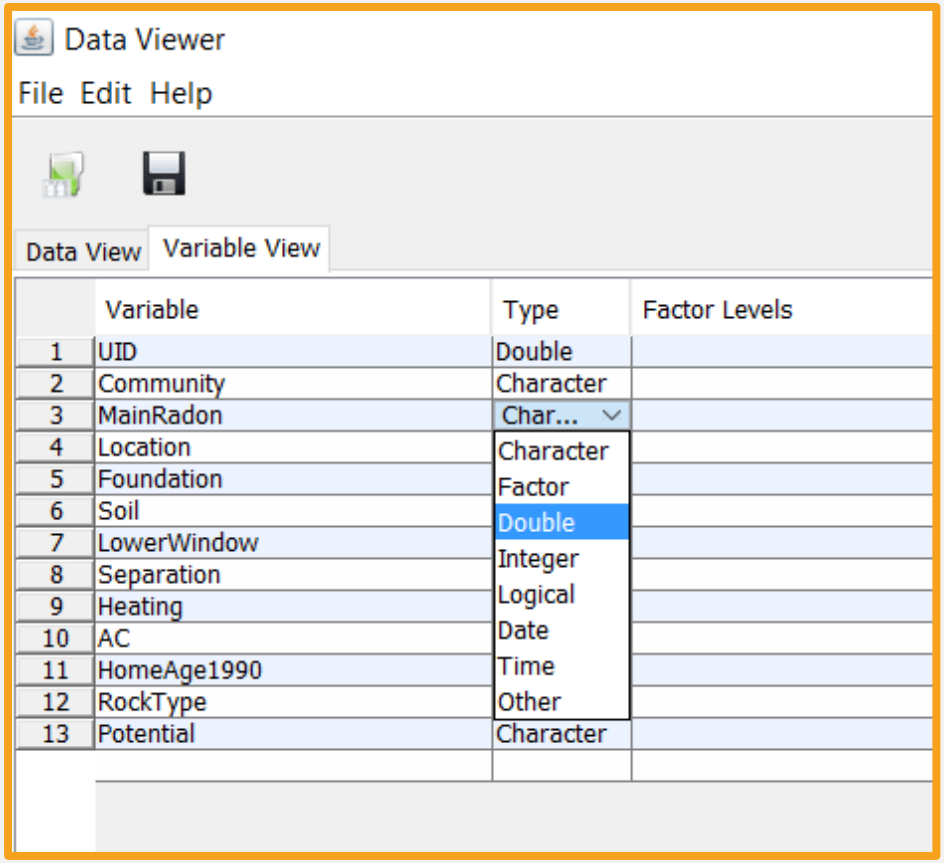
SPPH567 TUTORIAL II

TODAY'S TASKS

- Make changes to the data
- Calculate summary statistics
- Make frequency distribution plot (histogram)
- Make Q-Q plot
- Do goodness of fit test

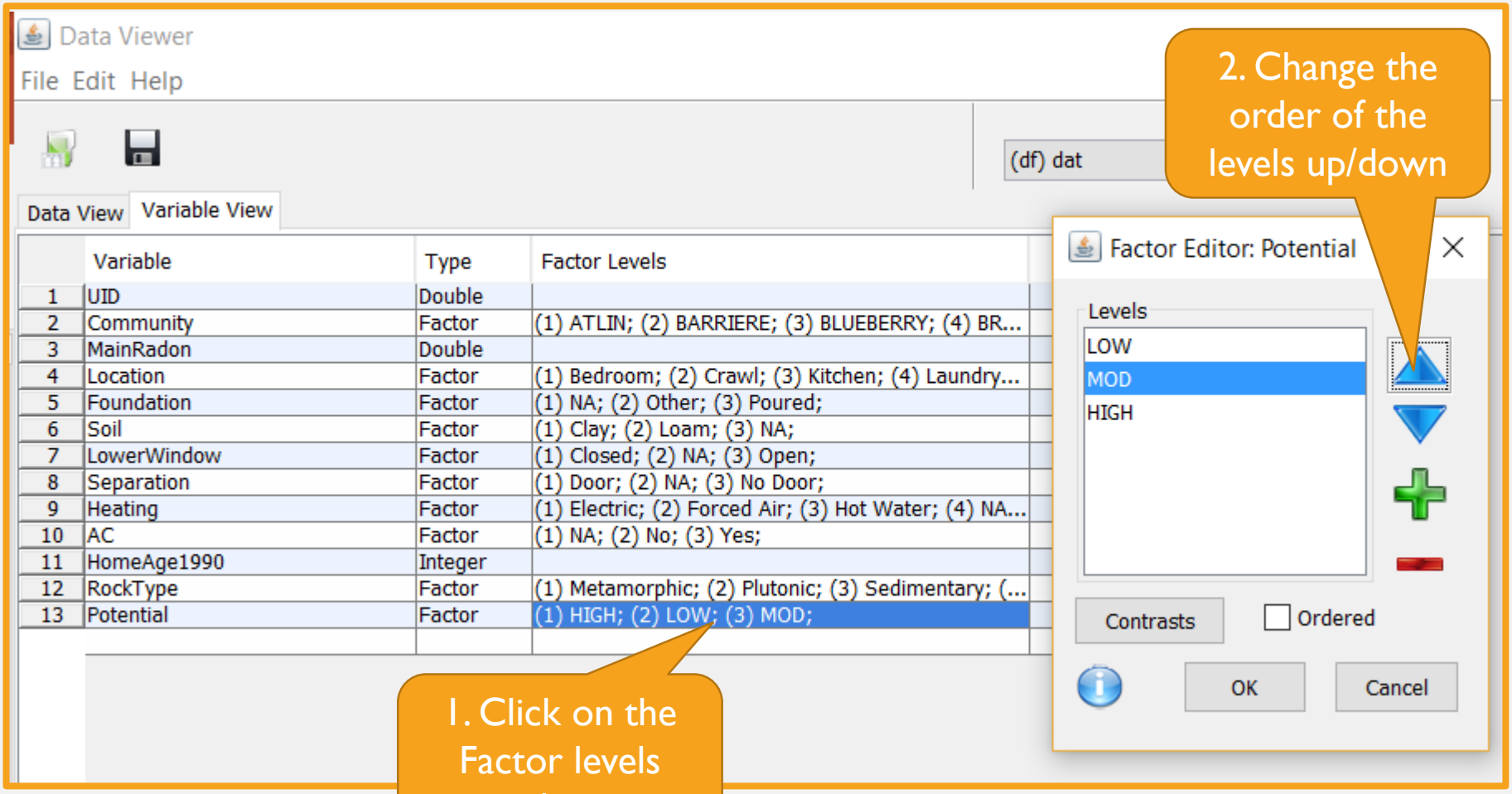
CHANGE DATA TYPE AND VARIABLE NAME

- Go to Variable View in the Data Viewer
- Click on the cell in the column Type for the variable you want to change
- Select the correct type in the drop-down menu.
- You can also change the name of the variable by double clicking the variable name

