# How Do You Teach Stats?


Chad Shepherd

Mathematics Teacher/ Department Chair

Pontiac Township High School

Pontiac IL

cshepherd@pontiac90.org

# OBJECTIVES

2-3 Statistics and Probability Activities

Focus on the Statistics/Probability Standards for Grades 9-11

# In general.....

## Year 1:

Students will use statistics to compare two or more different data sets, including the use of scatter plots, tograms, box plots, and standard deviation. Students will terpret outliers and recognize relationships and trends in data. They will use technology to examine data sets and tential models, as well as support the appropriate choice of model.

# The Distribution of the Rainbow

- **S.ID.1** Represent data plots on the real number line (dot plots, histograms, and box plots).
- **S.ID.2** Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.
- **S.ID.3** Interpret difference in shape center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).

## PART A: Collect the Data on Favorite Colors

Go to the board and rate your favorite color of Skittles. After every student has gone to the board, record the sum for each color below.

### Top Rated Skittles Colors of the Class

**Red:**

**Orange:**

**Yellow:**

**Green:**

**Purple:**

1. What was the most popular Skittle? What was the least popular Skittle?

2. Are there any outliers? Explain.

3. Do you feel that there is a relationship between our favorite Skittle and the amount that we will find in each bag? Explain.

## PART A: Collect the Data on Favorite Colors

Go to the board and rate your favorite color of Skittles. After every student has gone to the board, record the sum for each color below.

### Top Rated Skittles Colors of the Class

**Red:** 7

**Orange:** 2

**Yellow:** 1

**Green:** 4

**Purple:** 13

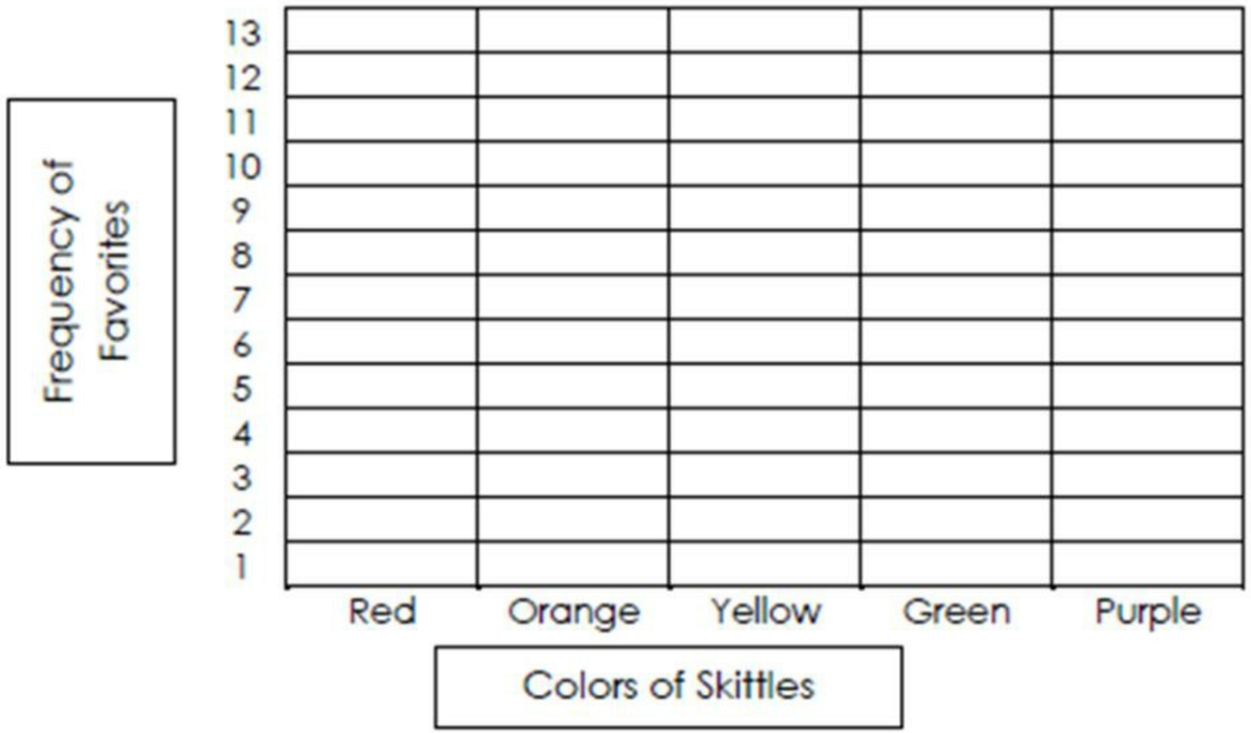
1. What was the most popular Skittle? **Purple** What was the least popular Skittle? **Yellow**

