The visualizations in this section help you answer questions like these:

 What is the peak value of the distribution?

 How many peaks are there in the distribution (unimodality versus bimodality)?

 How normal (or lognormal) is the data? We’ll discuss normal and lognormal

distributions in appendix B.

 How much does the data vary? Is it concentrated in a certain interval or in a certain

category?

many summary statistics assume that the data is

approximately normal in distribution (at least for continuous variables), so you want

to verify whether this is the case.

**Data Preparation**

**Data Visualization**

**Data Splitting**