

# summary\_table1

dk

4/13/2017

## 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:

```
# This code calculates a mean, standard deviation, and N for each measured variable by region & lake type

# read the summary file
# need to figure out how to read file directly from GitHub
nes_all <- read.csv("../nes_data.csv", header=TRUE, na.strings="NA", stringsAsFactors = FALSE)

# add a column to specify region
nes_all[nes_all$state %in% c("VERMONT", "CONNECTICUT", "RHODE ISLAND", "NEW HAMPSHIRE", "NEW YORK", "MA",
nes_all[nes_all$state %in% c("ALABAMA", "DELAWARE", "FLORIDA", "GEORGIA", "ILLINOIS", "INDIANA", "KENTU
nes_all[nes_all$state %in% c("ARKANSAS", "IOWA", "KANSAS", "LOUISIANA", "MISSOURI", "NEBRASKA", "NORTH I
nes_all[nes_all$state %in% c("ARIZONA", "CALIFORNIA", "COLORADO", "IDAHO", "MONTANA", "NEVADA", "NEW ME

# make REGION a factor
nes_all$region <- as.factor(nes_all$region)
# reorder region levels so final table reads from west to east
nes_all$region <- factor(nes_all$region, levels = c("WESTERN", "CENTRAL", "NORTHEASTERN", "EASTERN"))

# reshape data to [n]es_[a]ll [l]ong format
library(reshape2)

# nes_all[,9:34] <- apply(nes_all[,9:34], 2, as.numeric)

nal <- melt(nes_all,
            id.vars = c("state", "name", "region", "lake_type"),
            measure.vars = c((names(nes_all)[c(9:34)])))
nal$value <- as.numeric(nal$value)

## Warning: NAs introduced by coercion

# mean(
#   as.numeric(nal[nal$variable == "p_surface_area_loading" &
#             nal$region == "EASTERN" &
#             nal$lake_type == "NATURAL", "value"])
#   , na.rm = TRUE)

# calculate mean, standard dev, and N for each variable by region and lake type
library(plyr)
summary_nes = ddply(nal, .(variable, region, lake_type), summarize,
                    mean = mean(value, na.rm=TRUE),
                    sd = sd(value, na.rm=TRUE),
                    N = length(value[!is.na(value)]))
```

```

# any(!is.na(summary_nes$mean))

# now how to report numbers? let's convert them to character strings & round means & sd's to 2 decimal
summary_nes$mean <- as.character(round(summary_nes$mean, digits=2))
summary_nes$sd <- as.character(round(summary_nes$sd, digits=2))
summary_nes$N <- as.character(summary_nes$N)

# add parentheses around N
summary_nes$N <- paste("(", summary_nes$N, ")", sep="")

# now create a column with the 3 values & the +- symbol (using paste?)
summary_nes$stat <- paste(summary_nes$mean, "\u00b1", summary_nes$sd, summary_nes$N, sep=" ")
# Didn't need to do this so far, but might need to execute the following if +- symbol doesn't work
Encoding(summary_nes$stat) <- "UTF-8"

# now name classify variables into morphometry, physicochemical, & loading variables
summary_nes[summary_nes$variable %in% c(levels(summary_nes$variable)[1:5]), "variable_type"] <- "MORPHOLOGY"
summary_nes[summary_nes$variable %in% c(levels(summary_nes$variable)[6:12]), "variable_type"] <- "PHYSIOCHEMICAL"
summary_nes[summary_nes$variable %in% c(levels(summary_nes$variable)[13:28]), "variable_type"] <- "LOADING"

# let's make three different dataframes for each of the variable types (each will be a separate table)
morph <- summary_nes[summary_nes$variable_type=="MORPHOMETRY", ]
physchem <- summary_nes[summary_nes$variable_type=="PHYSIOCHEMICAL", ]
load <- summary_nes[summary_nes$variable_type=="LOADING", ]

# let's reshape these
morph_wide <- dcast(morph, variable + lake_type ~ region, value.var="stat")
physchem_wide <- dcast(physchem, variable + lake_type ~ region, value.var="stat")
load_wide <- dcast(load, variable + lake_type ~ region, value.var="stat")

# let's make the tables
# install.packages("devtools")
# devtools::install_github("rstudio/rmarkdown")
# For dev version
# devtools::install_github("haozhu233/kableExtra")

# load libraries
library(knitr)

## Warning: package 'knitr' was built under R version 3.3.2

library(kableExtra)
library(magrittr)

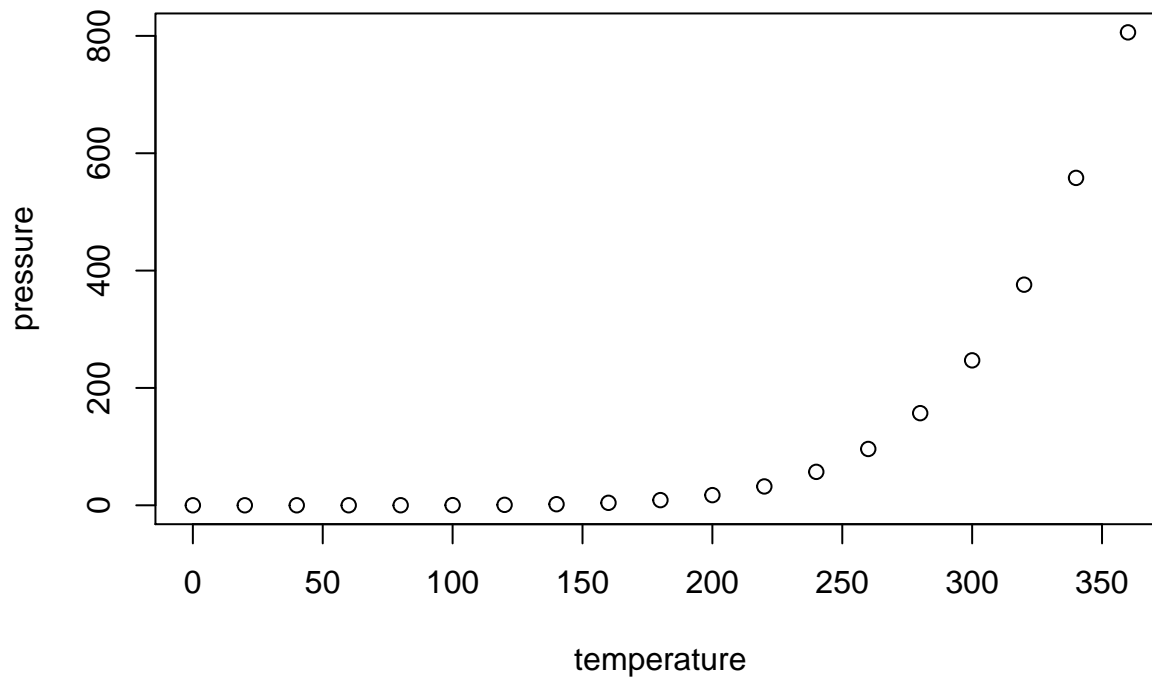
kable(morph_wide, format = "latex", booktabs = T) %>%
  kable_styling()

