

Hitwise AudienceView™

Data & Methodology Overview

Table of Contents

1.0	Introduction.....	3
1.1	About Hitwise and AudienceView™	3
1.2	Our Consumer Profiling Approach	4
2.0	Data Methodology Overview.....	6
2.1	Primary Data Sources	6
2.2	Behavioral Signature Linkage	8
3.0	Estimation through Weighting and Sample Projection.....	10
3.1	The Sample.....	10
3.2	Understanding weighting and scaling	12
3.3	Estimating Audience Behavior	12
3.4	Estimating Persons / Unique Users	12
4.0	Understanding Reported Metrics	14
4.1	Unique Users / Persons.....	14
4.2	Visits.....	15
4.3	Search.....	15

1.0 Introduction

1.1 About Hitwise and AudienceView™

Hitwise, a division of Connexity, is a leading consumer intelligence business. We serve marketers, media brands and advertising agencies globally and help performance-driven brands understand and engage their best consumer segments through expertly applied intelligence. By integrating behavioral, attitudinal and psychographic data, we provide a vivid and complete understanding of how consumers think and what they do in a multi-channel world.

As a result of the changing consumer environment and the rapid proliferation of content, channels and connected devices and as a result of the evolution of digital marketing and data-driven marketing, the clients we serve need solutions that provide a holistic view of the consumer in terms of their multi-channel usage, attitudinal mind-set and purchasing behaviors.

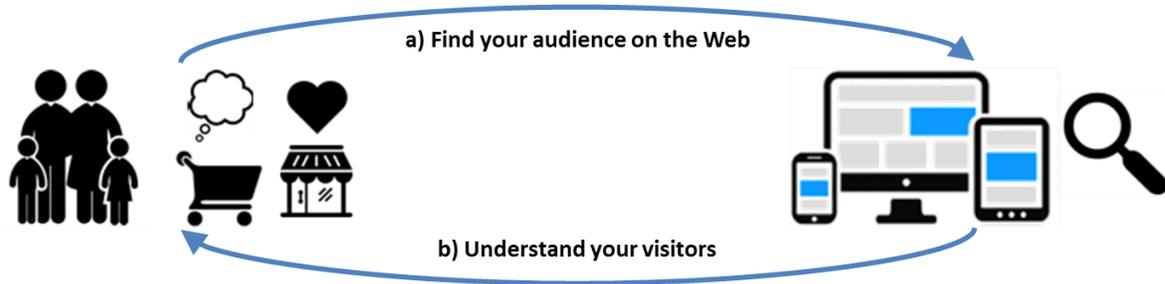
In response, we have developed the *AV360 database*, created through the integration of data from Hitwise and data from leading consumer research surveys and panels. This unique combined dataset is delivered to clients via AudienceView, an analytic platform that offers tools for creating unique audience segments on-the-fly and powerful analytic reporting to understand how those consumer segments think and behave.

AudienceView helps marketers, media and agencies achieve meaningful engagement with customers and audiences by fueling the planning process with the most complete insight into who consumers are, what they think and how they behave with media. The AV360 database links the vast store of online website visit and search measures from Hitwise (the largest online panel in each of our markets; a sum of approximately 12 Million users monthly across all three Hitwise markets) with the rich and vivid consumer profile data reported in leading market research survey and panel data, such as the Simmons National Consumer Study (one of the largest and most respected surveys of the American consumer).

Marketers have the ability to draw upon offline and online insights of specific consumer targets via AudienceView to develop more distinct and complete cross-channel insights on a single customer segment or consumer audience target.

1.2 Our Consumer Profiling Approach

AudienceView enables marketers, media and agencies to better understand who consumers are and what they do in both online and offline environments. AudienceView delivers insight that can describe: a) digital and online behavior of a target customer or consumer segment, or b) a multi-dimensional profile of an online audience.



a. Digital and online behavior of a target segment

AudienceView extends knowledge of a defined consumer segment to include deep insight into that segment’s online and digital behavior. This type of analysis helps our clients better engage customers with online content, and better plan campaigns to reach prospective customers.

Research Question	Consumer Segment Definition	Insights Learned
What do my customers do online and across media channels?		