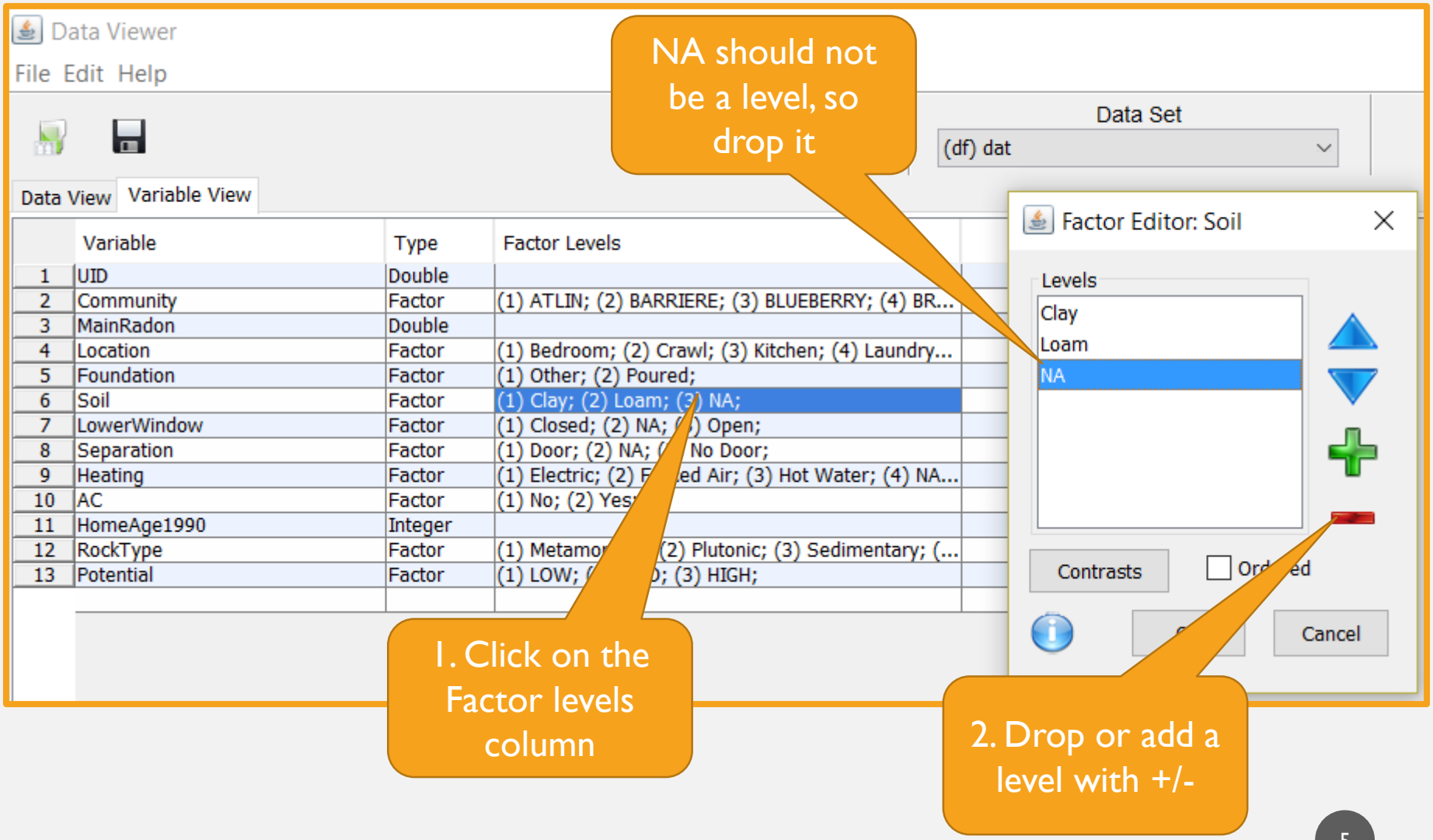


Categorical, dichotomous, Binary	➡ Factor
Continuous	➡ Double, Integer

CHANGE THE ORDER OF LEVELS FOR CATEGORICAL VARIABLES

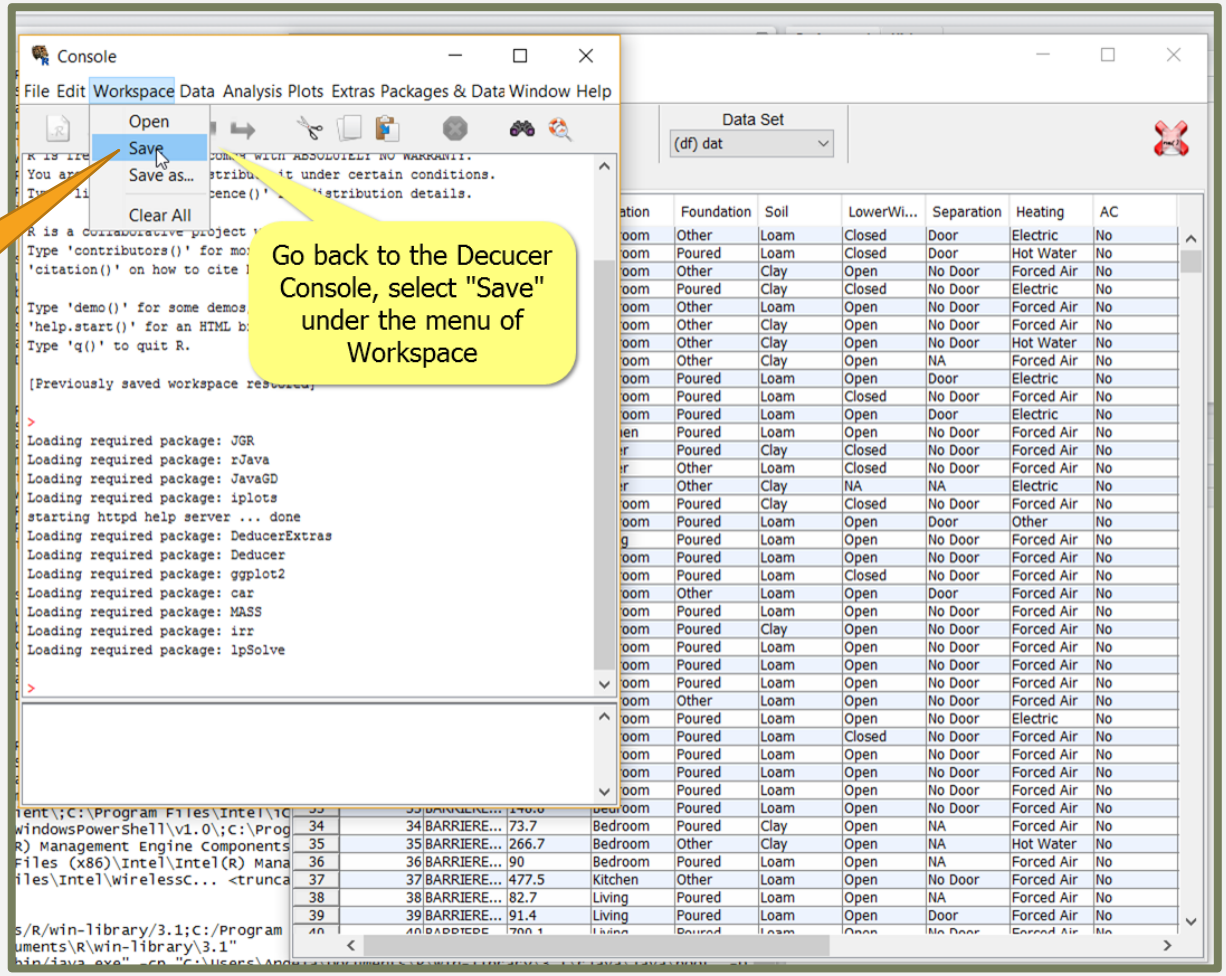


DROP/ADD A LEVEL FOR CATEGORICAL VARIABLES

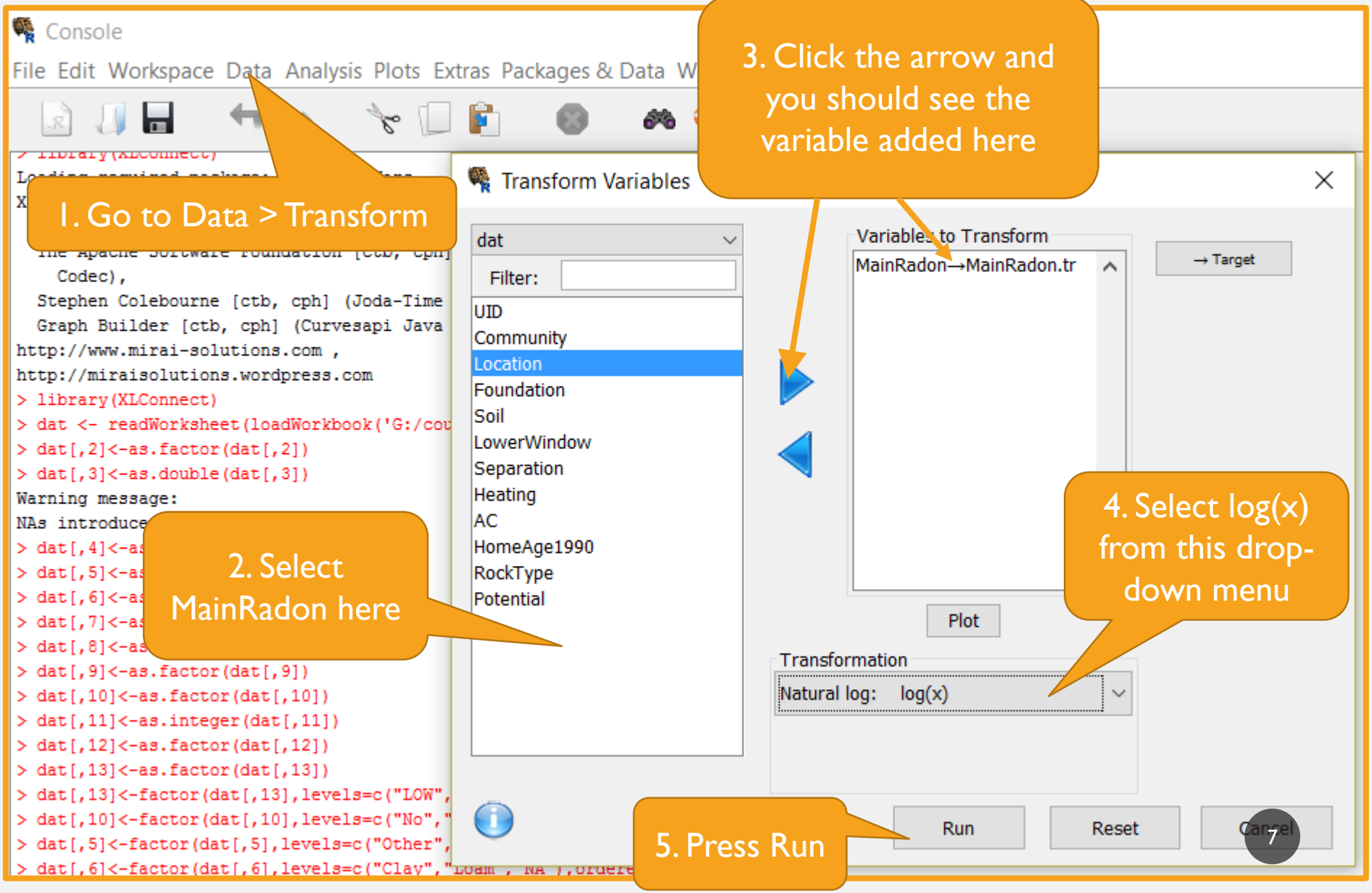


SAVE YOUR WORK ALONG THE WAY!

