

Case Study by Justin Maxfield

Chevlour: Advanced Hair Coloring for Women

Challenge

AdveriCorp's EVP of Sales has a meeting coming up with "*Chevlour: Advanced Hair Coloring for Women*", a hair coloring brand of a large upscale cosmetics company. He has asked the Sales Engineering team to come up with "something interesting" that will demonstrate to the client that AdveriCorp can:

- (1) Help inform the client's understanding of the *Chevlour* brand consumer.
- (2) Activate media that will allow *Chevlour* to drive sales.

A list of consumers who have recently purchased *Chevlour* has already been processed through AdveriCorp's internal "Audience Insights" tool to produce a comparison of how prevalent those shoppers are across AdveriCorp's audience segments versus AdveriCorp's 257 million records of people in the U.S.

Approach

1. Check data integrity of client's list of consumers.
2. Complete exploratory data analysis to gauge the differences between consumers and AdveriCorp's database.
3. Identify interesting segments by accounting for base rate information.
4. Define cut off values based on client needs to craft an actionable digital advertising strategy.

Outcome

By comparing the clients audience to AdveriCorp's base audience, which is 187x larger, the cosmetics company was able to identify what segments made their customers unique. These segments not only provided a clearer picture of their Purchaser Profile, they can also be used to activate media that will allow *Chevlour* to drive sales.

Executive Summary

► Chevlour customers are typically women between the ages of 51 and 57, who are married and live in a single family house for which they often do the grocery shopping. They earn between \$0 and \$25k annually and are registered voters. They enjoy gardening, photography, scrapbooking, and caring for their pets. They also frequently purchase hair care products. While they may shop at stores focused on value like Walmart, Target, and Sam's club, they are also concerned with creating an appearance of luxury.

► While Chevlour's customers read magazines like Health, In Style, and People slightly more than the general population, caution should be taken when using this as an avenue to stimulate sales. This is because magazine subscribers represent a relatively low percentage of consumers and AdveriCorp's audience.

► AdveriCorp could provide access to over 8.5 billion segment records of potential customers to grow Chevlour sales.

Task 1 – Sample Size

A list of the 3rd party segments that have a large enough sample size to be meaningful can be found in the Excel sheet titled “Task 1”.

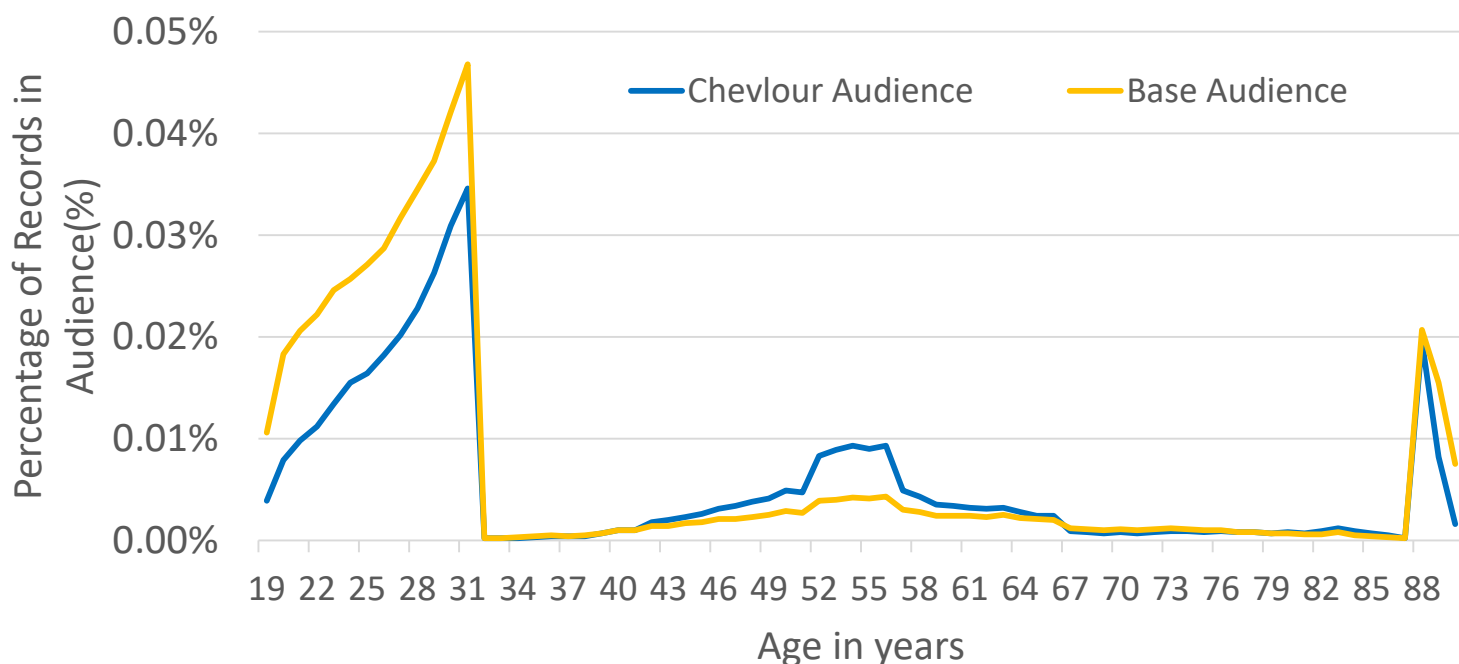
There are a few factors that go into determining how large of a sample size you need for it to be meaningful. Specifically, an estimate of the population, the margin of error, and confidence level. Given a population the size of the US (or AdveriCorp’s base audience), with 5% margin of error and a 95% confidence level, the suggested sample size is 385. While these are the standard values, they can always be made more conservative. For example, a 3% margin of error would require 1,068 samples. While this would exclude nearly 1/5 of the segments, it would be more precise.

