Homework 3 Instructions

FISH 552 - Kristin Privitera-Johnson

10/06/2021

Part 1 - Programming Assignment

Open a new script in RStudio and create an assignment header using comments.

```
# Name: First Last
# Homework 3
```

Complete the tasks below. Please label each question and task with comments.

```
#==Question 1a ####
# Your R code here

#==Question 1b ####
# Your R code here
```

When your script is complete, save it as `LastName_Homework3.R`, then clear your workspace (Workspace/Clear all) and run through your script again to make sure you don't have any `object not found` errors. Then go to the course website to submit your R script.

Question 1

Ram Myers compiled a very substantial stock-recruitment database of hundreds of fish from all over the world. This data was used in many of his and colleagues' notable papers appearing in Ecology, Science and Nature. This database can be found at https://www.ramlegacy.org. For this question we will focus on just one species, Atlantic Mackerel. The data files are in the Data folder on Canvas.

a) Read in the three data sets found on our Canvas for Homework 3. Call each data set `mack.black`, `mack.nafo` and `mack.ices` respectively. The column names are `"Year" "spawners" "recruits" "catch" "fishMortality"` Be sure that these are used instead of `V1, V2, . . . . .`

b) Create a single data set from `mack.black` and `mack.nafo` that contains all of the variables in each data set (`spawners recruits catch fishMortality`), but is restricted to years that the two data sets have in common. One function will do this. Call this data frame `mack.partial`. The column names of `mack.partial` should have suffices corresponding to the specific data set. By that I mean: `"spawners.black" "recruits.black" . . . "spawners.nafo" "recruits.nafo"`. Specifying one option in this function will do this for you.

c) Create a single data set from `mack.partial` and `mack.ices` that contains all of the variables in each data set, but is restricted to years that the two data sets have in common. One function will do this. Call this data frame `mack`.

d) Change the column names in `mack` that are uninformative to `"spawners.ices"` `"recruits.ices"` `"catch.ices"` `"fishMortality.ices"`.

e) Create this graph with the data in `mack`. Use `plot()` and then `lines()`.

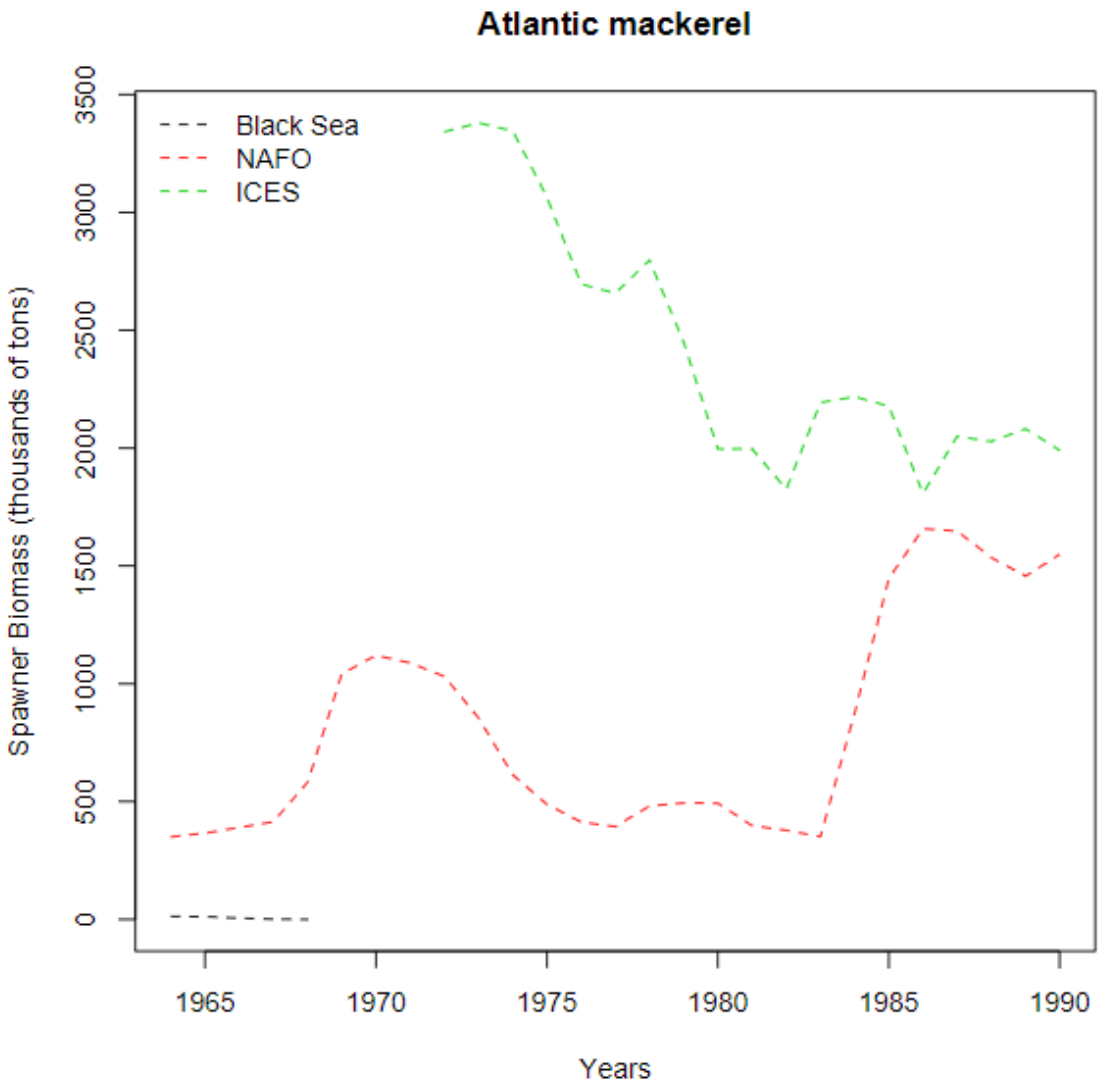

Figure 1: Figure to recreate

Question 2

For this question we'll be creating our own set of simulated data.

a) Create a data frame named temperature which has 2 columns: the dates Jan 1 2010 through Jun 30 2010 and a randomly generated temperature for each day. Use `rnorm` to generate the temperatures with the means listed below and a standard deviation of 5, then round to the nearest whole number. Hint: you only have to call `rnorm` once if you first create a vector of means.

```
##   Month Mean
## 1   Jan   40
## 2   Feb   42
## 3   Mar   51
## 4   Apr   55
## 5   May   58
## 6   Jun   62
```

b) Calculate the observed means for each month (based on the data you generated) and compare to the values specified in part a.

c) Check if you have any duplicate temperatures. On which days does this happen? (Note that this answer will change if you regenerate the data)

d) Now create a second data frame named observations which has 3 columns: the dates Jan 1 2010 through Jul 31 2010, but only every other day, the conditions (`sunny, cloudy, or partly cloudy`) as a `factor` (how you assign conditions to days is up to you), and the wind speed. Use `rnorm` to generate the wind speed with a mean of 5 and a standard deviation of 3, but change any negative values to 0.

e) Combine the 2 data frames into a single data frame `weather`, omitting any rows that don't match.

f) Calculate the `min` and `max` temperature for each of the conditions, as well as how many days of each condition there were.