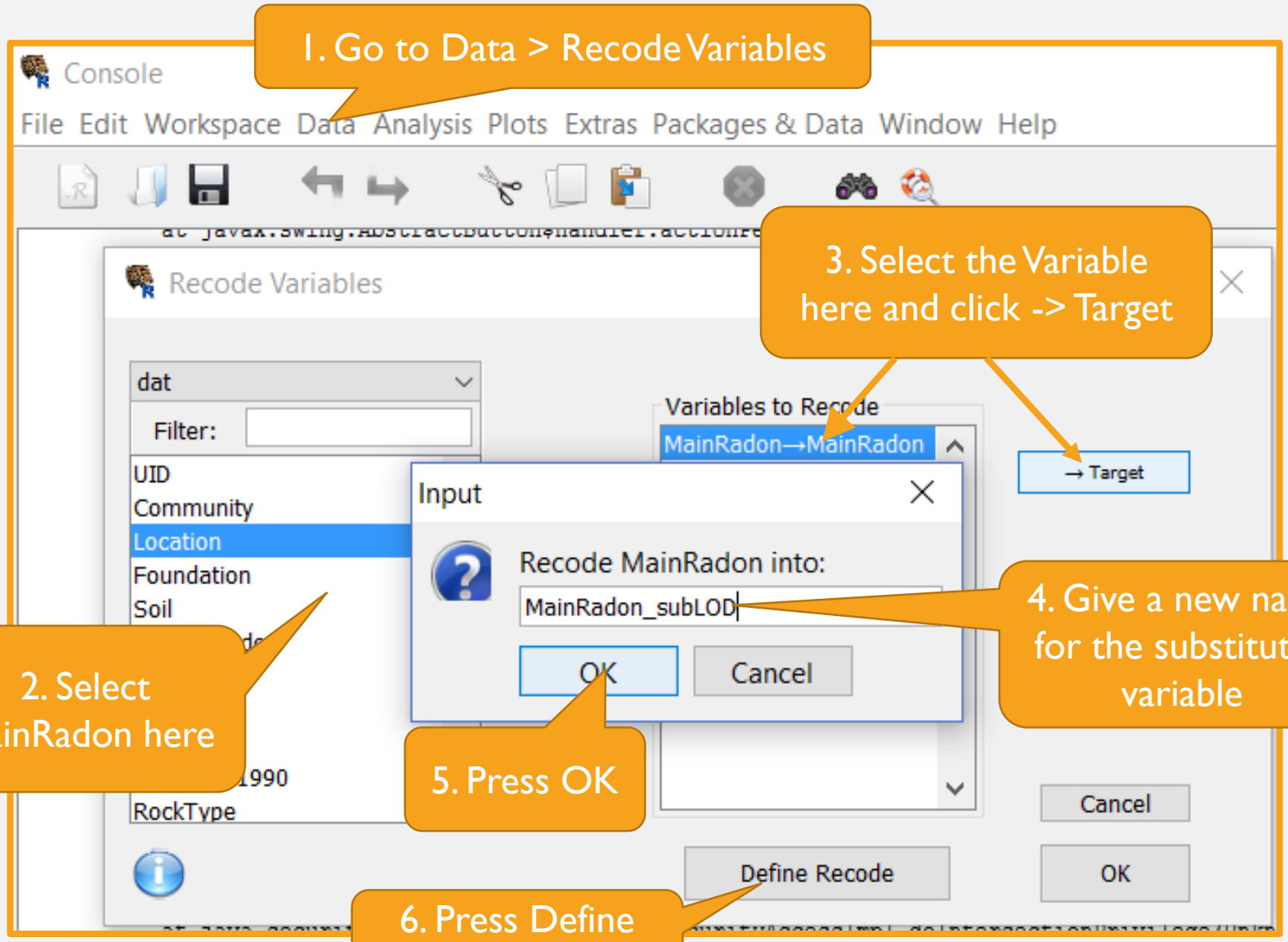
After making changes to your dataset, save the workspace to avoid accidents!



