Effect of nursery habitat availability to coral reef species

Dirkje Verhoeven, Dolf Andringa, Annelies Andringa: Marine Conservation Philippines

April 9, 2016

Read the spreadsheet (open document spreadsheet) and make a copy that we can use for processing the data.

df.raw <- read\_ods(path='../Data\_Surveys\_dates\_corrected.ods',sheet=2,col\_types = c(text,numeric,numeric,numeric,numeric,numeric,numeric,numeric,numeric,date))  
df.processed <- df.raw

Check the summary of the data. All variables are text variables which they shouldn't be.

summary(df.raw)

## SPECIES 0-20 25-45   
## Length:1896 Length:1896 Length:1896   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
## 50-70 75-95 LOCATION   
## Length:1896 Length:1896 Length:1896   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
## DEPTH OBSERVER REPLICA   
## Length:1896 Length:1896 Length:1896   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
## DATE   
## Length:1896   
## Class :character   
## Mode :character

Next we'll correct the values next. First convert the DATE column to a Date value. Then loop over the variables with integers (whole numbers) and correct those to integers. Lastly, the variables that are actually grouping variables (factors) should be delared as such. So loop over those variables and correct them too.

#Set the date value as a Date.  
df.processed$DATE=as.Date(df.processed$DATE)  
  
for (col in c('0-20','25-45','50-70','75-95','LOCATION','DEPTH','OBSERVER')){  
 df.processed[,c(col)] <- as.integer(df.processed[,c(col)])  
}  
for (col in c('LOCATION','DEPTH','OBSERVER')){  
 df.processed[,c(col)] <- as.factor(df.processed[,c(col)])  
}

And get another summary.

summary(df.processed)

## SPECIES 0-20 25-45 50-70   
## Length:1896 Min. : 0.000 Min. : 0.000 Min. : 0.00   
## Class :character 1st Qu.: 0.000 1st Qu.: 0.000 1st Qu.: 0.00   
## Mode :character Median : 0.000 Median : 0.000 Median : 0.00   
## Mean : 1.376 Mean : 1.353 Mean : 1.25   
## 3rd Qu.: 2.000 3rd Qu.: 2.000 3rd Qu.: 1.00   
## Max. :60.000 Max. :57.000 Max. :100.00   
##   
## 75-95 LOCATION DEPTH OBSERVER REPLICA   
## Min. : 0.000 9 :311 1:920 5 :486 Length:1896   
## 1st Qu.: 0.000 8 :272 2:976 3 :417 Class :character   
## Median : 0.000 4 :229 4 :389 Mode :character   
## Mean : 1.208 3 :222 6 :302   
## 3rd Qu.: 1.250 6 :206 8 :102   
## Max. :54.000 5 :193 7 : 76   
## (Other):463 (Other):124   
## DATE   
## Min. :2016-03-16   
## 1st Qu.:2016-03-24   
## Median :2016-03-30   
## Mean :2016-03-29   
## 3rd Qu.:2016-04-02   
## Max. :2016-04-08   
##

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

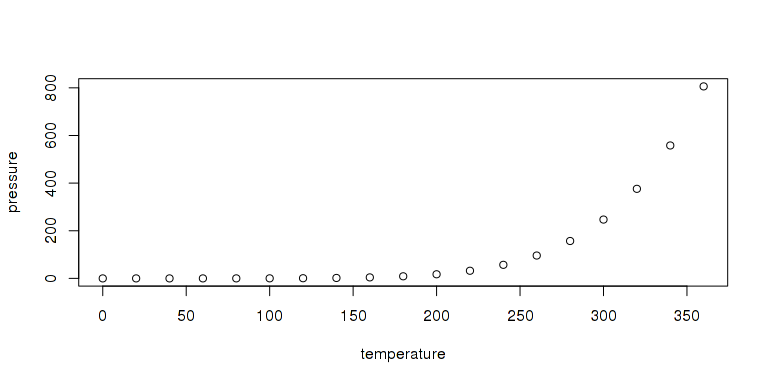
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

summary(cars)

## speed dist   
## Min. : 4.0 Min. : 2.00   
## 1st Qu.:12.0 1st Qu.: 26.00   
## Median :15.0 Median : 36.00   
## Mean :15.4 Mean : 42.98   
## 3rd Qu.:19.0 3rd Qu.: 56.00   
## Max. :25.0 Max. :120.00

## Including Plots

You can also embed plots, for example:



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.