Chapter 11
Data Analysis and Statistics

1. **Using Normal Distributions**
2. Populations, Samples, and Hypotheses
3. Collecting Data
4. Experimental Design
5. Making Inferences from Sample Surveys
6. Making Inferences from Experiments

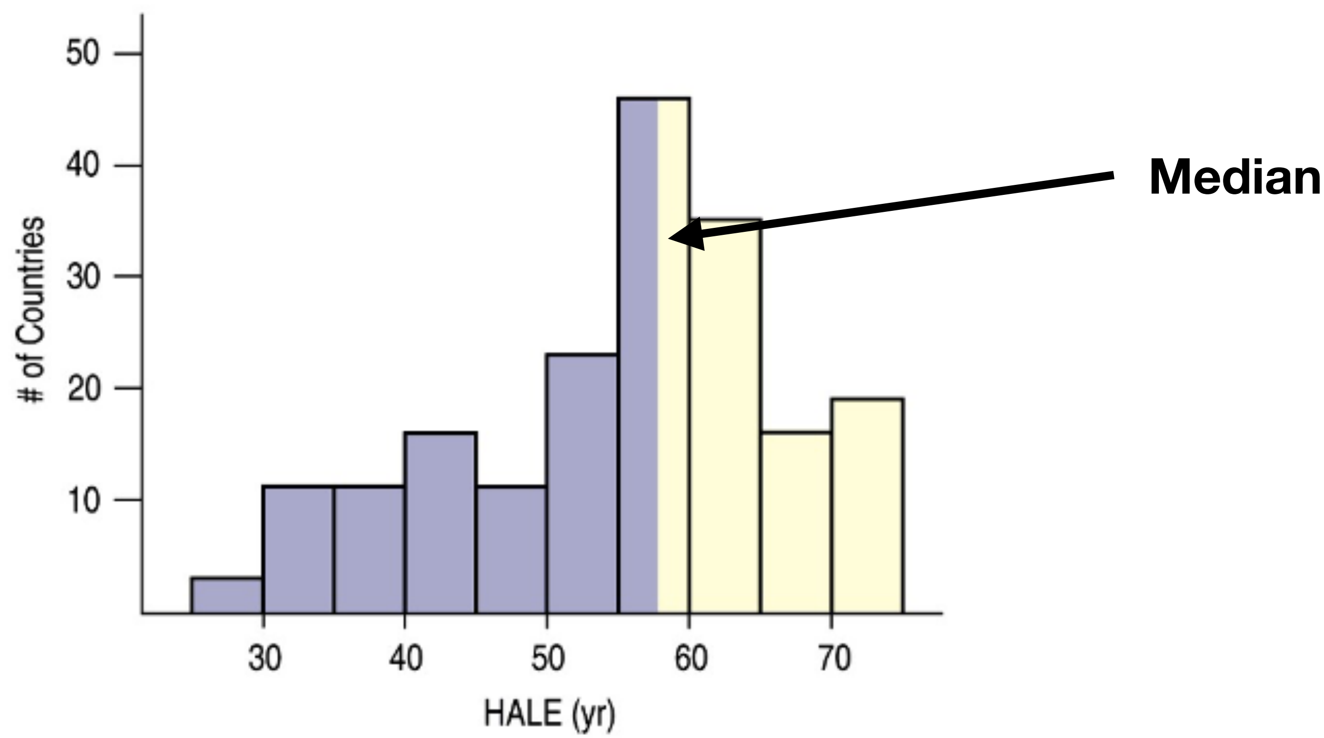
11.1 - Using Normal Distributions

Mean (μ) - is the average value in a dataset

$$\mu = \frac{\sum_{k=1}^{n} x_k}{n}$$

Median - the value in a dataset that has exactly half the data values above it and half below it.

