SUMMARY

================================================================================

These files contain 1,000,209 anonymous ratings of approximately 3,900 movies

made by 6,040 MovieLens users who joined MovieLens in 2000.

USAGE LICENSE

================================================================================

Neither the University of Minnesota nor any of the researchers

involved can guarantee the correctness of the data, its suitability

for any particular purpose, or the validity of results based on the

use of the data set. The data set may be used for any research

purposes under the following conditions:

\* The user may not state or imply any endorsement from the

University of Minnesota or the GroupLens Research Group.

\* The user must acknowledge the use of the data set in

publications resulting from the use of the data set

(see below for citation information).

\* The user may not redistribute the data without separate

permission.

\* The user may not use this information for any commercial or

revenue-bearing purposes without first obtaining permission

from a faculty member of the GroupLens Research Project at the

University of Minnesota.

If you have any further questions or comments, please contact GroupLens

<grouplens-info@cs.umn.edu>.

CITATION

================================================================================

To acknowledge use of the dataset in publications, please cite the following

paper:

F. Maxwell Harper and Joseph A. Konstan. 2015. The MovieLens Datasets: History

and Context. ACM Transactions on Interactive Intelligent Systems (TiiS) 5, 4,

Article 19 (December 2015), 19 pages. DOI=http://dx.doi.org/10.1145/2827872

ACKNOWLEDGEMENTS

================================================================================

Thanks to Shyong Lam and Jon Herlocker for cleaning up and generating the data

set.

FURTHER INFORMATION ABOUT THE GROUPLENS RESEARCH PROJECT

================================================================================

The GroupLens Research Project is a research group in the Department of

Computer Science and Engineering at the University of Minnesota. Members of

the GroupLens Research Project are involved in many research projects related

to the fields of information filtering, collaborative filtering, and

recommender systems. The project is lead by professors John Riedl and Joseph

Konstan. The project began to explore automated collaborative filtering in

1992, but is most well known for its world wide trial of an automated

collaborative filtering system for Usenet news in 1996. Since then the project

has expanded its scope to research overall information filtering solutions,

integrating in content-based methods as well as improving current collaborative

filtering technology.

Further information on the GroupLens Research project, including research

publications, can be found at the following web site:

http://www.grouplens.org/

GroupLens Research currently operates a movie recommender based on

collaborative filtering:

http://www.movielens.org/

RATINGS FILE DESCRIPTION

================================================================================

All ratings are contained in the file "ratings.dat" and are in the

following format:

UserID::MovieID::Rating::Timestamp

- UserIDs range between 1 and 6040

- MovieIDs range between 1 and 3952

- Ratings are made on a 5-star scale (whole-star ratings only)

- Timestamp is represented in seconds since the epoch as returned by time(2)

- Each user has at least 20 ratings

USERS FILE DESCRIPTION

================================================================================

User information is in the file "users.dat" and is in the following

format:

UserID::Gender::Age::Occupation::Zip-code

All demographic information is provided voluntarily by the users and is

not checked for accuracy. Only users who have provided some demographic

information are included in this data set.

- Gender is denoted by a "M" for male and "F" for female

- Age is chosen from the following ranges:

\* 1: "Under 18"

\* 18: "18-24"

\* 25: "25-34"

\* 35: "35-44"

\* 45: "45-49"

\* 50: "50-55"

\* 56: "56+"

- Occupation is chosen from the following choices:

\* 0: "other" or not specified

\* 1: "academic/educator"

\* 2: "artist"

\* 3: "clerical/admin"

\* 4: "college/grad student"

\* 5: "customer service"

\* 6: "doctor/health care"

\* 7: "executive/managerial"

\* 8: "farmer"

\* 9: "homemaker"

\* 10: "K-12 student"

\* 11: "lawyer"

\* 12: "programmer"

\* 13: "retired"

\* 14: "sales/marketing"

\* 15: "scientist"

\* 16: "self-employed"

\* 17: "technician/engineer"

\* 18: "tradesman/craftsman"

\* 19: "unemployed"

\* 20: "writer"

MOVIES FILE DESCRIPTION

================================================================================

Movie information is in the file "movies.dat" and is in the following

format:

MovieID::Title::Genres

- Titles are identical to titles provided by the IMDB (including

year of release)

- Genres are pipe-separated and are selected from the following genres:

\* Action

\* Adventure

\* Animation

\* Children's

\* Comedy

\* Crime

\* Documentary

\* Drama

\* Fantasy

\* Film-Noir

\* Horror

\* Musical

\* Mystery

\* Romance

\* Sci-Fi

\* Thriller

\* War

\* Western

- Some MovieIDs do not correspond to a movie due to accidental duplicate

entries and/or test entries

- Movies are mostly entered by hand, so errors and inconsistencies may exist