User Based CF and Item Based CF

Mritunjay And Sunil

12/10/2019

Recommended System Techniques on Airbnb (Amsterdam Hotel Recommendation)

1. User Based CF

2. Item Based CF

i. Cosine Simarity

ii. Pearson’s Similarity

iii. Euclidean Similarity

iv. Jacard Similarity

Load All User Defined functions

```
# Make recommendations for the target user using User-based CF
getrecommendations_UUB <- function(targetuser, users, topN=5, simfun=jacardsim) {
  sims = apply(users,1,function(user) simfun(user,targetuser))
  sims = sims[!is.na(sims) & sims >=0]
  wavrats = apply(users[names(sims),is.na(targetuser),drop=FALSE],2,function(unseenrats) sum(sims*unseenrats,na.rm=TRUE))
  s = sort(wavrats[!is.na(wavrats)], decreasing = TRUE)
  if (topN == FALSE) s else s[1:min(topN,length(s))] # get topN items
}

# get recommedations for the target user using Item-based CF
getrecommendations_IIB <- function(targetuser, itemsims, topN=5) {
  targetuser = targetuser[colnames(itemsims)] # ensure the item order is the same
  seenitems  = !is.na(targetuser)
  unseenitems = is.na(targetuser)
  seenrats = targetuser[seenitems]
  #not much difference between below two options
  preds = apply(itemsims[unseenitems,seenitems, drop=FALSE], 1, function(sims) my.weighted.mean(sims, seenrats))
  #preds = apply(itemsims[unseenitems,seenitems, drop=FALSE], 1, function(sims) sum(sims*seenrats,na.rm=TRUE))
  s = sort(preds[!is.na(preds)] , decreasing = TRUE)
  s[1:min(topN,length(s))]  # get topN items
}

# evaluate recomendations (if trainusers != NULL then do User-based CF else do Item-based CF)
# computes #testitems in topN recommendations (hits) for each testuser across a set of hold-out testitems
evalrecs = function(testusers, trainusers=NULL, itemsims=NULL, numtestitems=10, random=FALSE, topN=3, simfun=jacardsim) {
  res = sapply(1:nrow(testusers),function(i) {
    cat(".")
    testuserI(testusers[i,],trainusers=trainusers,itemsims=itemsims,numtestitems=numtestitems,random=random,topN=topN,simfun=simfun)})
  colnames(res) = rownames(testusers)
  res
}

# may give inaccurate results if testuser is in trainusers (trainuser ratings on testitem are not hidden)
testuserI <- function(testuser, trainusers=NULL, itemsims=NULL, numtestitems=10, random=FALSE, topN=3, simfun=jacardsim) {
  seenitemnames   = names(testuser)[!is.na(testuser)]
  unseenitemnames = names(testuser)[is.na(testuser)]  # may be null
  if (random) testitemnames = sample(seenitemnames,min(numtestitems,length(seenitemnames))) # test random N items
  else testitemnames = seenitemnames[1:min(numtestitems,length(seenitemnames))] # test first N items

  recs = ranks = list()
  rand = is.null(trainusers) & is.null(itemsims)
  for (testitemname in testitemnames) {
    truerating = testuser[testitemname]
    testuser[testitemname] = NA
    unseenitems = c(testitemname, unseenitemnames)
    if (!is.null(trainusers)) {
      # user-based CF
      usersims = apply(trainusers,1,function(trainuser) simfun(trainuser,testuser))
      usersims = usersims[!is.na(usersims) & usersims >=0]
      uitemsims = apply(trainusers[names(usersims),unseenitems,drop=FALSE],2,function(itemrats) sum(usersims*itemrats,na.rm=TRUE))
    }
    else if (!is.null(itemsims)) {
      # item-based CF
      seenitems = setdiff(seenitemnames, testitemname)
      seenrats = testuser[seenitems]
      uitemsims = apply(itemsims[unseenitems,seenitems,drop=FALSE],1,function(sims) my.weighted.mean(sims,seenrats))
    }
    else {
      # random prediction
      topNitems = sample(unseenitems,min(topN,length(unseenitems)))
      recs = c(recs,as.integer(is.element(testitemname,topNitems)))
    }
    if(!rand) {
      names(uitemsims) = unseenitems
      ssims = sort(uitemsims[!is.na(uitemsims)], decreasing = TRUE)
      ssims = ssims[1:min(topN,length(ssims))]
      res = as.integer(is.element(testitemname,names(ssims))) # test if the testitem is in the topN recommendations
      #res = paste(testitemname, paste(names(ssims[1:5]),collapse=",") ) # output item names only
      recs = c(recs,res)

      rk  = rank(uitemsims, na.last=NA)   # removes NA's
      rkpc = ((length(rk) - rk[testitemname] + 1)*100)/length(rk)
      ranks = c(ranks,rkpc)
    }
    testuser[testitemname] = truerating # restore the actual rating
  }
  # ensure outut is fixed length array
  if (length(recs)==0)  m1 = matrix(NA,numtestitems)
  else {
    m1 = as.matrix(recs)
    if (length(m1) < numtestitems) for (i in (length(m1)+1):(numtestitems)) {m1=rbind(m1,NA)}
  }
  if (length(ranks)==0) m2 = matrix(NA,numtestitems)
  else {
    m2 = as.matrix(ranks)
    if (length(m2) < numtestitems) for (i in (length(m2)+1):(numtestitems)) {m2=rbind(m2,NA)}
  }
  return(cbind(m1,m2))
}

meanHR = function(recs) {mean(unlist(recs[1:nrow(recs)/2,]),na.rm=TRUE)}
meanPR = function(recs) {mean(unlist(recs[(nrow(recs)/2+1):nrow(recs),]),na.rm=TRUE)}



# Make recommendations for the target user using User-based CF
getrecommendations_UU <- function(targetuser, users, topN=5, simfun=pearsonsim) {
  sims = apply(users,1,function(user) simfun(user,targetuser))
  sims = sims[!is.na(sims) & sims >=0]
  wavrats = apply(users[names(sims),is.na(targetuser), drop=FALSE],2,function(rats) weighted.mean(rats, sims, na.rm=TRUE))
  s = sort(wavrats[!is.na(wavrats)], decreasing = TRUE)
  if (topN == FALSE) s else s[1:min(topN,length(s))] # get topN items
}

# get recommedations for the target user using Item-based CF
getrecommendations_II <- function(targetuser, itemsims, topN=5) {
  targetuser = targetuser[colnames(itemsims)] # ensure the item order is the same as simmatrix
  seenitems  = !is.na(targetuser)
  unseenitems = is.na(targetuser)
  seenrats = targetuser[seenitems]
  preds = apply(itemsims[unseenitems,seenitems, drop=FALSE], 1, function(simrow) my.weighted.mean(seenrats, simrow))
  sp = sort(preds[!is.na(preds)] , decreasing = TRUE)
  sp[1:min(topN,length(sp))]  # get topN items
}

# compute the item-item similarity matrix (the matrix is symmetric so can compute half & then copy)
# (setting dir=1 generates the user similarity matrix)
getitemsimsmatrix = function(users, simfun=cosinesim, dir=2) {
  rw <<- 1;
  itemsims = apply(users, dir, function(itemA) {
    rw <<- rw + 1 ; cl <<- 1;
    apply(users,dir,function(itemB) {cl<<-cl+1; if (cl<rw) NA else if (cl==rw) NA else simfun(itemA,itemB)})
  })
  m = forceSymmetric(itemsims,uplo="L") # copy lower half to upper half
  as.matrix(m)
}

# similarity functions
euclidsim = function(x,y) { z=(y-x)^2; sz=sqrt(sum(z,na.rm=TRUE));
                            if (sz!=0) 1/(1+sz) else if (length(which(!is.na(z)))==0) NA else 1/(1+sz)}

euclidsimF= function(x,y) { z=(y-x)^2; sz=sum(z,na.rm=TRUE);
                            if (sz!=0) 1/(1+sz) else if (length(which(!is.na(z)))==0) NA else 1/(1+sz)}

cosinesim = function(x,y) { xy = x*y; sum(xy, na.rm=TRUE)/(sqrt(sum(x[!is.na(xy)]^2)*sum(y[!is.na(xy)]^2)))}

pearsonsim= function(x,y) { suppressWarnings(cor(unlist(x),unlist(y),use="pairwise.complete.obs")) }

mypearsim = function(x,y) { xy = x*y; x=x[!is.na(xy)]; y=y[!is.na(xy)];
                            mx=mean(x); my=mean(y);
                            sum((x-mx)*(y-my))/(sqrt(sum((x-mx)^2)*sum((y-my)^2)))}

pearsonRM = function(x,y) { mx=mean(x,na.rm=TRUE);my=mean(y,na.rm=TRUE);
                            xy=x*y;x=x[!is.na(xy)]; y=y[!is.na(xy)]
                            sum((x-mx)*(y-my))/(sqrt(sum((x-mx)^2)*sum((y-my)^2)))}

jacardsim = function(x,y) { validx= !is.na(x); validy= !is.na(y);
                            sum(as.integer(validx&validy))/sum(as.integer(validx|validy))}

##############################################################################
```

```
# For testing, we split the data by user, so test users are not in the trainset
# This is clean but does not test the situation where partial information
# is known about a user (as may be the case in User-based scenario).
# For item-based having partial info will make very little difference (since simmatrix is pre
computed)
################################################################################

# make predicted ratings for a sample of items for each test user
# if trainusers is defined then do User-based CF else do Item-based CF
# Note: if Item-based CF is to be performed them the itemsimilarity matrix (itemsims) must be
defined
predictCF = function(testusers, trainusers=NULL, itemsims=NULL, numtestitems=10, random=FALSE
, simfun=cosinesim) {
  preds = sapply(1:nrow(testusers),function(i) {
    cat(".")
    predictuser(testusers[i,],trainusers=trainusers,itemsims=itemsims,numtestitems=numtestite
ms,random=random,simfun=simfun)})
  colnames(preds) = rownames(testusers)
  preds
}

predictuser <- function(testuser, trainusers=NULL, itemsims=NULL, numtestitems=10, random=FAL
SE, simfun=cosinesim) {
  seenitemnames   = names(testuser)[!is.na(testuser)]
  if (random) testitemnames = sample(seenitemnames,min(numtestitems,length(seenitemnames))) #
test a random N items
  else testitemnames = seenitemnames[1:min(numtestitems,length(seenitemnames))] # test first
 N items
  preds = list()
  for (testitemname in testitemnames) {
    truerating = testuser[testitemname]
    testuser[testitemname] = NA
    if (!is.null(trainusers)) {
      # do user-based CF
      usersims = apply(trainusers,1,function(trainuser) simfun(trainuser,testuser))
      usersims = usersims[!is.na(usersims) & usersims >=0]
      predictedrating = my.weighted.mean(trainusers[names(usersims),testitemname], usersims)
    }
    else {
      # do item-based CF
      predictedrating = my.weighted.mean(testuser[seenitemnames], itemsims[seenitemnames,test
itemname])
    }
    testuser[testitemname] = truerating # restore the actual rating
    preds = c(preds,predictedrating,truerating)
  }
  preds = unname(preds)
  m = as.matrix(preds)
  if (length(m) < numtestitems*2) for (i in (length(m)+1):(numtestitems*2)) { m = rbind(m,NA
)}
  return(m)
}

# a weighted mean that handles NA's in both arguments (ratings and similarities)
my.weighted.mean = function(x,y) {
    xy = x*y;
    z = sum(abs(y[!is.na(xy)]))
    if (z == 0) as.numeric(NA) else sum(xy,na.rm=TRUE)/z
```

```
}

# computes average, mean absolute error
# each row contains prediction, actual, prediction, actual etc, hence errors are just the diff between consecutive cells
avgMAE = function(preds) {
  plist = unlist(preds)
  errors = sapply(1:(length(plist)/2),function(i) abs(plist[i*2-1]-plist[i*2]))
  errors = errors[errors != Inf]
  mean(errors,na.rm=TRUE)
}

showCM = function(preds, like) {
  plist = unlist(preds)
  cnts = sapply(1:(length(plist)/2), function(i) {
    pred = plist[i*2-1] ; actual = plist[i*2]
    if (!is.na(pred) & !is.nan(actual)) {
      if (pred>=like) {if(actual>=like) c(1,0,0,0) else c(0,1,0,0)}
      else if(actual<like) c(0,0,1,0) else c(0,0,0,1)
    } else c(0,0,0,0)
  })
  s = rowSums(cnts)   #returns cnts for: TP, FP, TN, FN

  cat(sprintf("TN=%5d FP=%5d\n",s[3],s[2]))
  cat(sprintf("FN=%5d TP=%5d  (total=%d)\n",s[4],s[1], sum(s)))
  cat(sprintf("accuracy  = %0.1f%%\n",(s[1]+s[3])*100/sum(s)))
  cat(sprintf("precision = %3.1f%%\n",s[1]*100/(s[1]+s[2])))
  cat(sprintf("recall    = %3.1f%%\n",s[1]*100/(s[1]+s[4])))
}

#######################
# miscellaneous aids
#######################

maketraintest = function(users,numtestusers) {
  testnames  = sample(rownames(users), min(numtestusers,nrow(users))) # identify N users randomly for testing
  trainnames = setdiff(rownames(users),testnames) # take remaining users for training
  trainusers <<- users[trainnames,]
  testusers  <<- users[testnames,]
  list(trainusers,testusers)
}

# extract only prediction or only actual ratings from the output of predictCF()
listpreds= function(results) {unlist(results)[c(TRUE,FALSE)]}
listrats = function(results) {unlist(results)[c(FALSE,TRUE)]}
validcnt = function(x) length(which(is.finite(x)))

# How sparse is the data in a data frame? Compute % of non-blank entries
fillrate = function(df) {cat((length(which(!is.na(df)))*100)/(nrow(df)*ncol(df)),"%")}

# same as above but also works on vectors
fillratev = function(df) {t=unlist(df); cat((length(which(!is.na(t)))*100)/length(t),"%")}

# how many values are > 0? Compute % of entries > 0
fillrateG = function(df,thresh) {t=unlist(df); cat((length(which(!is.na(t) & t > thresh))*100)/length(t),"%")}
fillrateL = function(df,thresh) {t=unlist(df); cat((length(which(!is.na(t) & t < thresh))*100
)/length(t),"%")}
fillrateE = function(df,thresh) {t=unlist(df); cat((length(which(!is.na(t) & t == thresh))*10
0)/length(t),"%")}
```

Load all the relevant libraries and Get the working directory and Load the Amsterdam Hotel Airbn data set

```
pacman::p_load(tidyverse, purrr, stringr, data.table, modelr, readxl,caret, corrplot, broom,
ggpubr, MASS,relaimpo, car,interplot, caTools, mice, gbm, reshape2, compiler, recommenderlab,
Matrix, knitr,tidyr, dplyr)
getwd()
```

```
## [1] "C:/Users/Rapsy/Desktop/Recommender_Assignment/MJ/Collaborative Filtering Recommendati
on System"
```

```
airbnb = read.csv("airbnb.csv", header=TRUE, sep=",") # transaction format!
names(airbnb) = c(colnames(airbnb))
head(airbnb,1)
```

```
##   Hotel_Id        Host_Name User_Id User_Name
## 1     2818 Erik And Mary Jo 2914515     Ivana
##                                 Hotel_name
## 1 Quiet Garden View Room & Super Fast WiFi
##                                    summary
## 1 Quiet Garden View Room & Super Fast WiFi
##
space
## 1 I'm renting a bedroom (room overlooking the garden) in my apartment in Amsterdam,  The r
oom is located to the east of the city centre in a quiet, typical Amsterdam neighbourhood the
"Indische Buurt". AmsterdamÃ¢\200\231s historic centre is less than 15 minutes away by bike o
r tram. The features of the room are: - Twin beds (80 x 200 cm, down quilts and pillows)  - 2
pure cotton towels for each guest  - reading lamps - bedside table - wardrobe - table with ch
airs - tea and coffee making facilities - mini bar - alarm clock - Hi-Fi system with cd playe
r, connection for mp3 player / phone - map of Amsterdam and public transport - Wi-Fi Internet
connection  Extra services: - Bike rental
##
description
## 1 Quiet Garden View Room & Super Fast WiFi I'm renting a bedroom (room overlooking the gar
den) in my apartment in Amsterdam,  The room is located to the east of the city centre in a q
uiet, typical Amsterdam neighbourhood the "Indische Buurt". AmsterdamÃ¢\200\231s historic cen
tre is less than 15 minutes away by bike or tram. The features of the room are: - Twin beds
(80 x 200 cm, down quilts and pillows)  - 2 pure cotton towels for each guest  - reading lamp
s - bedside table - wardrobe - table with chairs - tea and coffee making facilities - mini ba
r - alarm clock - Hi-Fi system with cd player, connection for mp3 player / phone - map of Ams
terdam and public transport - Wi-Fi Internet connection  Extra services: - Bike rental Indisc
he Buurt ("Indies Neighborhood") is a neighbourhood in the eastern portion of the city of Ams
terdam, in the Dutch province of Noord-Holland. The name dates from the early 20th century an
d is derived from the fact that the neighbourhood's streets are named after islands a
##   host_id host_name property_type    room_type accommodates
## 1 4070804    Daniel     Apartment Private room   Two Person
##             bathrooms    bedrooms    beds bed_type
## 1 One attach bathroom One bedroom One bed Real Bed
##
amenities
## 1 {Internet,Wifi,"Paid parking off premises","Buzzer/wireless intercom",Heating,Washer,"Sm
oke detector","Carbon monoxide detector","First aid kit","Safety card","Fire extinguisher",Es
sentials,Shampoo,"Lock on bedroom door","24-hour check-in",Hangers,"Hair dryer",Iron,"Laptop
friendly workspace","translation missing: en.hosting_amenity_49","translation missing: en.hos
ting_amenity_50","Private entrance","Hot water","Bed linens","Extra pillows and blankets","Si
ngle level home","Garden or backyard","No stairs or steps to enter","Flat path to guest entra
nce","Well-lit path to entrance","No stairs or steps to enter","Accessible-height bed","No st
airs or steps to enter","Host greets you","Handheld shower head","Paid parking on premises"}
##           cancellation_policy Ratings
## 1 strict_14_with_grace_period       3
```

