NCTM Regional Conference
Cleveland, Ohio
October 17, 2008
10:30 – noon

The Really Fun Part of Math: Data Analysis and Probability!

Jerry Moreno

John Carroll University

moreno@jcu.edu

www.jcu.edu/math/faculty/MORENO/moreno.htm

What we hope to do:

- The GAISE Report and OACS
  - Guidelines for Assessment and Instruction in Statistics Education

    A Curriculum Framework for Pre-K-12 Statistics Education
- Probability for Level A
  - Using spinners, disks, coins, cards, number cubes, Austin animal crackers
- Data Analysis for Level A
  - Types of data and appropriate graphs
  - Incorrect graphing
  - The Ohio Statistics Poster Competition

OACS Probability …highlights

Very similar to GAISE recommendations

- K-2: Describe the probability of chance events as more, less, or equally likely to occur.
- 3-4: Conduct a simple probability experiment, identify and represent possible outcomes, draw conclusions about their likelihood.
- 5-7:Find all possible outcomes of simple experiments using lists, arrays, tree diagrams; describe their probabilities as fractions.

PROBABILITY

For k-12, the ultimate goal of learning probability should be toward understanding that probability is a tool used to **understand the reasoning of statistical decision making and drawing conclusions** (GAISE Level C). (Unfortunately many state standards look at probability only as a subject in its own right focused on counting techniques.)

How do we get to the ultimate goal?

PROBABILITY – GAISE A

Levels of development of probability in grades k-4:

1. Identifying when one event compared to another is **less likely**, **equally likely**, or **more likely** to happen.

2. Providing a **measure** of how likely an event can happen, **in words**, namely:

Impossible	Unlikely	Equally Likely	Likely	Certain

3. Providing a **measure** of how likely an event can happen, **as a fraction**, for example:

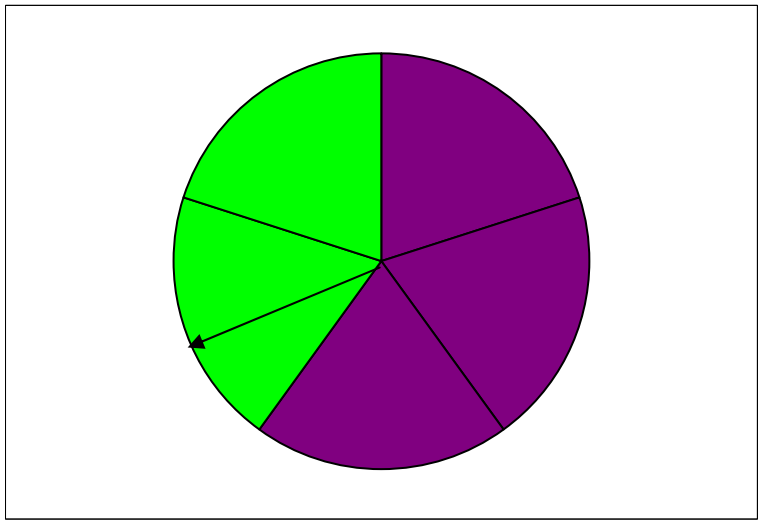
Impossible	Unlikely	Equally Likely	Likely	Certain
0	**1/4**	**1/2**	**3/4**	**1**

4. Estimating probabilities using empirical data.

5. Long-run relative frequency; Law of Large Numbers.

PROBABILITY – First Notion

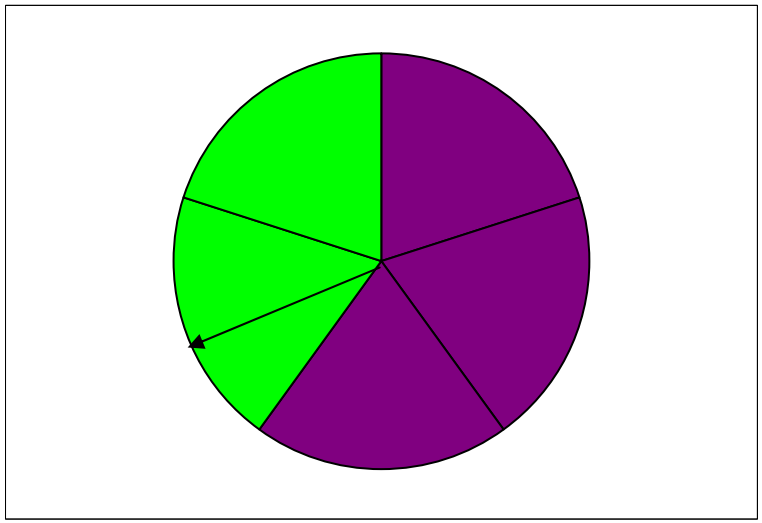
Identifying when one event compared to another is **less likely, equally likely**, or **more likely** to happen.



There are two colors, green and purple. In one spin of the spinner, is it less likely, equally likely, or more likely for the spinner to land on green rather than land on purple? Why?

PROBABILITY – First Notion

Identifying when one event compared to another is **less likely, equally likely**, or **more likely** to happen.

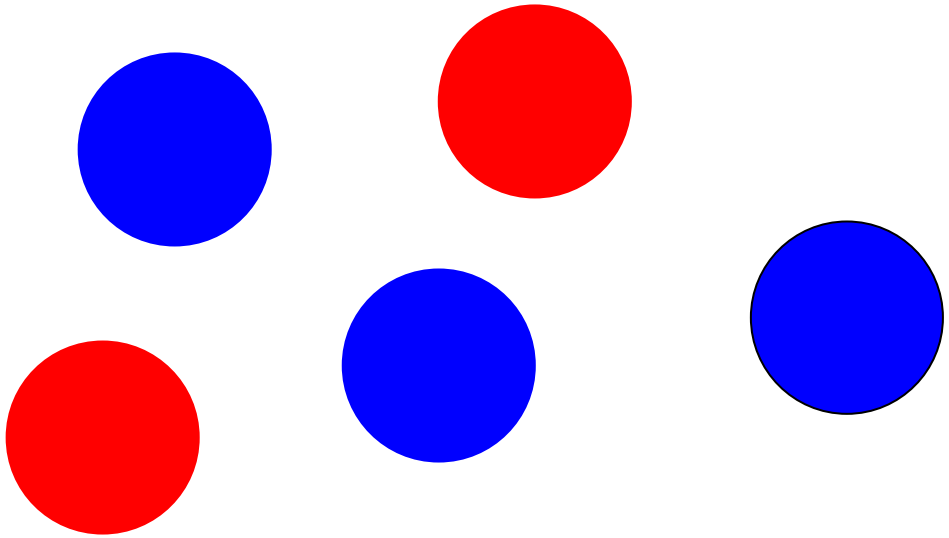


There are two colors, green and purple. In one spin of the spinner, is it less likely, equally likely, or more likely for the spinner to land on green rather than land on purple? Why?

Less likely because there are fewer green sectors (2) than purple ones (3).

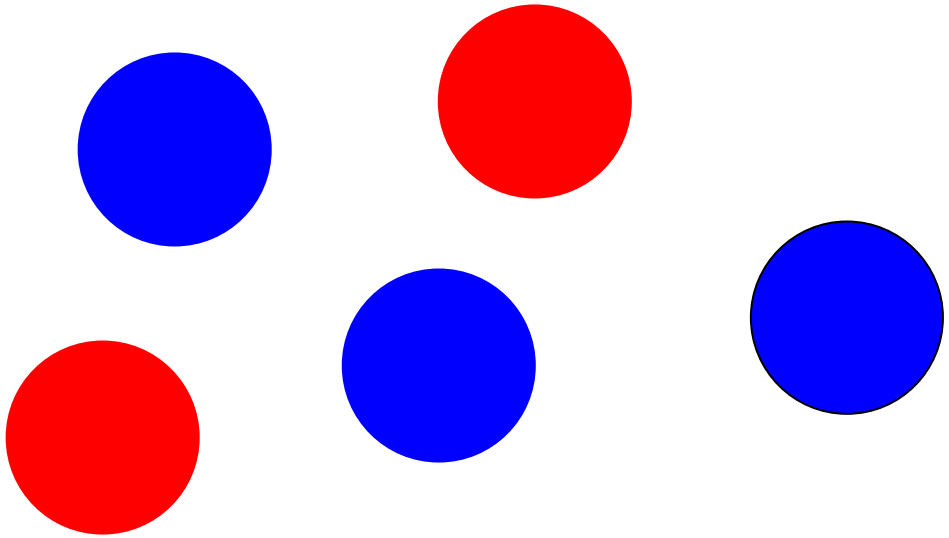
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 chip without looking. Is getting a blue chip less likely, equally likely, or more likely than getting a red chip? Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

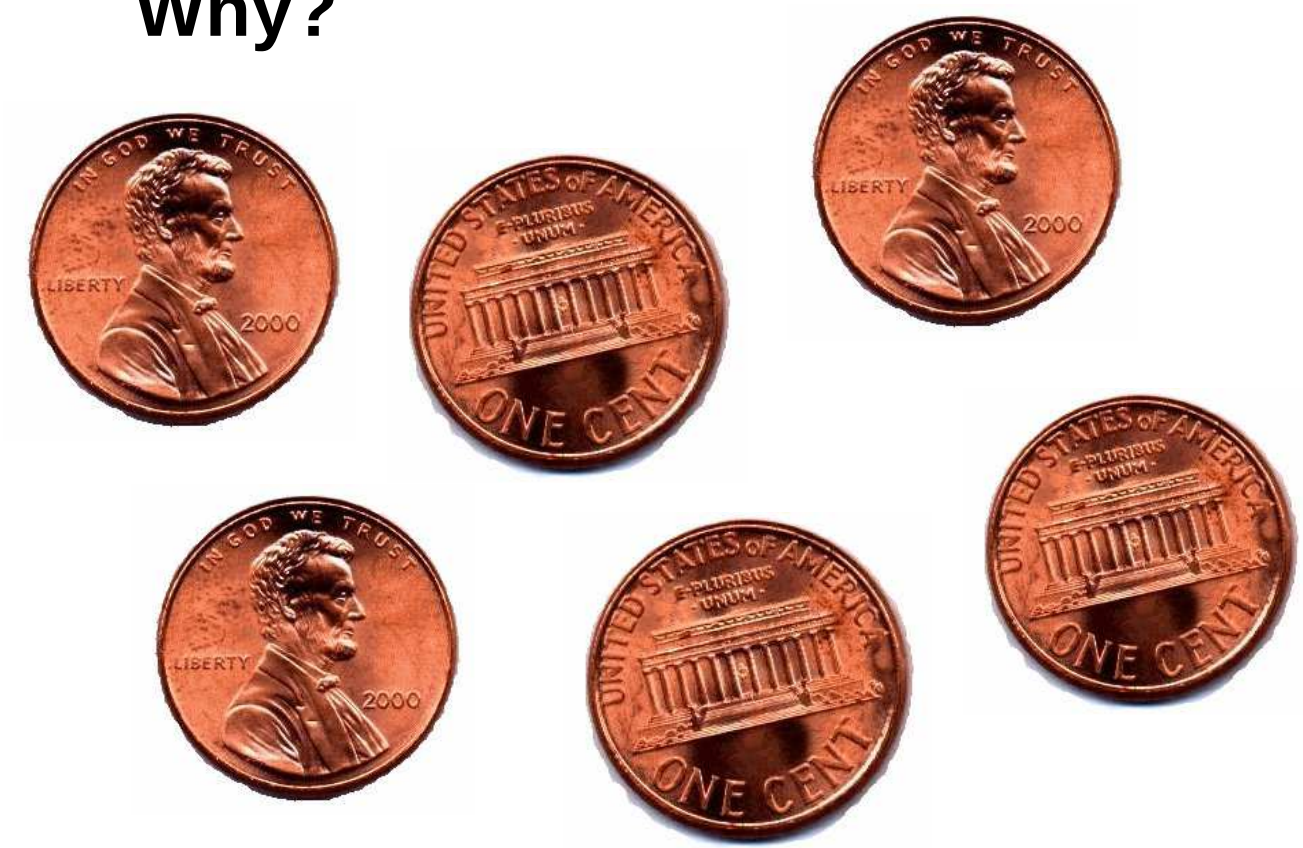
Choose 1 chip without looking. Is getting a blue chip less likely, equally likely, or more likely than getting a red chip? Why?



