

2.5 Exercise: Categorical variables

In order to get insights from your data you will need to graph it.

In this exercise you will construct the types of graphs and summaries we used to interpret data in the previous two videos.

You will learn how to use iNZight to:

1. Get bar charts and summary tables for your categorical variables.
2. Colour the bars.
3. Re-order the bars.

Note: Categorical variables will always produce a bar chart.

[iNZight Lite version [linked here](#)]

We advise that you watch the *iNZight instructional video* [linked here](#) before doing the Exercise (*deals with numeric and categorical variables*).

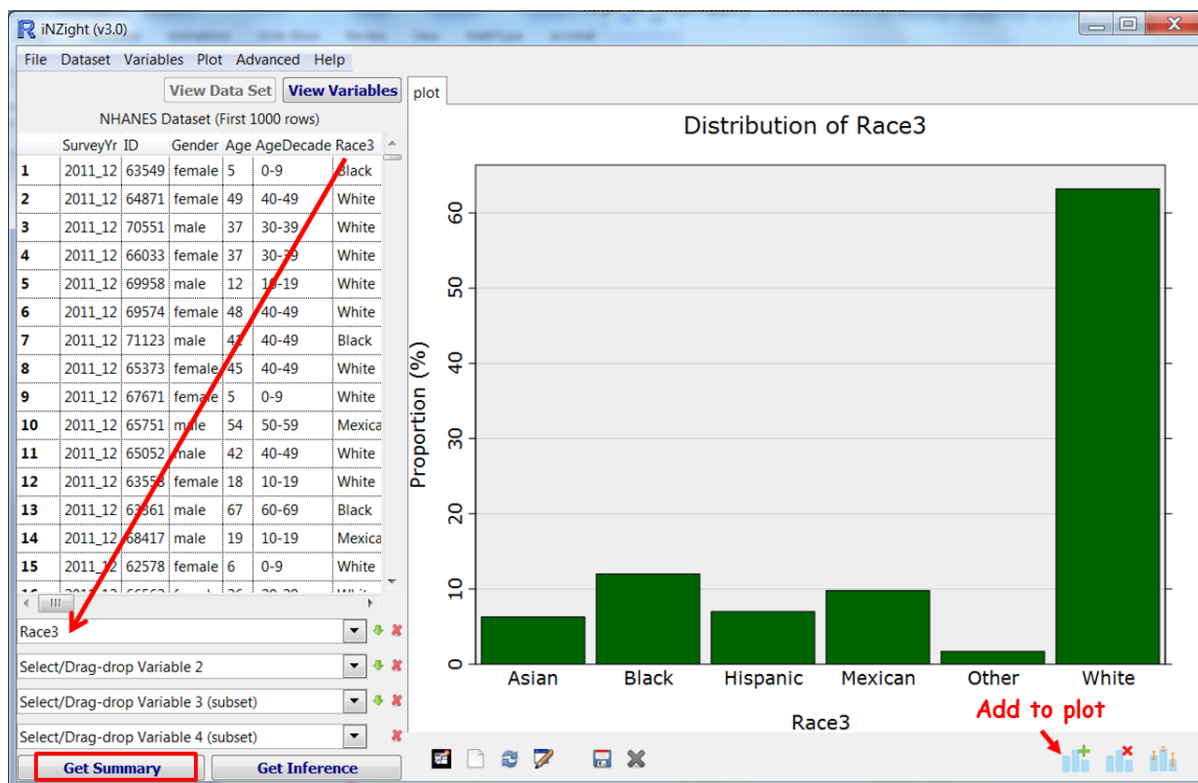
INSTRUCTIONS

Follow the instructions below to generate the graphs. Or you may prefer to [print the instructions](#). If you have a problem doing the exercise, scroll down to **Common questions**.

Load the `nhanes_1000` dataset into iNZight using **File > Example data** You will find the data set in **Module (package) FutureLearn**.

Make bar charts of categorical variables

Drag the variable name **Race3** into the **Variable1** slot. iNZight will automatically produce a bar chart for a categorical variable when you drag its name into the **Variable1** slot.



Generate the Summary window

Click **Get Summary** to get the associated summary table of counts and percentages.