Structure of Datasets

```
#airbnb$Hotel_Id = as.factor(airbnb$Hotel_Id)
# airbnb$User_Id = as.factor(airbnb$User_Id)
#airbnb$Hotel_Id = as.character(airbnb$Hotel_Id)
#length(unique(airbnb$Hotel_Id))
#airbnb$Hotel_Id = factor(airbnb$Hotel_Id,levels=c(unique(airbnb$Hotel_Id)), ordered = FALSE)
str(airbnb)
```

```
## 'data.frame':    20677 obs. of  20 variables:
##  $ Hotel_Id          : int  2818 2818 2818 2818 2818 2818 2818 2818 2818 2818 ...
##  $ Host_Name         : Factor w/ 508 levels "Aafje","Adriana",..: 136 136 136 136 136 136
136 136 136 136 ...
##  $ User_Id           : int  2914515 5711109 2944771 4620679 373226 2200958 1348274 543307
6 2847616 857406 ...
##  $ User_Name         : Factor w/ 2932 levels "(Email hidden by Airbnb)",..: 1205 1153 287
5 1130 2021 2308 413 2823 569 1964 ...
##  $ Hotel_name        : Factor w/ 507 levels "'Westerpark Sanctuary', Office-Apartmen
t",..: 383 383 383 383 383 383 383 383 383 383 ...
##  $ summary           : Factor w/ 382 levels "","'LORE'S PLACE' A lovely, open writers hom
e in the fun 'Indische Buurt' in Amsterdam! We are offering a open pla"| __truncated__,..: 24
2 242 242 242 242 242 242 242 242 242 ...
##  $ space             : Factor w/ 504 levels "","- 100 m2 floor space - private garden of
45 m2  - living room with a '30s bar, 55 inch QLED TV and home cinema "| __truncated__,..: 15
8 158 158 158 158 158 158 158 158 158 ...
##  $ description       : Factor w/ 506 levels "'LORE'S PLACE' A lovely, open writers home i
n the fun 'Indische Buurt' in Amsterdam! We are offering a open pla"| __truncated__,..: 317 3
17 317 317 317 317 317 317 317 317 ...
##  $ host_id           : int  4070804 4070804 4070804 4070804 4070804 4070804 4070804 40708
04 4070804 4070804 ...
##  $ host_name         : Factor w/ 404 levels "Aafje","Adriana",..: 81 81 81 81 81 81 81 81
81 81 ...
##  $ property_type     : Factor w/ 15 levels "Apartment","Bed and breakfast",..: 1 1 1 1 1
1 1 1 1 1 ...
##  $ room_type         : Factor w/ 3 levels "Entire home/apt",..: 2 2 2 2 2 2 2 2 2 2 ...
##  $ accommodates      : Factor w/ 10 levels "Five Person",..: 10 10 10 10 10 10 10 10 10 1
0 ...
##  $ bathrooms         : Factor w/ 11 levels "Four attach bathroom",..: 4 3 3 3 3 3 3 3 3 3
...
##  $ bedrooms          : Factor w/ 7 levels "Five bedroom",..: 3 3 3 3 3 3 3 3 3 3 ...
##  $ beds              : Factor w/ 7 levels "Five bed","Four bed",..: 3 3 3 3 3 3 3 3 3 3
...
##  $ bed_type          : Factor w/ 4 levels "Couch","Futon",..: 4 4 4 4 4 4 4 4 4 4 ...
##  $ amenities         : Factor w/ 508 levels "{\"Cable TV\",Internet,Wifi,\"Paid parking o
ff premises\",\"Buzzer/wireless intercom\",Heating,\"Family/kid fri"| __truncated__,..: 16 16
16 16 16 16 16 16 16 16 ...
##  $ cancellation_policy: Factor w/ 3 levels "flexible","moderate",..: 3 3 3 3 3 3 3 3 3 3
...
##  $ Ratings           : int  3 2 5 3 3 3 3 3 2 3 ...
```

Summary of Dataset

```
#summary(airbnb)
```

Create a dataset for CF from main airbnb dataset (User_ID, Hotel_ID, Ratings)

```
colnames(airbnb)
```

```
##  [1] "Hotel_Id"             "Host_Name"            "User_Id"
##  [4] "User_Name"            "Hotel_name"           "summary"
##  [7] "space"                "description"          "host_id"
## [10] "host_name"            "property_type"        "room_type"
## [13] "accommodates"         "bathrooms"            "bedrooms"
## [16] "beds"                 "bed_type"             "amenities"
## [19] "cancellation_policy" "Ratings"
```

```
airbnbCF = airbnb[,c("User_Id","Hotel_Id","Ratings")]
head(airbnbCF,4)
```

```
##   User_Id Hotel_Id Ratings
## 1 2914515     2818       3
## 2 5711109     2818       2
## 3 2944771     2818       5
## 4 4620679     2818       3
```

Unique User and Hotel

```
length(unique(airbnbCF$User_Id))
```

```
## [1] 2932
```

```
length(unique(airbnbCF$Hotel_Id))
```

```
## [1] 508
```

```
dim(airbnbCF)
```

```
## [1] 20677     3
```

Removing all those users corresponding to missing ratings and

Extract only the explicit ratings and visualize the histogram of Ratings

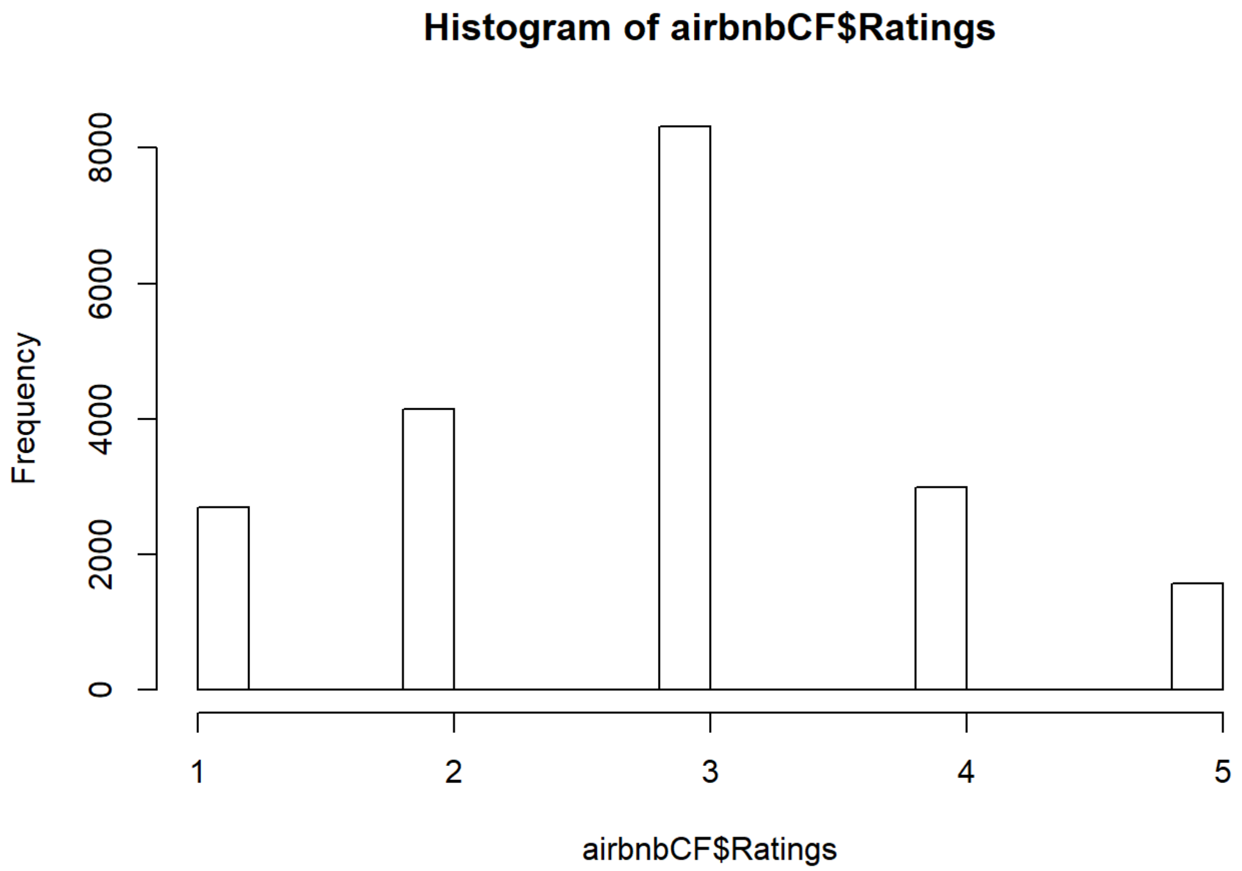
```
sapply(airbnbCF, function(x){sum(is.na(x))})
```

```
##  User_Id Hotel_Id  Ratings
##        0        0      967
```

```
airbnbCF$Ratings[is.na(airbnbCF$Ratings)] = 0
airbnbCF = airbnbCF[airbnbCF$Ratings > 0,]
sapply(airbnbCF, function(x){sum(is.na(x))})
```

```
##  User_Id Hotel_Id  Ratings
##        0        0        0
```

```
hist(airbnbCF$Ratings)
```



Eliminate users with too few ratings and Consider Activer users who had rated hotels more than and equal to 10 hotels

```
cnts = aggregate(Hotel_Id ~ User_Id, data = airbnbCF, FUN = length)
colnames(cnts) = c("user","numitems")
activeusers = cnts$user[cnts$numitems >= 10] ; length(activeusers)
```

```
## [1] 422
```

```
evCF = airbnbCF[airbnbCF$User_Id %in% activeusers,]
dim(evCF)
```

```
## [1] 4672    3
```

Eliminate Hotels with too few ratings and Consider Active Hotels who had been rated more than and equal to 10 users

```
cnts = aggregate(User_Id ~ Hotel_Id, data = airbnbCF, FUN=length)
colnames(cnts) = c("item","numusers")
popularhotels = cnts$item[cnts$numusers >= 10] ; length(popularhotels)
```

```
## [1] 508
```

```
ev = evCF[evCF$Hotel_Id %in% popularhotels,]
dim(ev)
```

```
## [1] 4672    3
```

```
str(ev)
```

```
## 'data.frame':    4672 obs. of  3 variables:
##  $ User_Id : int  2944771 2847616 2807294 4489932 5461945 4380449 4644013 913549 5039682 4
017740 ...
##  $ Hotel_Id: int  2818 2818 2818 2818 2818 2818 2818 20168 20168 20168 ...
##  $ Ratings : num  5 2 4 3 3 2 5 2 4 2 ...
```

Remove duplicate records from the datasets

```
ev_Final = ev %>% distinct(User_Id,Hotel_Id,.keep_all = TRUE)
dim(ev_Final)
```

```
## [1] 4621    3
```

```
str(ev_Final)
```

```
## 'data.frame':    4621 obs. of  3 variables:
##  $ User_Id : int  2944771 2847616 2807294 4489932 5461945 4380449 4644013 913549 5039682 4
017740 ...
##  $ Hotel_Id: int  2818 2818 2818 2818 2818 2818 2818 20168 20168 20168 ...
##  $ Ratings : num  5 2 4 3 3 2 5 2 4 2 ...
```

Convert the dataframe from long to wide format

```
users = acast(ev_Final, User_Id ~ Hotel_Id, value.var = "Ratings")
users = sweep(users, 1, rowMeans(users, na.rm=TRUE) )  # normalise the data
dim(users)
```

```
## [1] 422 508
```

Check the sparsity and fill rate of the matrix

```
fillrate(users)
```

```
## 2.155558 %
```

setup the train/test scheme

```
numtestusers = 84
test  = sample(rownames(users), min(numtestusers,nrow(users)))
train = setdiff(rownames(users),test)
```

User Based Collaborative Filtering

1. Cosine similarity

Test UBCF Cosine similarity metrics

```
preds = predictCF(users[test,], users[train,], numtestitems=10, random=FALSE, simfun=cosinesi
m) ; preds
```

