HACKATHON(KDM)

Name:Vishnu Praveen Gude

Number : 16103389

Application Name : Movies and Jokes Recommender

**Technical Aspects :**

R Language,C#.NET,R.NET ,RecommenderLab for R,Ggplot2 for R graphics,Collaborative Filtering in R.

**DataSets :**

1. <http://grouplens.org/datasets/movielens/>
2. <http://goldberg.berkeley.edu/jester-data/>

**Assumptions and Data Cleaning :**

1.MovieLens dataset consists of user Ids,ratings,movieIds,Movies parameters(Rating 0.0 to 5.0)

2.Jester data set consist of user Ids,ratings,JokeIds (Rating : -10.0 to 10.0)

3.We assume User Id is same for both to merge Two datasets for our Application Need.

**Implementation:**

1.Idea is to develop web application in which movies and jokes are recommended to users based on training datasets.

2.We need to embed R in c#.After a extensive web search I found R.NET library in which we can run R commands from c# and we can run R scripts from c#.

3.I started with C# console application because we need command prompt rights to run commands.

4.I have written four R script files movie.R,jester.R,visual.R and jester\_v.R

a)movie.R - Movie Recommendation algorithm using RecommenderLab package in R

b)jester.R - Joke Recommendation script uses same RecommenderLab package

c)visual.R - Visualisations of Movielens dataset using Ggplot2 package for R graphics

d)jester\_v.R – Visualisations of jester jokes dataset using Ggplot2 package.

5.From the console application ,With the help of R.NET library I am able to run R scripts.

6.After debugging the application I got the EXE process file.

7.Now I started Web application and designed two pages one for recommendations based on user Ids and one for visualisations.

8.Inside the web application I have written code to run EXE process internally whenever user clicks submit button.

9.When ever user enters ID it is saved into txt file and that txt file is input to R scripts so that our recommendation algorithm recommends movies and jokes for that ID and saves output into txt file.

10.Whenever EXE file gets done running,we will get results loaded from txt file and displayed them on the web page.

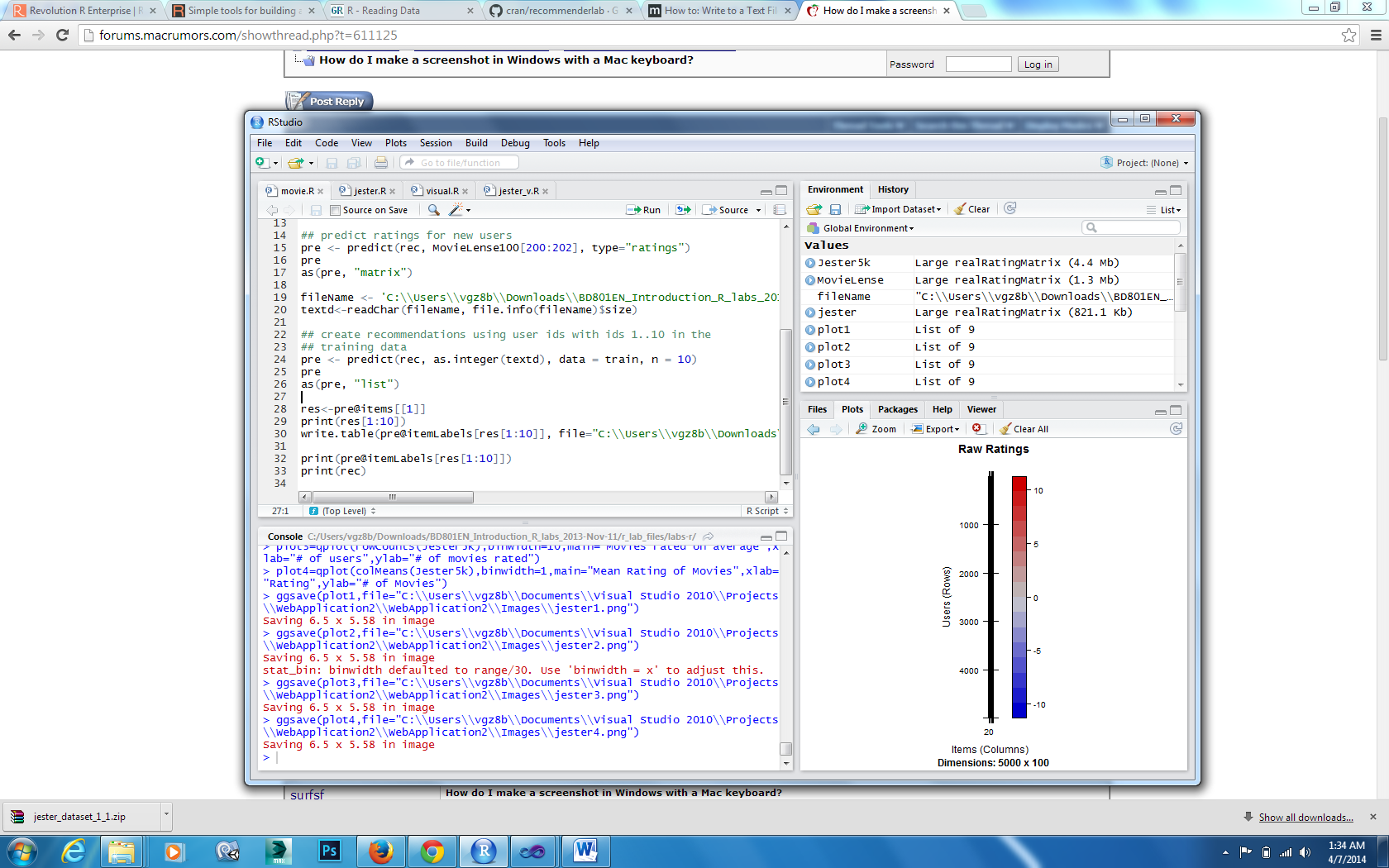
11.we can process continuous data too,assume a user rated one movie so that we will add that data to dataset which is input to R script .Whenever EXE file process is started our algorithm generates new recommendations based on growing datasets.

