0.0.1   Question 2biv:

We analyzed the data but something seems odd. On a closer look, there are many negative values for `profit`. For example, the movie `102 Dalmations` looks to have lost ~$18M but it was a widely successful film! What may account for this issue?

There are many possible reasons why it is happening. One possible I can think of is that when we deal with a large dataset, we may encounter errors in the system when mapping data. We sometimes may lose data or get incorrect data. An extra 0 could have been added or a decimal could have been pushed in the budget value. Another possible reason could be how we contrainted the data. We only selected the valeus from US excluding all other region. The profit might come from the regions outside of the US. Those are the possible cause for this issue.

0.0.2 Question 3c

What do you notice about the summary values generated in `earnings_summary`? Identify two properties about the histogram of the data.

Hint: Think in terms of about concepts from statistics like spread, modality, skew, etc. and how they may apply here.

The fact that median is greater than mean means that the histogram is left-skewed. The histogram is unimodal based on the five number summary. Also, the fact that maximum value is far greater than 75-percentile's value might indicate the possible outlier in our data.