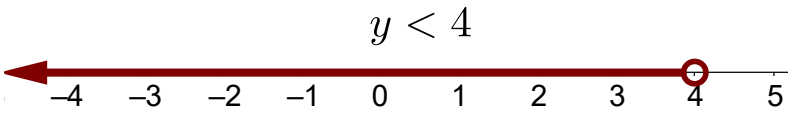
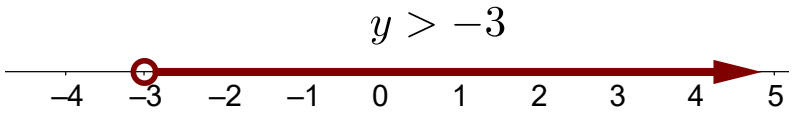
Module 1: Basic Numeracy & Calculation Skills

Intervals: Pages 1.05-1.05.1

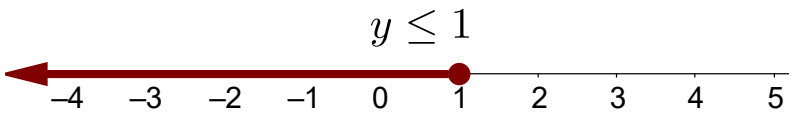
- An **open circle** denotes that the value is **not** to be included.
  - Example for "Less Than"



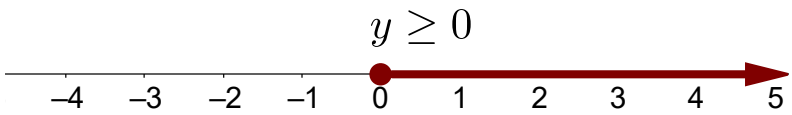
  - Example for "Greater Than"



- A **closed** circle means that the value **is** included.
  - Example for "Less Than or Equal To"



  - Example for "Greater Than or Equal To"



Sign rule for multiplication and division: Page 1.08

- Multiplying the same signs will result in a positive number
  - $+ \times + = +$
  - $- \times - = +$
- Multiplying different signs will result in a negative number
  - $- \times + = -$
  - $+ \times - = -$

Order of Operations: Pages 1.12-1.13

- Remember the acronym **PEMDAS**
  - **P**arentheses
  - **E**xponents
  - **M**ultiplication or **D**ivision (Left to Right)
  - **A**dditions or **S**ubtraction (Left to Right)

Module 2: Fractions, Decimals, & Percentages

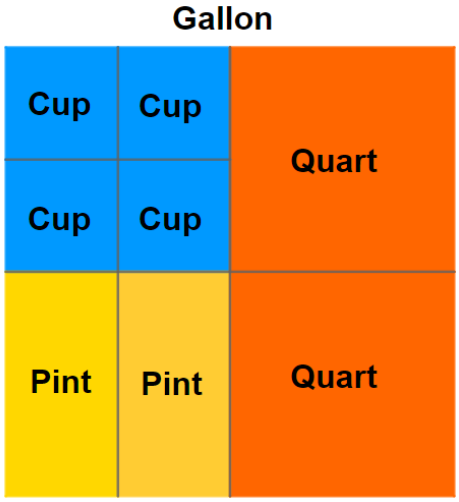
Converting Decimals, Fractions, and Percentages: Page 2.14

- Decimals, Fractions, and Percentages are just different ways of showing the same value
  - $\frac{1}{4}$ = 0.25 = 25%

Unit Conversions: Pages 2.15-2.15.3

(only the most common; a complete list can be found in MindEdge on page 2.15)

- **Common Unit Conversions for Household Measures of Volume:**
  - 1 tablespoon = 3 teaspoons
  - 1 fluid ounce = 2 tablespoons
  - The conversions involving one gallon can be visualized below:



- **Common Metric Conversions:**
  - 1 L = 1000 mL
  - 1 kg = 1000 g
  - 1 g = 1000 mg (milligrams)

Module 3: Basic Algebra

Like terms: Pages 3.04, 3.04.1, 3.04.2, and 3.04.3

- Terms that have the same variable(s) raised to the same exponent(s); they can be combined using addition and subtraction
  - Example: The expression $7x + 10 - 2x + 3$ is equal to $5x + 13$

