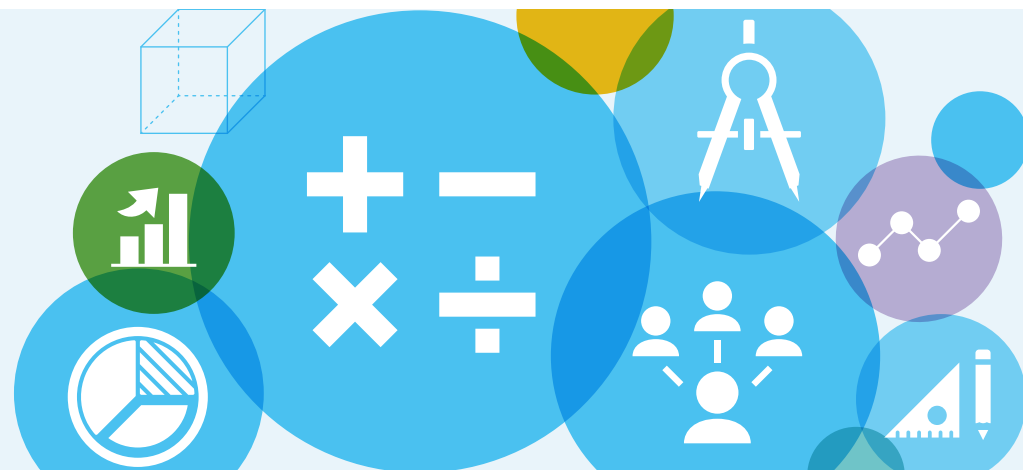




OVER THE HILL — AGING ON A NORMAL CURVE

Activity Items

The following items are part of this activity and appear at the end of this student version.

- Item 1: Percentage of Population Aged 65 and Older
- Item 2: Percentage of Family Households
- Item 3: Average Hours of Student Sleep on a School Night

Student Learning Objectives

- I will be able to estimate percentages based on a normal distribution using the “empirical rule.”
- I will be able to use a data set’s mean and standard deviation to fit it to a normal distribution.

