

Lec2

HL

8/30/2019

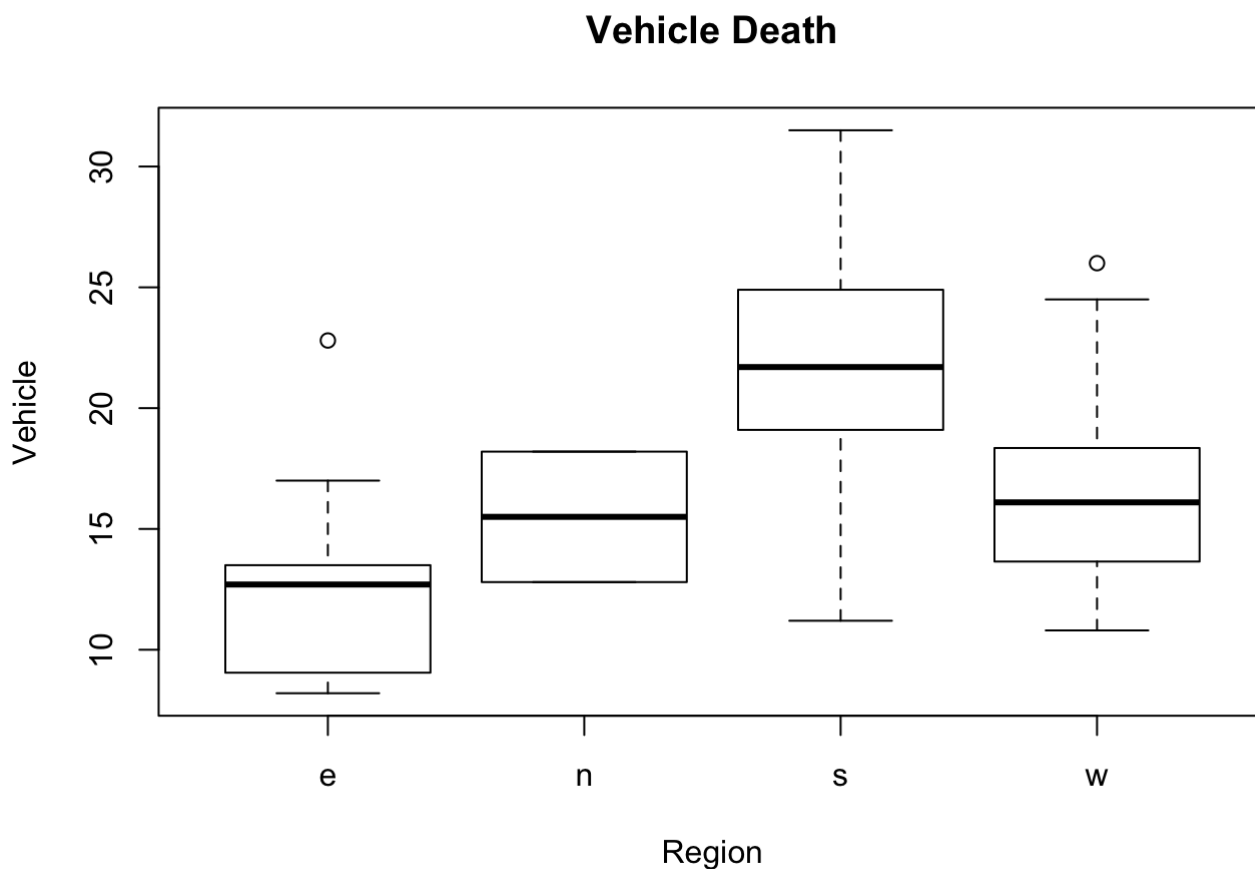
Box plot using vehicle death example

This is draw the box plot showing vehicle death by regions

```
library(ggplot2)

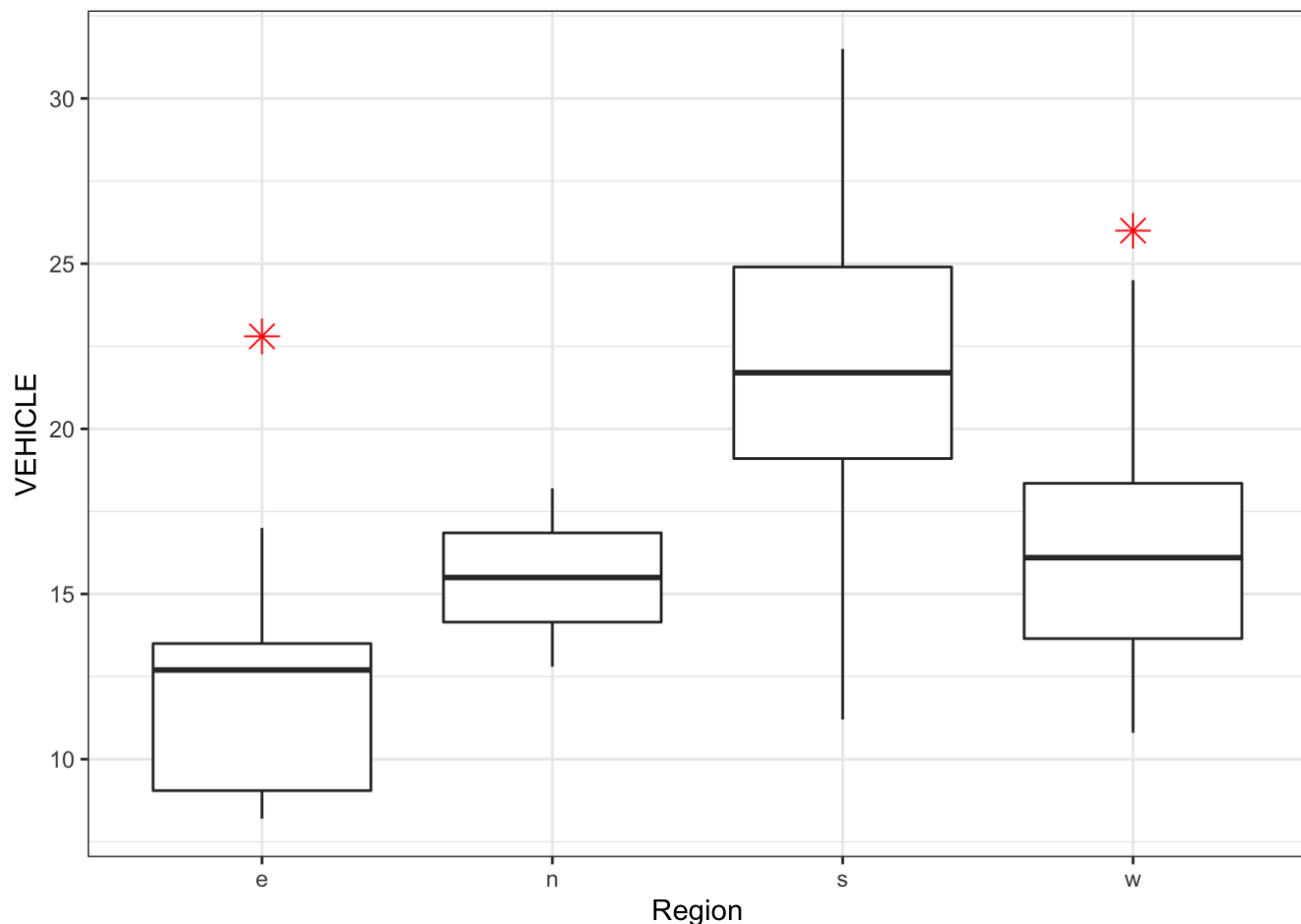
setwd("/Users/hongyeliu/Desktop/CS361_R_MarkDown")
Data_Dt <- read.table("DeathData.txt", header=TRUE)
df2 <- Data_Dt

### do a plot using standard boxplot
boxplot(VEHICLE~Region, data=df2, main="Vehicle Death", xlab="Region", ylab="Vehicle")
```



```
### do a plot with ggplot2 for death data wrt vehicle

# Change outlier, color, shape and size
p<-ggplot(df2, aes(x=Region, y=VEHICLE)) +
  geom_boxplot(outlier.colour="red", outlier.shape=8,
              outlier.size=4)+theme_bw()+
  theme(legend.position="top")
p
```

