

# **PMB / COMSCORE FUSED DATABASE**

## **USER MANUAL**



comSCORE.

# **PMB / COMSCORE FUSED DATABASE**

## **USER GUIDE**

The PMB / comScore fused data base provides new information for media buyers and sellers, linking the product and media usage data in PMB to the Internet viewing habits measured by comScore.

Website owners will find new information about their viewers, the products and services they use, their attitudes and the other media they are exposed to.

For the publications and other media reported in PMB, there will be new information on the Internet habits of their readers and viewers.

For Advertising Agencies and their client Advertisers, the fused database provides a single source for information on all media. It shows in unparalleled depth the media exposure patterns for users of advertiser's own product, and also media patterns for competitive brand users.

### **INFORMATION CONTAINED IN THE FUSED DATABASE**

- **READERSHIP** - for all major Canadian publications, using the standard readership and qualitative metrics, as regularly reported in PMB surveys.
- **WEBSITE VIEWING** – UV's and Page views for over 1,000 major websites, as regularly reported in comScore.
- **MEDIA USAGE** - General information on usage of other media including TV channels and programs, Radio, Community Newspapers, and others.
- **DEMOGRAPHICS** - Comprehensive list of demographics.
- **PRODUCTS / BRANDS** - Comprehensive list of products and brands used by Canadians.
- **RETAIL** - Consumers' patronage of over 350 major retail stores in Canada.
- **ATTITUDES** - Consumers' social views, health and environmental views and attitudes in 10 different product areas

The database is created by a fusion process that matches PMB respondents with comScore panelists based on their common demographic characteristics. The database provides an estimate of website viewers' usage of products and services, and their exposure to other media. The estimates are based on demographic characteristics and their accuracy is limited according to the effectiveness of the demographics at predicting consumers' behaviour.

## TABLE OF CONTENTS

Section		Page
I	Data Sources / Sample sizes	4
	Bases – 18+, Internet usage, Priority to comScore Internet usage level	
II	Demographics – Control variables, Use of PMB demographic information, Improved Income information	5
III	Computer Access <ul style="list-style-type: none"> <li>• Data dictionary</li> <li>• What can / can't be done</li> <li>• Reach / Frequency</li> <li>• Magazine / Website unduplicated audience</li> </ul>	7
IV	How to produce and analyze a crosstab	9
V	Minimum reporting standards, asterisking small samples	10
VI	Reporting and Access policies	11
VII	comScore Definitions	13

## SECTION 1

### DATA SOURCES / SAMPLE SIZES

- PMB sample size of approximately 22,000 respondents aged 18+, based on 24 months of interviewing.
- comScore monthly UV and Page Views averaged across three months with sample size of approximately 65,000 panelists aged 18+.
- The two databases combine to produce a fused database with an apparent sample size of 67,000 adults 18+.

### BASES

The fused database contains media and product information for Canadian adults aged 18 years and over. Although the tabulation software permits thousands of different subgroups as a base, two bases are given priority position in the data dictionary:

- *All 18+*. This corresponds to the Adults 18+ base in PMB.
- *18+ Internet Users*. This corresponds to the comScore 18+ Internet monthly user base.

### Internet User base :

The number of 18+ Internet users in the fused database is determined by the comScore file, since it is the more recent reading of Internet usage than the PMB measure. PMB users will notice minor difference in Internet usage level.

## **SECTION II DEMOGRAPHICS**

Congruent fusion preserves original data within controlled demographic cells. Analysis of PMB or comScore data using these cell boundaries will replicate very closely the data available in other PMB and comScore reports. (Small variations exist due to rounding and the need to match PMB Internet usage levels to comScore).

### **Control Variables**

For this fusion, control cells were based on 5 regions x 4 age groups x 2 genders, resulting in 40 cells overall :

- *Five region cells:* Atlantic, Quebec, Ontario, Prairies, BC
- *Four age cells:* 18-24, 25-34, 35-54, 55+
- *Gender:* Male, Female

Congruent fusion does not control for boundaries within a cell. For example Manitoba, Saskatchewan and Alberta are treated as a unit. Likewise, the age 55+ group is treated as a unit, although in PMB there are age breaks available at 65 and 70, and in comScore at 65.

### **Target Demographics**

Within the above control cells, the fusion process selects the best available match using information common to both PMB and comScore. These common variables include the following:

- Household Income
- Household Size
- Presence of Children

These are not control variables. So, for example, a comScore panelist in a one person household will most likely be fused with a PMB respondent in a one person household, but may be matched with someone in a two person household and vice versa.

