

Registration Number- 19BCE2119

Name- Gaurav Kumar Singh

Course- Data Visualization (L13+L14)

Digital Assignment 3

- Importing libraries and dataset

```
library(ggplot2)
```

```
library(dplyr)
```

```
library(plotly)
```

```
library(hrbrthemes)
```

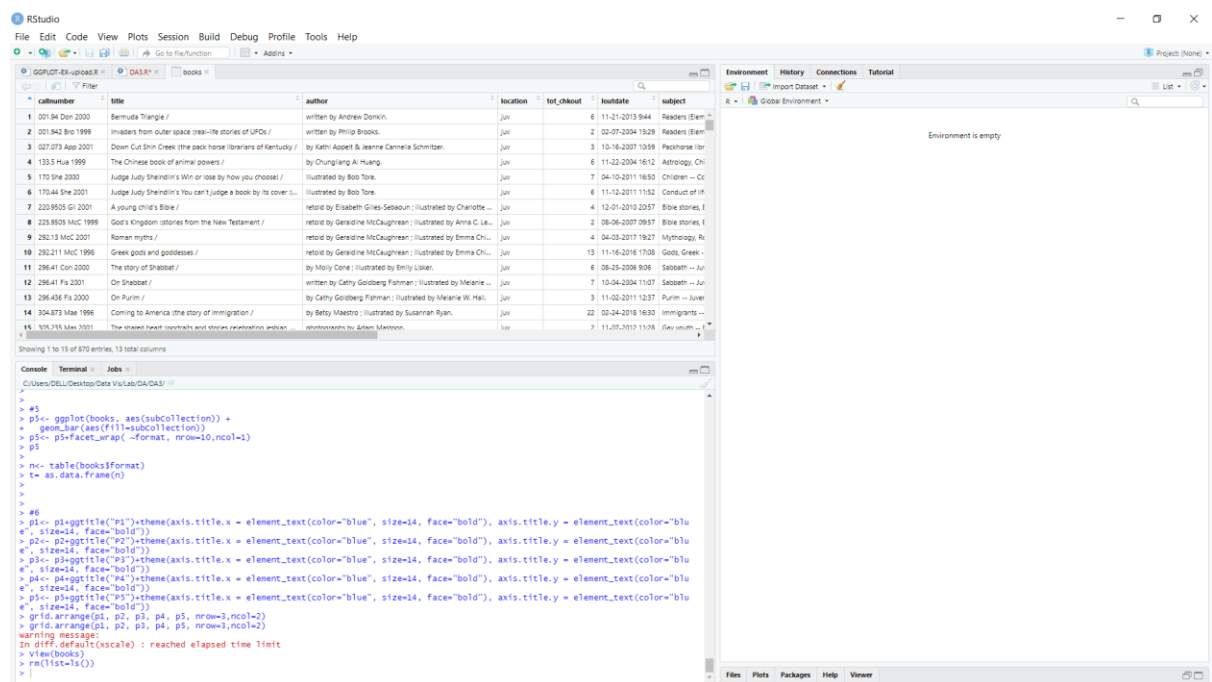
```
library(rlang)
```

```
library(tidyr)
```

```
library(gridExtra)
```

```
books <- read.csv("C:\\Users\\DELL\\Desktop\\Data Vis\\Lab\\DA\\DA3\\books.csv")
```

```
View(books)
```



