comScore Disaggregate Data User Guide

PURPOSE

The comScore Disaggregate Data offers web-wide visitation and transaction behavior based on a random sample from a cross-section of more than 2 million Internet users in the United States who have given comScore explicit permission to confidentially capture their Web-wide activity. This includes panelist-level browsing and buying behavior from tens of thousands of Internet users.

CREATING A REPORT

A. Demographics module

Step 1 – How would you like to search this dataset?

Step 1: How would you like to search this dataset?

What format are your ID codes?

- Machine ID

How does this work?

- Manually enter ID codes

ID Codes

Please enter ID codes separated by a space.

Example: 6365661 11478354 51815826

Save selected code list to myWRDS

Code List Name

- Upload file containing ID codes
- Retrieve saved codes from myWRDS
- Search entire database

Type URLs here (without http://www.)
Step 2 – What variables do you want in your query?

Step 3 – How would you like the query output?

Output format:
- fixed-width text (*.txt)
- comma-delimited text (*.csv)
- XML Excel Spreadsheet (*.xls)
- tab-delimited text (*.txt)
- HTML table (*.html)
- SAS Windows_32 dataset (*.sas7bdat)
- SAS Solaris_64 dataset (*.sas7bdat)
- dBASE file (*.dbf)
- STATA file (*.dta)
- SPSS file (*.sav)

E-Mail Address (Optional)

Save this query to myWRDS
Query Name

SUBMIT QUERY

Compression Type:
- None
- zlib (*.zlib)
- gzip (*.gz)

Recommended
B. Sessions module

Step 1 - What date range do you want to use?

Step 1: What date range do you want to use?
I would like data from Jan 1, 2009 to Jan 5, 2009.

Step 2 - How would you like to search this dataset?

Step 2: How would you like to search this dataset?

What format are your ID codes?
- Domain Name (Recommended)
- Domain ID
- Machine ID
- Site Session ID

How does this work?

- Manually enter ID codes

ID Codes [ Code Lookup ]

Please enter ID codes separated by a space.
Example: amazon.com ebay.com msn.com

- Save selected code list to myWRDS

- Upload file containing ID codes
- Retrieve saved codes from myWRDS
- Search entire database

Enter URLs here (without http://www.)
Step 3 - What variables do you want in your query?

**Step 3: What variables do you want in your query?**

How does this work?

**Household Information (0 of 10 selected)**

Select the items you would like to include in your search. Corresponding help links are available for more information on selected codes.

- [ ] Household Highest Education
- [ ] Household Oldest Age
- [ ] Household Income
- [ ] Household Size
- [ ] Racial Background
- [ ] Country of Origin
- [ ] Children
- [ ] Connection Speed
- [ ] Census Region
- [ ] Zip Code

Selected Items

Check All | Uncheck All

Demographic data point selections

**Transaction Information (0 of 5 selected)**

Select the items you would like to include in your search. Corresponding help links are available for more information on selected codes.

- [ ] Product Name
- [ ] Product Category ID
- [ ] Product Quantity
- [ ] Product Total Price
- [ ] Basket Total

Selected Items

Check All | Uncheck All

Are you analyzing eCommerce data?
Step 4 – How would you like the query output?

**How do you want to sort the variables?**

Select the desired format of the output file. For large data requests, select a compression type to expedite downloads. If you enter your email address, you will receive an email that contains a URL to the output file when the data request is finished processing.

**Compression Type**
- None
- zip (*.zip)
- gzip (*.gz)

**Date Format**
- YMDmDn8. (e.g. 19840725)
- DATE9. (e.g. 25JUL1984)
- DDMYY6. (e.g. 250784)
- MMDYY10. (e.g. 07/25/1984)
- DDMYyY10. (e.g. 07/25/1984)
- YYYMDMDS10. (e.g. 1984/07/25)

**Output format**
- fixed-width text (*.txt)
- comma-delimited text (*.csv)
- XML Excel Spreadsheet (*.xls)
- tab-delimited text (*.txt)
- HTML table (*.html)
- SAS Windows_32 dataset (*.sas7bdat)
- SAS Solaris_64 dataset (*.sas7bdat)
- dBase file (*.dbf)
- STATA file (*.dta)
- SPSS file (*.sav)

**E-Mail Address (Optional)**

**Save this query to myWRDS**

**Query Name**

**SUBMIT QUERY**
C. Transactions module

Step 1 – What date range do you want to use?

Step 1: What date range do you want to use?
I would like data from Jan 1 2009 to Mar 31 2009

Step 2 – How would you like to search this dataset?

Step 2: How would you like to search this dataset?
What format are your ID codes?
- Domain Name
- Domain ID
- Machine ID
- Site Session ID

How does this work?

- Manually enter ID codes

ID Codes [Code Lookup]
Enter URLs here (without http://www.)

- Upload file containing ID codes
- Retrieve saved codes from myWRDS
- Search entire database

Select Product Categories (optional)
---All Products---
1. Apparel
2. Shoes
3. Accessories
4. Jewelry & Watches
5. Other Apparel Items
6. Home Furniture
Step 3 – What variables do you want in your query?

**Household Information (0 of 10 selected)**

Select the items you would like to include in your search. Corresponding help links are available for more information on selected codes.

