**Section 2.4 and 2.5** **Describing Variability and Spread**

**Example 1.** According to a recent report from the US National Center for Health Statistics, females between 25 and 34 years of age have a bell-shaped distribution for height with a mean of 65 inches and standard deviation of 3.5 inches. Give an interval within which about 95% of the heights fall. Draw a labeled picture of a bell curve that illustrates this.

## PRACTICE! (From Online HW 03)

Height has an approximately bell-shaped distribution. For a sample of heights of college students collected, the males had $\bar{x} = 71$ and $s = 4$ and the females had $\bar{x} = 67$ and $s = 4$.

Use this information to answer parts a and b.

**a.** Use the empirical rule to describe the distribution of heights for females.

Approximately 68% of the observations fall within the interval (_____, _____).
(Type a whole number.)

Approximately 95% of the observations fall within the interval (_____, _____).
(Type a whole number.)

All or nearly all of the observations fall within the interval (_____, _____).
(Type a whole number.)

The following table summarizes responses of 4382 subjects in a recent survey to the question, "Within the past month, how many people have you known personally that were victims of homicide?"

Use this information to answer parts a through c.

| Number of Victims | Frequency |
|---|---|
| 0 | 3938 |
| 1 | 287 |
| 2 | 90 |
| 3 | 45 |
| 4 or more | 22 |
| **Total** | **4382** |

**a.** To find the mean, it is necessary to give a score to the "4 or more" category. Find it, using the score 4.5.

The mean is _____ victim(s)
(Round to two decimal places as needed.)

**b.** Find the median. Note the "4 or more" category is not problematic.

The median is _____ victim(s).

**Example 2**. The following table gives the losses in the principal battles of the Civil War. The figures are the total for killed, wounded, and missing.

| Battle | Union |
|---|---|
| Nashville | 2,140 |
| Franklin | 2,326 |
| Bull Run | 2,952 |
| Atlanta | 3,641 |
| Perryville | 4,348 |
| Chattanooga | 5,616 |
| Seven Pines and Fair Oaks | 5,739 |
| Second Bull Run | 7,800 |
| Murfreesboro | 11,578 |

| Battle | Union |
|---|---|
| Fredericksburg | 12,353 |
| Antietam | 12,469 |
| Shiloh | 13,573 |
| Seven Days Battles | 15,249 |
| Chickamauga | 15,851 |
| Chancellorsville | 16,030 |
| Gettysburg Campaign | 23,186 |
| Spotsylvania | 26,461 |
| Wilderness | 37,737 |

Create the Five Number Summary for this quantitative data:

Minimum Q0 = _____
1st Quartile Q1 = _____
Median Q2 = _____
3rd Quartile Q3 = _____
Maximum Q4 = _____

Find the Interquartile Range (IQR) for this set of data: _____

Sketch a box plot with well-labeled axis.

The range of values within 1.5*IQR of the median is considered unremarkable data, while values that fall outside this range are often flagged as **potential outliers**.

What is the value of 1.5*IQR? _____
What is the range of data within 1.5*IQR of the median? _____
Are there any potential outliers? _____

Another way to characterize potential outliers is by identifying values that are **more than 3 standard deviations from the mean**. In the Civil War, data the mean is about 12,110 and the standard deviation is about 9430. Are there any data values more than 3 standard deviations from the mean?

**PRACTICE.** Repeat the problems on this page using the GPA data provided on the previous worksheet.