More likely because there are three blue and two red chips, and three is greater than two.

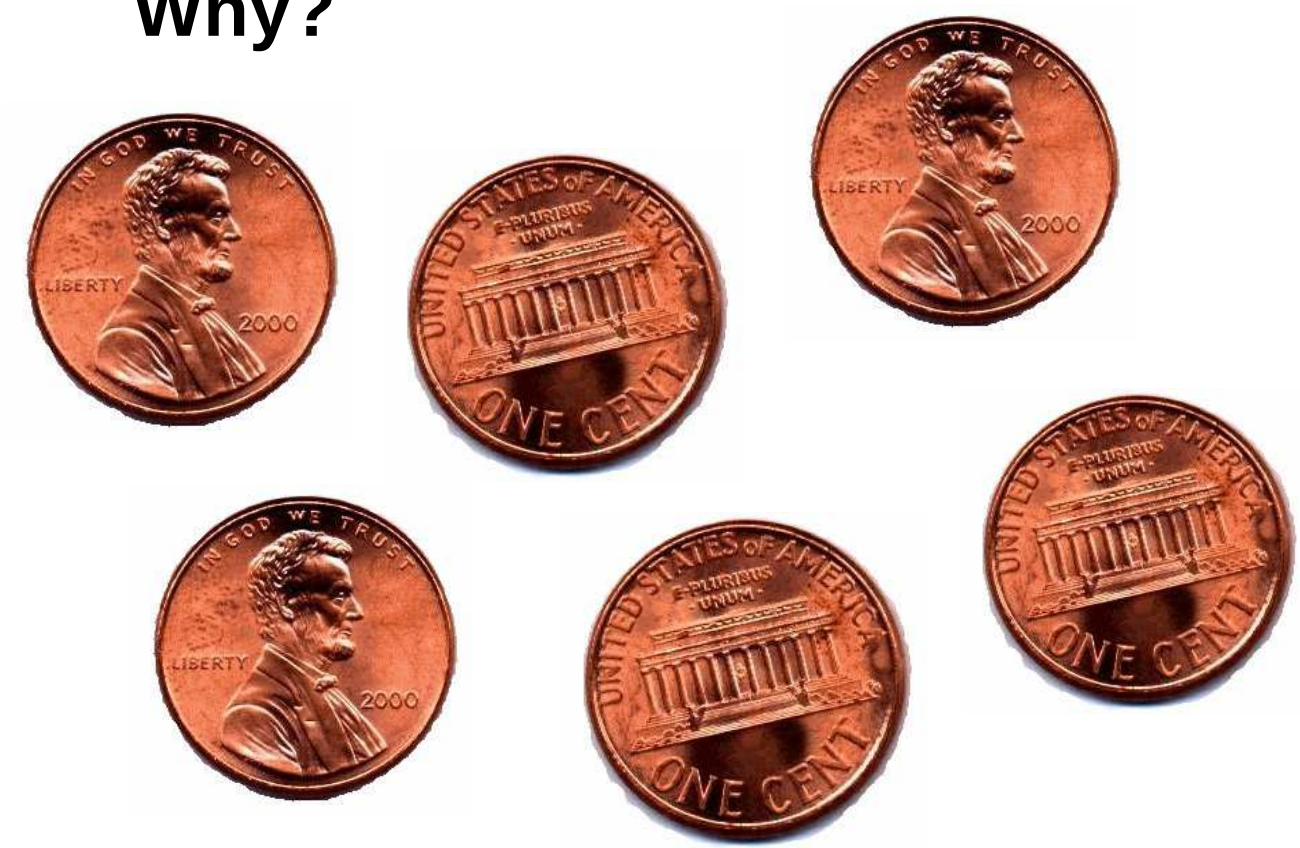
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 penny without looking. Is getting Lincoln Head up less likely, equally likely, or more likely than getting Lincoln Memorial up? Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

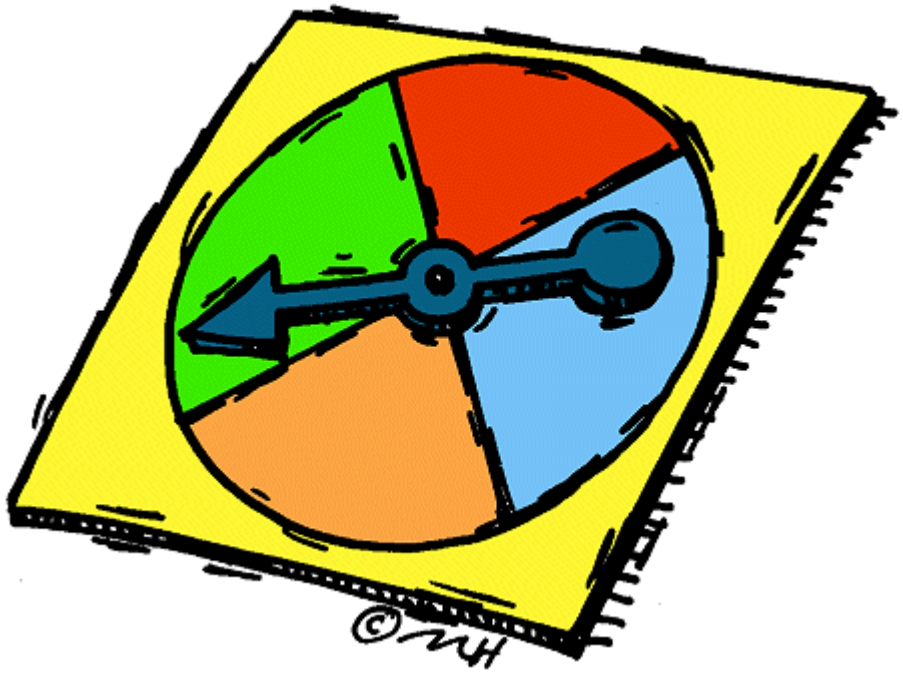
Choose 1 penny without looking. Is getting Lincoln head up less likely, equally likely, or more likely than getting Lincoln Memorial up? Why?



Equally likely because there are three Lincoln heads and three Lincoln Memorials, the same number of each.

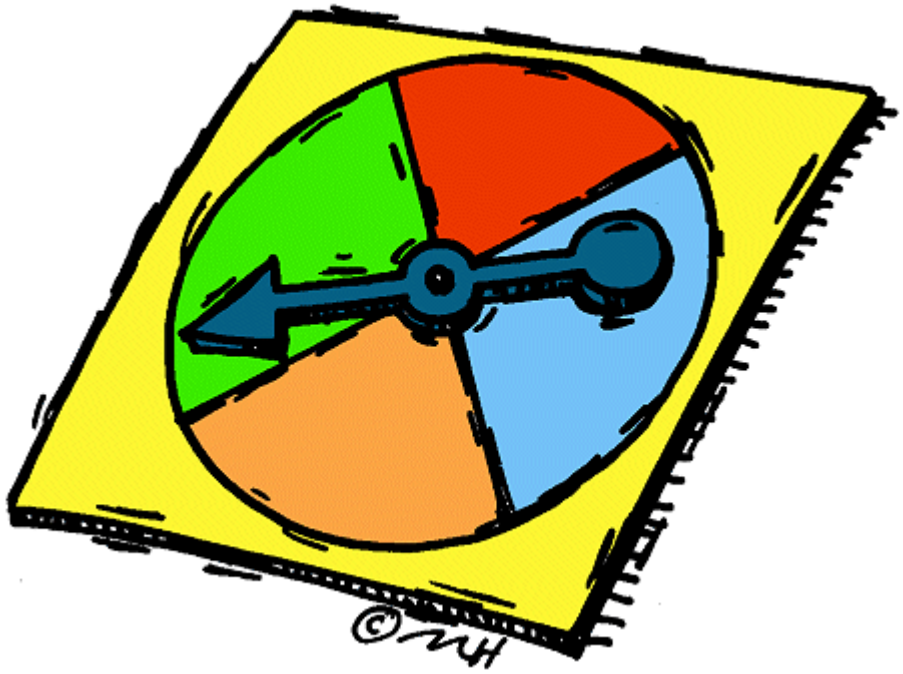
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Spin the spinner once. Is landing on green less likely, equally likely, or more likely than landing on a non-green color? Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

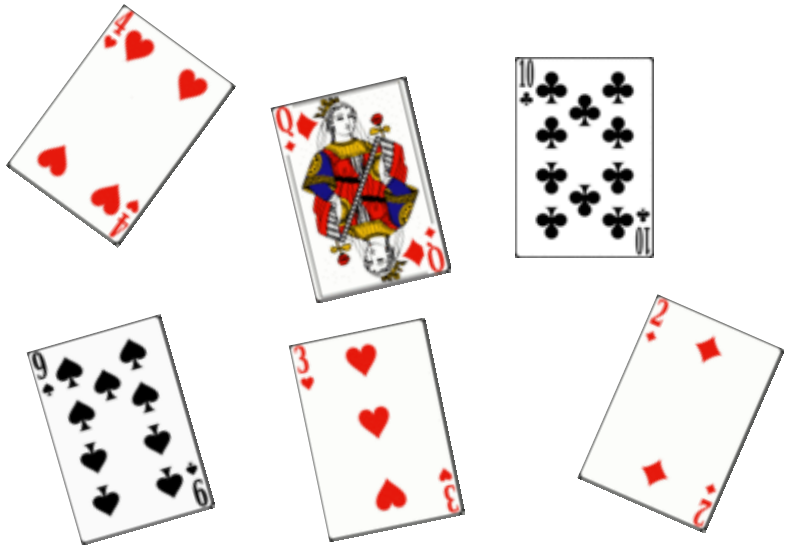
Spin the spinner once. Is landing on green less likely, equally likely, or more likely than landing on a non-green color? Why?



Less likely because there is one green sector and three other equally sized sectors that are not green.