12.R is very powerful and super fast compared to mahout recommendation system and ofcourse,all the popular algorithms are present in R too.

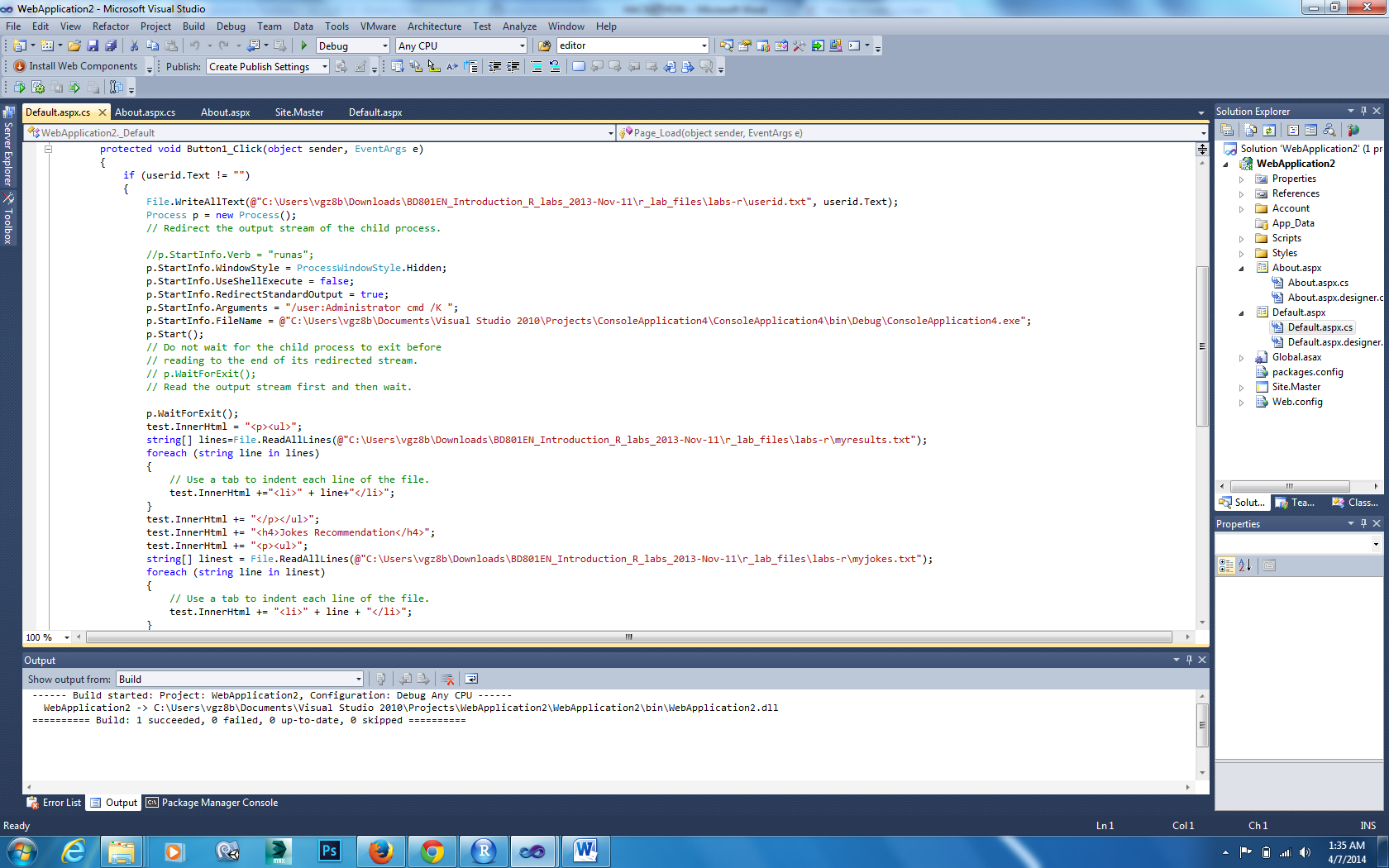
13.I did visualisations of Ratings,normalized ratings,average movie ratings etc in second page