2. Are there any outliers? Explain. **(Answers vary based on your results). Yes, there is a big gap between the number of people who prefer purple over other colors.**

3. Do you feel that there is a relationship between our favorite Skittle and the amount that we will find in each bag? Explain. **Yes, but I don't think it will be skewed as much as our class recorded.**

## PART B: Visual and Analyze the Qualitative Data

Create a **histogram** illustrating the favorite Skittle colors of the class that you recorded on the previous page. Make sure to label your graph accordingly.



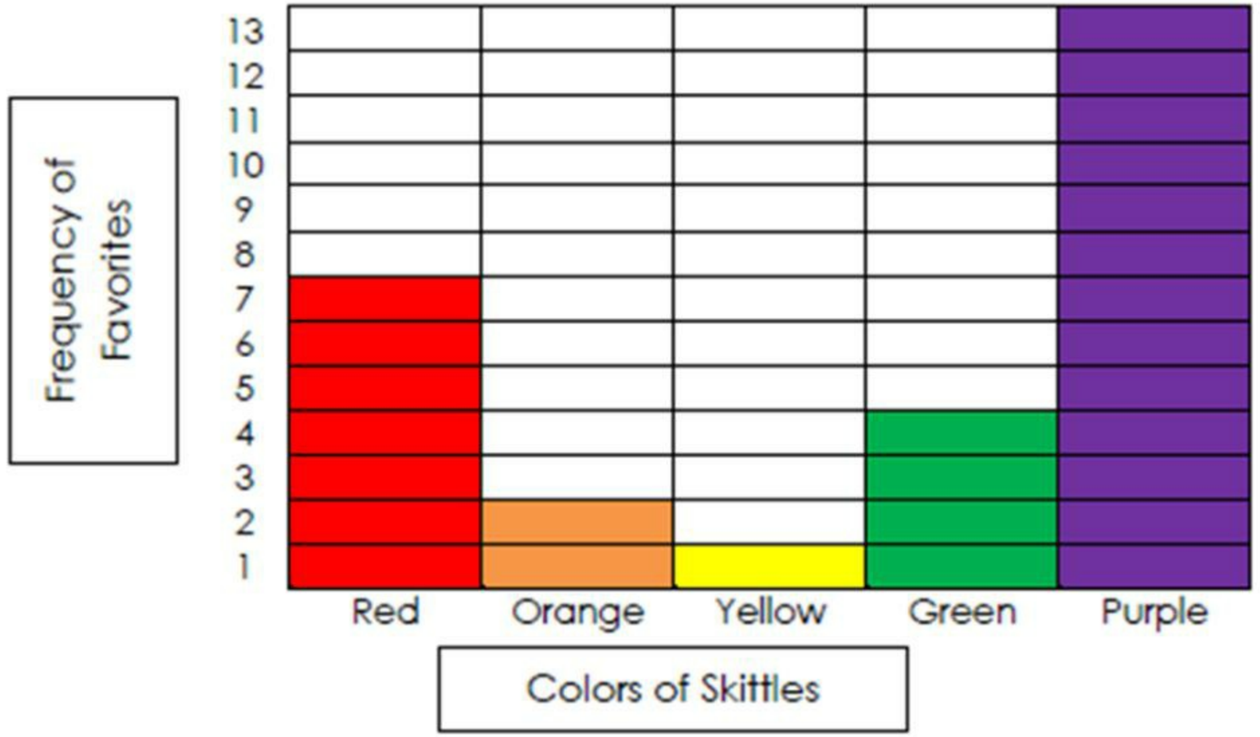
1. Identify the following measures of central tendency based on the number of students who prefer each color.

| Maximum Frequency | Minimum Frequency | Mode (Color) |
|---|---|---|
| | | |

2. Explain what the range of this situation represents.

## PART B: Visual and Analyze the Qualitative Data

Create a **histogram** illustrating the favorite Skittle colors of the class that you recorded on the previous page. Make sure to label your graph accordingly.