To answer research questions about the online or digital behaviors of a consumer segment, the target segment is defined by demographic, lifestyle and/or attitudinal characteristics (for example, 18-34 year-old mothers who buy Coca-Cola products). AudienceView creates the link to a consumer population segment in the online panel. Online behavior such as websites visited using both mobile and desktop devices, the clickstream of those website visits, and online searches provide insight into the target segment’s online behaviors.

b. Multi-dimensional profile of an online audience

AudienceView enables multi-dimensional profiling of online audience targets. This type of analysis helps our clients better understand and communicate with the audiences they plan to engage, whether visitors to their own online properties or to competitive properties.

Research Question	Consumer Segment Definition	Insights Learned
<p>Who are my visitors and what are their motivations?</p>		

To answer research questions about the profile of online consumers, the target audience is generally defined by their online behavior (for example, people who searched for “new cars” and who visited www.ebay.com). AudienceView creates the link to a highly correlated consumer population segment measured in the linked consumer research database to provide insight about that audience segment’s consumer profile. The profile includes demographics and lifestyles, attitudes and opinions, product and brand preferences, and cross-channel device behavior.

2.0 Data Methodology Overview

2.1 Primary Data Sources

The AV360 database that powers AudienceView leverages two primary data sources – web-based behavioral data sourced by Hitwise and consumer survey research sourced from well-respected and market-leading research providers in the United States, the United Kingdom and Australia.

Hitwise

Hitwise provides data describing mobile and desktop behaviors for millions of Internet users and billions of events (visits, searches, clickstream behavior, conversions, etc.) passively collected on a continuous basis.

The Hitwise database is built using data from a number of long-term, strategic partnerships with companies who provide their customers with a variety of web products and services on desktop and mobile devices. The companies providing this data offer their customers a wide range of software applications and content such as mobile data use optimization, specialist content aggregations (hobbies and interests etc.) and Internet security. In return, customers formally agree to an anonymous analysis of their browsing behaviors. The very broad appeal of the software and content offered by our data partners helps insure the breadth of coverage in our sample.

The Hitwise database captures key behaviors by consumers including searches from all major search engines and visits to all web pages. The methodology distills these behaviours out of the mass of detailed traffic data that constitutes interactions on the Internet. Irrelevant data is removed such as, third-party advertising content, and non-human traffic as well as any personally identifiable information. What's left provides a clear picture of the searches that individual consumers perform and the websites visited, together with an anonymized user identifier, a timestamp and other key data including the device that they used.

USA - Simmons National Consumer Study

The Simmons NCS provides survey measures over 60,000 variables on what American consumers do in the course of their daily lives such as media usage, lifestyle and descriptors, brand buying habits and psychographics. (See table 1 below). The Survey is based on a national probability sample of approximately 25,000 U.S. adults, aged 18+ collected over each year. The key focus of the NCS is to provide an in-depth description of consumer behavior in the top advertising spend categories including national media, automotive, financial services, retail shopping, healthcare and electronics. The NCS also includes comprehensive measures on the influential Hispanic population.

The Simmons NCS methodology uses a stratified, multi-frame sample design to accurately measure buying behaviors, usage and attitudes to products, brands and media among both English and Spanish-speaking U.S. adults. The NCS leverages multi-modal data collection techniques in which households are contacted by either telephone or mail to obtain permission to participate, and then receive survey booklets mailed to eligible household members for self-completion.

UK – Lightspeed GMI Global Market Research

In the United Kingdom consumer lifestyle data is captured from a large scale online panel provided by the highly respected Lightspeed GMI organization. It captures key consumer demographics, a core set of buying behaviors, lifestyle descriptors and self-reported web usage for 20,000 adults aged 18+. See table 1.

Note – these data are scheduled to be available within AudienceView mid-2016.

Table 1 Summary of consumer research attributes from survey and panel-based sources

Category	Survey items		Category	Survey items	
	NCS	GMI		NCS	GMI
Apparel	✓		Health and beauty aids	✓	✓
Automotive	✓	✓	Hispanics only - Spanish-language media	✓	
Beverages - alcoholic drinks	✓	✓	Hispanics only - special demographics	✓	✓
Beverages - non-alcoholic drinks	✓		Home furnishings & appliances	✓	
Cable/television/radio	✓		Home improvement	✓	
Cell/smartphones/tablets	✓	✓	Household products	✓	
Cleaning products	✓		Intermedia	✓	✓
Computers	✓		Internet	✓	✓
Daily activity	✓		Lawn & garden	✓	
Dining	✓		Lifestyle (demographics)	✓	✓
Electronics	✓	✓	Lifestyle statements	✓	
Entertainment/leisure	✓		Medicine/drugs/ailments	✓	✓
Financial services	✓	✓	Pets & pet food	✓	✓
Food – baking	✓		Print media	✓	
Food - condiments	✓		Segmentation solutions	✓	
Food – dairy	✓		Shopping	✓	
Food – frozen	✓		Telephones & srvc (no cell/smartphones)	✓	
Food – general	✓		Tobacco	✓	✓
Food – meat	✓		Travel	✓	✓
Food - snack/dessert	✓				