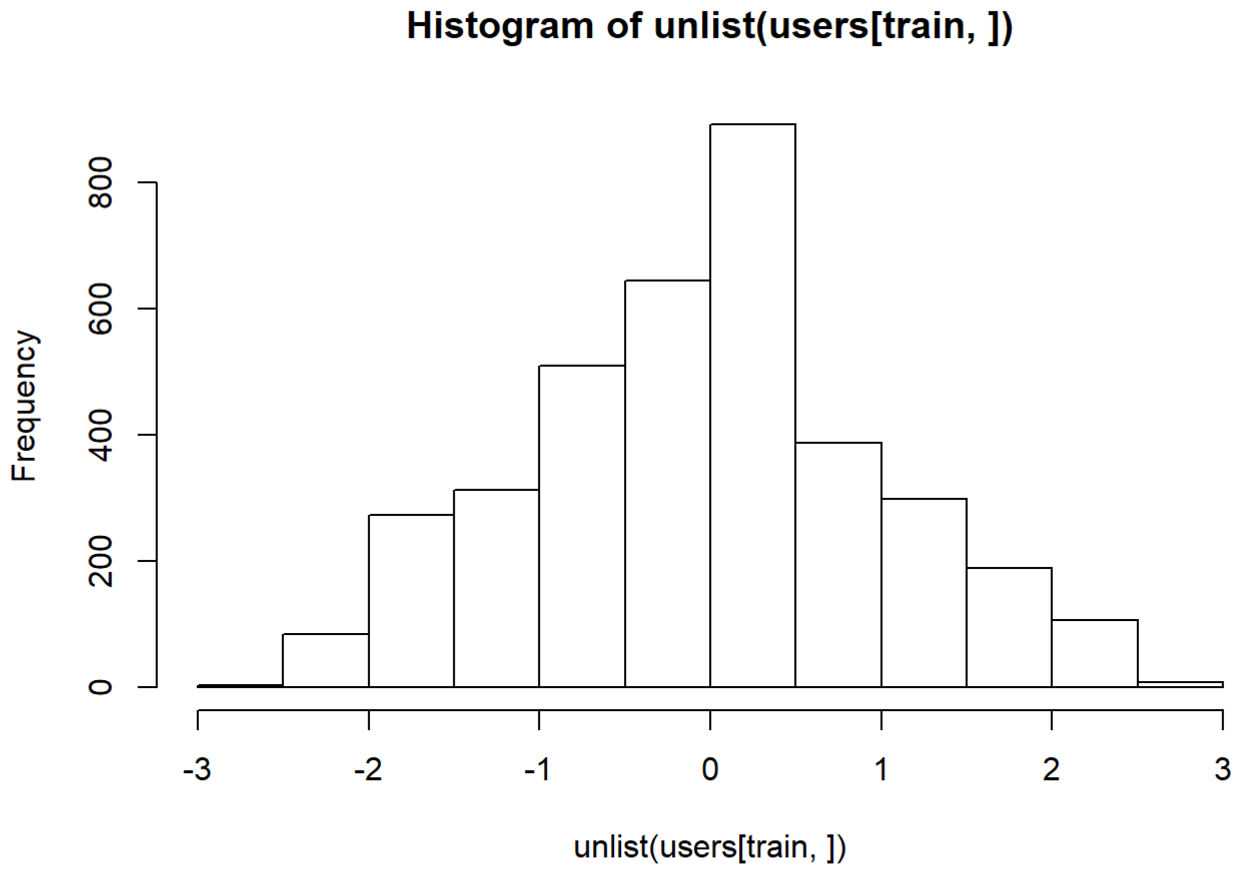
```
## ....................................................................................
##       4364462    4875280    1283321    4117053    1486411    2210060
##  [1,] 1.384615   -0.9055556 0.4        -0.6666667 NA         NA
##  [2,] 0.09090909 0.3        -1.75      0.7272727  -0.2       1.666667
##  [3,] NA         NA         -1.692308  -0.8       NA         NA
##  [4,] -0.9090909 0.3        0.25       -1.272727  -0.2       -0.3333333
##  [5,] NA         0.2        NA         NA         NA         -1.67803
##  [6,] -0.9090909 -1.7       -1.75      -0.2727273 -0.2       1.666667
##  [7,] NA         0.2        0.09090909 -0.7333333 0.1666667  2.083333
##  [8,] 2.090909   -1.7       -0.75      -0.2727273 -1.2       -2.333333
##  [9,] NA         0.1666667  NA         NA         -0.8333333 NA
## [10,] -1.909091  0.3        0.25       -2.272727  1.8        0.6666667
## [11,] NA         -0.2230769 0.1666667  2.090909   NA         0.8
## [12,] -1.909091  -1.7       0.25       1.727273   -0.2       -0.3333333
## [13,] 0.6        NA         -1.454545  0.5333333  -0.5090909 0
## [14,] 2.090909   1.3        2.25       -0.2727273 -1.2       -0.3333333
## [15,] NA         0.1538462  0.1666667  0.2        0.1818182  -0.5714286
## [16,] 0.09090909 0.3        0.25       0.7272727  0.8        -0.3333333
## [17,] 1          NA         NA         1.052525   -0.2       NA
## [18,] 1.090909   0.3        1.25       0.7272727  -1.2       -0.3333333
## [19,] -0.6153846 0.1666667  -0.3       NA         NA         NA
## [20,] -0.9090909 2.3        1.25       -1.272727  1.8        -0.3333333
##       5495353     5446448     2969181     4314129     1898757
##  [1,] NA          -0.2285714  0.1666667   -0.1        NA
##  [2,] -1.4        1.7         0.2         -0.1        0.5
##  [3,] -0.45       0.2445887   NA          -0.7211538  NA
##  [4,] 0.6         0.7         0.2         1.9         1.5
##  [5,] 0.3         -0.07692308 NA          NA          NA
##  [6,] 1.6         1.7         2.2         -1.1        -1.5
##  [7,] 0.3         -0.2727273  0.9090909   -0.1        NA
##  [8,] 1.6         -0.3        0.2         0.9         1.5
##  [9,] -0.8333333  NA          0.1666667   NA          -0.4681818
## [10,] -0.4        -0.3        0.2         -2.1        0.5
## [11,] -1.6        -0.5        NA          NA          0.3636364
## [12,] -0.4        -0.3        -0.8        -1.1        0.5
## [13,] NA          -0.2727273  NA          -0.06730769 0.4
## [14,] 0.6         -0.3        -0.8        -0.1        -0.5
## [15,] -0.07727273 NA          -0.09090909 1.5         NA
## [16,] -0.4        0.7         -0.8        -0.1        0.5
## [17,] -0.3777778  NA          NA          0.8         NA
## [18,] -0.4        -2.3        0.2         1.9         -1.5
## [19,] 0.4090909   NA          0.2727273   -0.15       NA
## [20,] -1.4        -1.3        -0.8        -0.1        -1.5
##       2691000      3399045    2536879     4447646    3308880   5039682
##  [1,] -0.007211538 NA         NA          -1.692308  0.6363636 NA
##  [2,] -1.615385    0.07692308 -0.3333333  0.3636364  -0.1      1
##  [3,] NA           NA         -1.384615   NA         NA        NA
##  [4,] 0.3846154    0.07692308 -0.3333333  1.363636   -0.1      -1
##  [5,] NA           NA         NA          0.1        NA        NA
##  [6,] 0.3846154    0.07692308 1.666667    -1.636364  1.9       -1
##  [7,] NA           NA         NA          NA         0.1818182 NA
##  [8,] -1.615385    2.076923   -0.3333333  -0.6363636 -1.1      1
##  [9,] NA           NA         NA          NA         -0.1      NA
## [10,] -1.615385    -0.9230769 0.6666667   -1.636364  0.9       -2
## [11,] NA           NA         NA          0.09090909 NA        NA
## [12,] 0.3846154    -0.9230769 -1.333333   -0.6363636 0.9       0
## [13,] -0.6         NA         NA          NA         -0.75     NA
## [14,] 2.384615     -1.923077  1.666667    1.363636   -0.1      0
## [15,] NA           NA         -0.04230769 NA         NA        NA
## [16,] 0.3846154    1.076923   -1.333333   1.363636   -2.1      0
## [17,] -1           NA         NA          NA         0.6       -0.07692308
## [18,] 0.3846154    0.07692308 -0.3333333  -2.636364  -0.1      1
## [19,] NA           NA         NA          NA         -0.75     NA
## [20,] -0.6153846   -0.9230769 NA          1.363636   -0.1      1
##       5037898    1602026    3631890    3628337    5537923 1542832
##  [1,] NA         -0.5832035 -0.2       -1.007692 NA      NA
##  [2,] 1.5        -1.714286  -1.272727  0.3333333 1       -0.1818182
##  [3,] NA         NA         1.384615   NA        NA      NA
##  [4,] -0.5       -1.714286  -2.272727  0.3333333 0       0.8181818
##  [5,] NA         -0.2751826 -1.7       NA        NA      NA
##  [6,] -0.5       1.285714   0.7272727  0.3333333 0       -0.1818182
##  [7,] NA         -0.5807309 NA         -0.1      NA      NA
##  [8,] -0.5       0.2857143  1.727273   0.3333333 0       -1.181818
##  [9,] -1.363636  -0.1428571 NA         NA        NA      0.3645299
## [10,] 1.5        0.2857143  -0.2727273 -1.666667 -1      -0.1818182
## [11,] -1.833333  NA         -0.9090909 0.5454545 NA      1.333333
## [12,] -1.5       -1.714286  1.727273   -1.666667 -1      -1.181818
## [13,] NA         -0.5863636 NA         -1.975    NA      NA
## [14,] 0.5        1.285714   -1.272727  1.333333  0       1.818182
## [15,] NA         -0.1095892 2          NA        NA      NA
## [16,] -0.5       -1.714286  -0.2727273 0.3333333 1       -0.1818182
## [17,] NA         NA         NA         NA        NA      NA
## [18,] -1.5       1.285714   1.727273   0.3333333 0       -0.1818182
## [19,] 0.08333333 0.2507177  0.4030303  NA        NA      -0.9887703
## [20,] 1.5        1.285714   0.7272727  NA        0       0.8181818
##       4046204    2686549 5732915  532034     4412129    808981
##  [1,] -0.7857143 NA      NA       NA         -0.6619102 NA
##  [2,] -1         0       -1       0.4545455  2.3        -0.7272727
##  [3,] -0.75      NA      NA       -0.6363636 -0.05      1.2
##  [4,] 1          1       0        0.4545455  0.3        1.272727
##  [5,] NA         NA      NA       NA         NA         NA
##  [6,] 0          0       1        1.454545   -0.7       2.272727
##  [7,] 0.3141026  NA      NA       -0.5       0.1666667  -0.9090909
##  [8,] -1         2       1        0.4545455  -0.7       -1.727273
##  [9,] NA         NA      1.307692 1.8        NA         1.041667
## [10,] 1          0       0        0.4545455  1.3        -0.7272727
## [11,] -0.7479119 NA      2.25     NA         NA         0.5714286
## [12,] 0          2       1        -1.545455  0.3        -1.727273
## [13,] 0.3671329  NA      -0.75    NA         NA         -0.08333333
## [14,] 2          -1      -1       -0.5454545 -1.7       -1.727273
## [15,] -1.545455  NA      NA       NA         0.1666667  NA
## [16,] 1          -2      0        -0.5454545 1.3        0.2727273
## [17,] -0.7135544 NA      NA       NA         0.8        0.2
## [18,] -2         -1      0        -0.5454545 -0.7       0.2727273
## [19,] 0.2142857  NA      NA       NA         NA         0.09090909
## [20,] 1          -1      0        -0.5454545 -1.7       2.272727
##       3281295     4415881   2941133    1539253    4175055     4180511
##  [1,] NA          NA        NA         NA         NA          -0.7
##  [2,] -1.363636   1.5       -0.2       1.1        -2.461538   -1.5
##  [3,] NA          NA        -0.1833333 -1.818182  -1.048701   0.7
##  [4,] -0.3636364  0.5       -2.2       -1.9       0.5384615   0.5
##  [5,] -0.8260218  NA        NA         NA         0.01428571  1.9
##  [6,] 0.6363636   0.5       1.8        1.1        1.538462    1.5
##  [7,] 0.3846154   NA        NA         2.181818   NA          0.8333333
##  [8,] 0.6363636   0.5       -0.2       0.1        -2.461538   0.5
##  [9,] NA          NA        NA         0.8333333  0.3153846   NA
## [10,] 0.6363636   -0.5      -1.2       1.1        1.538462    0.5
## [11,] 0.008333333 NA        -0.2730769 -0.1666667 -0.1        -0.2
## [12,] 0.6363636   0.5       0.8        -0.9       -0.4615385  0.5
## [13,] 0.2727273   1.1       0.9038462  -0.6       0.008333333 -0.1666667
## [14,] -1.363636   -1.5      -0.2       0.1        -0.4615385  -0.5
## [15,] -0.0810392  NA        NA         NA         0.8461538   NA
## [16,] 0.6363636   0.5       1.8        -0.9       1.538462    0.5
## [17,] -1.4        NA        NA         -0.4444444 NA          0.4
## [18,] -0.3636364  -1.5      -2.2       0.1        -0.4615385  -1.5
## [19,] 0.5454545   0.4545455 NA         NA         1.664286    0.4545455
## [20,] -0.3636364  0.5       1.8        0.1        1.538462    -0.5
##       3723952     3061513    4309450    1468271 2510107    1259411
##  [1,] -1.923077   NA         -1.2       NA      NA         NA
##  [2,] 1.916667    1.545455   0.2        0       0          0.2727273
##  [3,] 0.9         NA         -1.3       NA      NA         NA
##  [4,] 0.9166667   -0.4545455 2.2        0       0          -0.7272727
##  [5,] 0.5969697   NA         NA         NA      NA         NA
##  [6,] -0.08333333 0.5454545  1.2        0       0          1.272727
##  [7,] NA          NA         0.06666667 NA      1.25       0
##  [8,] -0.08333333 -0.4545455 -0.8       0       2          1.272727
##  [9,] NA          NA         NA         NA      0.25       1.5
## [10,] -2.083333   -1.454545  0.2        1       2          0.2727273
## [11,] NA          NA         NA         NA      -0.8461538 0.4
## [12,] -2.083333   0.5454545  -0.8       -2      -2         0.2727273
## [13,] -0.3846154  NA         NA         NA      NA         0.2142857
## [14,] 1.916667    -0.4545455 0.2        0       0          1.272727
## [15,] 0.35        NA         0.4675325  NA      0.1538462  NA
## [16,] -2.083333   -0.4545455 -0.8       0       1          -0.7272727
## [17,] NA          NA         NA         NA      NA         -1.7
## [18,] -0.08333333 -1.454545  0.2        1       -2         -1.727273
## [19,] -0.6363636  NA         NA         NA      NA         0.2142857
## [20,] 1.916667    1.545455   -1.8       0       -1         0.2727273
##       294984     950010     1211966    3146203     5385464    5039919
##  [1,] NA         NA         0.7692308  -1.916667   NA         NA
##  [2,] 0          -0.1       0.4        1.166667    -0.3333333 0.3
##  [3,] -0.3321678 -0.9590909 1.517148   -0.5189394  NA         NA
##  [4,] 0          0.9        0.4        0.1666667   0.6666667  -0.7
##  [5,] NA         NA         NA         -0.04924242 0.1666667  1.2
##  [6,] -1         0.9        0.4        2.166667    1.666667   0.3
##  [7,] NA         NA         NA         -1.7        NA         NA
##  [8,] 0          -0.1       -0.6       0.1666667   -2.333333  0.3
##  [9,] -0.5098071 NA         0.3333333  0.4166667   -0.5       NA
## [10,] 1          0.9        -0.6       -0.8333333  0.6666667  -0.7
## [11,] NA         0.4861111  NA         0.7         -0.8333333 NA
## [12,] 1          -1.1       0.4        -0.8333333  -0.3333333 -0.7
## [13,] NA         -0.1111111 NA         0.3076923   -1.090909  NA
## [14,] 0          -1.1       -0.6       -1.833333   0.6666667  -1.7
## [15,] NA         NA         0.3333333  0.6         NA         0.2
## [16,] 2          1.9        1.4        -0.8333333  0.6666667  0.3
## [17,] -0.6428571 NA         0          NA          NA         -0.5454545
## [18,] -1         -0.1       -0.6       0.1666667   -0.3333333 2.3
## [19,] NA         NA         -0.6666667 1.083333    -0.1833333 NA
## [20,] -2         -2.1       -0.6       0.1666667   -1.333333  0.3
##       1178877    4830630    1325697    3441513     2460925    3375808
##  [1,] -1.833333  0.07692308 NA         NA          -1.384615  NA
##  [2,] 0.07142857 0.6        0.6        -1          -1.181818  -1.5
##  [3,] NA         -1.15      NA         0.7923077   1.416667   NA
##  [4,] -1.928571  0.6        -0.4       2           1.818182   0.5
##  [5,] NA         NA         NA         NA          0.2        NA
##  [6,] 0.07142857 1.6        0.6        -1          1.818182   -0.5
##  [7,] NA         -0.9       -0.2030303 -0.15       NA         NA
##  [8,] 0.07142857 -0.4       0.6        2           -1.181818  1.5
##  [9,] 1.7        NA         NA         -0.8        1.8        NA
## [10,] 0.07142857 0.6        1.6        -1          -0.1818182 -1.5
## [11,] -0.75      NA         NA         -0.04545455 -1.923077  NA
## [12,] -0.9285714 -1.4       -1.4       0           -0.1818182 -0.5
## [13,] NA         NA         NA         NA          -0.8       NA
## [14,] 1.071429   0.6        -1.4       0           -2.181818  1.5
## [15,] NA         NA         0.2269231  -1          1.076923   NA
## [16,] 0.07142857 -1.4       -1.4       2           0.8181818  -0.5
## [17,] 0.25       NA         NA         -0.6153846  1.13141    NA
## [18,] 2.071429   0.6        0.6        -1          0.8181818  1.5
## [19,] -0.7285714 -0.2115385 NA         NA          -2.153846  0.2454545
## [20,] 0.07142857 -1.4       0.6        -2          -0.1818182 -0.5
##       505271     3846003    4938413    3146037     4735650    510307
##  [1,] NA         NA         NA         1.5         -0.9166667 0.1
##  [2,] 0.6363636  -0.8888889 0.1        -0.8571429  0.6        -0.9
##  [3,] NA         -0.4       NA         -1.444444   NA         NA
##  [4,] -0.3636364 1.111111   1.1        0.1428571   0.6        0.1
##  [5,] NA         NA         NA         NA          NA         -0.1
##  [6,] -1.363636  0.1111111  -1.9       -0.8571429  0.6        1.1
##  [7,] NA         0.55       -0.3423077 NA          NA         NA
##  [8,] 2.636364   -1.888889  0.1        1.142857    -1.4       2.1
##  [9,] 1.454545   NA         NA         1.675       -0.5454545 NA
## [10,] -1.363636  -0.8888889 -0.9       -1.857143   -1.4       0.1
## [11,] 1.923077   NA         0.1        2.166667    NA         -0.4923077
## [12,] -0.3636364 0.1111111  0.1        -0.8571429  1.6        -1.9
## [13,] NA         -0.35      NA         -1.454545   NA         -0.2583333
## [14,] -0.3636364 0.1111111  2.1        1.142857    0.6        0.1
## [15,] 0.1538462  -0.1111111 NA         NA          NA         0.1666667
## [16,] -1.363636  0.1111111  -1.9       -0.8571429  -0.4       -0.9
## [17,] NA         NA         NA         -0.07692308 0.2818182  2.615385
## [18,] -1.363636  2.111111   1.1        0.1428571   -0.4       0.1
## [19,] 1.454545   NA         0.1        1.636364    0.7777778  NA
## [20,] 2.636364   NA         0.1        1.142857    -0.4       0.1
##       2667507     5011601    4715209 4380449    1793895     5268852
##  [1,] 1.772727    0.4615385  0.5     NA         NA          NA
##  [2,] -1.090909   -0.7272727 -1.2    -0.7777778 1.8         0.9
##  [3,] NA          -0.3583333 NA      -0.3636364 -0.08333333 NA
##  [4,] -0.09090909 0.2727273  0.8     1.222222   -0.2        1.9
##  [5,] 0.4636364   -0.8181818 NA      -0.8       0.5909091   NA
##  [6,] -0.09090909 1.272727   1.8     0.2222222  -1.2        0.9
##  [7,] NA          -1         0.25    NA         0.1181818   -1.1
##  [8,] -0.09090909 0.2727273  0.8     -0.7777778 -0.2        -0.1
##  [9,] 1.166667    0          NA      NA         1.7         NA
## [10,] 1.909091    -0.7272727 0.8     0.2222222  0.8         -2.1
## [11,] 0.2113892   NA         NA      NA         NA          NA
## [12,] 0.9090909   -0.7272727 -0.2    1.222222   -0.2        -2.1
## [13,] NA          NA         NA      NA         -0.08333333 NA
## [14,] -1.090909   -0.7272727 -1.2    0.2222222  -0.2        1.9
## [15,] 0.8510526   -0.5583333 NA      NA         -0.3333333  NA
## [16,] 0.9090909   0.2727273  -0.2    0.2222222  -1.2        -1.1
## [17,] 0.8214286   NA         -0.75   NA         0.25        -0.2832168
## [18,] -0.09090909 1.272727   -0.2    -1.777778  -0.2        -2.1
## [19,] NA          0.4545455  NA      NA         0.6666667   1.757692
## [20,] -0.09090909 -1.727273  -1.2    NA         0.8         1.9
##       2865695    3093794       5647495 4688823     3821129    5058598
##  [1,] 0.4545455  0.9           NA      0.9062678   0.7833333  -0.4
##  [2,] 0.1666667  1.153846      -0.25   -0.07692308 0.4166667  -0.08333333
##  [3,] NA         NA            NA      NA          NA         NA
##  [4,] -0.8333333 0.1538462     1.75    -0.07692308 0.4166667  -0.08333333
##  [5,] NA         2.225         NA      NA          NA         NA
##  [6,] 0.1666667  0.1538462     -0.25   -0.07692308 0.4166667  -0.08333333
##  [7,] NA         -6.938894e-18 NA      0.2333333   NA         NA
##  [8,] -1.833333  -1.846154     -0.25   -0.07692308 -1.583333  1.916667
##  [9,] -0.280303  -0.4954545    NA      0.07179487  NA         -1.384615
## [10,] 0.1666667  -0.8461538    -1.25   -0.07692308 0.4166667  0.9166667
## [11,] NA         NA            NA      -0.2727273  0.1666667  -0.8
## [12,] 0.1666667  0.1538462     1.75    -0.07692308 -0.5833333 -1.083333
## [13,] 0.02097902 -0.9166667    NA      NA          1.55       1.9
## [14,] 0.1666667  1.153846      -0.25   -0.07692308 -0.5833333 -1.083333
## [15,] 0.4545455  -0.4424242    NA      NA          -0.5       NA
## [16,] 1.166667   0.1538462     -0.25   -0.07692308 -1.583333  1.916667
## [17,] NA         0.25          NA      -0.2727273  -0.9090909 0.55
## [18,] 0.1666667  2.153846      -0.25   -0.07692308 0.4166667  -2.083333
## [19,] NA         -0.5454545    NA      1.416667    NA         NA
## [20,] 0.1666667  -0.8461538    -1.25   -0.07692308 0.4166667  -0.08333333
##       995697     3278419    4046972    4935243  2965087    3886540
##  [1,] NA         NA         -0.6153846 0.25     -0.1666667 NA
##  [2,] -0.2       0.2727273  0.4285714  -2       -1.8       -0.4
##  [3,] NA         NA         -1.923077  NA       NA         0.3
##  [4,] -1.2       -1.727273  0.4285714  0        0.2        0.6
##  [5,] NA         NA         0.3        NA       0.6363636  NA
##  [6,] -1.2       0.2727273  -0.5714286 -2       2.2        -0.4
##  [7,] NA         NA         0.3        1.833333 NA         0.3
##  [8,] 1.8        2.272727   0.4285714  0        0.2        0.6
##  [9,] -0.4961538 NA         -0.1666667 NA       -1         0.1
## [10,] -0.2       -1.727273  -0.5714286 0        1.2        -0.4
## [11,] 0.7        NA         NA         NA       NA         -1.2
## [12,] 0.8        -0.7272727 -0.5714286 0        0.2        0.6
## [13,] NA         NA         -0.6969697 NA       0.3846154  NA
## [14,] -0.2       1.272727   0.4285714  1        -0.8       -0.4
## [15,] NA         -0.8       -0.1217949 NA       1.636364   0.08333333
## [16,] -0.2       0.2727273  0.4285714  2        0.2        0.6
## [17,] NA         0.1666667  NA         NA       -0.1666667 NA
## [18,] -0.2       1.272727   0.4285714  0        -1.8       -1.4
## [19,] -0.2727273 NA         NA         NA       -0.2       0.08333333
## [20,] 0.8        -1.727273  1.428571   -1       0.2        0.6
##       662664     2341442    3315379   4976553  270282     2991842
##  [1,] NA         0.1666667  0.5863636 NA       NA         0.04545455
##  [2,] -0.1111111 1.083333   -1.1      -0.2     -0.2727273 0.7333333
##  [3,] NA         0.2820513  NA        NA       2.214286   -1.679487
##  [4,] 0.8888889  1.083333   -0.1      0.8      -1.272727  1.733333
##  [5,] NA         -0.4346154 NA        NA       NA         -0.8
##  [6,] 0.8888889  0.08333333 0.9       -0.2     -2.272727  0.7333333
##  [7,] NA         -0.625     NA        NA       NA         NA
##  [8,] -0.1111111 0.08333333 -0.1      -0.2     1.727273   -1.266667
##  [9,] NA         -0.8333333 NA        NA       0.25       NA
## [10,] -0.1111111 -0.9166667 1.9       -0.2     -2.272727  1.733333
## [11,] NA         NA         NA        NA       NA         0.4
## [12,] 1.888889   -0.9166667 -0.1      -0.2     1.727273   0.7333333
## [13,] 0.09090909 NA         NA        NA       -0.3333333 0.3538462
## [14,] -1.111111  2.083333   -2.1      0.8      1.727273   0.7333333
## [15,] NA         NA         NA        NA       NA         NA
## [16,] -2.111111  0.08333333 -1.1      0.8      -0.2727273 -1.266667
## [17,] NA         -0.4444444 NA        1.727273 NA         NA
## [18,] -0.1111111 -0.9166667 0.9       -1.2     -0.2727273 -0.2666667
## [19,] NA         NA         NA        NA       NA         0.8
## [20,] NA         0.08333333 0.9       -0.2     -0.2727273 -1.266667
##       3984496
##  [1,] -0.85
##  [2,] 0.4
##  [3,] -0.4545455
##  [4,] 0.4
##  [5,] NA
##  [6,] -1.6
##  [7,] 0.1666667
##  [8,] 1.4
##  [9,] 2.454545
## [10,] -0.6
## [11,] 0.4307692
## [12,] -1.6
## [13,] 0.1666667
## [14,] 0.4
## [15,] NA
## [16,] 0.4
## [17,] -0.4
## [18,] 1.4
## [19,] -0.3846154
## [20,] -0.6
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)), "tests")
```

```
## avg MAE = 1.16354 from 375 tests
```

viewing a histogram of the ratings helps gauge the impact of different MAE’s and decide what the likethresh should be

```
hist(unlist(users[train,]))
```



```
likethresh = 2.5 # insert your own value here
showCM(preds,likethresh)
```

```
## TN=  373 FP=    1
## FN=    1 TP=    0  (total=375)
## accuracy  = 99.5%
## precision = 0.0%
## recall    = 0.0%
```