Solving Linear Equations: Pages 3.12 and 3.13

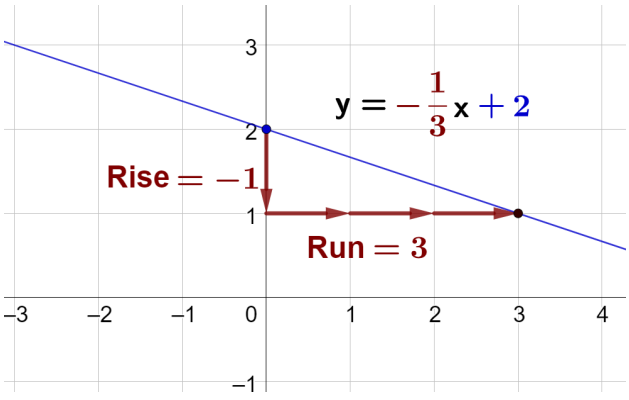
- If we add/subtract/multiply/divide the same quantity to both sides of an equation, the result will remain equal. (÷ by zero is not allowed)
  - Example: Starting with $x - 2 = 7$, we can add 2 to both sides of the equation to conclude $x = 9$.
- **Linear Inequalities: Page 3.19**: We solve linear inequalities the same way, but with one important exception – Whenever we multiply or divide both sides of an inequality by a negative number, we must switch the direction of the inequality.

Slope-intercept equation of a line: Pages 3.17, 3.17.1

- **y = mx + b**, where **m** is **slope** of the line and **b** is the **y-intercept**.
- The slope of a line is a description of its steepness. Positive slopes create "uphill" lines while negative slopes create "downhill" lines.

Graphing Linear Equations: Page 3.18

- First plot the y-intercept, where the line crosses the y-axis.
- Next, use Slope = $\frac{\text{Rise } (\updownarrow)}{\text{Run } (\leftrightarrow)}$, to locate a 2nd point on the line.
- Example: Graph $y = -\frac{1}{3}x + 2$



Module 4: Descriptive Statistics for a Single Variable

Types of Data: Page 4.02

Quantitative (numerical) data - consists of data values that are numerical, quantities that can be counted or measured (additions/subtractions make sense)

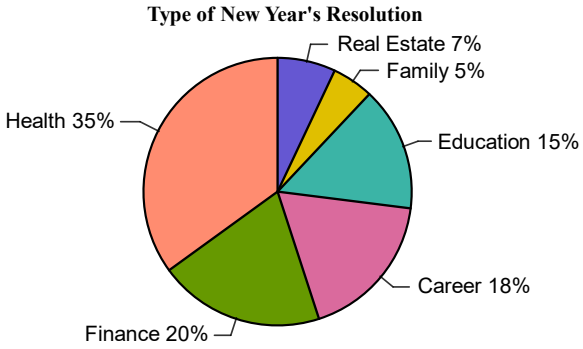
Examples: Height, Salary, Chance of rain, Weight

Categorical (qualitative) data - consist of data that are groups or labels, and are not necessarily numerical (additions/subtractions do not make sense)

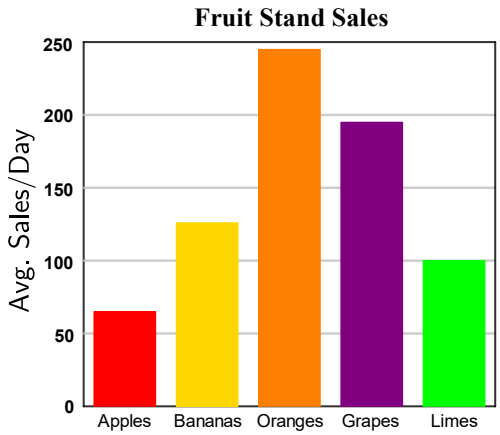
Examples: Hair Color, Country of Origin, Blood Type, Zip Codes