I created a column “suff_sample” where ‘1’ represents compare_n values ≥ 385 , and ‘0’ for values < 385 . Of the 4,819 original segments, 3,639 remained.

As a check on the integrity of the data, I examined it for any missing values or inconsistent numbers of total records (compare_n/compare_perc) and marked them in the check_range column. While there were a number of typos in the segment names, most were understandable in their current state. For visualization purposes, I changed “Macey’s” to “Macy’s” and “Jenneys” to “JC Penneys”.

Task 2 – Age Comparison

Initially, I came across the segments from Experian that had binned age ranges. However, I ended up not using those segments for two reasons. The first is that they were split into ‘Person 1’ and ‘Person 2’ which is a difference I am unclear on. Second, I discovered there was a segment for birth years between 1926 and 1997. As age is a continuous variable I opted to plot this data instead to make my comparison.



From this figure, it is clear that there are fewer customers in the 19-31 age range compared to the general population. The key age range for Chevlour customers is between 41-65 years old with the largest differences between 51-57.

Task 3 – Interesting Segments

A list of the 3rd party segments that could be considered interesting for the presentation to Chevlour can be found in the Excel sheet titled “Task 3”.

To identify interesting segments you could look at `compare_n` or `compare_perc` and sort their values in a descending order. However, this probably isn't the best way to identify unique characteristics of the target audience who are interested in ChevLOURS product. This is essentially a base rate problem. For example, if you sort by `compare_perc`, the top segment is being a regular buyer of salty snacks (99.6%). However, based on AdveriCorp's database, we see that `base_perc` is also high (98.2%). This tells us that buying salty snacks isn't the best indicator of what makes ChevLOURS consumers unique, just that most people like salty foods! A better way to generate a short list is by comparing the difference between the percentage of ChevLOURS consumers and the general population. Subtracting `base_perc` from `compare_perc` gives us a measure called `perc_diff` (Being sure, of course, to filter out segments with an insufficient sample size). By targeting segments that are considerably higher, as indicated by `perc_diff`, we can be more confident that the media we activate will drive sales.

We can make an arbitrary cut off value depending on the number of segments the client is looking to target. Segments with $> 1.5\%$ `perc_diff` would return 247 segments. A more conservative cutoff of $>5\%$ would return 128. I chose to use the more liberal 1.5% cutoff as it includes more segments from all 3 sources, but created the formula in such a way where it could be easily updated based on discussions with AdveriCorp's sales team. From these segments we can start to build a better understanding of the ChevLOURS brand consumer.

The segment with the largest difference is unsurprisingly regular buyers of hair coloring products (51.2% more than general population). This is a positive indicator though that calculating the percentage difference is a useful metric for identifying interesting segments. The other top 50 segments all relate to buying behavior. A few highlights of interest are cleaning implements and supplies, hair spray and hair styling products, dog food, cosmetics, coffee, refrigerated baked goods and meals, and yogurt.

While the majority of selected segments are about buying behavior, other sources reveal more about the average consumer. They are typically female, registered voters who are married and live in a single family house and are of Western European descent. They enjoy gardening, photography, scrapbooking, and caring for their pets.

