1.5.4   Visualizing Distributions with Box Plots

Louise mentioned earlier that she's also interested in Great Britain's theater market, especially musicals. While she's committed to creating a play in the U.S., she's also interested in researching musicals in Great Britain for a future project with an estimated budget of £4,000. To present Louise with the big picture, we will create box plots using statistical computations.

Taken together, measures of central tendency, measures of spread, and identification of outliers tell us about the distribution of our data, but *showing* the distribution makes the data story easier to understand—which has an even greater impact. **Box plots** are an effective way to show large amounts of information about a distribution in a small amount of space.

Box Plots

We'll use **box plots**, also called **box and whisker plots**, to compare the distribution of campaign goals and the distribution of total amounts pledged for plays in Great Britain. Remember that Louise estimates she'll need to raise £4,000 for her future project.

First, clear all the filters from your data. Set a filter on "country" to only show rows with GB as the country. Then, set a filter on "Subcategory" to only show rows with "musical" as the subcategory.

Select the Goal and Pledged columns and then go to the Insert tab. Click the Statistical Charts button and select "Box and Whisker."
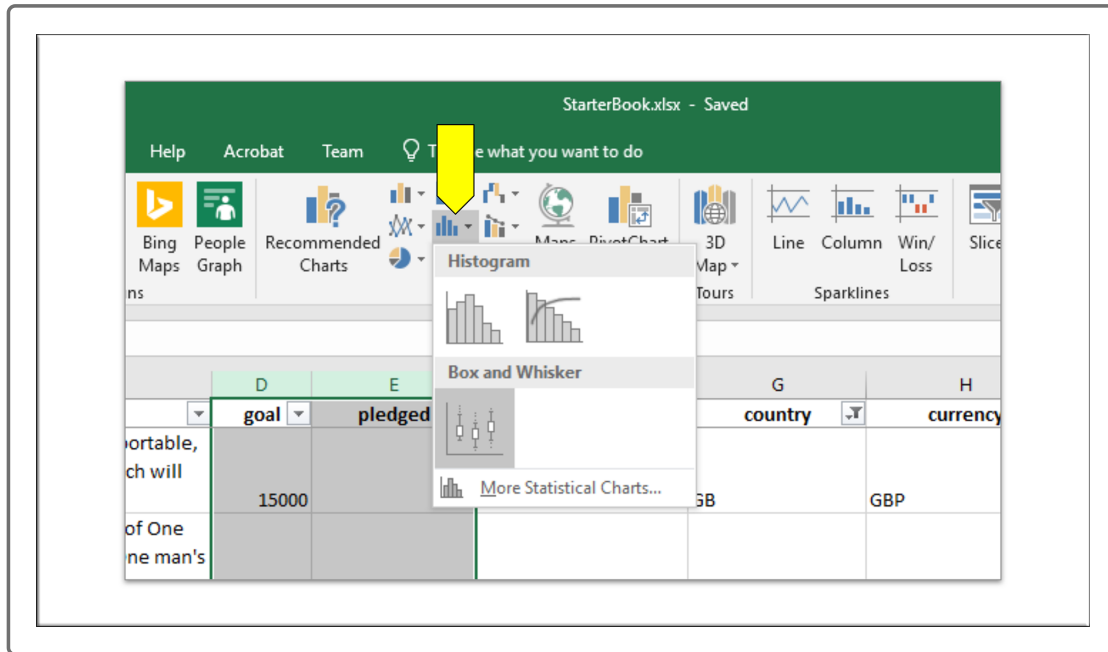
Now move the chart to its own sheet by going to the Design tab and selecting the Move Chart button.
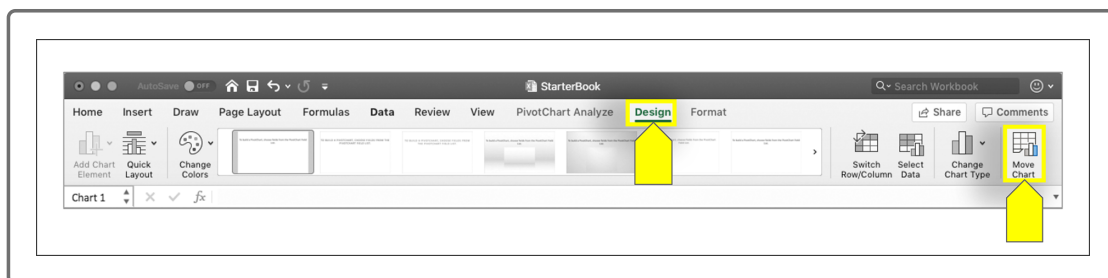
How do we read a box and whisker plot? The box shows the interquartile range with a line for the median and an "X" to indicate the mean. The whiskers show the extreme values within 1.5 times the interquartile range. Outliers are represented by labeled dots.
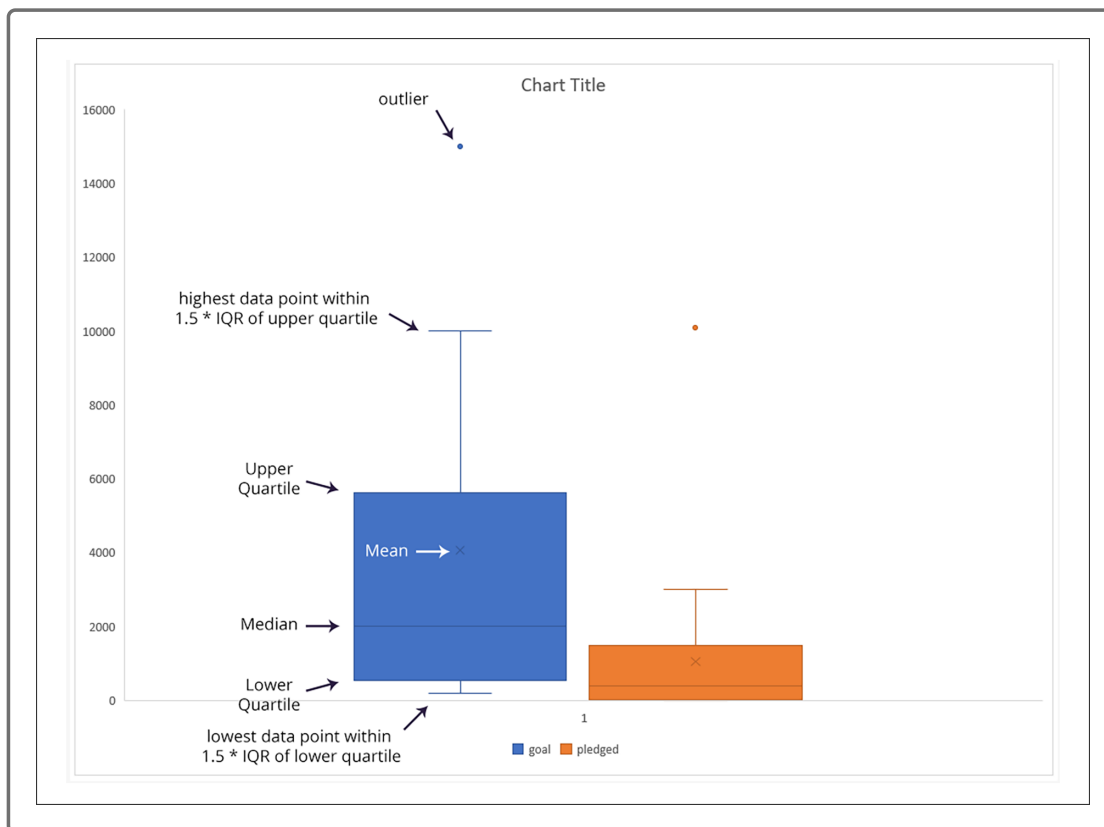
From these plots, we can see that the mean campaign goal is around £4,000. This is outside of the range of outliers for amount pledged, so Louise should probably try to get her play produced for less than £4,000. Half of the campaign goals are less than £2,000, which is just over the 3rd quartile for amounts pledged.

↻ Retake

↻ Retake