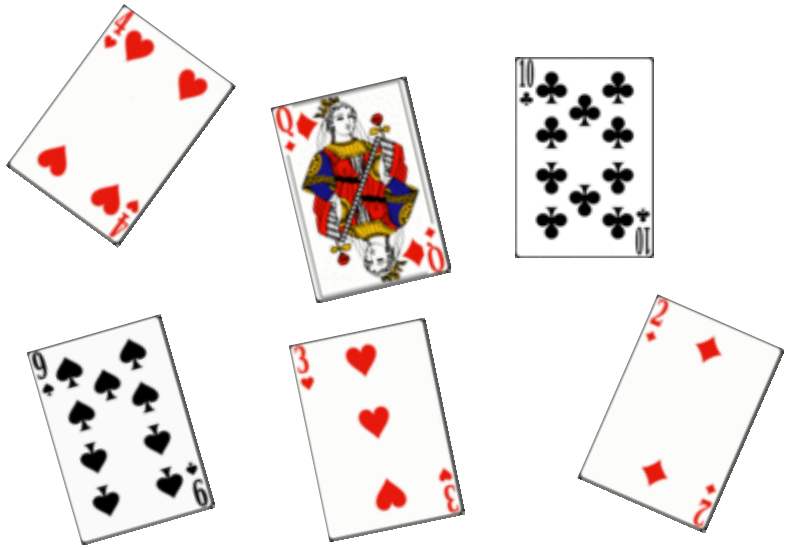
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 card without looking. Is getting a red card less likely, equally likely, or more likely than getting a black card? Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

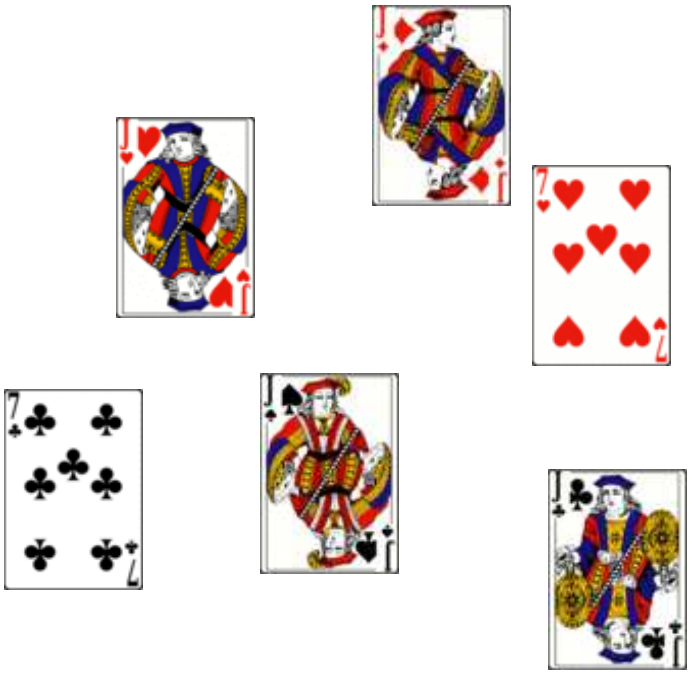
Choose 1 card without looking. Is getting a red card less likely, equally likely, or more likely than getting a black card? Why?



More likely because there are four red cards and only two black cards, and four is greater than two.

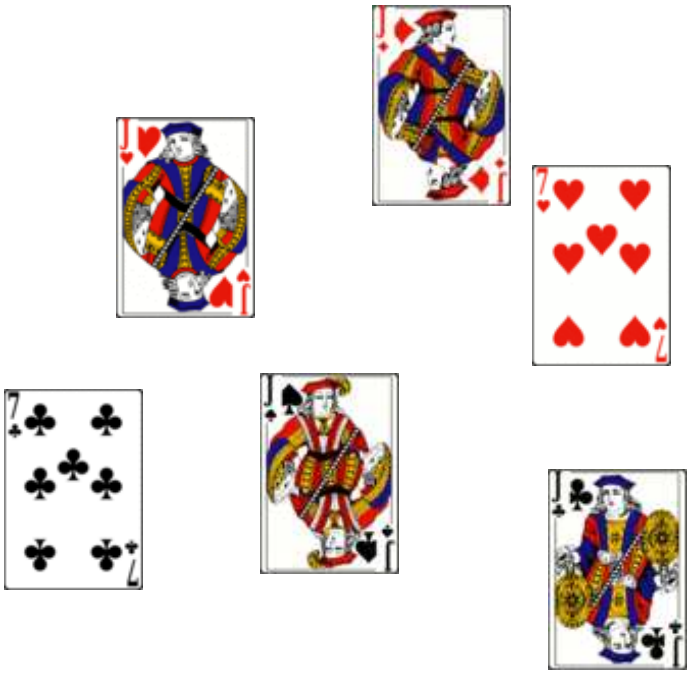
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 card without looking. Is getting a “7” less likely, equally likely, or more likely than getting a “Jack?” Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

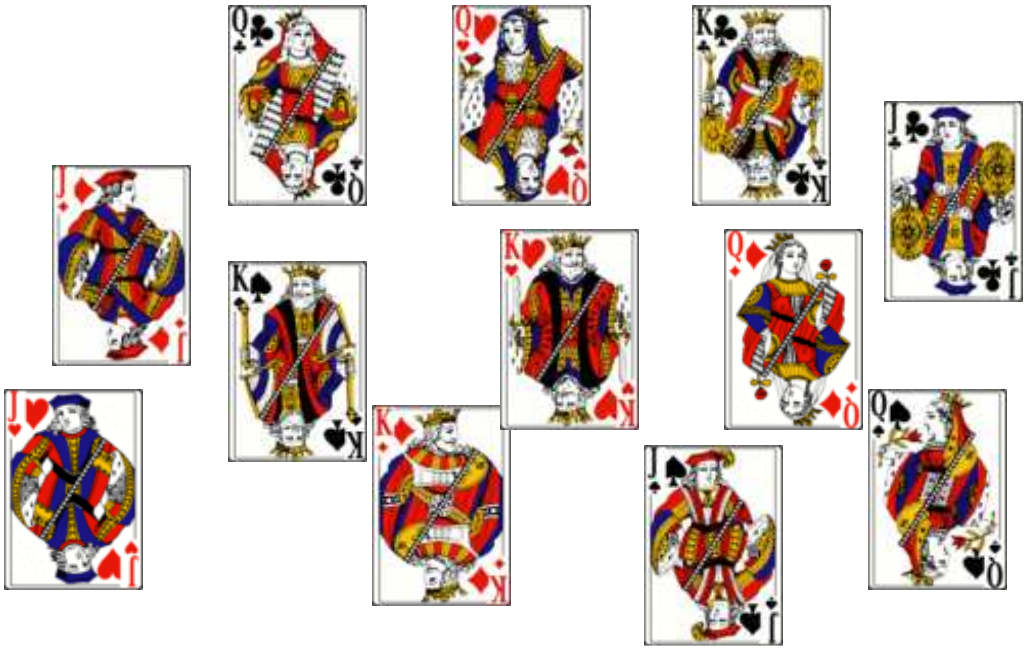
Choose 1 card without looking. Is getting a "7" less likely, equally likely, or more likely than getting a "Jack?" Why?



Less likely because there are only two "7's" but there are four "Jack's" and two is less than seven.

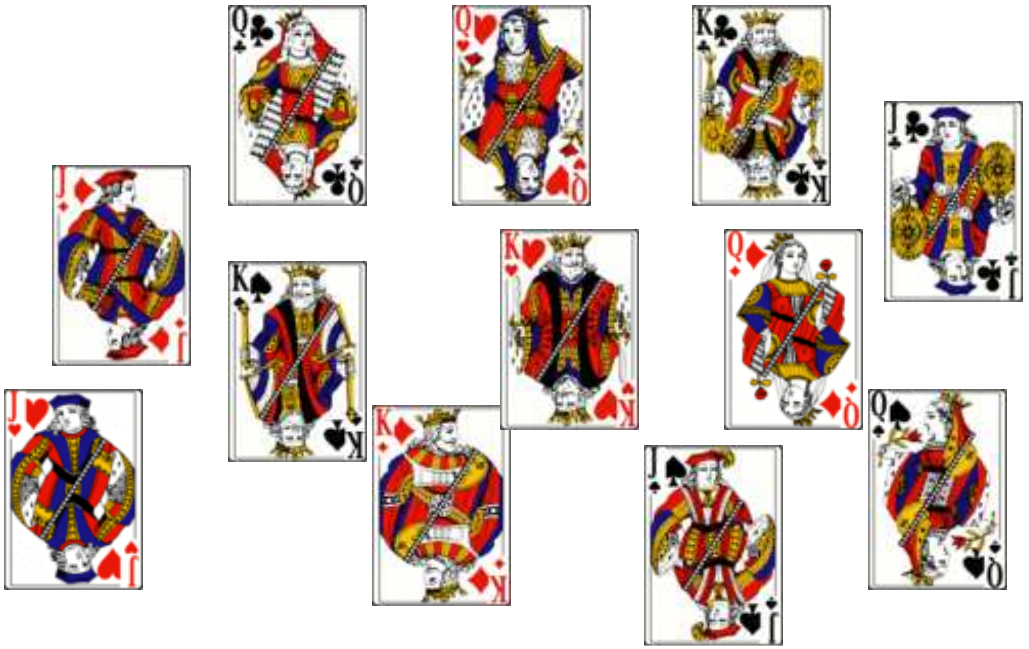
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 card without looking. Is getting a “one-eyed face card” less likely, equally likely, or more likely than getting a “two-eyed face card?” Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

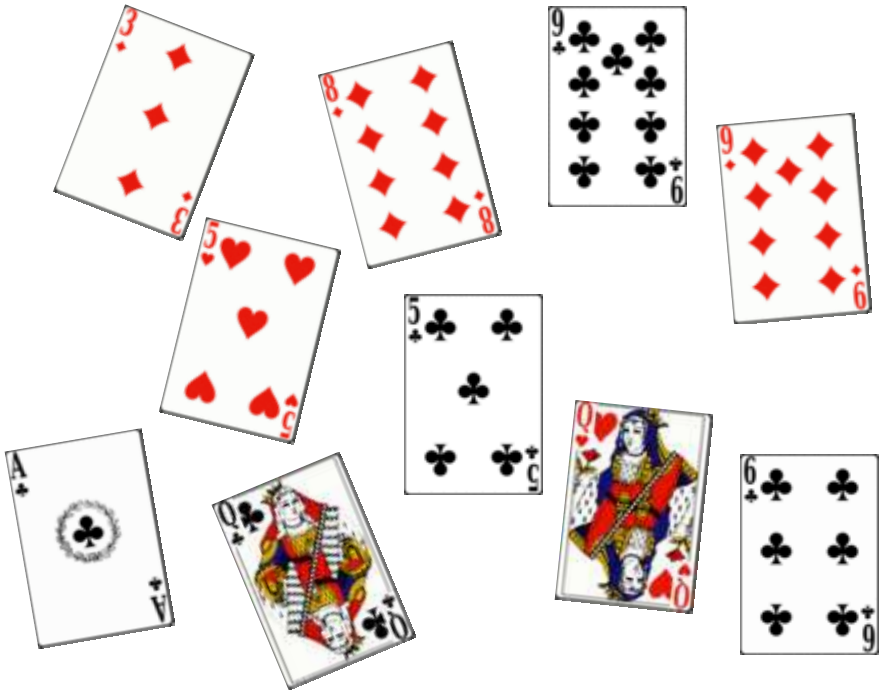
- **Choose 1 card without looking. Is getting a “one-eyed face card” less likely, equally likely, or more likely than getting a “two-eyed face card?” Why?**



Less likely because there are three one-eyed face cards but nine two-eyed face cards.