### **PMB Demographics**

The demographics reported on the fused database are in all cases the demographics of the PMB respondent to whom the comScore panelist has been fused. These demographics include the full PMB list including:

- *Control variables:* Age, gender, region
- *Target demographics:* Household income, Household size, Presence of children
- *Other demographics:* City and community size, Employment and Occupation, Education, Personal income, Age of children, Family structure, Grand children, Type of housing, Ethnic background, Sexual orientation

PMB characteristics have been preserved through fusion. Since the fused database uses PMB demographics, comScore users will note some small variances vs. original data. This results from the fact that target demographics are not always exactly matched, and an analysis that uses different age/sex/region boundaries may vary to some extent.

## **Improved Income Reporting**

Merging the PMB and comScore Media Metrix databases created the opportunity to improve the household income data in the fused product.

The fused database uses PMB income data, which has been weighted to match Statistics Canada census data for the total Canadian population. This resulted in some changes to the household income breaks, notably for instance the fact that the over \$100,000 grouping is now larger and is in line with the PMB data. This in turn changes the web entity audience composition compared to its original comScore income distribution.

comScore and PMB are pleased to announce that this is one of the first improvements accruing to the measurement of the digital universe from the joint partnership.

## SECTION III

### COMPUTER ACCESS

The fused database is available via PMB's licensed software providers, IMS/Nielsen and Telmar/Harris. The objective of the fusion is to permit cross tabulations of comScore web entities vs. PMB data including publication readership, demographics, product usage and attitudes.

#### **Data dictionary:**

Selecting a comScore web entity for analysis involves a selection process familiar to PMB users. The data dictionary includes a folder: "comScore". Within this folder are UV's (Unique Visitors) and PV's (Page Views) for website categories "Top Entities" as well as individual websites "All Entities".

#### **IMPORTANT**

The software packages have a wide range of functionalities, however, some are not available when a cross tabulation involves a comScore reported website entity.

#### **Permitted Cross tabulations**

When a cross tab involves a web entity, the following types of cross tabulations are permitted :



OK

- Recommended pattern is web entities in columns, one entity per column.
- Rows may contain any PMB data, excepting averages.
- If a web entity is selected as a base; no web entities in rows, none in columns

An example of a crosstabulation is shown in next Section .

#### **Cross tabulations that do NOT work.**



NO

- Cross tabulations that have web entities in columns AND in rows will not work.
- Cross tabulations that have web entities in the base and in a row or in a column will not work.
- Using logical operators (AND, OR) will not work with web entities
- Cross tabulations where web entities are crossed with averages, medians or total volume for PMB measured products or media.

The software licensees have a warning in the software to alert users to these limitations.

(There are two reasons for these limitations, either of which would be sufficient on its own. First, the web entity information is of a different data type from the regular PMB data. It is an average. Second, to carry forward millions of comScore duplication estimates onto the fused database would not be manageable from a data production standpoint.)

**Reach / Frequency** – not available. At this point in time there are many unresolved issues in estimating combined Internet / magazine reach frequency, not the least of which are the differing definitions of frequency between magazines and the Internet. The PMB Research

Committee and comScore Advisory Council will review this aspect in advance of the next fused database release.

### **How to obtain the net audience of print and web properties**

Each of the two software suppliers has a customized procedure for providing duplicated and unduplicated readership between a measured print vehicle and a measured website. Contact your supplier for further information.

## SECTION IV HOW TO PRODUCE AND ANALYZE A CROSTAB

This procedure will provide information on how many website viewers have characteristics measured by PMB:

**STEP 1 - Define the base :** The default base is “All adult Canadians”. Or, you can select the Digital Universe aged 18+, or any PMB data item.  
(You could select a comScore web entity, but that would impose restrictions as explained in Section III)

**STEP 2 – Define the columns:** We recommend you select Unique Visitors for comScore web entities. Each entity in its own column.

**STEP 3 – Define the rows :** Put PMB data in the rows. You may select from almost all available PMB data, including demographics, media usage, product usage, attitudes and lifestyle. One restriction - no averages.

PMB DATA	Web entities . . .	
	Website entity A UV's*	Website entity B UV's*
Demographics		000
Media usage		Unweighted
Lifestyle		V%
Product usage		H%
Attitudes		Index

**HOW TO READ THE CROSS TAB :** The output is a table. Each cell in the table contains statistics describing the people who are viewers of the web entity and have the characteristic defined in the row :

- 000 – number of web entity viewers (in thousands) who also have the row characteristic.
- Unweighted – sum of comScore panellists and PMB respondents in this cell
- V% - number in this cell as % of all entity B viewers.
- H% - number in this cell as % of all Canadians in the selected base with the PMB characteristic.
- Index – Measure of whether the number of viewers with this characteristic is above or below average.

\*Cross tabs with Page views work the same way, except that H% has no meaning.