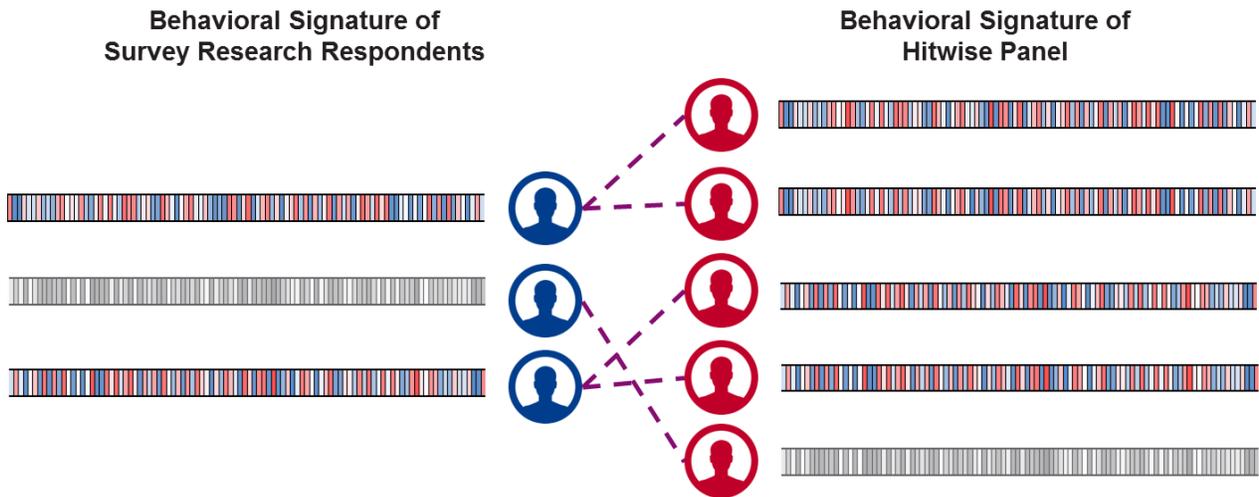
Australia – Shortly to be announced

Connexity is currently finalizing a partnership with a major provider of Australian consumer research data that will be incorporated into AudienceView. We are targeting market release around November 2016.

2.2 Behavioral Signature Linkage

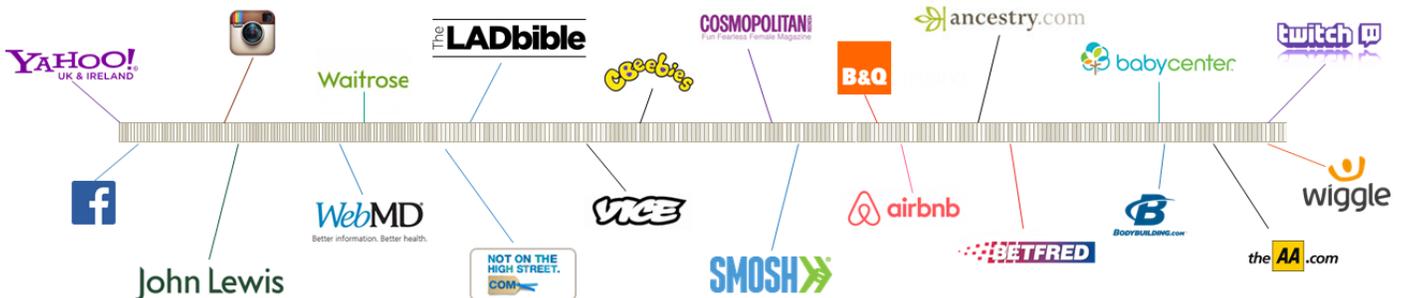
The AV360 database integration methodology relies on the identification of like-for-like records in the Hitwise and survey research databases using a type of imputation. Behavioral Signature Linkage relies on finding characteristic patterns of website visit behaviors across a core set of hundreds of websites found in both datasets. The characteristic visit patterns can then be used to match respondents across the Hitwise and survey research panels. Respondents in both data sets with the same behavioral patterns are matched on a 'look-alike' basis (see Figure 1).

