Homework 1

The purpose of this document is to get you some basic practice using *R Markdown*, as well as to get some practice reading in data and working with vectors and matrices. As mentioned in the syllabus, I will be asking for all homeworks to be completed with *R Markdown* so it's worth getting used to now.

To receive full credit, please create and render an *R Markdown* document in html or pdf format that has the following elements.

- Title, author and date
  - Use a YAML header in your document to do this
- An R code chunk displaying how to load data into R and store it into an object, along with text explaining the code.
- A generic ordered or unordered list with at least one level of nesting (could even be a shopping list)
- Some R output of some sort
- Calculate the mean of a variable, via `mean()` and display it in text. If there are missing data in the vector that you are trying to calculate the mean from, you will also need to include the additional argument `na.rm = TRUE`. Make some manipulation to the data (e.g., remove a few random cases) and report the mean in text again, using code. For example:

```
data(ChickWeight)
head(ChickWeight)
```

```
##   weight Time Chick Diet
## 1     42    0     1    1
## 2     51    2     1    1
## 3     59    4     1    1
## 4     64    6     1    1
## 5     76    8     1    1
## 6     93   10     1    1
```

```
# globally round all output to two digits
options(digits = 2)
```

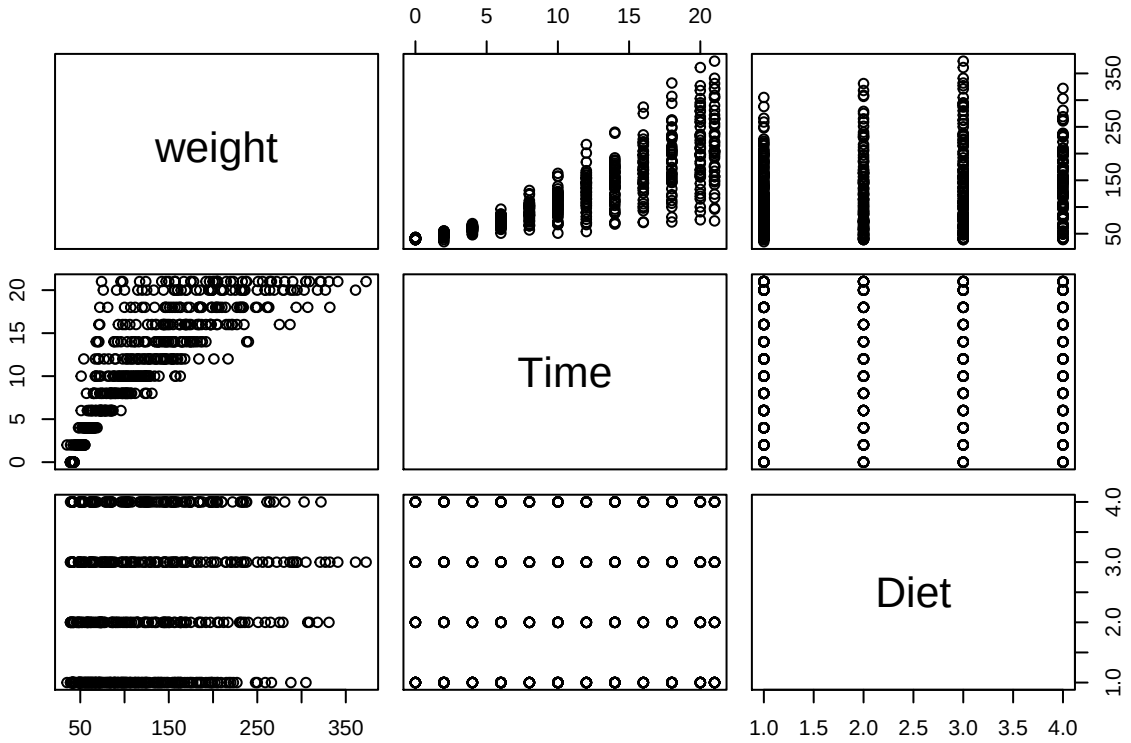
The average chick weight was 121.82 grams. At time point 0 the average weight was 41.06 grams.

Plot

Use the dataset you loaded, or pick a new dataset from base R to produce a `pairs()` plot (scatterplot matrix). To see the datasets available in R, type `data()` at the command prompt. You should see a summary of many different datasets. To get information on any one of these datasets, type `?` and the name of the dataset. For example `?ChickWeight` will tell you some information about the dataset above. If you choose to use a base R dataset, please use one other than the ChickWeight dataset.

To produce a pairs plot, you just need to feed it a data frame with the variables you want plotted. I encourage you to play around with this function and check out the documentation. For example, to get a scatterplot matrix of all but the third variable in the ChickWeight dataset, use

```r
pairs(ChickWeight[ ,-3])
```



Matrices and Vectors

Create the following matrix (which we created during the first class, if you saved your syntax)

$$\mathbf{m} = \begin{bmatrix} 18 & 32 & 11 & 41 & 73 \\ 61 & 47 & 22 & 87 & 63 \\ 44 & 52 & 23 & 42 & 77 \\ 23 & 17 & 5 & 72 & 83 \end{bmatrix}$$

1) Use three different methods to subset the matrix and obtain the following vector

$$\mathbf{v} = \begin{bmatrix} 52 & 42 & 77 \end{bmatrix}$$

2) Subset **m** again to get $\mathbf{v2} = \begin{bmatrix} 87 & 42 \end{bmatrix}$. Then create a new matrix using the following steps:

- Drop the final column of matrix **m**
- Add **v2** to the second and fourth rows (this will utilize the property of recycling).

Render the document

Once you have fulfilled all the requirements of the homework, you need to render it to html (recommended) or pdf (requires additional software installations) using either the `knit` function from the `knitr` package, or

`render` from the `rmarkdown` package. The `knit2html` button in Rstudio will essentially just run the code for you.

If you're using RStudio, you shouldn't have to worry about much to get it rendered. If not, make sure you have the correct packages installed. The code below should help.

```
#### Via the knitr package
# install package
install.packages("knitr") # note, this only needs to be done once
library(kntir)
knit2html("yourFile.Rmd")

#### Via the RMarkdown package
install.packages("rmarkdown")
library(rmarkdown)
render("yourFile.Rmd", "html_document")
```