## SECTION V

### MINIMUM REPORTING STANDARDS / ASTERISKING

The minimum reporting level for web entities in the fused database is 50,000 average monthly Unique Visitors for the three month period used in the study.

#### **Asterisking Small Samples**

The software access companies use an asterisking system approved by comScore and PMB to indicate data that may be unstable or unreliable due to small sample sizes:

- Single Asterisk (\*) indicates data is unstable because it is based on a small sample. This occurs when the unweighted count of split weight respondents is between 50 and 100.
- Pound (#) indicates data is unreliable and is reported only for consistency. This occurs when the unweighted count of split weight respondents is less than 50.

The cut off points (50 and 100) for the unweighted number of split weight respondents correspond to the cut off points historically used in PMB (20 and 40 respectively).

## SECTION VI

### REPORTING AND ACCESS POLICY

PMB and comScore have different data collection methods and reporting policies. Policies for this fused database reflect the interests of the two partners.

#### Data Collection

- comScore collects data for all websites accessed by Canadian consumers.
- PMB collects media data only for media that subscribe to the service. Full members are entitled to all data collected. Some restrictions apply to associate members of PMB in line with normal PMB access guidelines.

#### Fused Database Reporting

Decisions about including or excluding individual web entities on the fused database reflect the interests of both PMB members and comScore’s commercial interests.

- Web entities under 50,000 UV’s are excluded from the file: This relates to minimum reporting standards as explained in Section V.
- Web entities in the range 50,000 to 500,000 monthly UV’s : These are included on the file. However most are blocked unless the entity subscribes to both PMB and comScore. This is done in order to keep the data dictionary at a manageable size. Blocking is not applied to web entities considered to be of general interest to the membership.
- Web entities with 500,000 or more monthly UV’s : These are on the file. However, some are blocked unless they subscribe to both PMB and comScore. This is done in order to protect PMB members’ interests.

The above principles are summarized in the chart below:

**WEB ENTITIES ON FILE / BLOCKED / RELEASED (monthly UV’s)**

Under 50,000	50,000 – 500,000	500,000 +
Not on the file (below minimum reporting standards)	On file	On file
	Generally blocked (to keep data dictionary at a manageable size)  Exceptions : <ul style="list-style-type: none"> <li>• Released if subscribe to PMB and comScore</li> <li>• Released if of general interest to membership</li> </ul>	Generally released  Exceptions : Some blocked unless they subscribe to PMB and comScore to protect PMB members’ interests.

The full list of approximately 1,000 websites is shown on [www.PMB.ca](http://www.PMB.ca).

### **What if a blocked web entity subscribes after release?**

PMB and comScore have approved a system with the licensed software suppliers to override the blocking after the official release of the fused database.

#### **Access rights**

Access to the fused database requires both subscription to comScore, and membership of PMB:

- i. PMB full members who also subscribe to comScore Canada have access to all reported data, except brand data (see point [iii] below).
- ii. PMB Associate members who subscribe to the fused database, and also subscribe to comScore Canada have access to most of the fused data, but, in line with normal PMB access policy, they are restricted from accessing individual print titles and their web entities. Contact PMB for more information.
- iii. Brand usage data on the fused database are accessible only to brand sponsor(s), and their designated agency or agencies.

## SECTION VII

### COMSCORE DEFINITIONS

**WEB ENTITY:** comScore reports audiences for websites (identified by the URL) and for groups of websites known as web entities.

**PROPERTY:** The highest level of reporting in the Client Focus structure, Properties represent all Full Domains (i.e. felmont.com), Pages (i.e. sports.felmont.com/tennis), Applications or Online Services under common ownership or majority ownership for a single legal entity or majority ownership for a single legal entity.

A Property may also contain any digital media content that is not majority owned but has been legally signed over for reporting purposes by the majority owner.

**MEDIA TITLE:** An editorially and brand consistent collection of content in the digital landscape. This may represent a domain, a group of domains, online service or application.

**CHANNEL, SUBCHANNEL, GROUP, and SUBGROUP:** Many of the larger Media Titles contain URLs of editorially consistent content where the URLs together make up an entire Channel. The majority owner of a Media Title may be able to break out Channel-level detail.

For some of the largest Media Titles, Channels themselves may be broad, and Subchannels, Groups and Subgroups within the larger Channels may prove useful. Subchannels, Groups, and Subgroups are subject to the same rules as Channels.

**NETWORK:** A company that syndicates its content on other websites. The components of a network are the sites and pages where the content is syndicated.

**Ad Networks:** A company that sells advertising on a group of domains primarily for third party owned websites.

**AD NETWORK CHANNEL:** Similar to the relationship of a Channel to a Media Title, Ad Network channels are subsets of the Ad Network that usually include websites and/or web pages with similar content or that attract similar type of people.