## 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.

## 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  $\frac{1}{2}$
- Movies are mostly entered by hand, so errors and inconsistencies may exist  $% \left( 1\right) =\left( 1\right) +\left( 1\right$