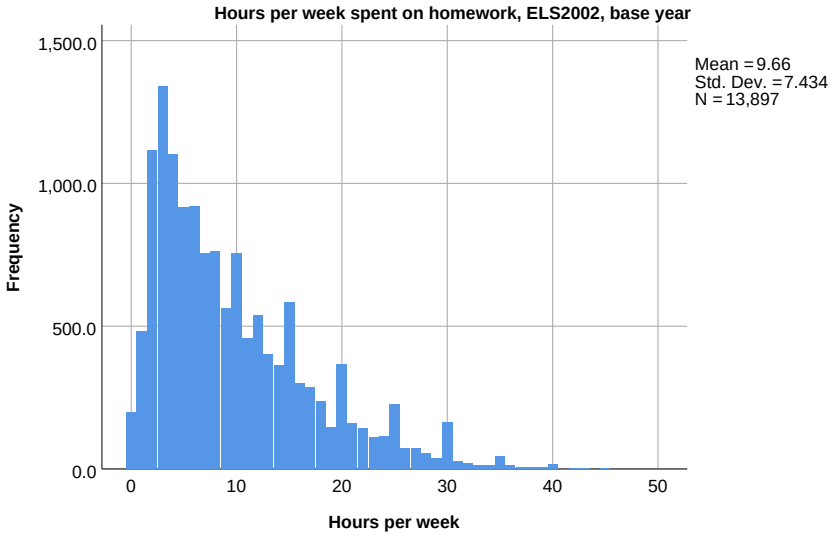
Histograms

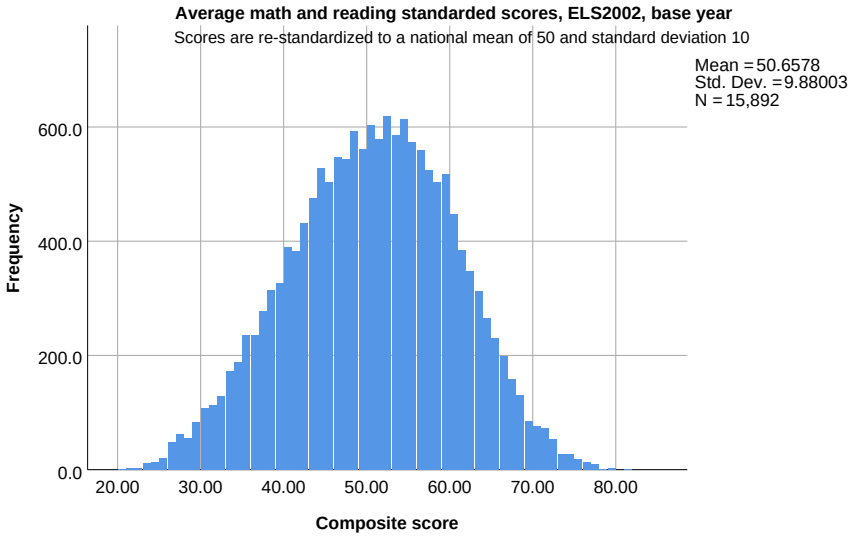
Histograms

- We are in sections 2.2, 2.6 of the textbook
- *Histograms* display quantitative data. Data values are separated into continuous classes (bins, bars), typically of the same width. The length of each bar is the frequency or relative frequency of the class.
- A histogram with a single mound is *unimodal*, two mounds is *bimodal*.
  - The mode(s) occur at the peak(s).
- A histogram is *left-skewed* if it has a longer tail on the left side.
  - The long left tail will pull the mean value to the left of the median.
- A histogram is *right-skewed* if it has a longer tail on the right.
  - The long right tail will pull the mean value to the right of the median.
- A histogram is *symmetric* if the two sides are similar.
  - The mean and the median will be the same.
- http://mathcenter.oxford.emory.edu/site/math117/shapeCenterAndSpread/

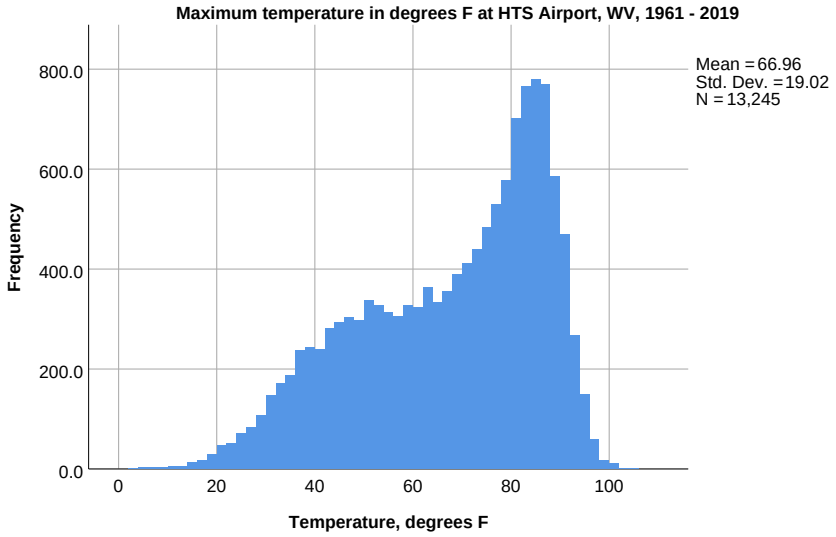
1. For each histogram shown,
   (a) identify the data as symmetrical and mound-shaped, symmetrical and rectangular (uniform), skewed to the left, skewed to the right, bimodal, or other;
   (b) estimate the location of the mode, median, and mean of the data.