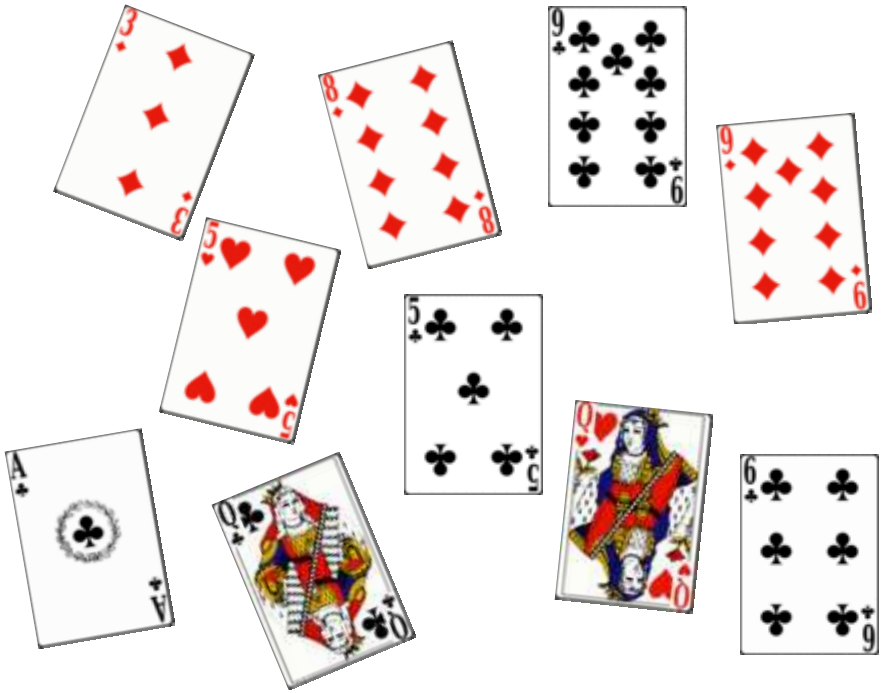
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 card without looking. Which suit is most likely to be chosen? Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

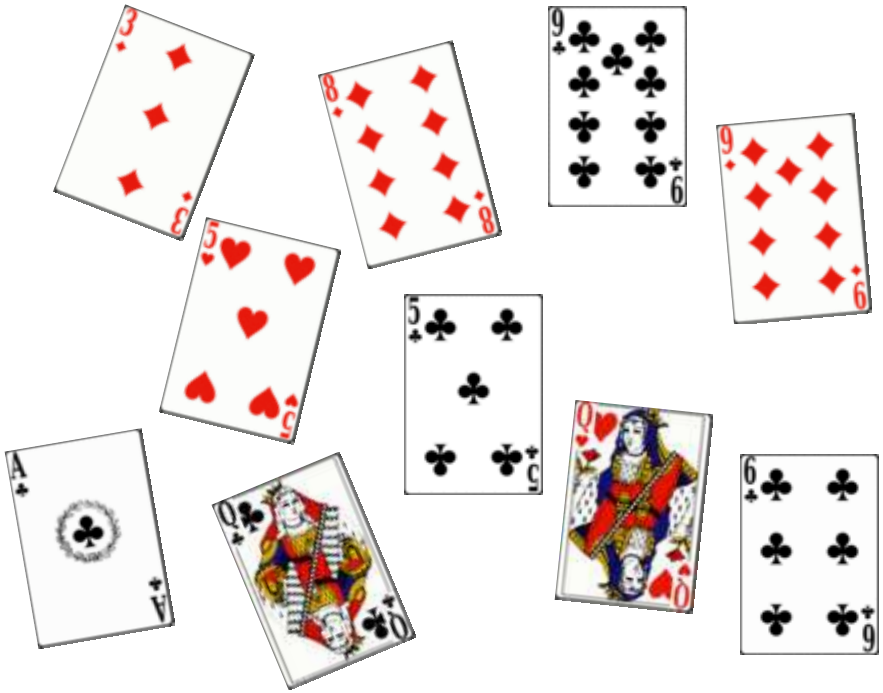
Choose 1 card without looking. Which suit is most likely to be chosen? Why?



Clubs because there are five clubs which is more than the number of diamonds (three) or the number of hearts (two).

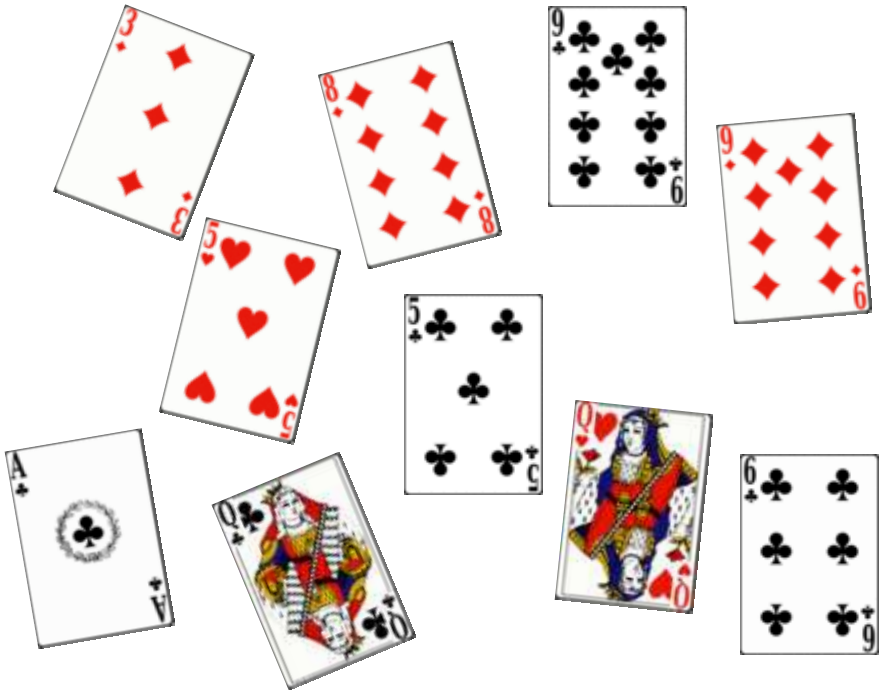
LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 card without looking. Is a club less likely, equally likely, or more likely to be chosen compared to a non-club? Why?



LESS LIKELY, EQUALLY LIKELY, MORE LIKELY

Choose 1 card without looking. Is a club less likely, equally likely, or more likely to be chosen compared to a non-club? Why?



Equally likely because there are the same number of Clubs (five) as non-Clubs (five, namely, three diamonds plus two hearts).

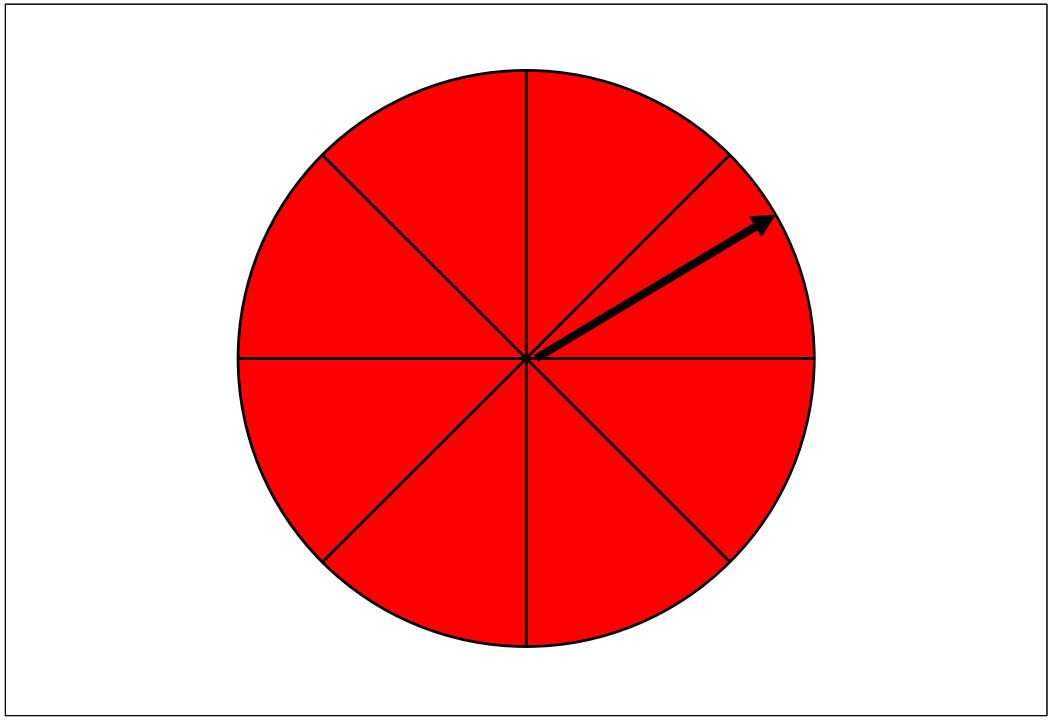
Probability as a Measure – in words

The second level of probability development at Level A is to consider the probability of an event as a measure **in words** from **impossible** to **certain** with increments in between.

Impossible **Unlikely** **Equally Likely** **Likely** **Certain**

Probability as a Measure – in words

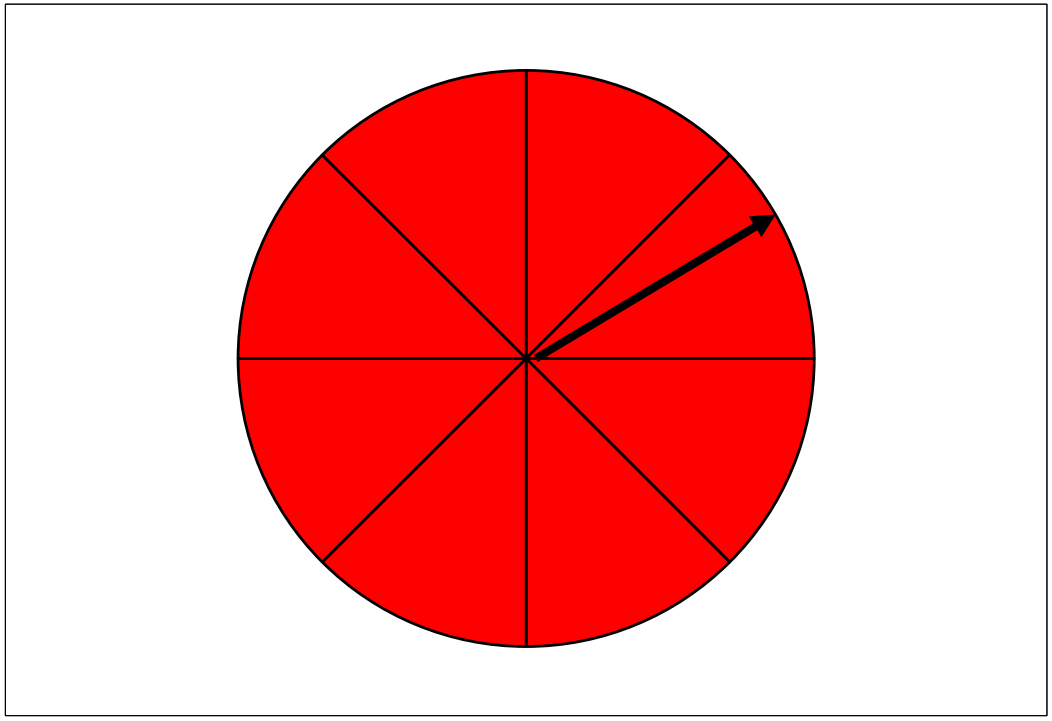
Impossible Unlikely Equally Likely Likely Certain



Spin the spinner once. How likely is it that it will land on a red space? Why?

