

By Jure Leskovec

STANFORD
UNIVERSITY



SNAP graph library
Large networks
What's new
People
Papers
Links
About

[License](#)

Web data: Amazon reviews

Dataset information

This dataset consists of reviews from [amazon](#). The data span a period of 18 years, including ~35 million reviews up to March 2013. Reviews include product and user information, ratings, and a plaintext review.

Dataset statistics

Number of reviews	34,686,770
Number of users	6,643,669
Number of products	2,441,053
Users with > 50 reviews	56,772
Median no. of words per review	82
Timespan	Jun 1995 - Mar 2013

Source (citation)

- J. McAuley and J. Leskovec. [Hidden factors and hidden topics: understanding rating dimensions with review text](#). RecSys, 2013.

Files

File	Description	Size
all.txt.gz	All product reviews (34,686,770 reviews)	11G
Amazon_Instant_Video.txt.gz	Amazon Instant Video reviews (717,651 reviews)	252M
Arts.txt.gz	Arts product reviews (27,980 reviews)	5.3M
Automotive.txt.gz	Automotive product reviews (188,728 reviews)	36M
Baby.txt.gz	Baby product reviews (184,887 reviews)	42M
Beauty.txt.gz	Beauty product reviews (252,056 reviews)	46M
Books.txt.gz	Book reviews (12,886,488 reviews)	4.4G
Cell_Phones_&_Accessories.txt.gz	Cell Phone reviews (78,930 reviews)	20M
Clothing_&_Accessories.txt.gz	Clothing reviews (581,933 reviews)	78M
Electronics.txt.gz	Electronics product reviews (1,241,778 reviews)	325M
Gourmet_Foods.txt.gz	Gourmet Food reviews (154,635 reviews)	30M
Health.txt.gz	Health product reviews (428,781 reviews)	87M
Home_&_Kitchen.txt.gz	Home & Kitchen product reviews (991,794 reviews)	210M
Industrial_&_Scientific.txt.gz	Industrial & Scientific product reviews (137,042 reviews)	13M
Jewelry.txt.gz	Jewelry reviews (58,621 reviews)	7.8M
Kindle_Store.txt.gz	Kindle Store reviews (160,793 reviews)	59M
Movies_&_TV.txt.gz	Movie & TV reviews (7,850,072 reviews)	2.8G
Musical_Instruments.txt.gz	Musical Instrument reviews (85,405 reviews)	20M
Music.txt.gz	Music reviews (6,396,350 reviews)	2.1G
Office_Products.txt.gz	Office product reviews (138,084 reviews)	30M
Patio.txt.gz	Patio product reviews (206,250 reviews)	45M
Pet_Supplies.txt.gz	Pet Supply reviews (217,170 reviews)	47M
Shoes.txt.gz	Shoe reviews (389,877 reviews)	51M
Software.txt.gz	Software reviews (95,084 reviews)	30M
Sports_&_Outdoors.txt.gz	Sports & Outdoor product reviews (510,991 reviews)	100M
Tools_&_Home_Improvement.txt.gz	Tools & Home Improvement product reviews (409,499 reviews)	90M
Toys_&_Games.txt.gz	Toy & Game reviews (435,996 reviews)	89M
Video_Games.txt.gz	Video Game reviews (463,669 reviews)	152M
Watches.txt.gz	Watch reviews (68,356 reviews)	15M
categories.txt.gz	Category information for all products	45M
titles.txt.gz	Titles for all products	61M

Data format

```
product/productId: B00006HAXW
review/userId: A1RSDE90N6RSZF
review/profileName: Joseph M. Kotow
review/helpfulness: 9/9
review/score: 5.0
review/time: 1042502400
review/summary: Pittsburgh - Home of the OLDIES
review/text: I have all of the doo wop DVD's and this one is as good or better than the
1st ones. Remember once these performers are gone, we'll never get to see them again.
Rhino did an excellent job and if you like or love doo wop and Rock n Roll you'll LOVE
this DVD !!
```

where

- product/productId: [asin](#), e.g. [amazon.com/dp/B00006HAXW](https://www.amazon.com/dp/B00006HAXW)
- review/userId: **id of the user**, e.g. [A1RSDE90N6RSZF](#)
- review/profileName: **name of the user**
- review/helpfulness: **fraction of users who found the review helpful**
- review/score: **rating of the product**
- review/time: **time of the review** (unix time)
- review/summary: **review summary**
- review/text: **text of the review**