1. Identify the following measures of central tendency based on the number of students who prefer each color.

| Maximum Frequency | Minimum Frequency | Mode (Color) |
|---|---|---|
| 13 | 1 | Purple |

2. Explain what the range of this situation represents.

The range represents the difference in the number of people who preferred purple over yellow.

## PART C: Collecting Data on the Distribution of the Colors

1. Open your bag of Skittles and record how many of each color is in your bag.

### Total Colors of Skittles Counted from Your Bag

| Red: | Orange: | Yellow: | Green: | Purple: | Total: |
|---|---|---|---|---|---|
| | | | | | |

2. Find members of your class with the same color sheet as you, collect the necessary data for your color and complete the information.

## PART C: Collecting Data on the Distribution of the Colors

Open your bag of Skittles and record how many of each color is in your bag.

### Total Colors of Skittles Counted from Your Bag (answers vary)

| Red: 4 | Orange: 3 | Yellow: 4 | Green: 3 | Purple: 5 | Total: 19 |
|---|---|---|---|---|---|

Find members of your class with the same color sheet as you, collect the necessary data for your color and complete the information.

Distribution of the Rainbow

Red

# Color Group Data

| Student Name | Number of my color in the bag |
|---|---|
| | |
| | |
| | |

Distribution of the Rainbow

Orange

# Color Group Data

| Student Name | Number of my color in the bag |
|---|---|
| | |
| | |
| | |

Distribution of the Rainbow

Yellow

# Color Group Data

| Student Name | Number of my color in the bag |
|---|---|
| | |
| | |
| | |

Distribution of the Rainbow

Green

# Color Group Data

| Student Name | Number of my color in the bag |
|---|---|
| | |
| | |
| | |

Distribution of the Rainbow

Purple

# Color Group Data

| Student Name | Number of my color in the bag |
|---|---|
| | |
| | |
| | |

## PART D: Organizing Data on the Color Group

1. Identify the following measures of central tendency based on the number of students who prefer each color.

| Q1 | Median | Q3 | Maximum | Minimum | IQR | Mode |
|---|---|---|---|---|---|---|
| | | | | | | |

2. Create a Box-plot with your information above

Distribution of the Rainbow

## PART D: Organizing Data on the Color Group

1. Identify the following measures of central tendency based on the number of students who prefer each color.

| Q1 | Median | Q3 | Maximum | Minimum | IQR | Mode |
|---|---|---|---|---|---|---|
| 3 | 3.5 | 5 | 6 | 1 | 2 | 3 |

2. Create a Box-plot with your information above. Suppose you are the Red Color Group. They will make a box-plot based on their assigned color.

## PART 2: Compare & Contrast the Data

### PART A: Record the results from all other groups.

| | Red | Orange | Yellow | Green | Purple |
|---|---|---|---|---|---|
| Maximum: | 6 | 6 | 6 | 6 | 6 |
| Minimum: | 1 | 1 | 2 | 2 | 3 |
| Median: | 3.5 | 3 | 4 | 3 | 4.5 |
| Mode: | 3 | 2 | 4 | 3 | 4 and 5 |
| $Q_1$: | 3 | 2 | 3 | 3 | 4 |
| $Q_3$: | 5 | 4 | 5 | 4 | 5 |
| Inter Quartile Range: | 2 | 2 | 2 | 1 | 1 |

**PART B: Create FIVE separate box plots** illustrating the central tendencies from each of the colors above. Make sure to label your plots accordingly. (Minimum, $Q_1$, Median, $Q_3$, & Maximum)

**Red**


**Orange**


**Yellow**


**Green**


**Purple**

## PART C: ANALYZE THE DATA

1. Compare and contrast the data from your color and the other colors? Describe your observations.

2. Do you feel there is a correlation between the colors of Skittles distributed vs. the colors that are favored by the consumer? Explain.

3. If you were the distributor, how can you utilize statistics to help you ensure customer satisfaction?

4. Do you feel the data you collected accurately represents the most popular Skittles flavors among all consumers? Explain your reasoning.