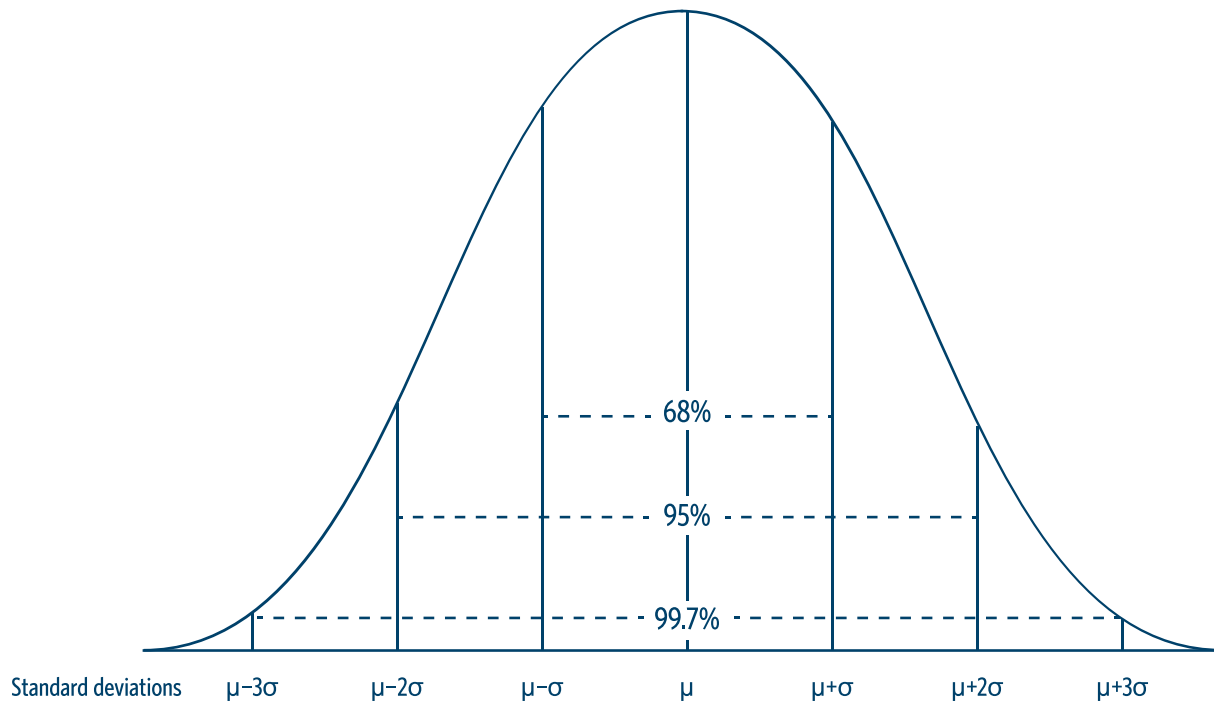
NAME: _____ DATE: _____

Part 1 – Examine a Normal Distribution

1. Examine the histogram in **Item 1: Percentage of Population Aged 65 and Older**, which shows data for all counties that are designated by the U.S. Census Bureau in the United States. What do you notice?
2. Approximately how many counties have populations in which roughly 24 percent of people are 65 and older?
3. Knowing that **Item 1** represents a normal distribution, what values would you estimate for the median and for the mean? Explain.

Part 2 - Estimate Values Using the “Empirical Rule”

One reason to use a normal distribution model is to make estimates about data — especially data that are not readily available. The model gives statisticians confidence that the mean is equal to the median and that the empirical rule holds true. This rule, also called the “68-95-99.7 rule,” states that for normal distributions:

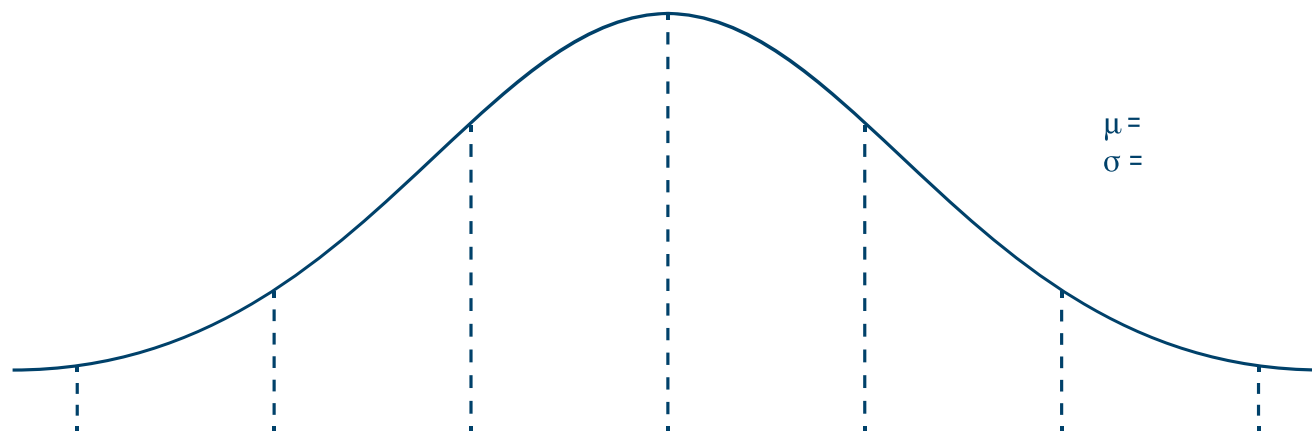


- Approximately 68 percent of the data are within one standard deviation (σ) of the mean (μ).
- Approximately 95 percent of the data are within two standard deviations (σ) of the mean (μ).
- Approximately 99.7 percent of the data are within three standard deviations (σ) of the mean (μ).

In this activity, you will use the normal model and the empirical rule to find estimates for various data.

