Rmarkdown

Environmental Informatics

Naomi

2017-04-24

Reproducible research

Rmarkdown

A really useful way to organize, store and present your workflow: the steps that you take in a data analysis project - from input, to analysis, to presentation

Gandrud (2013) Reproducible Research with R and Rstudio

Functionality (can include)

- text
- headings
- fonts
- equations
- images
- R code
- other code
- links to website
- output to different formats

Create a R markdown file

In RStudio, to menu File > New File > Rmarkdown...

Knit HMTL button to render

Syntax

Write a simple sentance in your R markdown file
>Now add some flare to it

Try `**making text bold **` which should look like this

making text bold

or `_ ** a combinination ** of different fonts _`

that should look like this

a combinination *of different fonts*

use a *backtick* ‘ to have something not render

More on Rmarkdown syntax

Try rendering with *Knit*

Note if things don't look right - a common problem is not putting a space between line so RMarkdown doesnt' know you are on to a new thing - so try adding a return

Images

`<div align="center"> <img src="../img/Rmarkdown.jpg" height=400> </div>` Pay attention to the path name -R is starting with in your working directory

Links

you can use '[]' to name and link

Latex

`[Latex](https://en.wikibooks.org/wiki/LaTeX)`

Equations in Rmarkdown - Chunk

Surrount your math in **$** symbols

Uses Latex markup language for equations as input.

Some nice examples of commonly used greek symbols and functions more latex

$$\frac{1}{n} \sum_{i=i}^{n} x_{i}$$

this came from `$$ \frac{1}{n} \sum_{i=i}^{n} x_{i} $$`

Including R code

You can also include R code: Use backticks, r, and then give your code *chunk* a name

```
cc = function(gmax=0.5, LAI=3, b=0.5, optT=24, Tair) {

  airm = exp(-b*(Tair-optT))
  cc = gmax * LAI * airm
  cc
}
cc(Tair=24)
```

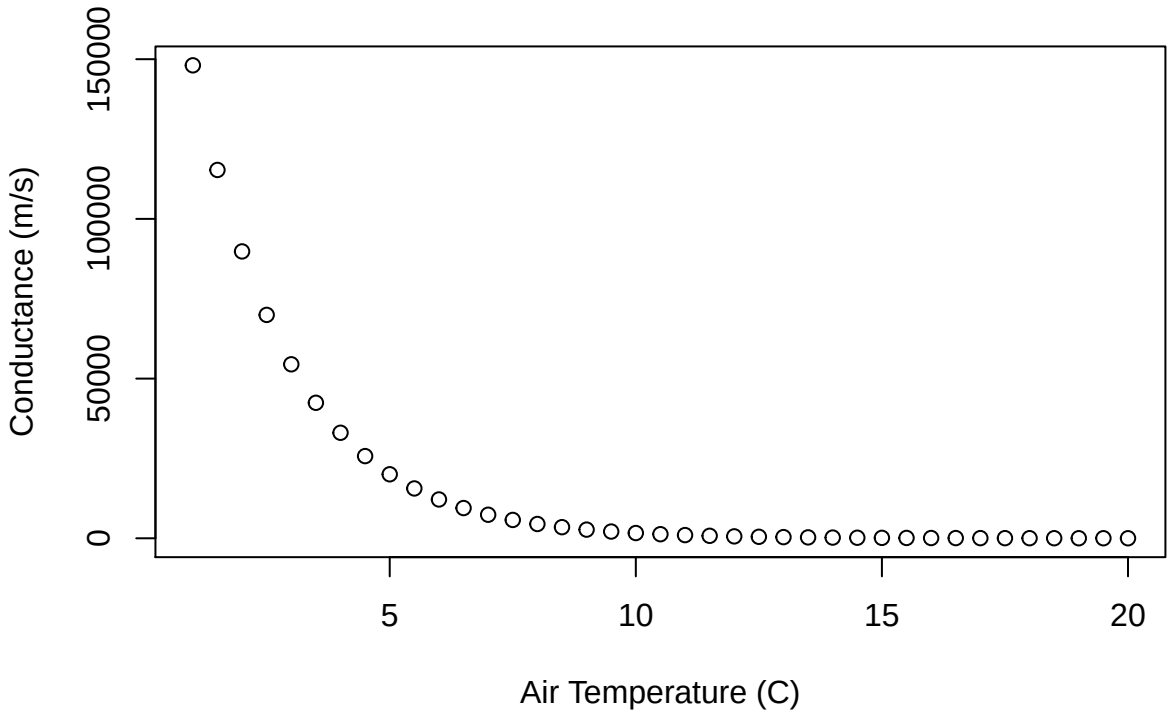
```
## [1] 1.5
```

```
optT = seq(from=15,to=24)
T = seq(from=1, to=20, by=0.5)

cc(gmax=0.5, LAI=3, optT=24, Tair=T)
```

```
##  [1] 148073.65652 115319.87965  89811.21257  69945.04268  54473.25401
##  [6]  42423.81288  33039.69869  25731.34321  20039.59024  15606.84857
## [11]  12154.62589   9466.03216   7372.15326   5741.43873   4471.43698
## [16]   3482.35862   2712.06362   2112.15727   1644.94974   1281.08814
## [21]    997.71245    777.01924    605.14319    471.28599    367.03790
## [26]    285.84940    222.61974    173.37643    135.02570    105.15812
## [31]     81.89723     63.78162     49.67318     38.68551     30.12831
## [36]     23.46395     18.27374     14.23160     11.08358
```

```
res=data.frame(Tair=T)
res$cc = cc(gmax=0.5, LAI=3, Tair=res$Tair)
plot(res$Tair, res$cc, ylab="Conductance (m/s)",xlab="Air Temperature (C)")
```



