

## 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.

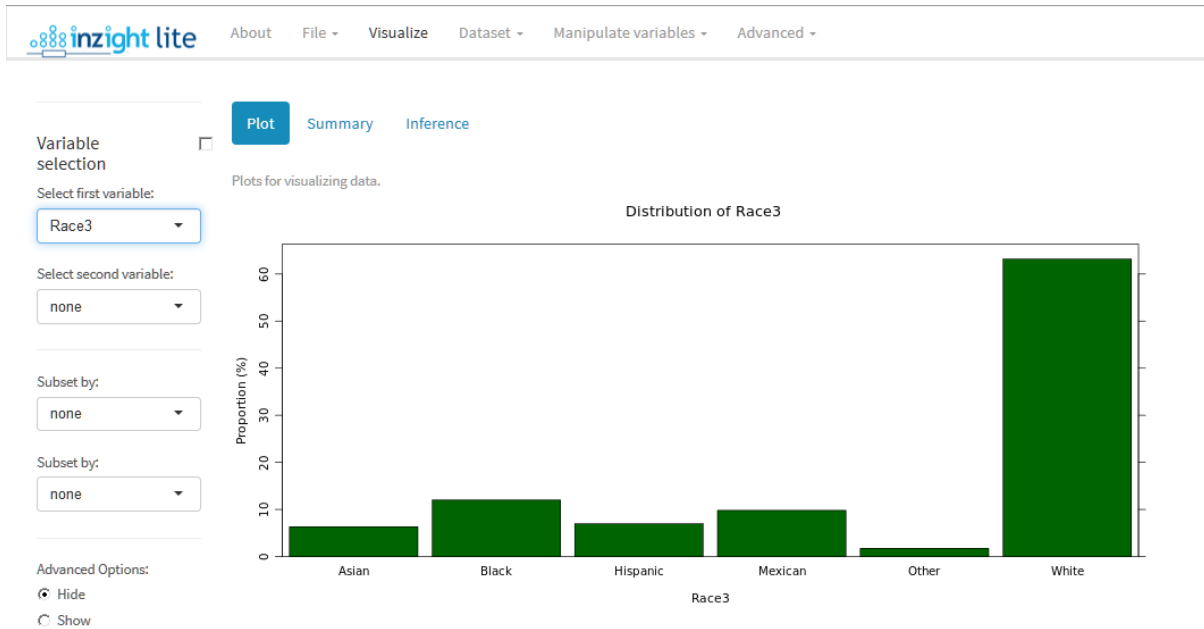
Import the **NHANES-1000** dataset into iNZight Lite:

- 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**.

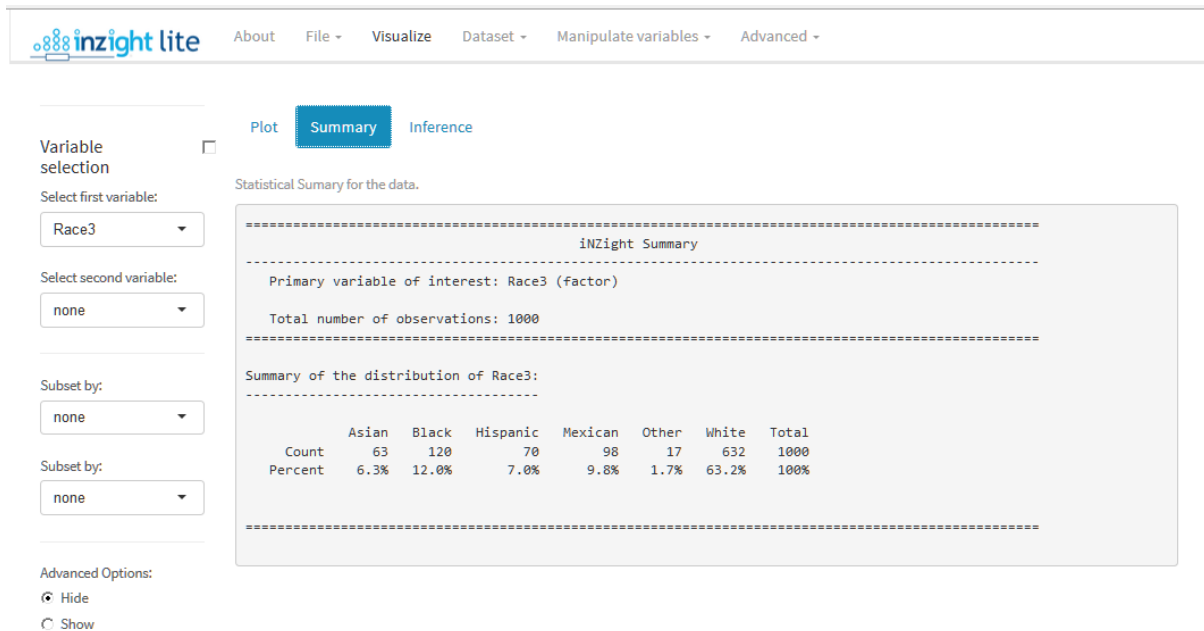
## Making bar charts of categorical variables