2. Euclidean similarity

```
preds = predictCF(users[test,], users[train,], numtestitems=10, random=FALSE, simfun=euclidsi
m) ; preds
```

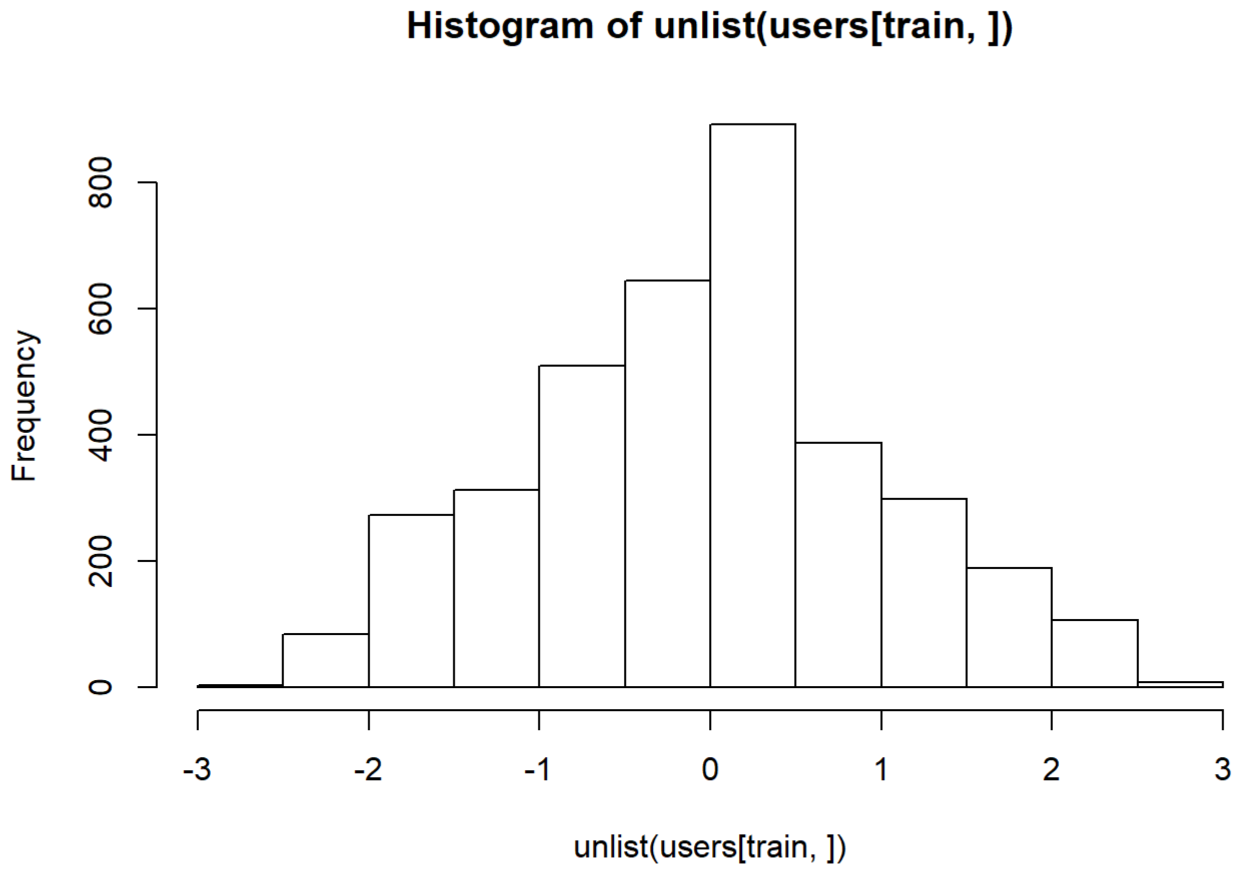
```
## ..................................................................................
##       4364462    4875280    1283321    4117053    1486411    2210060
##  [1,] 1.384615   -0.8904999 0.3719512  -0.7970778 0.1666667  0.2743862
##  [2,] 0.09090909 0.3        -1.75      0.7272727  -0.2       1.666667
##  [3,] NA         NA         -1.692308  -0.8       NA         -0.9166667
##  [4,] -0.9090909 0.3        0.25       -1.272727  -0.2       -0.3333333
##  [5,] NA         -0.2840518 -1.3       -0.6429216 0.07142857 -0.9301074
##  [6,] -0.9090909 -1.7       -1.75      -0.2727273 -0.2       1.666667
##  [7,] NA         -0.497389  0.09090909 -0.1036447 0.1333739  2.083333
##  [8,] 2.090909   -1.7       -0.75      -0.2727273 -1.2       -2.333333
##  [9,] NA         0.1666667  2.307692   NA         -0.8333333 NA
## [10,] -1.909091  0.3        0.25       -2.272727  1.8        0.6666667
## [11,] NA         -0.3772727 -0.2330097 2.090909   NA         0.8
## [12,] -1.909091  -1.7       0.25       1.727273   -0.2       -0.3333333
## [13,] 0.7272727  0.2233792  -1.454545  0.4627244  -0.621696  0
## [14,] 2.090909   1.3        2.25       -0.2727273 -1.2       -0.3333333
## [15,] NA         0.1538462  0.1666667  0.2        0.1818182  0.02748523
## [16,] 0.09090909 0.3        0.25       0.7272727  0.8        -0.3333333
## [17,] 1          2.2        0.5409091  1.108361   -0.2       -0.2
## [18,] 1.090909   0.3        1.25       0.7272727  -1.2       -0.3333333
## [19,] -0.6153846 -0.5584518 -0.3       NA         2.166667   NA
## [20,] -0.9090909 2.3        1.25       -1.272727  1.8        -0.3333333
##       5495353    5446448     2969181     4314129    1898757    2691000
##  [1,] NA         -0.4930921  -0.2341637  -0.2059147 NA         0.02002928
##  [2,] -1.4       1.7         0.2         -0.1       0.5        -1.615385
##  [3,] -0.48      0.4013808   0.1655172   -0.7384342 NA         0
##  [4,] 0.6        0.7         0.2         1.9        1.5        0.3846154
##  [5,] 0.3        -0.07692308 0           -0.7915114 0.8809524  -1
##  [6,] 1.6        1.7         2.2         -1.1       -1.5       0.3846154
##  [7,] -0.678811  -0.2727273  0.9090909   -0.1       1.7        -0.75
##  [8,] 1.6        -0.3        0.2         0.9        1.5        -1.615385
##  [9,] -0.8333333 -0.07692308 0.1666667   0.8        -0.4455253 0
## [10,] -0.4       -0.3        0.2         -2.1       0.5        -1.615385
## [11,] -0.9680442 -0.6670103  NA          0.42       0.582817   0.1640923
## [12,] -0.4       -0.3        -0.8        -1.1       0.5        0.3846154
## [13,] 0.2480152  -0.1989038  NA          0.1695992  0.4        -0.708502
## [14,] 0.6        -0.3        -0.8        -0.1       -0.5       2.384615
## [15,] -0.1536622 0.4         -0.09090909 1.5        NA         NA
## [16,] -0.4       0.7         -0.8        -0.1       0.5        0.3846154
## [17,] -0.4974553 NA          NA          0.8        0          -1
## [18,] -0.4       -2.3        0.2         1.9        -1.5       0.3846154
## [19,] 0.1620637  NA          0.4247505   -0.1322581 NA         NA
## [20,] -1.4       -1.3        -0.8        -0.1       -1.5       -0.6153846
##       3399045    2536879    4447646    3308880    5039682     5037898
##  [1,] 0.4166667  NA         -0.8001798 0.6363636  NA          NA
##  [2,] 0.07692308 -0.3333333 0.3636364  -0.1       1           1.5
##  [3,] 2.3        -1.093453  NA         -0.4       NA          NA
##  [4,] 0.07692308 -0.3333333 1.363636   -0.1       -1          -0.5
##  [5,] 0          -0.5       0.1        NA         -2.076923   NA
##  [6,] 0.07692308 1.666667   -1.636364  1.9        -1          -0.5
##  [7,] NA         NA         NA         0.3028169  -0.7272727  -1.363636
##  [8,] 2.076923   -0.3333333 -0.6363636 -1.1       1           -0.5
##  [9,] 1.545455   NA         NA         -0.1       NA          -1.363636
## [10,] -0.9230769 0.6666667  -1.636364  0.9        -2          1.5
## [11,] NA         NA         0.09090909 0.6363636  1.272727    -0.1028708
## [12,] -0.9230769 -1.333333  -0.6363636 0.9        0           -1.5
## [13,] NA         NA         NA         -0.5588985 NA          NA
## [14,] -1.923077  1.666667   1.363636   -0.1       0           0.5
## [15,] NA         0.3294446  -0.5454545 1.025148   NA          NA
## [16,] 1.076923   -1.333333  1.363636   -2.1       0           -0.5
## [17,] -1         NA         NA         1.10397    -0.07692308 -1
## [18,] 0.07692308 -0.3333333 -2.636364  -0.1       1           -1.5
## [19,] NA         NA         1.307692   0.2052632  NA          0.0560166
## [20,] -0.9230769 NA         1.363636   -0.1       1           1.5
##       1602026     3631890    3628337   5537923     1542832     4046204
##  [1,] -0.5199863  0.07162666 -1.064159 NA          -0.09090909 -0.7857143
##  [2,] -1.714286   -1.272727  0.3333333 1           -0.1818182  -1
##  [3,] NA          1.384615   0.25      -1.9        0           -0.6512054
##  [4,] -1.714286   -2.272727  0.3333333 0           0.8181818   1
##  [5,] -0.02101985 -1.7       NA        NA          0.8461538   NA
##  [6,] 1.285714    0.7272727  0.3333333 0           -0.1818182  0
##  [7,] 0.06750306  0.5        0.2253521 -0.2167563  1           0.6129032
##  [8,] 0.2857143   1.727273   0.3333333 0           -1.181818   -1
##  [9,] 0.08543756  -1.319815  NA        NA          0.5945291   NA
## [10,] 0.2857143   -0.2727273 -1.666667 -1          -0.1818182  1
## [11,] NA          -0.4088757 0.5454545 0.6428571   0.9413698   -0.5893192
## [12,] -1.714286   1.727273   -1.666667 -1          -1.181818   0
## [13,] -0.6799549  0.511436   -1.374716 0.08333333  0.4030376   0.196788
## [14,] 1.285714    -1.272727  1.333333  0           1.818182    2
## [15,] 0.2701557   1.292442   0.3846154 NA          -0.4        -0.6175178
## [16,] -1.714286   -0.2727273 0.3333333 1           -0.1818182  1
## [17,] 0.7272727   NA         1.311579  -0.04342312 NA          -0.1457571
## [18,] 1.285714    1.727273   0.3333333 0           -0.1818182  -2
## [19,] 0.4236108   0.5415186  NA        0.1         -1.051669   -0.4727479
## [20,] 1.285714    0.7272727  NA        0           0.8181818   1
##       2686549    5732915   532034     4412129    808981     3281295
##  [1,] 0          NA        0.5        -0.5470649 1.3        -0.4690181
##  [2,] 0          -1        0.4545455  2.3        -0.7272727 -1.363636
##  [3,] NA         NA        -0.5454545 -0.3586328 1.2        NA
##  [4,] 1          0         0.4545455  0.3        1.272727   -0.3636364
##  [5,] -0.6923077 0.3076923 NA         -0.1       NA         -0.7659617
##  [6,] 0          1         1.454545   -0.7       2.272727   0.6363636
##  [7,] NA         NA        -0.2121212 0.1666667  0.735392   0.01718679
##  [8,] 2          1         0.4545455  -0.7       -1.727273  0.6363636
##  [9,] -1.692308  1.307692  1.8        NA         1.438257   0.01086555
## [10,] 0          0         0.4545455  1.3        -0.7272727 0.6363636
## [11,] NA         2.25      NA         -1.328483  0.5714286  -0.0001390565
## [12,] 2          1         -1.545455  0.3        -1.727273  0.6363636
## [13,] 2          -0.75     NA         0.6        -0.5307642 0.562805
## [14,] -1         -1        -0.5454545 -1.7       -1.727273  -1.363636
## [15,] NA         NA        NA         0.1666667  NA         -0.3392636
## [16,] -2         0         -0.5454545 1.3        0.2727273  0.6363636
## [17,] NA         NA        0.6363636  1.528477   0.1928005  -0.6443906
## [18,] -1         0         -0.5454545 -0.7       0.2727273  -0.3636364
## [19,] NA         NA        NA         -0.7263158 -0.3012478 0.5454545
## [20,] -1         0         -0.5454545 -1.7       2.272727   -0.3636364
##       4415881   2941133     1539253    4175055    4180511     3723952
##  [1,] NA        -1.692308   NA         -0.1       -0.2079692  -1.923077
##  [2,] 1.5       -0.2        1.1        -2.461538  -1.5        1.916667
##  [3,] NA        -0.2754717  -1.818182  -1.144483  0.7         0.6598967
##  [4,] 0.5       -2.2        -1.9       0.5384615  0.5         0.9166667
##  [5,] -1.7      0           NA         -0.5522401 1.9         0.6120461
##  [6,] 0.5       1.8         1.1        1.538462   1.5         -0.08333333
##  [7,] NA        NA          1.923846   -1.101065  0.8333333   NA
##  [8,] 0.5       -0.2        0.1        -2.461538  0.5         -0.08333333
##  [9,] 0.3       -0.8461538  0.8333333  0.1759914  NA          NA
## [10,] -0.5      -1.2        1.1        1.538462   0.5         -2.083333
## [11,] NA        -0.443513   -0.1666667 -0.2612757 0.06538462  -1.909091
## [12,] 0.5       0.8         -0.9       -0.4615385 0.5         -2.083333
## [13,] 1.1       0.4942021   -0.6       -0.1629243 -0.06953642 -0.06829209
## [14,] -1.5      -0.2        0.1        -0.4615385 -0.5        1.916667
## [15,] 1.1       -0.09090909 NA         0.6563657  1.454545    0.1920213
## [16,] 0.5       1.8         -0.9       1.538462   0.5         -2.083333
## [17,] NA        1.909091    -0.4444444 -0.5       0.4         NA
## [18,] -1.5      -2.2        0.1        -0.4615385 -1.5        -0.08333333
## [19,] 0.4545455 -1          0.4        1.461168   0.4545455   -0.8513155
## [20,] 0.5       1.8         0.1        1.538462   -0.5        1.916667
##       3061513   4309450     1468271    2510107     1259411    294984
##  [1,] -0.3636364 -1.2        -0.7821612 0           NA         NA
##  [2,] 1.545455   0.2         0          0           0.2727273  0
##  [3,] NA         -1.3        -1.108108  -0.3471074  NA         -0.6160855
##  [4,] -0.4545455 2.2         0          0           -0.7272727 0
##  [5,] NA         0.9027027   -0.9166667 -0.8        -0.7       -0.4
##  [6,] 0.5454545  1.2         0          0           1.272727   -1
##  [7,] NA         -0.09347826 1.083333   -0.2307692  0          -0.6666667
##  [8,] -0.4545455 -0.8        0          2           1.272727   0
##  [9,] NA         NA          NA         0.25        0.8571429  -0.4907617
## [10,] -1.454545  0.2         1          2           0.2727273  1
## [11,] NA         NA          NA         0.5814402   0.3631285  NA
## [12,] 0.5454545  -0.8        -2         -2          0.2727273  1
## [13,] NA         NA          -0.5       -1.214286   0.1461988  -0.172093
## [14,] -0.4545455 0.2         0          0           1.272727   0
## [15,] NA         0.4908156   -1.7       -0.04255319 NA         0.5136611
## [16,] -0.4545455 -0.8        0          1           -0.7272727 2
## [17,] 0.6363636  NA          -1.727273  0.6         -1.793388  -0.6428571
## [18,] -1.454545  0.2         1          -2          -1.727273  -1
## [19,] NA         -0.6363636  -0.1398417 -0.4        -0.3628186 -1.846154
## [20,] 1.545455   -1.8        0          -1          0.2727273  -2
##       950010     