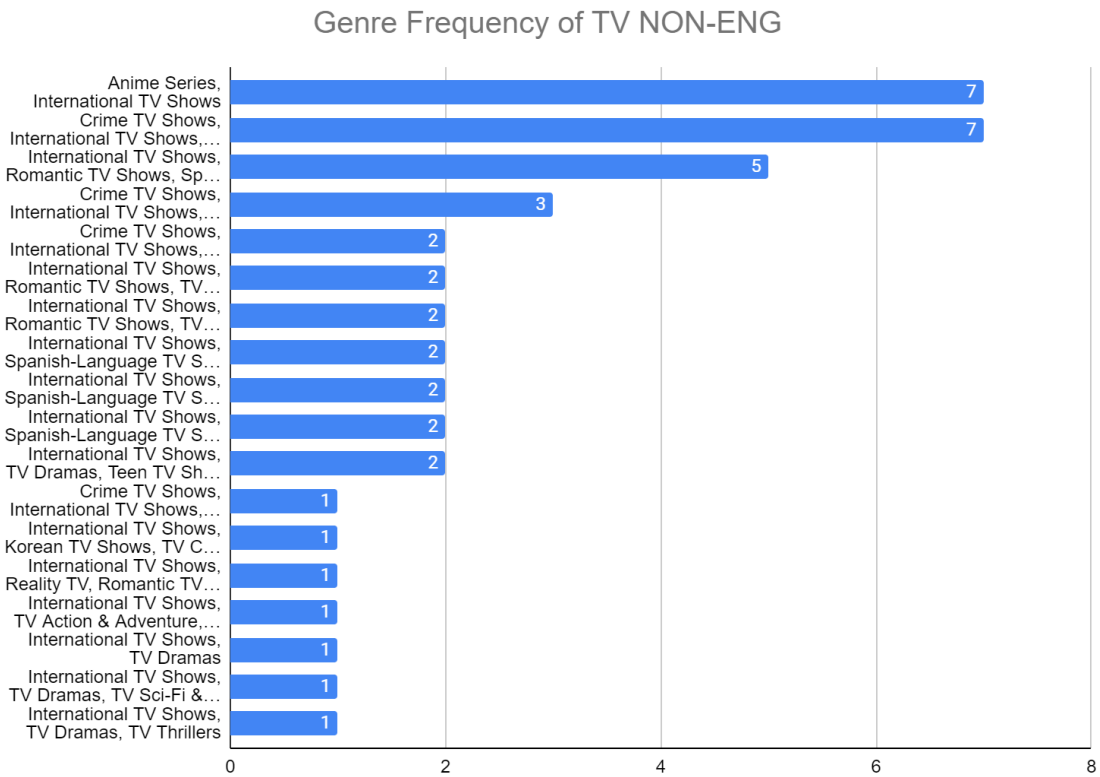
When looking at non-English media, we will immediately see the difference in watch time.