Midrange - the average of the min and max values in a dataset

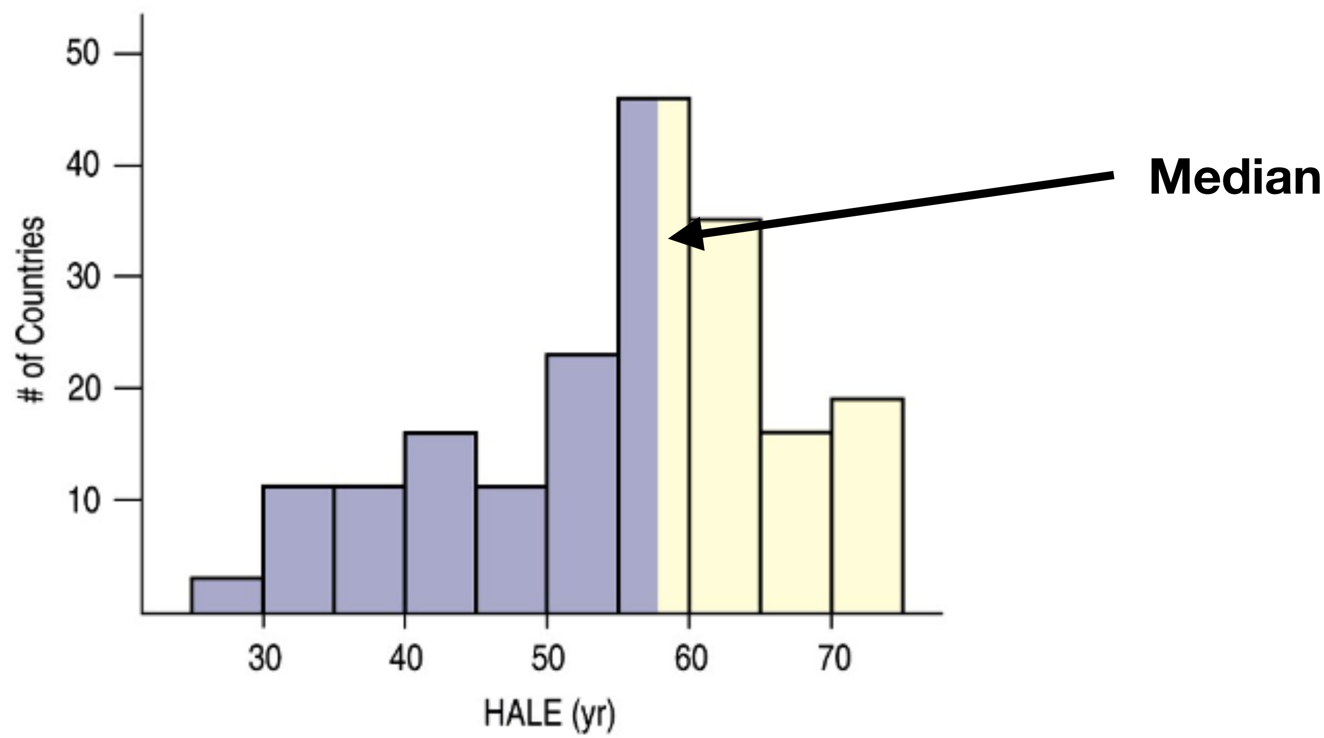


11.1 - Using Normal Distributions

The **median** is equal to the **mean** (μ) when the distribution of data is symmetric.

The **median** is a better choice for **center** than midrange.

For **unimodal** (one mode), symmetrical distribution, it is easy to find the center. It is just the center of symmetry.



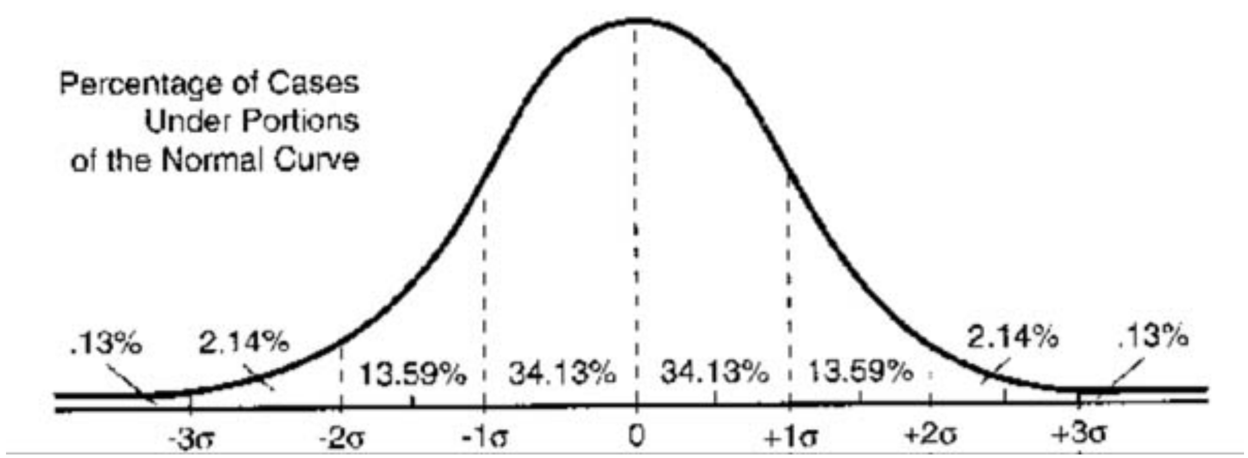
11.1 - Using Normal Distributions

The spread of the data is also important.