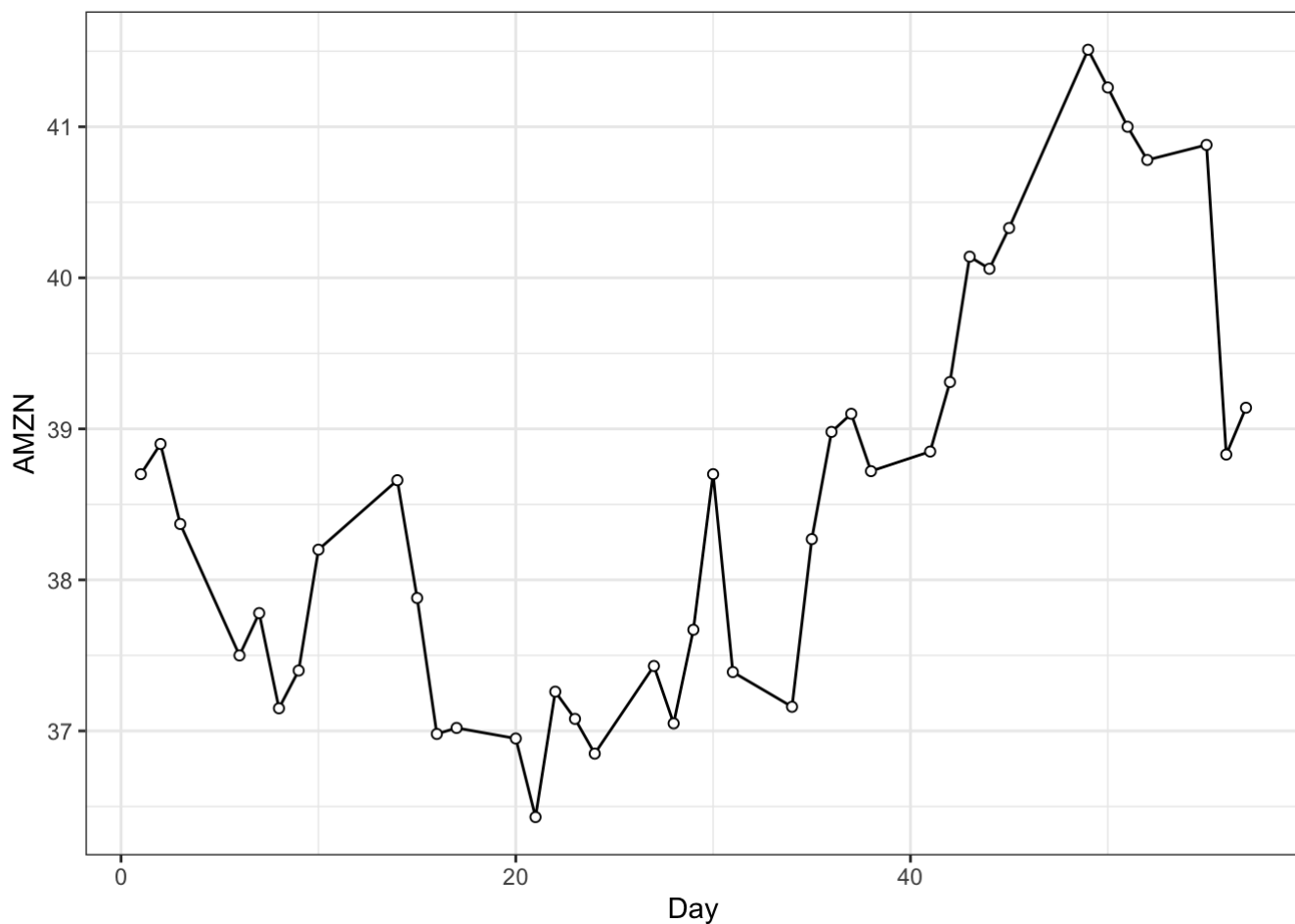


```
### end box plots
```

Plot time series using stock example

```
### time series of stocks
setwd("/Users/hongyeliu/Desktop/CS361_R_MarkDown")
Data_Stk <- read.table("Stock_sht.txt", header=TRUE)
df2 <- Data_Stk

p<-ggplot(df2, aes(x=Day, y=AMZN)) +
  geom_line(aes())+
  geom_point(shape = 21, colour = "black", fill = "white") +theme_bw()+
  theme(legend.position="top")
p
```



Example of scatter plots using body fat example

```
##### scatter plots
Data_Fat <- read.table("BodyFat.txt", header=TRUE)
plotdata <- Data_Fat
fig <- ggplot(plotdata, aes(HEIGHT, WEIGHT))

fig <- fig + geom_point(aes(x = HEIGHT, y = WEIGHT,color= "blue"), alpha = 0.6, size = 2
)

fig <- fig + ggtitle(paste(" "))
fig <- fig + xlab("HEIGHT") + ylab("WEIGHT")
fig <- fig + theme_bw() + theme( panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), axis.line = element_line(colour = "black"),legen
d.position = "none")
print(fig, newpage = FALSE)

which(Data_Fat$HEIGHT<40)
```

```
## [1] 42
```

```
which(Data_Fat$WEIGHT >300)
```

```
## [1] 39
```

```
Data_Fat[42,]
```

```
##      IDNO BODYFAT DENSITY AGE WEIGHT HEIGHT ADIPOSITIVITY NECK CHEST ABDOMEN
## 42     42     31.7   1.025  44    205   29.5         29.9 36.6   106   104.3
##      HIP THIGH KNEE ANKLE BICEPS FOREARM WRIST
## 42 115.5   70.6 42.5  23.7   33.6    28.7  17.4
```

```
Data_Fat <- Data_Fat[-c(39,42),]
plotdata <- Data_Fat
### normalize to standard coordinates

plotdata$HEIGHT <- plotdata$HEIGHT - mean(plotdata$HEIGHT)
plotdata$HEIGHT <- plotdata$HEIGHT /sd(plotdata$HEIGHT)

plotdata$WEIGHT <- plotdata$WEIGHT - mean(plotdata$WEIGHT)
plotdata$WEIGHT <- plotdata$WEIGHT /sd(plotdata$WEIGHT)

fig <- ggplot(plotdata, aes(HEIGHT, WEIGHT))

fig <- fig + geom_point(aes(x = HEIGHT, y = WEIGHT,color= "blue"), alpha = 0.6, size = 2
) +stat_density2d(color = "gray45")

fig <- fig + ggtitle(paste(" "))
fig <- fig + xlab("HEIGHT") + ylab("WEIGHT")
fig <- fig + theme_bw() + theme( panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(), axis.line = element_line(colour = "black"),legen
d.position = "none")
print(fig, newpage = FALSE)
```

