**Data management** (note: replace highlighted values as needed)

***Caution:*** Be aware that Excel likes to replace missing values (a blank cell) with zeros when copying values through formulas. This can corrupt your statistical results. See Process #4.

**Process #1. Deleting rows based on variable values**

* Click in a random cell that isn’t blank
* Under the Data tab of the ribbon, click the Filter button
* Click the tab at the top of the “student” column, and select only “(Blanks)”
* When you see only the data with blank values for “student”, click and drag across the row numbers to select all of the rows shown except for row 1 (the row with variable names)
* Right click, and select “Delete Row”
* Click the Filter button again to see all your (remaining) data

**Process #2. Calculations**: To calculate approximate age from birthyear: =2023-K2

* For missing values, follow the Process #4, but find-and-replace “2023” with blank

**Process #3. Merging variables**: To merge the two versions of the pro-choice (T2) /abortion rights (U2) question: =IF(ISBLANK(T2),U2,T2)

* For missing values, follow the Process #4

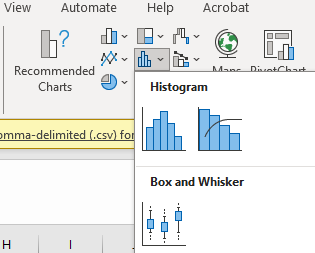
**Process** A screenshot of a computer

Description automatically generated**#4. Recovering missing values**

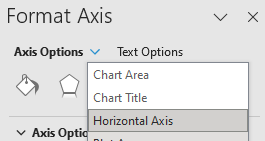
* First, you need to replace the column values with actual numbers rather than formulas. To do so, copy-paste the contents of the column to replace itself, but select to only “paste values” (see image on the right)
* Then, select the column (by clicking on the letter at the top of the column, just above the variable name), and use a find-and-replace (Ctrl-H), replacing the 0s with blanks (note: this won’t help if there are actual 0 values for the variable)

**Creating histograms or box-and-whisker plots for continuous variables**

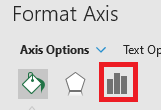
* Select the column you’re describing (e.g., click on the letter just above the variable name)
* Select the statistical charts and choose either Histogram or Box and Whisker



* Click on the graph
* Under the Format tab on the top ribbon, click Format Selection (all the way to the left)
  + For histograms, click to edit the Horizontal Axis:



* + For box and whisker plots, click to edit the Vertical Axis instead
* Select the button that looks like a bar chart:



* Make adjustments to settings, as desired:
  + For histograms, adjust the bin widths
  + For box and whisker plots, under “Axis Options,” adjust the minimum and maximum

**Descriptive statistics for continuous variables**

***Before you start,*** make sure you install the Analysis toolpak (instructions on how to do so here: <https://support.microsoft.com/en-us/office/load-the-analysis-toolpak-in-excel-6a63e598-cd6d-42e3-9317-6b40ba1a66b4>)

* To check if it’s installed, click Data on the ribbon, and then see if there is a button for Data Analysis on the right:

A screenshot of a computer

Description automatically generated

**Get the descriptive statistics**

* Under the Data tab, click “Data Analysis”
* Select “Descriptive Statistics” and click OK
* Click inside the box for “Input Range,” then click to select the column you’re describing
* Make sure the box “Labels in first row” is selected
* For “Output options,” I recommend “New Worksheet Ply”
* Make sure the box “Summary statistics” is selected

**Frequency tables + Histogram for Qualitative Variables**

**1. Make the initial pivot table**

* Make sure your variable has a name entered in the first row
* Under the Insert tab, click “PivotTable”
* Select the column you want
* Click inside the box for “Table/Range,” then click to select the column you’re describing
* For choice of where you want the PivotTable, I recommend “New Worksheet”
* On the right, click-and-drag your variable name down to the box called “Rows”

A screenshot of a computer

Description automatically generated

* Now, click-and-drag your variable name down to the box called “Values”
* In the Values box, click the small triangle to select, then choose “Value Field Settings”
* Select the “Count” option, then click OK

**2. Make the histogram**

* Click any cell inside the pivot table
* Under the Insert tab, click to add a 2-D Column bar chart (Clicking “Recommend Charts” may give you this as the default option)

**3. Add percentages to the pivot table**

* To add percentages, click-and-drag your variable name down to the box called “Values” again (the variable needs to be listed twice under “Values”)
* In the Values box, click the small “select” triangle for this second listing of your variable, and choose “Value Field Settings “
* Select the “Count” option, and then click on the “Show Values As” tab
* Select “% of Column Total”, then click OK