# Coalition for Innovative Media Measurement

# Television Data Match Rate & Bias Study

## September 2023

For People Who Are Exploring Big Data Integrations
For Advanced Advertising Solutions

*What to expect, what you might see,*
*what you might want to do about it*

*Reading Time: 10 Minutes*

# CONTENTS

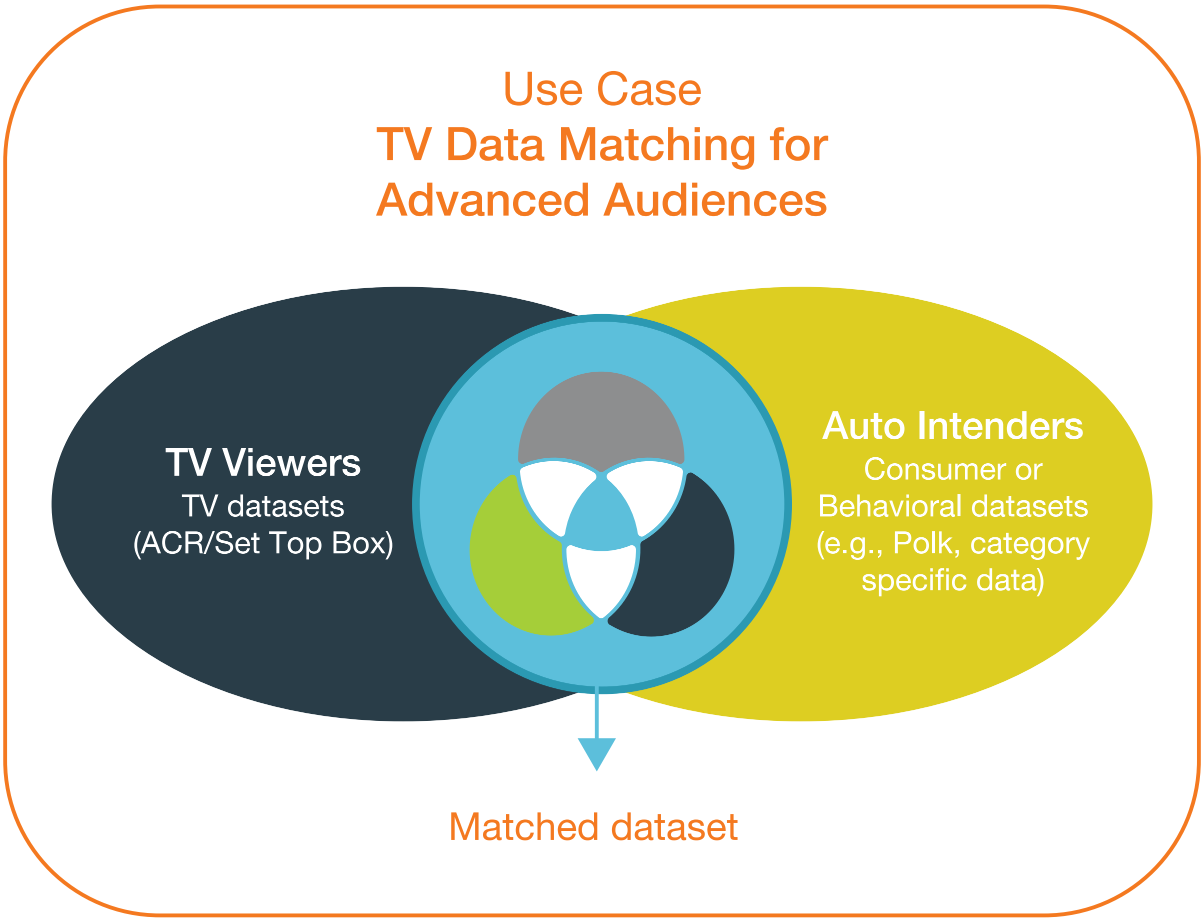
# Introduction

Identity resolution is crucial to advanced television targeting and attribution. These techniques often depend on successfully matching device-level or household-level TV viewing data with other datasets to target based on audience characteristics, retaining the scale and integrity of those demographic or behavioral characteristics. No matches are perfect; matching is a complicated blend of art and science. High match rates are very beneficial in TV planning and activation processes. However, assessing the accuracy of these matches is essential to create strategies to address any potential measurement biases. Combining high and precise match rates can improve the reliability and efficiency of the entire measurement process.