Probability as a Measure – in words

Impossible Unlikely Equally Likely Likely Certain

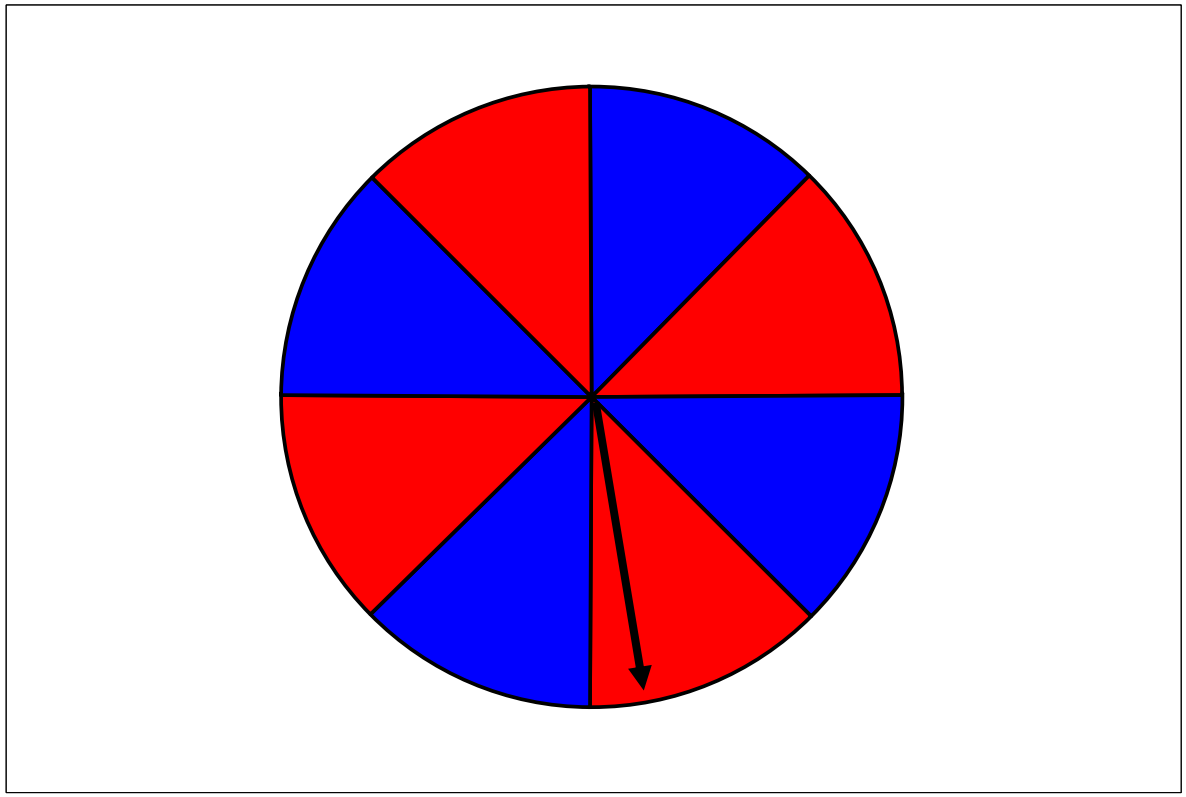


Spin the spinner once. How likely is it that it will land on a red space? Why?

Certain because all the sectors are red.

Probability as a Measure – in words

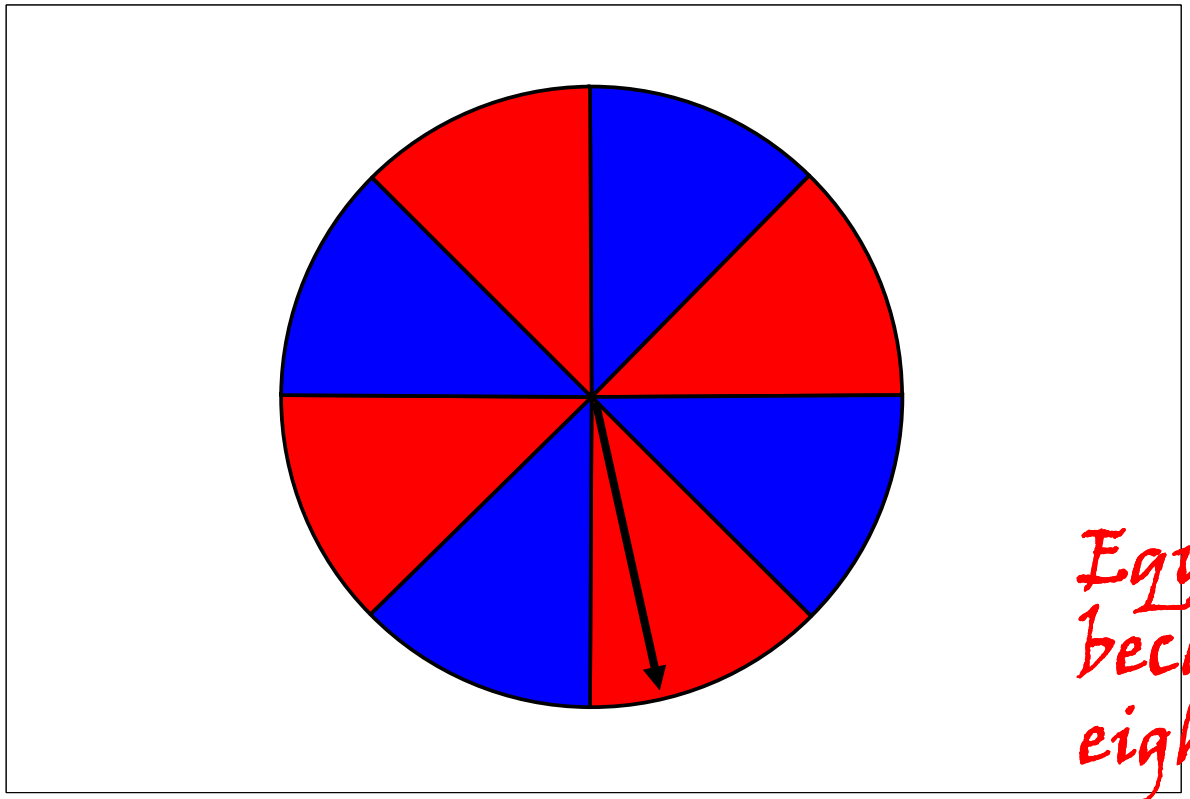
Impossible Unlikely Equally Likely Likely Certain



Spin the spinner once. How likely is it that it will land on a red space? Why?

Probability as a Measure – in words

Impossible Unlikely Equally Likely Likely Certain

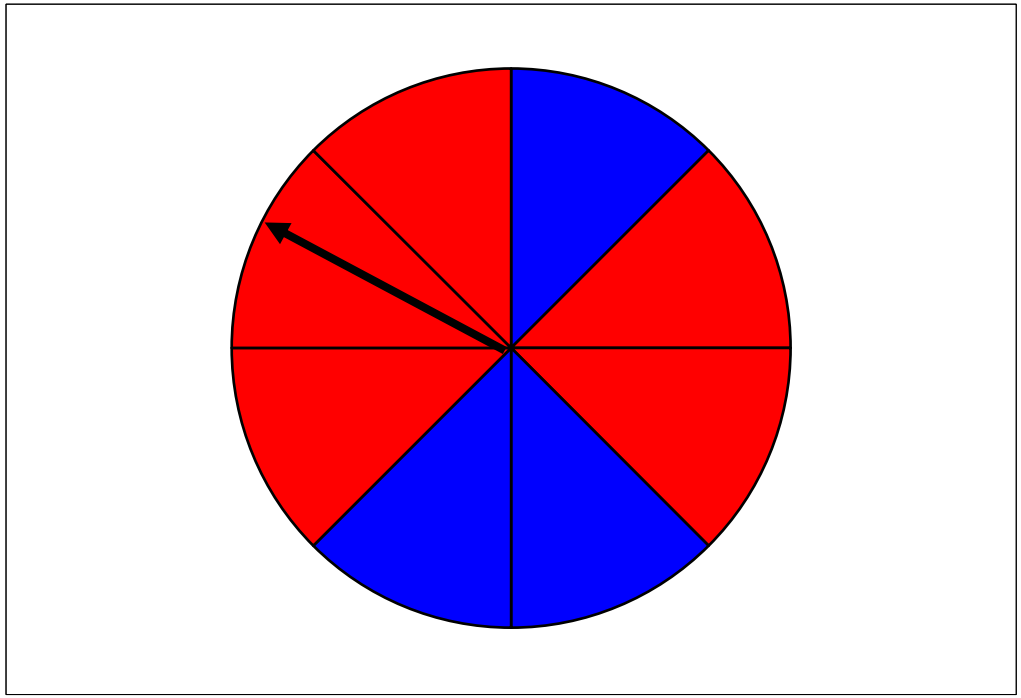


Spin the spinner once. How likely is it that it will land on a red space? Why?

Equally likely because four of the eight congruent sectors are red and four are blue, the same number.

Probability as a Measure – in words

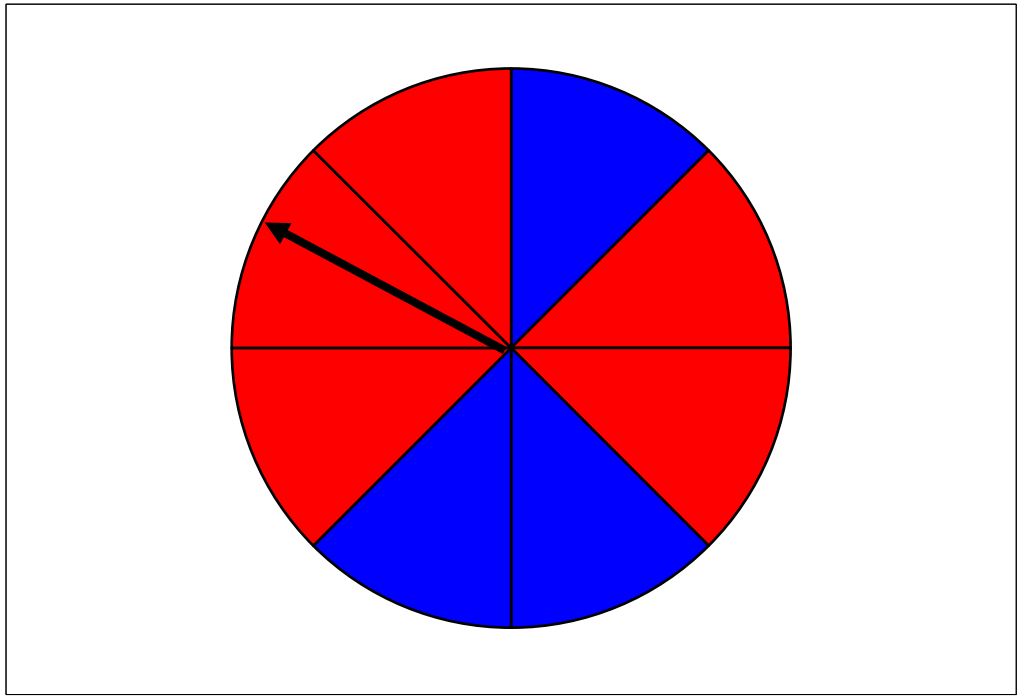
Impossible Unlikely Equally Likely Likely Certain



Spin the spinner once. How likely is it that it will land on a red space? Why?

