Session 2 practical solution

Q1

```
x <- rnorm(10000,0,sqrt(2))
y <- rnorm(10000,2,sqrt(3))
z <- x + y

(a)-(d)
par(mfrow=c(2,2))
plot(density(x), xlab="x", main="Distribution of X", ylim=c(0,0.4) , las=1)
plot(density(y), xlab="x", main="Distribution of Y", ylim=c(0,0.4), xlim=c(-6,10), las=1)
plot(x,y, xlab="x", ylab="y", main="Scatterplot between X and Y", cex=0.1, las=1)
plot(density(z) , xlab="x + y", main="Distribution of X + Y", ylim=c(0,0.4) , xlim=c(-
6,10), las=1)

(e)
mean(x)
var(x)

mean(y)
var(y)

mean(z)
var(z)

(f)
plot(x,z, xlab="x", ylab="y", main="Scatterplot between X and X+Y", cex=0.1, las=1)

[alternative plot]
library(RColorBrewer)
library(hexbin)
rf <- colorRampPalette(rev(brewer.pal(11,'Spectral')))
r <- rf(32)

rand.xz <- data.frame(x,z)
h <- hexbin(rand.xz)
plot(h, colramp=rf)
```

Q2

```
mvc <- read.csv("http://web.hku.hk/~ehylau/mvc.csv")

(a)
mvc$height.cat <- cut(mvc$height, c(155,167,172,180), include.lowest=T)
```

(b)

```
summary(lm(MVC~age+height.cat, data=mvc))
```

(c)

```
mvc$height.cat2 <- relevel(mvc$height.cat, ref="(167,172]")
summary(lm(MVC~age+height.cat2, data=mvc))
```

or

```
summary(lm(MVC~age+relevel(height.cat,ref=2), data=mvc))
```

(d)

Estimated association between MVC and age, height from a linear regression model

Variable	Estimate	95% CI	p-values
Age	-3.3	(-6.2, -0.5)	0.022
Height (cm)			
155-167	38.0	(-45.8, 121.7)	0.364
168-171	ref		
172-180	98.6	(23.3, 174.0)	0.012
R^2= 0.32	adj. R^2= 0.26		

(e)

```
summary(lm(MVC~age+relevel(height.cat,ref=2), data=mvc, subset=(age<=40)))
```

It suggests that there may be a U-shape relation between height and MVC in younger adults.