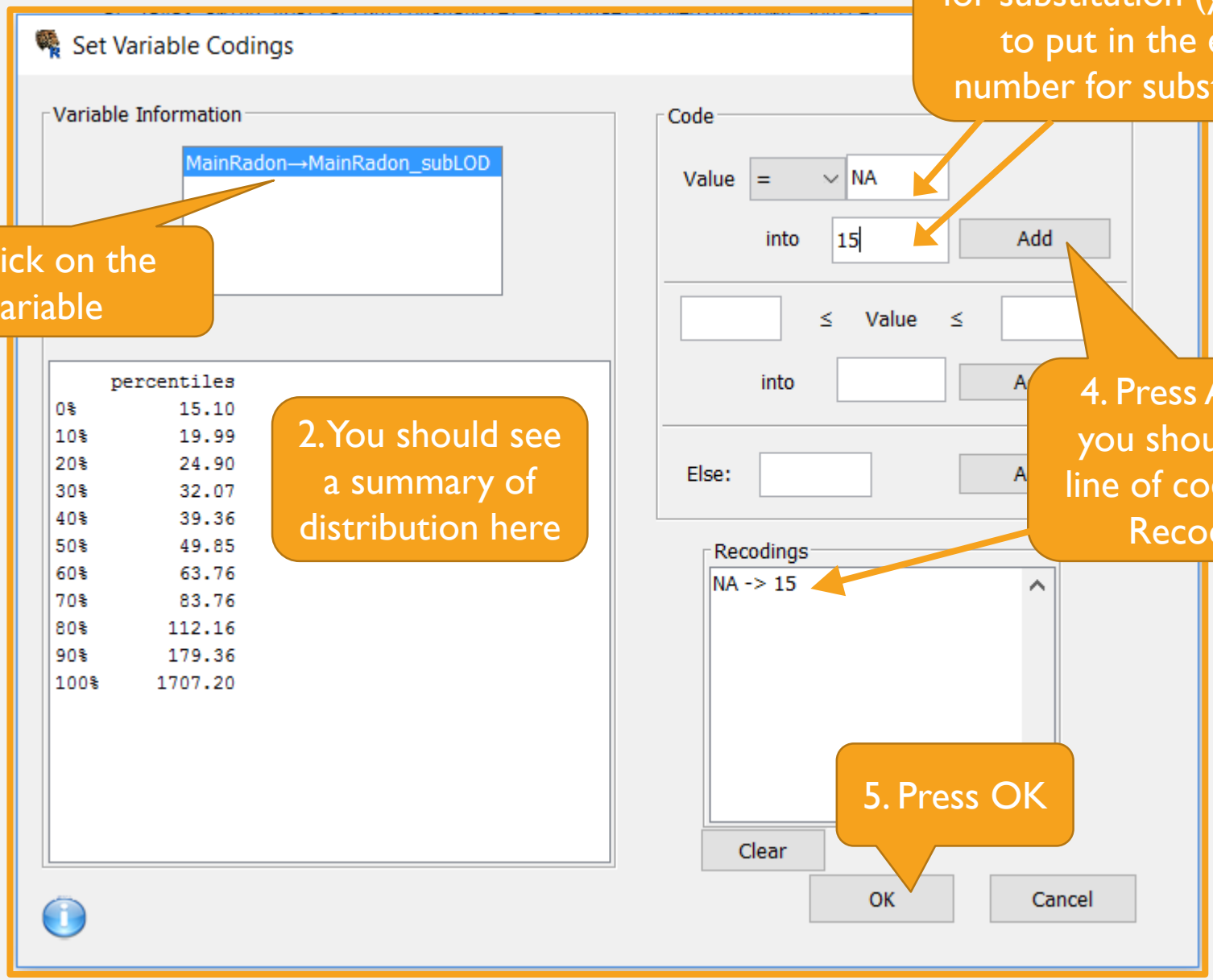
LOG TRANSFORM



SUBSTITUTE <LOD VALUES – I



SUBSTITUTE <LOD VALUES – 2



1. Click on the variable

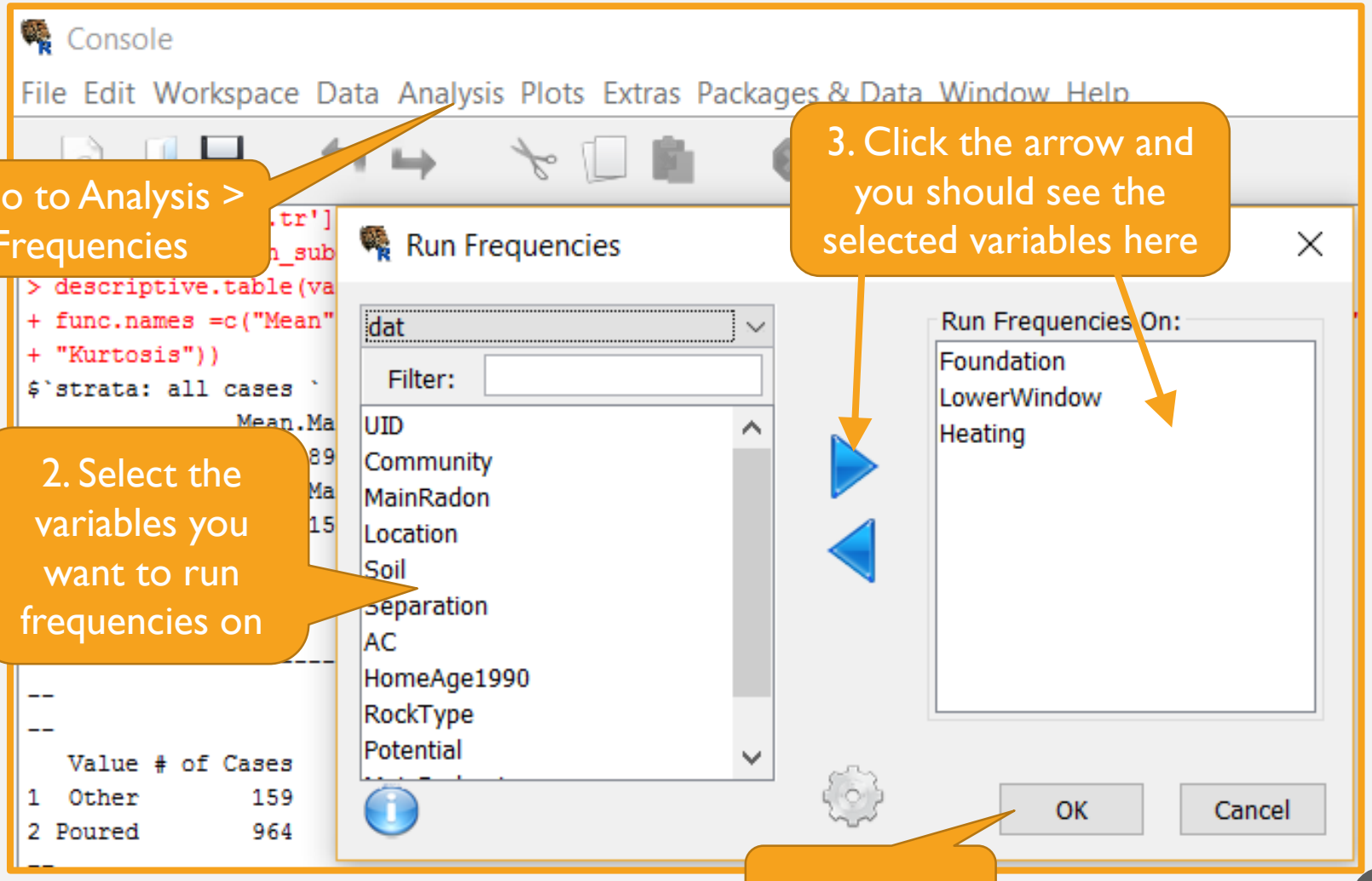
2. You should see a summary of distribution here

3. Setup here to define values of NA to the value for substitution (you need to put in the exact number for substitution)

4. Press Add and you should see a line of code under Recodings

5. Press OK

SUMMARY STATISTICS -- FREQUENCIES



SUMMARY STATISTICS – FREQUENCIES (CON'T)

In the Console window, you should see a table like this

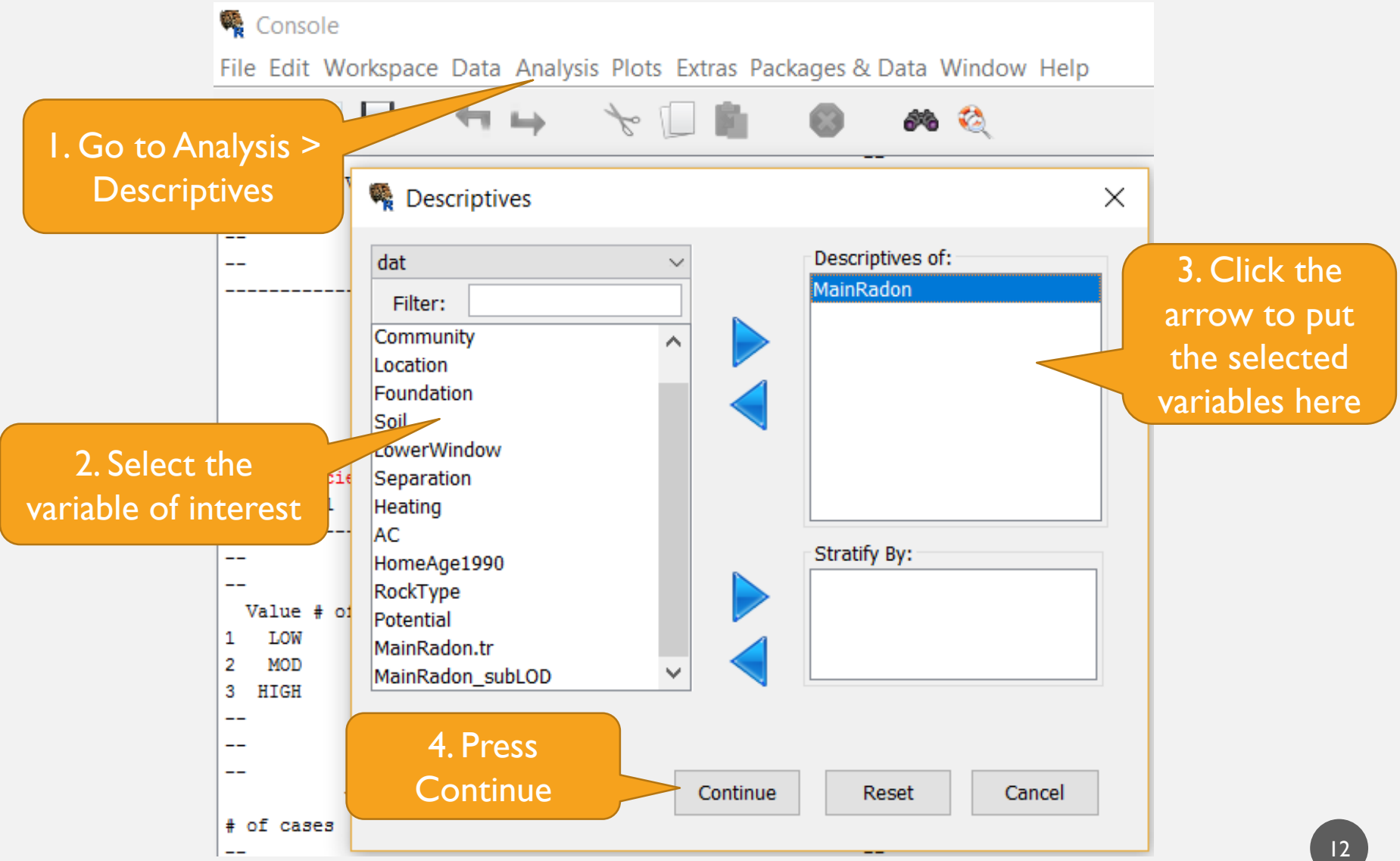
```
> frequencies(dat[c("Potential")] , r.digits = 1)
$Potential
-------------------------------------------------------------
--                        Frequencies                      --
--                                                         --
  Value # of Cases     %  Cumulative %
1   LOW         422  37.2          37.2
2   MOD         290  25.6          62.8
3  HIGH         422  37.2         100.0
--                                                         --
--                      Case Summary                       --
--                                                         --
           Valid Missing Total
# of cases  1134       0  1134
--                                                         --
--                                                         --
-------------------------------------------------------------
```