Graphical Displays for a single Categorical Variable: Pages 4.03-4.03.2

Pie Chart – Displays parts of the whole, percentages

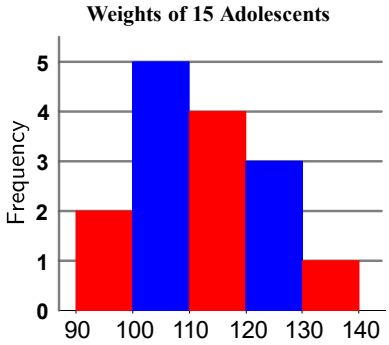


Bar Chart – Displays counts or frequencies of each category

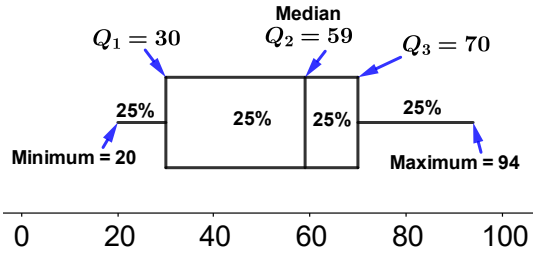


Graphical Displays for a single Quantitative Variable: Pgs. 4.04-4.04.3, 4.07

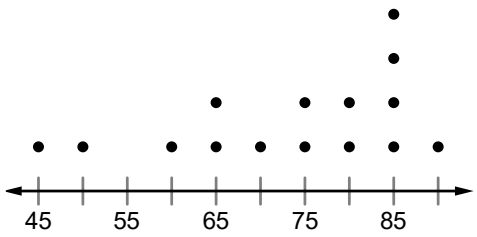
Histogram – displays the shape and spread of data



Box Plot – displays center, spread and outliers. Each section covers 25% of the data regardless of length. Can be horizontal or vertical.



Dot Plot – displays clusters, gaps, and outliers for smaller data sets. Each data value is seen in a dot plot.



Test Scores

Stem Plot – Display shape according to place values. Each data value if seen in a stem plot.

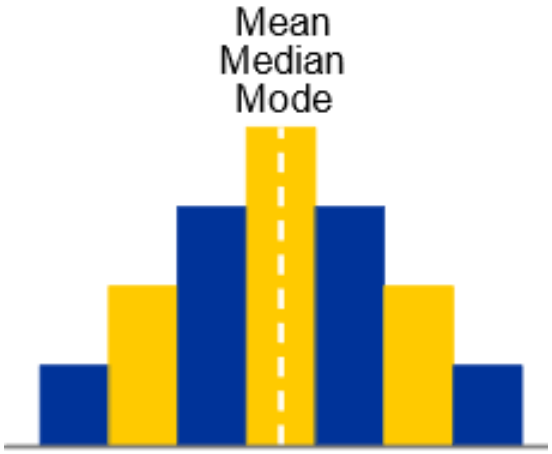
Data Set: {7, 9, 11, 13, 13, 20, 25, 46}

Stem	Leaf
0	7 9
1	1 3 3
2	0 5
3	
4	6

Histogram Shape: Page 4.04.2

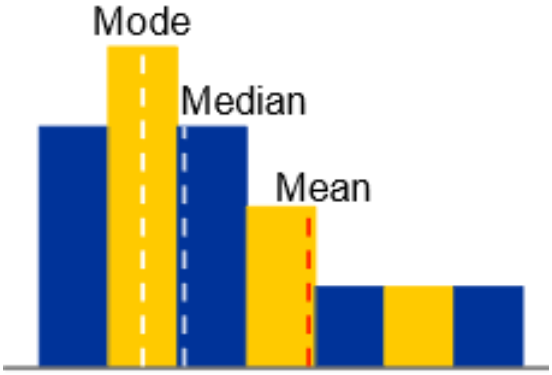
Symmetric Normal - left half is (roughly) same as right half.

- Mean, Median, and Mode are approximately equal.



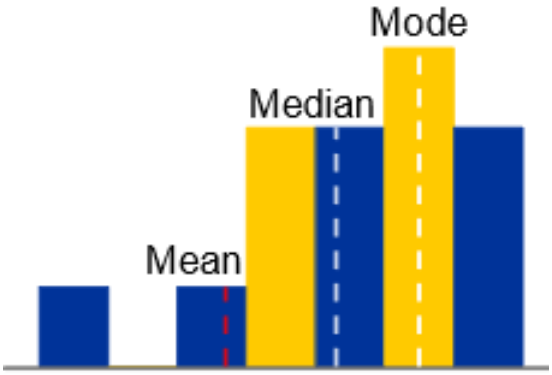
Skewed Right (positively skewed) – tail stretches to the right of the peak.

- Mode < Median < Mean.



Skewed left (negatively skewed) – tail stretches to the left of the peak.