- Remove NA values

#Method1

```
books<-na.omit(books, na.action = "omit", cols="call_class")
```

#Method2

```
books<-subset(books,!is.na(call_class))
```

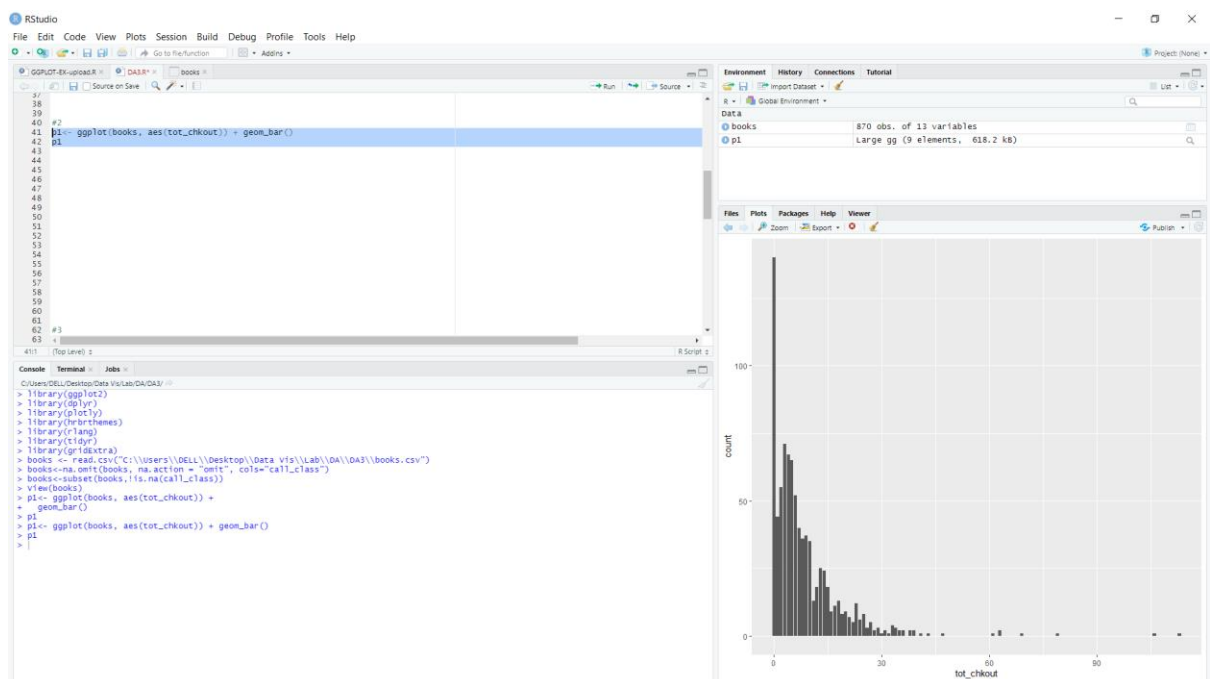
The screenshot shows the RStudio interface. The Environment pane on the right displays the 'books' dataset with 870 observations and 13 variables. The console on the left shows the following code being executed:

```
library(ggplot2)
library(dplyr)
library(plotly)
library(hrbrthemes)
library(clang)
library(tidy)
library(ggExtra)
books <- read.csv("C:\\Users\\DELL\\Desktop\\Data Vis\\Lab\\DA3\\books.csv")
books<-na.omit(books, na.action = "omit", cols="call_class")
books<-subset(books,!is.na(call_class))
view(books)
```

- Visualize frequency distribution of checkouts in the book dataset

```
p1<- ggplot(books, aes(tot_chkout)) + geom_bar()
```