# Introduction

The purpose of this paper is to provide more insight and explanations into the processes and results generated by data matching and identity resolution, focused exclusively on matches with television viewing data.

They also matched TV data to consumer and household demographic cluster data provided by Claritas, as well as to behavioral outcomes data (visits to casual and quick serve restaurants in a two-week period) provided by PlaceIQ (now Precisely).




The study had a relatively simple objective-investigate the extent to which post-match television audience viewing profiles and levels differed from the original profiles and levels. And then determine what caused the differences.

We also investigated whether the resulting matched dataset is sufficient for typical use cases like targeting and attribution.

The study was undertaken between 2021-2022 and involved a comparative analysis of five identity resolution (IDR) providers who matched TV data from one ACR and one STB provider.

The IDR companies conducted six matches in total-their graph to the two TV datasets (one ACR, one STB), the two TV datasets to the demographic cluster data and finally, the two TV datasets to the behavioral outcome data. So, each of those three matches were executed for the two TV data sources. This process resulted in four matched data sets for each IDR provider: TV viewing by demo and TV viewing by outcomes, for each of the two TV viewing sources.

# Introduction

## Study Participants

The main criteria used to evaluate the matches performed by the IDR providers in this study were match rates and match bias. Bear in mind, these are just two of many factors used in evaluating data matches. Match rates are important because they demonstrate the overlap between two datasets that can be found with a common identifier, such as hashed email, device address, mobile Ad IDs or cookies. It is a measure simply of the number of many records the two datasets have in common.

While the size/scale of the matched dataset is a key consideration, this study also considered match bias, looking at the representativeness of the matched database and its suitability for TV planning and attribution. A key consideration beyond the scope of this study is whether the households, devices within the households and TV data had been paired up accurately. Other considerations include determining whether the ID graphs reflect all the key demos and other population segments and the freshness of the match keys being used. Matching with match keys that have expired, such as an IP address that has been reassigned to another household, or an old email address no longer in use, will result in an incorrect, inaccurate match. Match rate issues are especially important when it comes to attribution analysis. For CTV campaigns, the dependence on IP addresses for matching can lead to low match rates, as IP address used for activation may have changed by the time the attribution analysis takes place. As a result, the device actually reached by the campaign may not be included in the attribution analysis. Instead, the analysis may include a different device that has been assigned the original IP address, potentially resulting in a misreported outcome (e.g. no response) being attributed to the campaign.

All of these issues are important dimensions of match quality or accuracy and matter a great deal. This study, however, evaluates a thinner slice of the issue, match rate and match bias. It does not specifically assess match accuracy. To do so would have required an independent validation dataset of IP addresses and names/addresses which, despite our best efforts, was not available.

Another thing to know about this study is that the source datasets (TV, Demographic and Product Usage) were taken at face value. They were not the subject of this analysis; we did not assess the accuracy of the base data. We only studied the impact of the match process. That's an important point. Of course, data users should always consider the quality of the source data they use. For instance, there is a pre-existing matching process embedded in TV data - devices have to be grouped into their

# Introduction

respective households. However, all the data used in this study was already householded and the strengths/weaknesses of that process was not taken into account in this study.

The IDR providers were not given any guidance about how to match the datasets, which household IDs to include, or which identifiers/ match keys to use. These decisions were left to their discretion and reflect their actual real-world processes, as far as we are aware. Some identity resolution providers utilized pre-existing crosswalks, or data exchange standards, across the different datasets, but others didn't. The presence of a crosswalk indicates that all the hard data integration work, the technical data science nitty-gritty, has already been worked out before the study. Interestingly, the study suggests that these crosswalks and the increased proficiency with specific datasets improve match rate.

Please keep in mind the results here are specific to the data sets and IDR processes used. While the results definitely prove the potential for biases due to the matching process of various degrees, they do not offer a generalizable finding that relate to *any or all* matches of *any* data sets. These are not norms. Other match study results will undoubtedly vary. However, the findings are useful in offering insights and helping to identify areas to investigate and questions to ask as you match datasets as part of advanced television planning, activation and evaluation.

# Key Findings

# Key Findings

## IDR ID Graphs to TV Data Match Rate

This analysis answered questions such as: Is there a sufficient TV dataset for target audience reporting after the match? Are matched TV viewing estimates adequate for planning of a TV campaign? How much variability is there across provider match rates?

