**1. In class time, we learn about government data analysis.**

# (1) Compare the precipitation from 1999 to 2005. Are they different?

# read in data

P <- read.csv("C:\\Users\\User\\Desktop\\UNdata\_precipitation.csv")

P

attach(P)

Value

tapply(Value, Year, mean)

y1999 <- P[Year==1999,]

y2005 <- P[Year==2005,]

hist(log10(y1999$Value))

hist(log10(y2005$Value))

t.test(log10(y1999$Value),log10(y2005$Value))

# (2) Precipitation for all countries over time.

m <- tapply(Value,Year,mean)

plot(x=levels(factor(Year)), y=m)

abline(lm(Value ~ Year, data=P))

**2. I find another government data analysis example.**

**Data : Proportion of land area covered by forest, percentage.**

**896 records, from Millennium Development Goals Database | United Nations Statistics Division**

# read in data

yt=read.csv("C:\\Users\\User\\Desktop\\UNdata\_Export\_20150702\_030426519.csv")

yt

# (1) Compare the proportion of forest area from 1990 to 2010. Are they different?

attach(yt)

Value

tapply(Value, Year, mean)

y1990 <- P[Year==1990,]

y2010 <- P[Year==2010,]

hist(y1990$Value)

hist(y2010$Value)

t.test(y1990$Value,y2010$Value)

# (2) The proportion of forest area for all countries over time.

m <- tapply(Value,Year,mean)

plot(x=levels(factor(Year)), y=m)

abline(lm(Value ~ Year, data=yt))