Data cleaning (Excel) - Homework 1

Excel File - Marriage Rates

Brenden Arias

This was a good introduction back into excel and a good way to show you how tedious wrangling data manually can be. However, you do have a strong understanding of the data when you look at almost every entry. You can see outliers and easily find null values or possible data entry errors. The visualization is very interesting because it shows a spike in some states that until you think about it don’t make sense. However, the spike of marriage rates in Nevada is probably an outlier because of all of the “shotgun” weddings that happen there.

The cleaning process is summarized below.

* Deleted top three rows of header meta data information.
* Deleted empty row between header and data
* Deleted bottom 4 rows of meta data in the footer
* Determined best way to convert data from a wide format to long format, using columns, State Year and Marriage Rate
* Then transposed the following data to convert it to a long format