

Assignment – SQL and R

Binish Kurian Chandy

2/8/2018

How to build the postgres database?

All files used to build database can be found at : [link](#)

Use backup file or execute the sql commands from pgAdmin to load the database.

Load required drivers to connect to database server.

```
#install.packages("RPostgreSQL")
library(RPostgreSQL)
```

Open connection to database

```
driver <- dbDriver("PostgreSQL")
con <- dbConnect(driver, host="localhost", port=5432, dbname="movies", user="postgres")
```

Fetch movie ratings from database.

The movie ratings are fetched by join query across three tables namely movie, rating and reviewer.

Movie table contains movie name, movie id and director name.

Reviewer contains reviewer id and name.

Rating contains movie id, reviewer id and stars.

```
query <- "select m.mid, m.title, m.director, re.name, ra.stars from movie m inner join rating ra on m.mid = ra.mid
ra.rid = re.rid"

rs <- dbSendQuery(con, query)
movies <- dbFetch(rs, n=-1)
head(movies)
```

| ## | mid | title | director | name | stars |
|------|-----|--------------|-------------------|----------------|-------|
| ## 1 | 106 | LOGAN | James Mangold | Sarah Martinez | 4 |
| ## 2 | 103 | DUNKIRK | Christopher Nolan | Sarah Martinez | 1 |
| ## 3 | 101 | GET OUT | Jordan Peele | Sarah Martinez | 5 |
| ## 4 | 105 | LADY BIRD | Greta Gerwig | Sarah Martinez | 4 |
| ## 5 | 102 | THE BIG SICK | Michael Showalter | Sarah Martinez | 5 |
| ## 6 | 104 | WONDER WOMAN | Patty Jenkins | Sarah Martinez | 2 |

```
dbDisconnect(con)
```

```
## [1] TRUE
```

```
dbUnloadDriver(driver)
```

```
## [1] TRUE
```

How many movies received 5 star rating?

```
unique(movies[which(movies$stars == 5), ]$title)
```

```
## [1] "GET OUT" "THE BIG SICK" "WONDER WOMAN" "LOGAN"
## [5] "DUNKIRK"
```