2.5 Exercise: Categorical variables (iNZight Lite)

In order 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 Lite 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.

INSTRUCTIONS

Follow these instructions to generate the graphs. If you have a problem doing the exercise, see **Common questions** on page 7.

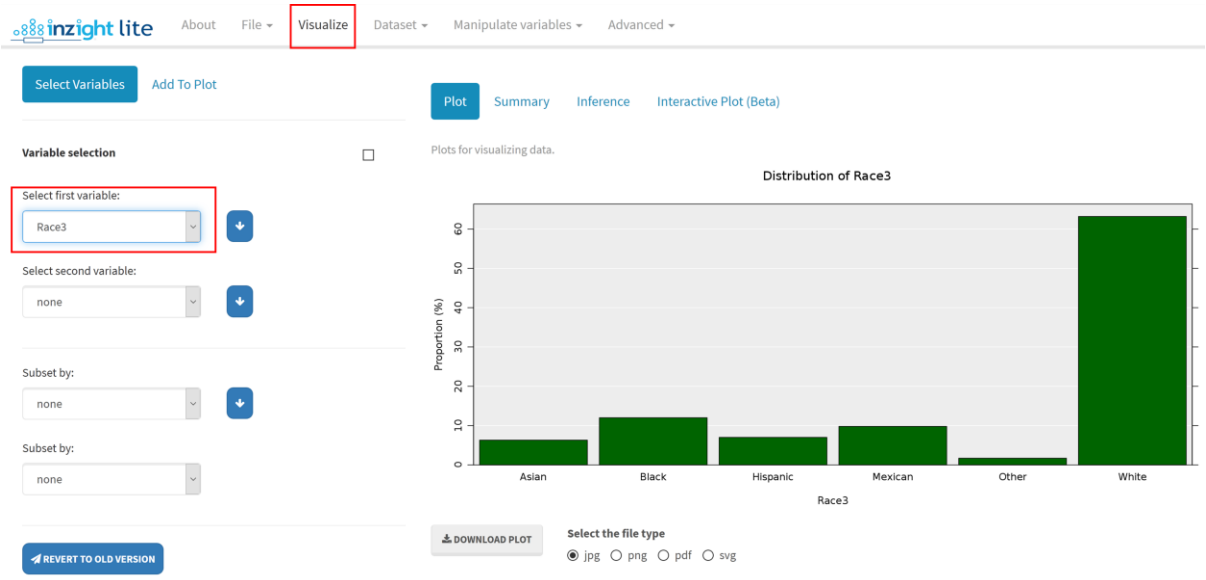
Import the **NHANES-1000** dataset into iNZight Lite:

- Select **File** > **Dataset Examples**
- Select Data set category: **Future-Learn**
- Select **NHANES-1000**
- Click on **Select Set.**

We will start by clicking on the top menu item **Visualize.**

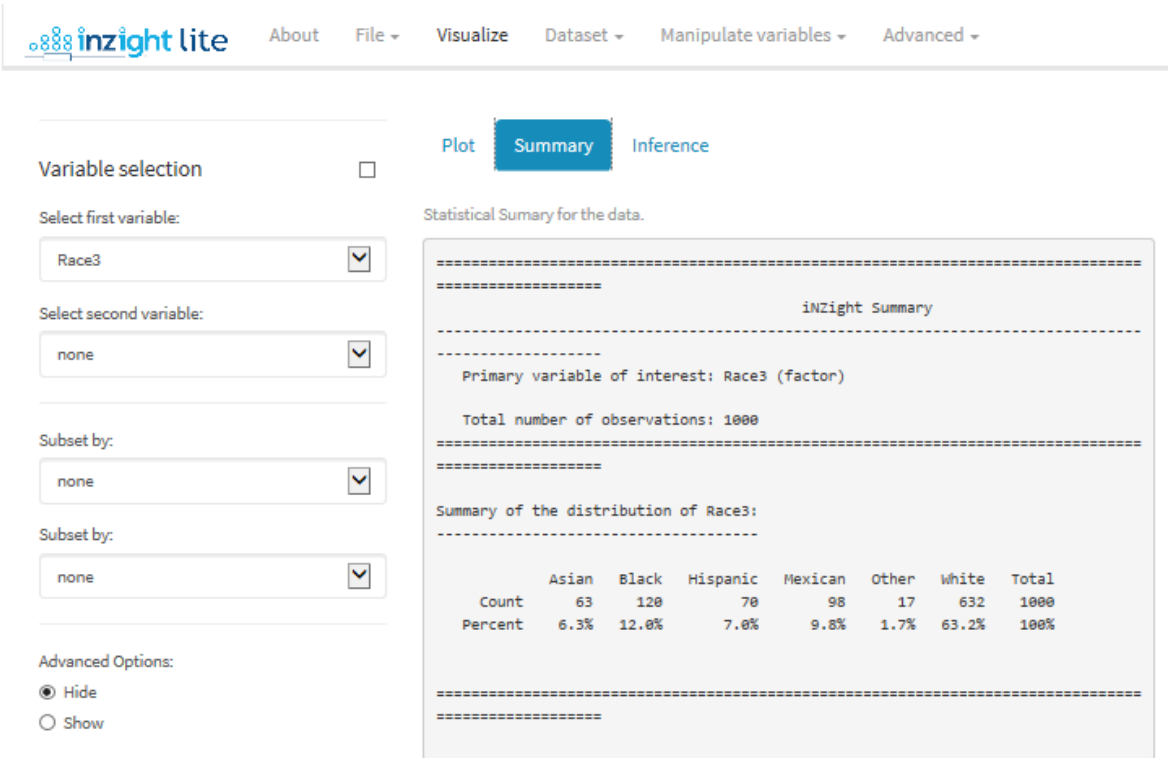
Making bar charts of categorical variables

