Module03 homework, Q2

Steve Simon

Dates: This file was created on 2020-07-06 and last modified on 2021-06-26.

Purpose: Answer M03-Q02. There are only two fields in the database that have null values. Get a count of the number of missing values for the indicator for teaching hospital (TEACHING_IND).

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

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 David Blackwell.

```r
library(sqldf)
```

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

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

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

```r
blackwell <- dbConnect(SQLite(),
  dbname="../data/ehr.sqlite")
david <- dbGetQuery(conn=blackwell, "
    select
      count(*) as number_missing_teaching_ind
      from hospital
      where TEACHING_IND is null
")

david
```

```
##   number_missing_teaching_ind
## 1                          42
```

```r
dbDisconnect(conn=blackwell)
```