- This Slideshow was developed to accompany the textbook
  - *Larson Algebra 2*
  - *By Larson, R., Boswell, L., Kanold, T. D., & Stiff, L.*
  - *2011 Holt McDougal*
- Some examples and diagrams are taken from the textbook.

Slides created by
Richard Wright, Andrews Academy
rwright@andrews.edu

11.1 Find Measures of Central Tendency and Dispersion

- **Measure of central tendency**
  - A number used to represent the center or middle of a set of data values.
- M**ean** , or *average*, of n numbers is the sum of the numbers divided by n.

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

11.1 Find Measures of Central Tendency and Dispersion

- **Median**
  - middle number when the numbers are written in order. (If n is even, the median is the mean of the two middle numbers.)
- **Mode**
  - number or numbers that occur most frequently. There may be one mode, no mode, or more than one mode.

11.1 Find Measures of Central Tendency and Dispersion

- The winning scores of 6 baseball games are
  - 5, 7, 8, 5, 10, 3
- Find the mean, median, and mode.

11.1 Find Measures of Central Tendency and Dispersion

- **Measure of dispersion**
  - Statistic that tells you how dispersed, or spread out, data values are.
- **Range**
  - difference between the greatest and least data values.

$$Range = max - min$$

- Find the range of the following data sets.
  - 14,17,18,19,20,24,30,32
  - 12,16,18,18,18,20,23

11.1 Find Measures of Central Tendency and Dispersion

- **Standard deviation**
  - describes the typical differences (or deviation) between a data's value and the mean.

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}}$$

11.1 Find Measures of Central Tendency and Dispersion

- Find the standard deviation of the following data set.
  - 4,8,12,15,3
- Finding the standard deviation on a TI calculator
  - [STAT] → Edit, Enter data values in L1 (clear list first)
  - [STAT] → CALC → 1-Var Stats, [ENTER] x2 , Find σx

11.1 Find Measures of Central Tendency and Dispersion

- Outliers
  - Value that is much greater than or much less than most of the values in a data set.
  - Can skew measures of central tendency and dispersion

11.1 Find Measures of Central Tendency and Dispersion

- **Air Hockey** You are competing in an air hockey tournament. The winning scores for the first 10 games are given below.

14,15,15,17,11

a. Find the mean, median, mode, range, and standard deviation of the data set.
b. The winning score in the next game is an outlier, 25. Find the new mean, median, mode, range, and standard deviation.
c. Which measure of central tendency does the outlier affect the most? the least?
d. What effect does the outlier have on the range and standard deviation?

11.1 Find Measures of Central Tendency and Dispersion

- *On the homework, do one standard deviation by hand. You can use your calculator to do the rest.*

11.2 Apply Transformations to Data

- **Adding a Constant to Data Values**

- When a constant is added to every value in a data set, the following are true:
  - The mean, median, and mode of the new data set can be obtained by adding the same constant to the mean, median, and mode of the original data set.
  - Range and standard deviation are unchanged.

11.2 Apply Transformations to Data

- The data below give the weights of 5 people. At the end of a month, each person had lost 3 pounds. Give the mean, median, mode, range, and standard deviation of the starting weights and the weights at the end of the month.

138, 142, 155, 140, 155

11.2 Apply Transformations to Data

- **Multiplying Data Values by a Constant**
- When each value of a data set is multiplied by a positive constant, the new mean, median, mode, range, and standard deviation can be found by multiplying each original statistic by the same constant.

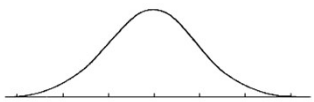
11.2 Apply Transformations to Data

- The data below give the weights of 5 people. Give the mean, median, mode, range, and standard deviation for the weights of the 5 people in kilograms.
- (*Note: 1 pound ≈ 0.45 kilogram)*

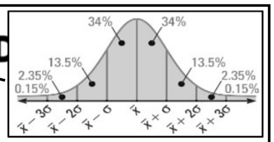
138, 142, 155, 140, 155

11.3 Use Normal Distributions

- A normal distribution is modeled by a bell-shaped curve called a normal curve that is symmetric about the mean.



