

Outliers and Robust Measures

3 min

For our income data, the difference between the mean (\$34,795) and median (\$32,978) was only about \$2,000. You may be wondering: Is the difference ever larger?

Let's imagine some very famous celebrity musicians have all decided to move to Melody Metropolis. We know celebrities make much more money than the typical musician in our dataset. In the learning environment, we've added three new incomes to the distribution:

- **\$48 million:** Paul McCartney, British musician of the Beatles
- **\$57 million:** BTS, South Korean K-pop band
- **\$81 million:** Beyoncé, American singer-songwriter

The second plot shows that the median appears almost unaffected by the addition of these three gigantic incomes: the median moves from \$32,978 to \$33,011. However, the mean makes a drastic change from \$34,795 to \$228,235. The mean is now well beyond even the maximum in the original distribution. An income of \$228,235 is definitely not a great measure of the center of our income distribution.

These celebrity incomes are examples of **outliers**, extreme values that are distant from the rest of the distribution. Just as with skewness, outliers tend to more heavily influence the mean than the median. This same pattern occurs with measures of spread: the standard deviation is more influenced by outliers and skewness than the interquartile range (IQR).

Because the median and IQR are NOT heavily influenced by extreme values, we say they are **robust**. Robust

Preview: Docs Statistics is the science that is concerned with methods for collecting, organizing, analyzing, and interpreting data.

[statistics](#)

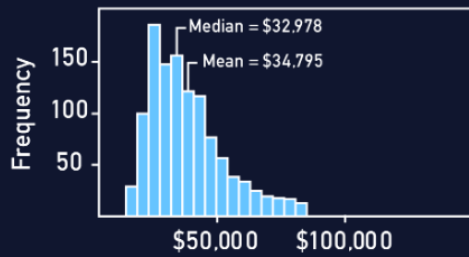
are often a better choice to measure the center and spread of a distribution that is skewed or has outliers.

Source:

[The World's Top-Earning Musicians Of 2019](#)

Income Distribution

Current Population



If Paul McCartney, BTS, and Beyoncé move to the city

