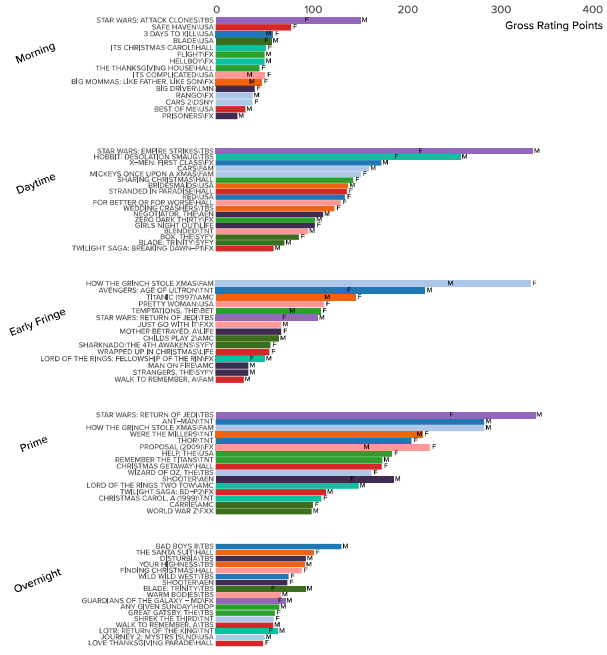
NETWORK TV MOVIE TRENDS

From Q4 2016-Q4 2017, women in the 18-49 audience watched more movies than men in most genres and across all dayparts. Examining viewership by network & daypart reveals that a select few networks dominate GRPs; genre data offers insight into the content that draws viewers to popular networks.

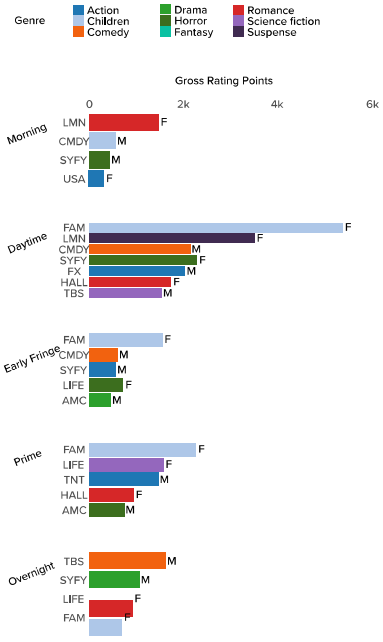
MOST POPULAR FILMS: MIDDLE AGE ADULTS

What is the most popular film on television by genre and daypart for 35-44 year old male and female audiences?



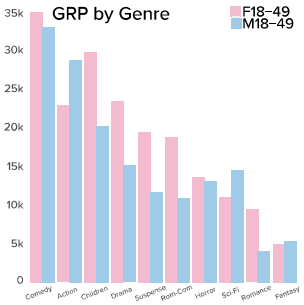
HIGHEST RATED NETWORKS

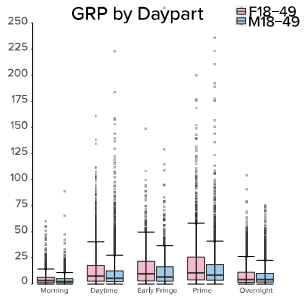
What are the most popular networks by genre, daypart, and sex?

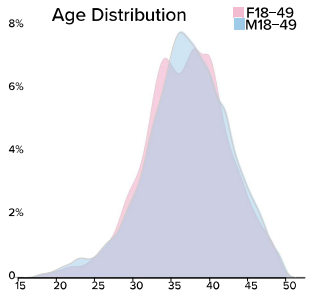


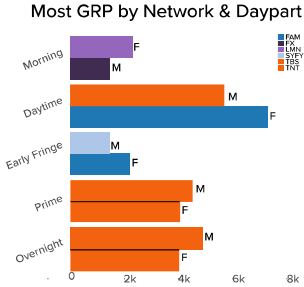
MOTIVATION

Network groups can sell airtime to advertisers at a premium when both parties understand, with high granularity, who is watching what when. Although viewership of linear programming is lower in the digital era, television advertising remains a surefire way to connect with large audiences with shared interests. Hypertargeting extends beyond marketing; content can be personally tailored when network groups understand their viewers.









