Project 9 Solutions

(Abhimanyu Agarwal)

Collaborators: N/A

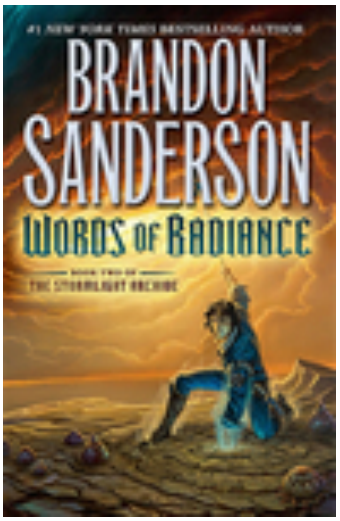
TA help:

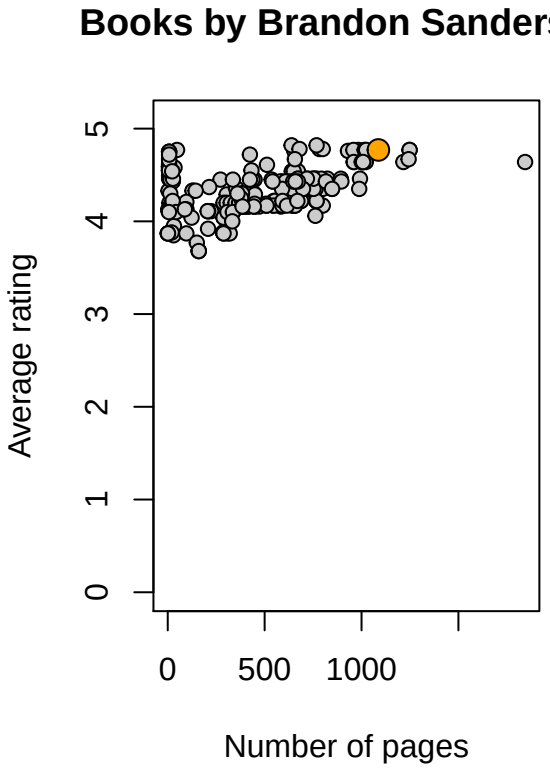
1) Melissa : Helped me go through Question 4 and 5.

Online resources used: N/A

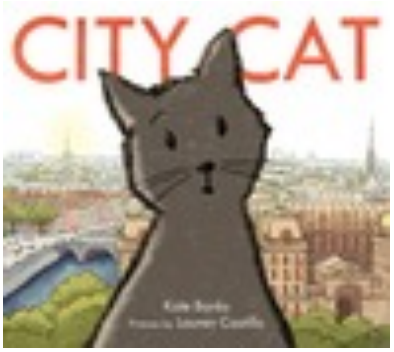
Question 1

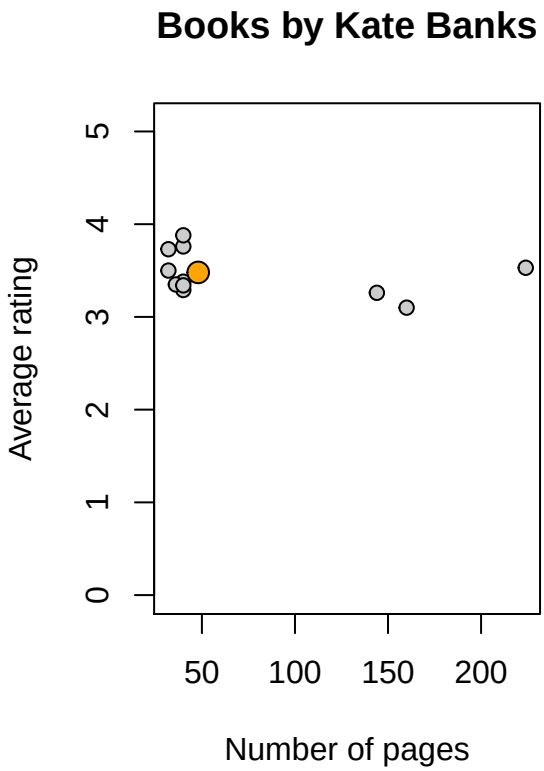
```
books <- read.csv("/class/datamine/data/goodreads/csv/goodreads_books.csv")
authors <- read.csv("/class/datamine/data/goodreads/csv/goodreads_book_authors.csv")

get_author_name <- function(my_authors_dataset, my_author_id){
  return(my_authors_dataset[my_authors_dataset$author_id==my_author_id,'name'])
}

fun_plot <- function(my_authors_dataset, my_books_dataset, my_book_id, display_cover=T) {
  book_info <- my_books_dataset[my_books_dataset$book_id==my_book_id,]
  all_books_by_author <- my_books_dataset[my_books_dataset$author_id==book_info$author_id,]
  author_name <- get_author_name(my_authors_dataset, book_info$author_id)

  img <- load.image(book_info$image_url)
  if(display_cover)
    {
      par(mfrow=c(1,2))
      plot(img, axes=FALSE)
    }
  plot(all_books_by_author$num_pages, all_books_by_author$average_rating,
       ylim=c(0,5.1), pch=21, bg='grey80',
       xlab='Number of pages', ylab='Average rating',
       main=paste('Books by', author_name))

  points(book_info$num_pages, book_info$average_rating,pch=21, bg='orange', cex=1.5)
}

#Function Call
fun_plot(authors, books, 17332218, display_cover = T)
```






```
fun_plot(authors, books, 17332219, display_cover = T)
```






```
#There are two functions provided as seen in the given code.
#The first one being - 'get_author_name ()' and the other being - 'fun_plot ()'.
```