The screenshot shows the 'Summary' window in iNZight. It displays the following information:

- Primary variable of interest: Race3 (factor)
- Total number of observations: 1000
- Summary of the distribution of Race3:

	Asian	Black	Hispanic	Mexican	Other	White	Total
Count	63	120	70	98	17	632	1000
Percent	6.3%	12.0%	7.0%	9.8%	1.7%	63.2%	100%

PRACTICE (~5 min)

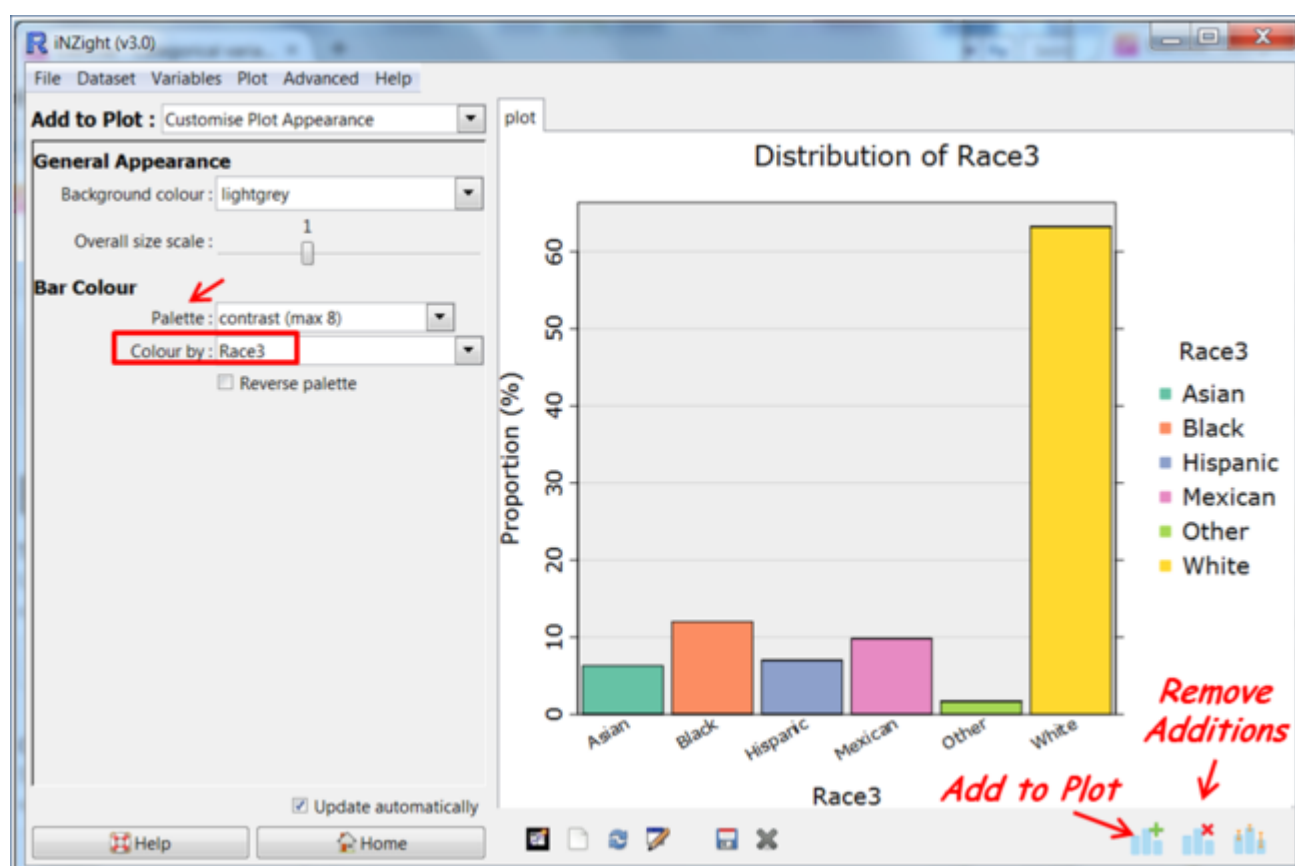
Choose another **categorical variable** from the dataset and plot it. Try a couple more and do this until you are confident that you can get simple graphs and tables like these quickly.