The **standard deviation** (σ) is how far each data value is from the mean.

$$\sigma = \sqrt{\frac{\sum (x-\mu)^2}{n}}$$

The **variance** (σ^2) is just standard deviation squared. Variance will come into play later in the unit.

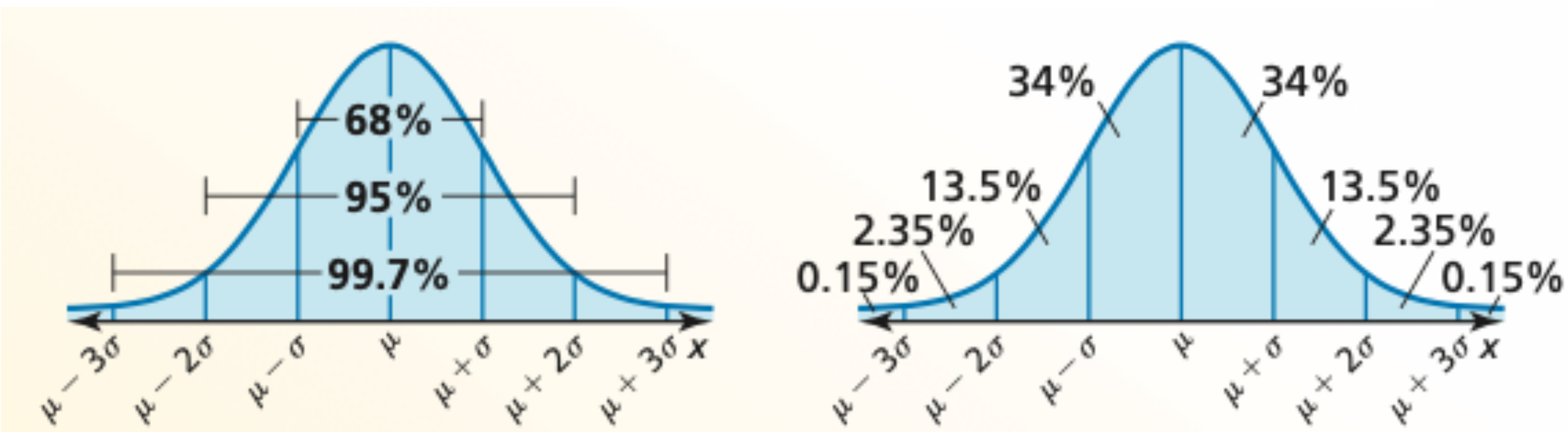


11.1 - Using Normal Distributions

The graph of a **normal distribution** is a bell-shaped curve called a **normal curve** that is symmetric about the mean.

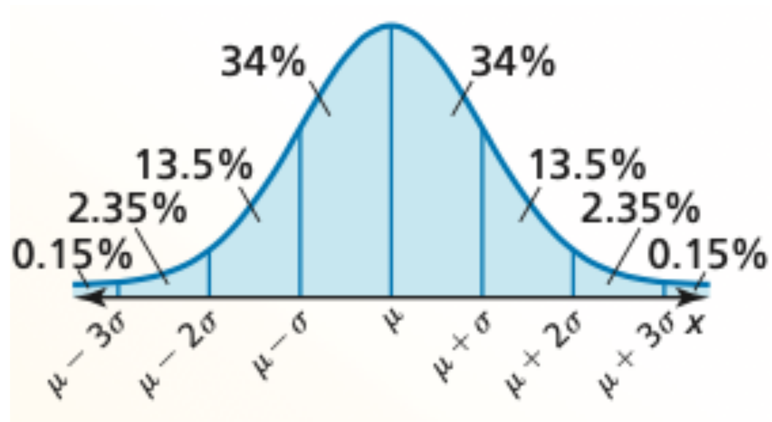
Areas Under a Normal Curve

A normal distribution with mean μ (the Greek letter *mu*) and standard deviation σ (the Greek letter *sigma*) has these properties.