Figure 1: Behavioral Signature Linkage



The websites used in the linked imputation process are a carefully selected and regularly reviewed set of sites for each market that reliably segment populations through clear signals of demographic differentiation, such as age, gender, income, and education. Figure 2 below shows an example of sites employed in the signature model in the United Kingdom.

Figure 2: Select Examples of Differentiating Websites Used in Establishing the Behavioral Signature



Note that the approach does not use any PII (personally identifiable information) on either survey respondents or individual panel members.

The methodology allows for tuning of the model by weighting individual websites to vary how much influence each website has in the match. Sites that should be more influential in the overall matching get a higher weight.

The behavioral signatures represent model, ideal patterns. Matching actual behavioral patterns against those models uses a standard algorithm called “nearest neighbor”. Nearest neighbor loosens the requirements for what is a match: allows patterns of behaviors to be matched that are most similar to the model, even if they do not exactly match.

Instead of using the matching process to generate a very large dataset including all attributes, the algorithm produces a lookup table that links the Hitwise and survey research respondents. The lookup table is regularly updated reflecting changes in both data sets and supports bi-directional linking – that is, segments targeted in either the Hitwise or the survey research data set can be used to identify look-alike records in the other data set. The survey research data set can donate demographic characteristics to Hitwise panelists, or Hitwise panelists can donate behaviors, such as search, back to survey research respondents. Figures 3a and 3b summarize this process.

Figure 3a: Bi-Directional Data Attribute Availability across Behavioral Signature Linkage

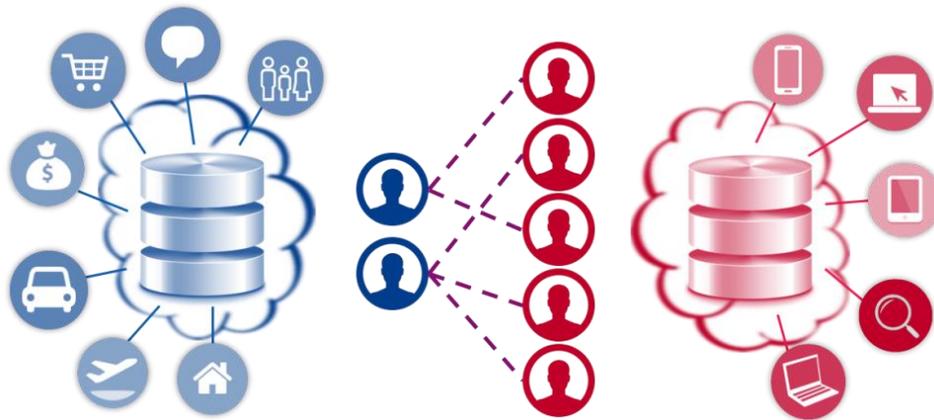
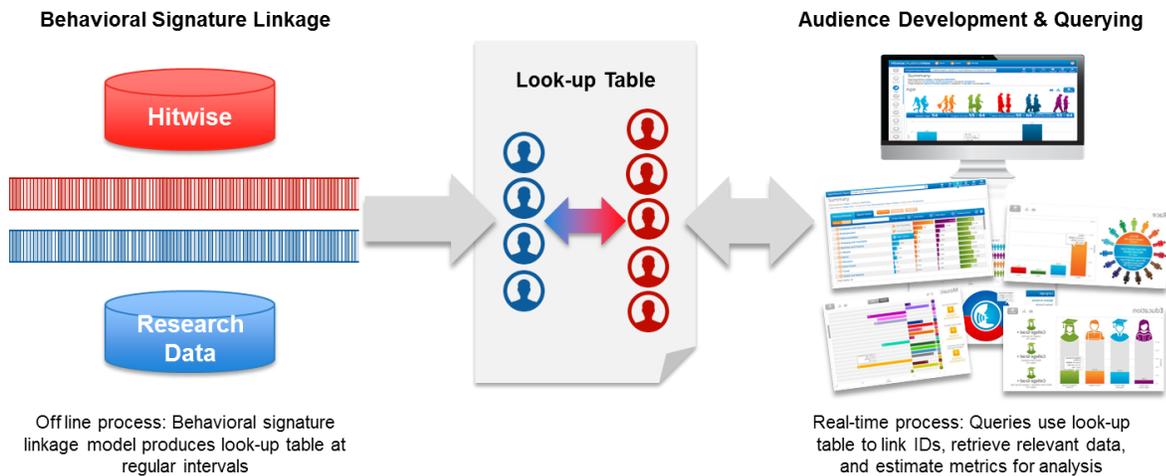