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot Try playing with echo=TRUE and eval=FALSE

inline R code

Simple inline R code is also possible, use backticks and starting code with r Lets say you want to evaluate something in line the circumference of a circle is π times R**2

- for a radius of 4 we get 50.2654825

Output formats

"easy" HTML

- ***output*** various formats:
  - **html**: share on public website

  - **pdf**: polished publication
  - **docx**: share with co-authors to track changes
  - **slidy_presentation** slides

- specify this at the top of the R markdown file
- if you add slides (*slidy_presentation*), you can use *** or a header to denote a new slide
- each output format has different options that can also be specified at the top of the R Markdown file in the **output:** section

Find these here

There are many other ways to tailor your R-Markdown - explore!

Some Helpful R functions

Ok now lets do some data analysis and communicate results in R Lets go back to our radiation example

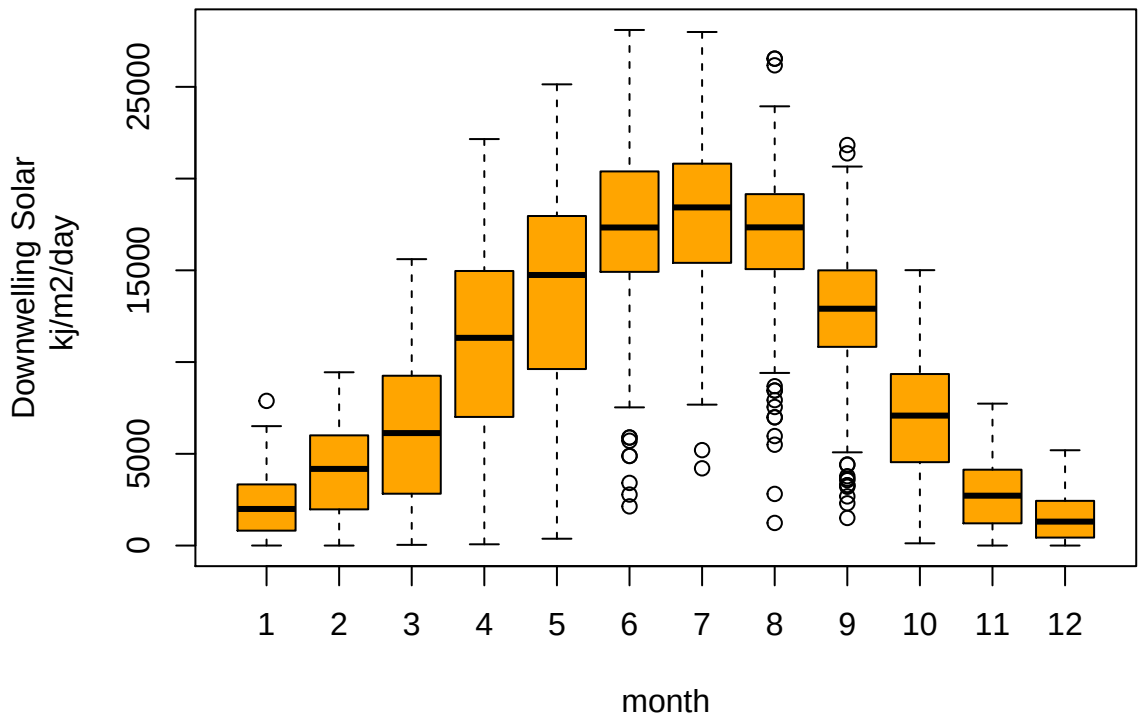
Lets imbedded the code in our R document

```
sunlight = read.table("sun.txt", header=T)
```

```
## Warning in if (!header) rlabp <- FALSE: the condition has length > 1 and
## only the first element will be used
```

```
## Warning in if (header) {: the condition has length > 1 and only the first
## element will be used
```

```
par(mar=c(5,6,3,2))
boxplot(sunlight$Kdown_direct~sunlight$month,
  ylab="Downwelling Solar\n kj/m2/day",
  xlab="month", col="orange")
```



