Chap1

Anjali Krishnan and Richard Troise

First, set working directory. 'data' is a table with two columns and same number of rows, and should be numeric. Columns have headers indicating the names of the variables. User will also input desired variable names in double quotes

```
data <- read.csv("chap1.csv", header = FALSE, skip = 1)
colnames(data) = c("Length", "Width")
```

View the data.

```
knitr::kable(xtable(data))
```

Length	Width
3	8
6	4
2	10
6	1
2	11
9	1
6	4
5	3
9	1
4	6
7	2
11	1
5	9
4	3
3	4
9	1
10	3
5	3
4	3
10	2

Calculate the mean for all columns

```
summarise_all(data,mean)
```

```
##   Length Width
## 1      6     4
```

Calculate the standard deviation for all columns

```
summarise_all(data,sd)
```

```
##     Length   Width
## 1 2.809757 3.14559
```

Replace 'var1' and 'var2' to an appropriate first and second column name using CTRL+F. **Only check off 'Match case' to avoid overwriting additional code.**

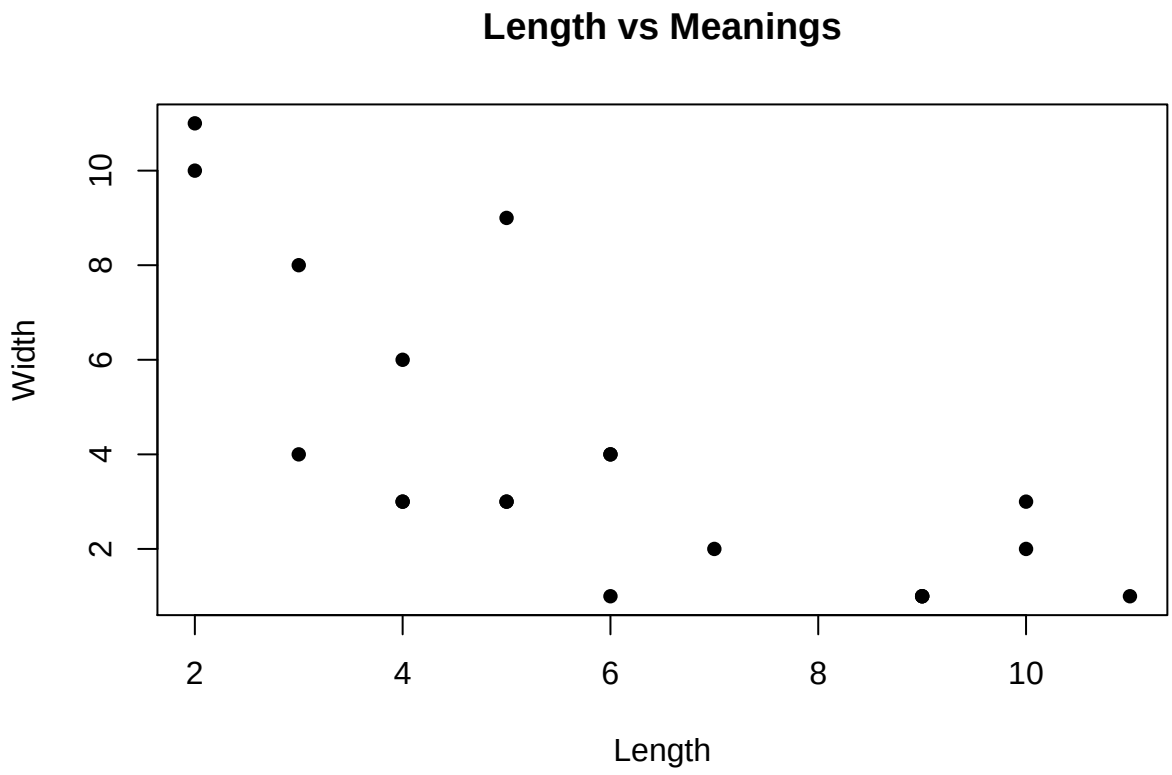
```
column = colnames(data)
colnames(data) <- c("V1", "V2")
var1 = data$V1
var2 = data$V2
```

We now perform a correlation and a test on the data which gives confidence intervals, regression analysis on the data, an ANOVA on the data. Replace 'var1' and 'var2' to an appropriate first and second column names using CTRL+F

```
cor1=cor.test(var1,var2, method = c("pearson"))
reg1=lm(var1~var2)
aov1=aov(var1~var2)
```

We now print the data and all the results. Adjust title before plotting in double quotes

```
plot(data, main = "Length vs Meanings",
     xlab = column[1], ylab = column[2], pch = 16)
```



Extracting the correlation test data. Nothing needs to be adjusted.

```
cor_table = data.frame(df = cor1$parameter,
                       t_value = cor1$statistic,
                       p_value = cor1$p.value,
                       r = cor1$estimate,
                       LowC.I. = cor1$conf.int[1],
                       UpperC.I = cor1$conf.int[2])
```

```
rownames(cor_table) = ""
```

Correlation Test

```
knitr::kable(xtable(cor_table))
```

df	t_value	p_value	r	LowC.I.	UpperC.I
18	-4.564434	0.0002403	-0.7324543	-0.8873588	-0.4289759

Regression Analysis

```
summary(reg1)
```

```
##
## Call:
## lm(formula = var1 ~ var2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0000 -1.6543 -0.0372  1.0372  3.3457
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   8.6170     0.7224  11.928 5.56e-10 ***
## var2         -0.6543     0.1433  -4.564  0.00024 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.965 on 18 degrees of freedom
## Multiple R-squared:  0.5365, Adjusted R-squared:  0.5107
## F-statistic: 20.83 on 1 and 18 DF,  p-value: 0.0002403
```

ANOVA table

```
summary(aov1)
```

```
##             Df Sum Sq Mean Sq F value  Pr(>F)
## var2         1  80.47   80.47   20.83 0.00024 ***
## Residuals   18  69.53    3.86
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```