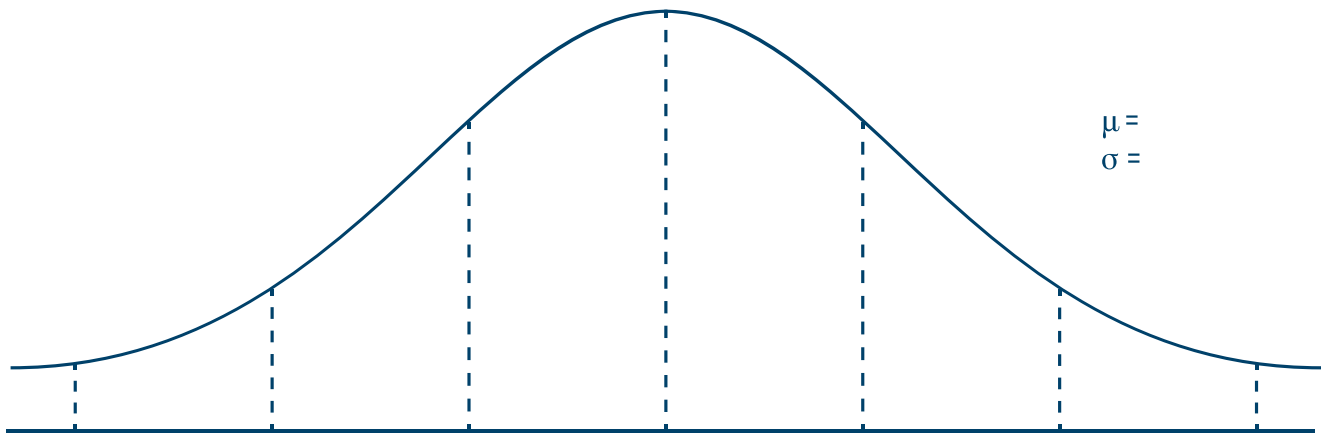
1. Examine the following hypothetical data distribution. The superintendent of a large county school system believes that the SAT mathematics scores of high school seniors in her county follow an approximately normal distribution. She knows that the mean score in her county is 510 and that the standard deviation (SD) is 90, so she can use the empirical rule to make other estimates.

- a. Fill in the normal curve below with values for μ and σ , and label each interval and the percentage of data each comprises, based on the normal approximation of those values.



- b. The superintendent wants to know: In this county, what percentage of students would you predict would get an SAT math score between 600 and 690? Explain, and show your work.
2. Use **Item 1**, where the mean proportion of people 65 and older in populations across all U.S. counties is 0.18 and the SD is 0.04, to answer the following questions and prompts. Note: Percentages are written as proportions in this case — so, for example, 22 percent would be 0.22.

- a. Fill in the normal curve below with values for μ and σ , and label each interval and the percentage of data each comprises, based on the normal approximation of those values.



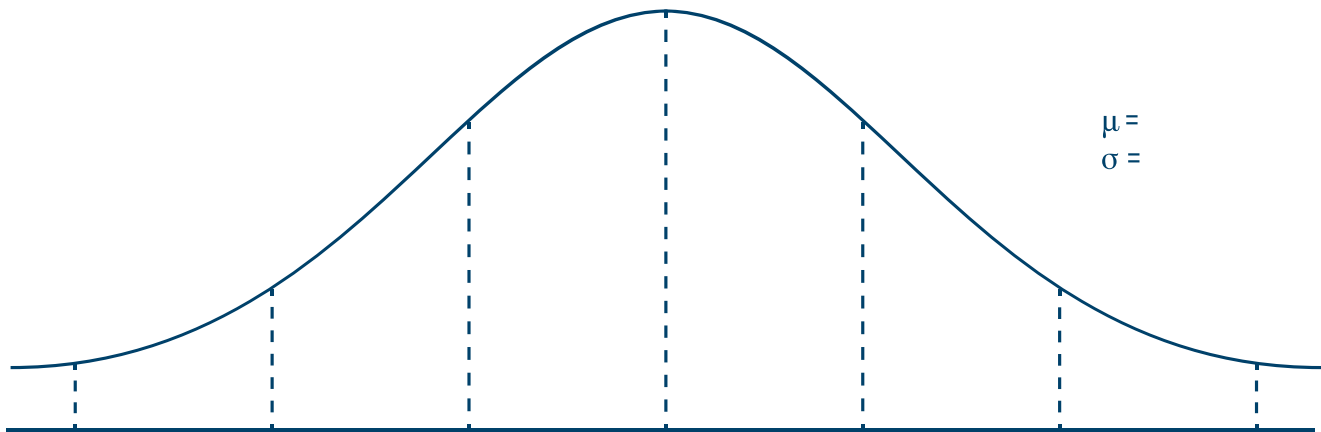
- b. Now find the values of $\mu - 3\sigma$, $\mu - 2\sigma$, and so on, showing your work below. Write these values in the correct places on the normal curve as well.
- c. If we assume the normal model for these data, in about what percentage of counties would the proportion of people 65 and older be more than 0.22? Show your work and explain your answer in the context of the data.
- d. In about what percentage of counties would the proportion of people 65 and older be less than 0.14? Explain, showing your work.