[Note: When you drag the name of a numeric variable you get a different form of plot – a dotplot – and a different kind of summary.]

Colour the bars

We will colour the bars of the **Race3** bar chart so ensure you have **Race3** in the first variable slot.

Use the **Add to Plot** function on the bottom menu under the plot window, **Customise plot appearance**. Under **Colour by** choose **Race3**. (Add to Plot can also be found under *Plot > Add to plot ...*)



Try changing the colour **Palette**. (*Speed Tip: Click into the Palette box and use the up and down arrow keys to explore the choices.*)

If you have coloured by a variable, that colour-by instruction will remain in force until you click **Remove Additions** and select **Remove colour by** or colour by another variable.

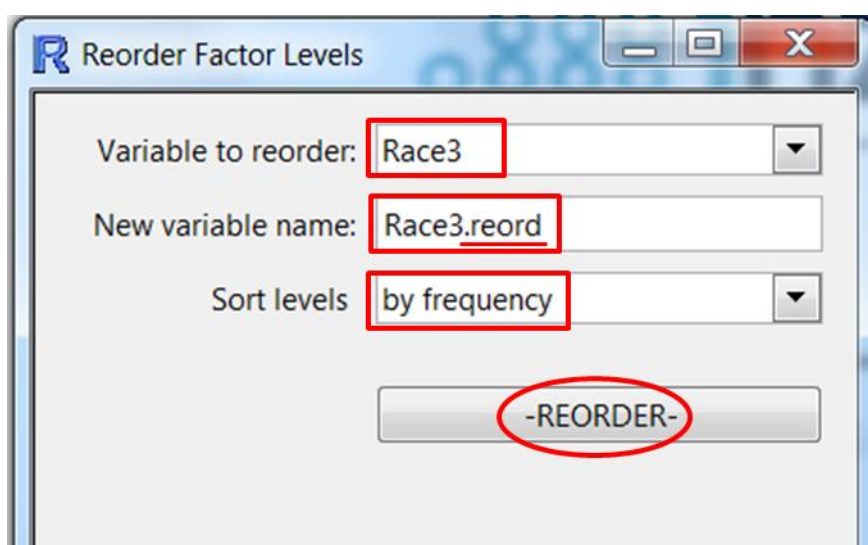
Reorder the bars by frequency

iNZight initially puts the bars of a graph in alphanumeric order. This is not always the best way to present a graph. It is often better to reorder your bars by either frequency or natural order.