**Screenshots:**

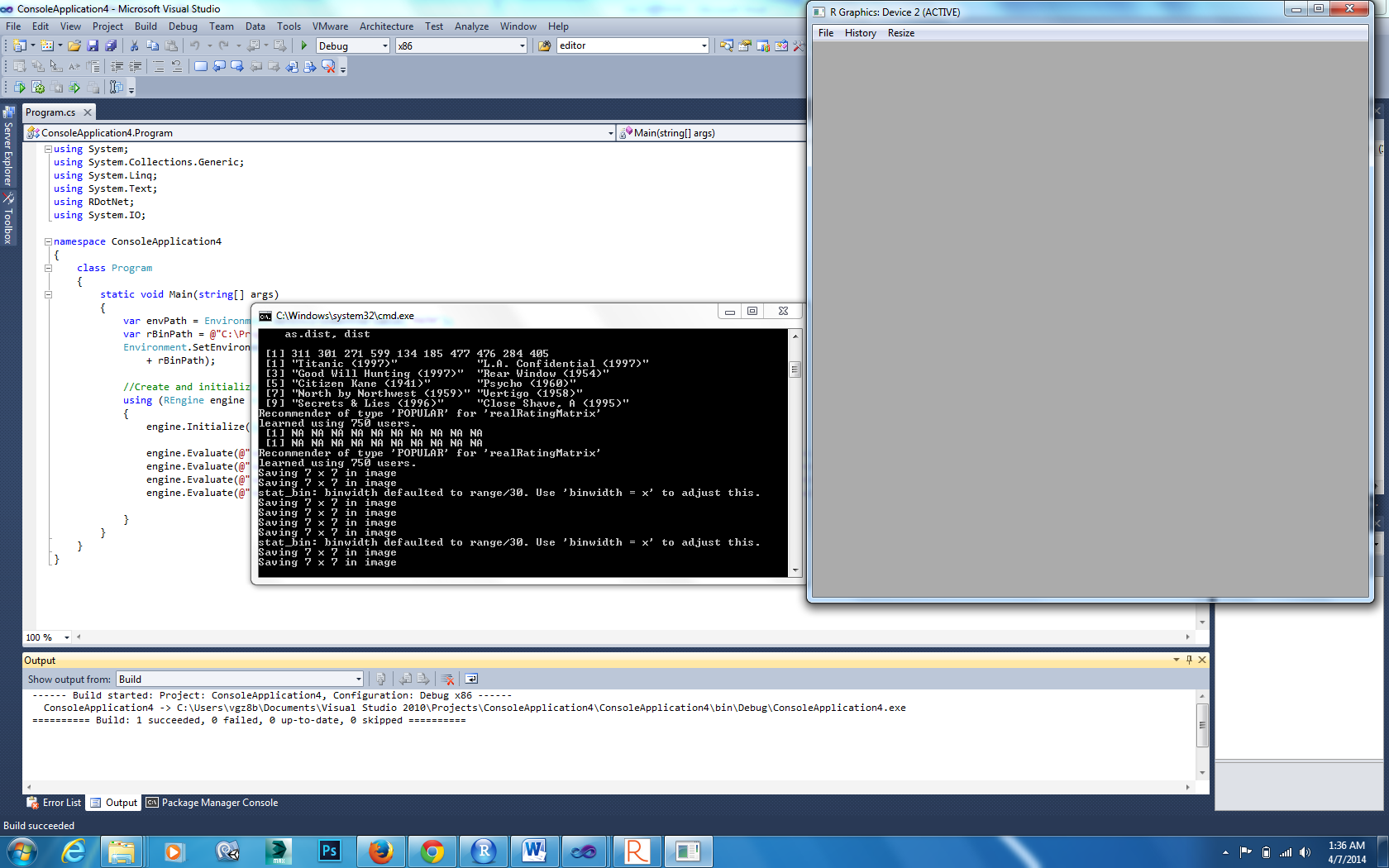
Rstudio with R scripts