11.1 - Using Normal Distributions

If an x-value is randomly selected from a distribution, what is its probability?



a. $P(x \geq \mu)$ $0.5 = 50\,\%$

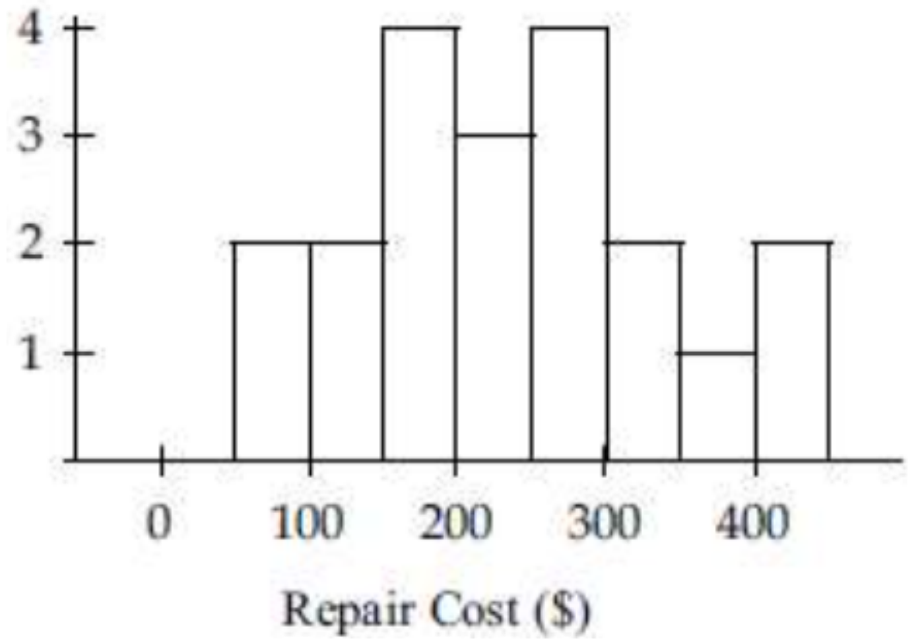
b. $P(\mu \leq x \leq \mu + 2\sigma)$ $0.475 = 47.5\,\%$

c. $P(\mu - \sigma \leq x \leq \mu)$ $0.34 = 34\,\%$

11.1 - Using Normal Distributions

An automobile brake and muffler shop reported the repair bills for their customers yesterday.

88	203
283	118
312	143
290	252
172	227
154	56
400	192
381	292
346	213
181	422



a. Find the mean and standard deviation of the repair costs.

$$\mu = \$236.25 \qquad \sigma = \$100.81$$

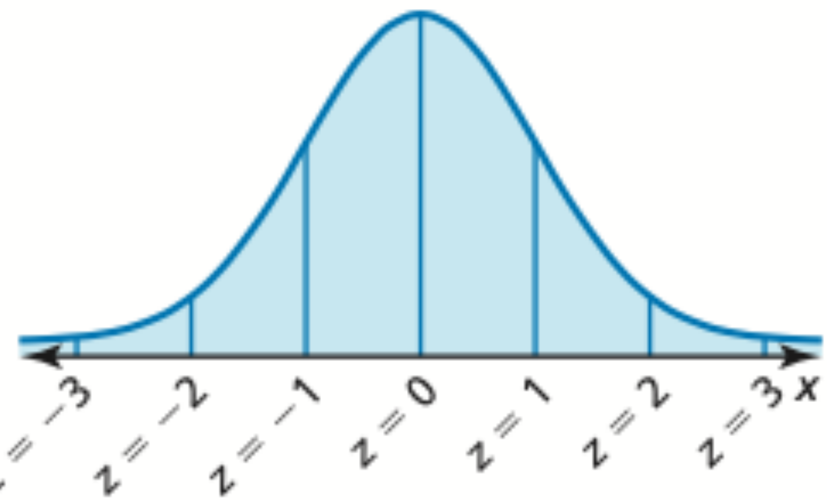
b. Is it appropriate to use the mean and standard deviation to summarize these data?

Yes. The data are roughly symmetric, with no outliers.

11.1 - Using Normal Distributions

The **standard normal distribution** is the normal distribution with mean 0 and standard deviation 1.

$$z = \frac{x - \mu}{\sigma}$$



The z-value for a particular x-value is called the **z-score**. It is the number of standard deviations the x-value lies above or below the mean μ .

11.1 - Using Normal Distributions

The standard normal table can be used to determine the probability of a z-score.