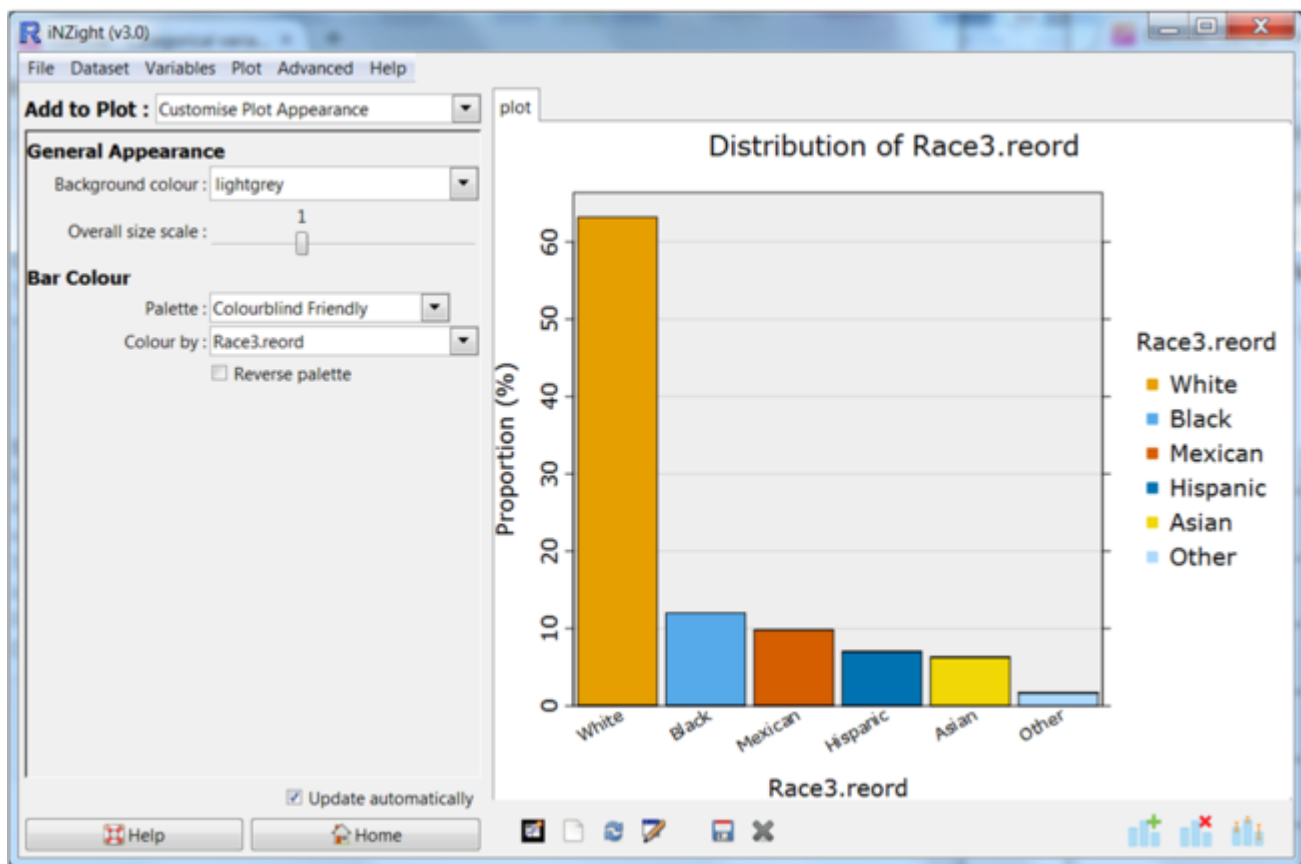
From the top menu, go:

Variables > Categorical Variables > Reorder Levels

- Select **Race3** from the drop down list
- Give the variable a new name or accept the name offered, **Race3.reord**
- Select Sort levels: **by frequency** (the default) and click **REORDER**. This creates a new variable with the desired ordering.



- ***This does not change the graph.*** It just creates a new variable.
- ***You will need to press the Home button*** (at bottom) to return to the Home command panel so that you can plot other variables.
- Now drag **Race3.reord** (or your new variable name) into the **Variable1** slot.
- Recolour by **Race3.reord**.



(I've chosen the so-called Colourblind friendly palette.)