```

## Including Plots

You can also embed plots, for example:

variable	lake_type	WESTERN	CENTRAL	NORTHEASTERN
drainage_area	IMPOUNDMENT	32620.28 $\pm$ 88819.8 (91)	22251.46 $\pm$ 77036.81 (130)	7659.8 $\pm$ 25000.22 (48)
drainage_area	NATURAL	1551.67 $\pm$ 3835.43 (31)	378.75 $\pm$ 259.86 (8)	1409.83 $\pm$ 3179.97 (123)
surface_area	IMPOUNDMENT	42.74 $\pm$ 92.81 (104)	67.97 $\pm$ 161.1 (135)	10.66 $\pm$ 21.43 (54)
surface_area	NATURAL	48.52 $\pm$ 114.55 (48)	10.71 $\pm$ 13.57 (42)	33.38 $\pm$ 114.65 (146)
mean_depth	IMPOUNDMENT	14.9 $\pm$ 12.04 (104)	6.81 $\pm$ 4.69 (134)	4.39 $\pm$ 3.91 (48)
mean_depth	NATURAL	20.9 $\pm$ 45.84 (45)	3.15 $\pm$ 1.95 (40)	7.99 $\pm$ 10.6 (126)
total_inflow	IMPOUNDMENT	62.13 $\pm$ 124.72 (92)	31.49 $\pm$ 70.1 (130)	45.59 $\pm$ 106.44 (48)
total_inflow	NATURAL	23.28 $\pm$ 65.09 (32)	37.18 $\pm$ 101.04 (8)	14.25 $\pm$ 35.62 (122)
retention_time	IMPOUNDMENT	105 $\pm$ 115.44 (92)	78.15 $\pm$ 120.21 (131)	20.95 $\pm$ 41.45 (46)
retention_time	NATURAL	83.86 $\pm$ 102.8 (32)	39.46 $\pm$ 87.43 (9)	64.75 $\pm$ 93.34 (112)



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