rm(list=ls(all=TRUE))

prc <- read.csv("C:/Users/Gayan/Desktop/SKKU\_DataScience\_2015-master/data\_sets/UNdata\_precipitation.csv")

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

y1999 <- prc[prc$Year==1999,]

y2005 <- prc[prc$Year==2005,]

hist(log10(y1999$Value), main = "Percipitaion in 1999", xlab="Precipitation")

hist(log10(y2005$Value), main = "Percipitaion in 2005", xlab="Precipitation")

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

Welch Two Sample t-test

data: y1999$Value and y2005$Value

t = -1.1474, df = 101.52, p-value = 0.2539

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-371883.91 99319.18

sample estimates:

mean of x mean of y

177543.9 313826.3

According to the result of the t-test, it is found that the precipitations are different. But not

significant

**# (2) Precipitation for all countries over time.**

aggCountry <- aggregate(prc$Value, list(prc$Country), mean)

hist(log10(as.double(aggCountry$x)), main = "Percipitaion of All countries", xlab="Precipitation")