

# Homework 1

## Introduction to Parallel Computing

J.D. Peiffer

2/1/2021

In this assignment we will develop functions for visualizing missing data in timeseries data. We will begin with an example dataset and then scale up our function to 50 years worth of data. Last, we will accelerate our computations with parallel programming packages. This assignment can be performed on any computer, although using the Lewis computing cluster may provide greater accelerations in the final, parallel computing, section.

### Section 1: An Example

Hydrological datasets, collected over decades, may contain missing recordings. Depending on the recording issue, this could be an hour, a day, or a year. Let us define a simple timeseries with data from 4 different weather stations over 5 different days.

```
test=matrix(nrow=5,ncol=4)
test[,1]=1:5
test[2:5,2]=2:5
test[3:5,3]=3:5
test[4:5,4]=4:5
test=data.frame(test)
colnames(test)=c("Station1","Station2","Station3","Station4")
rownames(test)=c("Day1","Day2","Day3","Day4","Day5")
View(test)
```

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>.

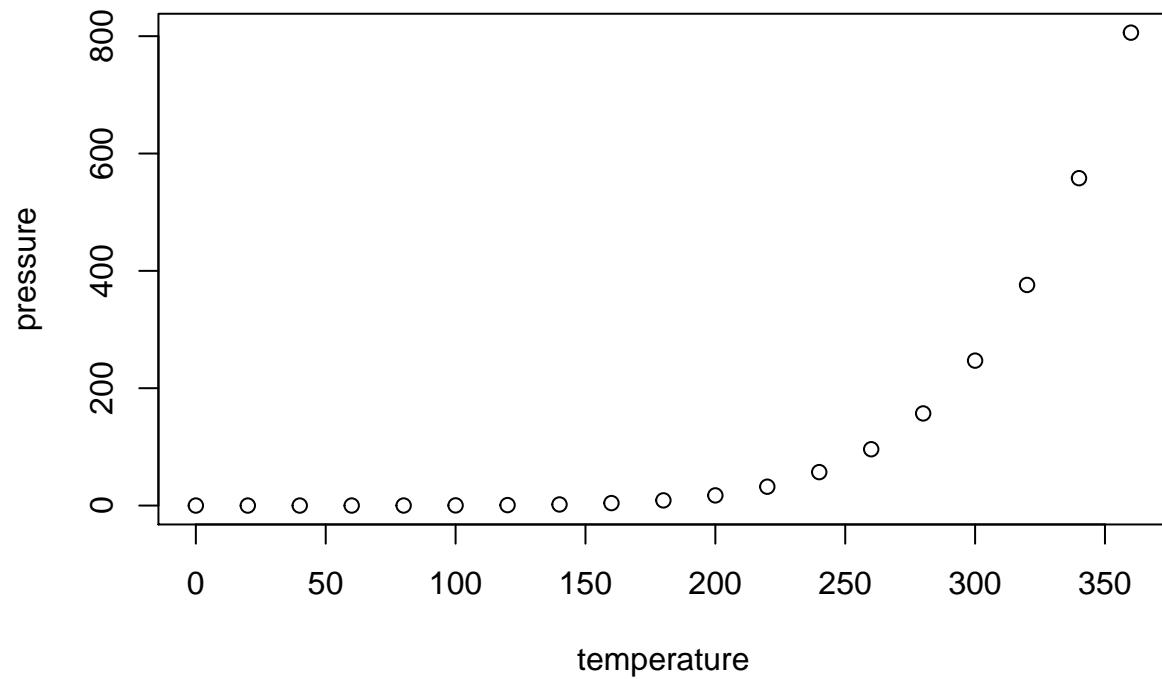
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
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.