Homework Week 2

CUNY MSDA DATA 607

Duubar Villalobos Jimenez mydvtech@gmail.com

February 12, 2017

Assignment - SQL a nd R

Choose six recent popular movies. Ask at least five people that you know (friends, family, classmates, imaginary friends) to rate each of these movie that they have seen on a scale of 1 to 5. Take the results (observations) and store them in a SQL database. Load the information into an R dataframe. Your deliverables should include your SQL scripts and your R Markdown code, posted to GitHub. This is by design a very open ended assignment. A variety of reasonable approaches are acceptable. You can (and should) blank out your SQL password if your solution requires it; otherwise, full credit requires that your code is "reproducible," with the assumption that I have the same database server and R software. You may work in a small group.

Solution

Need to type Local MySQL Root password

```
# Need to type root password for the local database
myLocalPassword <- 'pswrd'
```

Connect to database

```
# Connect to database
my.database = dbConnect(MySQL(), user='root', password = myLocalPassword, dbname='villalobos-movies', h
dbListTables(my.database)
```

```
## [1] "tblmovies"  "tblreviews" "tblusers"
```

Display Users Table

```
# -- ----------------------------------------------
# -- Table `users`
# -- ----------------------------------------------
my.users <- dbSendQuery(my.database, "SELECT * FROM tblUsers;")
dbFetch(my.users)
```

```
##   user_id   fname
## 1       1   Sarah
## 2       2   Maria
## 3       3   Elena
## 4       4   Diana
## 5       5 Michael
## 6       6   Heidy
```

Display Movies Table

```
# -- ------------------------------------------------
# -- Table `movies`
# -- ------------------------------------------------
my.movies <- dbSendQuery(my.database, "SELECT * FROM tblmovies;")
dbFetch(my.movies)
```

```
##   movie_id            title lenght
## 1        1            Logan     60
## 2        2              Si3    120
## 3        3     Kung Fu Yoga    141
## 4        4 The Ghazi Attack     90
## 5        5            Space    121
## 6        6            Raees     90
```

Display Reviews Table Raw info

```
# -- ------------------------------------------------
# -- Table `reviewss`
# -- ------------------------------------------------
my.reviews <- dbSendQuery(my.database, "SELECT * FROM tblreviews;")
dbFetch(my.reviews)
```

```
##   review_id movie_id user_id rating           review
## 1         1        1       1      5         The Best
## 2         2        1       5      5             Hala
## 3         3        4       4      3         Horrible
## 4         4        4       4      4       Historical
## 5         5        3       3     NA      Under rated
## 6         6        3       1      4 Really impressed
```

Read data from MySQL all tables combined into one query

```
movie.data = dbSendQuery(my.database, "SELECT
M.title As 'Title',
M.lenght AS 'Lenght',
U.fname As 'User',
R.rating As 'Rating',
R.review AS 'Review'
FROM tblMovies AS M
JOIN tblReviews AS R
ON M.movie_id = R.movie_id
JOIN tblUsers AS U
ON U.user_id = R.user_id;")
dbFetch(movie.data)
```

```
##              Title Lenght    User Rating           Review
## 1            Logan     60   Sarah      5         The Best
## 2            Logan     60 Michael      5             Hala
## 3 The Ghazi Attack     90   Diana      3         Horrible
## 4 The Ghazi Attack     90   Diana      4       Historical
```

```
## 5     Kung Fu Yoga    141   Elena     NA      Under rated
## 6     Kung Fu Yoga    141   Sarah      4 Really impressed
```

Only one user who gave reviews

```
movie.dataSingleUser = dbSendQuery(my.database, "SELECT
M.title As 'Title',
M.lenght AS 'Lenght',
U.fname As 'User',
R.rating As 'Rating',
R.review AS 'Review'
FROM tblMovies AS M
JOIN tblReviews AS R
ON M.movie_id = R.movie_id
JOIN tblUsers AS U
ON U.user_id = R.user_id
WHERE U.user_id = 1;")
dbFetch(movie.dataSingleUser)
```

```
##          Title Lenght  User Rating           Review
## 1        Logan     60 Sarah      5         The Best
## 2 Kung Fu Yoga    141 Sarah      4 Really impressed
```

Only one Rating score

```
movie.dataRating = dbSendQuery(my.database, "SELECT
M.title As 'Title',
M.lenght AS 'Lenght',
U.fname As 'User',
R.rating As 'Rating',
R.review AS 'Review'
FROM tblMovies AS M
JOIN tblReviews AS R
ON M.movie_id = R.movie_id
JOIN tblUsers AS U
ON U.user_id = R.user_id
WHERE R.rating = 5;")
dbFetch(movie.dataRating)
```

```
##   Title Lenght    User Rating   Review
## 1 Logan     60   Sarah      5 The Best
## 2 Logan     60 Michael      5     Hala
```

Database disconnect

```
### Disconnect
#dbGetQuery(my.database, "show processlist")
#dbGetQuery(my.database, "kill 14")
```