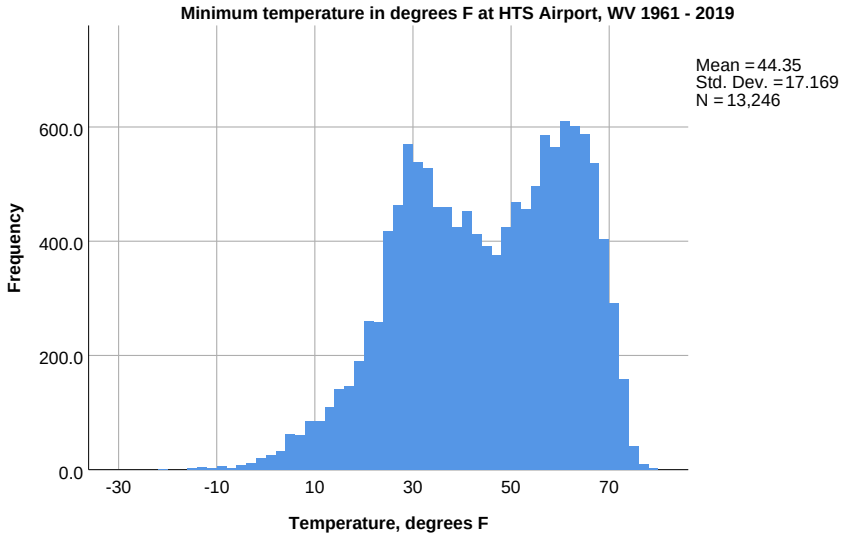
ELS:2002 features a nationally representative, longitudinal study of 10th graders in 2002 and 12th graders in 2004. The study consisted of surveys of students, their parents, math and English teachers, and school administrators. Data was collected in 2002 (base year), 2004 (first follow-up), 2006 (second follow-up), 2012 (third follow-up).

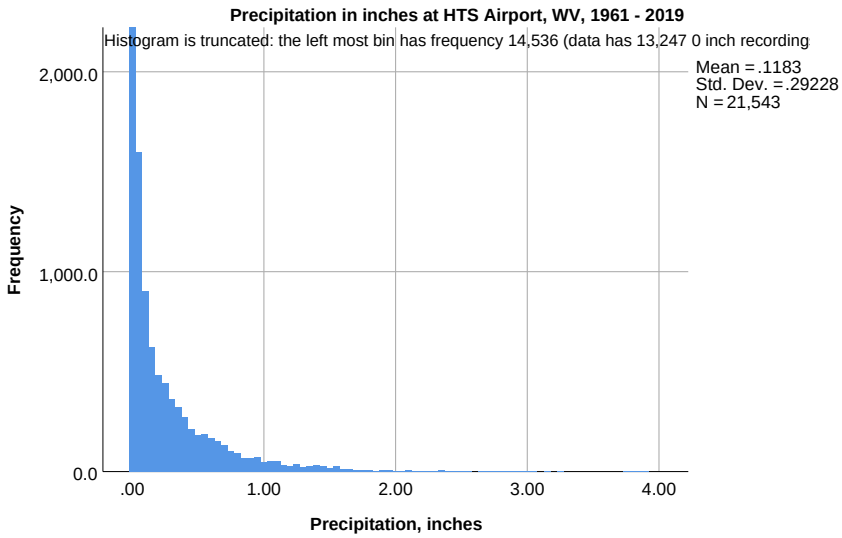




Data on precipitation and temperature from the Huntington Tri-state Airport, WV, Station USW00003860, from 1961 to 2019 is available from NOAA. This is the same data you will use in your STA 150L lab class to calculate a new column (average temperature), filter data, and find basic descriptive statistics.







2. Precipitation and temperature are displayed using a histogram. Explain why it makes sense to use a histogram for these variables. Explain your reasoning in two to three sentences.

3. A data set has a sample mean of 42 and sample median of 65. What would be the likely shape of the distribution of the data?

4. A set of data has sample mean of 55 and sample standard deviation of 15. If 5 was added to each observation, what would be the mean and standard deviation of the new data set?