==

Inferencial Data Analysis

==

1. Load the ToothGrowth data and perform some basic exploratory data analyses

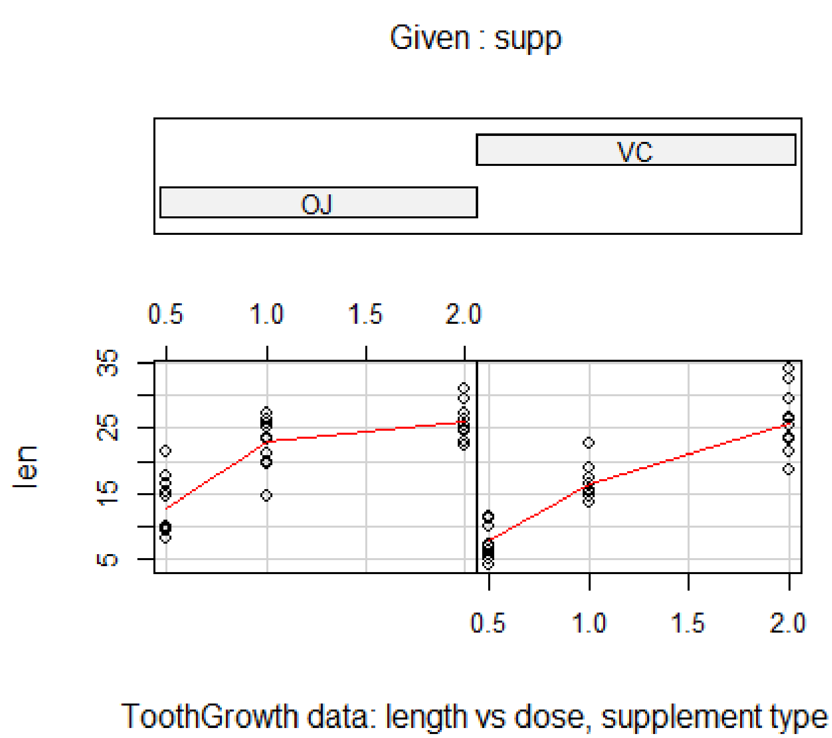
Load and see the data

```
library(datasets)
data = ToothGrowth
str(data)

## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

Basic Exploratory Data Analysis of the data

```
require(graphics)
plot1 = coplot(len ~ dose | supp, data = ToothGrowth, panel = panel.smooth,
       xlab = "ToothGrowth data: length vs dose, supplement type")
```



2. Basic Assumptions

1. Poplulations are independent, that the variances between populations are different, a random population was used.
2. The population was comprised of similar guinea pigs, measurement error was accounted for.

3. Null hypothesis & Confidence Interval

Supplement as a factor

Assumption: Null Hypothesis says that there is no corellation between toothgrowth and supplement and therefore the difference in mean of toothgrowth between two supplements is zero.

Lets test the null hypothesis

```
t.test(len ~ supp, paired = F, var.equal = F, data = ToothGrowth)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by supp
## t = 1.9153, df = 55.309, p-value = 0.06063
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1710156  7.5710156
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

Conclusion: A confidence interval of [-0.171, 7.571] does not allow us to reject the null hypothesis. Therefore our hypothesis that there is no correlation between delivery method and tooth length is strong.

Dosage as a factor

Assumption: Null Hypothesis says that there is no corellation between toothgrowth and dosage and therefore the difference in mean of toothgrowth on different dosages is zero.

To test dosage as a foactor, we test the toothgrowth data between different dosages.

```
dose1 <- subset(ToothGrowth, dose %in% c(0.5, 1.0))
dose2 <- subset(ToothGrowth, dose %in% c(0.5, 2.0))
dose3 <- subset(ToothGrowth, dose %in% c(1.0, 2.0))
```

First we test between dosage 0.5 & 1

```
t.test(len ~ dose, paired = F, var.equal = F, data = dose1)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -6.4766, df = 37.986, p-value = 1.268e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983781  -6.276219
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

Between dosage 0.5 & 2

```
t.test(len ~ dose, paired = F, var.equal = F, data = dose2)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -11.799, df = 36.883, p-value = 4.398e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -18.15617 -12.83383
## sample estimates:
## mean in group 0.5   mean in group 2
##            10.605            26.100
```

Between dosage 1 & 2

```
t.test(len ~ dose, paired = F, var.equal = F, data = dose3)
```

```
##
##  Welch Two Sample t-test
##
## data:  len by dose
## t = -4.9005, df = 37.101, p-value = 1.906e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.996481 -3.733519
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

Conclusions: The confidence intervals ([-11.98, -6.276] for doses 0.5 and 1.0, [-18.16, -12.83] for doses 0.5 and 2.0, and [-8.996, -3.734] for doses 1.0 and 2.0) allow for the rejection of the null hypothesis.

Hence we reject the null hypothesis that there is no relation between diferrent dosages.

Infact there is a strong corellation netween dosage and toothgrowth.