**73%** The average match rate between the two TV data sets and the five IDR providers' ID graphs.

**86%** Highest match rate across the providers and TV data sets.

**63%** Lowest match rate.

### Post-Match Scale

Across the six IDR providers, an average of

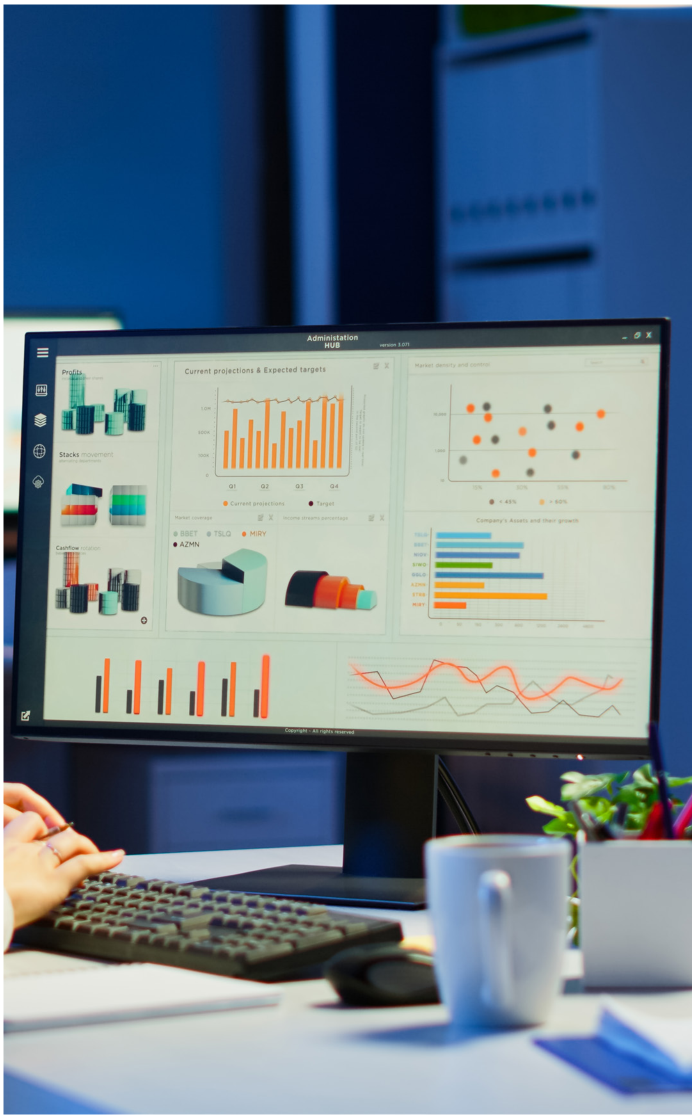
**1.9M and 1.5M Households**

remained across post-match with the two different TV data sets, from an initial set of approximately 2M households from each TV data provider.

This scale is robust and would be more than adequate for the most common national television use cases - for example, to reach planning for a target of consumers who intend to be in the market for a new or used car in the next 3-6 months, the target would be about 486K. A program with a 0.1 rating would have almost 500 viewers and a 9% relative error around that rating estimate at the 95% confidence level.

### What this means...

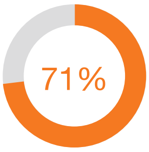
The initial match only linked TV data with ID providers' ID graphs. In practice, this is probably the most common match. What we learned is that the match process retains robust datasets for advanced TV analysis. It's also clear that individual IDR provider match performance varies a lot. We believe the variation is a function of the prevalence of specific match keys like hashed emails, IP addresses or Mobile Ad IDs as well as experience with specific datasets that are being matched, including existing crosswalks.



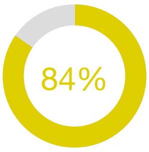
# Key Findings

## IDR ID Graphs to TV Data & Demographic Dataset Match Rate

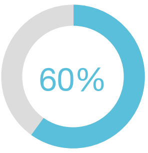
This analysis addressed whether there is a sufficient TV dataset for target audience reporting after the demographic match. Are there demographic shifts in viewing data after the match? Are there skews in household demographic representativeness vs. total US or regional geography?




The average match rate between Claritas demographic data and the two TV viewing data sets, among all five IDR providers, as a percent of the initial TV data.




Highest match rate among providers, TV Datasets and Claritas data.




Lowest match rate.