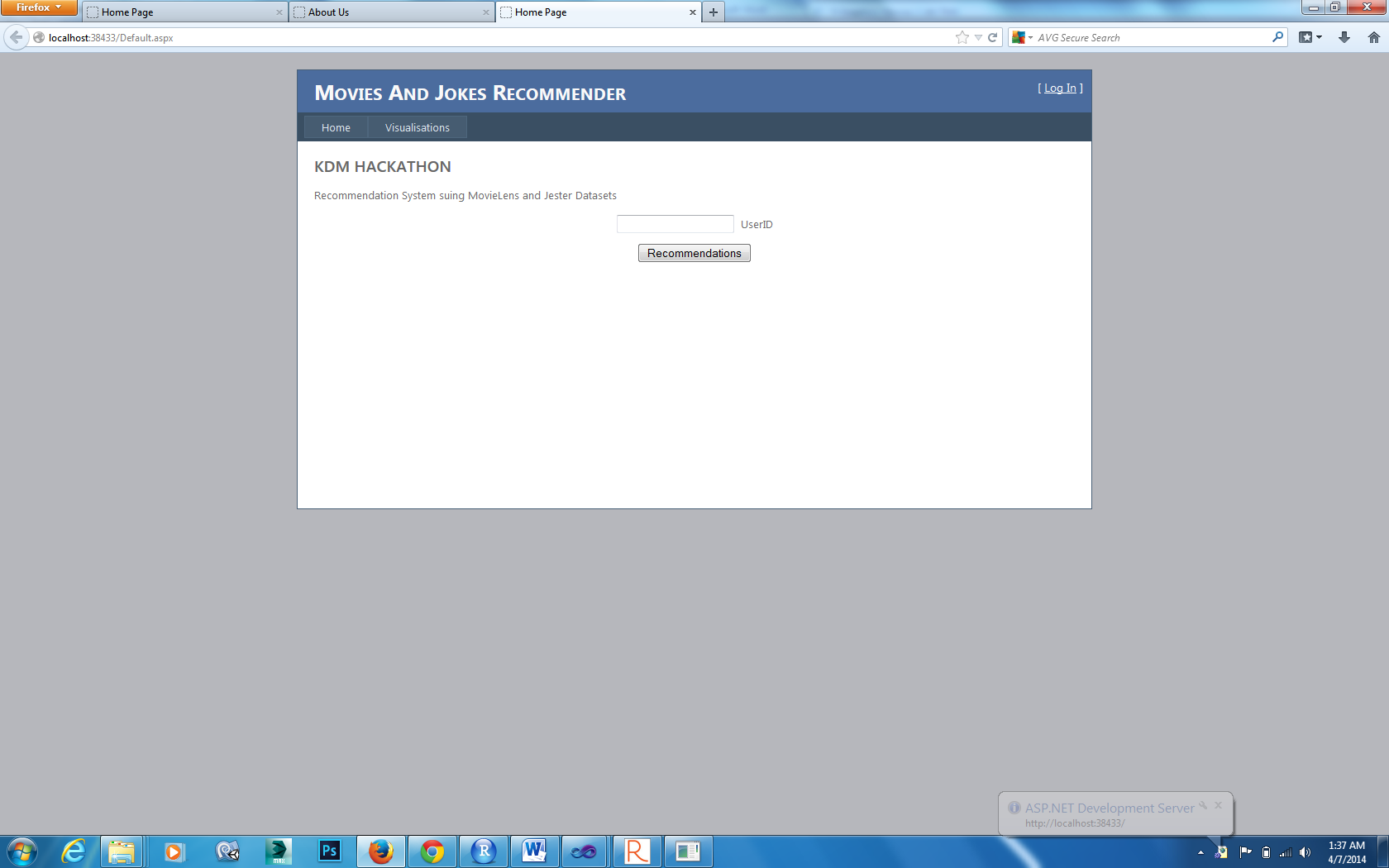


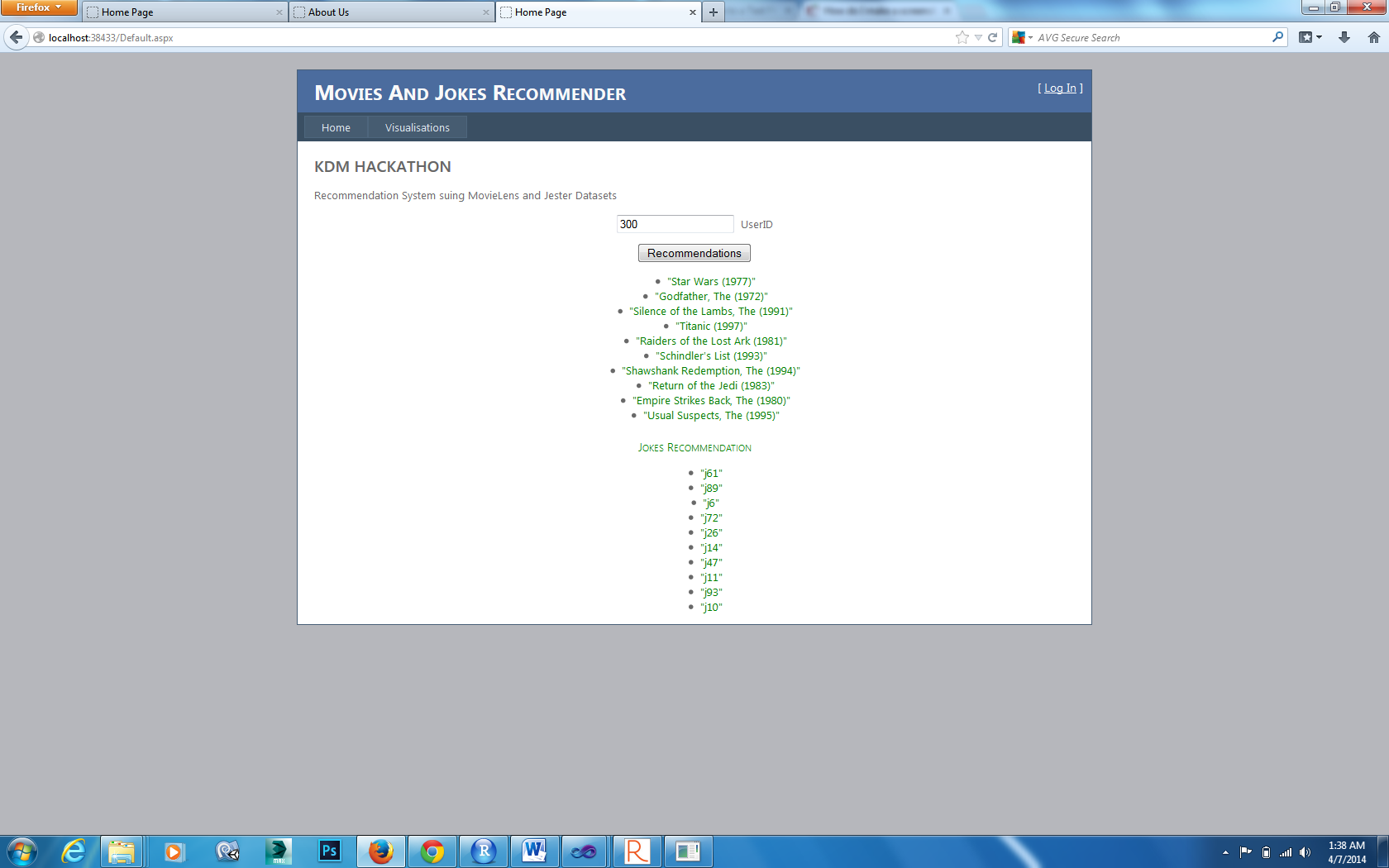
Visual studio