1211966    3146203     5385464    5039919    1178877
##  [1,] 0.1638418  0.7692308  -0.5163099  NA         NA         -1.475224
##  [2,] -0.1       0.4        1.166667    -0.3333333 0.3        0.07142857
##  [3,] -0.778261  1.486276   -0.5112994  -0.75      NA         NA
##  [4,] 0.9        0.4        0.1666667   0.6666667  -0.7       -1.928571
##  [5,] -0.3258065 -1.454545  -0.05909413 0.1437787  1.256618   NA
##  [6,] 0.9        0.4        2.166667    1.666667   0.3        0.07142857
##  [7,] NA         NA         0.180234    NA         -0.8       NA
##  [8,] -0.1       -0.6       0.1666667   -2.333333  0.3        0.07142857
##  [9,] 0.1818182  0.4305985  0.4166667   -0.5       1          1.487578
## [10,] 0.9        -0.6       -0.8333333  0.6666667  -0.7       0.07142857
## [11,] 0.63777    NA         0.7785714   -0.1865355 0.2        -0.2072621
## [12,] -1.1       0.4        -0.8333333  -0.3333333 -0.7       -0.9285714
## [13,] 0.3409091  -1         -0.2732484  -1.090909  NA         NA
## [14,] -1.1       -0.6       -1.833333   0.6666667  -1.7       1.071429
## [15,] NA         0.3333333  -0.3830716  -0.2261985 0.2        NA
## [16,] 1.9        1.4        -0.8333333  0.6666667  0.3        0.07142857
## [17,] NA         0          1.029988    NA         -1.191011  0.25
## [18,] -0.1       -0.6       0.1666667   -0.3333333 2.3        2.071429
## [19,] NA         -0.6666667 0.2059577   -0.2104709 NA         -0.208412
## [20,] -2.1       -0.6       0.1666667   -1.333333  0.3        0.07142857
##       4830630    1325697    3441513     2460925    3375808    505271
##  [1,] 0.4876404  0.09090909 -0.754717   -0.9527505 0.1        1.2
##  [2,] 0.6        0.6        -1          -1.181818  -1.5       0.6363636
##  [3,] -0.9714286 NA         0.425425    1.416667   0.4        NA
##  [4,] 0.6        -0.4       2           1.818182   0.5        -0.3636364
##  [5,] 0.1757188  NA         NA          -0.8725146 NA        -0.07692308
##  [6,] 1.6        0.6        -1          1.818182   -0.5      -1.363636
##  [7,] -0.2486056 -0.3187532 -0.1636364  0.1538462  2.090909  NA
##  [8,] -0.4       0.6        2           -1.181818  1.5       2.636364
##  [9,] NA         NA         -0.4189602  1.8        1.1       1.454545
## [10,] 0.6        1.6        -1          -0.1818182 -1.5      -1.363636
## [11,] NA         NA         -0.08108108 -1.286416  NA        0.6238675
## [12,] -1.4       -1.4       0           -0.1818182 -0.5      -0.3636364
## [13,] NA         NA         NA          -0.8       NA        NA
## [14,] 0.6        -1.4       0           -2.181818  1.5       -0.3636364
## [15,] NA         0.1064641  -0.3055556  -0.3743316 NA        0.1538462
## [16,] -1.4       -1.4       2           0.8181818  -0.5      -1.363636
## [17,] NA         NA         -0.07847401 1.128819   NA        NA
## [18,] 0.6        0.6        -1          0.8181818  1.5       -1.363636
## [19,] -0.3248408 0.1538462  NA          -2.153846  0.2736486 1.03901
## [20,] -1.4       0.6        -2          -0.1818182 -0.5      2.636364
##       3846003     4938413    3146037     4735650    510307     2667507
##  [1,] NA          -1.8       0.9389264   -0.4817606 0.1952381  1.796905
##  [2,] -0.8888889  0.1        -0.8571429  0.6        -0.9       -1.090909
##  [3,] -0.4        NA         -1.149677   0.3        NA         1.5
##  [4,] 1.111111    1.1        0.1428571   0.6        0.1        -0.09090909
##  [5,] 1.083736    0.02249589 -0.75       0.0871407  -0.1       0.3320084
##  [6,] 0.1111111   -1.9       -0.8571429  0.6        1.1        -0.09090909
##  [7,] 0.2986483   -0.3302554 0.4166667   NA         0.6219217  -0.4394869
##  [8,] -1.888889   0.1        1.142857    -1.4       2.1        -0.09090909
##  [9,] NA          NA         1.397105    -0.1318659 NA         0.4739113
## [10,] -0.8888889  -0.9       -1.857143   -1.4       0.1        1.909091
## [11,] NA          0.1        2.166667    NA         0.02536015 0.149233
## [12,] 0.1111111   0.1        -0.8571429  1.6        -1.9       0.9090909
## [13,] -0.4179431  -0.1       -1.434171   0.7777778  0.06928637 -0.03778446
## [14,] 0.1111111   2.1        1.142857    0.6        0.1        -1.090909
## [15,] 0.007946828 0.2        NA          1.454545   0.1666667  0.7519205
## [16,] 0.1111111   -1.9       -0.8571429  -0.4       -0.9       0.9090909
## [17,] NA          0.7        -0.07692308 0.4213852  1.478328   0.7762238
## [18,] 2.111111    1.1        0.1428571   -0.4       0.1        -0.09090909
## [19,] NA          0.1        1.1366      0.6984744  NA         -0.8834951
## [20,] NA          0.1        1.142857    -0.4       0.1        -0.09090909
##       5011601     4715209  4380449    1793895     5268852   2865695
##  [1,] 0.1876364   0.437223 NA         NA          NA        0.4545455
##  [2,] -0.7272727  -1.2     -0.7777778 1.8         0.9       0.1666667
##  [3,] -0.08001594 NA       -0.1945525 -0.4597098  NA        NA
##  [4,] 0.2727273   0.8      1.222222   -0.2        1.9       -0.8333333
##  [5,] -0.8181818  NA       -0.8       -0.08079678 NA        NA
##  [6,] 1.272727    1.8      0.2222222  -1.2        0.9       0.1666667
##  [7,] -0.2279496  -0.432   1          0.2754575   -1.1      NA
##  [8,] 0.2727273   0.8      -0.7777778 -0.2        -0.1      -1.833333
##  [9,] -0.04572018 -0.4     -0.9333333 1.7         NA        -0.2870808
## [10,] -0.7272727  0.8      0.2222222  0.8         -2.1      0.1666667
## [11,] NA          NA       NA         NA          NA        NA
## [12,] -0.7272727  -0.2     1.222222   -0.2        -2.1      0.1666667
## [13,] NA          NA       0.2        -0.08333333 NA        0.06712565
## [14,] -0.7272727  -1.2     0.2222222  -0.2        1.9       0.1666667
## [15,] -0.440194   1        NA         -0.836989   -1.181818 0.4545455
## [16,] 0.2727273   -0.2     0.2222222  -1.2        -1.1      1.166667
## [17,] NA          -0.44    -0.128959  0.25        -0.03375  -1.231017
## [18,] 1.272727    -0.2     -1.777778  -0.2        -2.1      0.1666667
## [19,] 0.4545455   0.437223 NA         0.5053723   1.823395  NA
## [20,] -1.727273   -1.2     NA         0.8         1.9       0.1666667
##       3093794     5647495    4688823     3821129     5058598
##  [1,] -0.3080513  0.3571429  0.5688928   0.5485807   -0.3075939
##  [2,] 1.153846    -0.25      -0.07692308 0.4166667   -0.08333333
##  [3,] 0.08333333  NA         NA          -0.08333333 0
##  [4,] 0.1538462   1.75       -0.07692308 0.4166667   -0.08333333
##  [5,] 1.894277    NA         2.214286    0.3486797   0.1342017
##  [6,] 0.1538462   -0.25      -0.07692308 0.4166667   -0.08333333
##  [7,] -0.02418605 0.25       0.1342277   0.9166667   -0.1818182
##  [8,] -1.846154   -0.25      -0.07692308 -1.583333   1.916667
##  [9,] -0.4691785  NA         -0.1419344  NA          -0.8443129
## [10,] -0.8461538  -1.25      -0.07692308 0.4166667   0.9166667
## [11,] NA          -0.9080882 0.1574623   0.6194639   -0.09450981
## [12,] 0.1538462   1.75       -0.07692308 -0.5833333  -1.083333
## [13,] -0.7255663  0.3571429  -1.303055   1.533163    0.6604278
## [14,] 1.153846    -0.25      -0.07692308 -0.5833333  -1.083333
## [15,] -0.3429305  NA         -0.718608   -0.7042608  NA
## [16,] 0.1538462   -0.25      -0.07692308 -1.583333   1.916667
## [17,] 0.25        0.7777778  0.03099852  -0.9090909  0.2717272
## [18,] 2.153846    -0.25      -0.07692308 0.4166667   -2.083333
## [19,] -0.5454545  NA         0.307097    -0.7748227  0.06666667
## [20,] -0.8461538  -1.25      -0.07692308 0.4166667   -0.08333333
##       995697     3278419    4046972     4935243    2965087    3886540
##  [1,] NA         NA         -0.6641509  0.8        -0.1884298 NA
##  [2,] -0.2       0.2727273  0.4285714   -2         -1.8       -0.4
##  [3,] NA         NA         -1.923077   0          NA         0.1981785
##  [4,] -1.2       -1.727273  0.4285714   0          0.2        0.6
##  [5,] NA         NA         -0.2667142  NA         0.6363636  -1.2
##  [6,] -1.2       0.2727273  -0.5714286  -2         2.2        -0.4
##  [7,] 0.3076923  NA         0.505422    1.017483   NA         0.3
##  [8,] 1.8        2.272727   0.4285714   0          0.2        0.6
##  [9,] -0.8077088 NA         -0.0133323  -0.797654  -1         0.1
## [10,] -0.2       -1.727273  -0.5714286  0          1.2        -0.4
## [11,] 1.216993   NA         0.6289164   NA         2          -1.659375
## [12,] 0.8        -0.7272727 -0.5714286  0          0.2        0.6
## [13,] NA         0.2        -0.8677897  -0.8461538 0.3846154  NA
## [14,] -0.2       1.272727   0.4285714   1          -0.8       -0.4
## [15,] NA         -0.8       -0.05412209 -0.7071683 1.636364   0.08333333
## [16,] -0.2       0.2727273  0.4285714   2          0.2        0.6
## [17,] NA         0.1666667  0.2727273   -1.933333  -0.1666667 NA
## [18,] -0.2       1.272727   0.4285714   0          -1.8       -1.4
## [19,] -0.2727273 NA         -1.9        0.1419726  0.1176849  0.09171271
## [20,] 0.8        -1.727273  1.428571    -1         0.2        0.6
##       662664     2341442    3315379   4976553    270282     2991842
##  [1,] NA         0.1666667  0.6356322 NA         0.3110048  0.3813953
##  [2,] -0.1111111 1.083333   -1.1      -0.2       -0.2727273 0.7333333
##  [3,] 0.09090909 0.2955975  NA        0.3226277  2.214286   -0.9918551
##  [4,] 0.8888889  1.083333   -0.1      0.8        -1.272727  1.733333
##  [5,] -1.833333  -0.4509085 -0.3      NA         1.5        -0.3777778
##  [6,] 0.8888889  0.08333333 0.9       -0.2       -2.272727  0.7333333
##  [7,] NA         -0.4261006 NA        NA         -1.5       -1.692308
##  [8,] -0.1111111 0.08333333 -0.1      -0.2       1.727273   -1.266667
##  [9,] 0.1666667  -0.2117324 NA        NA         0.1196172  2.333333
## [10,] -0.1111111 -0.9166667 1.9       -0.2       -2.272727  1.733333
## [11,] NA         1.888889   NA        NA         2.214286   0.4
## [12,] 1.888889   -0.9166667 -0.1      -0.2       1.727273   0.7333333
## [13,] 0.09090909 -0.157173  NA        -0.5       -0.3333333 0.3553903
## [14,] -1.111111  2.083333   -2.1      0.8        1.727273   0.7333333
## [15,] NA         NA         -1.1      -0.8333333 NA         NA
## [16,] -2.111111  0.08333333 -1.1      0.8        -0.2727273 -1.266667
## [17,] NA         -0.4444444 NA        1.067774   NA         NA
## [18,] -0.1111111 -0.9166667 0.9       -1.2       -0.2727273 -0.2666667
## [19,] NA         NA         0.45666   NA         NA         -0.2478873
## [20,] NA         0.08333333 0.9       -0.2       -0.2727273 -1.266667
##       3984496
##  [1,] -0.808
##  [2,] 0.4
##  [3,] -0.2542912
##  [4,] 0.4
##  [5,] -1.6125
##  [6,] -1.6
##  [7,] 0.1666667
##  [8,] 1.4
##  [9,] 1.4
## [10,] -0.6
## [11,] 0.2844523
## [12,] -1.6
## [13,] 0.4593103
## [14,] 0.4
## [15,] -0.8
## [16,] 0.4
## [17,] -0.4
## [18,] 1.4
## [19,] -0.01893889
## [20,] -0.6
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)), "tests")
```

```
## avg MAE = 1.119629 from 588 tests
```

