Module05 homework, Q1

Steve Simon

Dates: This file was created on 2020-07-12 and last modified on 2021-07-14.

Purpose: To answer M05-Q1. Use the encounter table in the ehr database. Classify age into two groups labelled 'Child' and 'Adult'. A child represents patients with age <= 20 and an adult represents patients with age >20. Hint: use the case function.

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 Barbara A. Bailar.

```
library(sqldf)
```

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

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

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

```
bailar <- dbConnect(SQLite(),
  dbname="../data/ehr.sqlite")
barbara <- dbGetQuery(conn=bailar,  "
  select
    age,
    case
      when age <= 20
        then 'Child'
        else 'Adult'
      end as age_group
  from encounter
  limit 20
")

barbara
```

```
##    AGE age_group
## 1   52     Adult
## 2   58     Adult
## 3   43     Adult
## 4   55     Adult
## 5   40     Adult
## 6   38     Adult
```

```
## 7   42     Adult
## 8   80     Adult
## 9   75     Adult
## 10  74     Adult
## 11  61     Adult
## 12  44     Adult
## 13  45     Adult
## 14  20     Child
## 15  27     Adult
## 16  29     Adult
## 17  42     Adult
## 18  52     Adult
## 19  40     Adult
## 20  41     Adult
```

```
dbDisconnect(conn=bailar)
```