Facebook Friends Birthday Data Analysis

## 1. Loading the data file and required packages

#install.packages('lubridate')  
library(lubridate)

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

fb <- read.csv('birthdays.csv')

## 2. Lets take look at the dataset using the head command.

head(fb)

## Title Start End  
## 1 Rahul Rao's birthday 06/28/2017 00:00 06/29/2017 00:00  
## 2 Bhargav Raj's birthday 06/28/2017 00:00 06/29/2017 00:00  
## 3 Salini Singh's birthday 06/28/2017 00:00 06/29/2017 00:00  
## 4 Harshitha Tadepalli's birthday 06/28/2017 00:00 06/29/2017 00:00  
## 5 Ashok Vishwanath's birthday 07/02/2017 00:00 07/03/2017 00:00  
## 6 Nikhil Voruganti's birthday 07/04/2017 00:00 07/05/2017 00:00  
## Duration  
## 1 24:00  
## 2 24:00  
## 3 24:00  
## 4 24:00  
## 5 24:00  
## 6 24:00

## 3. Lets remove the columns End and duration, as they might not be much useful for exploration.

fb <- subset(fb,select = -c(End,Duration))  
head(fb)

## Title Start  
## 1 Rahul Rao's birthday 06/28/2017 00:00  
## 2 Bhargav Raj's birthday 06/28/2017 00:00  
## 3 Salini Singh's birthday 06/28/2017 00:00  
## 4 Harshitha Tadepalli's birthday 06/28/2017 00:00  
## 5 Ashok Vishwanath's birthday 07/02/2017 00:00  
## 6 Nikhil Voruganti's birthday 07/04/2017 00:00

## 4. Lets convert the Start column data to date format.

fb$Start <- as.Date(fb$Start,format='%m/%d/%Y')  
head(fb$Start)

## [1] "2017-06-28" "2017-06-28" "2017-06-28" "2017-06-28" "2017-07-02"  
## [6] "2017-07-04"

## 5. Lets add month column to dataset.

fb$month <- format(fb$Start,'%b')  
head(fb)

## Title Start month  
## 1 Rahul Rao's birthday 2017-06-28 Jun  
## 2 Bhargav Raj's birthday 2017-06-28 Jun  
## 3 Salini Singh's birthday 2017-06-28 Jun  
## 4 Harshitha Tadepalli's birthday 2017-06-28 Jun  
## 5 Ashok Vishwanath's birthday 2017-07-02 Jul  
## 6 Nikhil Voruganti's birthday 2017-07-04 Jul

# Now Lets Explore The Dataset.

## 7. How many birthdays are in each month?

table(fb$month)

##   
## Apr Aug Dec Feb Jan Jul Jun Mar May Nov Oct Sep   
## 25 34 33 28 41 39 47 43 47 33 51 44

## 8. How many people share My birthday?

myBdayDate = as.Date('10/12/2017','%m/%d/%Y')  
shareMyBday <- subset(fb,fb$Start == myBdayDate)  
shareMyBday

## Title Start month  
## 140 Abhay Pratap Singh's birthday 2017-10-12 Oct  
## 141 Mohammad Umair Siddiqui's birthday 2017-10-12 Oct  
## 142 Chaitanya Kumar Madala's birthday 2017-10-12 Oct

## 9. Which month contains the most number of birthdays?

names(which(table(fb$month) == max(table(fb$month))))

## [1] "Oct"

## 10. Plot a bar chart to show number of birthdays in each month

#install.packages("ggplot2")  
library(ggplot2)

ggplot(data=fb,aes(x=month))+  
 geom\_bar(color='black',fill='#F79420')