- e. In about what percentage of counties would the proportion of people 65 and older be less than 0.10? Explain, showing your work.

- f. In about what percentage of counties would the proportion of people 65 and older be greater than 0.10? Explain, showing your work.

- g. According to the actual census data, in 12.6 percent of counties the proportion of residents 65 and older was more than 0.22; in 16.5 percent of counties it was less than 0.14; and in 2.2 percent it was less than 0.10. Compared with the predictions you made earlier, do the actual data support, or fail to support, the use of a normal model for this distribution? Explain.

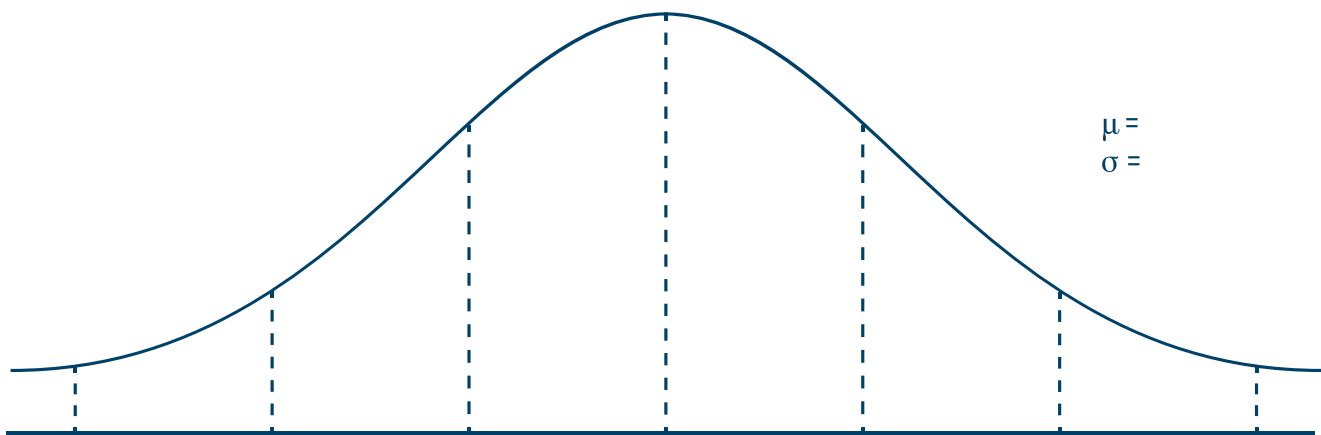
3. **Item 2: Percentage of Family Households** shows the percentages of households with children younger than 18, by county, for a random selection of 136 of the more than 3,000 counties in the United States. The mean of the data is 29.4 percent and the SD is 5.3 percent. Because the distribution is roughly bell-shaped, a normal model can be used to make predictions about all counties in the country.
- a. Fill in the normal curve below with values for μ and σ , and label each interval and the percentage of data each comprises, based on the normal approximation of those values.



- b. About what percentage of U.S. counties have children in 24.1 to 34.7 percent of their households? Explain how you know.
- c. About what percentage of U.S. counties have children in 34.7 to 40.0 percent of their households? Explain how you know.