Probability as a Measure – in words

Impossible Unlikely Equally Likely Likely Certain

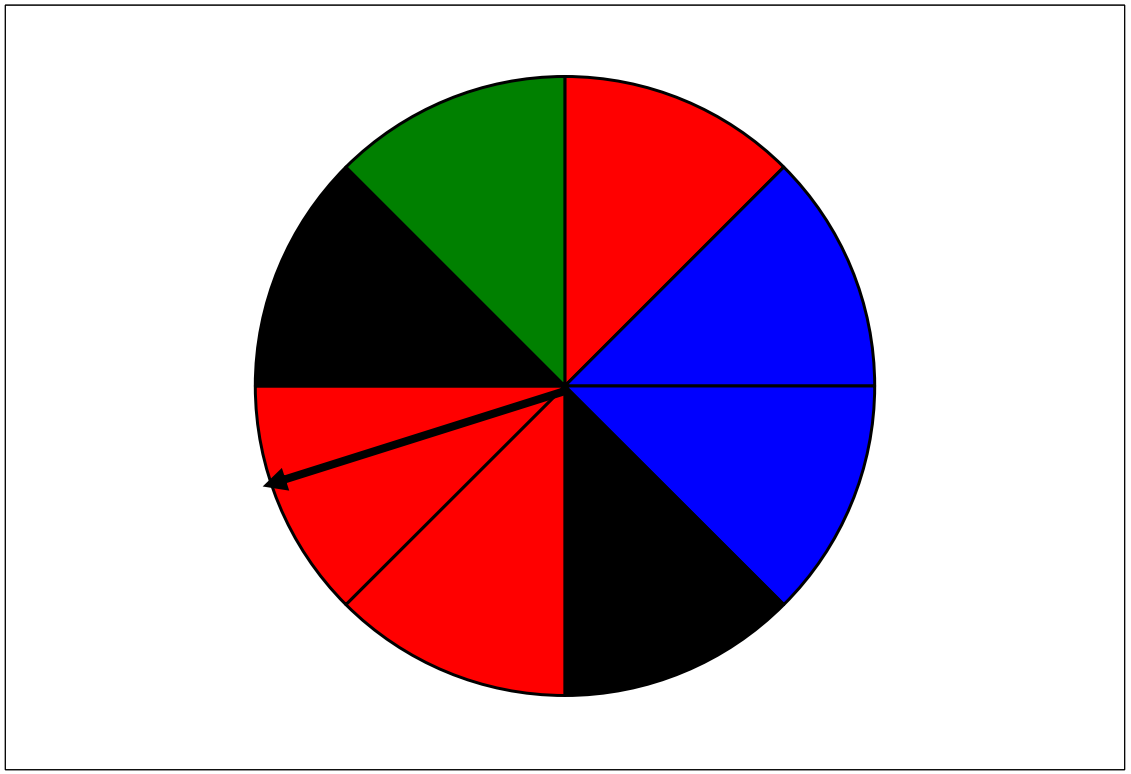


Spin the spinner once. How likely is it that it will land on a red space? Why?

Likely because the majority of the eight congruent sectors, five of them are red which is more than the three blue ones.

Probability as a Measure – in words

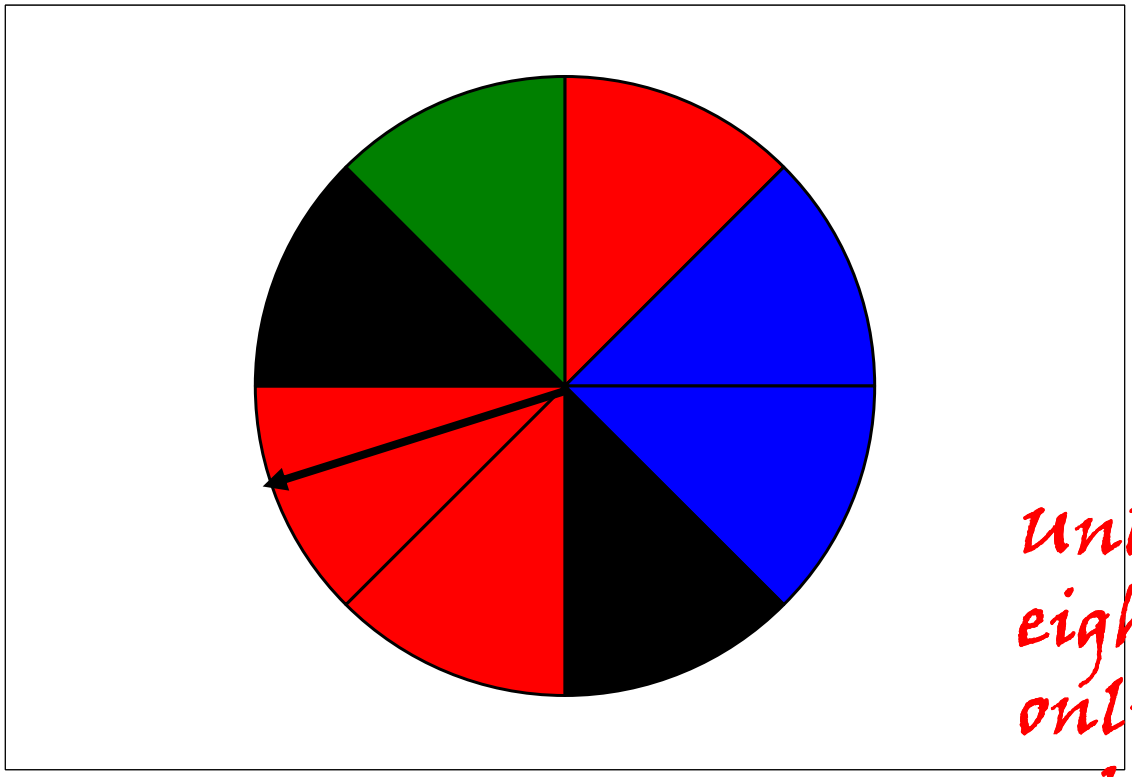
Impossible Unlikely Equally Likely Likely Certain



Spin the spinner once. How likely is it that it will land on a red space? Why?

Probability as a Measure – in words

Impossible Unlikely Equally Likely Likely Certain



Spin the spinner once. How likely is it that it will land on a red space? Why?

Unlikely because of the eight congruent sectors only three of them are red. That's less than half of them.

Probability as a Measure – as a number

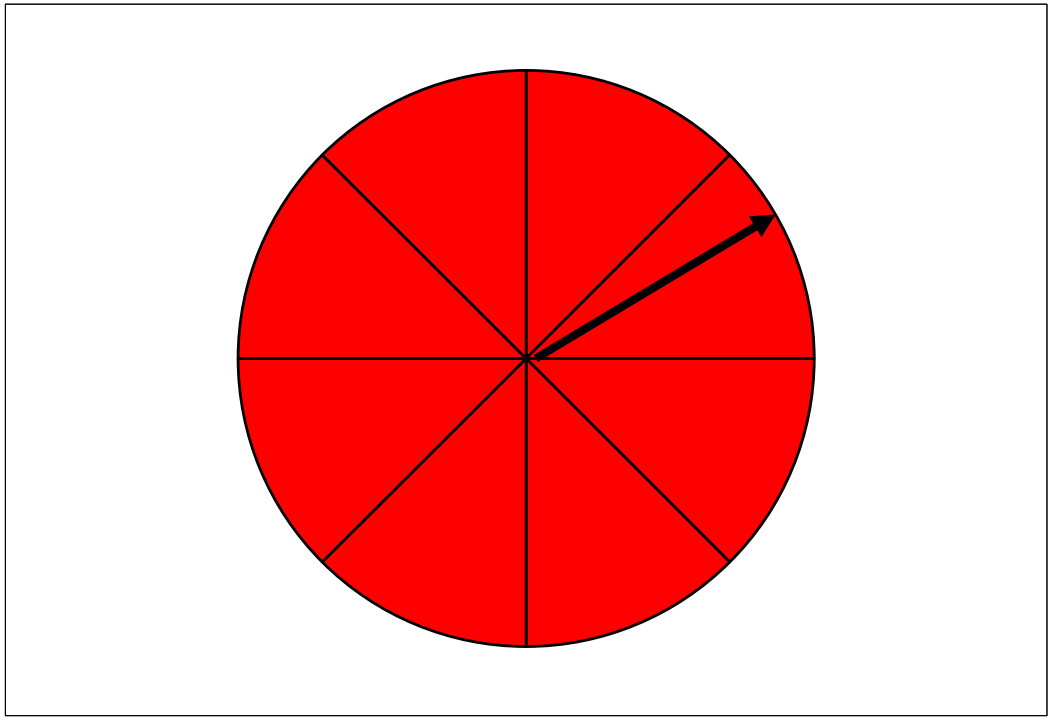
Impossible		Unlikely		Equally Likely		Likely		Certain
0		1/4		2/4 = 1/2		3/4		1
0	1/8	2/8	3/8	1/2	5/8	3/4	7/8	1

The third level of probability development is to consider the probability of an event as a measure **in fractions** from 0 to 1.

The initial assignment of fractions should be 0 to "impossible," 1/4 to "unlikely," 1/2 to "equally likely," 3/4 to "likely," and 1 to "certain." Moving to eighths, sixteenths, etc. is a natural progression (especially using spinners) that eventually leads to understanding that the scale is a continuum from 0 to 1.

Probability as a Measure – as a number

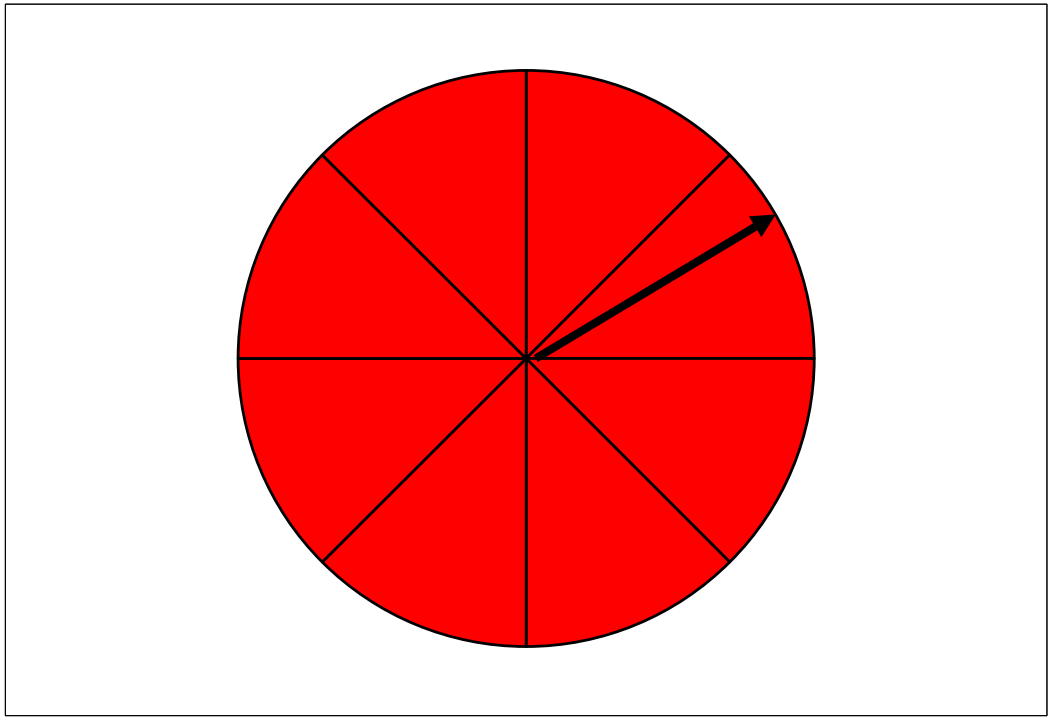
0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1



Spin the spinner once. How likely is it that it will land on a red space? Why?