## PART C: ANALYZE THE DATA

1. Compare and contrast the data from your color and the other colors? Describe your observations.

   Answers will vary depending on the data. For example: The ranges seem to be similar from color to color. There aren't any outiers. For the greens, the Q1 and the median were the same value. There seems to be about 18-19 Skittles in each back. Each color was represented ranging from 1-6 pieces of a particular color.

2. Do you feel there is a correlation between the colors of Skittles distributed vs. the colors that are favored by the consumer? Explain.

   Answers will vary depending on the data. For example, the class favored purple Skittles over the other colors. Purple was never the minimum color in each bag. Purple has the higher median and mode. This class didn't favor yellow, but yellow wasn't sparse in our bags. There seems to be a good balance of colors. Therefore, I don't think there is a correlation between our favorite Skittles and the data from our bags.

3. If you were the distributor, how can you utilize statistics to help you ensure customer satisfaction? Answers will vary. For example, the distributor could survey their customers through their website. Based off of the favorites, they can adjust their colors accordingly. If everyone hated a certain color, that color could be eliminated. Likewise, customers could vote on a new color.

4. Do you feel the data you collected accurately represents the most popular Skittles flavors among all consumers? Explain your reasoning. Answers will vary depending on the data. For example, I don't think the most popular Skittles are represented by our data because our class is a small sample size of students around the same age. For more accurate data, the sample would have to broaden.

# Year 3

Students will understand how visual displays and summary statistics relate to different types of data and to probability distributions. Also, they will understand methods of collecting data (including sample surveys, experiments, and simulations), organizing, summarizing, analyzing, and presenting data. Lastly, they will understand how randomness and design affect conclusions.

# *DISTRACTED DRIVING*

This activity is from the
2007 NCSSM Summer Statistics Writing Program.

- **S.IC.1** Understand statistics as a process for making inferences about population parameters based on a random sample from that population.
- **S.IC.2** Decide if a specified model is consistent with results from a given data-generating process, e.g. using simulations.
- **S.IC.4** Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling.
- **S.IC.5** Use data from a randomized experiment to compare two treatments; use simulations to decide if difference between parameters are significant
- **S.IC.6** Evaluate reports based on data

# Distracted driving

Are drivers more distracted when using a cell phone than when talking to a passenger in the car? Researchers wanted to find out, so they designed an experiment. Here are the details.

*In a study involving 48 people, 24 people were randomly assigned to drive in a driving simulator while using a cell phone. The remaining 24 were assigned to drive in the driving simulator while talking to a passenger in the simulator. Part of the driving simulation for both groups involved asking drivers to exit the freeway at a particular exit. In the study, 7 of the 24 cell phone users missed the exit, while 2 of the 24 talking to a passenger missed the exit.* (from the 2007 AP* Statistics exam, question 5)

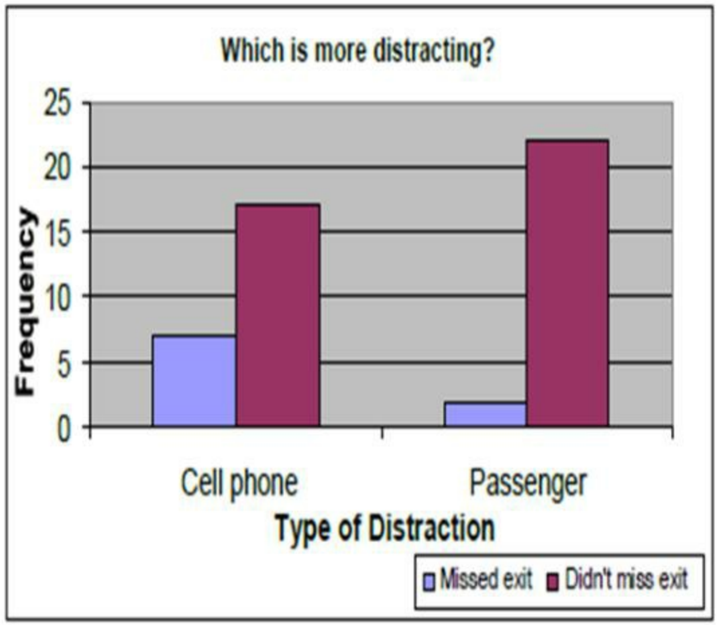
- Let's start by summarizing the data from this study. Each of the 48 people in the experiment can be classified into one of the four cells in the table below based on the experimental condition to which they were assigned and whether they missed the designated exit. Use information from the previous paragraph to complete the table.