- d. About what percentage of U.S. counties have children in less than 18.8 percent of their households? Explain how you know.
- e. Is your answer to question 3d consistent with the histogram in **Item 2**? Explain.
- f. According to the census data for just the 136 counties surveyed, 75.7 percent had children in 24.1 to 34.7 percent of their households; 8.9 percent had children in 34.7 to 40.0 percent of them; and 2.9 percent had children in less than 18.8 percent of them. Compared with the predictions you made earlier, do the actual data support, or fail to support, the use of a normal model for this distribution? Explain.
- g. List two reasons that the percentages you predicted earlier could differ from those actually reported by the Census Bureau.

4. **Item 3: Average Hours of Student Sleep on a School Night** is derived from a random sample of 441 high school students. The histogram is approximately bell-shaped and fairly symmetric. Its mean is 7.2 hours, its SD is 1.5 hours, and it approximately fits a normal model.
- a. Fill in the normal curve below with values for μ and σ , and label each interval and the percentage of data each comprises, based on the normal approximation of those values. Assume that the model applies to all high school students.

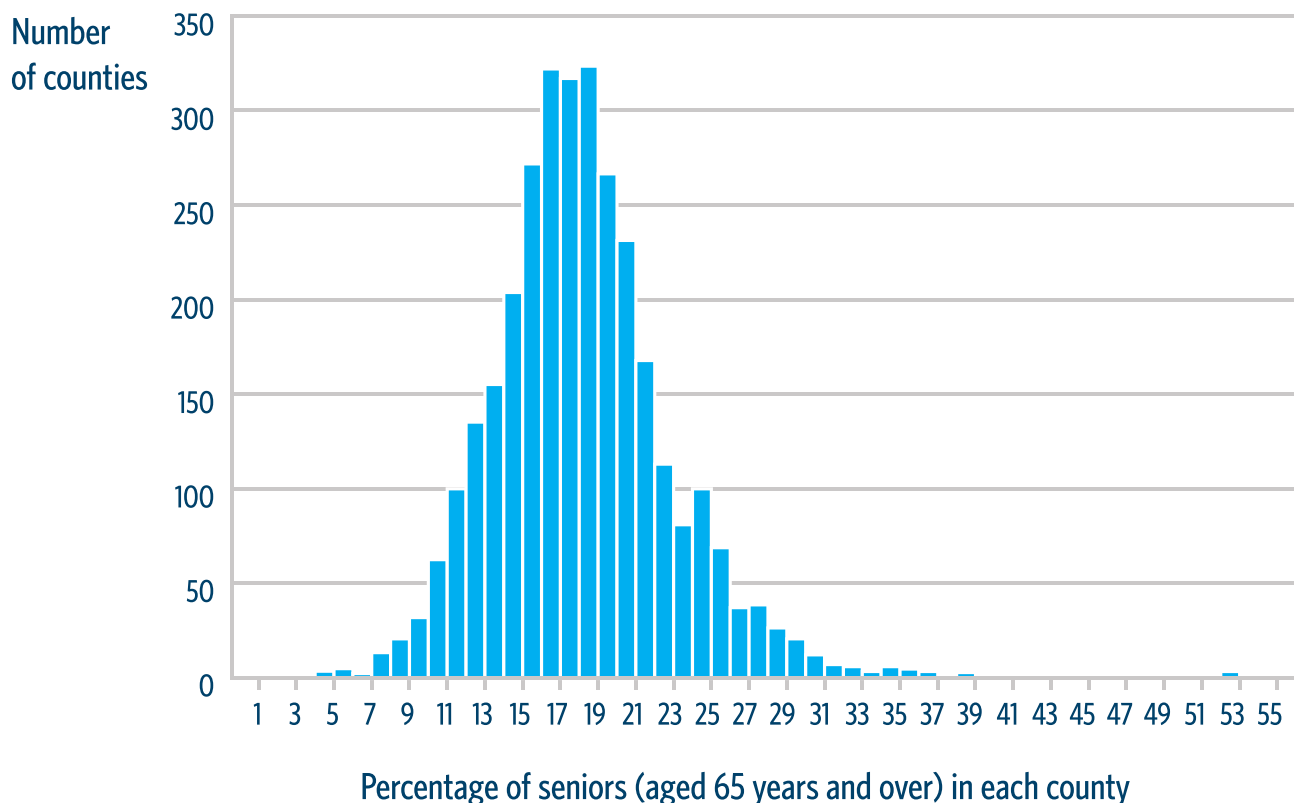


- b. About what percentage of all high school students would you predict to sleep between 5.7 and 8.7 hours each night? Explain how you know.
- c. About what percentage of all high school students would you predict to sleep more than 8.7 hours each night? Explain how you know.
- d. About what percentage of all high school students would you predict to sleep fewer than 8.7 hours each night? Explain how you know.

