Assignment 2

Umais Siddiqui

September 4, 2017

Git hub : https://github.com/umais/DATA-607/tree/master/Assignment2"

Rpub : http://rpubs.com/umais/305818"

Movies Database

This R Markdown is written to work with the MYSQL Movies database. The SQL file can be viewed from the githib link below.

Th database that is created has three tables . RecentMovies, Person and PersonMovieRating. RecentMovies will store the movie information for the recent movies. The PErson table stores all the people that were surveyed to rate the movie. The PersonMovieRating stores the ratings that were collected by the people in the person table. We will be joining these tables and gathering this information in one select query and then loading it in to a R data frame.

Step 1 Load the R MYSQL Library

```
library(RMySQL)
```

```
## Loading required package: DBI
```

Step 2 Connect to the MySQL Movies database

```
 In this step we will be connecting to the Movies database and retrieving the information that
 was collected in the survey and load that in the R Data Frame
```

```
mydb = dbConnect(MySQL(), user='root', password='Welcome@1', dbname='movies', host='localhost')

rs = dbSendQuery(mydb, "SELECT m.MovieName,p.FirstName,p.LastName,r.Rating
FROM PersonMovieRating r INNER JOIN RecentMovie m
on r.MovieId=m.MovieId
INNER JOIN Person p
ON r.personID=p.personid;")
```

```
## Warning in .local(conn, statement, ...): Decimal MySQL column 3 imported as
## numeric
```

```
df=fetch(rs, n=-1)
```

Step 3 Show the first few rows using the head function and show the summary of the ratings.

```
head(df)
```

```
##   MovieName FirstName LastName Rating
## 1   Get Out    Mussab    Aftab    3.0
## 2   Get Out   Ufariah   Midhat    2.5
## 3   Get Out       Sam    Anaum    2.0
## 4   Get Out     Steve  Attwell    2.0
## 5   Get Out      Bill   McCann    2.0
## 6   Get Out   Stephen  Jackson    2.0
```

```
summary(df$Rating)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.000   2.000   3.250   3.429   5.000   5.000
```