Select the variable name **Race3** into the **Select first variable** slot. iNZight Lite will automatically produce a bar chart for a categorical variable when you click its name in the **Select first variable** slot.



Generate the Summary window

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



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 select 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.

Select **Add to Plot** and using the default **Customise Plot Appearance**

- under **Bar Colour** see **Colour by** and using that select **Race3.**
    - At this point ***Colour palette*** *options will appear* and allow you to change the colour scheme.
- From **Colour palette** select **Colourblind Friendly**.

If you have coloured by a variable, that colour-by instruction will remain in force until you colour by another variable or set **Colour by** to the blank entry at the top of the list.

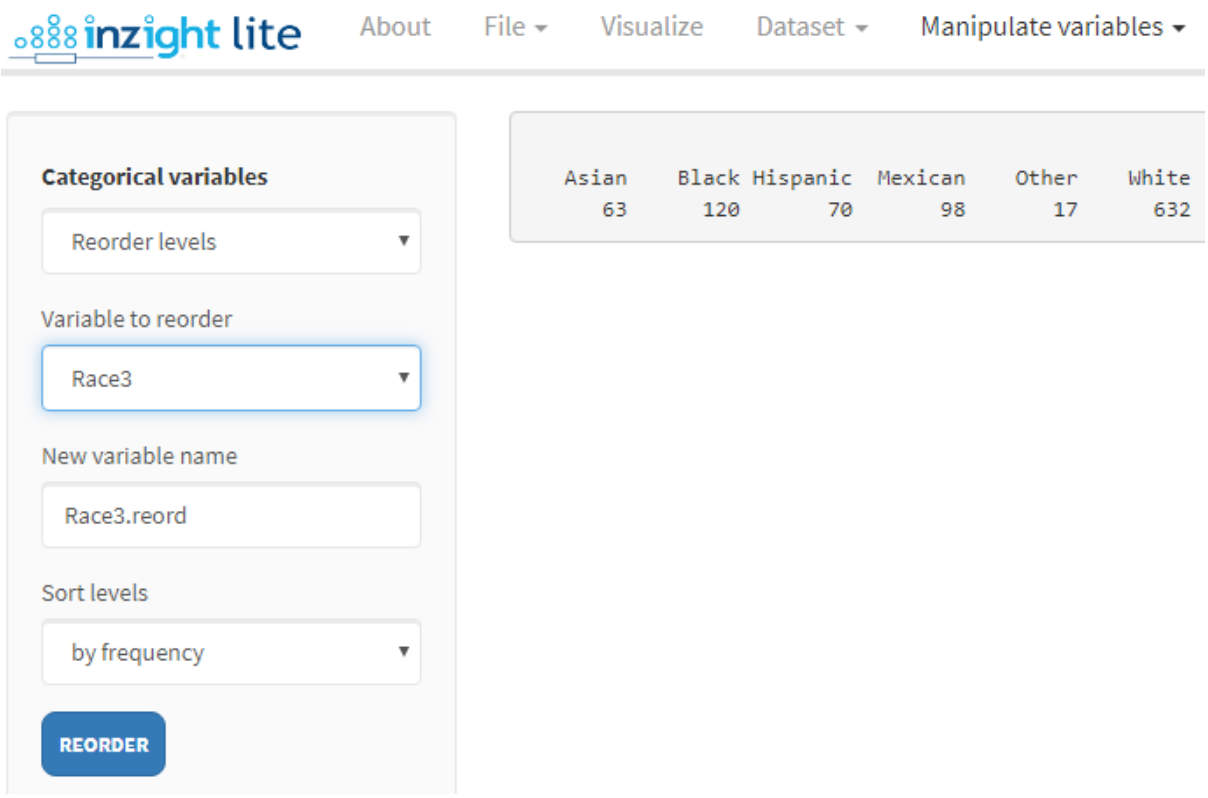
Reorder the bars by frequency

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:
Manipulate Variables > **Categorical Variables**

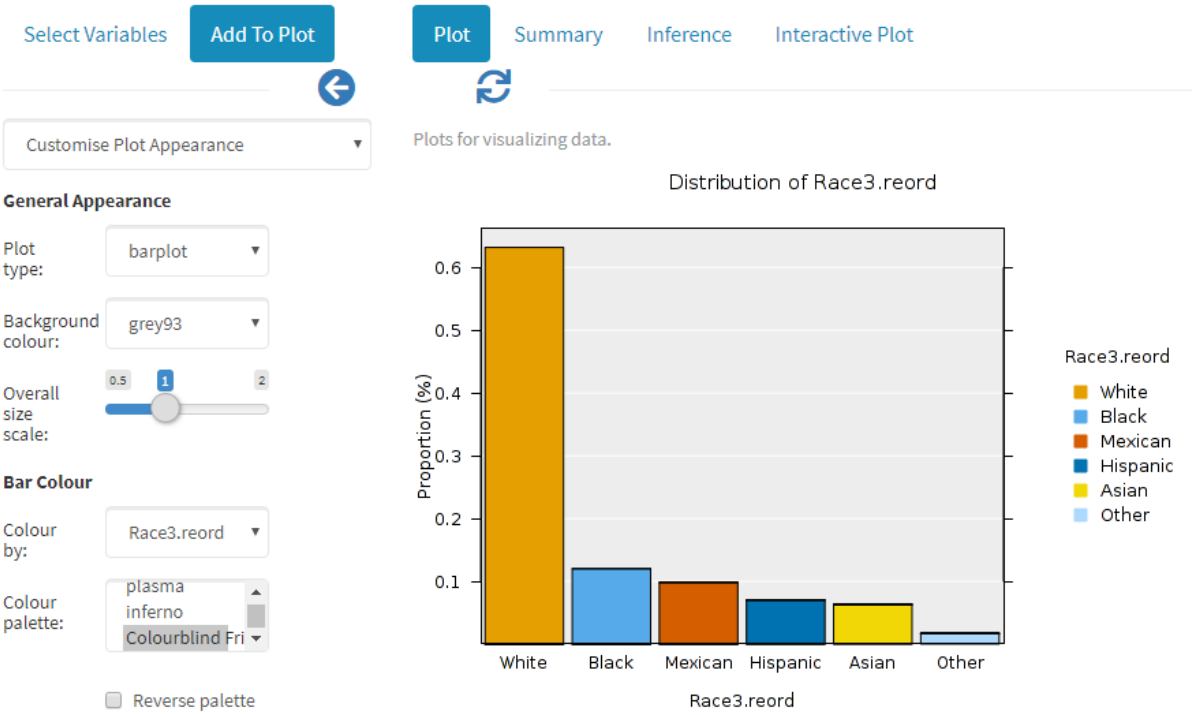
Under **Categorical Variables** select **Reorder Levels** (the default)

- Under Variable to reorder, select **Race3** from the drop down list
- A ***new variable*** will be created. Under **New variable name** you have the opportunity to give the variable a new name if you don't want to accept the name offered by default, **Race3.reord**
- Under **Sort levels**: select **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* return to **Visualize** to change the variable being plotted to the new one.