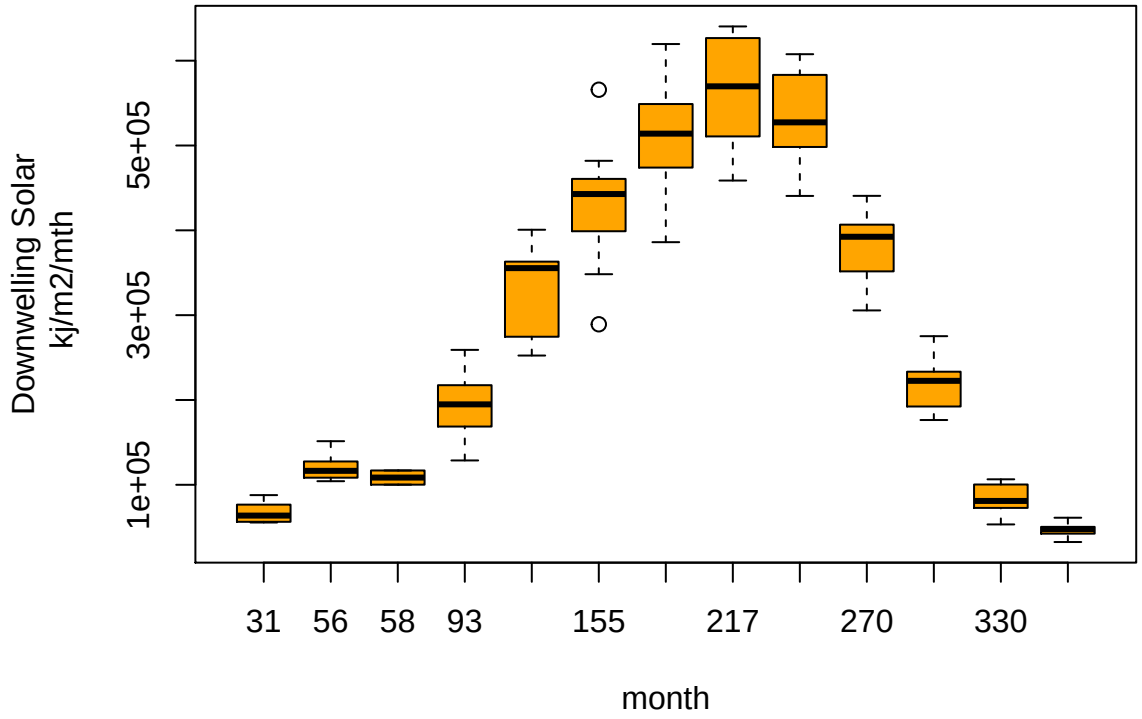
And lets demo a few useful R functions

1. *aggregate*
2. *which.max*

```
# total radiation monthly
sun.mth = aggregate(sunlight, by=list(sunlight$month, sunlight$year), sum)
sun.mth[1:3,]
```

```
##   Group.1 Group.2 day month  year Kdown_direct
## 1      10    1944 496   310 60264    275223.20
## 2      11    1944 465   330 58320     53304.39
## 3      12    1944 496   372 60264     46677.64
```

```
par(mar=c(5,6,3,2))
boxplot(sun.mth$Kdown_direct~sun.mth$month,
  ylab="Downwelling Solar\n kj/m2/mth",
  xlab="month", col="orange")
```




```
# what is the maximum radiation received in any month
result=which.max(sun.mth$Kdown_direct)
sun.mth[result,]
```

```
##    Group.1 Group.2 day month  year Kdown_direct
## 22       7    1946 496   217 60326     640355.4
```

```
# prettier
sun.mth[result,c("Group.1","Group.2","Kdown_direct")]
```

```
##    Group.1 Group.2 Kdown_direct
## 22       7    1946     640355.4
```