- Mean < Median < Mode.

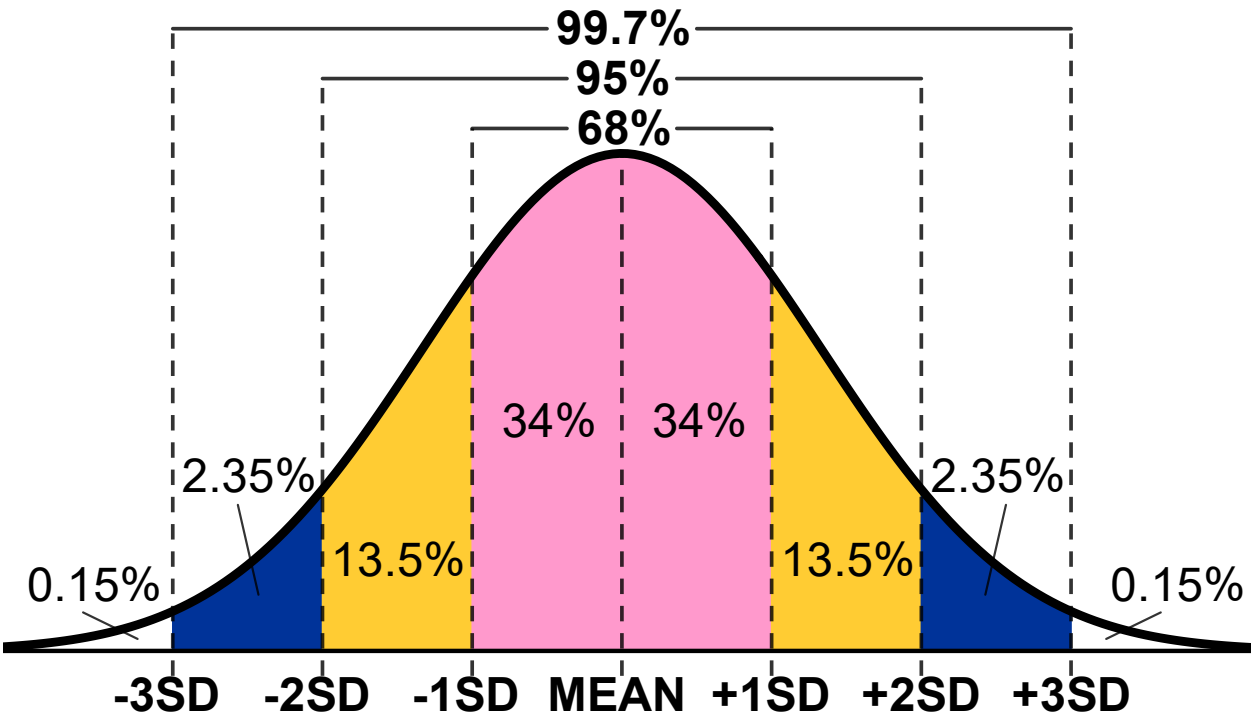


Measures of Center: Page 4.05.1 - value which represents the "typical" data point in a data set

- **Mode -** value that occurs most often in a data set
- **Median -** halfway point, equal number of data points above the median as below, *always order the data from smallest to largest first*
- **Mean** (*common average*) - add up all the data points and divide by how many data points there are

---

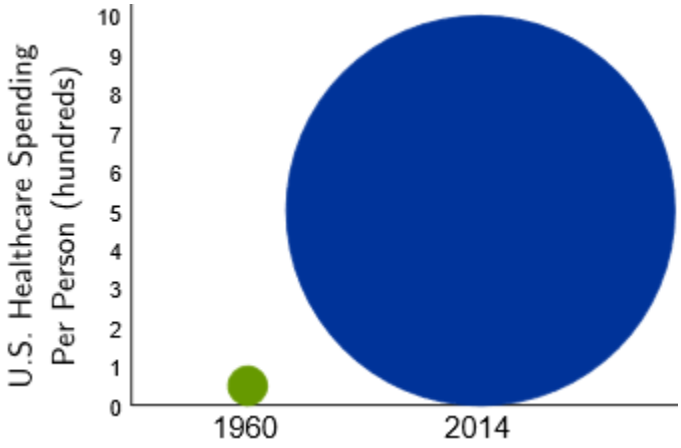
Standard Deviation Rule: Page 4.05.2 (68-95-99.7 Rule) - for Normal Distributions (*bell shaped curves*)



- 68% of the data is within 1 standard deviation of the mean.
- 95% of the data is within 2 standard deviations of the mean.
- 99.7% of the data is within 3 standard deviations of the mean.