- Household Highest Education
- Household Oldest Age
- Household Income
- Household Size
- Racial Background
- Country of Origin
- Children
- Connection Speed
- Census Region
- Zip Code

**Selected Items**

Select demographic information

**Transaction Information (3 of 8 selected)**

Select the items you would like to include in your search. Corresponding help links are available for more information on selected codes.

- Event Time
- Product Name
- Product Category ID
- Domain Name
- Domain ID
- Product Quantity
- Product Total Price
- Basket Total

**Selected Items**

Select transaction information to be included in report.
Step 4 – How would you like the query output?

Please select the variables by which you would like your results sorted.

1. **machine_id**
2. **site_session_id**
3. **None (default ordering)**

Select the desired format of the output file. For large data requests, select a compression type to expedite downloads. If you enter your email address, you will receive an email that contains a URL to the output file when the data request is finished processing.

Output format:
- fixed-width text (*.txt)
- comma-delimited text (*.csv)
- XML Excel Spreadsheet (*.xls)
- tab-delimited text (*.txt)
- HTML table (*.htm)
- SAS Windows_32 dataset (*.sas7dat)
- SAS Solaris_64 dataset (*.sas7bdat)
- dBase file (*.dbf)
- STATA file (*.dta)
- SPSS file (*.sav)

Compression Type:
- None
- zip (*.zip)
- gzip (*.gz)

Date Format:
- YYMDDn8. (e.g. 19840725)
- DATE9. (e.g. 25JUL1984)
- DDMMYY6. (e.g. 250784)
- MMDDYY10. (e.g. 07/25/1984)
- DDMMYY10. (e.g. 25/07/1984)
- YYMDDs10. (e.g. 1984/07/25)

E-Mail Address (Optional)

☐ Save this query to myWRDS

Query Name

Save query for easy web-access.

Submit Query
## comScore Web Behavior -- 2009 Transactions

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Data Type</th>
<th>Variable Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASKET_TOT</td>
<td>Data</td>
<td>Basket Total</td>
</tr>
<tr>
<td>CENSUS_REGION</td>
<td>Data</td>
<td>Census Region</td>
</tr>
<tr>
<td>CHILDREN</td>
<td>Data</td>
<td>Children</td>
</tr>
<tr>
<td>CONNECTION_SPEED</td>
<td>Data</td>
<td>Connection Speed</td>
</tr>
<tr>
<td>COUNTRY_OF_ORIGIN</td>
<td>Data</td>
<td>Country of Origin</td>
</tr>
<tr>
<td>DOMAIN_ID</td>
<td>Data</td>
<td>Domain ID</td>
</tr>
<tr>
<td>DOMAIN_NAME</td>
<td>Data</td>
<td>Domain Name</td>
</tr>
<tr>
<td>EVENT_DATE</td>
<td>Event</td>
<td>event_date</td>
</tr>
<tr>
<td>EVENT_TIME</td>
<td>Event</td>
<td>Event Time</td>
</tr>
<tr>
<td>HOH_MOST_EDUCATION</td>
<td>Data</td>
<td>Household Highest Education</td>
</tr>
<tr>
<td>HOH_OLEDEST_AGE</td>
<td>Data</td>
<td>Household Oldest Age</td>
</tr>
<tr>
<td>HOUSEHOLD_INCOME</td>
<td>Data</td>
<td>Household Income</td>
</tr>
<tr>
<td>HOUSEHOLD_SIZE</td>
<td>Data</td>
<td>Household Size</td>
</tr>
<tr>
<td>MACHINE_ID</td>
<td>Data</td>
<td>machine_id</td>
</tr>
<tr>
<td>PROD_CATEGORY_ID</td>
<td>Data</td>
<td>Product Category ID</td>
</tr>
<tr>
<td>PROD_NAME</td>
<td>Data</td>
<td>Product Name</td>
</tr>
<tr>
<td>PROD_QTY</td>
<td>Data</td>
<td>Product Quantity</td>
</tr>
<tr>
<td>PROD_TOTPRICE</td>
<td>Data</td>
<td>Product Total Price</td>
</tr>
<tr>
<td>RACIAL_BACKGROUND</td>
<td>Data</td>
<td>Racial Background</td>
</tr>
<tr>
<td>SITE_SESSION_ID</td>
<td>Data</td>
<td>site_session_id</td>
</tr>
<tr>
<td>ZIP_CODE</td>
<td>Data</td>
<td>Zip Code</td>
</tr>
</tbody>
</table>
Appendix B – Frequently Asked Questions

Q: The comScore dataset on WRDS has "visited URL" and "referral URL" fields but they show only the name of the site (i.e. google.com). Does comScore log the complete request string such as http://rds.yahoo.com/S=2766679/K=laptop/y=2/SID=e/TID=BVTC_11/

A: Complete URLs are not given for privacy and security reasons.

Q: How can I download a couple months of session data from ComScore? I'm trying to replicate on a small scale what Moe and Fader (2004) did in their article "Dynamic Conversion Behavior at E-Commerce Sites."

A: The session data from ComScore data is very big. CSV format will open in Excel if the file is small enough, otherwise you might need a more robust statistical package such as SAS, Stata, Matlab, etc.

Q: What time zone is used for the EVENT_TIME variable in Comscore?

A: The event_time data in Comscore uses Greenwich Mean Time, as per a phone call to Comscore Support (2/5/2009).