Post-match, significant dataset sizes remain:

**1.8M and 1.4M**

households with TV viewing and demos, on average across the two TV datasets and the five IDRs.

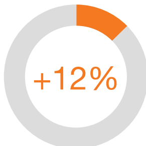
Claritas is a near-census dataset and its match rates to the IDRs' ID graphs were high. That's why there is relatively little difference in scale between the households remaining after the initial match (TV to IDR graph only) and the second match (TV to IDR Graph to Claritas).

### Post-Match Bias

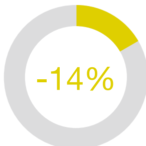
By matching the Claritas' demographic data to TV data, we were able to discern post-match changes in demographic profiles of the television data - in other words, did some of the IDR provider matches data spines miss out more of some socioeconomic or demographic segments than others?

Compared to the pre-matched television source data, we found the matching process produces a viewer profile that skews higher income and older.

### Income Bias




Average over representation of highest income segment (median HH income $122K) across providers.




Under-representation of lowest income segment (median HH income $22K).

Three IDR providers showed even higher over representation of high income - +24% to +25% and more under representation of lower income -27% to -31%.

# Key Findings

## Age Bias

+10% Average over representation of older half of population (TV Dataset 1).

+20% Average over representation of older half of population (TV Dataset 2).

-9% Average under representation of younger half of population (TV Dataset 1).

-18% Average under representation of younger half of population (TV Dataset 2).

The study showed that these biases are driven inherently by skews in the identity resolution graphs and match processes used by the different IDR providers. This is not surprising since the identity resolution datasets are not national probability datasets - they are just extremely large datasets that tend to be somewhat older, more upscale and established less mobile versus the national US norm.

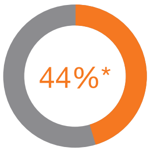
### What this means...

Matching TV data with IDR provider graphs and Claritas data maintains dataset sizes for analysis and highlights demographic skews. Upscale and older demographics are over represented in post-match data due to the matching process. Providers show similar overstatement patterns, though to varying degrees. We believe addressing these skews through weighting is crucial, assuming the use case calls for a nationally-representative data.

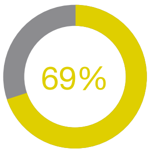
# Key Findings

## IDR ID Graphs to TV Data & Location Data Match Rate

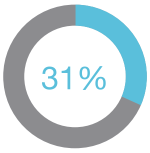
This part of the analysis answered the different questions: Is there a sufficient TV dataset for target audience reporting after the match? Are there shifts in share of restaurant visits after the match? Are there shifts in regional skews? These types of data might be used for targeting, or for outcome measurement.




Average match rate between PlaceIQ and the two TV viewing data sets.




Highest match rate across providers.

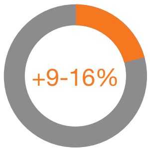



Lowest match rate.

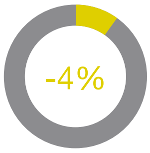
Post-match, significant datasets remain:
**1.1M and 0.9M**
households from two TV datasets.

*Importantly, only a subset of the population (23mm homes) visited Casual, Fast-Casual or Quick-Serve restaurants (QSR) in the two-week test period, so the dataset and match rates are smaller than the TV or demographic data matches.

As we saw with the matched demographic data, matching TV data to location data can change the restaurant visitor profile of the post-matched data. The matches to PlaceIQ increased the share of visits to Casual and Fast Casual restaurants compared to the unmatched visitation data. Quick Serve Restaurants saw a relatively small under representation of visits.




Over representation of Fast Casual Restaurant Visits, on average, with the TV data matches. We saw a significant range by IDR provider (from -38% fewer visits to +46% more visits.)




Under representation of QSR restaurant visits post-match, on average, for the TV data matches though there is a significant range by IDR provider (from 20% fewer visits to 5% more visits)

We also saw considerable variability among IDR providers when it came to individual restaurant chain share of visits. Three of the providers showed share of visits to McDonald's right on par with pre-matched data, though one provider over represented McDonald's share of visits by 14-17% in the TV data matches while another underrepresented the visits by 17-18%.

### What this means...

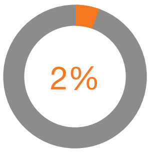
Matching behavioral visitation data with television data yields robust datasets for analysis. However, skews toward certain restaurant types and brands occur after matching. These skews can impact planning decisions and attribution outcomes, necessitating the use of weighting to mitigate their effects if the use case calls for a nationally-representative data.