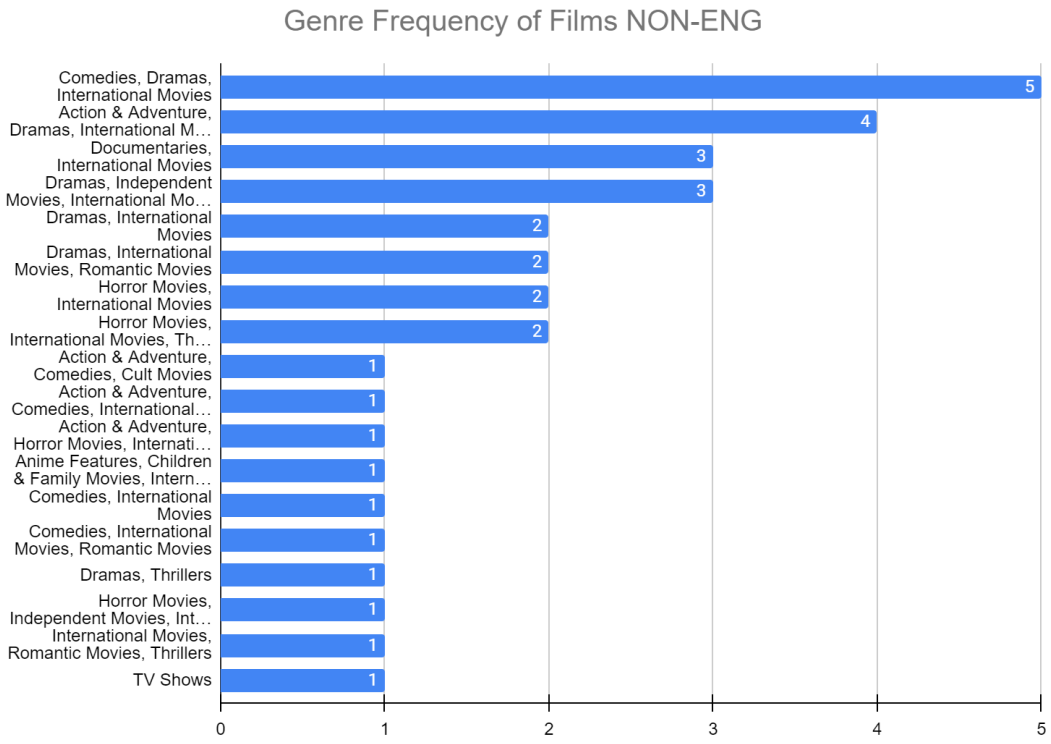


We can see the gap in watch time between films and tv is even larger here than it was in English media consumption for the top 10. It is even more apparent when we have them side by side.



From this data we can evidently see the discrepancy between English and non-English media as English media has a higher watch time in both categories. However, that does not mean we should not note the difference within non-English media. Here we can see a huge gap between non-English Film and TV categories. We can observe that the difference within English media is by a larger margin. The marginal difference for English media is 6,778,030,000 hours and 7,990,990,000 hours for non-English media. For both English and non-English media, both lean towards greater TV watch time. Then there must be something different about the genres of content for non-English media right? Well as our data will show,





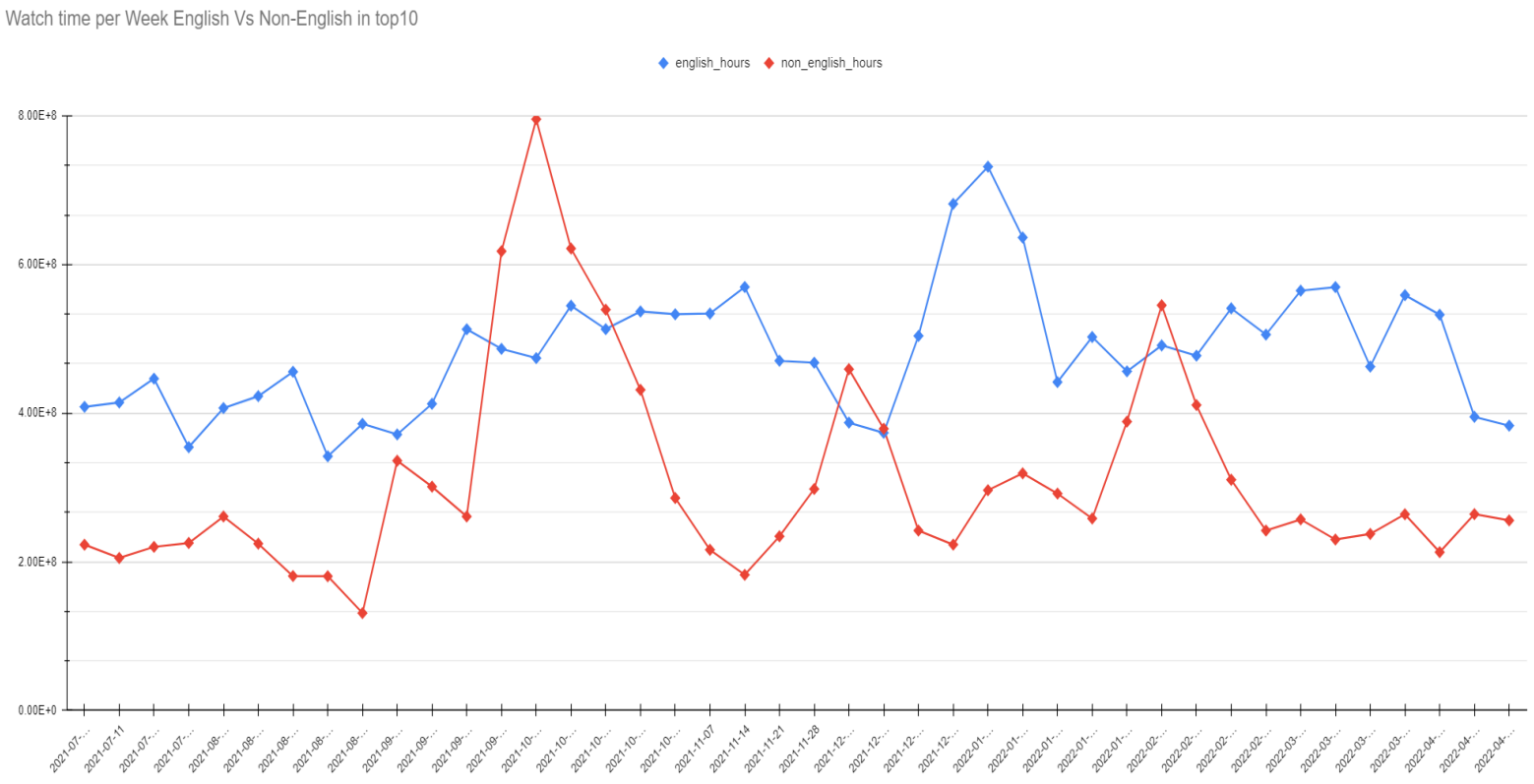
Although similar in categories preferred, the pecking order is slightly different with Dramas being one of the most frequent tags. Even so, we still see the tags of Comedy, and Action & Adventure being prominent throughout the genres, reflective of its English counterpart that also saw traction in these categories. Furthermore, looking at its more popular half, non-English TV sees the standout popularity of genres like Anime. Even so, it still frequents Crime, Romance, Action and Adventure, and Drama as English TV does. As a result, if we were to classify non-English media, we could say that similar to English media, they prefer TV to film and enjoy very similar genres of media. Although they may be in a slightly different order, the same building blocks are there. Then why is it that English media is consumed more? Is there just more of it? Or is it just being watched at a much higher rate. Is non-English media incapable of surpassing English media?