Task 4 – Interesting Segments

There are a number of interesting segments that contradict the Purchaser Profile and can be found in the Excel sheet titled “Task 4”.

For these analyses I returned to using `compare_perc` to look at Targeted Segments based on the Purchaser Profile. I did this since I cared less about if these segments were above, the same, or below the base audience, and more about understanding the Chevlor brand consumers. I examined the available segments and focused on things like Consumer Attitudes (TrueTouch), life stage + socioeconomic status (Mosaic), Entertainment, Shopping, Income, and Number of Children.

To get a broad overview, I made use of the touchpoints provided by TrueTouch. I found that in the compare audience, ‘No time like the present’ was the most prevalent (.6%). These are described as being “impulsive shoppers who want to keep up with the Joneses”; “Work Hard, Play Hard”, a phrase taken directly from the Profile was represented much less in the audience (.1%). “Stop and smell the roses”, eco-conscious buyers were last (.04%). In conjunction with the age data, it is unlikely Chevlor’s consumers are typically using their products for experimental and edgy neon colored hairstyles.

I discovered that the top 3 stores visited in last 3 months were Walmart, Target, and Sam's Club. I’ve plotted the top 20 stores that Chevlor consumers visited in the past 3 months. They are equally likely to visit the Dollar General Store as they are Vera Bradley.

3.6% of the audience make \$0-25k a year, while only .3% make 100-250k.

Reality TV makes up the largest percentage of TV shows viewed (.8%).

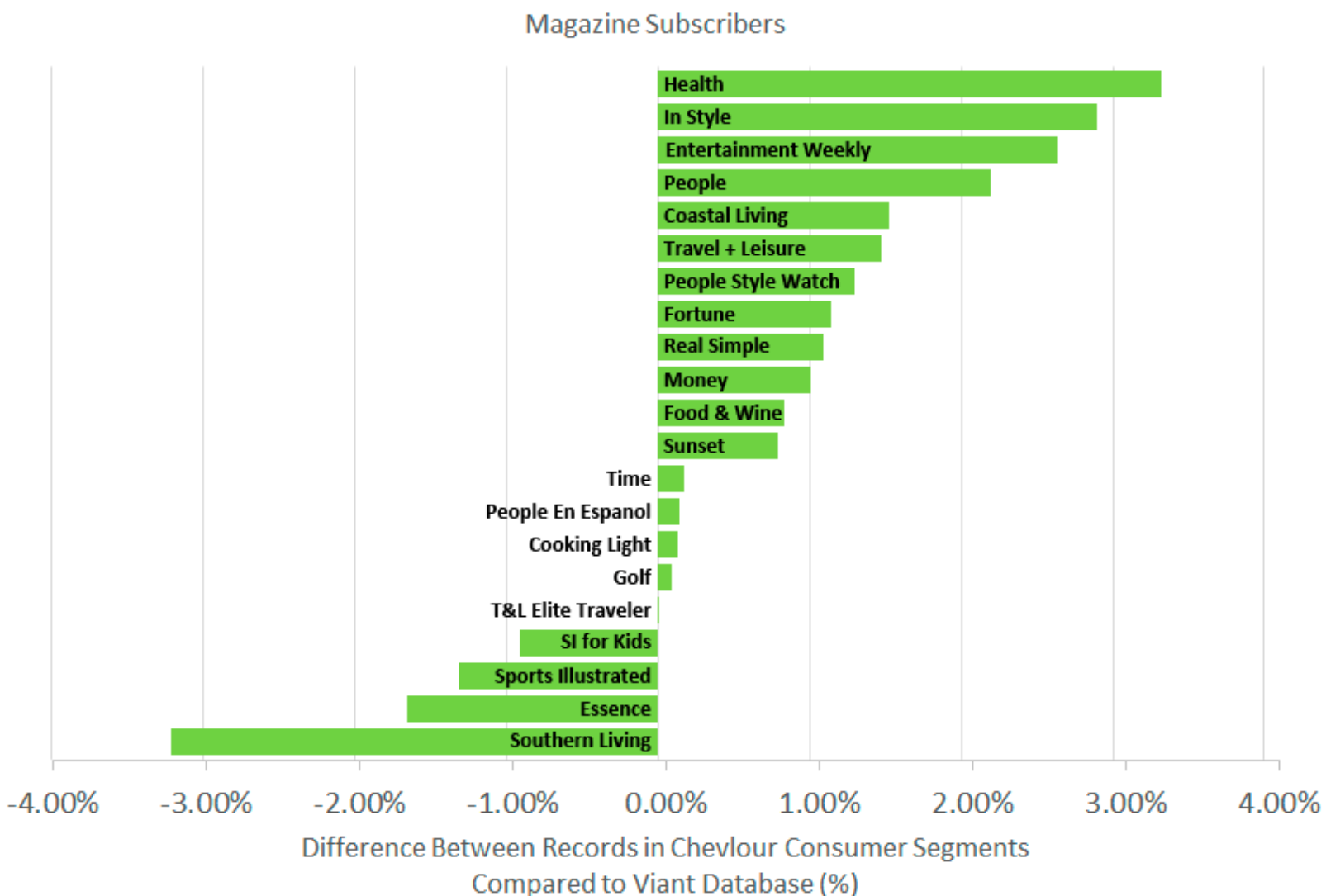
These segments and others like them don't necessarily paint a picture of an individual where “Quality and Luxury are key to their entire lives.” Instead, there is more evidence for a pragmatic sensibility overlaid with a concern for keeping up appearances. So while it may be wisest for the client to continue to market their product as a high end beauty product to maintain the appearance of their brand, they should consider media buys that reach the audience described above.



