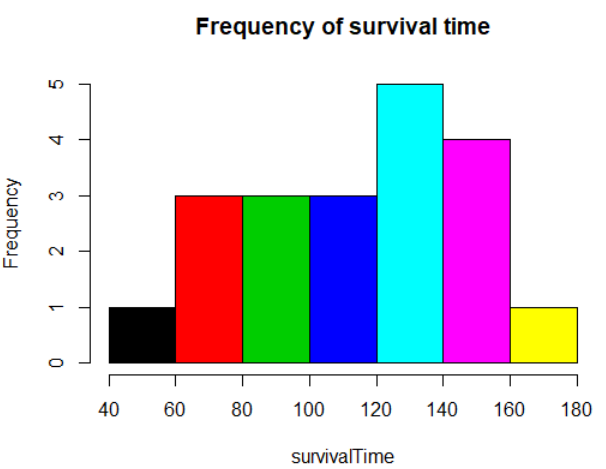
1. > install.packages("PASWR")

> library("PASWR")

> hist(Rat$survival.time,xlab = 'survivalTime',main='Frequency of survival time',col=rep(1:10))



* 1. > data = read.table("C:\\Users\\Administrator\\Desktop\\statistical computing\\HW02\\Gapminder.csv",header=T,sep=',')  
     > head(data,5)

country continent year lifeExp pop gdpPercap

1 Afghanistan Asia 1952 28.801 8425333 779.4453

2 Afghanistan Asia 1957 30.332 9240934 820.8530

3 Afghanistan Asia 1962 31.997 10267083 853.1007

4 Afghanistan Asia 1967 34.020 11537966 836.1971

5 Afghanistan Asia 1972 36.088 13079460 739.9811

* 1. > install.packages("ggplot2")

> library("ggplot2")

> table(data[,"continent"])

Africa Americas Asia Europe Oceania

618 300 396 360 24

> africa = data[(continent=="Africa"),"country"]

> americas = data[(continent=="Americas"),"country"]

> asia = data[(continent=="Asia"),"country"]

> europe = data[(continent=="Europe"),"country"]

> oceania = data[(continent=="Oceania"),"country"]

> africa\_unique = length(unique(africa))

> americas\_unique = length(unique(americas))

> asia\_unique = length(unique(asia))

> europe\_unique = length(unique(europe))

> oceania\_unique = length(unique(oceania))

> africa\_unique

[1] 57

> americas\_unique

[1] 25

> asia\_unique

[1] 33

> europe\_unique

[1] 30

> oceania\_unique

[1] 2

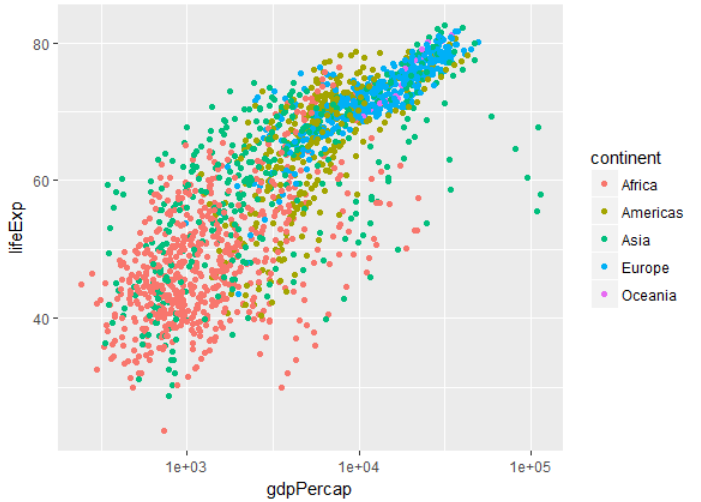
> p<-ggplot(data, aes(x = gdpPercap, y = lifeExp))

* 1. > p + geom\_point()

> p + geom\_point() + scale\_x\_log10()

> p <- p + scale\_x\_log10()

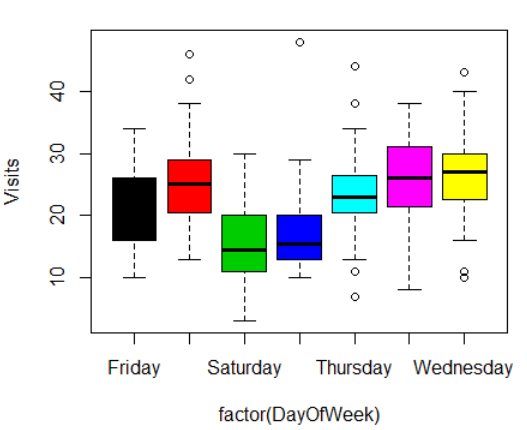
> p + geom\_point()