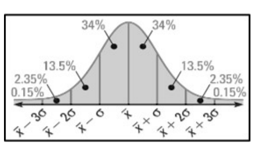
11.3 Use Normal Distributions



- A normal distribution with mean $\bar{x}$ and standard deviation σ has the following properties:
  - The total area under the related normal curve is 1.
  - About 68% of the area lies within 1 standard deviation of the mean.
  - About 95% of the area lies within 2 standard deviations of the mean.
  - About 99.7% of the area lies within 3 standard deviations of the mean.

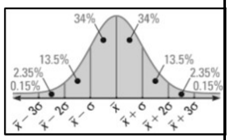
11.3 Use Normal Distributions

- A normal distribution has mean and standard deviation. For a randomly selected x-value from the distribution, find P($\bar{x} - \sigma \leq x \leq \bar{x} + 3\sigma$)



11.3 Use Normal Distributions

- The weight of strawberry packages is normally distributed with a mean of 16.18 oz and standard deviation of 0.34 oz. If you randomly choose 2 containers, what is the probability that both weigh less than 15.5 oz?



11.3 Use Normal Distributions

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

Formula = $Z = \frac{x-\bar{x}}{\sigma}$

- The z value for a particular x-value is called the **z-score** for the x-value and is the number of standard deviations the [...]lue lies above or below the mean $\bar{x}$.

11.3 Use Normal Distributions

- If a z-score is known, the probability of that value or less can be found from a **Standard Normal Table**.
- P(z ≤ -0.4) = 0.3446

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−

11.3 Use Normal Distributions

- **Finding Probabilities with Z-scores using a TI-graphing calculator**
- use the normalcdf function. It computes P($z_1 < z < z_2$), which is the area under the standard normal curve between $z_1 and z_2$.
- To calculate P($-1 < z < 2$), press **2nd DISTR, normalcdf(** and then press **ENTER.**
- After **normalcdf(** type **-1 , 2)** and then press **ENTER.**
- normalcdf(-1,2) = 0.8186

11.3 Use Normal Distributions

- A survey of 20 colleges found that the average credit card debt for seniors was $3450. The debt was normally distributed with a standard deviation of $1175. Find the probability that the credit card debt of the seniors was at most $3600.
- **Step 1: Find** the z-score corresponding to an x-value of $3600.
- **Step 2: Use** the table or normalcdf to find P($x \leq \$3600$).

11.4 Select and Draw Conclusions from Samples

- Population
  - A group of people or objects that you want information about.
- Sample
  - When it is too hard to work with everything, information is gathered from a subset of the population.
- There are 4 types of samples:
  - Self-selected – member volunteer
  - Systematic – rule is used to select members
  - Convenience – easy-to-reach members
  - Random – everyone has equal chance of being selected

11.4 Select and Draw Conclusions from Samples

- A manufacturer wants to sample the parts from a production line for defects. Identify the type of sample described.
  - The manufacturer has every 5th item on the production line tested for defects.
  - The manufacturer has the first 50 items on the production line d

11.4 Select and Draw Conclusions from Samples

- Unbiased Sample
  - Ensure accurate conclusions about a population from a sample.
  - An **unbiased sample** is representative of the population.
  - A sample that over- or underrepresents part of the population is a **biased sample.**
- Although there are many ways of sampling a population, a random sample is preferred because it is most likely to be representative of the population.

11.4 Select and Draw Conclusions from Samples

- A magazine asked its readers to send in their responses to several questions regarding healthy eating. Tell whether the sample of responses is biased or unbiased. Explain.

11.4 Select and Draw Conclusions from Samples

- The owner of a company with 300 employees wants to survey them about their preference for a regular 5-day, 8-hour workweek or a 4-day, 10-hour workweek. Describe a method for selecting a random sample of 50 employees to poll.

11.4 Select and Draw Conclusions from Samples

- Sample Size
  - When conducting a survey, the larger the sample size is, the more accurately the sample represents the population.
  - As the sample size increases, the margin of error decreases.
- Margin of error
  - Gives a limit on how much the responses of the sample would differ from the responses of the population.
- For a sample size n, the margin of error is:
- Margin of error = $\pm \frac{1}{\sqrt{n}}$

11.4 Select and Draw Conclusions from Samples

- **Survey** In a survey of 1535 people, 48% preferred Brand A over Brand B and Brand C.
  - What is the margin of error for the survey?
  - Give an interval that is likely to obtain the exact percent of all people who prefer Brand A.

