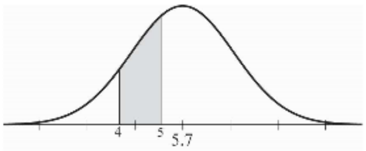
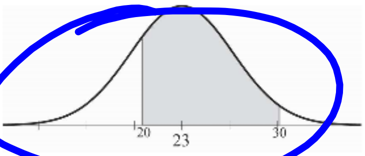
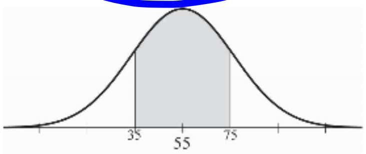
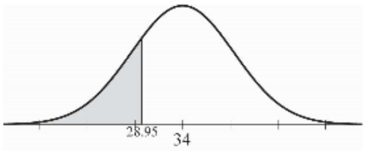
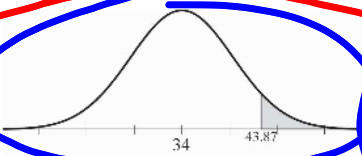
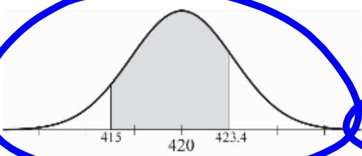
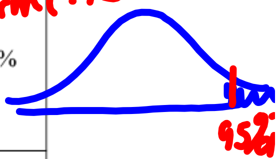
Statistics Chapter 5: Normal Practice 2 – KEY

Normal Model	Sketch	Left cutpoint and z-score	Right cutpoint and z-score	Percentage
1. N(5.7, 1.3)		4 z = -1.31	5 z = -.54	20.0%
2. N(23, 3.4)		20 z = - 0.88	30 z = 2.06	79.1%
3. N(55, 20)		35 z = - 1	75 z = 1	Center 68%
4. N(34, 6)		- ∞ z = - ∞	28.95 z = - 0.84	Lowest 20%
5. N(34,6)		43.87 z = 1.645	∞ * z = ∞ 10	Highest 5%
6. N(420, 3.4)		415 z = - 1.47	423.4 z = 1	77.1%

$1.645 = \frac{x-34}{6}$

invnorm(.95)



normalcdf (-1.47,1)

7. On the back write a word problem to give a possible context to one of the examples above.

Responses will vary.

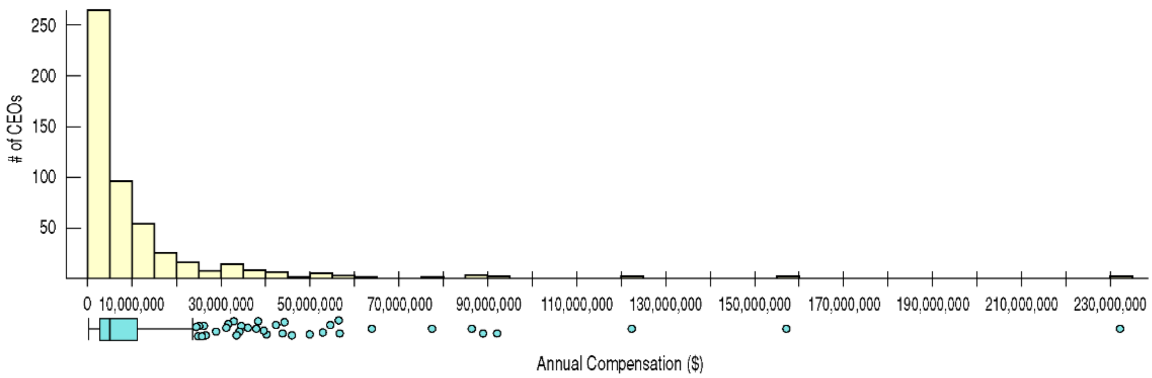
$z = \frac{415-420}{3.4}$

$1 = \frac{x_1 - 42}{3.4}$

Review Day

What Can Go Wrong?

- Don’t use a Normal model when the distribution is not unimodal and symmetric.



What Can Go Wrong? (cont.)

- Don't use the mean and standard deviation when outliers are present—the mean and standard deviation can both be distorted by outliers.
- Don't round your results in the middle of a calculation.
- Don't worry about minor differences in results.

What have we learned?

- We've learned the power of standardizing data.
  - Standardizing uses the SD as a ruler to measure distance from the mean (z-scores).
  - With z-scores, we can compare values from different distributions or values based on different units. Even if the data does not follow a normal model.
  - z-scores can identify unusual or surprising values among data.

 Chapter 5, Slide 29

What have we learned? (cont.)

±1SD ±2SD ±3SD

- We've learned that the 68-95-99.7 Rule can be a useful rule of thumb for understanding distributions:
  - For data that are unimodal and symmetric, about 68% fall within 1 SD of the mean, 95% fall within 2 SDs of the mean, and 99.7% fall within 3 SDs of the mean.