Frequencies – number of cases

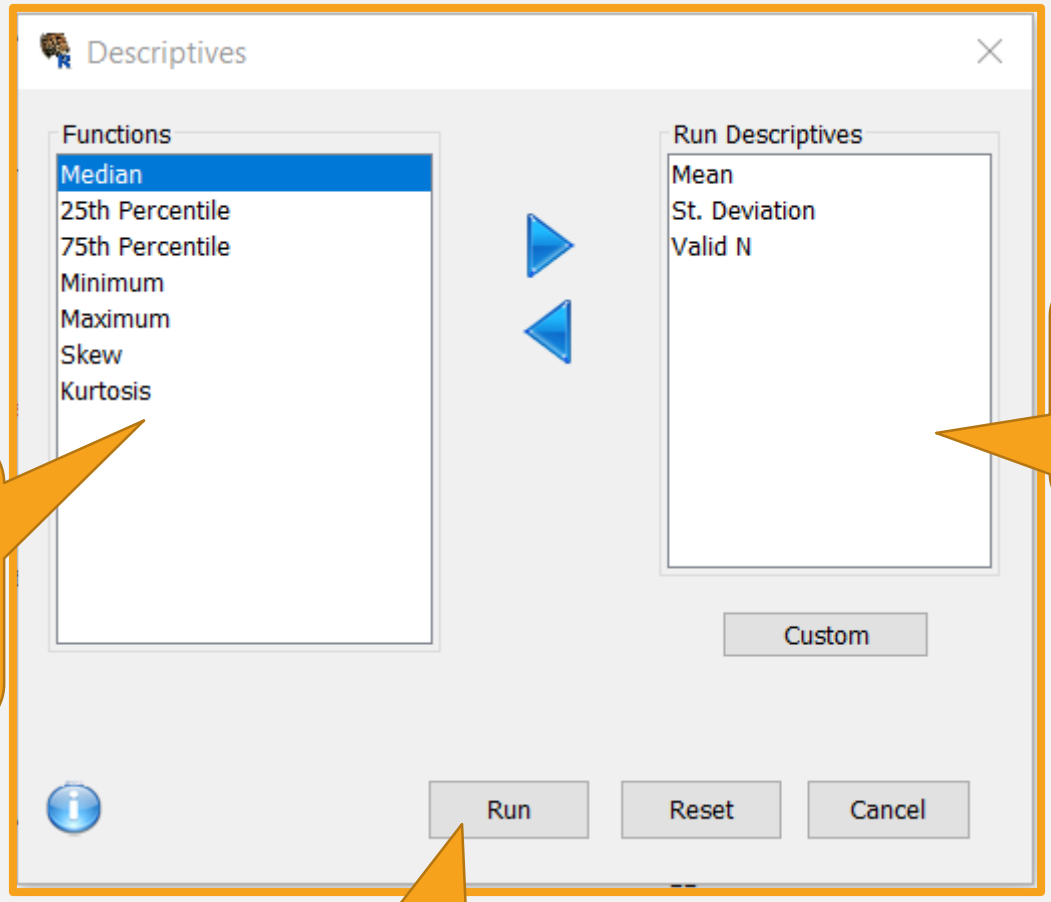
Proportion of the total number of valid cases

Cumulative proportion

SUMMARY STATISTICS – DESCRIPTIVES



SUMMARY STATISTICS – DESCRIPTIVES (CON'T)

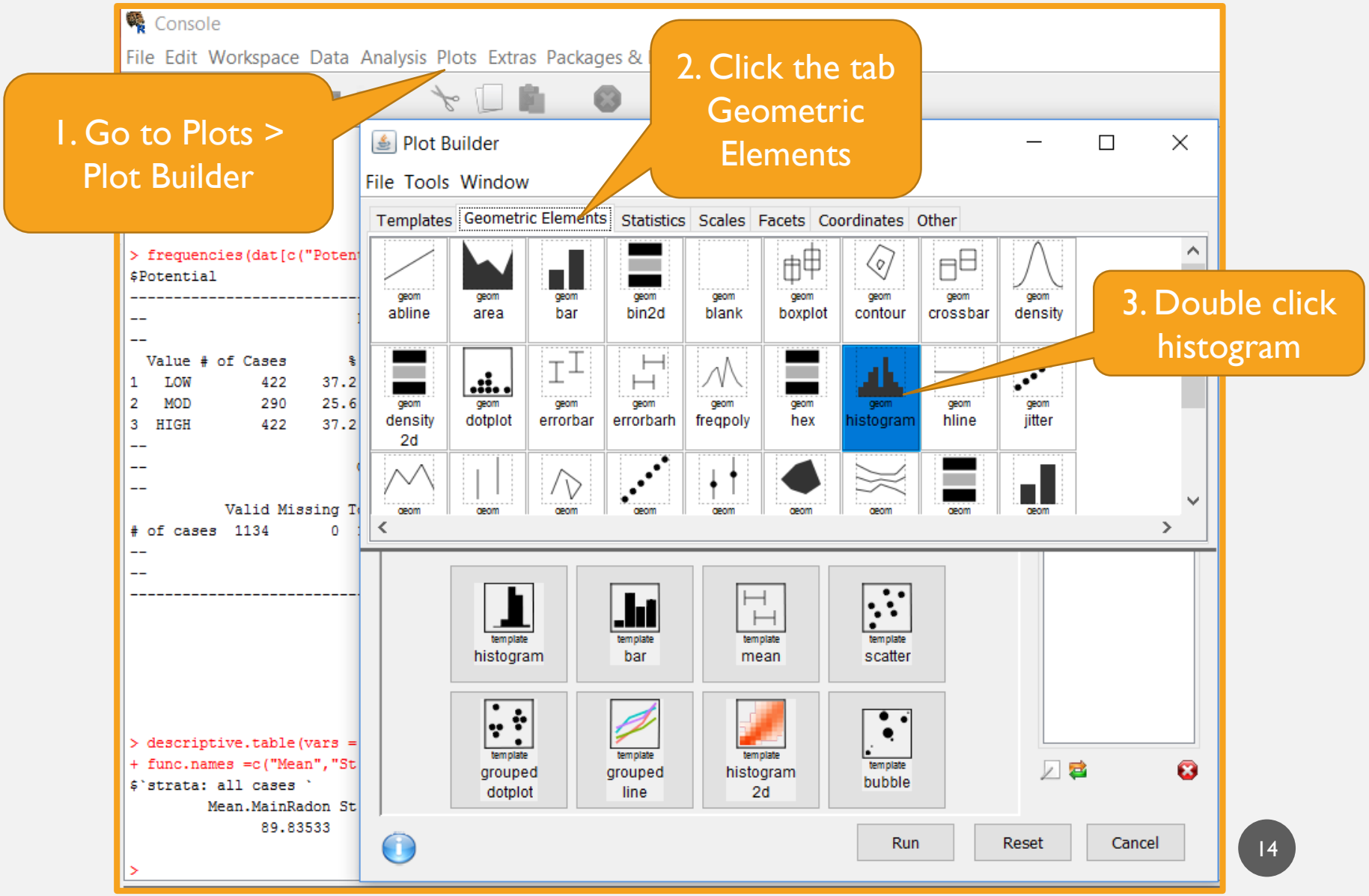


1. Select the descriptive statistics you want to calculate

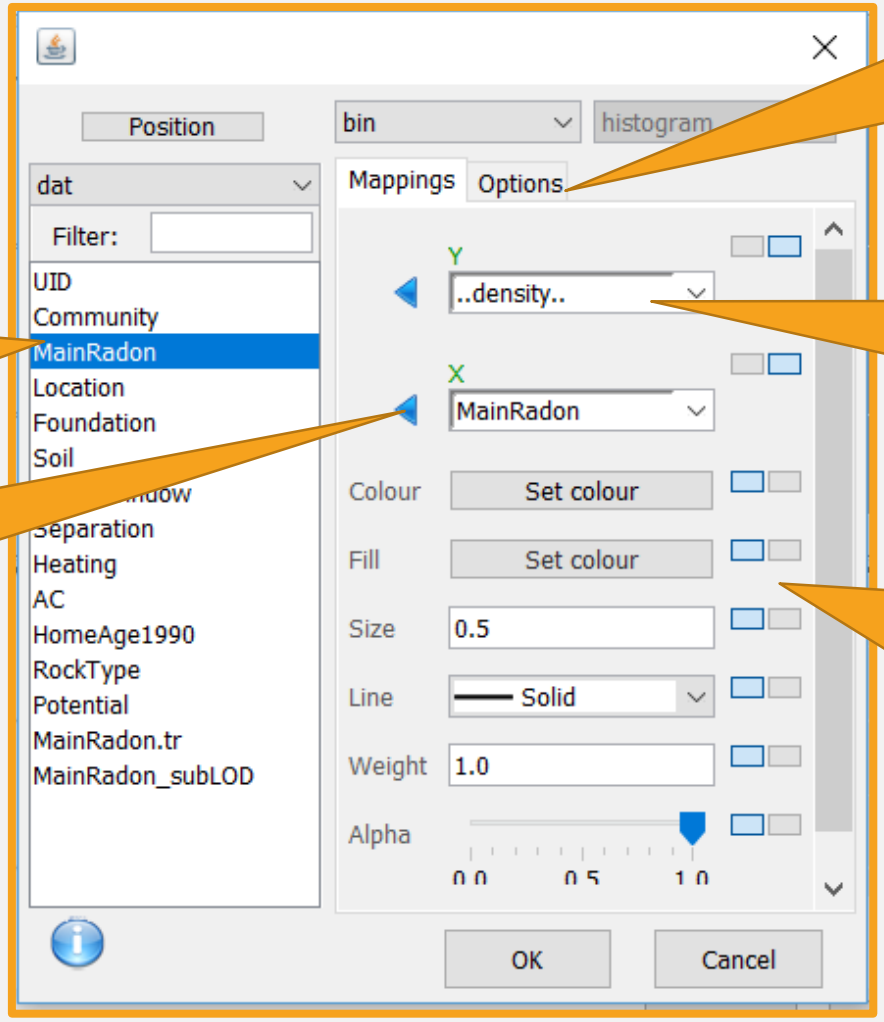
2. Click the arrow to put the selected variables here