Probability as a Measure – as a number

0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1

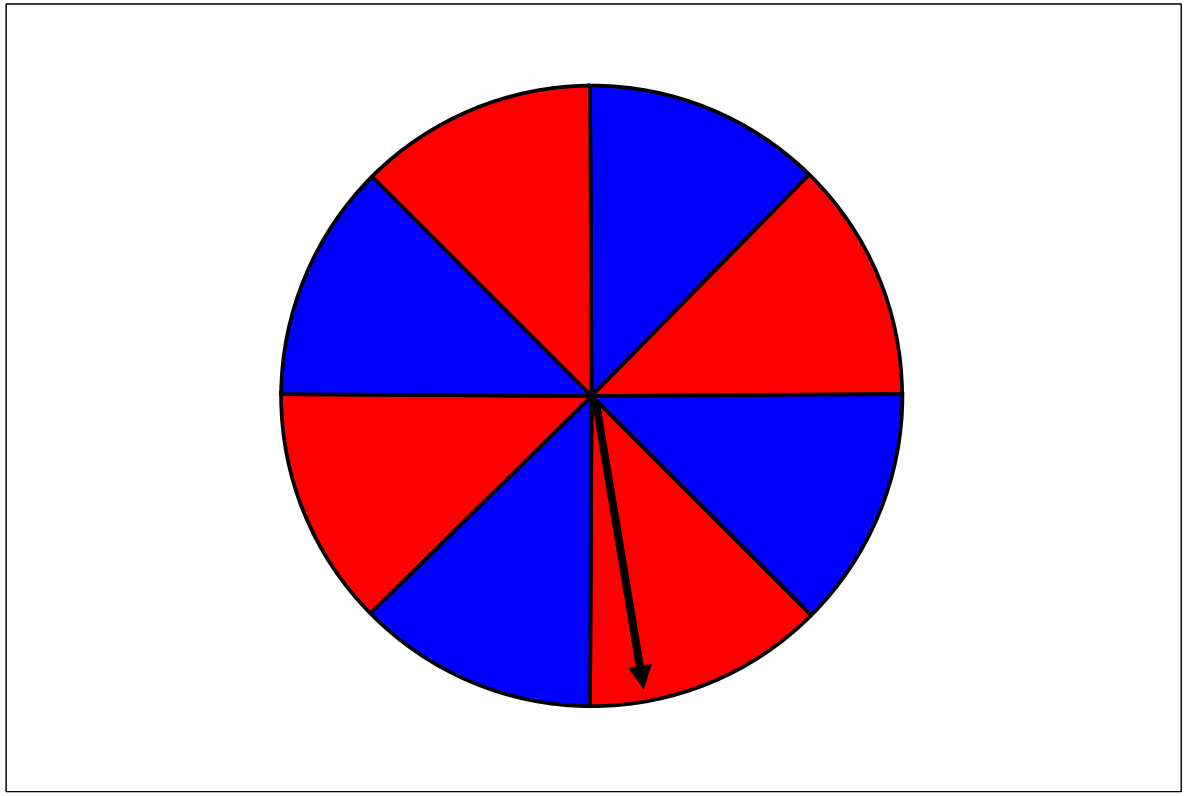


Spin the spinner once. How likely is it that it will land on a red space? Why?

The probability is 1 that it will land on a red space because getting red is the only possibility.

Probability as a Measure – as a number

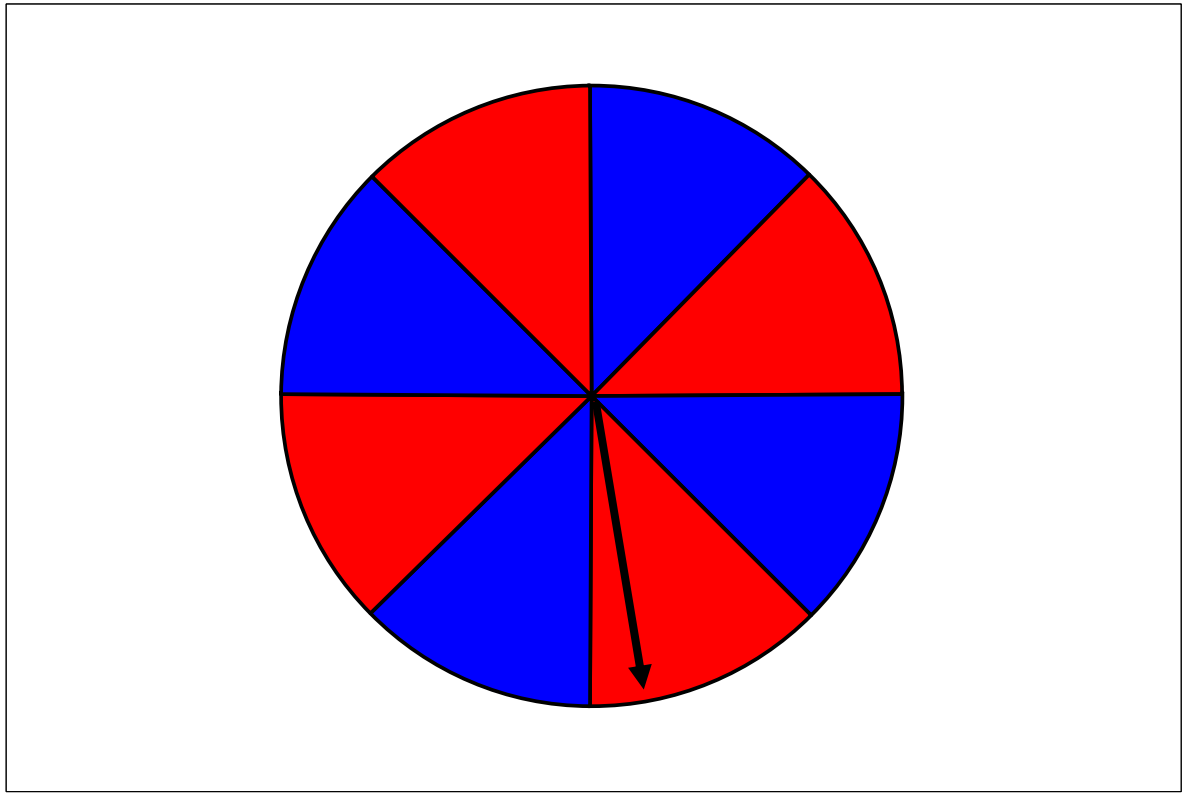
0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1



Spin the spinner once. How likely is it that it will land on a red space? Why?

Probability as a Measure – as a number

0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1

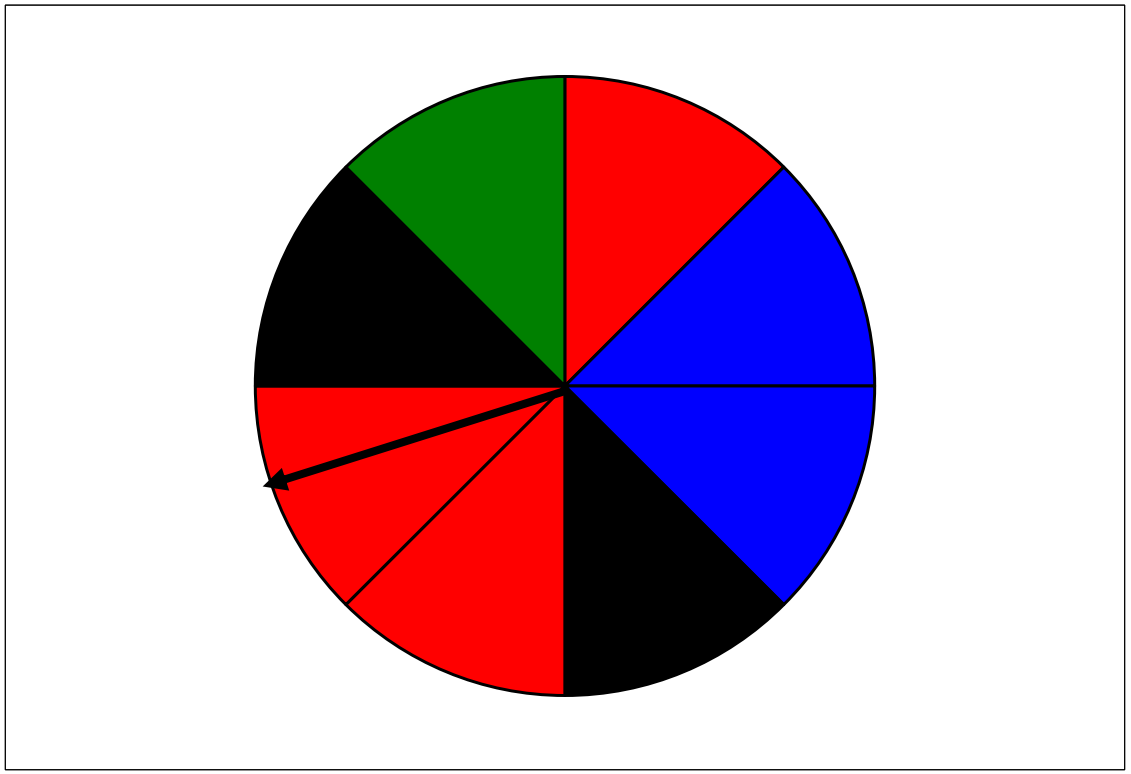


Spin the spinner once. How likely is it that it will land on a red space? Why?

The probability that it will land on red is 4/8 or 1/2 because of the eight congruent sectors, four are red.

Probability as a Measure – as a number

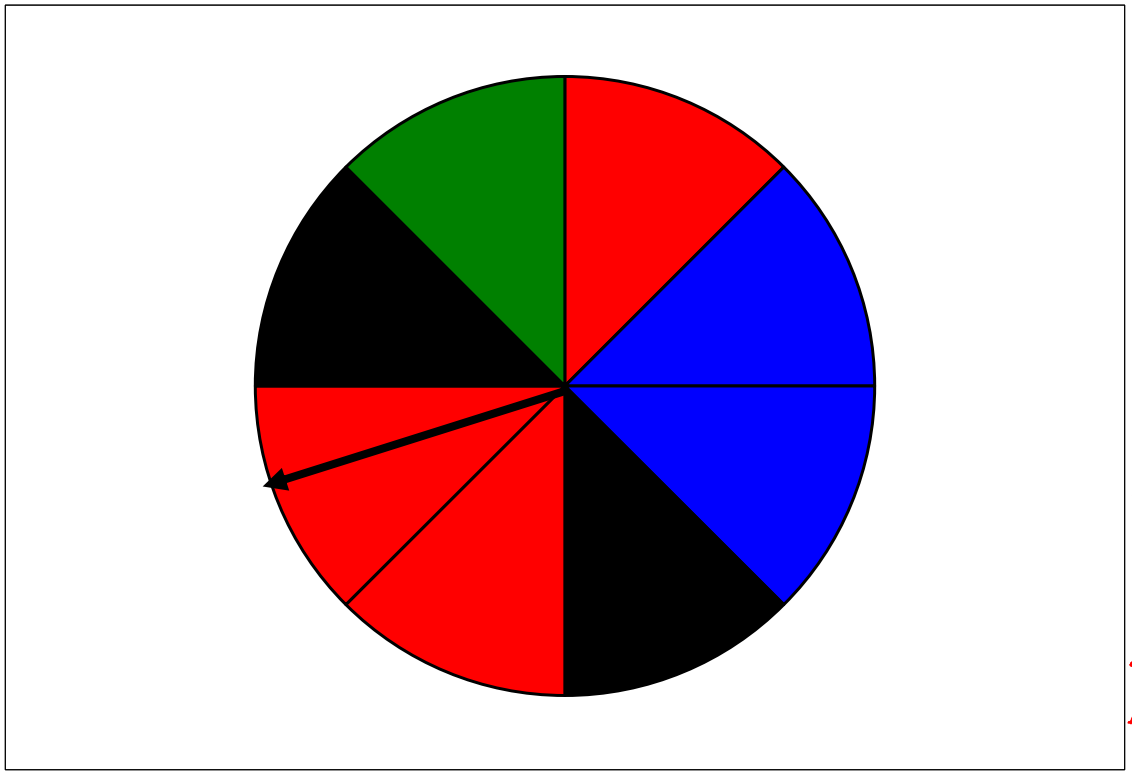
0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1



Spin the spinner once. How likely is it that it will land on a red space? Why?

