

```
#####
# L4.1. ANOVA
#####
## clear memory
rm(list = ls())

## load data
Data = read.table("http://edu.sablab.net/data/txt/depression2.txt", header=T, sep=
"\t")
str(Data)
DataGH = Data[Data$Health == "good", ]

## build 1-factor ANOVA model
res1 = aov(Depression ~ Location, DataGH)
summary(res1)

## build the ANOVA model
res2 = aov(Depression ~ Location + Health + Location*Health, Data)

## show the ANOVA table
summary(res2)

## Load function
source("http://sablab.net/scripts/drawANOVA.r")
x11()
drawANOVA(res2)

#####
# L4.2. Linear regression
#####

Data = read.table("http://edu.sablab.net/data/txt/cells.txt", header=T, sep="\t")
Data

## calculate LinModel
x = Data$Temperature
y = Data$Cell.Number
res = lm(y~x)
## another way:
res = lm(Cell.Number~Temperature, data=Data)

## see the results
res
summary(res)
cor(x, y)^2
# draw the data
x11()
plot(x, y)

# draw the regression and its confidence (95%)
lines(x, predict(res, int = "confidence")[, 1], col=4, lwd=2)
lines(x, predict(res, int = "confidence")[, 2], col=4)
lines(x, predict(res, int = "confidence")[, 3], col=4)
```

```

# draw the prediction for the values (95%)
lines(x, predict(res, int = "pred")[, 2], col=2)
lines(x, predict(res, int = "pred")[, 3], col=2)

##-----
## L4.2.2 Regression Analysis
##-----

Data = read.table("http://edu.sablab.net/data/txt/leukemia.txt",
                  header=T, sep="\t", quote="\")
str(Data)

x11()
plot(density(Data$WBC))

## log WBC and time
Data$WBC = log10(Data$WBC)

x11()
plot(density(Data$WBC))

## separate to two datasets
data1 = Data[Data$AG == "Positive", ]
data2 = Data[Data$AG == "Negative", ]

##-----
## Q1. Build scatter plot.

## build scatter plots
x11()
par(mfcol=c(1, 2))
plot(data1$WBC, data1$Survival, pch=19, col=3)
plot(data2$WBC, data2$Survival, pch=19, col=4)

##-----
## Q2. Linear model. Estimations.

## build linear models
lm1 = lm(Survival ~ WBC, data1)
lm2 = lm(Survival ~ WBC, data2)
summary(lm1)
summary(lm2)

## estimate survival for patient with WBC = 20000
wbc = log2(20000)
surv1 = lm1$coefficients[1]+lm1$coefficients[2]*wbc

## build scatter plots
x11()

```

```
plot(data1$WBC, data1$Survival, pch=19, col=3, main="Survival")
lines(data1$WBC, predict(lm1, int = "confidence")[, 1], col=4, lwd=2)
lines(data1$WBC, predict(lm1, int = "confidence")[, 2], col=4)
lines(data1$WBC, predict(lm1, int = "confidence")[, 3], col=4)
points(wbc, surv1, col=2, pch=19)

lines(data1$WBC, predict(lm1, int = "pred")[, 2], col=2)
lines(data1$WBC, predict(lm1, int = "pred")[, 3], col=2)
```