



# Weeding Your Data



## Will it be used?

Not all data will be used in a Power BI report. For example, will a customer's fax number be used in a visual report?

## Can it be simplified?

Do you need everything in the column? For example, if an order number *always* has the same letters in front of it (e.g. SO1234) you can probably safely delete the "SO" component.

## Is there missing data?

More importantly, can you fill in the missing values? If you can't, it is probably a good idea to weed the column, because you don't want a lot of missing values. It will throw off how your visuals will display.

## Can it be combined?

Sometimes there are columns that can be combined together in order to simplify the data. Make sure, though, that combining two columns does not create a high-cardinality column, especially if the original columns were low (or lower) cardinality.

## What type of report are you creating with the data?

Before weeding your data, determine if the report you are creating will be "visual" (meaning composed primarily of visuals) or "paginated" (meaning composed of rows of data). Some data that is suitable for a paginated report will not be suitable for a visual report and vice versa.



**Weeding data is reversible,  
so be ruthless!**