Misrepresenting Data with Graphical Displays: Page 4.09

- **Scale of Axis**- The vertical scale should start at zero. Each axis should have consistent scaling (For example, **don't use** 10, 60, 70, 80, 90 for an axis)
- **Omitting Labels or Units**- leaves size and categories unspecified
- **Using a 2-Dimensional Graph to Represent a 1-Dimensional Measurement**- In graphs like the one below, our eyes see area, which distorts the true differences we are trying to illustrate, the heights of each circle. **Avoid using such graphs!**



Module 5: Descriptive Statistics for Two Variables

Relationship Between Two Variables: Page 5.02

- **Explanatory Variable** – Influences the response variable.
- **Response Variable** – Is affected by the explanatory variable.

Graphical Displays of Two Variable Data: Pages 5.03, 5.04, 5.05

Variable Type: C → C

Graphical Display: Two Way Table

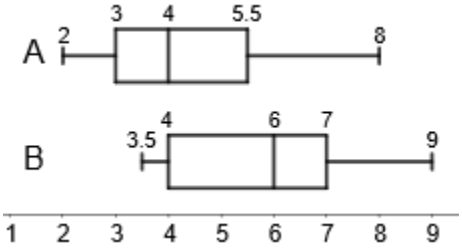
Numerical Measure: Conditional Percentages

	Smoker	Non Smoker	Total
Adult	70	43	113
Child	9	92	101
Total	79	135	214

Variable Type: C → Q

Graphical Display: Side by Side Boxplots

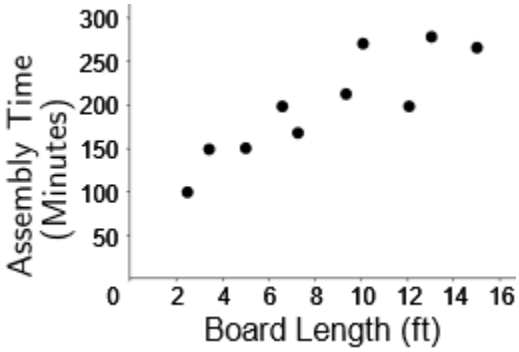
Numerical Measure: Five Number Summary



Variable Type: Q → Q

Graphical Display: Scatterplot

Numerical Measure: Correlation Coefficient (r value)

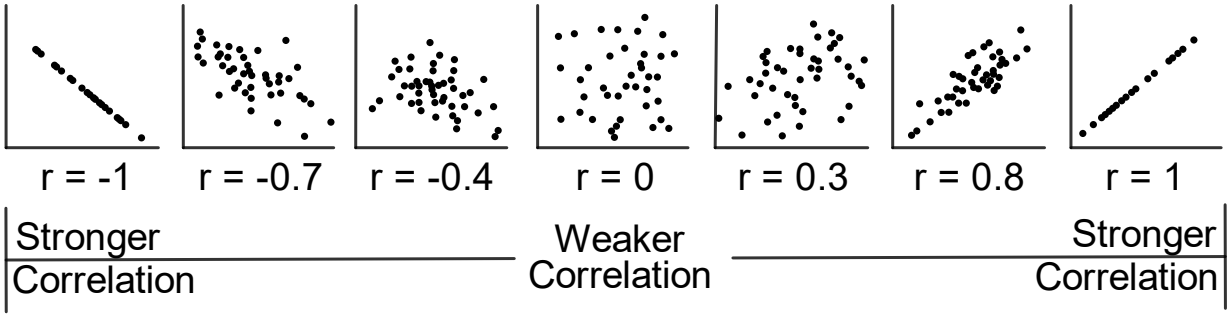


Correlation (Q → Q): Pages 5.05, 5.06, 5.07, 5.08, 6.09

- **Direction:**
  - **Positive Correlation** – scatterplot reveals an "uphill trend." As the explanatory variable increases, the response variable *increases.*
  - **Negative Correlation** - scatterplot reveals a "downhill trend." As the explanatory variable increases, the response variable *decreases.*
  - **No Correlation**- scatterplot reveals no trend between the variables

- **Strength:** On a scatterplot, the closer the points are laid out in a line, the stronger the correlation.
  - **Correlation Coefficient** (**r**) - measures the direction and strength of the *linear* relationship between the variables
    - The closer **r** is to +1, the stronger the positive correlation.
    - The closer **r** is to -1, the stronger the negative correlation.
    - The closer **r** is to 0, the weaker the correlation.