Sam Deery-Schmitt
IST 719, Information Visualization
MS Applied Data Science Candidate
Syracuse University

SOURCES

Data owner: RSG Media (url available in script)
R packages: cowplot, dplyr, ggplot2, gridExtra, grid, rgl, scatterplot3d

DATA

The initial dataset featured 114,373 observations of 17 variables; each record details data about the performance of one movie on one network. This dataset features the most popular movies during the time period. Aggregations were done by age, sex, daypart, genre, network, and network group.

Project Summary

Nielsen Data Analysis for RSG Media

github repo

In a real-world application that intersected with academic coursework, I analyzed Nielsen data for RSG Media as part of an interview process, later submitting this analysis for my Information Visualization, IST 719. I received and accepted an internship offer.

Project Overview

Objective

To analyze and report on Nielsen data provided by RSG Media, focusing on viewer demographics and behaviors.

Challenge

The dataset is enormous, with hundreds of thousands of records and dozens of variables, and the domain was unfamiliar.

Data Source

Nielsen dataset featuring detailed viewership data.

Submission

The analysis was submitted in the form of a written report with visualizations to RSG Media, and a poster along with an R script to my course professor.

Technical Process

Data Analysis

Employed R for in-depth data analysis, focusing on viewership patterns, genre preferences, and network performance.

Visualization

Created visualizations to highlight key findings, including audience preferences and viewership trends.

Report and Poster Creation

Compiled findings into a comprehensive report and an engaging poster, showcasing both analytical depth and visual communication skills.

Challenges and Problem-Solving

Navigating the complexity of Nielsen data to extract meaningful insights.
Translating complex data findings into visually compelling and easy-to-understand formats.

Impact and Results

The project's success led to an internship offer from RSG Media, affirming the practical application and relevance of the skills developed.
Demonstrated proficiency in data analysis and visualization, essential for a career in data science.

Reflection on Professional and Academic Integration

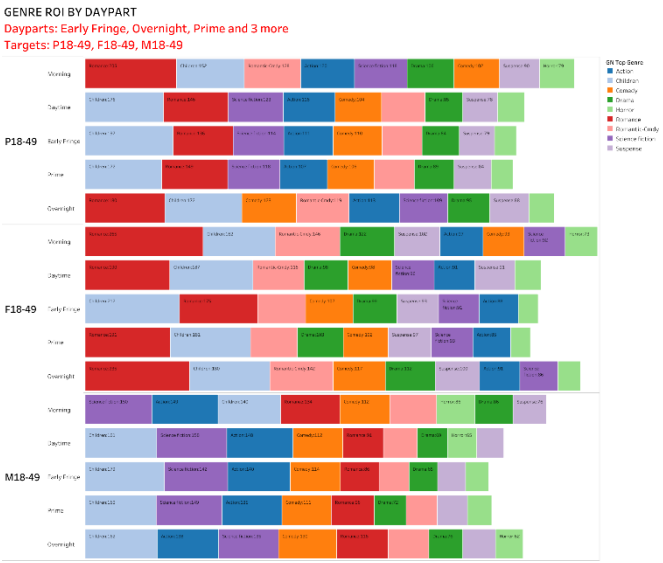
This project demonstrated direct application of academic learning in a professional context. I consulted frequently with RSG to better understand the domain, and took Nielsen’s educational courses to learn about the ratings.
It highlighted the importance of clear communication in data science, especially when translating technical insights for diverse audiences.
This project also helped me receive an internship offer from RSG media, which I accepted.

Full project analysis starts on next page.

