comScore Web Behavior Database

The comScore Web Behavior Database captures detailed browsing and buying behavior by 100,000 Internet users across the United States at the domain level. The panel is based on a random sample from a cross-section of more than 2 million global Internet users who have given comScore explicit permission to confidentially capture their Web-wide activity.

The unique panel identifier is Machine ID and there are cases where multiple configured machines exist within a household, but all demographic information is based upon the associated household. All sessions are aggregated by machine in the household, so that individual breakdowns are not available and a particular individual could use more than one machine. The items below are associated with each machine-household.

Demographics

Variables	Type	Label/Description
machine_id	num	machine identifier
hoh_most_education	num	most education – head of household
census region	num	census region
household size	num	household size
hoh_oldest_age	num	oldest age – head of household
household_income	num	household income
children	num	presence of children
racial_background	num	racial background
connection_speed	num	connection speed
country_of_origin	num	country of origin
zip_code	num	zip code

A product-transaction table shows online purchases if and only if there was a transaction in the session. The following items are records for each transaction.

Transaction Information

Variable	Type	Label/Description
prod_category_id	num	product category ID
prod_num	char	product name
prod_qty	num	product quantity
prod_totprice	num	product total price
basket_tot	num	basket total

User sessions are recorded with date and time stamps and the most detailed clickstream item is a Session ID, which represents a domain visit that may have more than one 'page view.' The items below are associated with each Session ID.

Session Information

Variable	Type	Label/Information
user_session_id	num	identifies a session of activity
domain_id	num	domain ID
ref_domain_name	char	referring domain name
pages_viewed	num	pages viewed
duration	num	duration at site
event_date	char	date of activity
event_time	char	time of activity

The items in the tables above are linked in a WRDS web query using a four step process:
1. The demographics table is linked to session table using 'machine_id'
2. The domain name lookup table is joined to the traffic and transaction tables on 'domain_id'
3. The session table is linked to the transaction table using 'site_session_id' ('machine_id' is also provided in the transaction table for convenience, but not needed to link the tables)
4. The product category lookup table is linked with the transaction table using 'prod_category_id'

NOTE: All .txt files are tab delimited

Most Educated Head of Household

0	Less than a high school diploma
1	High school diploma or equivalent
2	Some college but no degree
3	Associate degree
4	Bachelor's degree
5	Graduate degree
99	Missing

Household Income

1	Less than 15k
2	15k-24.999k
3	25k-34.999k
4	35k-49.999k
5	50k-74.999k
6	75k-99.999k
7	100k+

Age of Eldest Head of Household and Age of User

1	18-20
2	21-24
3	25-29
4	30-34
5	35-39
6	40-44
7	45-49
8	50-54
9	55-59
10	60-64
11	65 and over

Household Size

1	1
2	2
3	3
4	4
5	5
6	6+

Racial Background

1	White
2	Black
3	Asian
5	Other

Connection Speed

0	Not broadband
1	Broadband

Census Region of Residence

1	Northeast
2	North Central
3	South
4	West

Country of Origin

1	Hispanic
0	Non-Hispanic

Child Present

0	No
1	Yes

Product Category
1. Apparel
2. Shoes
3. Accessories
4. Jewelry & Watches
5. Other Apparel Items
6. Home Furniture
7. Home Appliances
8. Tools & Equipment
9. Kitchen & Dining
10. Bed & Bath
11. Garden & Patio
12. Pet Supplies
13. Food & Beverage
14. Automotive Accessories
15. Sport & Fitness
16. Health & Beauty
17. Art & Collectibles
18. Tobacco Products
19. Baby Supplies
20. Other Home & Living Items
21. Books & Magazines
22. Music
23. Movies & Videos
24. Other BMV
25. Desktop Computers
26. Laptop Computers
27. Handhelds, PDAs & Portable Devices
28. Printers, Monitors & Peripherals
29. Computer Software (excl. PC Games)
30. Other Computer Supplies
31. Audio & Video Equipment
32. Cameras & Equipment
33. Mobile Phones & Plans
34. Other Electronics & Supplies
35. PC Video Games
36. Console Video Games
37. Video Game Consoles & Accessories
38. Business Machines
39. Office Furniture
40. Office Supplies
41. Movie Tickets
42. Event Tickets
43. Air Travel
44. Hotel Reservations
45. Car Rental
46. Travel Packages
47. Other Travel
48. Online Content Sales
49. Online Service Subscriptions
50. Personals & Dating
51. Photo Printing Services
52. Shipping Services
53. Other Services
54. Toys & Games (excl. Video Games)
55. Arts, Crafts & Party Supplies

56. Other Toy & Game Items
57. Flowers
58. Greetings
59. Gift Certificates & Coupons
60. Other Flower & Gift Items
99. Unclassified

Terms of Use/Publication Guidelines:

➢ Researchers are required to submit to comScore a research abstract outlining the project scope and use of comScore data. Abstracts are to be submitted prior to the start of any new project that utilizes comScore data.
➢ All research and usage of comScore data must be submitted to comScore prior to publication; comScore reserves the right of review and approval.
➢ All usage must include proper attribution: "Source: comScore, Inc."
➢ comScore reserves the right to use excerpts of the institution's research for marketing, publicity and other purposes.
➢ Any public disclosure of the comScore data or materials or their derivatives shall <u>not</u> include the following:
   o Sales (transaction) data at the individual web site, or category level
   o Visitation (traffic) data for named web sites
   o Visitation (traffic) or sales (transaction) data for a group of sites described in such a way as to infer or imply the names of web sites included in the group.

Contact:
Please direct all research abstracts and questions regarding the comScore dataset to:

Eric Young
comScore, Inc.
eyoung@comscore.com
312-777-6610