3. Press Run

MAKE HISTOGRAM - 1



MAKE HISTOGRAM - 2



1. Select the variable to put at X axis

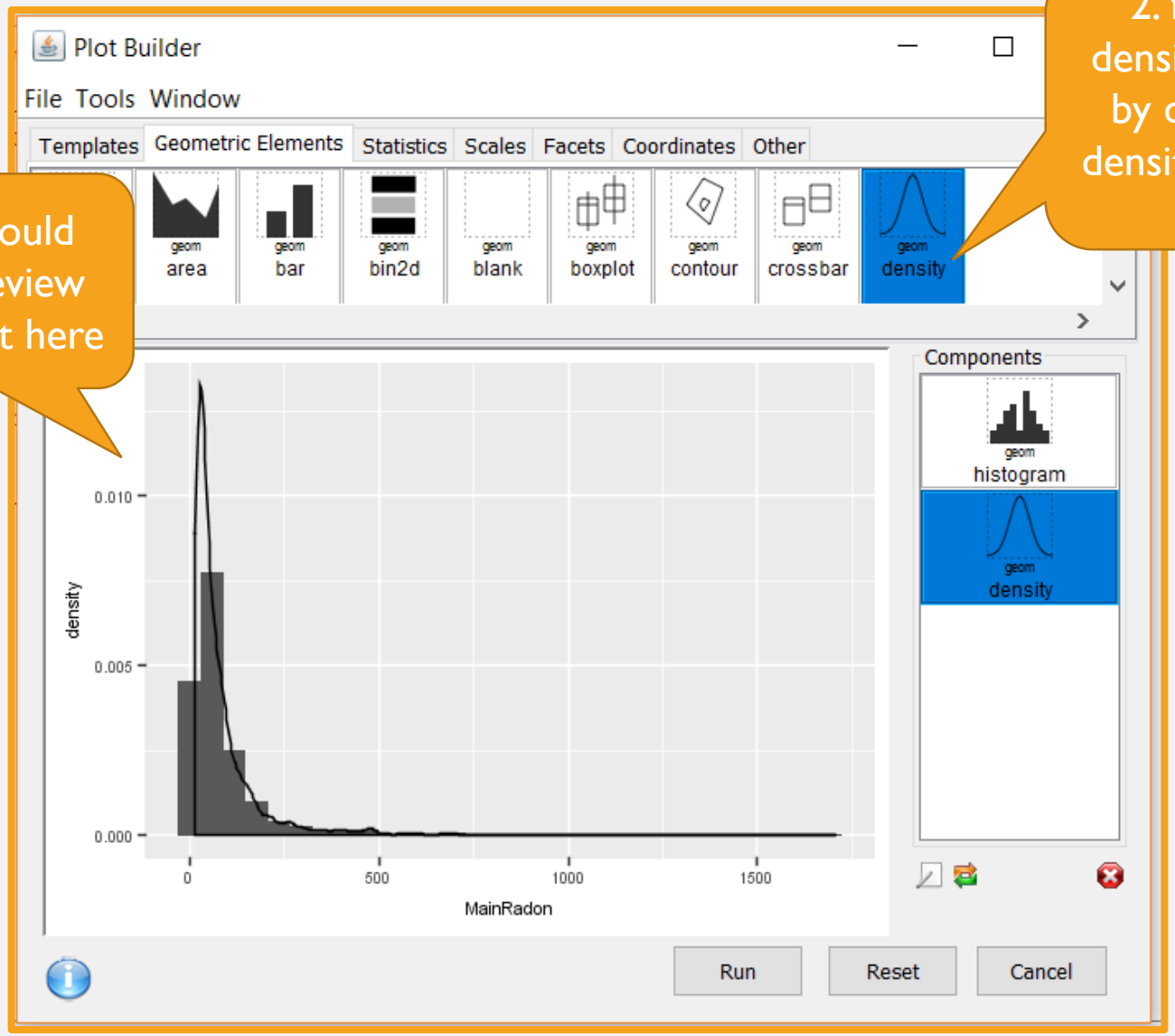
2. Click this arrow to add the variable to X