# Post-Match Television Viewing Profiles

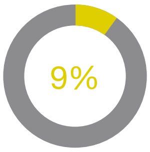
The study was designed to reveal the extent to which post-match television viewer profiles differ from the original viewer profile, e.g., how does the skew towards older, upper income consumers in the IDR graphs impact television viewing levels? Are the post-match households heavier viewers? Are some dayparts watched more than others? How much variability is there by IDR provider in these viewing hours?

# Post-Match Television Viewing Profiles

We found slightly higher viewing levels after the two sources of viewing data were matched to the IDR providers' graphs.




Higher post-match viewing for the TV1 dataset.




Higher post-match viewing for TV2, on average.

Analysis of the match data showed that TV2 dataset post-match viewing levels were higher due to the inclusion of fewer light viewing households (-12%) and more heavy viewing households (+6%) in the post-matched data. This produced more viewing hours in Primetime and to a small degree, sports.

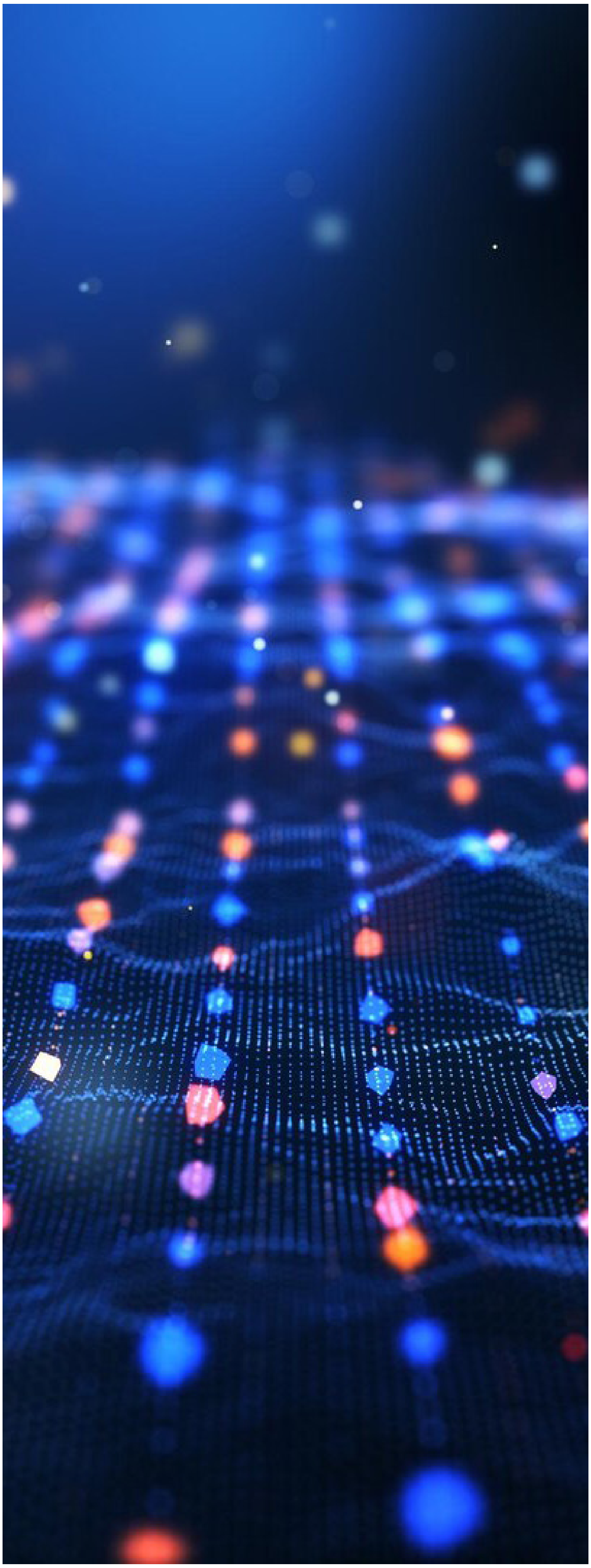
Individual IDR provider post-match TV data results varied from an understatement of **4%** of minutes viewed to an overstatement of **24%.**

These over representations were identical for Weekday and Weekend Total Day dayparts.

Compared to pre-matched television data, matches also produce higher viewing hours to African-American and Hispanic programming; especially with TV 2 dataset.

