

Introduction to R

Session 4 – Data manipulation

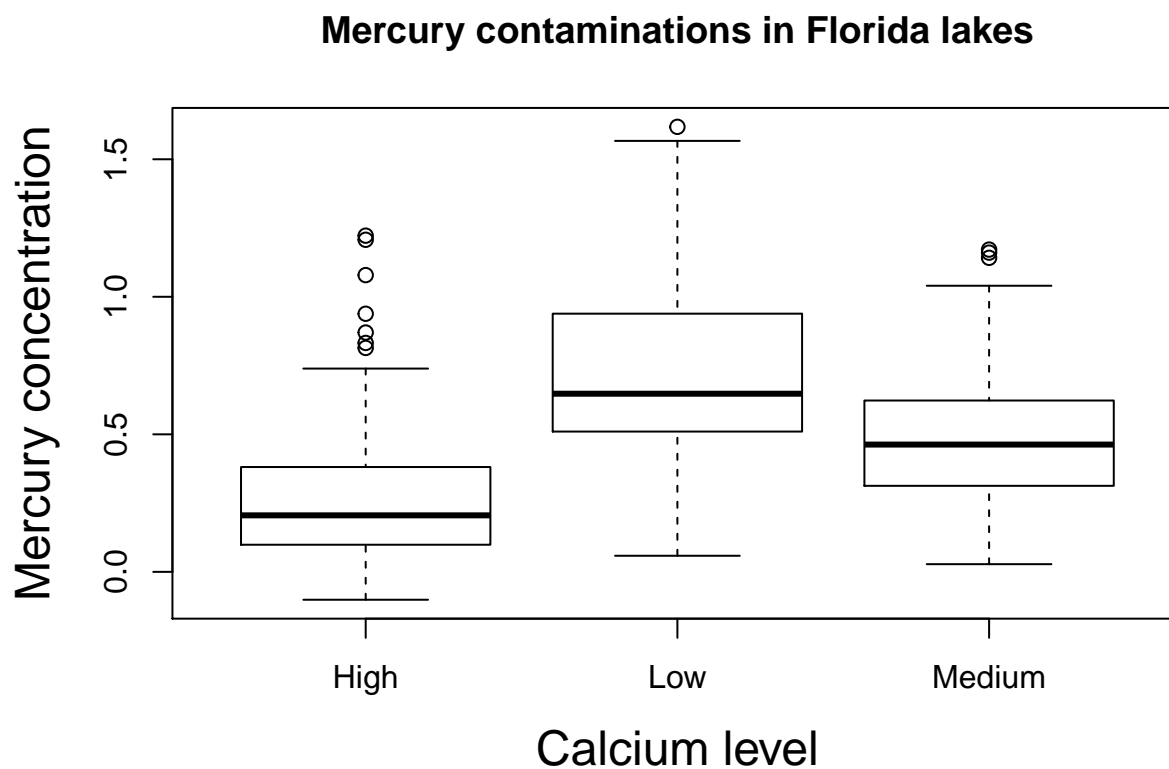
Statistical Consulting Centre

19 July, 2017

1. Boxplot

1. Draw boxplots showing the distribution of `Mercury` for each of the levels of `Calcium`. The boxplot should look exactly like

```
with(joined.long.df, plot(as.factor(Calcium), Mercury,
  xlab = "Calcium level", cex.lab = 1.5,
  ylab = "Mercury concentration",
  main = "Mercury contaminations in Florida lakes"))
```

