R Notebook

Problem1

```
#load library
library(RSQLite)

#use dbConnect to connect to the database
db <- dbConnect(SQLite(), dbname = "/Users/christinamathai/Downloads/projectdb.db")
#list tables in projectdb.db by using dbListTables
dbListTables(db)
```

```
## [1] "projectmgr" "projects"
```

```
#output: The two tables are "projectmgr" and "project"

#list all column from project table
sqlCmd1 <- "SELECT * FROM projects"
#dbGetQuery returns the result of the query
rs1 = dbGetQuery(db, sqlCmd1)
print (rs1)
```

```
##   pid    pname budget pmgr
## 1   1    TOGAF  25000  100
## 2   2  AirDrop  45000  103
## 3   3 WebQueue  55500  100
## 4   4  LastWay 250000  105
```

```
#list all column from projectmgr table
sqlCmd2 <- "SELECT * FROM projectmgr"
rs2 = dbGetQuery(db, sqlCmd2)
print (rs2)
```

```
##   pmid          pmname
## 1  100         John Wu
## 2  103      Ann Molloy
## 3  105 Sandeep Raghani
```

1.List the names and budgets of all projects.

```
sqlCmd3 <- "SELECT pname,budget FROM projects"
rs3 = dbGetQuery(db, sqlCmd3)
print (rs3)
```

```
##      pname budget
## 1    TOGAF  25000
## 2  AirDrop  45000
## 3 WebQueue  55500
## 4  LastWay 250000
```

2.How many projects are there?

```
sqlCmd4 <- "SELECT * FROM projects"
#dbGetQuery returns the result of the query
rs4 = dbGetQuery(db, sqlCmd4)
print (rs4)
```

```
##   pid    pname budget pmgr
## 1   1    TOGAF  25000  100
## 2   2  AirDrop  45000  103
## 3   3 WebQueue  55500  100
## 4   4  LastWay 250000  105
```

```
#the output gives the list of projects. There are 4 projects
```

3.List the names of all project managers.

```
sqlCmd5 <- "SELECT pmname FROM projectmgr"
rs5 = dbGetQuery(db, sqlCmd5)
print (rs5)
```

```
##            pmname
## 1         John Wu
## 2      Ann Molloy
## 3 Sandeep Raghani
```

4.List all project names and the name of the project manager only, sorted in alphabetical order by project name.

```
sqlCmd6 <- "SELECT pname, pmname FROM projects JOIN projectmgr ON projects.pmgr = projectmgr.pmid ORDER
rs6 = dbGetQuery(db, sqlCmd6)
print (rs6)
```

```
##      pname          pmname
## 1  AirDrop      Ann Molloy
## 2  LastWay Sandeep Raghani
## 3    TOGAF         John Wu
## 4 WebQueue         John Wu
```

5.What is the total budget of all project managed by "John Wu"?

```
sqlCmd7 <- "SELECT SUM (budget) FROM projects JOIN projectmgr ON projects.pmgr = projectmgr.pmid WHERE p
rs7 = dbGetQuery(db, sqlCmd7)
print (rs7)
```

```
##   SUM (budget)
## 1        80500
```

6.How many projects have a budget of less than $50,000 but more than $10,000?

```
sqlCmd8 <- "SELECT pname, budget FROM projects WHERE budget < 50000 AND budget > 10000"
query <- dbGetQuery(db, sqlCmd8)
print(query)
```

```
##     pname budget
## 1   TOGAF  25000
## 2 AirDrop  45000
```

7.List the name of each project manager, the number of projects they manage, and the total budget of their projects.

```
sqlCmd9 <- "SELECT pmname, SUM(budget), COUNT (pname) FROM projects JOIN projectmgr ON projects.pmgr = p
rs9 = dbGetQuery(db, sqlCmd9)
print (rs9)
```

```
##            pmname SUM(budget) COUNT (pname)
## 1      Ann Molloy       45000             1
## 2         John Wu       80500             2
```

```
## 3 Sandeep Raghani      250000              1
```

8.List all projects that have a project name that starts with "W".

```
sqlCmd10 <- "SELECT pname FROM projects WHERE pname LIKE 'W%'"
query <- dbGetQuery(db, sqlCmd10)
print(query)
```

```
##      pname
## 1 WebQueue
```

Problem 2

```
#1.Create a table named transactions  in SQLite that matches the CSV (column name and order must match)
#2.Import (load) the entire data from the CSV into that table using the SQLite .import command.

# sqlite> .mode csv
#sqlite> create table transactions (
#   ...> visits,
#   ...> transactions,
#   ...> OS,
#  ...> Gender,
#   ...> Revenue
#   ...> );
#sqlite> .import /Users/christinamathai/Downloads/customertxndata.csv transactions

#3.Create a SQL query that finds the total revenue by OS.
#SELECT OS, SUM (Revenue) FROM transactions GROUP BY OS;
#Android,4099918.38994199
#iOS,6272605.33393828

#4.Create a SQL query that finds the largest revenue and for which OS and gender it occurred.
#SELECT OS, Gender, MAX (Revenue) FROM transactions;
#Android,NA,990.306213040332

#5. Create a SQL query that finds the average number of visits by gender, excluding NA cases
#SELECT Gender, AVG(visits) FROM transactions WHERE Gender != "NA" GROUP BY Gender;
#Female,21.2134831460674
#Male,10.8718940936864
```