This is likely due to the post-match skew to older, more upscale HHs. But bear in mind that these results varied by TV source and IDR provider. The important finding here is that the matching process impacts TV viewing levels in a variety of ways, depending upon provider.

Do the differences matter enough? Would the same media planning decisions be made? Depending on the provider, CPMs could be 10-20% lower for audiences as a result of matching and match provider. Would this mean different media effectiveness decisions? Depending on the provider, the reported ROI/ ROAS could be 20% higher.



# Recommendations

Please keep in mind the results here were specific to the data sets and IDR processes used. While the matching process leaves sufficient data scale for most media measurement tasks, the results definitely prove the potential for biases to varying degrees, due to the matching process. But the study does not offer a generalizable finding that relates to any match of any data set. In short, your results may vary.

# Recommendations

Prior to selecting an IDR provider, have them share blinded historical studies to assist your IDR selection.

Evaluate post-match data carefully; make sure it sufficiently approximates the source data. If you're matching television data, key attributes like income, age, ethnicity, presence of children and region are extremely relevant. If the use case applies, determine if weighting post-match data is necessary to correct biases. We agree with the MRC, which calls for transparent disclosure of match rates and any systematic bias introduced by the match process that are relevant to the particular use case. (Outcomes and Data Quality Standards, 2022).

Work with providers to develop data crosswalks, or data exchange standards, across datasets to ensure quality matches with minimal biases. Matching in cleanrooms, which seems to be increasingly more prevalent, should also reduce variability and yield better results. And consider undertaking a match accuracy study, which can be conducted only with a provider with access to both television set IP addresses and physical addresses, the key to identity resolution.

When designing an attribution study, marketers should construct thoughtful test versus control groups to increase the probability that any IDR biases are applied equally to both groups.

CIMM also recommends the industry adopt a transparent, standardized reporting system that provides key facts about the quality and composition of matched data used in audience targeting and measurement.

The intent is to inform users about the relative skews of the post-matched data sources on standard television metrics - as well as their applicability for media strategy and investment decisions - and to address issues of transparency.

A sample of this report appears in the appendix.

# Conclusion

This work was shared with all the IDR providers and they have the benefit of our analysis. We hope that after seeing their performance benchmarked against others, they will seek improvements. We strongly encourage transparent disclosure of match rates and biases.

# Conclusion

We believe this study provides a glimpse into the match process as well as ways to navigate issues related to quality and comparability of IDR provider graphs as this space continues to evolve.

Data matching is hugely important, and hugely challenging for many reasons. We applaud CIMM for sponsoring the study and the efforts of all the identity resolution companies and data providers that participated in this study and helped in this industry educational effort.

# Appendix

# STANDARDIZED REPORT FOR TV DATA MATCHES - Audience Skew & Recommended Adjustments

A standard template for comparing pre-match and post-match TV data sample compositions as well as the weighting that must be done to bring the sample back into balance and reflect US population. Identity providers should report the TV data match rate (%), the resulting match count (000) as well as the household demographic composition percentages and the number of average household weekly viewing hours (000) for full transparency

| | | TV Data Source | ID Graph/TV Data Source | Consumer Data & ID Graph TV Data Source | Adjustments |
|---|---|---|---|---|---|
| | TV Data Set HH Match % | | | | |
| | HH Match Count (000) | | | | |
| | | Demographic Compostion | | | Recommended Weighting |
| Income | % | | | | |
| | Top 20% | | | | |
| | 20% | | | | |
| | Mid 20% | | | | |
| | 20% | | | | |
| | Low 25% | | | | |
| | | 100% | 100% | 100% | 100% |
| Age : Household Head | <35 | | | | |
| | 35-64 | | | | |
| | 65+ | | | | |
| | | 100% | 100% | 100% | 100% |
| Ethnicity/Race | Asian | | | | |
| | Hispanic | | | | |
| | African Amercian | | | | |
| | | 100% | 100% | 100% | 100% |
| Presence of Children | <Age 18 | | | | |
| | <Age 6 | | | | |
| | | 100% | 100% | 100% | 100% |
| Region | Northeast | | | | |
| | Midwest | | | | |
| | South | | | | |
| | West | | | | |
| | | 100% | 100% | 100% | 100% |
| TV Attributes | Average HH Wkly Viewing Hrs (000) | | | | |

cimm

coalition for innovative media measurement

https://cimm-us.org/