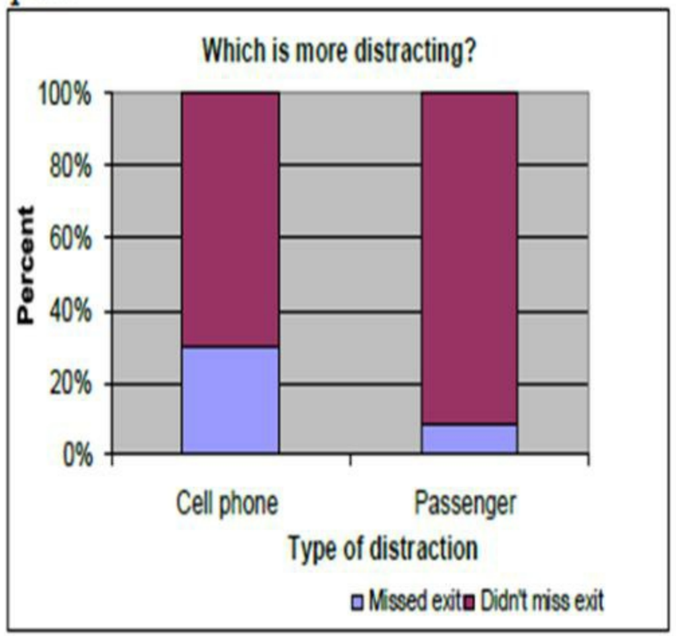
| | | *Distraction* | |
|---|---|---|---|
| | | Cell phone | Passenger |
| *Missed exit?* | Yes | | |
| | No | | |

- Let's start by summarizing the data from this study. Each of the 48 people in the experiment can be classified into one of the four cells in the table below based on the experimental condition to which they were assigned and whether they missed the designated exit. Use information from the previous paragraph to complete the table.

| | | *Distraction* | |
|---|---|---|---|
| | | Cell phone | Passenger |
| *Missed exit?* | Yes | 7 | 2 |
| | No | 17 | 22 |

To analyze data, we begin by making one or more graphs.




- Two types of Excel bar graphs are shown above. Explain the difference in what the two graphs display. Then tell which one you prefer and why.

Next, we add numerical summaries. We might be interested in comparing the counts, percents, or proportions of people in the two groups who missed the freeway exit.

- Fill in the missing entries in the table below for the passenger group.

| | *Missed exit* | | |
|---|---|---|---|
| | Number | Proportion | Percent |
| Cell phone group | 7 | 0.292 | 29.2 |
| Passenger group | | | |

Next, we add numerical summaries. We might be interested in comparing the counts, percents, or proportions of people in the two groups who missed the freeway exit.

- Fill in the missing entries in the table below for the passenger group.