Figure 3b: AudienceView Behavioral Linkage Usage and Update Process



3.0 Estimation through Weighting and Sample Projection

3.1 The Sample

The data sample is the most important factor in determining the quality of any market research database; it is the foundation on which everything is built. The AudienceView sample size and selection process provides richness and depth in our reporting that allows us to provide detailed information about audiences and their online behaviors.

The Hitwise Sample:

The Hitwise sample, drawing from a diverse range of user panels with full visibility of secure traffic and mobile browsing data, is designed to provide both accuracy and reliability:

- Accurate representation of the Internet population through a broadly based sample that includes a balanced selection of the complete range of Internet users, modes and means of access and websites; and
- Reliable outcome measures through a very large sample size and rigorous processing procedures.

Usage data from large-scale Internet user panels allows Hitwise to accurately report information on hundreds of thousands of websites across over 170 industries. This dataset provides a detailed view of online consumer usage within key industries, including insights into the way consumers browse the web from mobile devices.

The Survey Research Sample:

In every market where survey research data is employed in AudienceView, the consumer lifestyle data is captured via surveys from large samples of respondents sourced from well-respected online consumer panel suppliers. While these surveys are smaller in scale than the Hitwise panel, they capture key demographics, lifestyle descriptors and core buying behaviors that are important to describing the profile of online audiences. For more information on these samples, see section 2.0 of this document.

Sample Size and Estimate Confidence:

In order to judge the confidence of estimates reported in AudienceView, it is important to understand how samples can contribute to the variability of estimations and the confidence of those estimations as indicated in AudienceView.

Understanding Sample Size

With AudienceView, sample size for a given estimate depends on a number of factors and will become narrower corresponding to the depth of analysis, the time period of the analysis and dimensions requested in the analysis.

- **Monthly total panel size:** While the overall average monthly raw sample size contributing to AudienceView is very large (for example, approximately 8-9 Million

panelists per month on average in the United States), the portion of that sample available for reporting in any given month will vary depending upon the percentage of sample meeting minimum active time required in the panel.

- **Segment definition:** When defining audience segments, the size of the sample in the originating available panel will narrow as more dimensions in audience attributes or behavior are added to the definition of an audience segment, particularly with an “AND” operator.
- **Reporting period:** While an original sample size is provided upon creation of a segment in some AudienceView markets (such as in the United States), it is important to understand that this sample size represents the sample available for the segment definition in the most recent 4 weeks of data. Because users of AudienceView are able to alter the reporting period upon report creation, the total sample size for any given report may vary depending upon the reporting period being analyzed - the longer the reporting period, the greater the sample size.

Understanding Confidence Indicators

Where available, Confidence Indicators are reported as asterisks (*) on metrics within reports to identify levels of confidence in reported estimates based on variability in sample and sample weighting. Confidence indicators are provided to guide you in understanding where to put more or less reliance on the absolute reported estimate. The confidence indicators that appear in AudienceView are summarized below.

Indicator	Confidence
<none>	High
*	Medium
**	Low

These indicators are a straight forward approach to represent results from the complex process of estimating confidence. This process employs resampling to directly ‘observe’ the variability of results in a sample of analytical cases and then extends those results (based on a number of factors) to additional cases. The ranges are coded as asterisks (*) representing high confidence (little variability) to low confidence (higher variability).

3.2 Understanding weighting and scaling

Weighting helps re-align the AudienceView sample with the underlying population for each market to ensure the accuracy of results. Scaling then allows sample measurements (such as, website visit numbers, keyword searches, clicks, unique users, etc.) to be projected up to represent the actual volume occurring within the respective population. All metrics reported in AudienceView originate from our base sample and are projected in reports made available in the product.

Weighting of the sample corrects for both random and systematic biases in the relative contributions of different user subgroups. Weightings used for Internet users cover geographic, demographic and socio-economic factors.

