Homework12a

Steve Simon

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

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 George Box.

Use the longitudinal_example_db. It is available as on the Canvas website or you can find it on the Insights platform.

1. Do an inner join of baseline_table and one_year_table. Display the first ten rows of data only.

2. Explain why id 104 included in this inner join, but not the inner join shown in the video?

3. Count the number of records after an inner join of baseline_table and year_one_table. Compare this to the number of records in the year_one_table.

4. Compute the average pk score at baseline, the average score at one year, and the average change score.

5. Find and list the two labels in migraine_table that do not correspond to any codes in demog_table.

6. Show that there are no unmatched labels or unmatched codes for group_table.

```
library(sqldf)
```

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

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

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

```
box <- dbConnect(SQLite(),
  dbname="../data/longitudinal_example_db.sqlite")
george1 <- dbGetQuery(conn=box, "
    select tbl_name from sqlite_master
")

george1
```

```
## [1] tbl_name
## <0 rows> (or 0-length row.names)
```

```
dbDisconnect(conn=box)
```