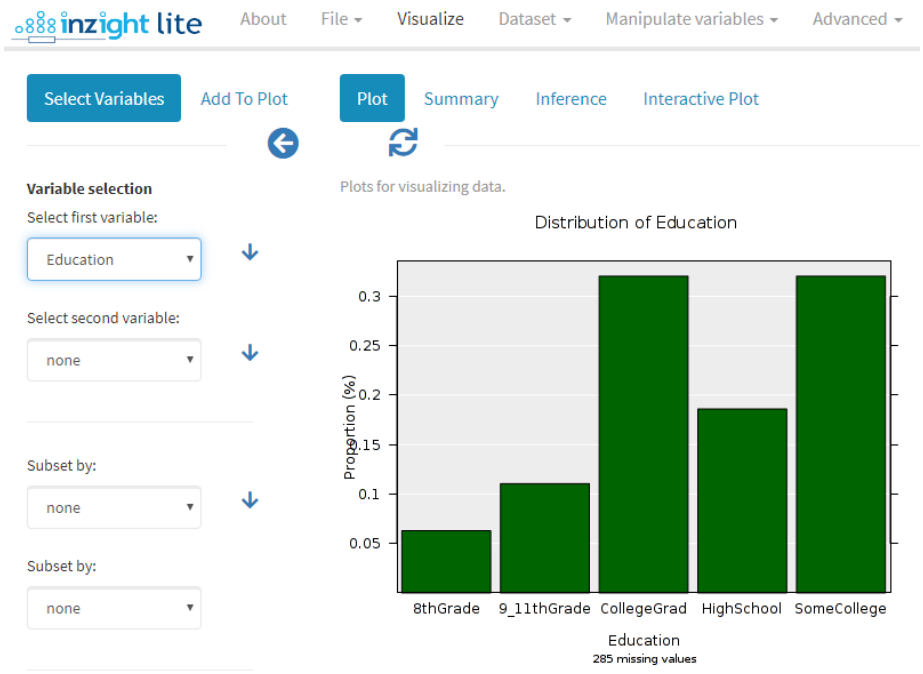
- Now select **Race3.reord** (or your new variable name) in the **Variable1** slot.
- Recolour by **Race3.reord**.



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

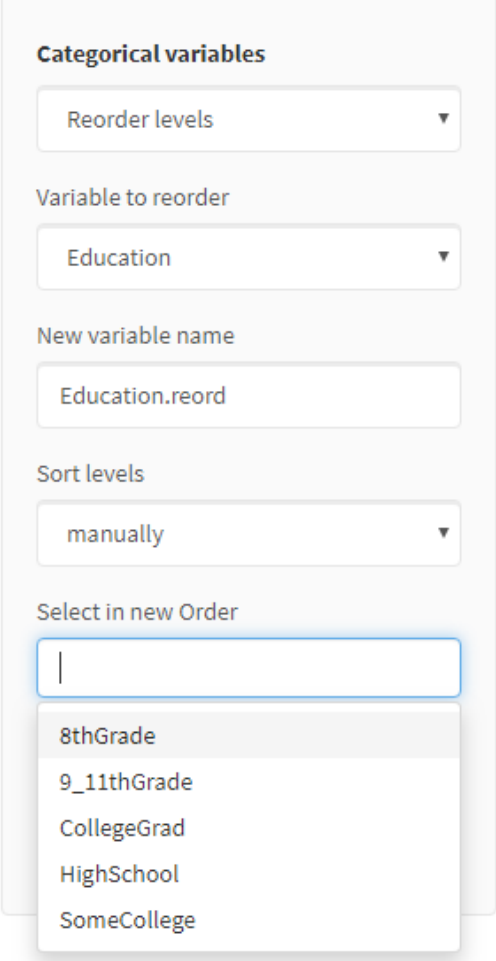
Manually reorder the bars

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, i.e., 8thGrade, 9-11thGrade, HighSchool, SomeCollege, CollegeGrad, but currently they are not.

To reorder Education:



Proceed as you did with **Race3** but this time …

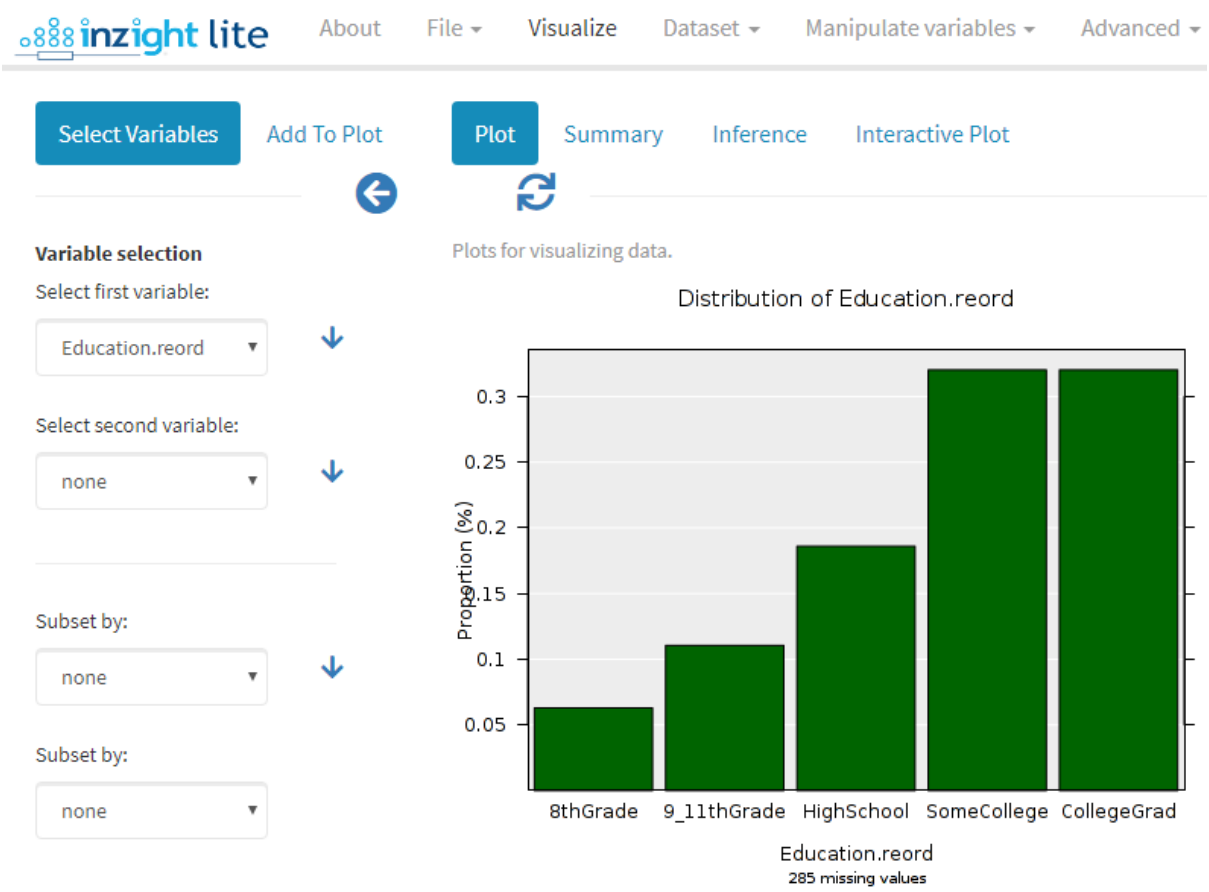
- Under **Sort levels**: select **manually**

When you click on **Select in new Order** it will show you the levels/categories. **Click** on them in the order you want.

They will enter the box in that order. If you make mistakes you can remove some of them from the box with the backspace key.

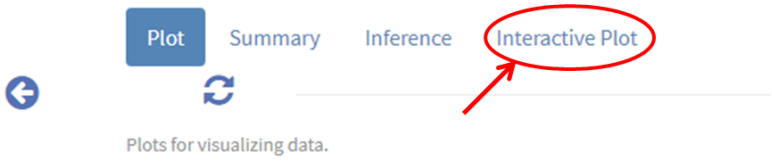
When you are happy, click **REORDER** to create a new variable with the desired ordering (here it will be called **Education.reord**).

Go back to **Visualize** and plot **Education.reord**



Optional:

Try this new feature (interactive plots)



Click on the **Interactive Plot** tab. This will give you an interactive version of your graph that lets you query it in various ways like hovering over bars or clicking them. Explore!

You can download these plots as Interactive HTML files which you can give to others. They do not need to be connected to iNZight lite to work.

Other ways of representing data on a categorical variable

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

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 with current settings** icon (underneath the **plot** window) to redraw the plot.

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

Use Add **to Plot > Customize Labels**.

I can't get rid of the segmentation in the bars. How do I do this?

Set **Colour by** to the blank entry at the top of the list

The reordered graph did not appear, it was the same graph unordered.

Click on a different variable and then click back on the variable you wanted the ordered bar graph for (there is sometime a small glitch in the online version of the software and you need to do this).