Task 5 – Magazine Subscriptions

Magazine reading habits, similar to the 3rd party data, is best viewed in comparison to AdveriCorp's larger database (perc_diff) as it provides a better approximation of the general population. So while People, Southern Living, and Cooking Light are the three most read magazines by Chevlor customers, by looking at the perc_diff, we see that Health (+3.3%), In Style (+2.9%), Entertainment Weekly (+2.6%), and People (+2.2%) are the four publications which customers read more on average than the broader sample. Even though Southern Living is in the top 3 for number of readers in the consumer audience, it has the largest difference compared to the base audience rate. Chevlor consumers are ~3.3% less likely to read Southern Living compared to the base audience rate.

In Style certainly matches with the Purchaser Profile of being fashion conscious. Entertainment Weekly and People are more entertainment and celebrity focused. Taken in conjunction with the TV habits examined previously, this suggests that Chevlor customers find keeping up with current pop culture important.



Task 6 – Final List

Taken together, the third party and magazine segments can be compiled into a final list that can be found in the Excel sheet titled “Task 6”.

As I completed this task, I was uncertain about the number of segments to include in the final list. I do not know what a reasonable number is to construct an advertising strategy. I used my personal judgment informed by histograms of the sources of the data to select the arbitrary values I did. I’m confident though that communicating with an experienced sales team would help inform this choice. For now, I did what I could by using pivot tables and a simple formula tied to the absolute cell reference so that if more or less segments should be included in the final list it would take almost no time to adjust.