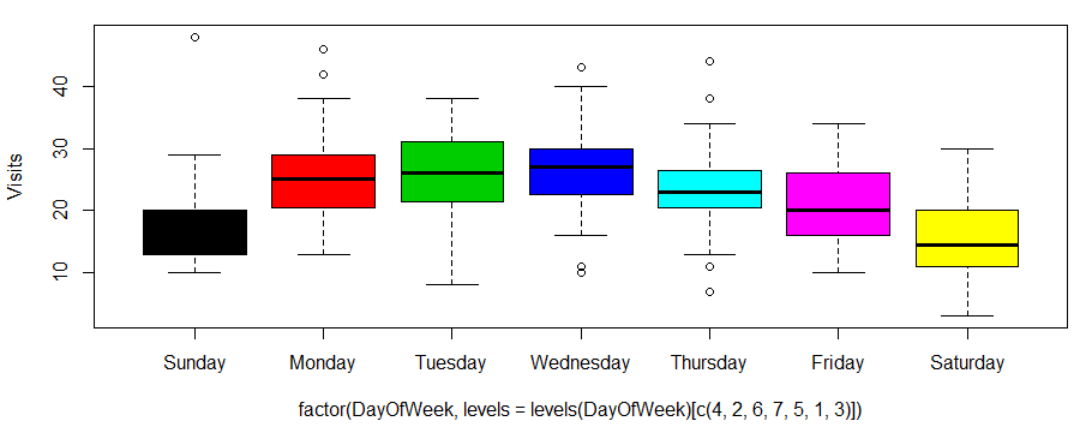
> p + geom\_point(aes(color=continent))  


> data = read.csv("C:\\Users\\Administrator\\Desktop\\statistical computing\\HW02\\website traffic.csv")

> attach(data)

> plot(Visits~factor(DayOfWeek),col=c(1:10))



>plot(Visits~factor(DayOfWeek,levels=levels(DayOfWeek)[c(4,2,6,7,5,1,3)]),col=c(1:10))  


* 1. > summary(data)

DayOfWeek MonthDay Year Visits

Friday :30 August 1 : 1 Min. :2009 Min. : 3.00

Monday :31 August 10: 1 1st Qu.:2009 1st Qu.:16.25

Saturday :30 August 11: 1 Median :2009 Median :22.00

Sunday :30 August 12: 1 Mean :2009 Mean :22.23

Thursday :31 August 13: 1 3rd Qu.:2009 3rd Qu.:27.75

Tuesday :31 August 14: 1 Max. :2009 Max. :48.00

Wednesday:31 (Other) :208

* 1. > name = c("Ana","Brian","Cathy","Dough","John","Lucas","Marcus","Nabin","William","Zoe")

> test1 = c(56,78,87,89,95,98,59,78,87,98)

> table1 = data.frame(name,test1)

> test2 = c(86,67,78,89,87,67,94,78,81,83)

> table2 = data.frame(name,test2)

> data = merge(table1,table2)

> data

name test1 test2

1 Ana 56 86

2 Brian 78 67

3 Cathy 87 78

4 Dough 89 89

5 John 95 87

6 Lucas 98 67

7 Marcus 59 94

8 Nabin 78 78

9 William 87 81

10 Zoe 98 83

> length(which(test2>test1))

[1] 2

> length(which(test1>test2))

[1] 6

> length(which(test1==test2))

[1] 2

> mean(test1)

[1] 82.5

> mean(test2)

[1] 81

> sd(test1)

[1] 14.96106

> sd(test2)

[1] 8.869423

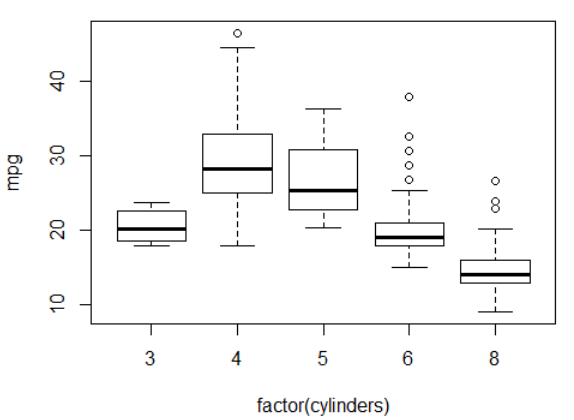
> data = read.table("https://archive.ics.uci.edu/ml/machine-learning-databases/auto-mpg/auto-mpg.data")

> names(data) <- c("mpg","cylinders",'displacement','horsepower','weight','acceleration','model year','origin','car name')

data = read.table("https://archive.ics.uci.edu/ml/machine-learning-databases/auto-mpg/auto-mpg.data",na.strings = '?')  
> newData = na.omit(data)

> dim(newData)

[1] 392 9

> plot(mpg~factor(cylinders))  


> vote = read.table("https://www.stat.berkeley.edu/~statlabs/data/vote.data",header=T)  
> ncol(vote)

[1] 4

> names(vote)

[1] "precinct" "candidate" "race" "income"

>hist(race,xaxt='n')

>axis(1,at=seq(0,5,by=1),labels=c('missing','white','hispanic','black','asian','Other'))  