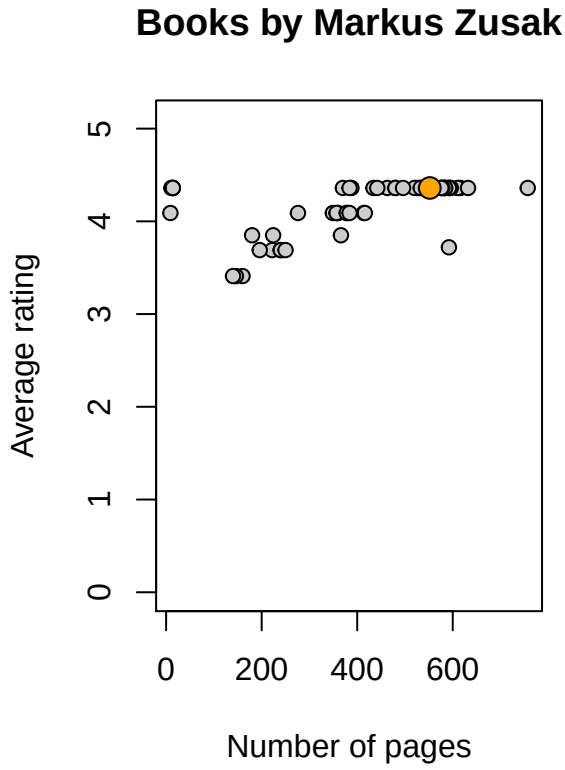
```
#Function - 'get_author_name()' takes in two arguments namely - 'my_authors_dataset' and 'my_author_id'
#Function - 'fun_plot()' takes in four arguments namely - 'my_authors_dataset, my_books_dataset, my_boo

#The function is displaying the display cover and giving a variation of the Average rate vs Number of P
#based on the books by the specific author the ID corresponds to.
#The display cover being displayed in an additional option given here. On setting that to false, the di
```

Question 2

```
books <- read.csv("/class/datamine/data/goodreads/csv/goodreads_books.csv")
authors <- read.csv("/class/datamine/data/goodreads/csv/goodreads_book_authors.csv")

get_author_name <- function(my_authors_dataset, my_author_id){
  return(my_authors_dataset[my_authors_dataset$author_id==my_author_id,'name'])
}

fun_plot <- function(my_authors_dataset, my_books_dataset, my_book_id, display_cover=T) {
  if(0 != sum(my_books_dataset$book_id==my_book_id))
  {
    book_info <- my_books_dataset[my_books_dataset$book_id==my_book_id,]
    all_books_by_author <- my_books_dataset[my_books_dataset$author_id==book_info$author_id,]
    author_name <- get_author_name(my_authors_dataset, book_info$author_id)

    img <- load.image(book_info$image_url)

    if(display_cover){
      par(mfrow=c(1,2))
      plot(img, axes=FALSE)
    }

    plot(all_books_by_author$num_pages, all_books_by_author$average_rating,
         ylim=c(0,5.1), pch=21, bg='grey80',
         xlab='Number of pages', ylab='Average rating',
         main=paste('Books by', author_name))

    points(book_info$num_pages, book_info$average_rating,pch=21, bg='orange', cex=1.5)
  }
  else
  {
    print('Book ID not found.')
  }
}

#Function Call
fun_plot(authors, books, 123, display_cover = T)
```

```
[1] "Book ID not found."
```

```
fun_plot(authors, books, 19063, display_cover = T)
```



Question 3

```
#Function to get_author_id corresponding to their names.
get_author_id <- function(my_authors_dataset,my_author_name){
  return(my_authors_dataset[my_authors_dataset$name==my_author_name,'author_id'])
}
#Function Call
get_author_id(authors, "Brandon Sanderson") # 38550
```

```
[1] 38550
```

```
get_author_id(authors, "J.K. Rowling") # 1077326
```

```
[1] 1077326
```

Question 4

```
books <- read.csv("/class/datamine/data/goodreads/csv/goodreads_books.csv")

#We add another parameter - 'my_books_dataset' which basically is the name of the dataset being passed
#The search parameter is word. This function looks for specific code in particular books and then retur
#It returns the title of the books that make use of that particular word.

search_books_for_word <- function(my_books_dataset, word) {
  return(my_books_dataset[grepl(word, my_books_dataset$description, fixed=T),]$title)
}
```

```
#Search word here: 'youth'
search_books_for_word(books, 'youth')
```

###Question 5

```r
#Returns the name of the book by taking two parameters namely 'my_books_dataset' and 'my_book_id'.
#Reloading the dataset here
books <- read.csv("/class/datamine/data/goodreads/csv/goodreads_books.csv")

get_book_name <- function(my_books_dataset, my_book_id){
  return(my_books_dataset[my_books_dataset$book_id==my_book_id, 'title'])
}

get_book_name(books, 5333265) # 5333265
```

```
[1] "W.C. Fields: A Life on Film"
```

```r
#The corresponding book name is "W.C. Fields: A Life on Film"
```

Submitting deliverables: project09.RMD, project09.R and project09.pdf

Pledge

By submitting this work I hereby pledge that this is my own, personal work. I've acknowledged in the designated place at the top of this file all sources that I used to complete said work, including but not limited to: online resources, books, and electronic communications. I've noted all collaboration with fellow students and/or TA's. I did not copy or plagiarize another's work.

> As a Boilermaker pursuing academic excellence, I pledge to be honest and true in all that I do. Accountable together - We are Purdue.