11.4 Select and Draw Conclusions from Samples

- A polling company conducts a poll for a U.S. presidential election. How many people did the company survey if the margin of error is $\pm 3\%$?

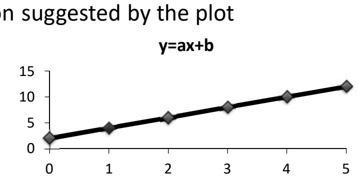
11.5 Choose the Best Model for Two-Variable Data

- To find the best model for a set of data pairs (x, y)…

1. Make a scatter plot
2. Determine the function suggested by the plot

Linear

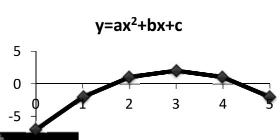
$y = ax + b$



11.5 Choose the Best Model for Two-Variable Data

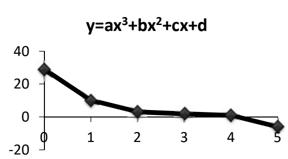
Quadratic

$y = ax^2 + bx + c$



Cubic

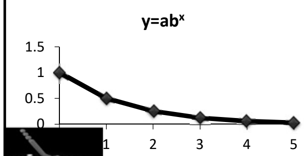
$y = ax^3 + bx^2 + cx + d$



11.5 Choose the Best Model for Two-Variable Data

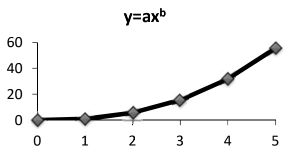
Exponential

$y = ab^x$



Power

$y = ax^b$



11.5 Choose the Best Model for Two-Variable Data

✸ To graph data on TI-Graphing Calculator

1. STAT → Edit…
2. Clear lists by highlighting L1 (or L2) and push CLEAR
3. Enter x-values in L1 and y-values in L2
4. Push Y= → clear any equations
5. In Y= hightlight Plot 1 and push ENTER
6. To zoom push ZOOM → ZoomStat
7. Choose type of graph (linear, quadratic, cubic, exponential, power)

This can be done on Excel if you don't have a graphing calculator.

11.5 Choose the Best Model for Two-Variable Data

✸ To see your regression with your data points

1. Select the type the regression from STAT→CALC
2. Specify the x-data (2^{nd} L1)
3. Comma
4. Specify the y-data (2^{nd} L2)
5. Comma
6. Name the regression Y1 (VARS→Y-VARS→Function…→Y1)
7. You should see "*your*Reg L1, L2, Y1"
8. Push Enter
9. Graph

This can be done on Excel if you don't have a graphing calculator.

11.5 Choose the Best Model for Two-Variable Data

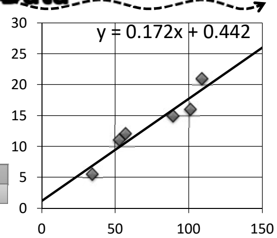
- Microsoft Excel
1. Enter your data in two columns
2. Highlight the columns and click Insert → Scatter
   - You should now have a scatter plot
3. To get a regression
   a. Select your graph and click Chart Tools Layout → Trendline → More Trendline Options
   b. Select your regression type (quadratic is polynomial order 2, cubic is polynomial order 3)
   c. Checkmark the Display Equation on Chart box

k OK and your regression and equation will be on the graph

11.5 Choose the Best Model for Two-Variable Data

- The table shows the cost of a meal x (in dollars) and the tip y (in dollars) for parties of 6 at a restaurant. Find a model for the data.

x	34.48	52.54	89.64	100.76	65.60	109.34
y	5.5	11	15	16	12	21



11.5 Choose the Best Model for Two-Variable Data

- The table shows amount y of money in your savings account after x weeks.

x	y
0	0
1	200
2	250
3	300
4	300
5	300
6	315
7	340
8	405