Manually reorder the bars

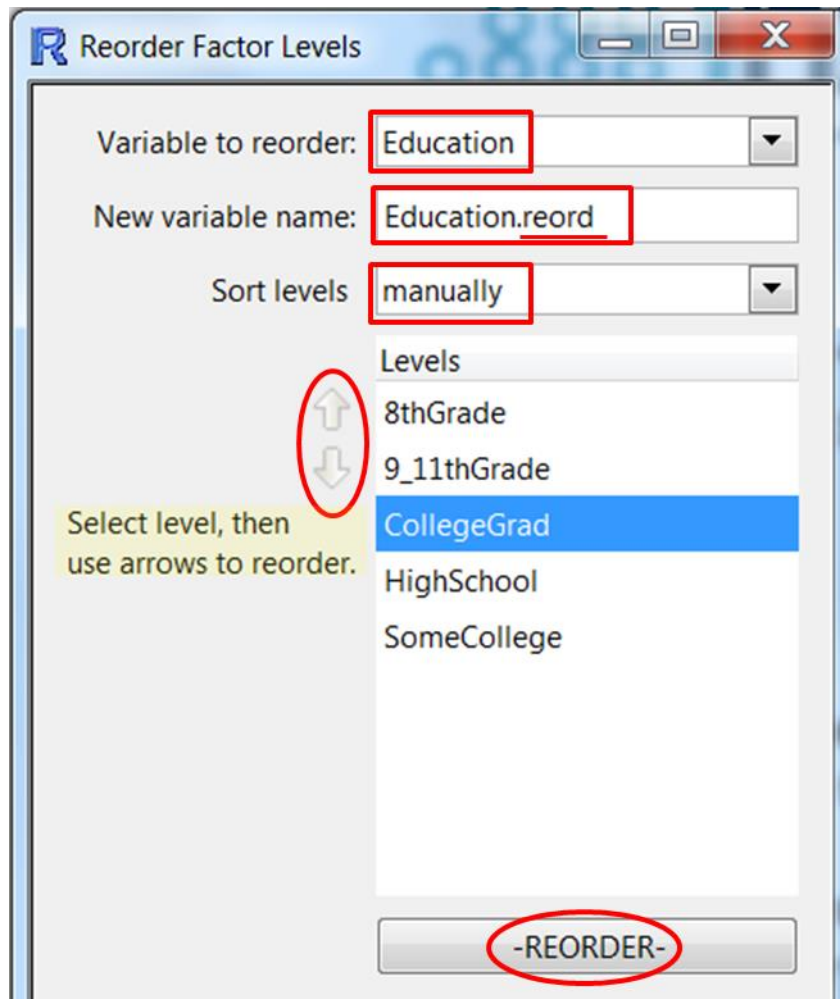
If the data categories are ordinal (have a natural order) you will have to manually reorder them.

Create a bar chart for **Education**. The categories should be displayed from lowest to highest level of education. That is, 8thGrade, 9-11thGrade, HighSchool, SomeCollege, CollegeGrad, but currently they are not. To reorder **Education**:

From the top menu, go to:

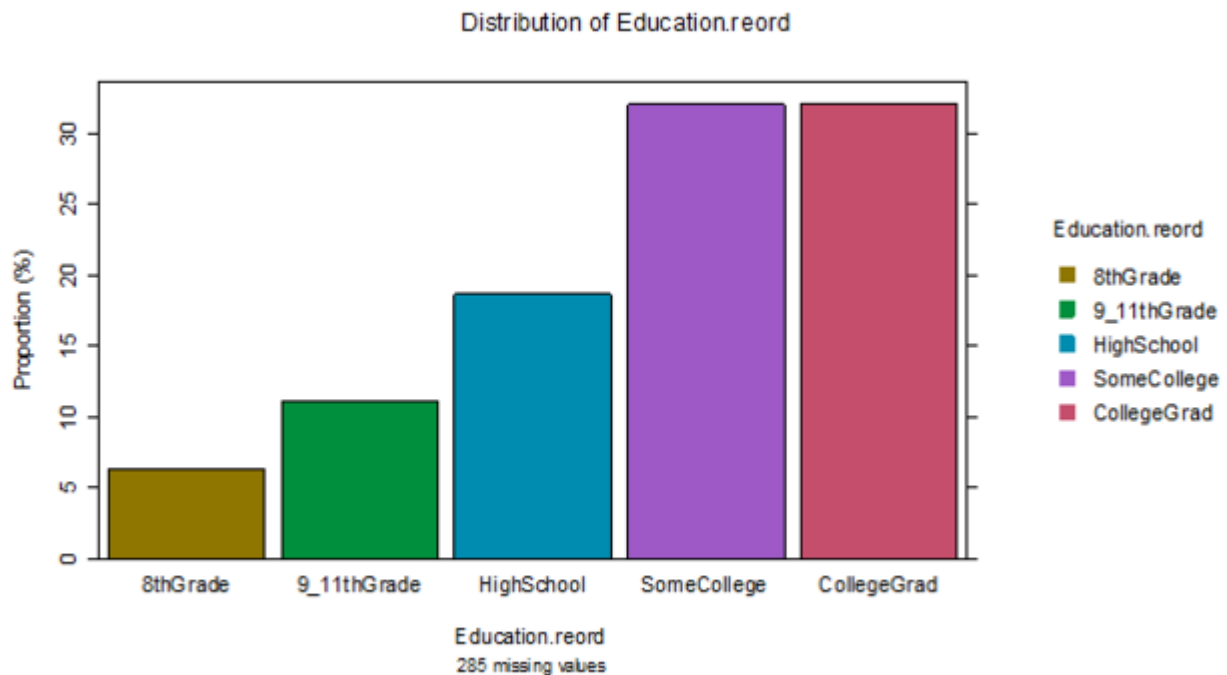
Variables > Categorical Variables > Reorder Levels

Under **Sort levels** click **manually**. To change the ordering, click on the name of one of the levels you want to move and then click on the up and down arrow keys (circled in red on the diagram below) to move it up or down the list. When you have finished re-arranging the list into the order you want, click **REORDER**.