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.



The screenshot shows the inZight lite software interface. The top navigation bar includes the logo and menu items: About, File, Visualize, Dataset, Manipulate variables, and Advanced. The main interface has three tabs: Plot, Summary (selected), and Inference. On the left, there are controls for variable selection: 'Select first variable:' is set to 'Race3', 'Select second variable:' is set to 'none', and two 'Subset by:' dropdowns are also set to 'none'. Below these are 'Advanced Options' with radio buttons for 'Hide' (selected) and 'Show'. The main content area displays a 'Statistical Summary for the data.' window. This window has a title 'inZight Summary' and contains the following text: 'Primary variable of interest: Race3 (factor)', 'Total number of observations: 1000', and 'Summary of the distribution of Race3:'. Below this is a table showing the distribution of counts and percentages for each race category.

	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)

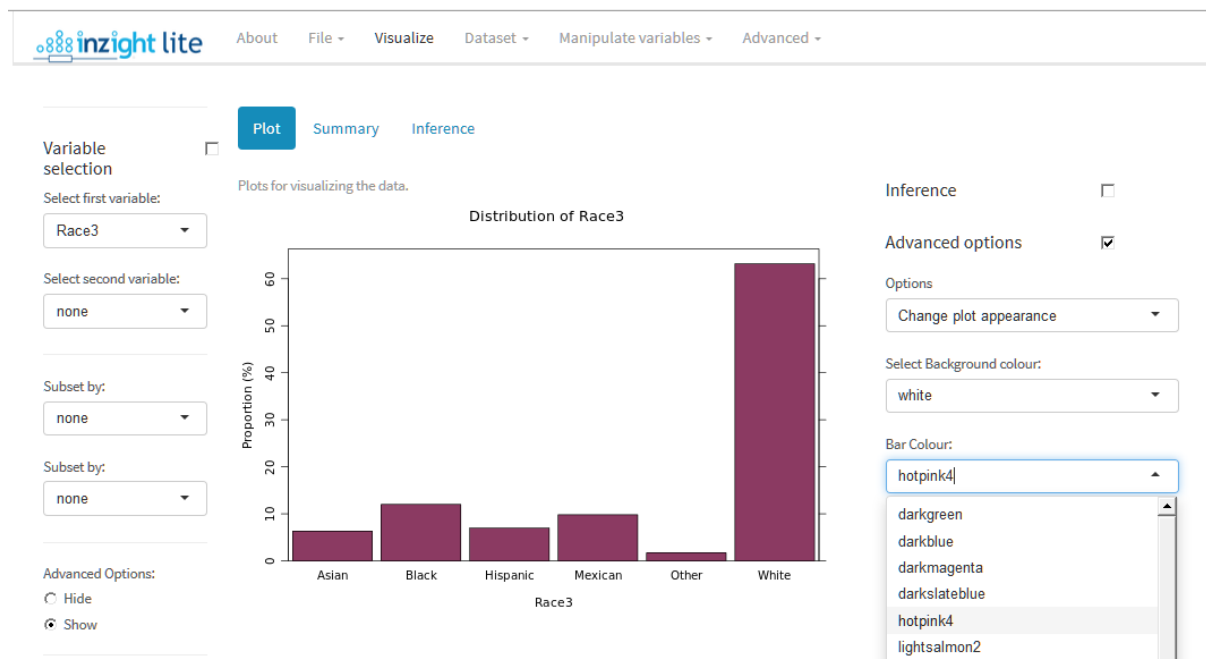
Choose another **categorical variable** from the dataset and plot it. Try a couple more and do this until you are confident that you can get simple graphs and tables like these quickly.

[Note: When you 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.