I understand that this population represents people ages 18-49 in the United States who own televisions and watch any of the network group channels included in the dataset. After looking at data across all quarters, I chose to look at Q4 2016 – Q4 2017 to focus on the most recent data. As I explored the data, I recognized a significant difference in viewing behavior among women and men.

Looking at "Average Movie Audience and Share by Daypart," it is clear that the average audience (numerical value) for women is higher than that for men across all dayparts, with the largest differences occurring in the Early Fringe and Prime dayparts. However, as "Top Genre by GRP's and Male/Female Skew" demonstrates, the higher average audience for women across all dayparts does not neatly translate to higher GRPs when genre is taken into account. Looking at GRP by sex, there is a nearly even split in comedy and horror across all dayparts, and men consistently contribute to more viewing of the action and science fiction genres than women. However, women account for significantly more children, drama, suspense, romance, and romantic comedy viewing, accounting for at least 60% of GRPs in each of those categories when looking at all dayparts together, with romance films yielding the largest divide at 71% viewership.

ROI follows a pattern that resembles the trend for male/female skew, with the genres that skew heavily towards women and men yielding high ROIs for each sex across dayparts, as evidenced by "GENRE ROI BY DAYPART." Specifically, romance and romantic comedy yield top-3 ROIs for women in every daypart, while science fiction and action yield top 3 ROIs across dayparts for men. However, there is one notable outlier: children.



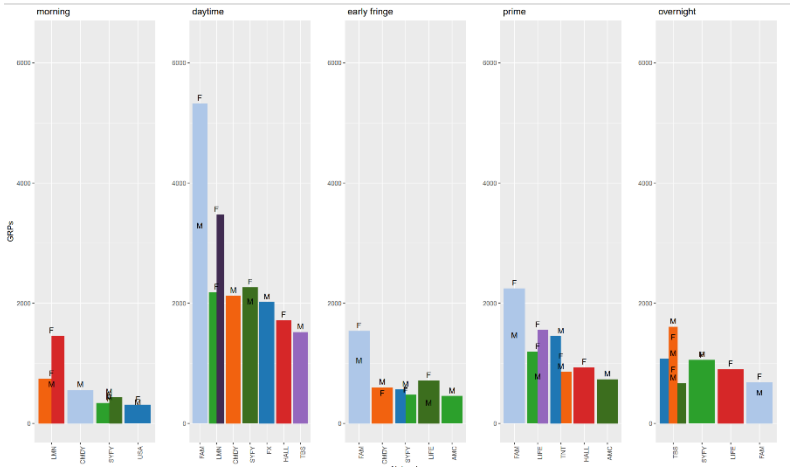
The visualization "GENRE ROI BY DAYPART" reveals that for men, whom on average make up 41% of children viewership, this genre actually has the highest ROI for every daypart except morning. This is especially significant because the children genre ranks no higher than 3rd in overall GRP in any given daypart, yet the ROI is between 140–179 for each daypart for men, and 162–212 for women. Children ranks 2nd in ROI by daypart for

women in every daypart save early fringe, where it overtakes romance, which otherwise dominates by daypart for women in ROI, with a range of 175–265.

An additional oddity this same visualization shows is that comedy, despite having the GRPs of any genre, has a lower ROI compared to other top genres in each daypart, ranking anywhere from 4th – 7th for women, and 4th – 5th for men. Comedy does appear to yield higher ROI for men by daypart, with a range of 111–130, while the ROI for women in the same genre is 93–117.

The median income of viewers for the top genres is higher for men across most genres (looking at Median Income By Female Skew, not included in document), with slight variation over dayparts, and above-average median incomes for both men and women in both romantic comedy and romance; men additionally surpass the average in science-fiction, comedy, drama and children. The highest median income for both sexes is in the romantic comedy genre, with just under $32,000 for women and near $33,000 for men.