viewing a histogram of the ratings helps gauge the impact of different MAE’s and decide what the likethresh should be

```
hist(unlist(users[train,]))
```



```
likethresh = 2.5 # insert your own value here
showCM(preds,likethresh)
```

```
## TN=   587 FP=     0
## FN=     1 TP=     0  (total=588)
## accuracy  = 99.8%
## precision = NaN%
## recall    = 0.0%
```

3. Pearson’s Simarity

```
preds = predictCF(users[test,], users[train,], numtestitems=10, random=FALSE, simfun=pearsons
im) ; preds
```

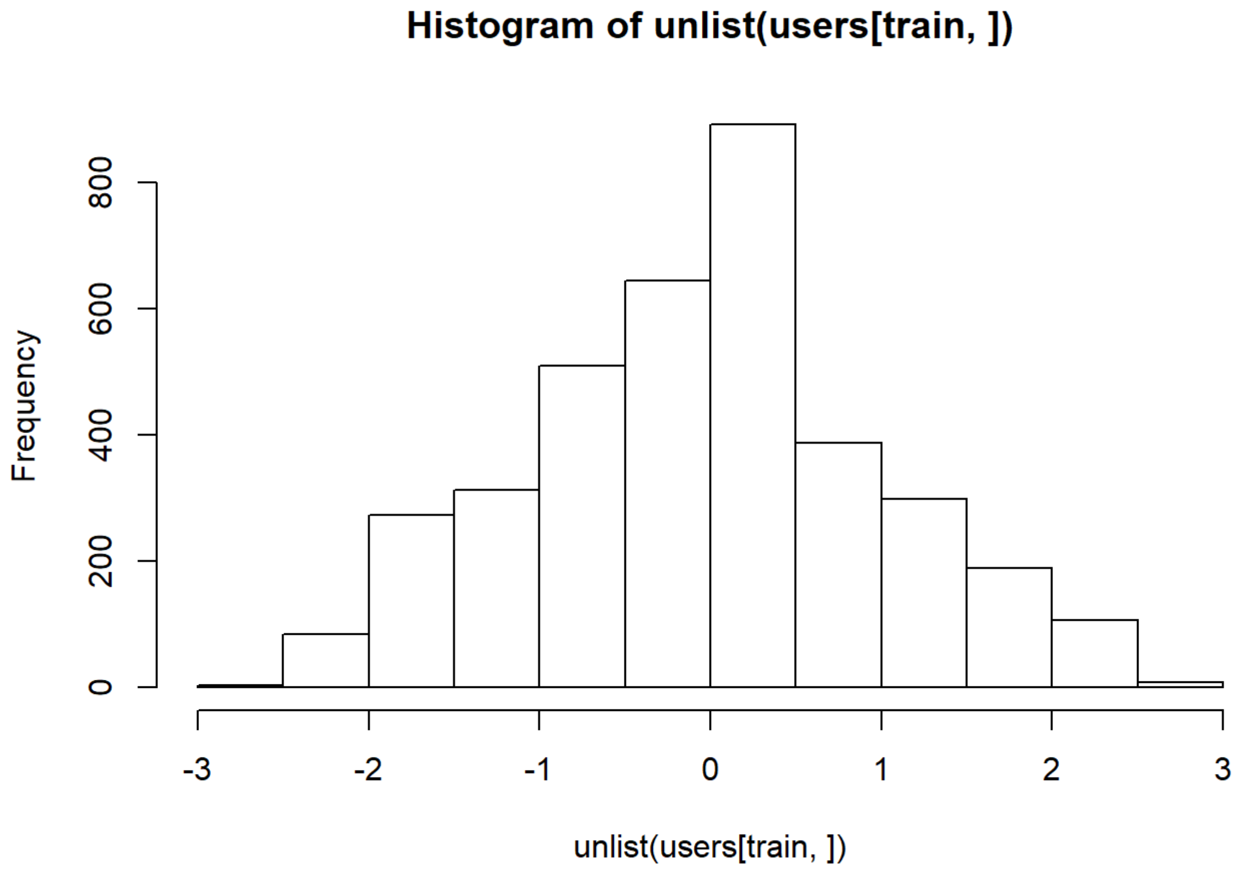
```
## ..................................................................................
##       4364462    4875280 1283321 4117053    1486411 2210060    5495353
##  [1,] NA         NA      NA      NA         NA      NA         NA
##  [2,] 0.09090909 0.3     -1.75   0.7272727  -0.2    1.666667   -1.4
##  [3,] NA         NA      NA      NA         NA      NA         NA
##  [4,] -0.9090909 0.3     0.25    -1.272727  -0.2    -0.3333333 0.6
##  [5,] NA         0.2     NA      NA         NA      NA         NA
##  [6,] -0.9090909 -1.7    -1.75   -0.2727273 -0.2    1.666667   1.6
##  [7,] NA         0.2     NA      NA         NA      NA         NA
##  [8,] 2.090909   -1.7    -0.75   -0.2727273 -1.2    -2.333333  1.6
##  [9,] NA         NA      NA      NA         NA      NA         NA
## [10,] -1.909091  0.3     0.25    -2.272727  1.8     0.6666667  -0.4
## [11,] NA         NA      NA      NA         NA      NA         NA
## [12,] -1.909091  -1.7    0.25    1.727273   -0.2    -0.3333333 -0.4
## [13,] NA         NA      NA      NA         NA      NA         NA
## [14,] 2.090909   1.3     2.25    -0.2727273 -1.2    -0.3333333 0.6
## [15,] NA         NA      NA      NA         NA      NA         NA
## [16,] 0.09090909 0.3     0.25    0.7272727  0.8     -0.3333333 -0.4
## [17,] NA         NA      NA      NA         NA      NA         NA
## [18,] 1.090909   0.3     1.25    0.7272727  -1.2    -0.3333333 -0.4
## [19,] NA         NA      NA      NA         NA      NA         NA
## [20,] -0.9090909 2.3     1.25    -1.272727  1.8     -0.3333333 -1.4
##       5446448 2969181 4314129 1898757 2691000    3399045    2536879
##  [1,] NA      NA      NA      NA      NA         NA         NA
##  [2,] 1.7     0.2     -0.1    0.5     -1.615385  0.07692308 -0.3333333
##  [3,] NA      NA      NA      NA      NA         NA         NA
##  [4,] 0.7     0.2     1.9     1.5     0.3846154  0.07692308 -0.3333333
##  [5,] NA      NA      NA      NA      NA         NA         NA
##  [6,] 1.7     2.2     -1.1    -1.5    0.3846154  0.07692308 1.666667
##  [7,] NA      NA      NA      NA      NA         NA         NA
##  [8,] -0.3    0.2     0.9     1.5     -1.615385  2.076923   -0.3333333
##  [9,] NA      NA      NA      NA      NA         NA         NA
## [10,] -0.3    0.2     -2.1    0.5     -1.615385  -0.9230769 0.6666667
## [11,] NA      NA      NA      NA      NA         NA         NA
## [12,] -0.3    -0.8    -1.1    0.5     0.3846154  -0.9230769 -1.333333
## [13,] NA      NA      NA      NA      NA         NA         NA
## [14,] -0.3    -0.8    -0.1    -0.5    2.384615   -1.923077  1.666667
## [15,] NA      NA      NA      NA      NA         NA         NA
## [16,] 0.7     -0.8    -0.1    0.5     0.3846154  1.076923   -1.333333
## [17,] NA      NA      NA      NA      NA         NA         NA
## [18,] -2.3    0.2     1.9     -1.5    0.3846154  0.07692308 -0.3333333
## [19,] NA      NA      NA      NA      NA         NA         NA
## [20,] -1.3    -0.8    -0.1    -1.5    -0.6153846 -0.9230769 NA
##       4447646    3308880 5039682 5037898 1602026   3631890    3628337
##  [1,] NA         NA      NA      NA      -1.333333 NA         NA
##  [2,] 0.3636364  -0.1    1       1.5     -1.714286 -1.272727  0.3333333
##  [3,] NA         NA      NA      NA      NA        NA         NA
##  [4,] 1.363636   -0.1    -1      -0.5    -1.714286 -2.272727  0.3333333
##  [5,] NA         NA      NA      NA      NA        NA         NA
##  [6,] -1.636364  1.9     -1      -0.5    1.285714  0.7272727  0.3333333
##  [7,] NA         NA      NA      NA      -1.692308 NA         NA
##  [8,] -0.6363636 -1.1    1       -0.5    0.2857143 1.727273   0.3333333
##  [9,] NA         NA      NA      NA      NA        NA         NA
## [10,] -1.636364  0.9     -2      1.5     0.2857143 -0.2727273 -1.666667
## [11,] NA         NA      NA      NA      NA        NA         NA
## [12,] -0.6363636 0.9     0       -1.5    -1.714286 1.727273   -1.666667
## [13,] NA         NA      NA      NA      NA        NA         NA
## [14,] 1.363636   -0.1    0       0.5     1.285714  -1.272727  1.333333
## [15,] NA         NA      NA      NA      -1.692308 NA         NA
## [16,] 1.363636   -2.1    0       -0.5    -1.714286 -0.2727273 0.3333333
## [17,] NA         NA      NA      NA      NA        NA         NA
## [18,] -2.636364  -0.1    1       -1.5    1.285714  1.727273   0.3333333
## [19,] NA         NA      NA      NA      NA        NA         NA
## [20,] 1.363636   -0.1    1       1.5     1.285714  0.7272727  NA
##       5537923 1542832    4046204   2686549 5732915 532034     4412129
##  [1,] NA      NA         NA        NA      NA      NA         NA
##  [2,] 1       -0.1818182 -1        0       -1      0.4545455  2.3
##  [3,] NA      NA         NA        NA      NA      NA         NA
##  [4,] 0       0.8181818  1         1       0       0.4545455  0.3
##  [5,] NA      NA         NA        NA      NA      NA         NA
##  [6,] 0       -0.1818182 0         0       1       1.454545   -0.7
##  [7,] NA      NA         NA        NA      NA      NA         NA
##  [8,] 0       -1.181818  -1        2       1       0.4545455  -0.7
##  [9,] NA      NA         NA        NA      NA      NA         NA
## [10,] -1      -0.1818182 1         0       0       0.4545455  1.3
## [11,] NA      NA         0.2142857 NA      NA      NA         NA
## [12,] -1      -1.181818  0         2       1       -1.545455  0.3
## [13,] NA      NA         NA        NA      NA      NA         NA
## [14,] 0       1.818182   2         -1      -1      -0.5454545 -1.7
## [15,] NA      NA         NA        NA      NA      NA         NA
## [16,] 1       -0.1818182 1         -2      0       -0.5454545 1.3
## [17,] NA      NA         0.575     NA      NA      NA         NA
## [18,] 0       -0.1818182 -2        -1      0       -0.5454545 -0.7
## [19,] NA      NA         0.2142857 NA      NA      NA         NA
## [20,] 0       0.8181818  1         -1      0       -0.5454545 -1.7
##       808981     3281295    4415881 2941133 1539253 4175055    4180511
##  [1,] NA         NA         NA      NA      NA      NA         NA
##  [2,] -0.7272727 -1.363636  1.5     -0.2    1.1     -2.461538  -1.5
##  [3,] NA         NA         NA      NA      NA      NA         NA
##  [4,] 1.272727   -0.3636364 0.5     -2.2    -1.9    0.5384615  0.5
##  [5,] NA         -1.916667  NA      NA      NA      NA         NA
##  [6,] 2.272727   0.6363636  0.5     1.8     1.1     1.538462   1.5
##  [7,] NA         NA         NA      NA      NA      NA         NA
##  [8,] -1.727273  0.6363636  0.5     -0.2    0.1     -2.461538  0.5
##  [9,] NA         NA         NA      NA      NA      NA         NA
## [10,] -0.7272727 0.6363636  -0.5    -1.2    1.1     1.538462   0.5
## [11,] NA         NA         NA      NA      NA      -0.6666667 NA
## [12,] -1.727273  0.6363636  0.5     0.8     -0.9    -0.4615385 0.5
## [13,] NA         NA         NA      NA      NA      NA         NA
## [14,] -1.727273  -1.363636  -1.5    -0.2    0.1     -0.4615385 -0.5
## [15,] NA         NA         NA      NA      NA      0          NA
## [16,] 0.2727273  0.6363636  0.5     1.8     -0.9    1.538462   0.5
## [17,] NA         NA         NA      NA      NA      NA         NA
## [18,] 0.2727273  -0.3636364 -1.5    -2.2    0.1     -0.4615385 -1.5
## [19,] NA         NA         NA      NA      NA      NA         NA
## [20,] 2.272727   -0.3636364 0.5     1.8     0.1     1.538462   -0.5
##       3723952     3061513    4309450 1468271 2510107 1259411    294984
##  [1,] NA          NA         NA      NA      NA      NA         NA
##  [2,] 1.916667    1.545455   0.2     0       0       0.2727273  0
##  [3,] NA          NA         NA      NA      NA      NA         NA
##  [4,] 0.9166667   -0.4545455 2.2     0       0       -0.7272727 0
##  [5,] NA          NA         NA      NA      NA      NA         NA
##  [6,] -0.08333333 0.5454545  1.2     0       0       1.272727   -1
##  [7,] NA          NA         NA      NA      NA      NA         NA
##  [8,] -0.08333333 -0.4545455 -0.8    0       2       1.272727   0
##  [9,] NA          NA         NA      NA      NA      NA         -0.6428571
## [10,] -2.083333   -1.454545  0.2     1       2       0.2727273  1
## [11,] NA          NA         NA      NA      NA      NA         NA
## [12,] -2.083333   0.5454545  -0.8    -2      -2      0.2727273  1
## [13,] NA          NA         NA      NA      NA      NA         NA
## [14,] 1.916667    -0.4545455 0.2     0       0       1.272727   0
## [15,] NA          NA         NA      NA      NA      NA         NA
## [16,] -2.083333   -0.4545455 -0.8    0       1       -0.7272727 2
## [17,] NA          NA         NA      NA      NA      NA         -0.6428571
## [18,] -0.08333333 -1.454545  0.2     1       -2      -1.727273  -1
## [19,] NA          NA         NA      NA      NA      NA         NA
## [20,] 1.916667    1.545455   -1.8    0       -1      0.2727273  -2
##       950010 1211966   3146203    5385464    5039919 1178877    4830630
##  [1,] NA     NA        NA         NA         NA      NA         NA
##  [2,] -0.1   0.4       1.166667   -0.3333333 0.3     0.07142857 0.6
##  [3,] NA     NA        NA         NA         NA      NA         NA
##  [4,] 0.9    0.4       0.1666667  0.6666667  -0.7    -1.928571  0.6
##  [5,] NA     NA        NA         NA         NA      NA         NA
##  [6,] 0.9    0.4       2.166667   1.666667   0.3     0.07142857 1.6
##  [7,] NA     NA        NA         NA         NA      NA         NA
##  [8,] -0.1   -0.6      0.1666667  -2.333333  0.3     0.07142857 -0.4
##  [9,] NA     NA        NA         NA         NA      NA         NA
## [10,] 0.9    -0.6      -0.8333333 0.6666667  -0.7    0.07142857 0.6
## [11,] NA     NA        NA         -0.8333333 NA      NA         NA
## [12,] -1.1   0.4       -0.8333333 -0.3333333 -0.7    -0.9285714 -1.4
## [13,] NA     NA        NA         NA         NA      NA         NA
## [14,] -1.1   -0.6      -1.833333  0.6666667  -1.7    1.071429   0.6
## [15,] NA     0.3333333 NA         NA         NA      NA         NA
## [16,] 1.9    1.4       -0.8333333 0.6666667  0.3     0.07142857 -1.4
## [17,] NA     NA        NA         NA         NA      NA         NA
## [18,] -0.1   -0.6      0.1666667  -0.3333333 2.3     2.071429   0.6
## [19,] NA     NA        NA         -0.8333333 NA      NA         NA
## [20,] -2.1   -0.6      0.1666667  -1.333333  0.3     0.07142857 -1.4
##       1325697 3441513 2460925    3375808 505271     3846003    4938413
##  [1,] NA      NA      NA         NA      NA         NA         NA
##  [2,] 0.6     -1      -1.181818  -1.5    0.6363636  -0.8888889 0.1
##  [3,] NA      NA      NA         NA      NA         NA         NA
##  [4,] -0.4    2       1.818182   0.5     -0.3636364 1.111111   1.1
##  [5,] NA      NA      NA         NA      NA         NA         NA
##  [6,] 0.6     -1      1.818182   -0.5    -1.363636  0.1111111  -1.9
##  [7,] NA      NA      NA         NA      NA         NA         NA
##  [8,] 0.6     2       -1.181818  1.5     2.636364   -1.888889  0.1
##  [9,] NA      NA      NA         NA      NA         NA         NA
## [10,] 1.6     -1      -0.1818182 -1.5    -1.363636  -0.8888889 -0.9
## [11,] NA      NA      NA         NA      NA         NA         NA
## [12,] -1.4    0       -0.1818182 -0.5    -0.3636364 0.1111111  0.1
## [13,] NA      NA      NA         NA      NA         NA         NA
## [14,] -1.4    0       -2.181818  1.5     -0.3636364 0.1111111  2.1
## [15,] NA      NA      NA         NA      NA         NA         NA
## [16,] -1.4    2       0.8181818  -0.5    -1.363636  0.1111111  -1.9
## [17,] NA      NA      NA         NA      NA         NA         NA
## [18,] 0.6     -1      0.8181818  1.5     -1.363636  2.111111   1.1
## [19,] NA      NA      NA         NA      NA         NA         NA
## [20,] 0.6     -2      -0.1818182 -0.5    2.636364   NA         0.1
##       3146037    4735650   510307 2667507     5011601    4715209
##  [1,] NA         NA        NA     NA          NA         NA
##  [2,] -0.8571429 0.6       -0.9   -1.090909   -0.7272727 -1.2
##  [3,] NA         NA        NA     NA          NA         NA
##  [4,] 0.1428571  0.6       0.1    -0.09090909 0.2727273  0.8
##  [5,] NA         NA        NA     NA          NA         NA
##  [6,] -0.8571429 0.6       1.1    -0.09090909 1.272727   1.8
##  [7,] NA         NA        NA     NA          NA         NA
##  [8,] 1.142857   -1.4      2.1    -0.09090909 0.2727273  0.8
##  [9,] NA         NA        NA     NA          NA         NA
## [10,] -1.857143  -1.4      0.1    1.909091    -0.7272727 0.8
## [11,] NA         NA        NA     NA          NA         NA
## [12,] -0.8571429 1.6       -1.9   0.9090909   -0.7272727 -0.2
## [13,] NA         NA        NA     NA          NA         NA
## [14,] 1.142857   0.6       0.1    -1.090909   -0.7272727 -1.2
## [15,] NA         NA        NA     NA          NA         NA
## [16,] -0.8571429 -0.4      -0.9   0.9090909   0.2727273  -0.2
## [17,] NA         NA        NA     NA          NA         NA
## [18,] 0.1428571  -0.4      0.1    -0.09090909 1.272727   -0.2
## [19,] NA         0.4545455 NA     NA          NA         NA
## [20,] 1.142857   -0.4      0.1    -0.09090909 -1.727273  -1.2
##       4380449    1793895 5268852 2865695    3093794    5647495 4688823
##  [1,] NA         NA      NA      NA         NA         NA      NA
##  [2,] -0.7777778 1.8     0.9     0.1666667  1.153846   -0.25   -0.07692308
##  [3,] NA         NA      NA      NA         NA         NA      NA
##  [4,] 1.222222   -0.2    1.9     -0.8333333 0.1538462  1.75    -0.07692308
##  [5,] NA         NA      NA      NA         NA         NA      NA
##  [6,] 0.2222222  -1.2    0.9     0.1666667  0.1538462  -0.25   -0.07692308
##  [7,] NA         NA      NA      NA         NA         NA      NA
##  [8,] -0.7777778 -0.2    -0.1    -1.833333  -1.846154  -0.25   -0.07692308
##  [9,] NA         NA      NA      NA         NA         NA      NA
## [10,] 0.2222222  0.8     -2.1    0.1666667  -0.8461538 -1.25   -0.07692308
## [11,] NA         NA      NA      NA         NA         NA      NA
## [12,] 1.222222   -0.2    -2.1    0.1666667  0.1538462  1.75    -0.07692308
## [13,] NA         NA      NA      NA         NA         NA      NA
## [14,] 0.2222222  -0.2    1.9     0.1666667  1.153846   -0.25   -0.07692308
## [15,] NA         NA      NA      NA         NA         NA      NA
## [16,] 0.2222222  -1.2    -1.1    1.166667   0.1538462  -0.25   -0.07692308
## [17,] NA         NA      NA      NA         NA         NA      NA
## [18,] -1.777778  -0.2    -2.1    0.1666667  2.153846   -0.25   -0.07692308
## [19,] NA         NA      NA      NA         NA         NA      NA
## [20,] NA         0.8     1.9     0.1666667  -0.8461538 -1.25   -0.07692308
##       3821129    5058598     995697 3278419    4046972    4935243
##  [1,] NA         NA          NA     NA         NA         NA
##  [2,] 0.4166667  -0.08333333 -0.2   0.2727273  0.4285714  -2
##  [3,] NA         NA          NA     NA         NA         NA
##  [4,] 0.4166667  -0.08333333 -1.2   -1.727273  0.4285714  0
##  [5,] NA         NA          NA     NA         NA         NA
##  [6,] 0.4166667  -0.08333333 -1.2   0.2727273  -0.5714286 -2
##  [7,] NA         NA          NA     NA         NA         NA
##  [8,] -1.583333  1.916667    1.8    2.272727   0.4285714  0
##  [9,] NA         NA          NA     NA         NA         NA
## [10,] 0.4166667  0.9166667   -0.2   -1.727273  -0.5714286 0
## [11,] NA         NA          NA     NA         NA         NA
## [12,] -0.5833333 -1.083333   0.8    -0.7272727 -0.5714286 0
## [13,] NA         NA          NA     NA         NA         NA
## [14,] -0.5833333 -1.083333   -0.2   1.272727   0.4285714  1
## [15,] NA         NA          NA     NA         NA         NA
## [16,] -1.583333  1.916667    -0.2   0.2727273  0.4285714  2
## [17,] NA         NA          NA     NA         NA         NA
## [18,] 0.4166667  -2.083333   -0.2   1.272727   0.4285714  0
## [19,] NA         NA          NA     NA         NA         0.06666667
## [20,] 0.4166667  -0.08333333 0.8    -1.727273  1.428571   -1
##       2965087 3886540 662664     2341442    3315379 4976553 270282
##  [1,] NA      NA      NA         NA         NA      NA      NA
##  [2,] -1.8    -0.4    -0.1111111 1.083333   -1.1    -0.2    -0.2727273
##  [3,] NA      NA      NA         NA         NA      NA      NA
##  [4,] 0.2     0.6     0.8888889  1.083333   -0.1    0.8     -1.272727
##  [5,] NA      NA      NA         NA         NA      NA      NA
##  [6,] 2.2     -0.4    0.8888889  0.08333333 0.9     -0.2    -2.272727
##  [7,] NA      NA      NA         NA         NA      NA      NA
##  [8,] 0.2     0.6     -0.1111111 0.08333333 -0.1    -0.2    1.727273
##  [9,] NA      NA      NA         NA         NA      NA      NA
## [10,] 1.2     -0.4    -0.1111111 -0.9166667 1.9     -0.2    -2.272727
## [11,] NA      NA      NA         NA         NA      NA      NA
## [12,] 0.2     0.6     1.888889   -0.9166667 -0.1    -0.2    1.727273
## [13,] NA      NA      NA         NA         NA      NA      NA
## [14,] -0.8    -0.4    -1.111111  2.083333   -2.1    0.8     1.727273
## [15,] NA      NA      NA         NA         NA      NA      NA
## [16,] 0.2     0.6     -2.111111  0.08333333 -1.1    0.8     -0.2727273
## [17,] NA      NA      NA         NA         NA      NA      NA
## [18,] -1.8    -1.4    -0.1111111 -0.9166667 0.9     -1.2    -0.2727273
## [19,] NA      NA      NA         NA         NA      NA      NA
## [20,] 0.2     0.6     NA         0.08333333 0.9     -0.2    -0.2727273
##       2991842    3984496
##  [1,] NA         NA
##  [2,] 0.7333333  0.4
##  [3,] NA         NA
##  [4,] 1.733333   0.4
##  [5,] NA         NA
##  [6,] 0.7333333  -1.6
##  [7,] NA         NA
##  [8,] -1.266667  1.4
##  [9,] NA         NA
## [10,] 1.733333   -0.6
## [11,] NA         NA
## [12,] 0.7333333  -1.6
## [13,] NA         NA
## [14,] 0.7333333  0.4
## [15,] NA         NA
## [16,] -1.266667  0.4
## [17,] NA         NA
## [18,] -0.2666667 1.4
## [19,] NA         NA
## [20,] -1.266667  -0.6
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)), "tests")
```

```
## avg MAE = 1.113358 from 18 tests
```

viewing a histogram of the ratings helps gauge the impact of different MAE’s and decide what the likethresh should be

```
hist(unlist(users[train,]))
```



```
likethresh = 2.5 # insert your own value here
showCM(preds,likethresh)
```

```
## TN=    18 FP=     0
## FN=     0 TP=     0  (total=18)
## accuracy  = 100.0%
## precision = NaN%
## recall    = NaN%
```

3. jacard Simarity

```
preds = predictCF(users[test,], users[train,], numtestitems=10, random=FALSE, simfun=jacardsi
m) ; preds
```