  - **Examples:**



Module 4/5 Summary of Graphical Displays: Pages 4.03-4.04.3, 4.07, 5.03-5.05

- **Just C** – Pie chart or bar chart
- **Just Q** – Histogram, stem plot, boxplot, or dot plot
- **C → C** – Two-way table with Conditional Percentages
- **C → Q** – Side-by-side boxplot with 5-number summary
- **Q → Q** – Scatterplot with correlation coefficient

Module 6: Correlation & Regression

Sampling Methods: Page 6.02, 6.02.1

- Collecting Data:
    - **Population** - the collection of people or objects that is the main focus of the research being done.
    - **Sampling Frame** - the list of people or objects that can *potentially* be included in the study
    - **Sample** - the subset of the sampling frame that is *actually* being studied

- Bias occurs when the Sampling Frame does not accurately represent the Population
    - Example: A manager wanted to know if all of their employees were satisfied with the company. They sent out a survey to all of the part-time employees asking them to rate their satisfaction from 1 to 10, 10 being the most satisfied.
        - This introduced bias since the population was **all** employees, but their sampling frame was only the **part-time** employees.

Study Design: Page 6.02

- **Experimental Study** - Researchers randomly assign participants to two or more groups. One group is designated as a control group where no treatment (placebo) is given while all other groups are given treatments to determine if there is **causation** between variables.
- **Observational Study** - There are no treatment or control groups because the participants self-select their groups. Researchers observe if there is an **association** between variables.

Association Vs Causation: Page 6.03

- **Association** means there is a relationship between two variables. Association does not necessarily imply causation.
    - We can use scatterplots to visualize the data and determine if there is at least an association, but we cannot determine causation from a scatterplot alone.
    - Can establish association through an observational study.
- **Causation -** A change in one variable creates a change in the other variable.
    - Can only be determined from an experiment.

Lurking Variables: Page 6.04

- A **lurking variable** is a variable not included in the study, but affects the variables that were included in the study
- Never assume that a causation exists just because there is an association between two variables – always be on the lookout for lurking variables.

Simpson's Paradox: Page 6.05

- A counterintuitive situation that occurs when a result that appears in individual groups of data disappears or reverses when the groups are combined.
- Can only occur when the sizes of the groups are inconsistent
- Example:

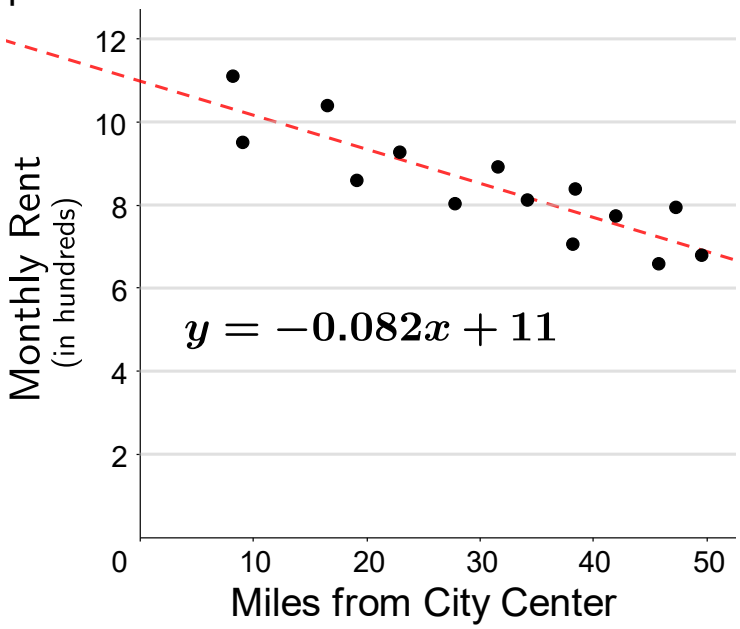
	Passed Prep Course A	Passed Prep Course B
Cohort #1	6/10 = 60%	580/1000 = 58%
Cohort #2	29/100 = 29%	1/10 = 10%
Total	35/110 = 31.2%	581/1010 = 57.5%

    - Prep Course A had a higher passing percentage in Cohort #1 and Cohort #2, but overall Prep Course B had a higher passing percentage.