If the problem is a lack of content for non-English media, then we should look at the difference in abundance between English and non-English media. By looking at our netflix_shows table as a reference for all media on Netflix, we see that there are roughly 8807 different shows, 4531 of those are of English speaking origin, and 4276 are of non-English origin. By those statistics, English media makes up 51.4% of all of the media available to us. If English media is roughly half, what of the categories that are most frequent on Netflix as a whole? With over 514 unique tag combinations (see Appendix C), the top 5 most prominent tag combinations being the following.

listed_in	Frequency
Dramas, International Movies	362
Documentaries	359
Stand-Up Comedy	334
Comedies, Dramas, International Movies	274
Dramas, Independent Movies, International Movies	252

With categories like Action & Adventure not appearing until the 15th most common by frequency, and categories like Thriller and Crime appearing even further down on the list. As a result we can see that non-English media isn't facing a content drought. In fact, it very much has substantial content. So are viewers just not watching as much non-English media as English media consumers?

In order to best judge how much media is being consumed, we should look at it week by week to see the differences in data.



As we can see from our line chart, non-English media is being consumed, on average, at a much lower rate than English media. However, this should not take away from the fact that in seven different weeks, non-English media has been able to match, if not, supersede the amount

of popular English media consumed. Suggesting that non-English media has the potential to rise in viewership. This possibility is further suggested by the fact that on average, the difference week by week in non-English media is 6090238.095. This implies that it is on the rise (see appendix D). This is even further supported by the fact that if we look at how long it takes for a show or film to trend after it has been added to Netflix, we can see that the median for non-English TV and film is less than that of English film and TV (see appendix E). We can argue that there exists an audience and desire for non-English content, making it a prime market to explore.

CONCLUSION

In conclusion, despite there being a clear difference in the amount of popular media consumed in English and non-English categories. We can see that there exists content that non-English consumers would enjoy and that they are actively watching said content. In order to further test the validity of our results, I would recommend we look into testing the weeks where non-English media outwatched English media via a Z - test. I would further recommend a logistical regression measuring how likely a viewer is to rewatch a show or movie. This is an important point because after a film or show has reached all audiences, in order for it to remain popular, it would have to maintain a large enough viewership to keep it in the top 10. By completing this analysis we can determine what factors lead to shows or movies being rewatched. I believe non-English media is a growing market that will prove to be a fruitful market to invest in as shown by the data.

