

comScore Web Behavior Database

The comScore Web Behavior Database is based on a sample of about 50,000 US Internet users who have given comScore explicit permission to confidentially capture their detailed browsing and buying behavior at the domain level.

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+

Connection Speed

0	Not broadband
1	Broadband

Racial Background

1	White
2	Black
3	Asian
5	Other

Country of Origin

1	Hispanic
0	Non-Hispanic

Census Region of Residence

1	Northeast
2	North Central
3	South
4	West

Child Present

0	No
1	Yes

Product Category

- | | |
|--|---------------------------------------|
| 1. Apparel | 31. Audio & Video Equipment |
| 2. Shoes | 32. Cameras & Equipment |
| 3. Accessories | 33. Mobile Phones & Plans |
| 4. Jewelry & Watches | 34. Other Electronics & Supplies |
| 5. Other Apparel Items | 35. PC Video Games |
| 6. Home Furniture | 36. Console Video Games |
| 7. Home Appliances | 37. Video Game Consoles & Accessories |
| 8. Tools & Equipment | 38. Business Machines |
| 9. Kitchen & Dining | 39. Office Furniture |
| 10. Bed & Bath | 40. Office Supplies |
| 11. Garden & Patio | 41. Movie Tickets |
| 12. Pet Supplies | 42. Event Tickets |
| 13. Food & Beverage | 43. Air Travel |
| 14. Automotive Accessories | 44. Hotel Reservations |
| 15. Sport & Fitness | 45. Car Rental |
| 16. Health & Beauty | 46. Travel Packages |
| 17. Art & Collectibles | 47. Other Travel |
| 18. Tobacco Products | 48. Online Content Sales |
| 19. Baby Supplies | 49. Online Service Subscriptions |
| 20. Other Home & Living Items | 50. Personals & Dating |
| 21. Books & Magazines | 51. Photo Printing Services |
| 22. Music | 52. Shipping Services |
| 23. Movies & Videos | 53. Other Services |
| 24. Other BMV | 54. Toys & Games (excl. Video Games) |
| 25. Desktop Computers | 55. Arts, Crafts & Party Supplies |
| 26. Laptop Computers | 56. Other Toy & Game Items |
| 27. Handhelds, PDAs & Portable Devices | 57. Flowers |
| 28. Printers, Monitors & Peripherals | 58. Greetings |
| 29. Computer Software (excl. PC Games) | 59. Gift Certificates & Coupons |
| 30. Other Computer Supplies | 60. Other Flower & Gift Items |
| | 99. Unclassified |

Terms of Use/Publication Guidelines:

Use of this data implies acknowledgement that the comScore data in WRDS are a random sample of panelists, and do not represent actual visitations or sales or comScore's estimates thereof. Researchers are encouraged to submit to comScore a research abstract outlining the project scope and use of comScore data for the purpose of assuring that comScore data is used correctly.

- All usage must include proper attribution: "Source: comScore, Inc."
- Any public disclosure of the comScore data or materials or their derivatives shall not include the following:
 - Sales (transaction) data at the individual web site, or category level
 - Visitation (traffic) data for named web sites
 - 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:

Brian Kirchner
comScore, Inc.
bkirchner@comscore.com
312-775-6513