

Coding assignment

Ajla Dzajic

Friday, May 27, 2016

Data

For calculating the expected results we're provided with two input files, namely

1. **tickets_sold.csv** - each row contains the booking information for one individual ticket
2. **row_capacity.csv** - each row contains the maximum capacity for an auditorium_row

For simplicity, we assume that there is only one auditorium in the cinema that we are looking at.

Exected results

Our task is to calculate three different results, namely

1. **Total_tickets_per_week_and_movie**
2. **Seat_load_factor_per_week_and_movie**
3. **Seat_load_factor_per_week**

```
# dir.create("data_assignment")
# setwd("C:/Users/ajla/Documents/data_assignment")
```

```
row_capacity <- read.csv("assignment/row_capacity.csv", header = TRUE)
tickets_sold <- read.csv("assignment/tickets_sold.csv", header = TRUE)

str(row_capacity)
```

```
## 'data.frame': 2 obs. of 2 variables:
## $ auditorium_row : int 1 2
## $ max_seats_per_row: int 10 10
```

```
str(tickets_sold)
```

```
## 'data.frame': 30 obs. of 8 variables:
## $ ticket_id : int 10001 10002 10003 10004 10005 10006 10007 10008 10009 10010 ...
## $ movie : Factor w/ 2 levels "Bond","Minions": 1 1 1 1 1 2 2 2 2 2 ...
## $ show_id : int 728 728 728 728 728 814 814 814 814 814 ...
## $ show_time : Factor w/ 4 levels "03/01/2016 20:00",...: 1 1 1 1 1 2 2 2 2 2 ...
## $ calendarweek : Factor w/ 2 levels "week1","week2": 1 1 1 1 1 1 1 1 1 1 ...
## $ auditorium_row: int 1 1 2 2 2 1 1 1 2 2 ...
## $ seat_number : int 5 6 3 7 8 1 2 3 5 6 ...
## $ booking_time : Factor w/ 12 levels "03/01/2016 19:43",...: 1 1 2 3 3 4 4 4 5 5 ...
```

Total_tickets_per_week_movie_row

Aggregate the number of individual tickets per show.

```
my_var <- tickets_sold[c("ticket_id", "movie", "calendarweek", "auditorium_row")]
head(my_var)
```

```
##   ticket_id  movie calendarweek auditorium_row
## 1    10001   Bond      week1           1
## 2    10002   Bond      week1           1
## 3    10003   Bond      week1           2
## 4    10004   Bond      week1           2
## 5    10005   Bond      week1           2
## 6    10006 Minions      week1           1
```

```
ft <- ftable(auditorium_row ~ calendarweek + movie, data = my_var)
ft
```

```
##               auditorium_row 1 2
## calendarweek movie
## week1      Bond              2 3
##           Minions            3 4
## week2      Bond              3 4
##           Minions            5 6
```

Seat_load_factor_per_week_movie_row

Divide the total tickets by the maximum capacity for that particular auditorium_row.

```
for(i in 1:length(row_capacity$auditorium_row)){
  ft[, i] <- ft[, i]/row_capacity$max_seats_per_row[row_capacity$auditorium_row == i]
}
ft
```

```
##               auditorium_row 1 2
## calendarweek movie
## week1      Bond              0.2 0.3
##           Minions            0.3 0.4
## week2      Bond              0.3 0.4
##           Minions            0.5 0.6
```

Seat_load_factor_per_week_row

Use the average seat_load_factor over the individual shows per calendarweek.

```
aggregate(Freq ~ calendarweek + auditorium_row, data = ft, FUN = mean)
```

```
##   calendarweek auditorium_row Freq
## 1      week1           1 0.25
## 2      week2           1 0.40
## 3      week1           2 0.35
## 4      week2           2 0.50
```