p1



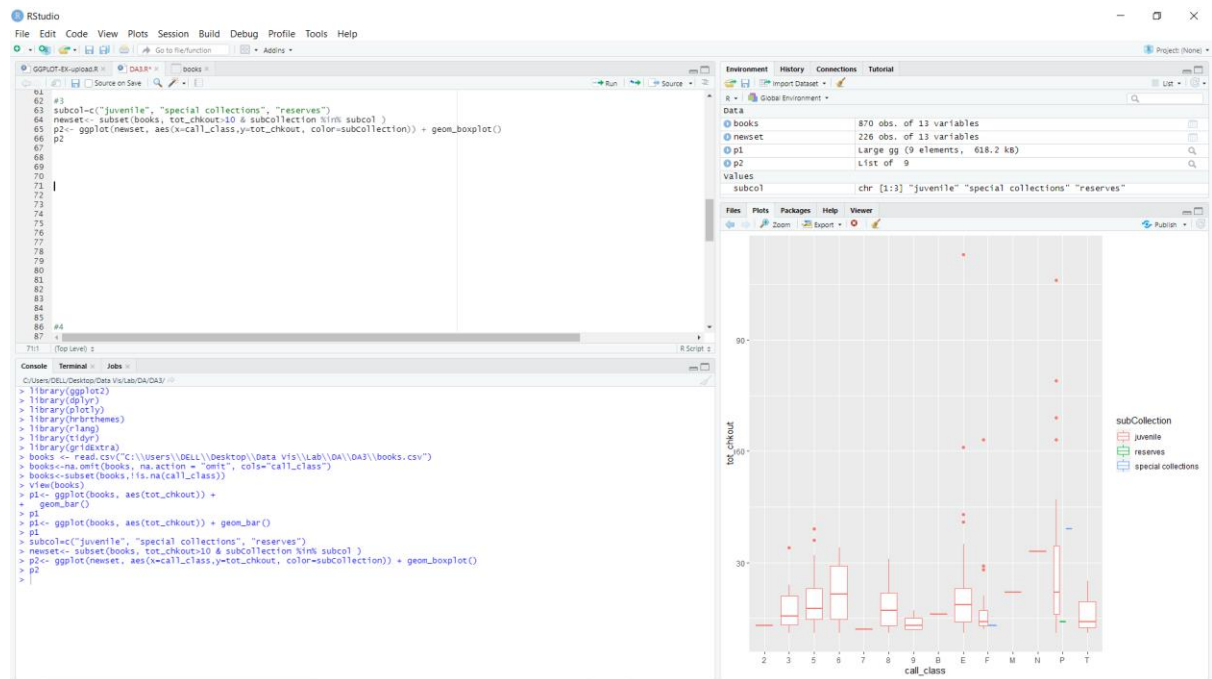
- Visualize boxplot plot high usage books by call number class [Hint: derive new variable high usage books which have more than 10 checkouts] and color by subCollection- any three category

```
subcol=c("juvenile", "special collections", "reserves")
```

```
newset<- subset(books, tot_chkout>10 & subCollection %in% subcol )
```

```
p2<- ggplot(newset, aes(x=call_class,y=tot_chkout, color=subCollection)) + geom_boxplot()
```

p2



- Visualize stacked bar and group chart high usage books by call number class [Hint: derive new variable high usage books which have more than 10 checkouts] and color by subCollection- any four category

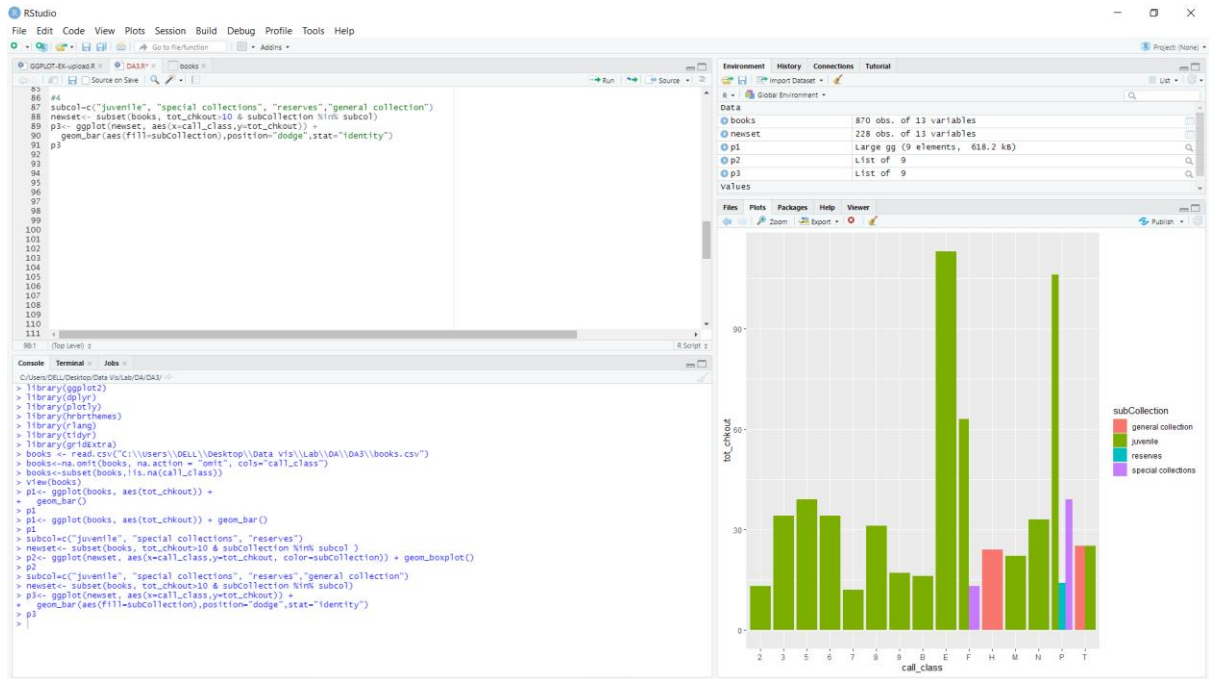
```
subcol=c("juvenile", "special collections", "reserves", "general collection")
```

```
newset<- subset(books, tot_chkout>10 & subCollection %in% subcol)
```

```
p3<- ggplot(newset, aes(x=call_class,y=tot_chkout)) +
```

```
  geom_bar(aes(fill=subCollection),position="dodge",stat="identity")
```