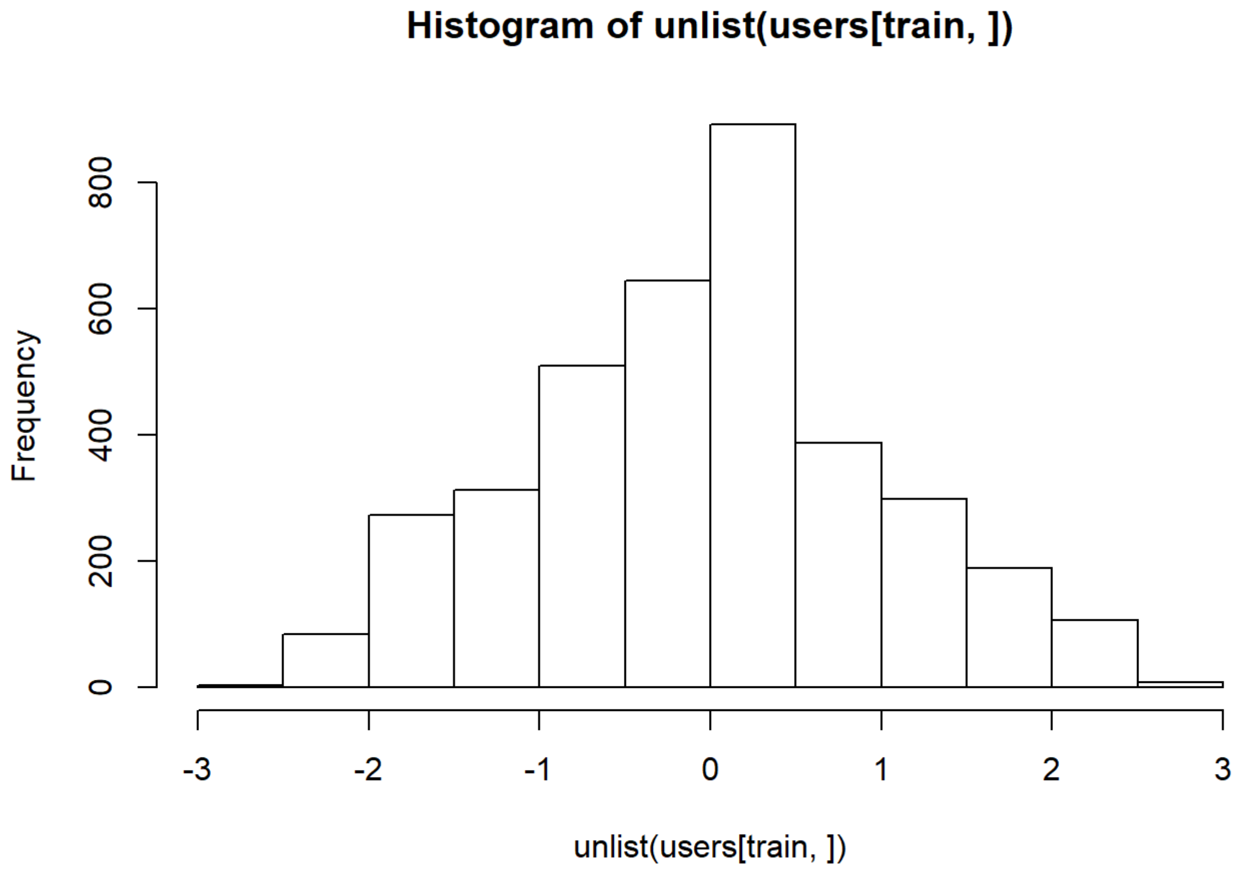
```
## .....................................................................................
##       4364462    4875280     1283321    4117053     1486411    2210060
##  [1,] 1.384615   -0.9051724  0.35       -0.611975   0.1666667  -0.3753278
##  [2,] 0.09090909 0.3         -1.75      0.7272727   -0.2       1.666667
##  [3,] NA         NA          -1.692308  -0.8        NA         -0.9166667
##  [4,] -0.9090909 0.3         0.25       -1.272727   -0.2       -0.3333333
##  [5,] NA         -0.09142857 -1.3       -0.6342342  0.07142857 -0.781653
##  [6,] -0.9090909 -1.7        -1.75      -0.2727273  -0.2       1.666667
##  [7,] NA         -0.1330409  0.09090909 -0.09090909 0.1369048  2.083333
##  [8,] 2.090909   -1.7        -0.75      -0.2727273  -1.2       -2.333333
##  [9,] NA         0.1666667   2.307692   NA          -0.8333333 NA
## [10,] -1.909091  0.3         0.25       -2.272727   1.8        0.6666667
## [11,] NA         -0.1751479  -0.4968254 2.090909    NA         0.8
## [12,] -1.909091  -1.7        0.25       1.727273    -0.2       -0.3333333
## [13,] 0.58       0.1710526   -1.454545  0.4269316   -0.6058646 0
## [14,] 2.090909   1.3         2.25       -0.2727273  -1.2       -0.3333333
## [15,] NA         0.1538462   0.1666667  0.2         0.1818182  1.199147
## [16,] 0.09090909 0.3         0.25       0.7272727   0.8        -0.3333333
## [17,] 1          2.2         0.65       1.048879    -0.2       -0.2
## [18,] 1.090909   0.3         1.25       0.7272727   -1.2       -0.3333333
## [19,] -0.6153846 -0.9583333  -0.3       NA          2.166667   NA
## [20,] -0.9090909 2.3         1.25       -1.272727   1.8        -0.3333333
##       5495353     5446448     2969181     4314129     1898757
##  [1,] NA          -0.4618995  -0.5150094  -0.3733728  NA
##  [2,] -1.4        1.7         0.2         -0.1        0.5
##  [3,] -0.45       0.2732328   0.1363636   -0.7858974  NA
##  [4,] 0.6         0.7         0.2         1.9         1.5
##  [5,] 0.3         -0.07692308 0           -0.6928994  0.8947368
##  [6,] 1.6         1.7         2.2         -1.1        -1.5
##  [7,] -0.1756049  -0.2727273  0.9090909   -0.1        1.7
##  [8,] 1.6         -0.3        0.2         0.9         1.5
##  [9,] -0.8333333  -0.07692308 0.1666667   0.8         -0.4906634
## [10,] -0.4        -0.3        0.2         -2.1        0.5
## [11,] -0.5838462  -0.6990172  NA          0.65        0.524026
## [12,] -0.4        -0.3        -0.8        -1.1        0.5
## [13,] 0.168595    -0.1246036  NA          -0.02179487 0.4
## [14,] 0.6         -0.3        -0.8        -0.1        -0.5
## [15,] -0.04508066 0.4         -0.09090909 1.5         NA
## [16,] -0.4        0.7         -0.8        -0.1        0.5
## [17,] -0.1778917  NA          NA          0.8         0
## [18,] -0.4        -2.3        0.2         1.9         -1.5
## [19,] 0.06945863  NA          0.5085664   -0.15       NA
## [20,] -1.4        -1.3        -0.8        -0.1        -1.5
##       2691000     3399045    2536879    4447646    3308880    5039682
##  [1,] -0.00559542 0.4166667  NA         -0.4532124 0.6363636  NA
##  [2,] -1.615385   0.07692308 -0.3333333 0.3636364  -0.1       1
##  [3,] 0           2.3        -1.014553  NA         -0.4       NA
##  [4,] 0.3846154   0.07692308 -0.3333333 1.363636   -0.1       -1
##  [5,] -1          0          -0.5       0.1        NA         -2.076923
##  [6,] 0.3846154   0.07692308 1.666667   -1.636364  1.9        -1
##  [7,] -0.75       NA         NA         NA         0.4090909  -0.7272727
##  [8,] -1.615385   2.076923   -0.3333333 -0.6363636 -1.1       1
##  [9,] 0           1.545455   NA         NA         -0.1       NA
## [10,] -1.615385   -0.9230769 0.6666667  -1.636364  0.9        -2
## [11,] 0.05        NA         NA         0.09090909 0.6363636  1.272727
## [12,] 0.3846154   -0.9230769 -1.333333  -0.6363636 0.9        0
## [13,] -0.6        NA         NA         NA         -0.5444043 NA
## [14,] 2.384615    -1.923077  1.666667   1.363636   -0.1       0
## [15,] NA          NA         0.5168691  -0.5454545 1.045455   NA
## [16,] 0.3846154   1.076923   -1.333333  1.363636   -2.1       0
## [17,] -1          -1         NA         NA         1.349675   -0.07692308
## [18,] 0.3846154   0.07692308 -0.3333333 -2.636364  -0.1       1
## [19,] NA          NA         NA         1.307692   0.6973684  NA
## [20,] -0.6153846  -0.9230769 NA         1.363636   -0.1       1
##       5037898     1602026    3631890    3628337    5537923    1542832
##  [1,] NA          -0.7788835 0.3760104  -0.95842   NA         -0.09090909
##  [2,] 1.5         -1.714286  -1.272727  0.3333333  1          -0.1818182
##  [3,] NA          NA         1.384615   0.25       -1.9       0
##  [4,] -0.5        -1.714286  -2.272727  0.3333333  0          0.8181818
##  [5,] NA          0.2905365  -1.7       NA         NA         0.8461538
##  [6,] -0.5        1.285714   0.7272727  0.3333333  0          -0.1818182
##  [7,] -1.363636   0.2475296  0.5        0.35       -0.3802198 1
##  [8,] -0.5        0.2857143  1.727273   0.3333333  0          -1.181818
##  [9,] -1.363636   0.05860806 -0.3040752 NA         NA         0.448477
## [10,] 1.5         0.2857143  -0.2727273 -1.666667  -1         -0.1818182
## [11,] -0.07596372 NA         -0.2307984 0.5454545  0.6428571  0.3527074
## [12,] -1.5        -1.714286  1.727273   -1.666667  -1         -1.181818
## [13,] NA          -0.7268657 0.6216216  -0.8122741 0.08333333 0.5185741
## [14,] 0.5         1.285714   -1.272727  1.333333   0          1.818182
## [15,] NA          -0.1801111 0.7227191  0.3846154  NA         -0.4
## [16,] -0.5        -1.714286  -0.2727273 0.3333333  1          -0.1818182
## [17,] -1          0.7272727  NA         1.25       -0.2055035 NA
## [18,] -1.5        1.285714   1.727273   0.3333333  0          -0.1818182
## [19,] 0.01242403  0.7123839  0.4114022  NA         0.1        -1.427273
## [20,] 1.5         1.285714   0.7272727  NA         0          0.8181818
##       4046204     2686549    5732915   532034      4412129    808981
##  [1,] -0.7857143  0          NA        0.5         -0.2018182 1.3
##  [2,] -1          0          -1        0.4545455   2.3        -0.7272727
##  [3,] -0.7261596  NA         NA        -0.5        -0.4142342 1.2
##  [4,] 1           1          0         0.4545455   0.3        1.272727
##  [5,] NA          -0.6923077 0.3076923 NA          -0.1       NA
##  [6,] 0           0          1         1.454545    -0.7       2.272727
##  [7,] 0.3535188   NA         NA        -0.07925408 0.1666667  0.7365967
##  [8,] -1          2          1         0.4545455   -0.7       -1.727273
##  [9,] NA          -1.692308  1.307692  1.8         NA         1.058619
## [10,] 1           0          0         0.4545455   1.3        -0.7272727
## [11,] -0.4431349  NA         2.25      NA          -1.338084  0.5714286
## [12,] 0           2          1         -1.545455   0.3        -1.727273
## [13,] 0.3284876   2          -0.75     NA          0.6        -0.7253788
## [14,] 2           -1         -1        -0.5454545  -1.7       -1.727273
## [15,] -0.1983724  NA         NA        NA          0.1666667  NA
## [16,] 1           -2         0         -0.5454545  1.3        0.2727273
## [17,] -0.05799158 NA         NA        0.6363636   1.424852   0.1841667
## [18,] -2          -1         0         -0.5454545  -0.7       0.2727273
## [19,] -0.2272727  NA         NA        NA          -0.5692308 -0.4090909
## [20,] 1           -1         0         -0.5454545  -1.7       2.272727
##       3281295     4415881   2941133    1539253     4175055    4180511
##  [1,] -0.5025728  NA        -1.692308  NA          -0.1       -0.1
##  [2,] -1.363636   1.5       -0.2       1.1         -2.461538  -1.5
##  [3,] NA          NA        -0.2628253 -1.818182   -0.7116967 0.7
##  [4,] -0.3636364  0.5       -2.2       -1.9        0.5384615  0.5
##  [5,] -1.186782   -1.7      0          NA          -0.3602557 1.9
##  [6,] 0.6363636   0.5       1.8        1.1         1.538462   1.5
##  [7,] -0.1958855  NA        NA         1.851291    -0.3372781 0.8333333
##  [8,] 0.6363636   0.5       -0.2       0.1         -2.461538  0.5
##  [9,] -0.02596066 0.3       -0.8461538 0.8333333   0.2124927  NA
## [10,] 0.6363636   -0.5      -1.2        1.1        1.538462   0.5
## [11,] 0.1494174   NA        -0.2289941  -0.1666667 -0.471875  0.1
## [12,] 0.6363636   0.5       0.8         -0.9       -0.4615385 0.5
## [13,] 0.6681449   1.1       0.3014056   -0.6       -0.2032533 0.02631579
## [14,] -1.363636   -1.5      -0.2        0.1        -0.4615385 -0.5
## [15,] -0.2447673  1.1       -0.09090909 NA         0.1766695  1.454545
## [16,] 0.6363636   0.5       1.8         -0.9       1.538462   0.5
## [17,] -0.6018809  NA        1.909091    -0.4444444 -0.5       0.4
## [18,] -0.3636364  -1.5      -2.2        0.1        -0.4615385 -1.5
## [19,] 0.5454545   0.4545455 -1          0.4        1.238363   0.4545455
## [20,] -0.3636364  0.5       1.8         0.1        1.538462   -0.5
##        3723952     3061513    4309450    1468271    2510107    1259411
##  [1,] -1.923077   -0.3636364 -1.2       -0.4090909 0          NA
##  [2,] 1.916667    1.545455   0.2        0          0          0.2727273
##  [3,] 0.4153638   NA         -1.3       -0.9090909 -0.4341085 NA
##  [4,] 0.9166667   -0.4545455 2.2        0          0          -0.7272727
##  [5,] 0.3666297   NA         0.8390625  -0.9166667 -0.8       -0.7
##  [6,] -0.08333333 0.5454545  1.2        0          0          1.272727
##  [7,] NA          NA         -0.186455  1.083333   -0.2083333 0
##  [8,] -0.08333333 -0.4545455 -0.8       0          2          1.272727
##  [9,] NA          NA         NA         NA         0.25       0.4
## [10,] -2.083333   -1.454545  0.2        1          2          0.2727273
## [11,] -1.909091   NA         NA         NA         0.5755857  0.4
## [12,] -2.083333   0.5454545  -0.8       -2         -2         0.2727273
## [13,] 0.3025045   NA         NA         -0.5       -0.8421053 0.09693878
## [14,] 1.916667    -0.4545455 0.2        0          0          1.272727
## [15,] 0.35        NA         0.4599303  -1.7       -0.1761787 NA
## [16,] -2.083333   -0.4545455 -0.8       0          1          -0.7272727
## [17,] NA          0.6363636  NA         -1.727273  0.6        -1.835714
## [18,] -0.08333333 -1.454545  0.2        1          -2         -1.727273
## [19,] -1.03696    NA         -0.6363636 -0.3733564 -0.4       -0.4506803
## [20,] 1.916667    1.545455   -1.8       0          -1         0.2727273
##        294984      950010      1211966    3146203     5385464    5039919
##  [1,] NA          0.1818182   0.7692308  -0.5460317  NA         NA
##  [2,] 0           -0.1        0.4        1.166667    -0.3333333 0.3
##  [3,] -0.6536684  -0.7086369  1.426573   -0.5174419  -0.75      NA
##  [4,] 0           0.9         0.4        0.1666667   0.6666667  -0.7
##  [5,] -0.4        -0.02432432 -1.454545  -0.05232558 0.127864   1.323833
##  [6,] -1          0.9         0.4        2.166667    1.666667   0.3
##  [7,] -0.5850951  NA          NA         -0.4466111  NA         -0.8
##  [8,] 0           -0.1        -0.6       0.1666667   -2.333333  0.3
##  [9,] -0.5533835  0.1818182   0.3961841  0.4166667   -0.5       1
## [10,] 1           0.9         -0.6       -0.8333333  0.6666667  -0.7
## [11,] NA          0.4376877   NA         0.8         -0.3611111 0.2
## [12,] 1           -1.1        0.4        -0.8333333  -0.3333333 -0.7
## [13,] -0.02179487 0.4285714   -1         -0.4709154  -1.090909  NA
## [14,] 0           -1.1        -0.6       -1.833333   0.6666667  -1.7
## [15,] 0.3605519   NA          0.3333333  -0.4396825  -1.058874  0.2
## [16,] 2           1.9         1.4        -0.8333333  0.6666667  0.3
## [17,] -0.6428571  NA          0          1.125575    NA         -1.236364
## [18,] -1          -0.1        -0.6       0.1666667   -0.3333333 2.3
## [19,] -1.846154   NA          -0.6666667 -0.1621693  -0.4431187 NA
## [20,] -2          -2.1        -0.6       0.1666667   -1.333333  0.3
##        1178877    4830630     1325697    3441513     2460925    3375808
##  [1,] -1.003588  0.8970414   0.09090909 -0.95       -0.5384615 0.1
##  [2,] 0.07142857 0.6         0.6        -1          -1.181818  -1.5
##  [3,] NA         -1.15       NA         0.2039886   1.416667   0.4
##  [4,] -1.928571  0.6         -0.4       2           1.818182   0.5
##  [5,] NA         0.2236842   NA         NA          -0.85      NA
##  [6,] 0.07142857 1.6         0.6        -1          1.818182   -0.5
##  [7,] NA         -0.03947368 -0.1575207 -0.15       0.1538462  2.090909
##  [8,] 0.07142857 -0.4        0.6        2           -1.181818  1.5
##  [9,] 1.4        NA          NA         -0.525      1.8        1.1
## [10,] 0.07142857 0.6         1.6        -1          -0.1818182 -1.5
## [11,] -0.2215279 NA          NA         -0.04545455 -0.3009381 NA
## [12,] -0.9285714 -1.4        -1.4       0           -0.1818182 -0.5
## [13,] NA         NA          NA         NA          -0.8       NA
## [14,] 1.071429   0.6         -1.4       0           -2.181818  1.5
## [15,] NA         NA          -0.2347494 0.1363636   -0.3076923 NA
## [16,] 0.07142857 -1.4        -1.4       2           0.8181818  -0.5
## [17,] 0.25       NA          NA         0.06827539  1.015936   NA
## [18,] 2.071429   0.6         0.6        -1          0.8181818  1.5
## [19,] -0.382074  -0.1568047  0.1538462  NA          -2.153846  0.2496314
## [20,] 0.07142857 -1.4        0.6        -2          -0.1818182 -0.5
##       505271      3846003    4938413    3146037     4735650    510307
##  [1,] 1.2         NA         -1.8       0.6666667   -0.0628655 0.25
##  [2,] 0.6363636   -0.8888889 0.1        -0.8571429  0.6        -0.9
##  [3,] NA          -0.4       NA         -0.9748104  0.3        NA
##  [4,] -0.3636364  1.111111   1.1        0.1428571   0.6        0.1
##  [5,] -0.07692308 1.185185   0.3071006  -0.75       0.08721834 -0.1
##  [6,] -1.363636   0.1111111  -1.9       -0.8571429  0.6        1.1
##  [7,] NA          0.1797297  -0.3390533 0.4166667   NA         0.6886169
##  [8,] 2.636364    -1.888889  0.1        1.142857    -1.4       2.1
##  [9,] 1.454545    NA         NA         1.035294    0.1331169  NA
## [10,] -1.363636   -0.8888889 -0.9       -1.857143   -1.4       0.1
## [11,] 0.1062271   NA         0.1        2.166667    NA         0.1646154
## [12,] -0.3636364  0.1111111  0.1        -0.8571429  1.6        -1.9
## [13,] NA          -0.35      -0.1       -1.291191   0.7777778  0.2460738
## [14,] -0.3636364  0.1111111  2.1        1.142857    0.6        0.1
## [15,] 0.1538462   0.06152515 0.2        NA          1.454545   0.1666667
## [16,] -1.363636   0.1111111  -1.9       -0.8571429  -0.4       -0.9
## [17,] NA          NA         0.7        -0.07692308 0.5787879  1.484615
## [18,] -1.363636   2.111111   1.1        0.1428571   -0.4       0.1
## [19,] 0.8351648   NA         0.1        1.078022    0.5664336  NA
## [20,] 2.636364    NA         0.1        1.142857    -0.4       0.1
##       2667507     5011601     4715209    4380449    1793895     5268852
##  [1,] 1.778555    -0.06593407 0.4112426  NA         NA          NA
##  [2,] -1.090909   -0.7272727  -1.2       -0.7777778 1.8         0.9
##  [3,] 1.5         0.04559355  NA         -0.1818182 -0.6893939  NA
##  [4,] -0.09090909 0.2727273   0.8        1.222222   -0.2        1.9
##  [5,] 0.333199    -0.8181818  NA         -0.8       -0.162943   NA
##  [6,] -0.09090909 1.272727    1.8        0.2222222  -1.2        0.9
##  [7,] -0.6060606  -0.2137443  -0.6184211 1          0.09431818  -1.1
##  [8,] -0.09090909 0.2727273   0.8        -0.7777778 -0.2        -0.1
##  [9,] 0.4165123   -0.06896552 -0.4       -0.9333333 1.7         NA
## [10,] 1.909091    -0.7272727  0.8        0.2222222  0.8         -2.1
## [11,] 0.196038    NA          NA         NA         NA          NA
## [12,] 0.9090909   -0.7272727  -0.2       1.222222   -0.2        -2.1
## [13,] -0.05204082 NA          NA         0.2        -0.08333333 NA
## [14,] -1.090909   -0.7272727  -1.2       0.2222222  -0.2        1.9
## [15,] 0.5454545   -0.318306   1          NA         -0.6738139  -1.181818
## [16,] 0.9090909   0.2727273   -0.2       0.2222222  -1.2        -1.1
## [17,] 0.8384354   NA          -0.375     -0.17      0.25        -0.2281469
## [18,] -0.09090909 1.272727    -0.2       -1.777778  -0.2        -2.1
## [19,] -0.75       0.4545455   0.3343195  NA         0.5229373   1.691716
## [20,] -0.09090909 -1.727273   -1.2       NA         0.8         1.9
##       2865695    3093794    5647495    4688823     3821129     5058598
##  [1,] 0.4545455  -0.5853659 0.3571429  0.6779357   0.4525727   -0.2754098
##  [2,] 0.1666667  1.153846   -0.25      -0.07692308 0.4166667   -0.08333333
##  [3,] NA         0.08333333 NA         NA          -0.08333333 0
##  [4,] -0.8333333 0.1538462  1.75       -0.07692308 0.4166667   -0.08333333
##  [5,] NA         1.846733   NA         2.214286    0.2154472   0.08181818
##  [6,] 0.1666667  0.1538462  -0.25      -0.07692308 0.4166667   -0.08333333
##  [7,] NA         0          0.3412698  -0.173061   0.9166667   -0.1818182
##  [8,] -1.833333  -1.846154  -0.25      -0.07692308 -1.583333   1.916667
##  [9,] -0.2312729 -0.2922228 NA         0.032       NA          -0.5370304
## [10,] 0.1666667  -0.8461538 -1.25      -0.07692308 0.4166667   0.9166667
## [11,] NA         NA         -0.8447293 -0.1608947  0.8126984   -0.176911
## [12,] 0.1666667  0.1538462  1.75       -0.07692308 -0.5833333  -1.083333
## [13,] 0.03242212 -0.6276923 0.3571429  -1.163858   1.55        0.6540984
## [14,] 0.1666667  1.153846   -0.25      -0.07692308 -0.5833333  -1.083333
## [15,] 0.4545455  -0.4534266 NA         -0.6474926  -0.6571757  NA
## [16,] 1.166667   0.1538462  -0.25      -0.07692308 -1.583333   1.916667
## [17,] -0.9678063 0.25       0.7777778  0.04727273  -0.9090909  0.1414246
## [18,] 0.1666667  2.153846   -0.25      -0.07692308 0.4166667   -2.083333
## [19,] NA         -0.5454545 NA         0.3067164   -0.95       0.06666667
## [20,] 0.1666667  -0.8461538 -1.25      -0.07692308 0.4166667   -0.08333333
##       995697     3278419    4046972    4935243    2965087    3886540
##  [1,] NA         NA         -0.627451  0.6341463  -0.1842105 NA
##  [2,] -0.2       0.2727273  0.4285714  -2         -1.8       -0.4
##  [3,] NA         NA         -1.923077  0          NA         0.109828
##  [4,] -1.2       -1.727273  0.4285714  0          0.2        0.6
##  [5,] NA         NA         -0.258156  NA         0.6363636  -1.2
##  [6,] -1.2       0.2727273  -0.5714286 -2         2.2        -0.4
##  [7,] 0.3076923  NA         0.4062758  0.5710383  NA         0.3
##  [8,] 1.8        2.272727   0.4285714  0          0.2        0.6
##  [9,] -0.5579882 NA         0.01338443 -0.787242  -1         0.1
## [10,] -0.2       -1.727273  -0.5714286 0          1.2        -0.4
## [11,] 1.199754   NA         0.6661211  NA         2          -1.633415
## [12,] 0.8        -0.7272727 -0.5714286 0          0.2        0.6
## [13,] NA         0.2        -0.976881  -0.8461538 0.3846154  NA
## [14,] -0.2       1.272727   0.4285714  1          -0.8       -0.4
## [15,] NA         -0.8       -0.0818943 -0.707305  1.636364   0.08333333
## [16,] -0.2       0.2727273  0.4285714  2          0.2        0.6
## [17,] NA         0.1666667  0.2727273  -1.933333  -0.1666667 NA
## [18,] -0.2       1.272727   0.4285714  0          -1.8       -1.4
## [19,] -0.2727273 NA         -1.9       0.1227472  0.06982249 0.09210526
## [20,] 0.8        -1.727273  1.428571   -1         0.2        0.6
##       662664     2341442    3315379   4976553    270282     2991842
##  [1,] NA         0.1666667  0.5948403 NA         0.3333333  0.04545455
##  [2,] -0.1111111 1.083333   -1.1      -0.2       -0.2727273 0.7333333
##  [3,] 0.09090909 0.2831909  NA        0.5835381  2.214286   -0.7024421
##  [4,] 0.8888889  1.083333   -0.1      0.8        -1.272727  1.733333
##  [5,] -1.833333  -0.550805  -0.3      NA         1.5        -0.3
##  [6,] 0.8888889  0.08333333 0.9       -0.2       -2.272727  0.7333333
##  [7,] NA         -0.625     NA        NA         -1.5       -1.692308
##  [8,] -0.1111111 0.08333333 -0.1      -0.2       1.727273   -1.266667
##  [9,] 0.1666667  0.0263687  NA        NA         0.1141304  2.333333
## [10,] -0.1111111 -0.9166667 1.9       -0.2       -2.272727  1.733333
## [11,] NA         1.888889   NA        NA         2.214286   0.4
## [12,] 1.888889   -0.9166667 -0.1      -0.2       1.727273   0.7333333
## [13,] 0.09090909 -0.2916667 NA        -0.5       -0.3333333 0.3566719
## [14,] -1.111111  2.083333   -2.1      0.8        1.727273   0.7333333
## [15,] NA         NA         -1.1      -0.8333333 NA         NA
## [16,] -2.111111  0.08333333 -1.1      0.8        -0.2727273 -1.266667
## [17,] NA         -0.4444444 NA        0.9669775  NA         NA
## [18,] -0.1111111 -0.9166667 0.9       -1.2       -0.2727273 -0.2666667
## [19,] NA         NA         0.4651106 NA         NA         -0.4
## [20,] NA         0.08333333 0.9       -0.2       -0.2727273 -1.266667
##       3984496
##  [1,] -0.85
##  [2,] 0.4
##  [3,] 0
##  [4,] 0.4
##  [5,] -1.64142
##  [6,] -1.6
##  [7,] 0.1666667
##  [8,] 1.4
##  [9,] 1
## [10,] -0.6
## [11,] 0.2136095
## [12,] -1.6
## [13,] 0.5
## [14,] 0.4
## [15,] -0.8
## [16,] 0.4
## [17,] -0.4
## [18,] 1.4
## [19,] 0.1296154
## [20,] -0.6
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)), "tests")
```

```
## avg MAE = 1.111389 from 588 tests
```