- e. According to **Item 3**, some students reported that they slept very little each night. Based on the normal model, about what percentage of all students would you estimate sleep fewer than 2.7 hours a night? Explain how you know.

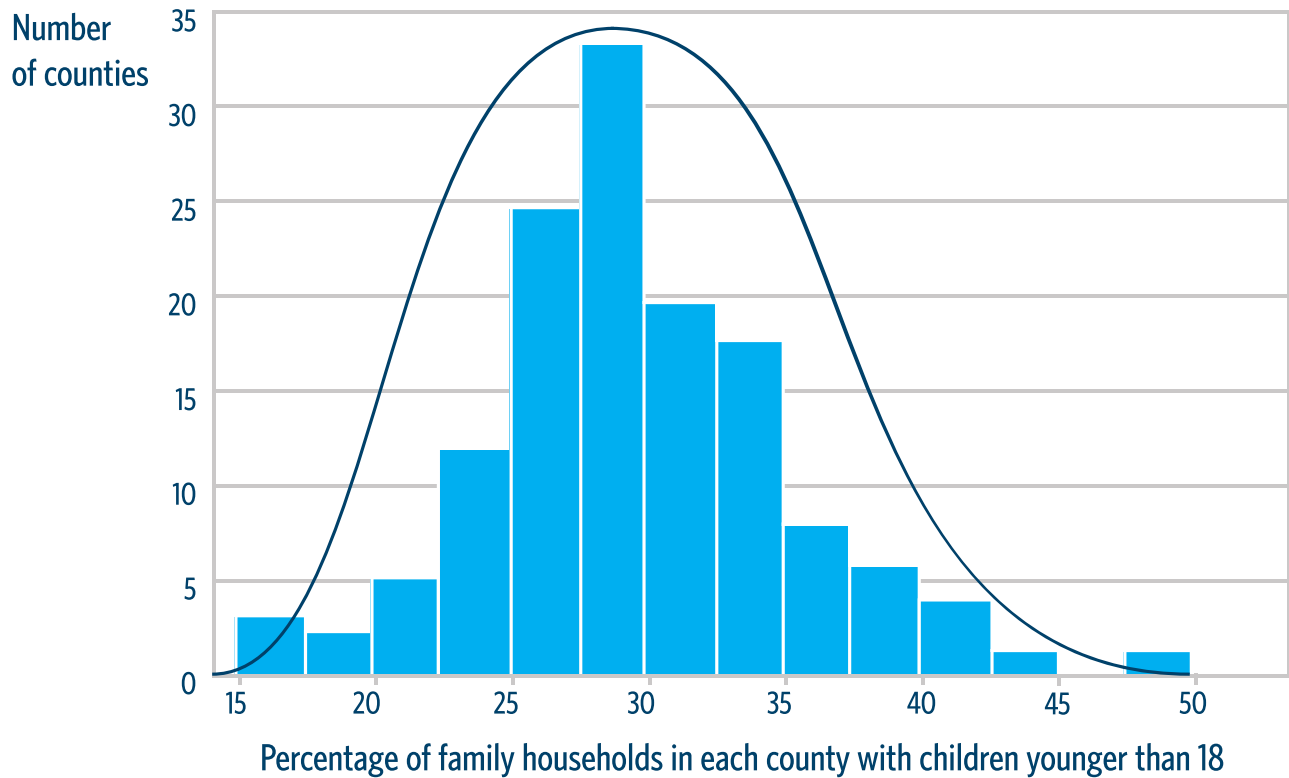
 - f. According to the actual survey data, 70 percent of the students reported sleeping between 5.7 and 8.7 hours each night; 15.6 percent reported sleeping more than 8.7 hours; and 0.01 percent reported sleeping 3 or fewer hours. Compared with the predictions you made earlier, do the actual data support, or fail to support, the use of a normal model for this distribution? Explain.
5. Not all data distributions can use the normal model to make estimates. Based on what you know about normality and what you learned in this activity, describe a hypothetical distribution of any variable that cannot be modeled using a normal distribution.