p3



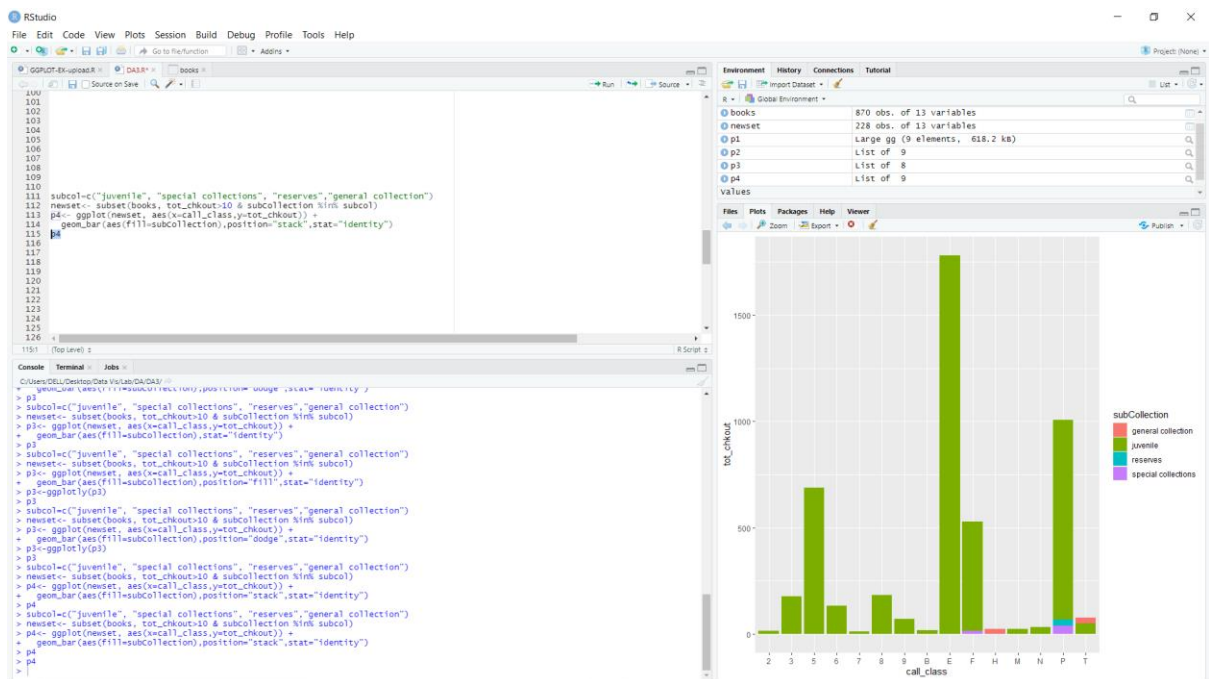
```
subcol=c("juvenile", "special collections", "reserves", "general collection")
```

```
newset= subset(books, tot_chkout>10 & subCollection %in% subcol)
```

```
p4<- ggplot(newset, aes(x=call_class,y=tot_chkout)) +
```

```
  geom_bar(aes(fill=subCollection),position="stack",stat="identity")
```

```
p4
```