3. Select density or count for Y

4. You can change the color/fill or other graphical setting

5. You can change the bin width in the Options

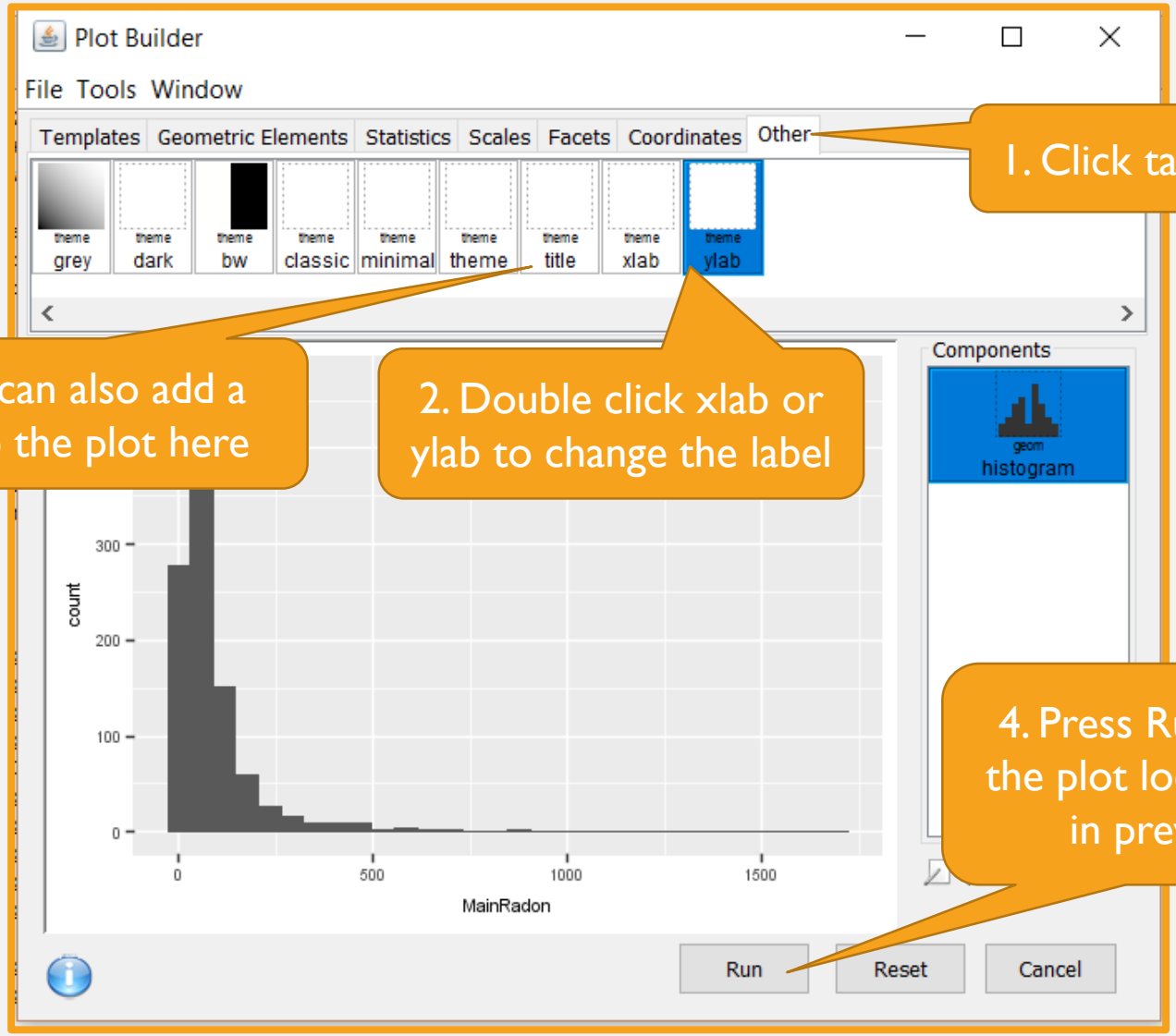
MAKE HISTOGRAM - 3



1. You would see a preview of the plot here

2. You can add a density function line by double clicking density in Geometric Elements

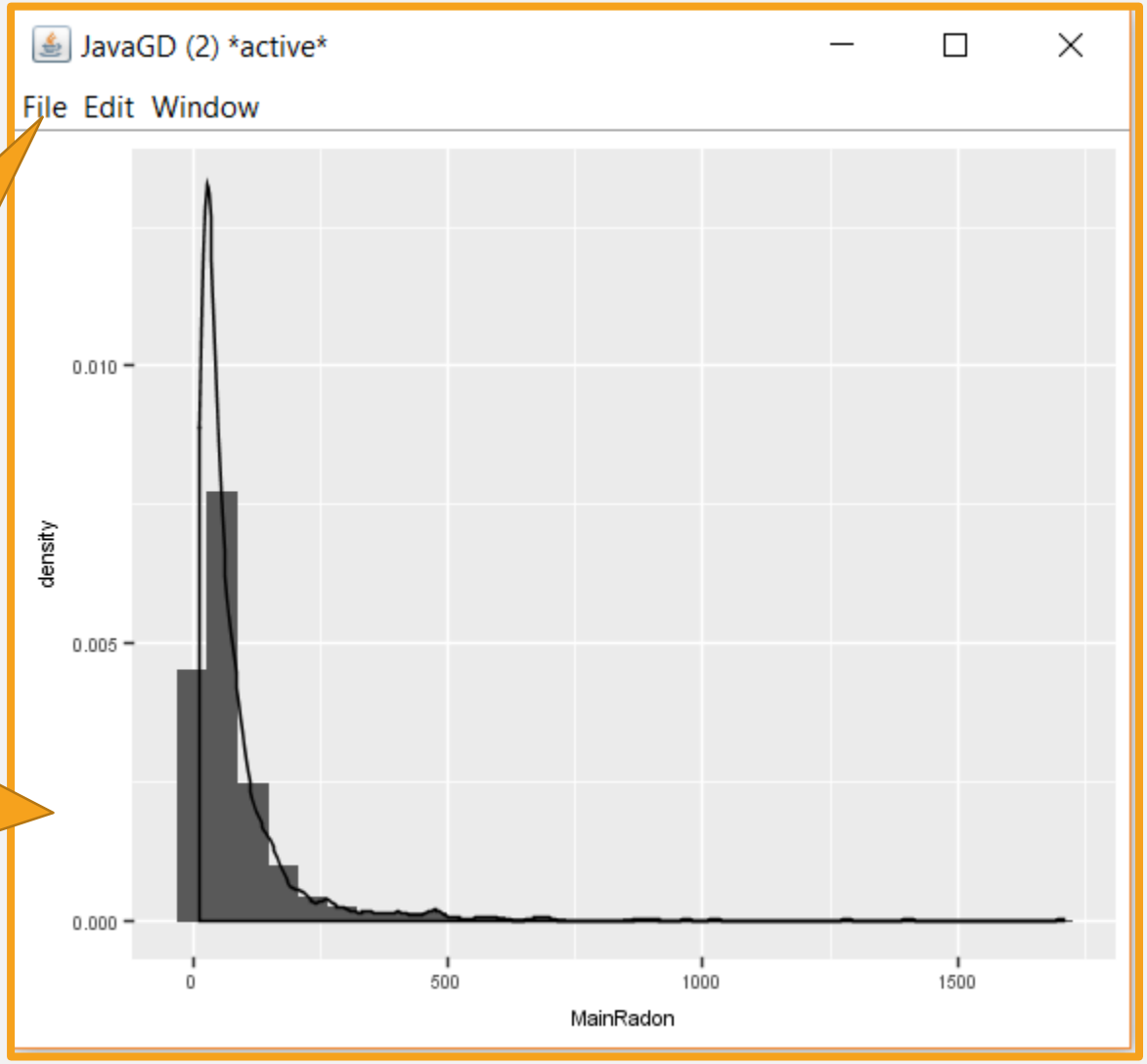
CHANGE X OR Y LABELS



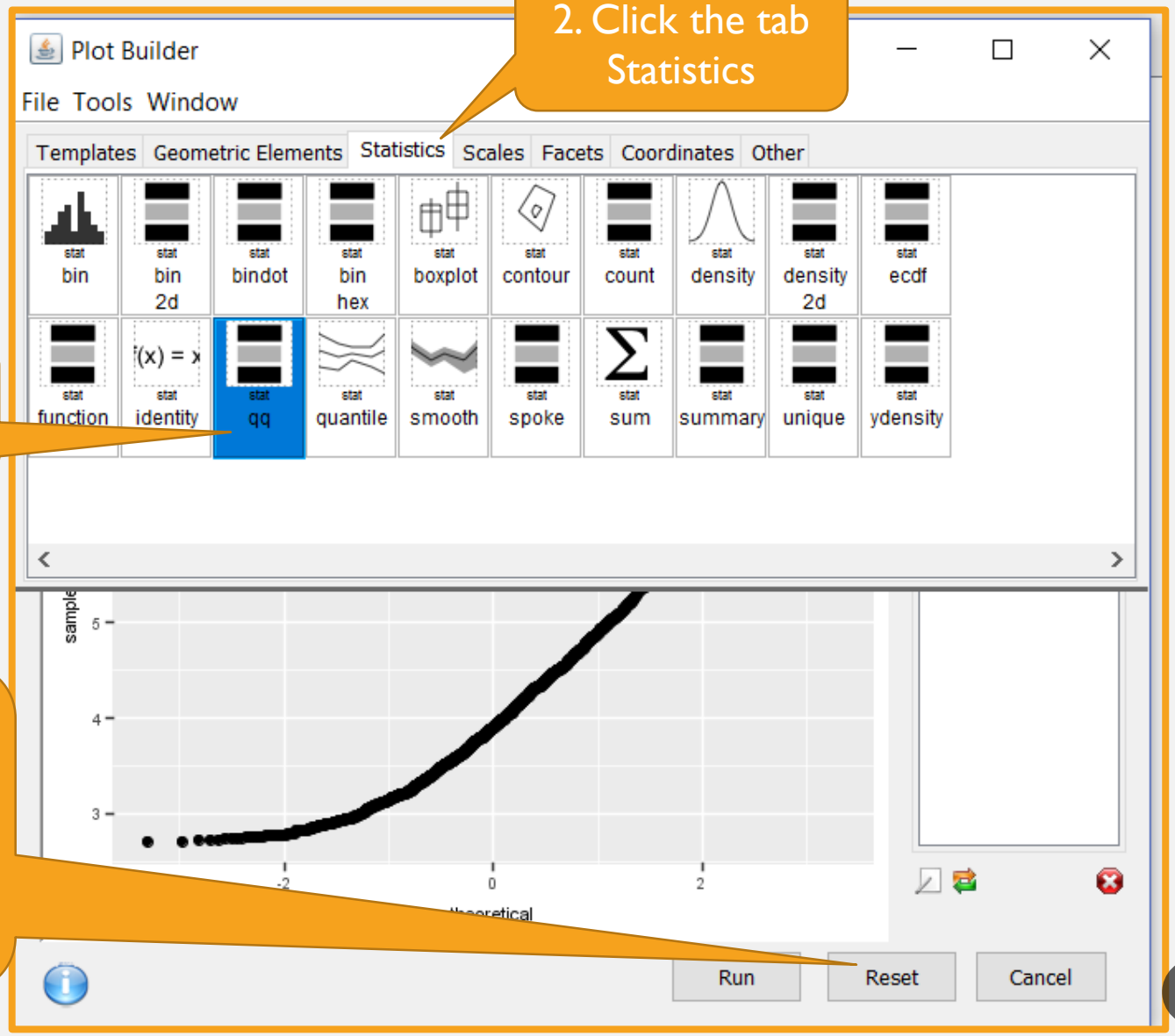
MAKE HISTOGRAM - 4

2. To save the plot, go to File > save as any format in the options

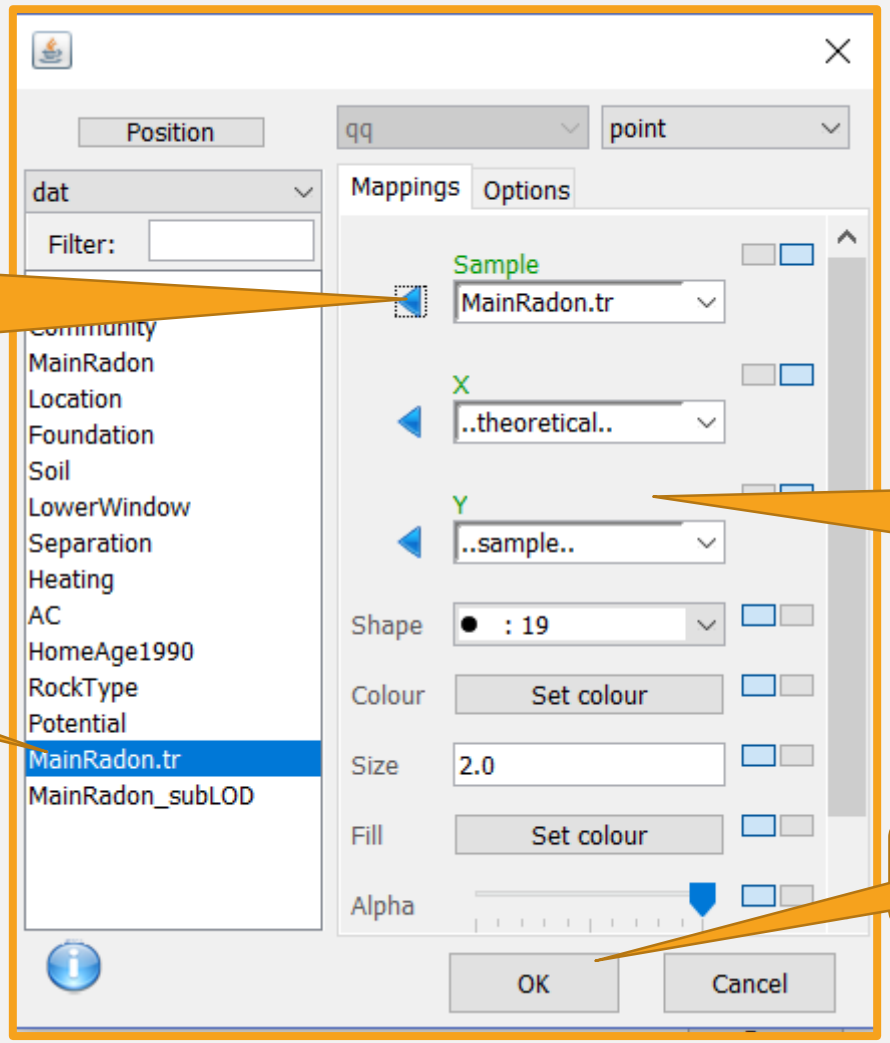
1. Now the plot is made and you can change the size by dragging the window