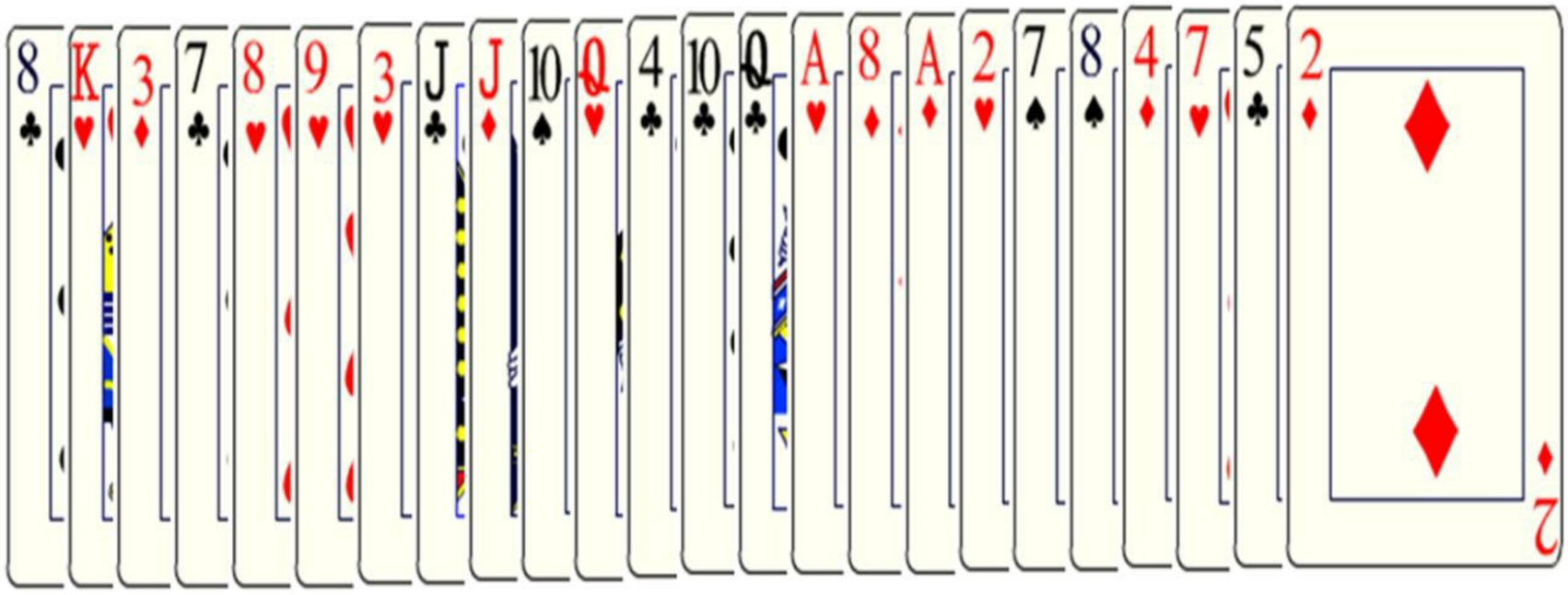
| | *Missed exit* | | |
|---|---|---|---|
| *Group* | Number | Proportion | Percent |
| Cell phone | 7 | 0.292 | 29.2 |
| Passenger | 2 | 0.083 | 8.3 |

# DIFFERENTIAL EFFECT OR SAME EFFECT?

**Activity: Could the observed difference be due to the chance assignment of people to groups?** *Materials: Standard deck of playing cards for each group of 3-4 students*

What would happen if we reassigned the 48 people in this experiment to the cell phone and passenger groups many times, assuming that the group assignment had no effect on whether each driver missed the exit? Let's try it and see.

One possible plan is to take out the J, Q, K, A of spades.
We'll let the other spades represent 'missed exit' drivers.



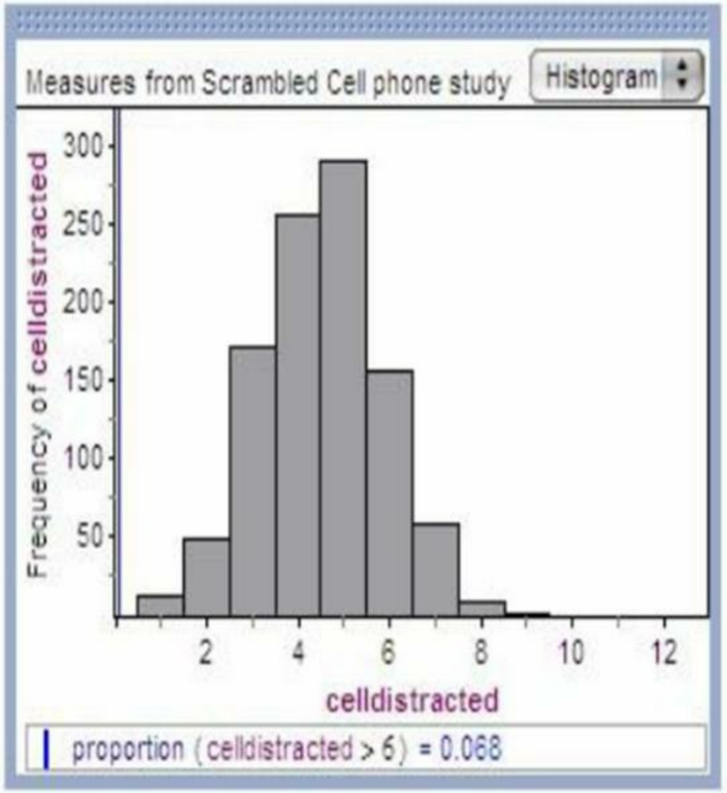
5. Repeat this process 9 more times so that you have a total of 10 trials. Record your results in the table provided.

| Trial | Number who missed exit in cell phone group | Number who missed exit in passenger group |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |

7. Combine results with your classmates. In what percent of the class's simulation trials did 7 or more people in the cell phone group miss the freeway exit?