This high-level overview of the data helped me discern what dimensions I should focus on as I drilled down, and additionally inspired me to look at a few additional dimensions, including network, network group, and movie / network. I downloaded the "Top Movie Ranker" data as a csv, capturing Q4 2016 – Q4 2017 for all dayparts, with 114,373 observations of 17 variables. When I looked at networks with the highest GRPs by daypart and sex, and compared this distribution to that of GRP by network group and sex over daypart, I saw an interesting outcome. Turner owns 5 of the 10 networks with the highest GRP for a sex within a daypart, with TBS ranking as the most popular (by this measure) channel for both men and women overnight, as well as the most popular channel for men during daytime; TNT rules primetime, accruing the most GRPs for both men and women. However, single network dominance does not translate to network group dominance; Viacom has the most GRPs for any one network group for both women and men during daytime, and edges out Turner for men overnight and during primetime. Disney's FAM network has the single highest GRP of any network for any daypart (more than 7000 during daytime for women), yet it is Viacom that has the highest GRPs for any given daypart during that same timeslot.

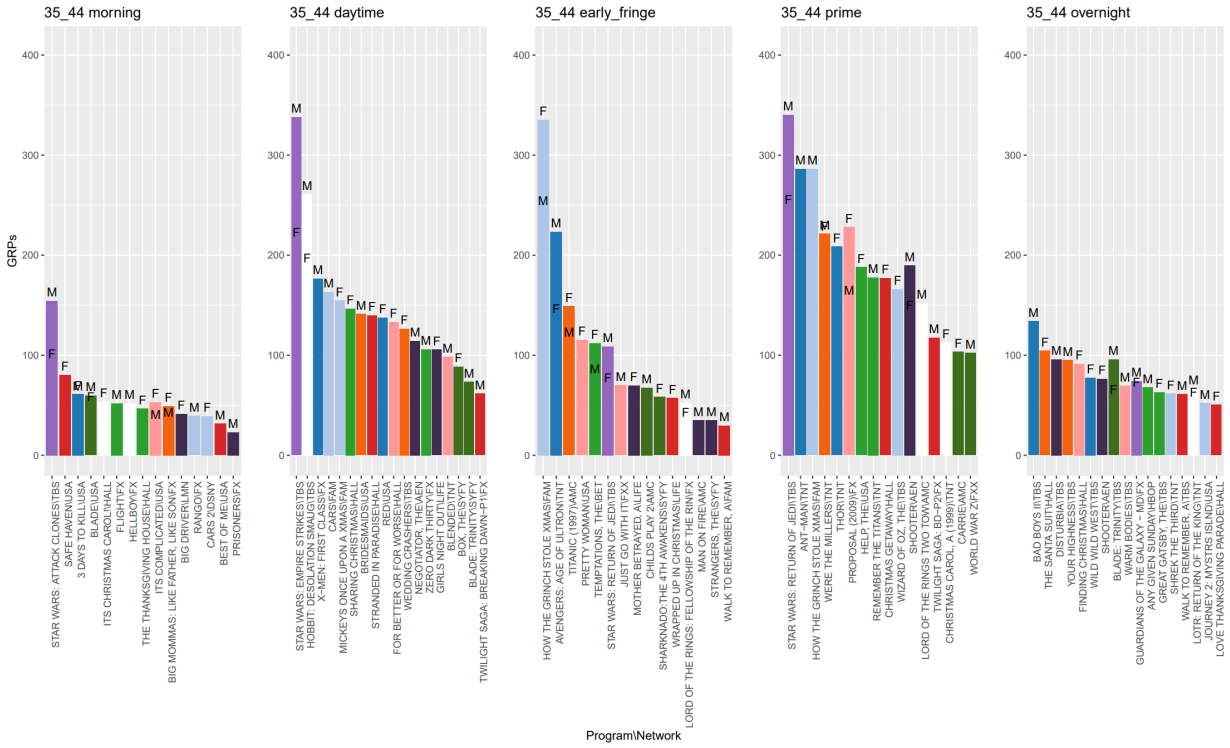


Going more granular, when we add genre to this aggregation, we see a more diverse representation of networks, as well as understand which genres give certain networks and network groups the advantage in different dayparts. For example, in the overnight daypart, TBS commands action for men, as well as comedy and horror for men and women. SYFY, however, has the highest GRP for both men and women in the overnight drama category, and additionally claims the highest GRPs for men and women in both drama and horror in the morning, horror for both sexes in the daytime, and action for men and women in early fringe, as well as drama for women. I highlighted SYFY in particular because this is an example of a niche network that has managed to carve out a space in multiple genres across multiple

dayparts. There are similar, interesting trends at the network group level, but at this point I want to wrap up this analysis with a focus on the most popular films.

I made several interactive 3D scatterplots to look at the most popular films, as measured by GRPs and average audience (numeric), adding a third dimension age, then daypart, and coloring by sex. This helped me see the overall distribution of movie / network pairs by these different dimensions. The most popular movies, as measured by GRPs and average audience (numeric), include many of the Star Wars films (in multiple dayparts), but a Bugs Life on FAM is just as popular. Interestingly, Star Wars was most popular on TBS. I added 2D planes to look at different percentiles and discern the truly best performing movies. Frankly, these 3D plots are best viewed interactively, so I did not include them, they can be recreated and examined using my script.