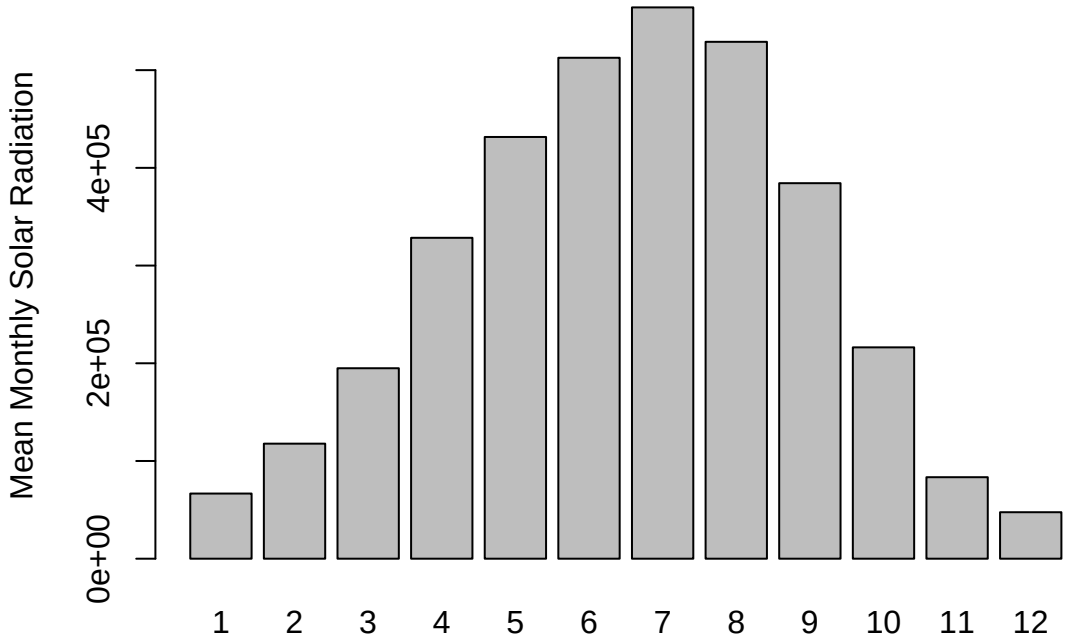
```
# clean up
sun.mth$month=as.integer(sun.mth$Group.1)
sun.mth$year= as.integer(sun.mth$Group.2)
sun.mth[result,c("month","year","Kdown_direct")]
```

```
##    month year Kdown_direct
## 22     7 1946     640355.4
```

```
# and then maybe find average values
sun.avg = aggregate(sun.mth, by=list(sun.mth$month), mean)
sun.avg[1:3,]
```

```
##   Group.1 Group.1 Group.2   day month   year Kdown_direct
## 1       1       1  1949.5 496.0     1 1949.5     66683.06
## 2       2       2  1949.5 411.8     2 1949.5    117743.43
## 3       3       3  1949.5 496.0     3 1949.5    194917.91
```

```
barplot(sun.avg$Kdown_direct, names=sun.avg$Group.1, ylab="Mean Monthly Solar Radiation")
```



Assigment

Given a climate data set (clim.txt), do the following tasks and document your work in an R-markdown document and store it in your repository- add the name of the repository to gauchospace

Tasks

- read in the file **read.table("clim.txt", header=TRUE)**
- graph precipitation and average temperature by month, using a box plot
- find wettest and driest years
- add two pictures to illustrate what a wet and dry year might look like for the ecosystem of your choice
- create a new variable that is 1 when its spring, and 2 when its summer, 3 with its fall and 4 when its winter (eg. split your year in to 4, 3 month periods)
- find wettest and driest seasons
- graph the relationship between winter precipitation and summer temperature
- add some text to describe the graph and comment on why this might be useful to look at
- make at least one edit to your file and commit changes
- make sure your work includes code, headings
- output your work to a pdf and then to slides

Help

- Rstudio menu Help > Markdown Quick Reference
- R Markdown Cheat Sheet

Further Resources

- rmarkdown.rstudio.com: authoritative reference site