8. Based on the class's simulation results, do you think it's possible that cell phones and passengers are equally distracting to drivers, and that the difference observed in the original experiment could have been due to the chance assignment of people to the two groups? Why or why not?

Here are the results of 1000 trials of a computer simulation, like the one you did with the playing cards, showing the number of drivers who missed the exit in the cell phone group.




9. In the computer simulation, how often did 7 or more drivers in the cell phone group miss the exit when there is no difference in the effects of the experimental conditions? Do you think the results of the original experiment could be due to chance and not to a difference in the effects of cell phone use and talking to a passenger on driver distraction? Explain your reasoning.

# Year 2

Students will understand independence and conditional probability and use them to interpret data. Students will be able to compute probabilities of independent, dependent, and compound events.

# *HUNGER GAMES PROBABILITY*

- **S.CP.3** Understand the conditional probability of A given B as P (A and B)/P(B), and interpret independence of A and B as saying that the conditional probability of A given B is the same as the probability of A, and the conditional probability of B given A is the same as the probability of B
- **S.CP.5** Recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations.
- **S.CP.6** Find the conditional probability of A given B as the fraction of B's outcomes that also belong to A, and interpret the answer in terms of the model.

***Introduction:***

*In the book and movie series, The Hunger Games ©, the author sets up a futuristic society that is separated into 12 districts and a capital. A rule of this society is that each of the 12 districts has to provide one male and one female tribute from the ages of 12-18 to fight to the death in an arena each year.*

*Only one individual from the tribute pool will survive the games and be crowned the champion. The individuals are chosen during a ceremony called "The Reaping." The reaping process is as follows:*

- *When an individual turns 12, their name is placed into the drawing one time*
- *Every birthday afterwards, their name is added into the drawing one additional time so, at the very least, an 18 year old would have their name in the drawing 7 times*
- *If an individual chooses to, they may exchange additional entries into the drawing for rations (food). Generally, the poorer a person is, the more entries that person has in the drawing*

*This year, the following entry count for District 12 was obtained:*

a. *There are 252 children between the ages of 12-18*
b. *There are 784 names in the female drawing*
c. *There are 978 names in the male drawing*

**Part I: Let's use the above situation and this year's entry data to calculate some probabilities:**

1. Calculate the chances of Katniss (the main character of the story) being chosen, if she is 16 years old and has 15 ration entries in the drawing.

   *P*(Katniss) = _____

2. If Prim, Katniss's little sister, just turned 12 and has no ration entries, what is the probability that she will be chosen?

   *P*(Prim) = _____

3. Gale, Katniss's best guy friend, is 18 years old. Because he has been single-handedly supporting his family since turning 12, he has exchanged rations for 35 additional entries. What are his chances of being chosen during The Reaping?

   *P*(Gale) = _____

4. Peeta, who is relatively well off, has no ration entries. If he is 16 years old, what are the chances that his name would be chosen?

   *P*(Peeta) = _____

5. Using what you know about the number of entries for Katniss and Prim, what is the probability that Katniss OR Prim is chosen?

   *P*(Katniss or Prim) = _____

6. Using what you know about the number of entries for Gale and Peeta, what is the probability that Gale OR Peeta is chosen?