I got increasingly granular in this analysis. I divided up the 18-49 target by sex and age according to buckets that seemed to be fairly standard in multiple industries: 18-24, 25-34, 35-44, and 44-49 (this bucket is normal bigger, but that was the limit with this data). You can see the 35-44 visual below. I looked at the single most popular move for both men and women for each genre for each daypart within these groups. This reveals interesting trends within and among age groups and sexes; for example, Bad Boys II on TBS is the most-watched action movie on any network in the overnight daypart for both M25-34 and 35-34. In the 35-34 category, a different star wars film claims the most GRPs in any one genre on any one network in three different timeslots—morning, daytime, and prime—for both men and women; How the Grinch Stole Christmas is singularly popular in the early fringe daypart. I went even deeper, aggregating films by median age instead of within age groups, and wrote a function to plot the most popular film for that median age for each sex for any one channel for each genre. This level of detail revealed that much of the Bad Boys II viewing activity comes from 36 year old men, who comprise most of the approximately 150 GRPs the movie has for 36-45 year-olds in this timeslot.



The granularity of these aggregations is best explored visually, and can provide an interesting way to discern popularity. For example, understanding that the most popular movie by network and genre for 36-year old women is Big Daddy on Bravo means that the popularity of this specific movie should be further scrutinized, and other Adam Sandler movies should be introduced when this demographic is likely to watch. Or, because Attack of the Clones on TBS is by far the most popular film by network and genre for the morning daypart for 35-44 year-olds, Turner may look to sell more advertising space to Capital One for the credit card advertisements that feature Samuel Jackson, who stars in the film. The science fiction genre does have a fairly high (relative to other genres) median income for men in that daypart.

By being able to understand who is watching what when, we can make better recommendations to network groups about who the likely audiences are at all times of the day and thus provider advertisers with highly specific information that can be used to decrease customer acquisition cost. Additionally, network groups can make strategic decisions about what movies they should license or original content they should create. Ultimately, however, the data I analyzed only represents part of the picture; to make truly insightful recommendations, we need to compare the linear programming data to digital data in order to understand how people are watching content. This approach yields the specificity necessary to make even better business decisions, for example what movies to stream to certain demographics, and what to continue to run on television, and the accompanying advertisements to run. Another level of analysis that would further segment the audience would be to look at the regional level, which could ultimately lead to extremely personalized content and advertisements.

I have included my script. I have not touched up my visualizations in Adobe Illustrator, but I can if necessary. I can also include all the plots my script generates, or explore other dimensions as necessary.