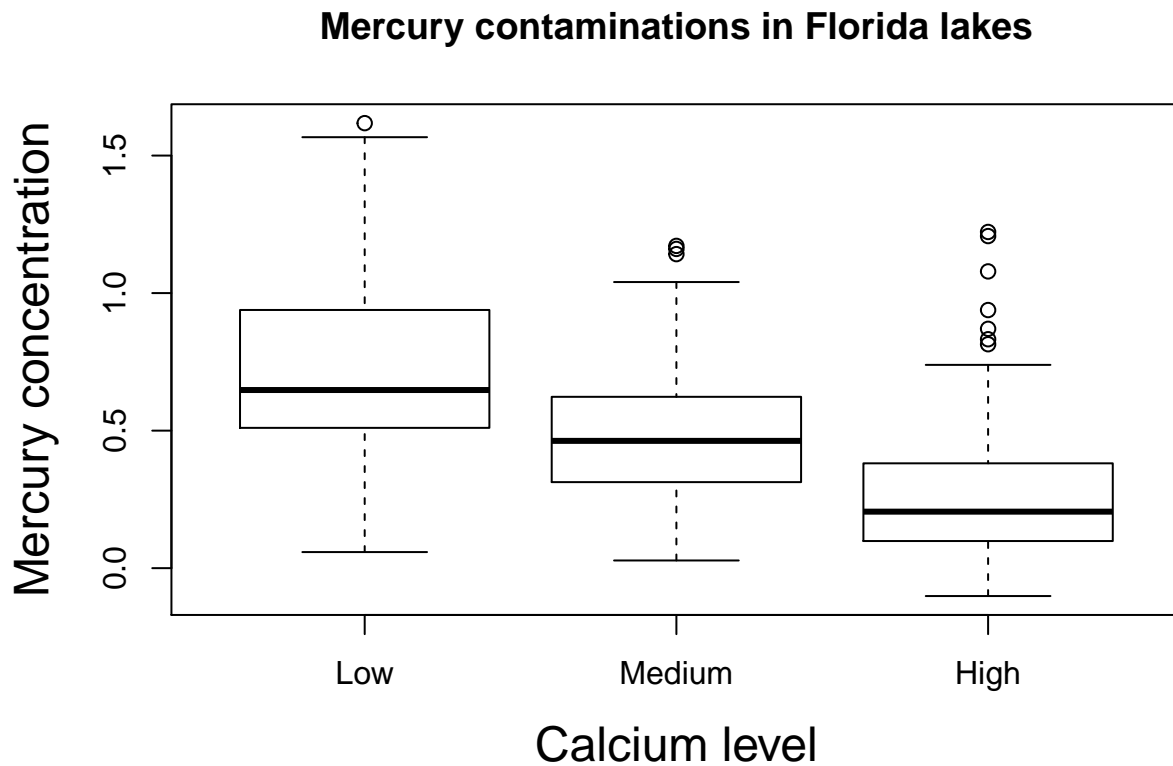


The boxplot should have:

- labels on the x-axis,
- an informative label on the y-axis (not `Mercury`),
- a title, and
- labels that are large enough to read without squinting.

- Now try to produce the boxplot so that Calcium is plotted in a sensible order with respect to Calcium level like

```
joined.long.df$Calcium <- factor(joined.long.df$Calcium, level = c("Low", "Medium", "High"))
with(joined.long.df, plot(Calcium, Mercury,
  xlab = "Calcium level", cex.lab = 1.5,
  ylab = "Mercury concentration",
  main = "Mercury contaminations in Florida lakes"))
```



Hint: use the following command to convert Calcium variable for plotting.

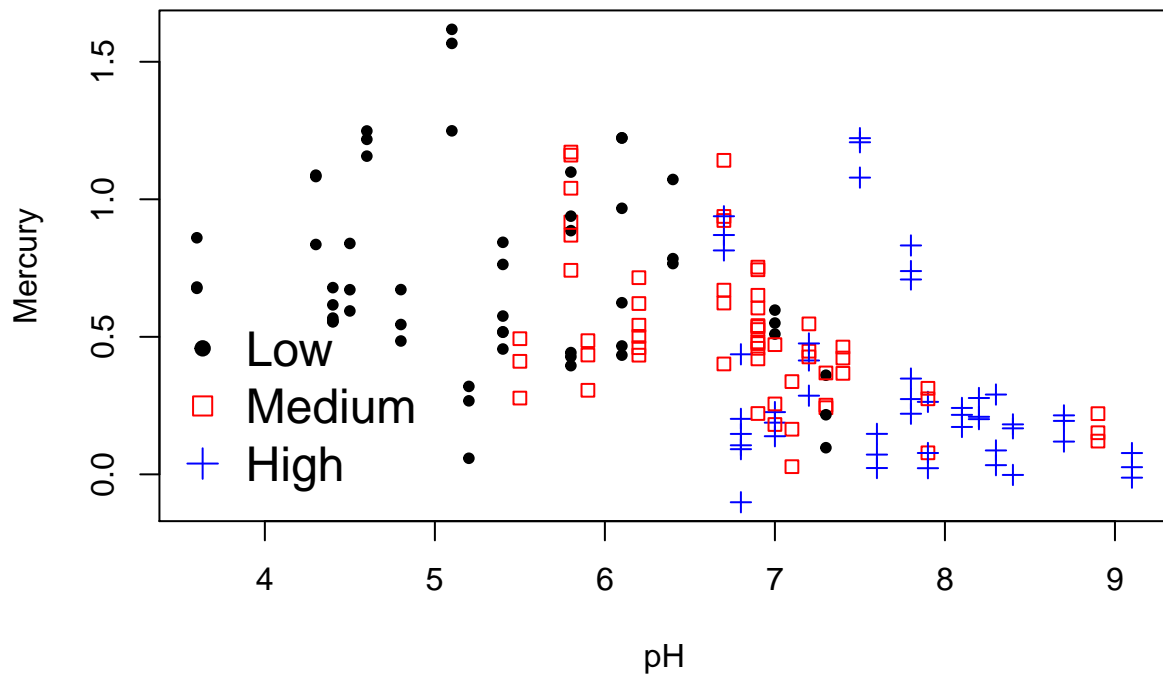
```
joined.long.df$Calcium <- factor(joined.long.df$Calcium, level = c("Low", "Medium", "High"))
```