Item 1: Percentage of Population Aged 65 and Older



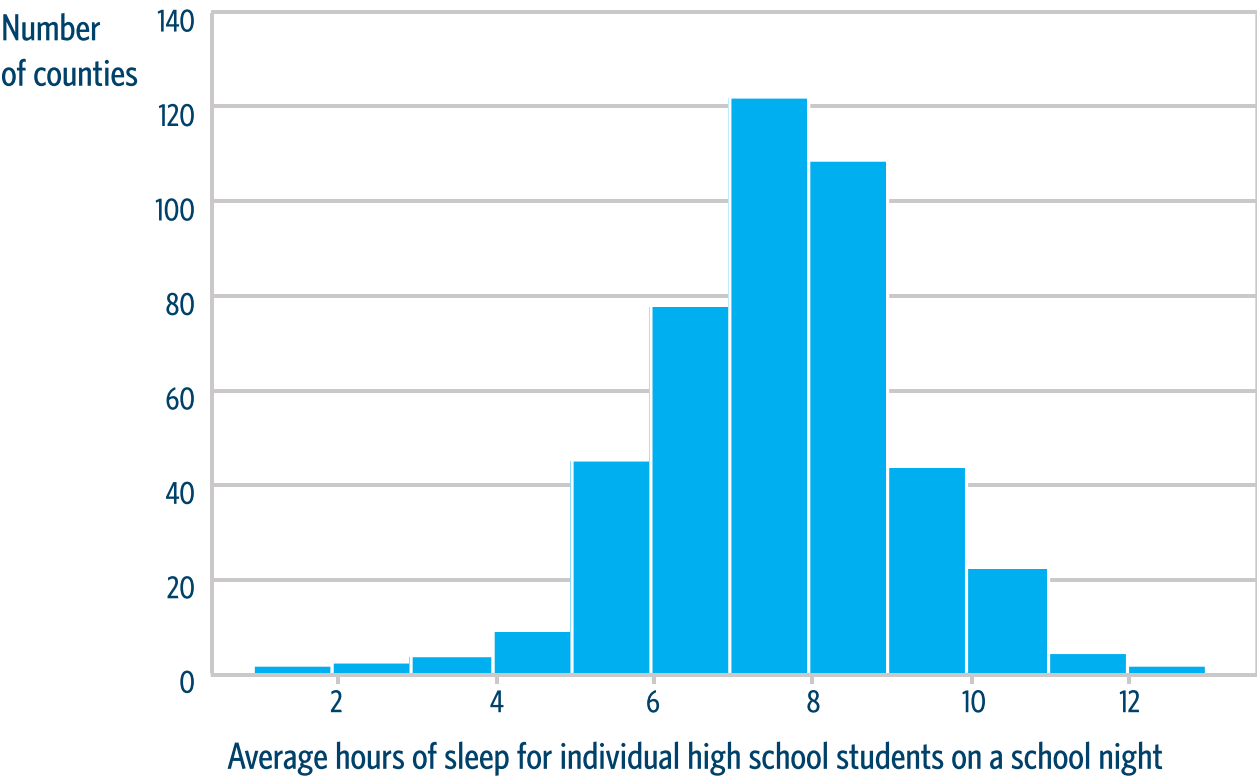
Source for data: U.S. Census Bureau, Population Division, Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties, and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2014.

Item 2: Percentage of Family Households



Source for data: U.S. Census Bureau, 2014 American Community Survey 1-Year Estimates, Selected Population Profile in the United States.

Item 3: Average Hours of Student Sleep on a School Night



Note: These data come from a random sample generated by the Census at School Random Sampler data tool:

www.amstat.org/censusatschool/RandomSampleExport.cfm