What have we learned? (cont.)

- We see the importance of *Thinking* about whether a method will work:
  - **Normality Assumption**: We sometimes work with Normal tables (Table Z). These tables are based on the Normal model.
  - Data can't be exactly Normal, so we check the **Nearly Normal Condition** by making a histogram (is it unimodal, symmetric and free of outliers?).

Statistics Chapter 5: Review B – KEY

1. Compile your knowledge about these ideas from this chapter in your own words and examples

Idea	Description	Example
Unimodal, symmetric distributions	A distribution with a single peak. The left half of the distribution is a near mirror image of the right.	The distribution of a sample of heights of adults males is likely to be unimodal and symmetric.
Standard deviation	A measure of spread based on distance from the mean. Standard deviation is a useful measure of spread for unimodal and symmetric distributions.	The standard deviation of the distribution of heights of adult males may be around 3".
Using the standard deviation as a ruler	If we measure in standard deviations, we can compare values from distributions that have different scales.	We can decide if it is more likely for males to be over 72" or females to be over 68" by comparing the number of standard deviations each measurement is from its mean.
z-score (standardizing)	A measure of how many standard deviations away from the mean a value is. $z = \frac{y-\mu}{\sigma}$	If male heights are distributed according to N(68", 3"), a man who is 65" would have a z-score of -1.
Normal models	Useful models for describing unimodal and symmetric distributions.	Male heights and weights are not Normal, but can be modeled by Normal distributions.
Empirical Rule for Normal Models	In a Normal model, 68% of values are expected to be within 1 SD of the mean, 95% within 2 SDs, and 99.7, or almost all observations, within 3 SDs of the mean.	68% of males heights are expected to be between 65" and 71", 95% of heights are expected to be between 62" and 74", and 99.7% are expected to be between 59" and 77".

2. Madeline's height has a z-score of -2.1 compared to the heights of girls her age. What does this mean?

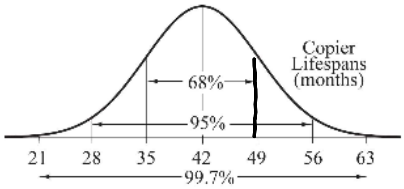
   If the distribution of heights is unimodal and symmetric, Madeline's height is 2.1 standard deviations ~~above~~ below the mean height of girls her age.

3. Martin is 41.5" tall. The mean and standard deviation of heights of boys his age are 46" and 1.3" respectively. Who is taller for their age, Madeline or Martin? Explain using statistics.

   Martin's z-score is ~~(46 – 44)/1.3 = 1.5~~ $(41.5-46)/1.3 = -3.46$. ~~Madeline~~ Martin is taller for ~~her~~ his age. ~~Her~~ His z-score of ~~2.1~~ -3.46 is ~~greater than~~ farther from the mean ~~Martin~~'s.

4. A manufacturer claims that lifespans for their copy machines (in months) can be described by a Normal model $N(42,7)$.

   a. Draw and clearly label the model, showing ± 1, 2, and 3 standard deviations from the mean and the corresponding months and percents.



   b. What does the mean and standard deviation describe about copier lifespans?

According to the Normal model, a typical copier might last about 42 months, 68% of copiers are expected to last between 35 and 49 months, 95% of copier are expected to last between 28 and 56 months, and 99.7% of copiers are expected to last between 21 and 63 months.

$z = \frac{28-42}{7} = -2$

$z = \frac{49-42}{7} = 1$

   c. According to this model, what percent of copiers are expected to last between 28 and 49 months?

normalcdf(−2, 1) find z z

According to the Normal model, 81.5% of copiers are expected to last between 28 and 49 months.

   d. According to this model, hat percent of copies last more than 36 months? More than 4 years? Less than 2 years?

	4 years	2 years
$z = \frac{36-42}{7} = -0.86$	$z = \frac{48-42}{7} = 0.86$	$z = \frac{24-42}{7} = -2.57$
normalcdf(-.86,10)	normalcdf(.86,10)	normalcdf (-10,-2.57)
$Pct(z > -0.86) = 80.4\%$	$Pct(z > 0.86) = 19.6\%$	$Pct(z < 2.57) = 0.5\%$
According to the Normal model, 80.4% of copiers are expect to last more than 36 months.	According to the Normal model, 19.6% of copiers are expected to last more than 4 years (48 months).	According to the Normal model, 0.5% of copiers are expected to last less than two years.

   e. According to this model what is the 3rd quartile of copier lifespans? Describe the meaning of this statistic in full context. Q3 = 75th percentile

z=invnorm(.75) = $0.6745 = \frac{y-42}{7}$

$y = 46.7$

According to the Normal model, the third quartile of copier lifespans is 46.7 months. In other words, 25% of copiers are expected to last more than 46.7 months.