References

Stoll, J. (2022, April 22). *Netflix: Number of subscribers worldwide 2022*. Statista. Retrieved June 23, 2022, from https://www.statista.com/statistics/250934/quarterly-number-of-netflix-streaming-subscribers-worldwide/

Stoll, J. (2022, May 11). *Netflix paid subscriber count by region 2022*. Statista. Retrieved June 23, 2022, from https://www.statista.com/statistics/483112/netflix-subscribers/

Appendix A

```
WITH top10_counts (title, category, times_appeared) AS (
SELECT DISTINCT show_title,
category,
count(show_title) as times_appeared
FROM netflix_top10
GROUP BY show_title
)
SELECT category,
count(category) as times_appeared
FROM top10_counts
GROUP BY category
```

Via this query I was able to get the following table

Category	Times_appeared
Films (English)	199
Films (Non-English)	170
TV (English)	106
TV (Non-English)	97

Appendix B

```
WITH season_count(show_title, category, seasons) AS (
SELECT show_title,
category,
COUNT(DISTINCT season_title)
FROM netflix_top10
WHERE category like '%TV%'
GROUP BY show_title
)

SELECT category,
COUNT(show_title)
FROM season_count

where seasons > 1
GROUP BY category
```

Via the following query I was able to get the following table

Category	shows with multiple seasons
TV (English)	25
TV (Non-English)	7

Appendix C

```
SELECT listed_in,
count(listed_in) AS frequency
FROM netflix_shows
GROUP BY listed_in
ORDER BY frequency DESC
```

Via this query I was able to get a list of all of the unique tag combinations and then by selecting the number of entries I was able to determine the number of unique tag combinations

Appendix D

```
WITH english_media(week, weekly_hours_viewed) AS (
SELECT week,
total(weekly_hours_viewed)
FROM netflix_top10
WHERE category LIKE '%(English)%'
GROUP BY week
),
non_english_media(week, weekly_hours_viewed) AS (
SELECT week,
total(weekly_hours_viewed)
FROM netflix_top10
WHERE category LIKE '%Non-English%'
GROUP BY week
)


SELECT english_media.week,
english_media.weekly_hours_viewed + non_english_media.weekly_hours_viewed
AS total_weekly_media_hours,
english_media.weekly_hours_viewed
AS english_media_hours,
english_media.weekly_hours_viewed - LAG(english_media.weekly_hours_viewed, 1, 0) OVER
(
ORDER BY english_media.week
) AS english_media_weekly_dif,

non_english_media.weekly_hours_viewed
AS non_english_media_hours,
non_english_media.weekly_hours_viewed - LAG(non_english_media.weekly_hours_viewed, 1,
0) OVER (
ORDER BY english_media.week
) AS non_english_media_weekly_dif
FROM english_media
INNER JOIN non_english_media ON english_media.week = non_english_media.week
```

Via this query I was able to get data on week to week consumption, further statistics can be found by applying aggregate functions to the selected columns. This is how the average, and average difference from week to week

Appendix E

```
WITH english_media(week, weekly_hours_viewed) AS (
SELECT week,
total(weekly_hours_viewed)
FROM netflix_top10
WHERE category LIKE '%(English)%'
GROUP BY week
),
non_english_media(week, weekly_hours_viewed) AS (
SELECT week,
total(weekly_hours_viewed)
FROM netflix_top10
WHERE category LIKE '%Non-English%'
GROUP BY week
)
SELECT english_media.week,
english_media.weekly_hours_viewed + non_english_media.weekly_hours_viewed
AS total_weekly_media_hours,
english_media.weekly_hours_viewed
AS english_media_hours,
english_media.weekly_hours_viewed - LAG(english_media.weekly_hours_viewed, 1, 0) OVER
(
ORDER BY english_media.week
) AS english_media_weekly_dif,

non_english_media.weekly_hours_viewed
AS non_english_media_hours,
non_english_media.weekly_hours_viewed - LAG(non_english_media.weekly_hours_viewed, 1,
0) OVER (
ORDER BY english_media.week
) AS non_english_media_weekly_dif
FROM english_media
INNER JOIN non_english_media ON english_media.week = non_english_media.week
```

Via this query I was able to find the week a show was added to netflix and when it first ended up on the top10. By finding the difference in days between those dates and taking the median, this was done as there were outliers that greatly influenced the average making the median a more accurate measure of time to trend.