MAKE Q-Q PLOT - 1



MAKE Q-Q PLOT - 2



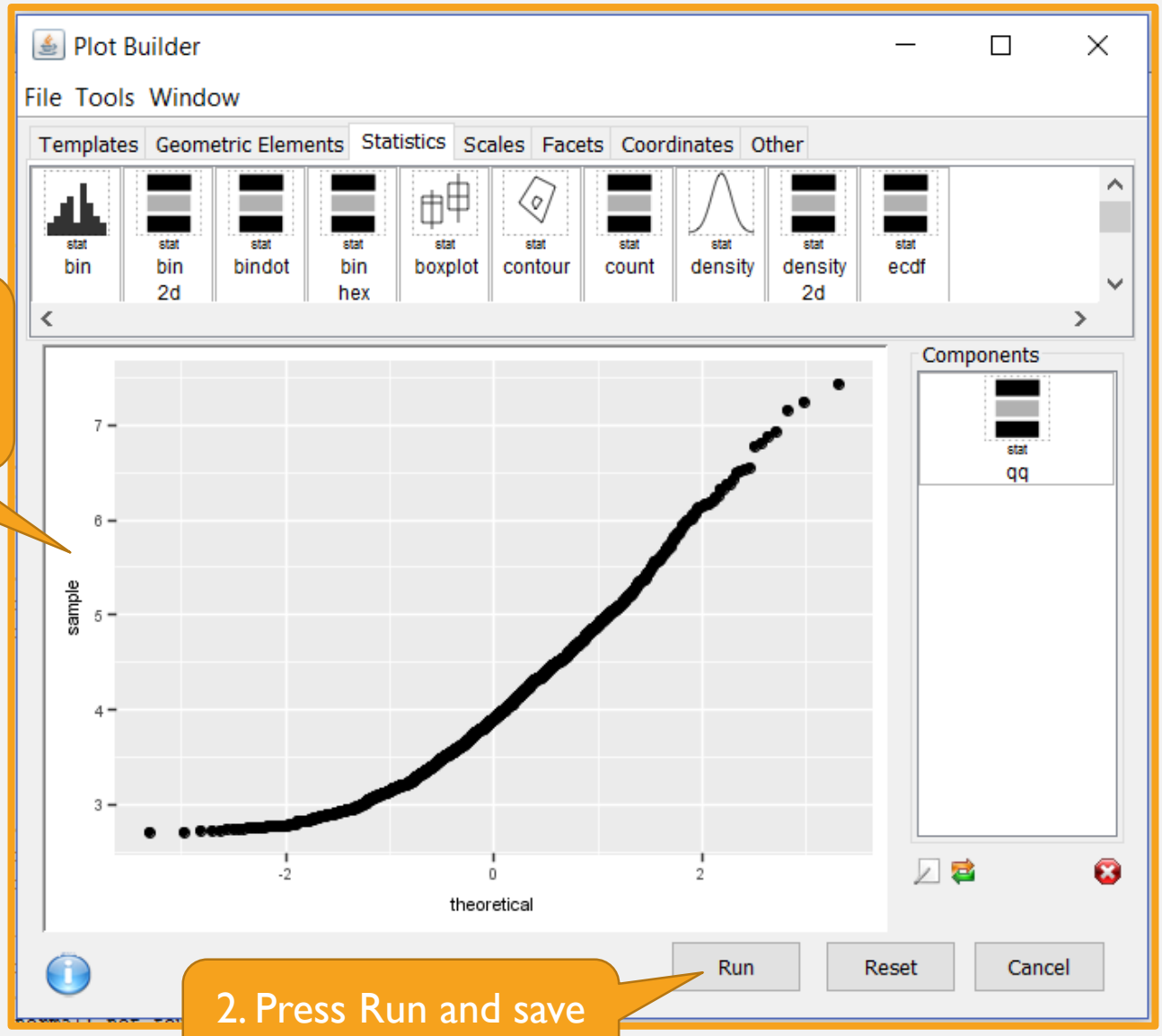
1. Select the variable to plot

2. Click the arrow here to add the variable to Sample

3. Make sure X is theoretical, and Y is sample

4. Press OK

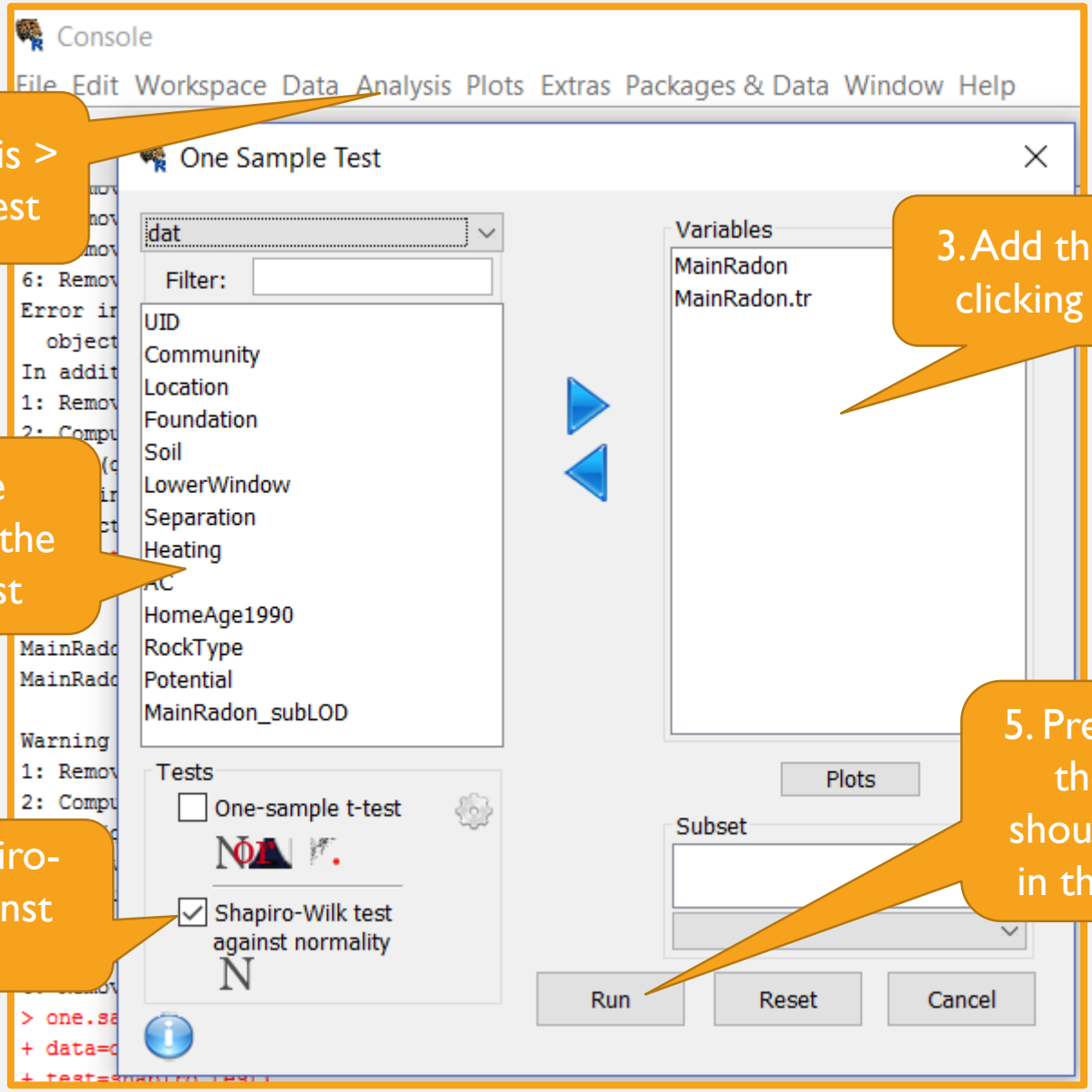
MAKE Q-Q PLOT - 3



1. Now we have the preview of the qq plot

2. Press Run and save the plot as in page 16

GOODNESS OF FIT TEST



1. Go to Analysis > One Sample Test

2. Select the variable(s) for the normality test

3. Add them here by clicking the arrow

4. Check Shapiro-Wilk test against normality

5. Press Run and the results should show up in the Console