   *P*(Gale or Peeta) = _____

**Part II: Using the probabilities you calculated in Part I, determine the following:**

7. What is the probability that both Gale and Katniss are chosen?\

   *P*(Katniss and Gale) = _____

8. What is the probability that Katniss and Peeta are chosen

   *P*(Katniss and Peeta) = _____

9. What is the probability that Prim and Katniss are chosen?

   *P*(Katniss and Prim) = _____

10. What is the probability that Prim and Peeta are chosen?

    *P*(Prim and Peeta) = _____

**Part III: After "The Reaping", the two tributes from each of the 12 districts make their way to the capital, are trained to fight, interviewed, and then placed into the arena.**

11. Not taking skill, age, or other factors into consideration, what is each tributes chance of winning the games?

    *P*(win) = _____

12. What is the probability that Katniss was chosen during "The Reaping" and then wins the games?

    *P*(Katniss chosen and Katniss wins) = _____

13. At the initial conflict at the cornucopia, 11 tributes were removed from the arena. What is the probability that a tribute survived this initial conflict?

    *P*(tribute survived) = _____

14. If Katniss survived the initial conflict, what is the probability of Katniss winning at this point?

    *P*(Katniss winning) = _____

15. During the 74th Hunger Games, there was a rule change: The President decided that two tributes from the same district could both win. At this point in the games, there were only 3 pairs of tributes from the same district left. What is the probability that both Peeta and Katniss win at this point?

    *P*(Katniss and Peeta win) = _____

| Tribute District (Name, if applicable) | Male? | Female? | Career? | Age 12-15 | Age 16-18 | Wealthy District? | Poor District? | Volunteer? | Ability Score |
|---|---|---|---|---|---|---|---|---|---|
| 1 (Glimmer) | | X | X | | X | X | | X | 9 |
| 1 (Marvel) | X | | X | | X | X | | X | 9 |
| 2 (Clove) | | X | X | X | | X | | X | 10 |
| 2 (Cato) | X | | X | | X | X | | X | 10 |
| 3 | | X | | | X | | X | | 9 |
| 3 | X | | | X | | | X | | 8 |
| 4 | | X | X | | X | X | | X | 5 |
| 4 | X | | X | X | | X | | X | 8 |
| 5 (Foxface) | | X | | X | | X | | | 7 |
| 5 | X | | | X | | X | | | 5 |
| 6 | | X | | | X | | X | | 9 |
| 6 | X | | | | X | | X | | 8 |
| 7 | | X | | | X | X | | | 7 |
| 7 | X | | | | X | X | | | 9 |
| 8 | | X | | X | | | X | | 8 |
| 8 | X | | | | X | | X | | 7 |
| 9 | | X | | X | | | X | | 5 |
| 9 | X | | | X | | | X | | 6 |
| 10 | | X | | X | | | X | | 4 |
| 10 | X | | | | X | | X | | 8 |
| 11 (Rue) | | X | | X | | | X | | 7 |
| 11 (Thresh) | X | | | | X | | X | | 10 |
| 12 (Katniss) | | X | | | X | | X | X | 11 |
| 12 (Peeta) | X | | | | X | | X | | 8 |

**Using the information above, calculate the following probabilities that a person is chosen at random with the indicated characteristics:**

16. $P(\text{Career} \mid \text{Female}) =$ _____

17. $P(\text{Ability score over 8} \mid \text{Male}) =$ _____

18. $P(\text{Ability score over 8} \mid \text{Female}) =$ _____

19. $P(\text{Ability score over 8} \mid \text{Poor}) =$ _____

20. $P(\text{Not career} \mid \text{Volunteer}) =$ _____

21. $P(\text{Under age 16} \mid \text{Volunteer}) =$ _____

22. $P(\text{Female or Volunteer}) =$ _____

23. $P(\text{Male or Career}) =$ _____

24. $P(\text{Career or Wealthy}) =$ _____

25. $P(\text{Age 12-15 or Poor}) =$ _____

**Part V: Determine whether or not the two events in each situation are independent, based on the information given and the probabilities you have calculated throughout the activity. Explain your answer.**

26. The probability that Prim and Peeta are chosen during The Reaping.

Are the events above independent (circle one)? YES or NO

Explain:

_____

_____

_____

27. The probability of Katniss being chosen and winning The Hunger Games.

Are the events above independent (circle one)? YES or NO

Explain:

_____

_____

_____

28. From the 24 tributes listed in part IV, the probability of choosing a male tribute and an ability score over 8.

Are the events above independent (circle one)? YES or NO

Explain:

_____

_____

_____

# THANK YOU NCTM ORLANDO!

# ANY QUESTIONS????

cshepherd@pontiac90.org