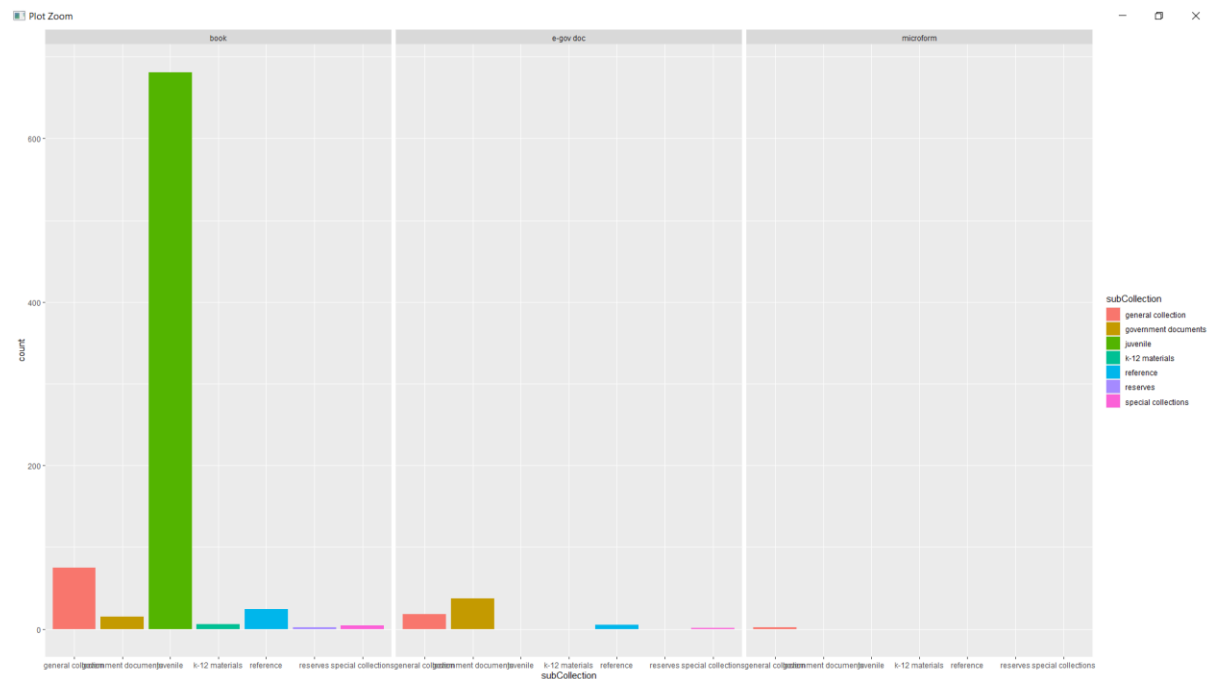


- Create a bar plot that depicts the number of items in each sub-collection, faceted by format. Arrange sub plots horizontally.

```
p5<- ggplot(books, aes(subCollection)) +
  geom_bar(aes(fill=subCollection))

p5<- p5+facet_wrap( ~format, nrow=1,ncol=3)

p5
```



- Apply various themes in the above plots

```
p1<- p1+ggtitle("P1")+theme(axis.title.x = element_text(color="blue", size=14, face="bold"),
axis.title.y = element_text(color="blue", size=14, face="bold"))

p2<- p2+ggtitle("P2")+theme(axis.title.x = element_text(color="blue", size=14, face="bold"),
axis.title.y = element_text(color="blue", size=14, face="bold"))

p3<- p3+ggtitle("P3")+theme(axis.title.x = element_text(color="blue", size=14, face="bold"),
axis.title.y = element_text(color="blue", size=14, face="bold"))

p4<- p4+ggtitle("P4")+theme(axis.title.x = element_text(color="blue", size=14, face="bold"),
axis.title.y = element_text(color="blue", size=14, face="bold"))

p5<- p5+ggtitle("P5")+theme(axis.title.x = element_text(color="blue", size=14, face="bold"),
axis.title.y = element_text(color="blue", size=14, face="bold"))

grid.arrange(p1, p2, p3, p4, p5, nrow=3,ncol=2)
```

