Q1

Username: 44373082

name of the public repository: ACST_s1_2019

File name: ACST890 test

Q2

Linear regression is a widely used statistical learning method today, which is simple and practical. A simple linear regression can be expressed mathematically as $Y \approx \beta 0 + \beta 1*X$, where X is the predictor variable, Y is the response variable, $\beta 0$ is the intercept term, and $\beta 1$ is the slope. In practice, $\beta 0$ and $\beta 1$ are unknown; therefore, before we make predictions, it is need to use the data to estimate the coefficients.

Simple linear regression is a useful way to predict response. However, in practice, we often have more than one predictor. Thus, a better way is to extend the simple linear regression model

Then the multiple Linear regression model takes the form

$Y = \beta 0 + \beta 1*X1 + \beta 2*X2 + \cdots + \beta p*Xp + \epsilon$,

Simple model: The true relationship may not be linear. there may be other variables that cause Y changes, and there may be measurement errors. It is usually assume that the error term is independent of X. The quality of the linear regression fit usually is evaluated using two related quantities: residual standard error (RSE) and R^2 statistics.

Q3

```
dataset <- read.csv("singapore.economy.csv")
newdata <- na.omit(dataset)
plot(dataset$time,dataset$gdp,main="Singapore GDP growth",xlab="Time",ylab="GDP(%)")
m <- tapply(newdata$gdp,newdata$period,mean)
sd <- tapply(newdata$gdp,newdata$period,sd)
install.packages("data.table")
library(data.table)
stat.table <- data.table(m,sd)
pairs(~gdp+exp+epg+hpr,newdataset)
fit <- lm(newdata$gdp ~ newdata$exp,data=newdata)
summary(fit)
fit <- lm((newdata$gdp ~
newdata$exp+newdata$epg+newdata$hpr+newdata$oil+newdata$gdpus+newdata$crd,data=ne
wdata)
quantile(newdata$gdp,0.05)
```