PSTAT 10 Worksheet 8

Sou Hamura

2024-07-22

```
library(RSQLite)
library(sqldf)
```

```
## Loading required package: gsubfn
```

```
## Loading required package: proto
```

```
library(DBI)
chinook_db <- dbConnect(SQLite(), "Chinook_Sqlite.sqlite")
```

Problem 1

1. What is the total length in minutes of all tracks with GenreId = 2?

```
dbGetQuery(chinook_db, "select GenreId, sum(Milliseconds)/60000 as minutes from Track
           where Track.GenreId = 2")
```

```
##   GenreId minutes
## 1       2     632
```

2. Using a join, add the name of the genre to the above query.

```
dbGetQuery(chinook_db, "select Genre.Name, Track.GenreId,  sum(Milliseconds)/60000 as minutes from Track
           inner join Genre on Track.GenreId = Genre.GenreId
           where Track.GenreId = 2")
```

```
##   Name GenreId minutes
## 1 Jazz       2     632
```

3. Using a group by, find this information for all genres. I've provided just the first 5 records of the output.

```
dbGetQuery(chinook_db, "select Genre.Name, Track.GenreId,  sum(Milliseconds)/60000 as minutes from Track
           inner join Genre on Track.GenreId = Genre.GenreId
           group by Track.GenreId, Genre.Name")
```

```
##                    Name GenreId minutes
## 1                  Rock       1    6137
## 2                  Jazz       2     632
## 3                 Metal       3    1930
## 4     Alternative & Punk     4    1296
## 5         Rock And Roll       5      26
## 6                 Blues       6     364
## 7                 Latin       7    2247
## 8                Reggae       8     238
## 9                   Pop       9     183
## 10           Soundtrack      10     175
## 11           Bossa Nova      11      54
## 12       Easy Listening      12      75
## 13          Heavy Metal      13     138
## 14              R&B/Soul      14     223
## 15      Electronica/Dance    15     151
## 16                World      16     104
## 17          Hip Hop/Rap      17     103
## 18      Science Fiction      18     568
## 19              TV Shows      19    3324
## 20      Sci Fi & Fantasy     20    1261
## 21                 Drama      21    2746
## 22                Comedy      22     449
## 23           Alternative      23     176
## 24             Classical      24     362
## 25                 Opera      25       2
```

Problem 2

```r
dbGetQuery(chinook_db, "select EmployeeId as RepId, e.FirstName as RepFirstName, e.LastName as RepLastNa
           WHERE c.FirstName = 'Heather' AND c.LastName = 'Leacock'")
```

```
##   RepId RepFirstName RepLastName CustomerId CustFirstName CustLastName
## 1     4     Margaret        Park         22       Heather      Leacock
```

Problem 3

1. How many difference countries are the customers from?

```r
dbGetQuery(chinook_db, "select distinct count(Country) from Customer")
```

```
##   count(Country)
## 1             59
```

2. Retrieve the unique titles of the employees.

```r
dbGetQuery(chinook_db, "select distinct title from Employee")
```

```
##                    Title
## 1     General Manager
## 2       Sales Manager
## 3 Sales Support Agent
## 4          IT Manager
## 5            IT Staff
```

3. The following determines how many unique genres there are. Was distinct necessary here? Explain.

No because the number of GenreId in Genre is already distributed according to the number of Genres