This will create a new variable **Education.reord** with the desired ordering. Now, close pop-up windows and :

- Plot **Education.reord** (by dragging its name to the variable 1 slot)
- Colour by **Education.reord**.




(This one is using the rainbow(hcl) palette.)

Optional:

Interactive versions of the plot

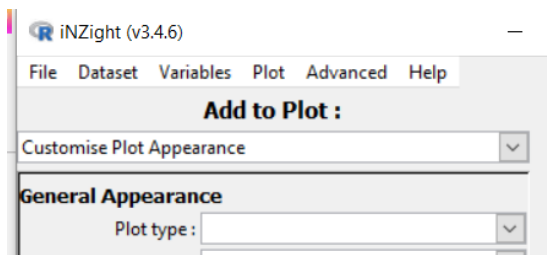
Save your graph as **File Type: Interactive HTML** (you will have to supply a name for the file). The file will open up in your default browser. If that is a modern browser like Chrome, Firefox or Safari (but not Internet Explorer) this will then give you an interactive version of the graph that lets you query it in various ways like hovering over bars or clicking them. Explore!

You can give such files to others. They do not need to be connected to iNZight to work.

Alternatively when this icon  appears and is blue, simply clicking it will fire up an interactive version of the current plot.

Other ways of representing data on a categorical variable (iNZight versions from 3.4.6)

There are several ways of plotting the data on a single categorical variable. Go to **Add to Plot** and look at what is delivered by the various options under **Plot type**. Can you see relationships between the ways the various types of graph represent the information? Play with some of the controls for each plot type.



Common questions

When I resized the window my plot disappeared.

Click the **Redraw plot** icon (you can hover over the icons to see what they do).

How do I get rid of the colour segmentation in the bars?

You have a colour-by setting that is still active. Go to **Remove Additions** underneath the plot and then use **Remove colour by**.

The plot didn't change when I used reorder categories and clicked "Reorder".

Clicking re-order does not create a new graph, it creates a new variable (e.g. race3.reord). To get the new (reordered) graph you have to drag/select the new variable.

[The behaviour of iNZight Lite is different ...](#)

My bar labels are on an angle. How do I make them horizontal?

Stretch the plot window horizontally until there is room for them to become horizontal and then click the **Redraw plot** icon (underneath the plot window) to redraw the plot. Another option is **Add to plot** and under **Overall Plot Appearance**, reduce **Overall size scale** (makes every plot element smaller).

I don't like the headings. How do I change them?

Use **Add to Plot > Axes and labels**.

How do I save my graph from iNZight for use in another document?

Click on the **Save Plot** icon at the bottom of your graph, or go **Plot > Save Plot**. There several formats available, e.g. pdf, png, ...

How do I copy and paste my graph from iNZight into another document?

Copy and paste only works in Windows. (Otherwise you have to use screen shots.) Click on the **New graphics window** icon at the bottom of your graph (you can hover over the icons to see what they do), then you will have another

copy of your graph in a separate graphics window. Right click on this new window, and **Copy as metafile** or **Copy as bitmap**.

How can I save the data file in a way that remembers the reordering I have done?

Going **File > Save ...** enables you to save the data in a R native format which remembers all new variables you have created including features like reorderings and renamings. You read it back in a new iNZight session using **File > Load ...**.

Remember, for *how to* instructions for iNZight you should consult iNZight's [User Guides](#). For *technical issues* with running iNZight, you should first consult iNZight's [frequently asked questions \(FAQ\)](#) page and if the answer you need is not there ask at [iNZight support](#).