Dataset used:

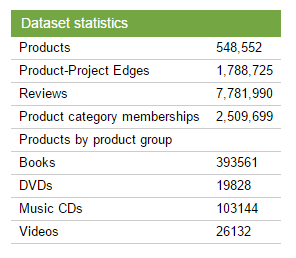
Amazon product co-purchasing network metadata.

The data was collected by crawling Amazon website and contains product metadata and review information about 548,552 different products (Books, music CDs, DVDs and VHS video tapes).

For each product the following information is available:

* Title
* Salesrank
* List of similar products (that get co-purchased with the current product)
* Detailed product categorization
* Product reviews: time, customer, rating, number of votes, number of people that found the review helpful.

The data was collected in summer 2006.



Data format:

* Id: Product id (number 0, ..., 548551)
* ASIN: Amazon Standard Identification Number
* title: Name/title of the product
* group: Product group (Book, DVD, Video or Music)
* salesrank: Amazon Salesrank
* similar: ASINs of co-purchased products (people who buy X also buy Y)
* categories: Location in product category hierarchy to which the product belongs (separated by |, category id in [])
* reviews: Product review information: time, user id, rating, total number of votes on the review, total number of helpfulness votes (how many people found the review to be helpful)