Established third-party population measures and population measures of Internet users are used for weighting. Where possible, those estimates are drawn from the most reliable government census sources with recognized industry measures to supplement and extend official statistics.

3.3 Estimating Audience Behavior

Estimates of audience behavior, such as visitation to websites, are based on a count of activity observed in our sample. These activity counts are weighted and scaled up (projected) from the sample. Behaviors 'inherit' the weights of the actors; that is, our weighting ensures an accurate representation of Internet users and that representation is carried forward to the actions. Scaling to produce final estimates relies on a sophisticated process of identifying a wide range of properties for which actual activity is known. For example, in producing web site visit estimates, a wide range of websites are used for which actual traffic numbers are known through site analytics. Scaling factors are then developed by comparing the range of those sites analytics against the equivalent sample estimates. Because traffic levels for many sites reliably vary between weekdays and weekends, separate scaling factors are computed across the week.

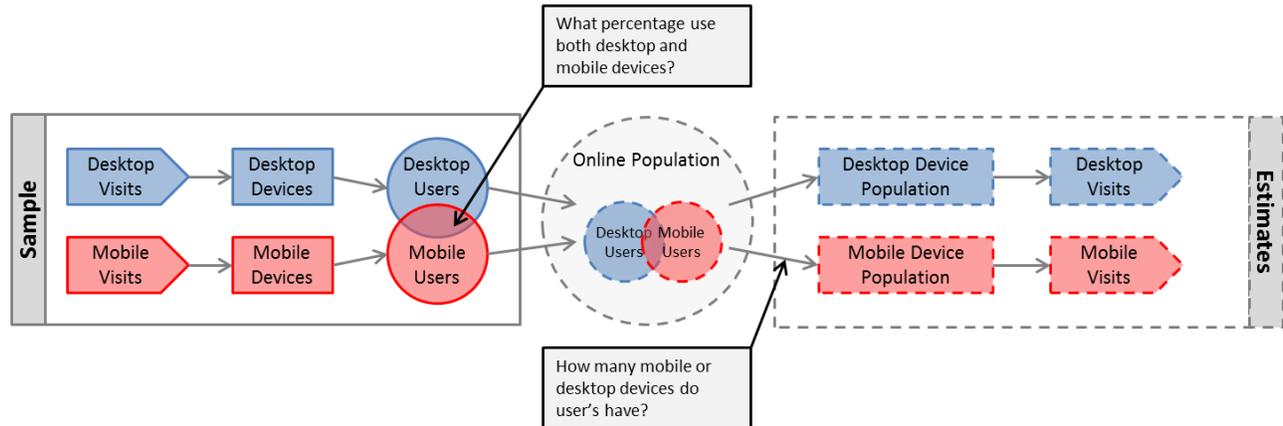
3.4 Estimating Persons / Unique Users

Estimation of Unique Users (UU's, previously termed Unique Visitors), is based on a projected count of panelists in our desktop and mobile samples. When an audience segment is defined, the portion of the sample meeting the segment's definition are identified and counted. That count is then weighted and projected to estimate the population value. The weighting and projection processes for Unique Users account for various factors:

- **Behaviors to people:** the Hitwise sample tracks behaviors that are performed on devices such as smartphones or desktop computers. The behaviors must be translated back through those devices to the individuals performing those actions
- **Desktop-Mobile overlap:** some proportion of people use both desktop and mobile devices, and that overlap is estimated for each website to ensure that we don't double count desktop and mobile users.
- **Multiple devices:** individuals often own multiple mobile and/or multiple desktop devices

- **Weights and scaling:** our sample must be weighted and scaled to reflect the true distribution and size of the population

The following figure sets out the logic of the sample and the modeling for Unique Users by device.



Estimates of unique users visiting a website, using a search term or doing any other measured behavior can be made by:

1. Defining the segment based on the digital behaviors in the sample seen in both the mobile and desktop samples
2. Netting the unique mobile and desktop devices and the unique mobile and desktop users
3. Adjusting the count by the estimated overlap
4. Applying the demographic weighting and population projection to scale up to the population of users

The estimates made by Hitwise of the device overlap together with the scaling and weighting factors (and the numbers of other devices held by users) are based on standards made available by industry and government sources. Examples of these sources include:

- United Kingdom
 - Office of National Statistics (ONS)
 - Ofcom (independent regulatory agency for communications)
- United States
 - US Census Bureau, American Community Survey
- Australia
 - Australian Bureau of Statistics (ABS)

4.0 Understanding Reported Metrics

AudienceView delivers metrics through a number of reports. These reports tell us who consumers are, how they feel about categories of interest, how they behave, their brand preferences, what they do online and how they access the Internet. AudienceView groups these reports into several report types: Profile Snapshots, Digital Snapshots, Cross-tabs, and Digital Behavior Analysis.