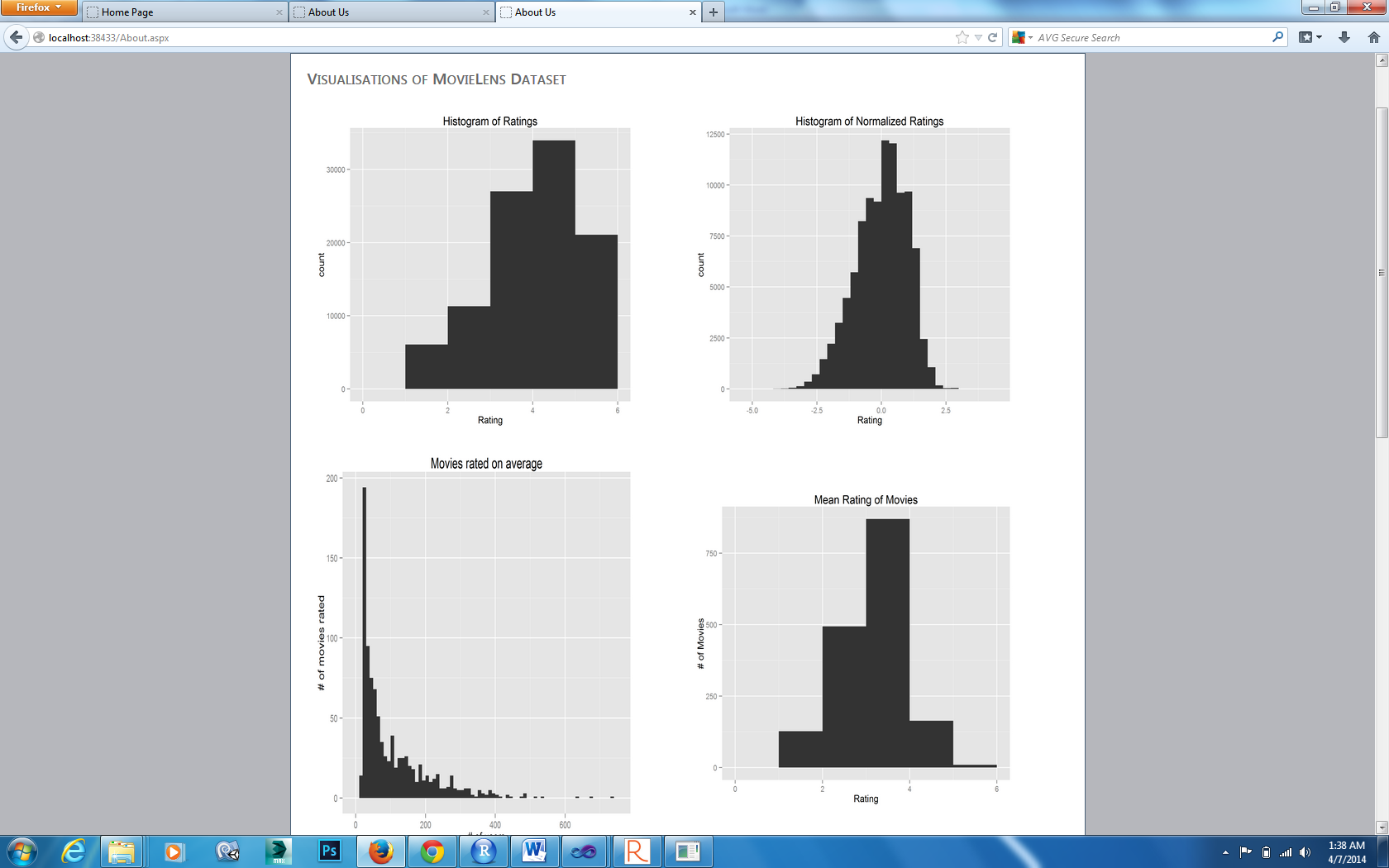


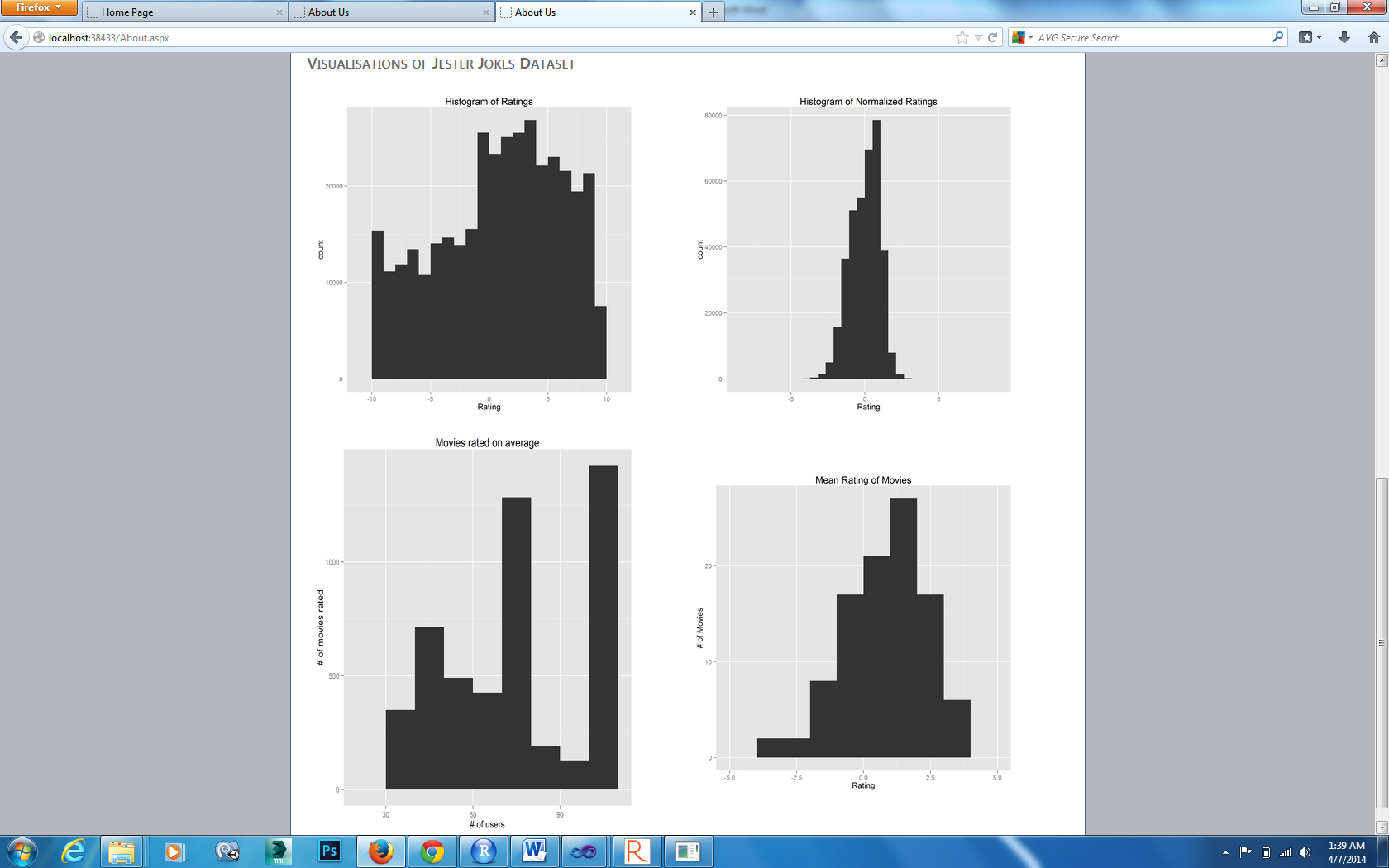
Console application output











**Execution Time**

**R took around 10 sec to process two large datasets and provide recommendations for users in web pages.**

Thanking You

Vishnu Praveen Gude