At the bottom left is the advanced options selector button. Click on **Show** then select your colour from the drop down on Bar Colour. The example shows the results for **hot pink**.

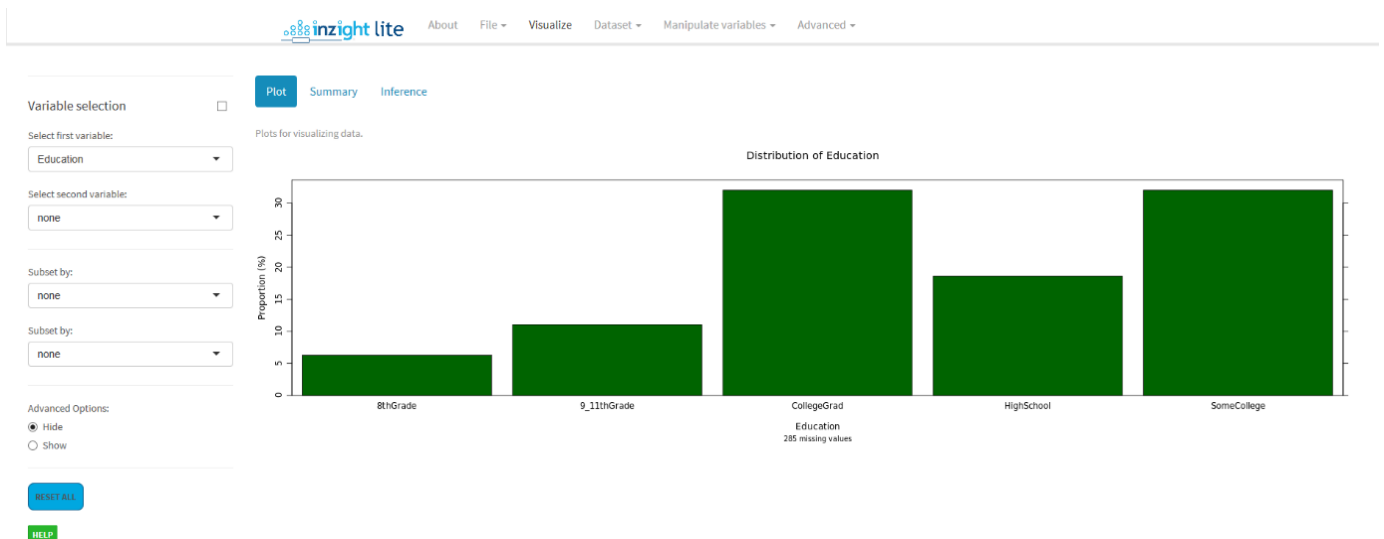


If you have coloured by a variable, that colour-by instruction will remain in force until you colour by another variable.

## 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.



### > Manipulate Variables

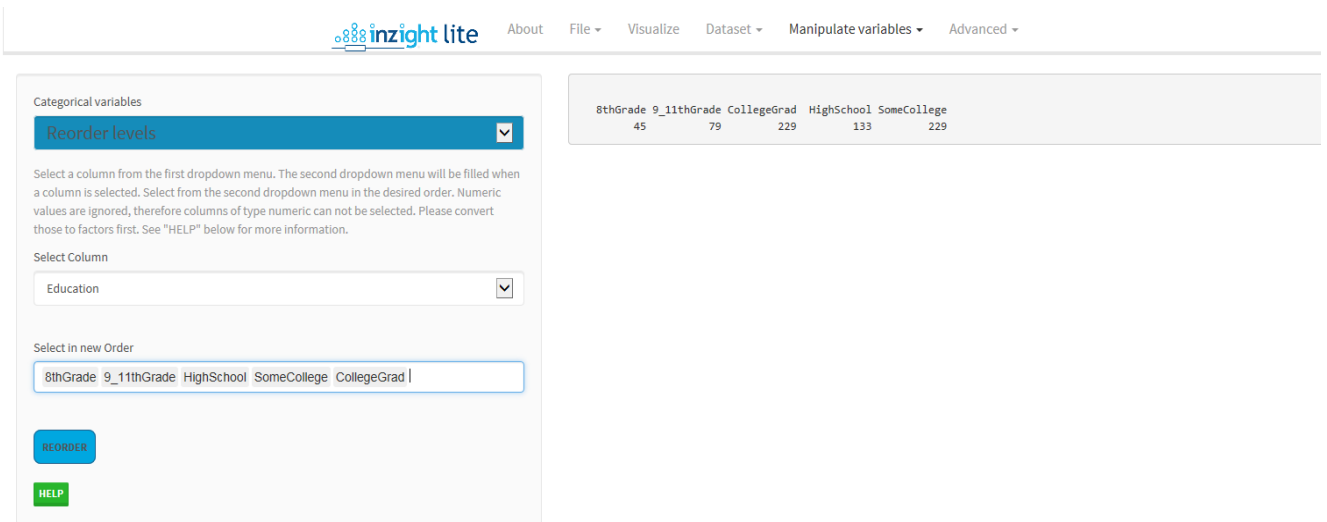
To reorder **Education**:

From the top menu, select

### > Manipulate Variables > Categorical Variables

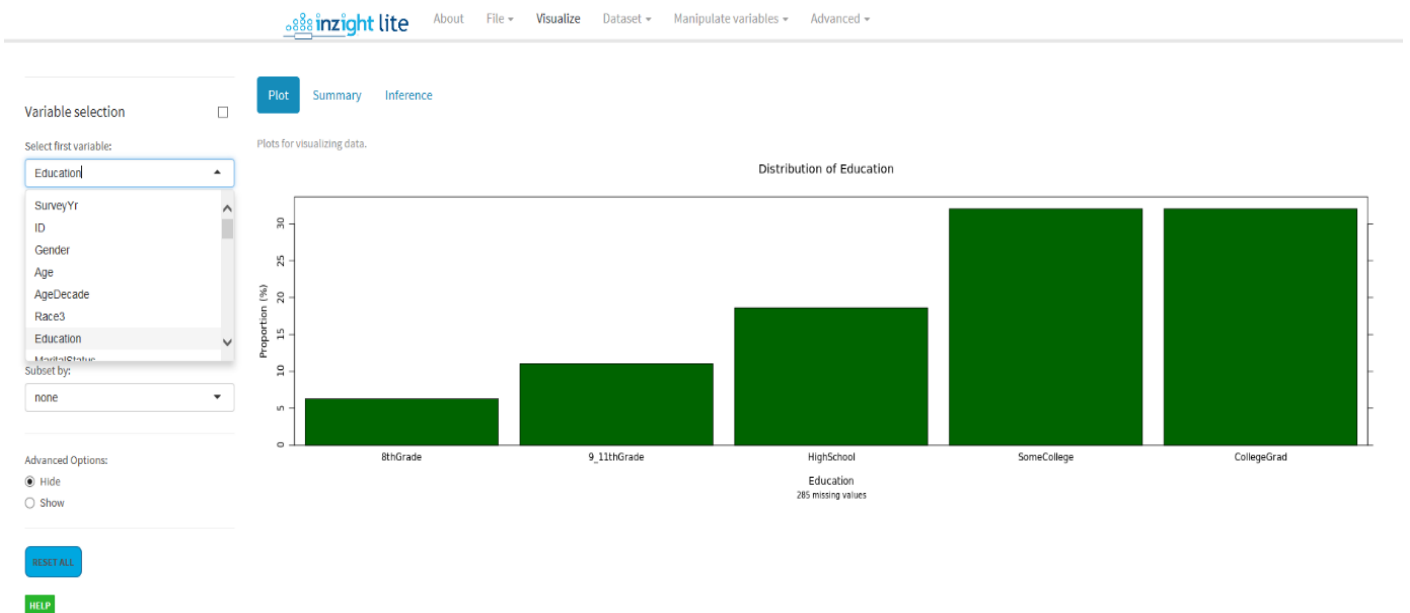
- In the first drop menu select: **Reorder Levels** (this should be set to this automatically)
- then, in the next drop down menu **Select Column**, select: **Education**
- Click on the variable names one at a time, in the order you would like them to be shown on the graph.

- then click: **Reorder** (Then go back to visualize and select the variable Education, this will update your preferences)



8thGrade	9_11thGrade	CollegeGrad	HighSchool	SomeCollege
45	79	229	133	229

If your graph does not show the new order once you select Visualize, then click on another variable and go back to your variable and the graph will update – it is meant to update immediately but there is a small glitch there.



Education	Proportion (%)
8thGrade	~15
9_11thGrade	~22
HighSchool	~28
SomeCollege	~32
CollegeGrad	~32

## 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?**

Click "Remove Additions" underneath the plot. There is an option called "Remove segmentation in bar charts".

**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).