The Profile reports tell us who consumers are – their age, gender, household income, socio-economic groupings and more. These reports, along with the cross-tab reports, primarily report metrics related to the size and composition of the audience in terms of unique users or persons, the percentage of persons and an index of that percentage in relation to the total population of analysis. Digital Behavior reports not only report on the size of the audience, but also on audience behavior, including the benchmarking of online site visitation and search.

A description of the key metrics available in these reports is provided below.

4.1 Unique Users / Persons

The Unique Users metric is an estimate quantifying the number of projected persons for a target audience segment. Unique Users or projected persons are reported in AudienceView as:

- **Unique Users:** In Digital Behavior Reports, the Unique Users metric represents the estimated number of different online adult consumers in the target audience segment that visited a site, a site category (industry) or that used a reported search term. Individuals who may have visited a site (industry, etc.) on different devices are accounted for. This metric is expected to be turned on in AudienceView to clients in mid-2016.
- **Projected Persons:** In Snapshot Profile Reports or in Crosstab Reports, the Projected Persons metric represents the estimated size of the target audience segment in terms of total adult consumers.
- **Index:** The Index metric aims to provide a sense of the target audience segment's likelihood to exhibit a reported behavior or have a reported consumer characteristic in relation to the population universe ("base") being analyzed. Index examines the percentage of a target audience segment vs. the percentage of the base population and is reported as:
 - 100 (same as the base)
 - >100 (more likely than in the base)
 - <100 (less likely than in the base)

For an understanding of how Unique Persons are estimated and projected from the AudienceView sample, please see "Estimating Persons / Unique Users" in section 3.0 of this document.

4.2 Visits

The Visits metric is a behavioral estimate quantifying a target audience segment's visitation activity to web sites or to web site categories (also referred to as an "industries"). Visits are reported to web sites and web site categories as:

- **Total Visits:** The estimated sum total of visits made to a site or site category by a target audience segment within the time period specified for the analysis
- **Visit Share:** The percentage of a target audience segment's total visits to all sites within a category that were made to a specific site or sub-category of sites
- **Visit Share by Device:** The percentage of a target audience segment's total visits to all sites within a category that were made to a specific site or sub-category of sites while on either a desktop device or mobile device
- **Device Rate:** The proportion of visits to any one web site or web site category the target audience made via a desktop device versus a mobile device

Note that Visits as estimated in AudienceView may differ slightly from Visits reported in other services or products, such as site-centric reporting tools. In AudienceView, a site Visit can span the user being active on other pages, such as when a user is viewing multiples sites on a tabbed browser (the exception to this occurs when the two episodes of activity on a site span a period of inactivity greater than 30 minutes or when the episodes span across a date change). Conversely, in many site-centric reporting tools, a site Visit punctuated by activity on another site will often be treated as a separate Visit. AudienceView aims to reflect the experience of the user. Interspersed activity, where the user may launch a new browser tab and view a different site temporarily and then return to viewing the original site, would generally not be experienced as two separate visit events; it might be experienced as temporarily suspended in the interim, but as a single visit among other visits.

4.3 Search

The Search metric is a behavioral estimate quantifying a target audience segment's use of search and clicks on search results. The power of Search is in this data's ability to reveal a target audience segment's intent when seeking content online and what the segment's intent was when visiting selected sites or industries. Search is reported for specific search terms (keywords) in AudienceView as:

- **Search Share:** For a specific search term, the percentage of a target audience segment's total search clicks made during the period of analysis that the specific search term represents
- **Paid Clicks:** The percentage of clicks the specific search term represents of all paid search clicks performed by the target audience segment during the period of analysis

- **Organic Clicks:** The percentage of clicks the specific search term represents of all organic search clicks performed by the target audience segment during the period of analysis
- **Paid Rate:** The proportion of all clicks the target audience segment performed against paid results for the specific search term
- **Organic Rate:** The proportion of all clicks the target audience segment performed against organic results for the specific search term