**The 30 Richest Americans**

The following table shows the net worth of the 30 richest Americans (according to *Forbes*, 2012)[[1]](#footnote-1).

|  |  |
| --- | --- |
| **Name** | **Worth (in billions)** |
| Bill Gates | 66 |
| Warren Buffett | 46 |
| Lawrence Ellison | 41 |
| Christy Walton and family | 27.9 |
| Charles Koch | 31 |
| David Koch | 31 |
| Jim Walton | 26 |
| Alice Walton | 26.3 |
| S. Robson Walton | 26.1 |
| Michael Bloomberg | 25 |
| Larry Page | 20.3 |
| Jeff Bezos | 23.2 |
| Sergey Brin | 20.3 |
| Sheldon Adelson | 20.5 |
| George Soros | 19 |
| Forrest Mars, Jr. | 17 |
| Jacqueline Mars | 17 |
| John Mars | 17 |
| Steve Ballmer | 15.9 |
| Paul Allen | 15 |
| Carl Icahn | 14.8 |
| Michael Dell | 14.6 |
| Phil Knight | 13.1 |
| Donald Bren | 13 |
| Len Blavatnik | 12.5 |
| Ronald Perelman | 12 |
| Abigail Johnson | 11.8 |
| John Paulson | 11 |
| Laurene Powell Jobs and family | 11 |
| James Simons | 11 |

**Discuss the Following Questions**

* Enter the data in your statistical software of your choice (e.g., R, SAS, StatCrunch).

1. Looking at the data, do you believe that there are any outliers?
2. Find the mean, median, and standard deviation for the worth variable. Write these values down.
3. Interpret the value of the standard deviation.

Let’s see what happens to our measures of center when you remove the outlier(s).

Example statistical software instructions for removing outlier(s):

* **SPSS:** To remove an outlier, click on the number on the left-hand side of the screen corresponding to the row you want to delete. When you do this, the row will become highlighted. Then, go to Edit—Cut. Repeat this for all outlying values you wish to remove. This removes those values from the data set.
* **StatCrunch:** To remove an outlier, click on the number of the row you want to delete (on the far left side of the spreadsheet) so that the entire row becomes highlighted. Then, go to Edit—Rows—Delete. You will be prompted to see if you really want to delete the row, and you should select Delete Rows!
* **R**: To remove an outlier, go back to the CSV file you saved and omit the row(s) with the outlying values. Re-save the data (perhaps by giving the file a new name so you don’t lose the original data) and then read it again into R.

Once you have removed the outlier(s), find the new mean, median, and standard deviation. Write those values down. (NOTE: If this was important data, a good practice would be to save the original data file and then make a copy of it where you are free to delete things that you want to without losing the original data, in case you need it later.)

1. How do the mean and median change when you remove the outlier(s)?
2. How does the standard deviation change when you remove the outlier(s)?
3. Which measure of center (the mean or median) is *resistant*? Why?

1. Source: <http://www.forbes.com/forbes-400/list/> [↑](#footnote-ref-1)