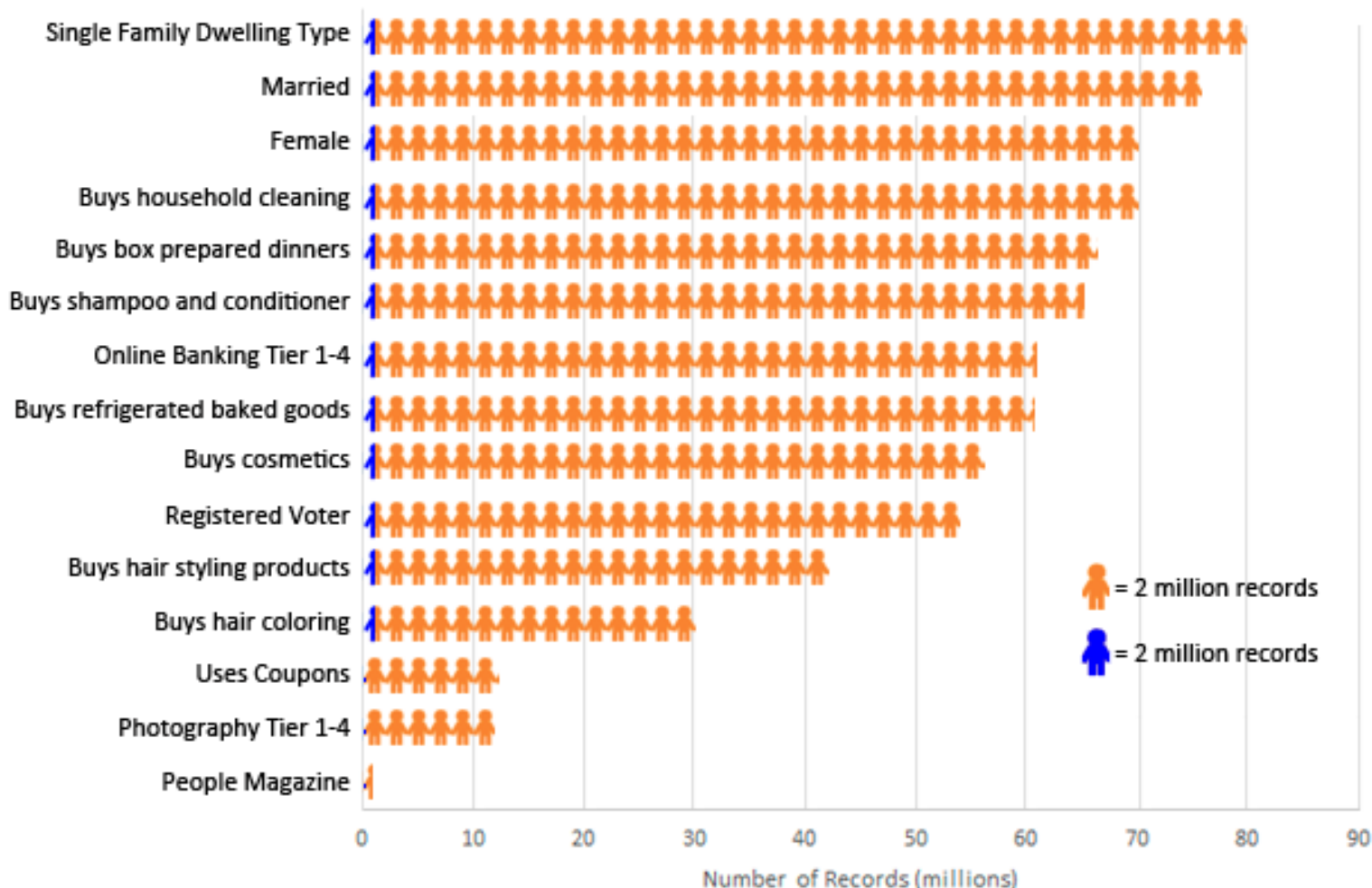
The 252 segments that remain from the original 4,849 represent consumers who can be targeted digitally. The requirements for the segments on the final list are that they have a sufficiently large sample size (≥ 385), have no inconsistencies in the total number of records in the audience, and most importantly differ from the general audience to account for base rate information ($\text{perc_diff} \geq 1.5\%$). By leveraging AdveriCorp’s base audience, Chevlor has the option to reach millions of potential customers.

The big question that remains is just how segments could be targeted by using AdveriCorp’s platform. The answer, using the final list of segments, is over 8.5 billion. This number was computed by comparing the number of records in the Chevlor consumer audience to the number in the base audience for that same segment. For example, there were 1.2 million records of consumers who lived in a single family dwelling unit. AdveriCorp’s base audience has a total of nearly 80 million records for that segment. Of course, there may be some overlap in the consumer audience and AdveriCorp’s, but by assuming 100% overlap and subtracting these numbers we can create a conservative estimate how just how much reach is possible by partnering with AdveriCorp. It’s interesting to note that this view of the data can be especially useful in identifying which segments to focus on. For instance, magazines may not be the best route to purchase advertisements in since the consumer audience already accounts for a larger proportion of the general population.

The visualization on the next page includes a selection of interesting segments comparing the consumer and base audiences. The advantage of AdveriCorp’s database is clear. By selecting segments that are prevalent amongst Chevlor consumers and digitally targeting advertisements through AdveriCorp, we can activate media to drive sales.

Task 6 – Final List (continued)

Potential Reach with Viant



Data that could be useful for additional analyses:

- Historical data of previous media buys. Knowing which segments were targeted previously and observing how revenue changed in response would provide a useful additional metric to weight each segment by. This would potentially allow for regression analyses to be performed to predict what magnitude of sales changes to expect for the proposed strategy.
- Updated list of consumers. As the client builds their business they will be gathering more and more records that can be processed through AdveriCorp's Audience Insight tool. Not only would this allow us to examine trends of consumer demographics and behavior (being able to answer questions like "are more young people buying our product than before?"), it would also demonstrate AdveriCorp's continued value as a provider of actionable strategies for driving sales.

I hope you have found this case study a useful tool in assessing my ability to complete a task that would be of value to AdveriCorp.