MarkGonsalvesAssignment2

Mark Gonsalves

9/6/2020

Overview of Project

1) I first mocked up a survey template in Excel.

https://github.com/mjgons/DATA607/blob/master/.github/workflows/Movie%20Reviews%20Survey%20Template.xlsx

2) Then I created a survey using google forms and emailed it out. https://docs.google.com/forms/d/e/1FAIpQLSeBOS6qeF43XUcsD6271HWFjzZ1HS7GzhfE_GMUKqI-9-y5IA/viewform?vc=0&c=0&w=1&flr=0 The survey includes an option for "unknown" if a person hasn't watched the movie. The survey requries a response to all questions so there won't be missing data.

3) I exported the survey from google forms to a CSV file on my computer. https://github.com/mjgons/DATA607/blob/master/.github/workflows/Movie%20Reviews.csv

4) I cleaned up the data in excel and saved it as a CSV file.

5) I uploaded the cleaned CSV file to my GitHub account.

https://github.com/mjgons/DATA607/blob/master/.github/workflows/Movie%20Reviews%20Clean.csv

6) I then created the table in MySQLWorkbench and populated the table.

Everything else is in the code below:

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(RMySQL)
```

```
## Loading required package: DBI
```

```
library(DBI)
```

Importing SQL into R and working with the data

I'm connecting the SQL database to R Studio. Looking at the data and then creating a new dataframe called db.

```
moviedb <- DBI::dbConnect(RMySQL::MySQL(), dbname = "Movies", user="root", port=3306, host='localhost',

dbListTables(moviedb)
```

```
## [1] "MovieSurvey"
```

```
glimpse(moviedb)
```

```
## Formal class 'MySQLConnection' [package "RMySQL"] with 1 slot
##   ..@ Id: int [1:2] 0 0
```

```
class(dbGetQuery(moviedb,"SHOW TABLES"))
```

```
## [1] "data.frame"
```

```
dbGetQuery(moviedb, "SHOW TABLES")
```

```
##   Tables_in_movies
## 1      MovieSurvey
```

```
dbGetQuery(moviedb, "EXPLAIN MovieSurvey")
```

```
##        Field        Type Null Key Default Extra
## 1  Survey_id         int   NO PRI    <NA>
## 2 first_name varchar(50)   NO        <NA>
## 3  last_name varchar(50)   NO        <NA>
## 4  Star_Wars varchar(50)   NO        <NA>
## 5      Sonic varchar(50)   NO        <NA>
## 6     Trolls varchar(50)   NO        <NA>
## 7     Onward varchar(50)   NO        <NA>
## 8  Spiderman varchar(50)   NO        <NA>
## 9      Mulan varchar(50)   NO        <NA>
```

```
db <-dbGetQuery(moviedb, "SELECT * FROM MovieSurvey")

colnames(db) <-c("ID", "First Name", "Last Name", "Star Wars", "Sonic", "Trolls", "Onward", "Spider-Man
head(db)
```

```
##   ID First Name Last Name Star Wars   Sonic  Trolls Onward Spider-Man   Mulan
## 1  1       Mark Gonsalves         3       3       1      1          2 Unknown
## 2  2       Josh Gonsalves         5       3 Unknown      2          4 Unknown
## 3  3     Justin Gonsalves         5       5 Unknown      3          2 Unknown
## 4  4      Becka Gonsalves         5 Unknown Unknown      2    Unknown Unknown
## 5  5    Rebecca   Chertok         5       4 Unknown      1          1 Unknown
## 6  6   Jee Hang       Lee         5       1       1      1          1 Unknown
```

Conclusion

The data shows that Becka hasn't watched Sonic. Most people enjoyed Sonic with 1 score of 5, 1 score of 4, 2 scores of 3 and 1 score of 1. Thus I would recommend Becka watches Sonic with her kids Josh and Justin while I do my homework.

```
#install.packages("kableExtra")
library(knitr)
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'
```

Table 1: Movie Reviews

First Name	Last Name	Star Wars	Sonic	Trolls	Onward	Spider-Man	Mulan
Mark	Gonsalves	3	3	1	1	2	Unknown
Josh	Gonsalves	5	3	Unknown	2	4	Unknown
Justin	Gonsalves	5	5	Unknown	3	2	Unknown
Becka	Gonsalves	5	Unknown	Unknown	2	Unknown	Unknown
Rebecca	Chertok	5	4	Unknown	1	1	Unknown
Jee Hang	Lee	5	1	1	1	1	Unknown

```
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
kable(db[1:6,2:9 ], caption = "Movie Reviews") %>%
kable_styling(bootstrap_options = c('striped', 'hover', 'responsive', 'condensed'))
```