Regression Analysis: Pages 6.06, 6.07, 6.07.1

- **Simple linear equation** (***regression line*** or line of best fit) - models the data on a scatter plot with a line
  - x is the explanatory variable, and y is the response variable
  - Equation is given by y = mx + b where m is the slope and b is the y-intercept

- Used to **predict** data
  - Plug explanatory values in for x and calculate corresponding response values for y.
  - Example:



    - Using the linear regression above, we can predict that the monthly rent, y, for a home 50 miles from the center will be about: $y = -0.082(50) + 11 = -4.1 + 11 = 6.9$ or $690.

Module 7: Probability

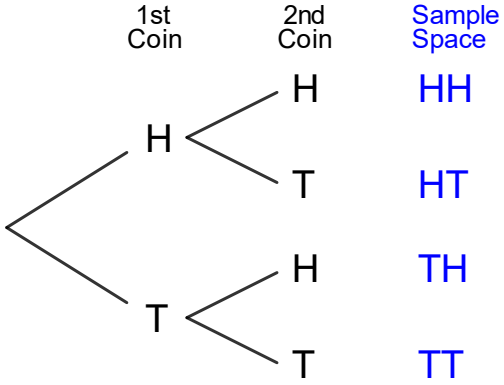
Probability: Page 7.02

- **Probability** is the chance of an event occurring and can be expressed as a percentage, decimal, or fraction.
  - The notation P(E) to represent "the probability of event E."
  - **Example:** The probability of flipping a tails on a coin can be represented as:
    - P(T) = 50% = 0.5 = $\frac{1}{2}$
- 0% probability – **Impossible**
- More than 0%, but less than 40% probability – **Unlikely**
- 40% to 60% probability – **As likely as unlikely**
- more than 60%, but less than 100% probability – **Likely**
- 100% probability – **Certain**

Sample Spaces and Probability: Pages 7.04 7.05, 7.05.1

- Theoretical Probability = $\frac{\text{Number of outcomes with the desired event}}{\text{Total number of outcomes}}$

- **Sample Space** - set of all possible outcomes.
- **Tree Diagram** – Used to determine the sample space.
  - **Example**: The sample space for flipping 2 coins can be found by created the diagram below. In the diagram H represents heads and T represents Tails.



    - The sample space is {HH, HT, TH, TT} and the total number of outcomes is 4.

Complementary Events: Page 7.07

- **Complementary events** are those that do not have any common outcomes and when combined they comprise the sample space.
    - **P(not A) = 1 – P(A)**

Probability Formulas: Pages 7.08, 7.09, 7.11.1, 7.11.2

- **Notation**
    - **P(A or B)** represents the probability that event A will occur, or event B will occur, or both A and B will occur.
    - **P(A and B)** represents the probability that events A and B will occur at the same time.
    - **P(A|B)** represents the probability that event A will occur, given that event B has already occurred.

- **Vocabulary**
    - **Disjoint Events** cannot occur at the same time.
        - P(A and B) = 0
        - Example:
            - A = Randomly selecting a person with type B blood.
            - B = Randomly selecting a person with type O blood.
    - **Independent Events** – We say events A and B are independent if the occurrence of one of them does not affect the probability that the other will occur.
        - P(A|B) = P(A) and P(B|A) = P(B)
            - "probability of A will be the same whether or not B has already occurred. Also, probability of B will be the same whether or not A has already occurred."
        - Example:
            - A = Flipping a coin and landing on tails
            - B = Rolling a die and landing on 3

- **Formulas**
  - **OR Rule (General Addition)**
    - P(A or B) = P(A) + P(B) – P(A and B)
    - Simplifies to P(A or B) = P(A) + P(B) for disjoint events
  - **AND Rule (General Multiplication)**
    - P(A and B) = P(A) x P(B|A)
    - Simplifies to P(A and B) = P(A) x P(B) for independent events
  - **Conditional Probability**
    - $P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$