2. Scatterplot

Draw a scatter plot which shows the relationship between pH and Mercury for each of the Calcium levels. The scatter plot should look exactly the same as

```
with(joined.long.df, plot(pH, Mercury,
  col = c(1,2,4)[Calcium],
  pch = c(20, 22,3)[Calcium]))

legend("bottomleft", pch = c(20, 22, 3),
  col = c(1, 2, 4),
  bty = "n", cex = 1.5,
  legend = c("Low", "Medium", "High"))
```



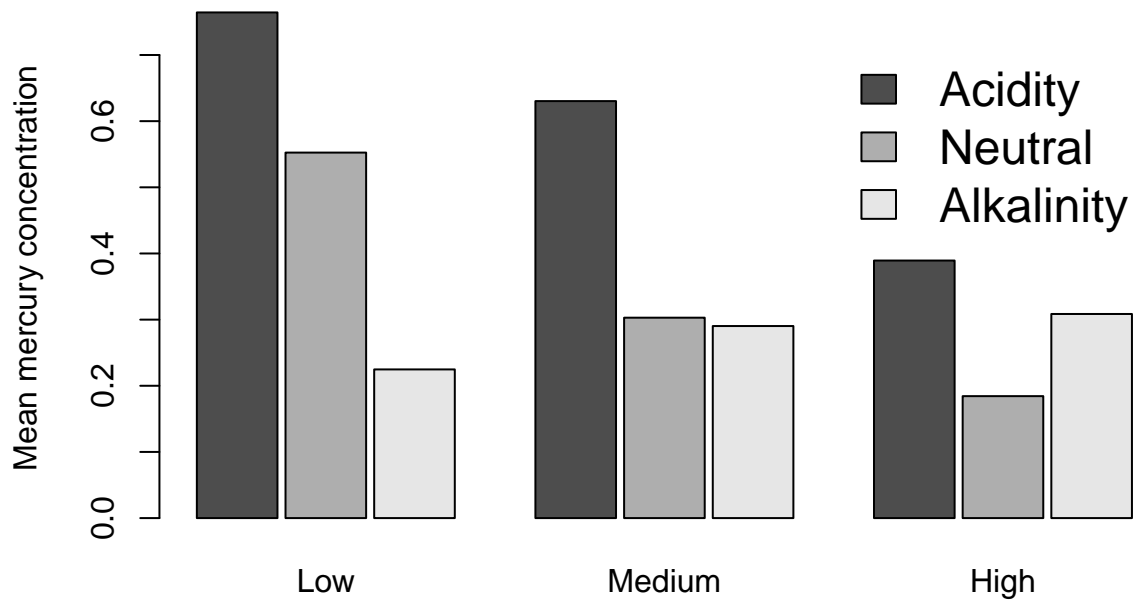
Hint:

- Low Calcium level: using `pch = 20` and `col = 1`.
- Medium Calcium level: using `pch = 22` and `col = 2`.
- High Calcium level: using `pch = 3` and `col = 4`.

3. Barplot

1. Draw a juxtaposed barplot that shows the average Mercury concentration for the six combinations of Calcium and pHtype. The graph should look exactly the same as

```
trt.mean <- with(joined.long.df, tapply(Mercury,
                                         list(pHtype, Calcium), mean))
barplot(trt.mean, beside = T, space = c(0.1, 1),
        legend.text= c("Acidity", "Neutral", "Alkalinity"),
        ylab = "Mean mercury concentration",
        args.legend = list(bty = "n", cex = 1.5))
```



Draw the barplot in the following steps:

- Calculate the mean `mercury` level for each combination of `pHtype` and `plot.cal`, using the `tapply()` function.
 - Save the calculated means in a new object called `trt.mean`.
 - Produce the barplot using `beside = TRUE`, putting a small space between the bars. (Reminder: the argument that does this is `space`).
 - Add a legend.
2. Mess up the plot by playing with `horiz`, `angle` `density` and any arguments that you want. Note: keep this to yourself.