viewing a histogram of the ratings helps gauge the impact of different MAE’s and decide what the likethresh should be

```
hist(unlist(users[train,]))
```



```
likethresh = 2.5 # insert your own value here
showCM(preds,likethresh)
```

```
## TN=   587 FP=     0
## FN=     1 TP=     0  (total=588)
## accuracy  = 99.8%
## precision = NaN%
## recall    = 0.0%
```

Recommendation

Using pearson’s similarity for user - 57920

```
users_r = acast(ev_Final, User_Id ~ Hotel_Id, value.var = "Ratings")
users_r = sweep(users_r, 1, rowMeans(users_r, na.rm=TRUE) )  # normalise the data
users_r = as.matrix(users_r)
rownames(users_r)[1]
```

```
## [1] "57920"
```

```
target = users_r[rownames(users_r)[1],]
names(getrecommendations_UU(target, users_r, simfun=pearsonsim))
```

```
## [1] "570753" "791690" "226121" "236180" "101618"
```

```
Top1_Hotel = integer(nrow(users_r))
Top2_Hotel = integer(nrow(users_r))
Top3_Hotel = integer(nrow(users_r))
Top4_Hotel = integer(nrow(users_r))
Top5_Hotel = integer(nrow(users_r))

for (i in 1:nrow(users_r)) {
  target = users_r[rownames(users_r)[i],]
  a = getrecommendations_UU(target, users_r, simfun=pearsonsim)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cf <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFact
ors = TRUE)
rownames(df_cf) = rownames(users)
write.csv(df_cf, file = "User Based Collaborative Recommended Hotel For each user using pears
ons.csv")
```

Using Cosine similarity for user - 57920

```
rownames(users_r)[1]
```

```
## [1] "57920"
```

```
getrecommendations_UU(target, users_r, simfun=cosinesim)
```

```
##   533623   516482   361342   601882   189754
## 2.307692 2.250000 2.181818 2.000000 1.900000
```

```
Top1_Hotel = integer(nrow(users_r))
Top2_Hotel = integer(nrow(users_r))
Top3_Hotel = integer(nrow(users_r))
Top4_Hotel = integer(nrow(users_r))
Top5_Hotel = integer(nrow(users_r))

for (i in 1:nrow(users_r)) {
  target = users_r[rownames(users_r)[i],]
  a = getrecommendations_UU(target, users_r, simfun=cosinesim)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cf <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFact
ors = TRUE)
rownames(df_cf) = rownames(users)
write.csv(df_cf, file = "User Based Collaborative Recommended Hotel For each user using cosin
e.csv")
```

Using Cosine similarity for user - 57920

```
rownames(users_r)[1]
```

```
## [1] "57920"
```

```
getrecommendations_UU(target, users_r, simfun=euclidsim)
```

```
##   408794   533623   619662   771343   361342
## 2.615385 2.307692 2.200000 2.111111 2.078838
```

```
Top1_Hotel = integer(nrow(users_r))
Top2_Hotel = integer(nrow(users_r))
Top3_Hotel = integer(nrow(users_r))
Top4_Hotel = integer(nrow(users_r))
Top5_Hotel = integer(nrow(users_r))

for (i in 1:nrow(users_r)) {
  target = users_r[rownames(users_r)[i],]
  a = getrecommendations_UU(target, users_r, simfun=euclidsim)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cf <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFact
ors = TRUE)
rownames(df_cf) = rownames(users)
write.csv(df_cf, file = "User Based Collaborative Recommended Hotel For each user using eucli
dean.csv")
```

Using Jacard similarity for user - 57920

```
rownames(users_r)[1]
```

```
## [1] "57920"
```

```
getrecommendations_UU(target, users_r, simfun=jacardsim)
```

```
##   408794   533623   619662   771343   361342
## 2.615385 2.307692 2.200000 2.111111 1.986253
```

```
Top1_Hotel = integer(nrow(users_r))
Top2_Hotel = integer(nrow(users_r))
Top3_Hotel = integer(nrow(users_r))
Top4_Hotel = integer(nrow(users_r))
Top5_Hotel = integer(nrow(users_r))

for (i in 1:nrow(users_r)) {
  target = users_r[rownames(users_r)[i],]
  a = getrecommendations_UU(target, users_r, simfun=jacardsim)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cf <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFact
ors = TRUE)
rownames(df_cf) = rownames(users)
write.csv(df_cf, file = "User Based Collaborative Recommended Hotel For each user using Jacar
d.csv")
```

Item Based Collaborative Filtering

Convert the dataframe from long to wide format

```
users_IBCF = acast(ev_Final, User_Id ~ Hotel_Id, value.var = "Ratings")
users_IBCF = sweep(users_IBCF, 1, rowMeans(users_IBCF, na.rm=TRUE) )  # normalise the data
dim(users_IBCF)
```

```
## [1] 422 508
```

Check the sparsity and fill rate of the matrix

```
fillrate(users_IBCF)
```

```
## 2.155558 %
```

setup the train/test scheme

```
numtestusers = 84
test  = sample(rownames(users_IBCF), min(numtestusers,nrow(users_IBCF)))
train = setdiff(rownames(users_IBCF),test)
```

compute the item similarity matrix

Cosine Similarity

```
st=Sys.time(); item_cosine_sims = getitemsimsmatrix(users_IBCF[train,], simfun=cosinesim); Sys.time()-st
```

```
## Time difference of 3.551247 secs
```

```
cat("Fill rate for cosine similarity : "); fillrate(item_cosine_sims); cat("\n\n");
```

```
## Fill rate for cosine similarity :
```

```
## 12.5589 %
```

```
st=Sys.time(); item_euclidean_sims = getitemsimsmatrix(users_IBCF[train,], simfun=euclidsim); Sys.time()-st
```

```
## Time difference of 3.204997 secs
```

```
cat("Fill rate for euclidean similarity : "); fillrate(item_euclidean_sims); cat("\n\n");
```

```
## Fill rate for euclidean similarity :
```

```
## 13.37885 %
```

```
st=Sys.time(); item_pearson_sims = getitemsimsmatrix(users_IBCF[train,], simfun=pearsonsim); Sys.time()-st
```

```
## Time difference of 5.284658 secs
```

```
cat("Fill rate pearson's similarity  : "); fillrate(item_pearson_sims); cat("\n\n");
```

```
## Fill rate pearson's similarity  :
```

```
## 0.9323269 %
```

```
st=Sys.time(); item_jacard_sims = getitemsimsmatrix(users_IBCF[train,], simfun=jacardsim); Sys.time()-st
```

```
## Time difference of 3.073435 secs
```

```
cat("Fill rate jacard similarity : "); fillrate(item_euclidean_sims); cat("\n\n");
```

```
## Fill rate jacard similarity :
```

```
## 13.37885 %
```

test IBCF Using Cosine similarity

```
preds = predictCF(users_IBCF[test,], itemsims=item_cosine_sims, numtestitems=10, random=FALSE
)
```

```
## .....................................................................................
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)),"tests")
```

```
## avg MAE = 1.052709 from 570 tests
```

test IBCF Using euclidean similarity

```
preds = predictCF(users_IBCF[test,], itemsims=item_euclidean_sims, numtestitems=10, random=FA
LSE)
```

```
## .....................................................................................
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)),"tests")
```

```
## avg MAE = 1.089706 from 587 tests
```

test IBCF Using pearson similarity

```
preds = predictCF(users_IBCF[test,], itemsims=item_pearson_sims, numtestitems=10, random=FALS
E)
```

```
## .....................................................................................
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)),"tests")
```

```
## avg MAE = 1.098366 from 77 tests
```

test IBCF Using jacard similarity

```
preds = predictCF(users_IBCF[test,], itemsims=item_jacard_sims, numtestitems=10, random=FALSE
)
```

```
## .....................................................................................
```

```
cat("avg MAE =",avgMAE(preds), "from", validcnt(listpreds(preds)),"tests")
```

```
## avg MAE = 1.085475 from 587 tests
```

Recommendation for a user - 57920, using Item - Based

We required minimum of 2 hotels for recommendations using Euclidean similarity

```
target = users_r[rownames(users_IBCF)[1],c("568005","789743")]
getrecommendations_II(target, item_euclidean_sims)
```

```
##  43980  55709  75444 118117 165833
##  -1.75  -1.75  -1.75  -1.75  -1.75
```

Recommendation using Cosine similarity for Item Based

```
target = users_r[rownames(users_IBCF)[1],]
getrecommendations_II(target, item_cosine_sims)
```

```
## 171631 182839 543930 675673 682461
##   2.25   2.25   2.25   2.25   2.25
```

```
Top1_Hotel = integer(nrow(users_IBCF))
Top2_Hotel = integer(nrow(users_IBCF))
Top3_Hotel = integer(nrow(users_IBCF))
Top4_Hotel = integer(nrow(users_IBCF))
Top5_Hotel = integer(nrow(users_IBCF))

for (i in 1:nrow(users_IBCF)) {
  target = users_r[rownames(users_IBCF)[i],]
  a = getrecommendations_II(target, item_cosine_sims)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cfib <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFa
ctors = TRUE)
rownames(df_cfib) = rownames(users)
write.csv(df_cfib, file = "Item Based Collaborative Recommended Hotel For each user using cos
ine.csv")
```

Recommendation using Euclidean similarity for Item Based

```
target = users_r[rownames(users_IBCF)[1],]
getrecommendations_II(target, item_euclidean_sims)
```

```
##  64769 171631 182839 221922 528594
##   2.25   2.25   2.25   2.25   2.25
```

```
Top1_Hotel = integer(nrow(users_IBCF))
Top2_Hotel = integer(nrow(users_IBCF))
Top3_Hotel = integer(nrow(users_IBCF))
Top4_Hotel = integer(nrow(users_IBCF))
Top5_Hotel = integer(nrow(users_IBCF))

for (i in 1:nrow(users_IBCF)) {
  target = users_r[rownames(users_IBCF)[i],]
  a = getrecommendations_II(target, item_euclidean_sims)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cfib <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFa
ctors = TRUE)
rownames(df_cfib) = rownames(users)
write.csv(df_cfib, file = "Item Based Collaborative Recommended Hotel For each user using euc
lidean.csv")
```

Recommendation using Pearson’s similarity for Item Based

```
target = users_r[rownames(users_IBCF)[1],]
getrecommendations_II(target, item_pearson_sims)
```

```
## 294468 304189 320689 327285 793845
##   1.75   1.75   1.25   1.25   1.25
```

```
Top1_Hotel = integer(nrow(users_IBCF))
Top2_Hotel = integer(nrow(users_IBCF))
Top3_Hotel = integer(nrow(users_IBCF))
Top4_Hotel = integer(nrow(users_IBCF))
Top5_Hotel = integer(nrow(users_IBCF))

for (i in 1:nrow(users_IBCF)) {
  target = users_r[rownames(users_IBCF)[i],]
  a = getrecommendations_II(target, item_pearson_sims)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cfib <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFa
ctors = TRUE)
rownames(df_cfib) = rownames(users)
write.csv(df_cfib, file = "Item Based Collaborative Recommended Hotel For each user using pea
rson.csv")
```

Recommendation using Jacard similarity for Item Based

```
target = users_r[rownames(users_IBCF)[1],]
getrecommendations_II(target, item_jacard_sims)
```

```
##  64769 171631 182839 221922 528594
##   2.25   2.25   2.25   2.25   2.25
```

```
Top1_Hotel = integer(nrow(users_IBCF))
Top2_Hotel = integer(nrow(users_IBCF))
Top3_Hotel = integer(nrow(users_IBCF))
Top4_Hotel = integer(nrow(users_IBCF))
Top5_Hotel = integer(nrow(users_IBCF))

for (i in 1:nrow(users_IBCF)) {
  target = users_r[rownames(users_IBCF)[i],]
  a = getrecommendations_II(target, item_jacard_sims)
  Top1_Hotel[i] = names(a[order(a,decreasing=TRUE)[1]])
  Top2_Hotel[i] = names(a[order(a,decreasing=TRUE)[2]])
  Top3_Hotel[i] = names(a[order(a,decreasing=TRUE)[3]])
  Top4_Hotel[i] = names(a[order(a,decreasing=TRUE)[4]])
  Top5_Hotel[i] = names(a[order(a,decreasing=TRUE)[5]])
}

df_cfib <- data.frame(Top1_Hotel, Top2_Hotel, Top3_Hotel, Top4_Hotel, Top5_Hotel, stringsAsFa
ctors = TRUE)
rownames(df_cfib) = rownames(users)
write.csv(df_cfib, file = "Item Based Collaborative Recommended Hotel For each user using Jac
ard.csv")
```

The End