Homework11a

Steve Simon

This file was created on 2020-07-24 and last modified on 2020-07-25.

Note: this solution uses R and SQLite. An alternate solution using SAS and Oracle is also available.

Note: Some of the names used in this code are arbitrary and you can choose whatever names you want. To emphasize which names can be modified at your discretion, I am using names of famous statisticians.

The statistician being honored in this code is Robert V. Hogg.

Use the same database shown in the video. It is available on the Insights platform, or you can download a sqlite file from Canvas.

1. List id and migraine_label for the first ten records after joining the results_table and migraine_table.
2. Get a count of the number of records in the database in the control group and the treatment group. Use the label for group and not the number code.
3. Get a count of the numbers of males and females where you restrict age to be less than 40. Use the label for sex, but convert it to all uppercase.

```
library(sqldf)
```

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

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

```
## Loading required package: RSQLite
```

```
hogg <- dbConnect(SQLite(),
  dbname="../data/acupuncture/acupuncture_db.sqlite")
robert1 <- dbGetQuery(conn=hogg,  "
    select r.id, g.group_label
      from results_table as r
      join group_table as g
      on r.grp=g.group_code
      limit 10
")
robert1
```

```
##     id group_label
## 1  100   Treatment
## 2  101     Control
## 3  104     Control
## 4  105     Control
## 5  108     Control
## 6  112   Treatment
## 7  113   Treatment
## 8  114   Treatment
## 9  126     Control
## 10 130   Treatment
```

```
dbDisconnect(conn=hogg)
```

```
library(sqldf)
hogg <- dbConnect(SQLite(),
  dbname="../data/acupuncture/acupuncture_db.sqlite")
robert2 <- dbGetQuery(conn=hogg,  "
    select g.group_label, count(r.id) as n
      from results_table as r
      join group_table as g
      on r.grp=g.group_code
      group by r.grp
")
robert2
```

```
##   group_label   n
## 1     Control 196
## 2   Treatment 205
```

```
dbDisconnect(conn=hogg)
```

```
library(sqldf)
hogg <- dbConnect(SQLite(),
  dbname="../data/acupuncture/acupuncture_db.sqlite")
robert3 <- dbGetQuery(conn=hogg,  "
    select
        upper(s.sex_label) as sex,
        count(r.id) as n
      from results_table as r
      join sex_table as s
        on r.sex=s.sex_code
      where age < 40
      group by r.sex
")
robert3
```

```
##      sex  n
## 1   MALE 15
## 2 FEMALE 92
```

```
dbDisconnect(conn=hogg)
```