Standard Normal Table										
z	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
−3	.0013	.0010	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000+
−2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
−1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
−0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159
1	.8413	.8643	.8849	.9032	.9192	.9332	.9452	.9554	.9641	.9713
2	.9772	.9821	.9861	.9893	.9918	.9938	.9953	.9965	.9974	.9981
3	.9987	.9990	.9993	.9995	.9997	.9998	.9998	.9999	.9999	1.0000−

11.1 - Using Normal Distributions

Real-Life Example

a. A study finds that the weights of infants at birth are normally distributed with a mean of 3270 grams and a standard deviation of 600 grams. An infant is randomly chosen. What is the probability that the infant weighs 4170 grams or less? Use the table from the previous slide.

$$z = \frac{x - \mu}{\sigma}$$

$$P(z \leq 1.5) = 0.9332$$

b. What is the probability that the infant weighs 3990 grams or more?

$$P(z \geq 1.2) = 1 - 0.8849 = 0.1151$$

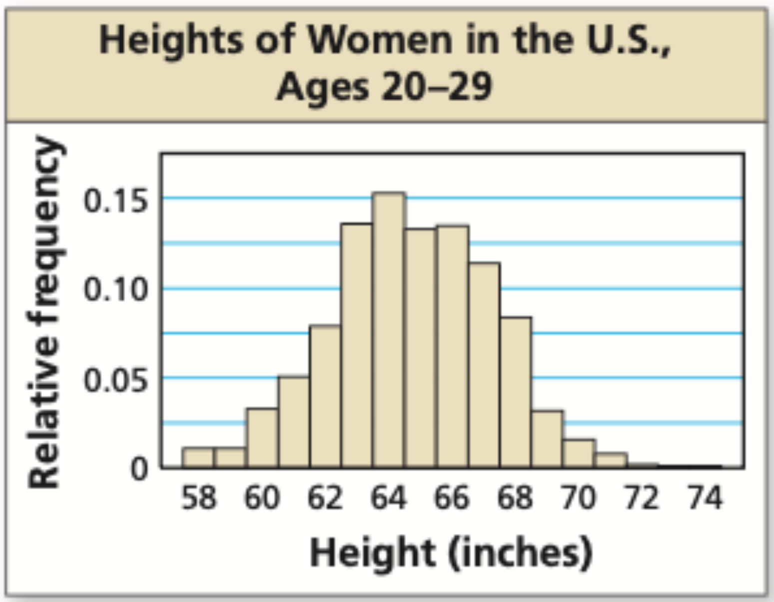
11.1 - Using Normal Distributions

Recognizing Normal Distributions

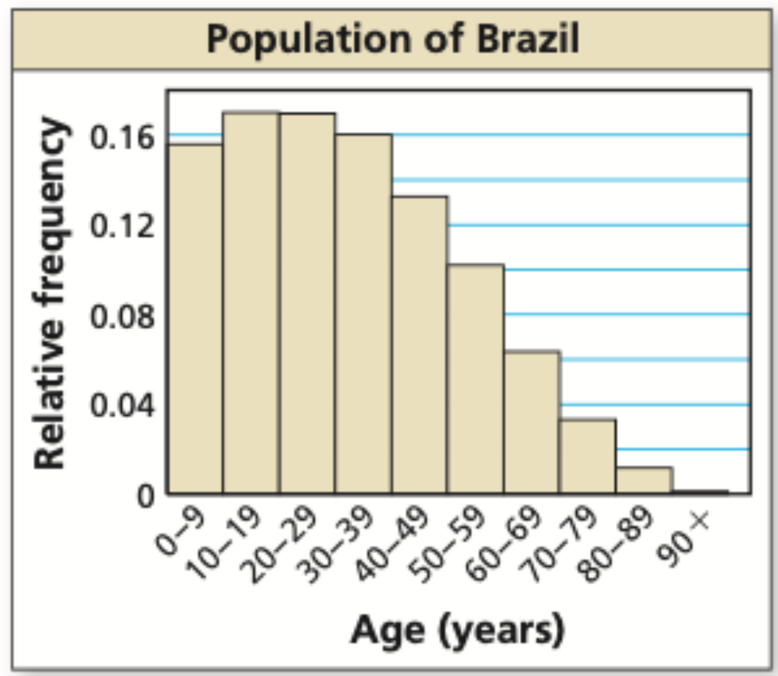
Not all distributions are normal.

Determine whether histogram has a normal distribution.

a.



b.



approx. normal distribution

skewed right,
not normal distribution