Probability as a Measure – as a number

0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1

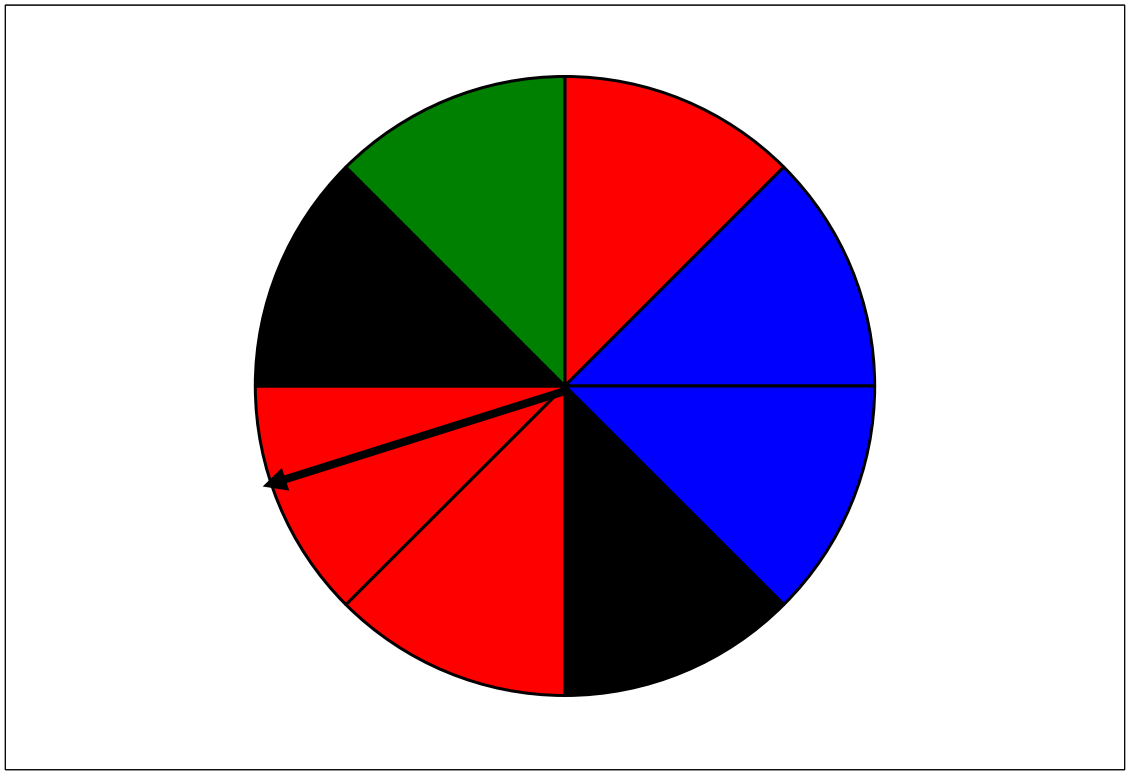


Spin the spinner once. How likely is it that it will land on a red space? Why?

The probability of landing on red is 3/8 because of the eight congruent sectors, three are red.

Probability as a Measure – as a number

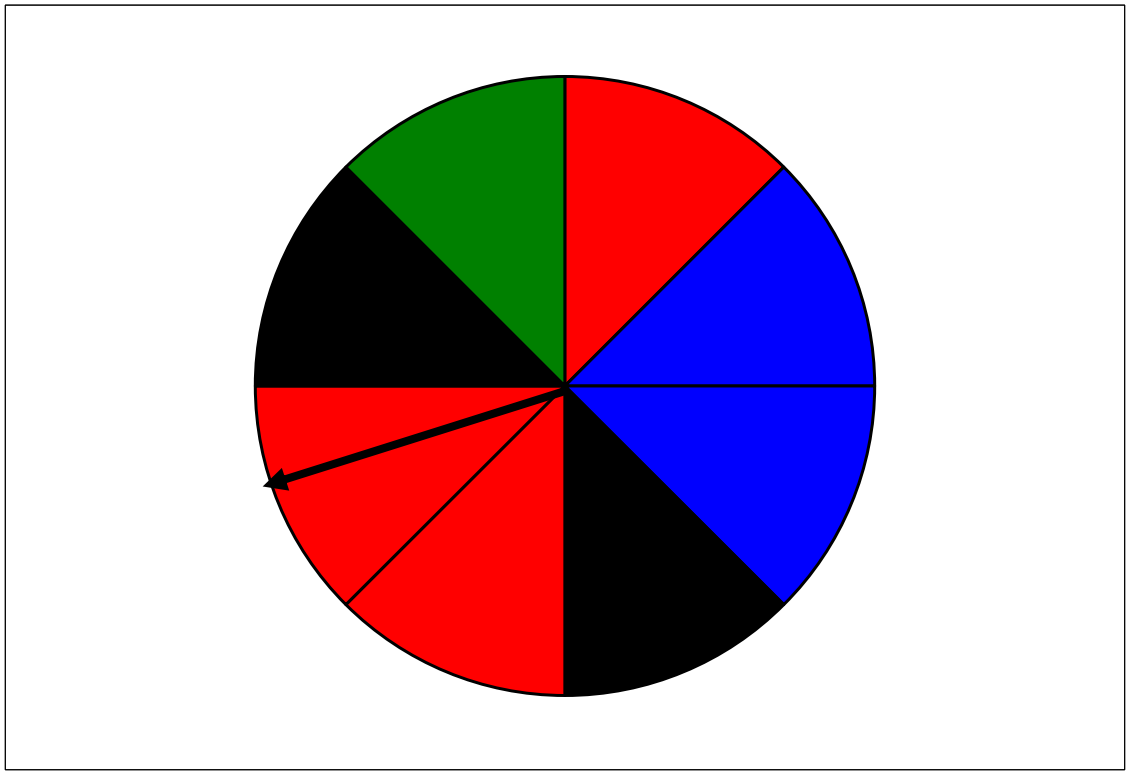
0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1



Spin the spinner once. How likely is it that it will land on a yellow space? Why?

Probability as a Measure – as a number

0	1/8	2/8	3/8	4/8	5/8	6/8	7/8	1
0	1/8	1/4	3/8	1/2	5/8	3/4	7/8	1

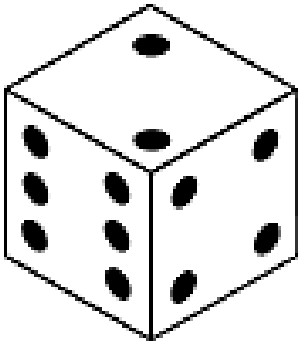


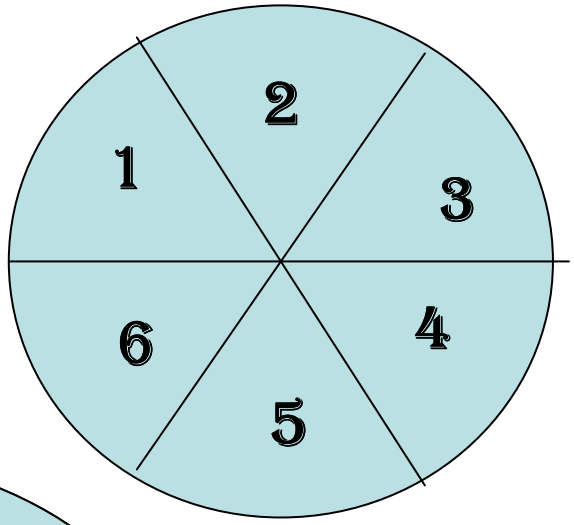
Spin the spinner once. How likely is it that it will land on a yellow space? Why?

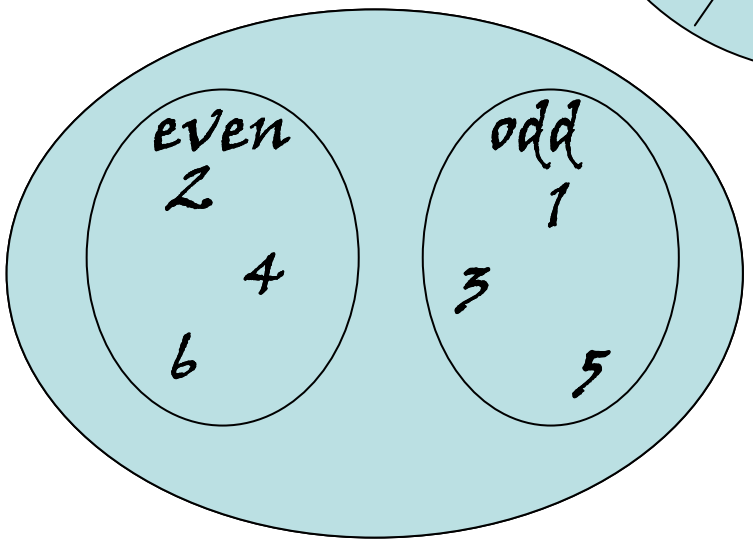
The probability the spinner will stop on yellow is 0 because there are no yellow sectors.

Probability as a Measure – as a number

0	1/6	2/6	3/6	4/6	5/6	1
0	1/6	1/3	1/2	2/3	5/6	1



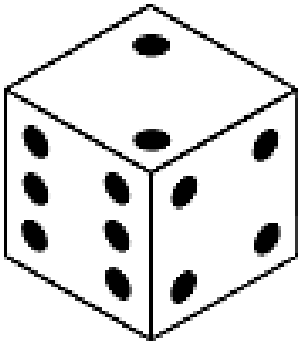


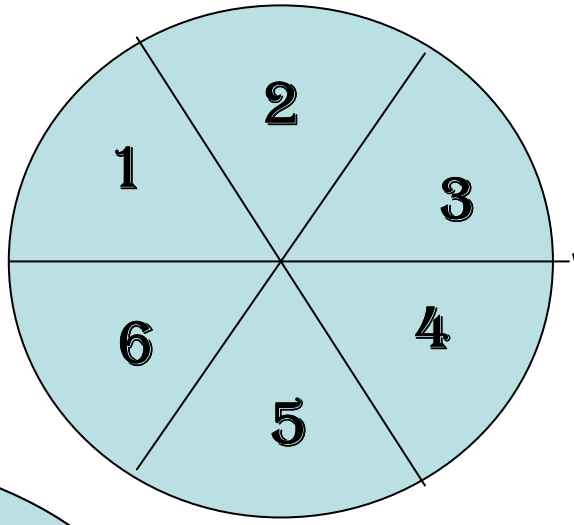


Roll the number cube once. How likely is it that it will land showing an even number up? Why?

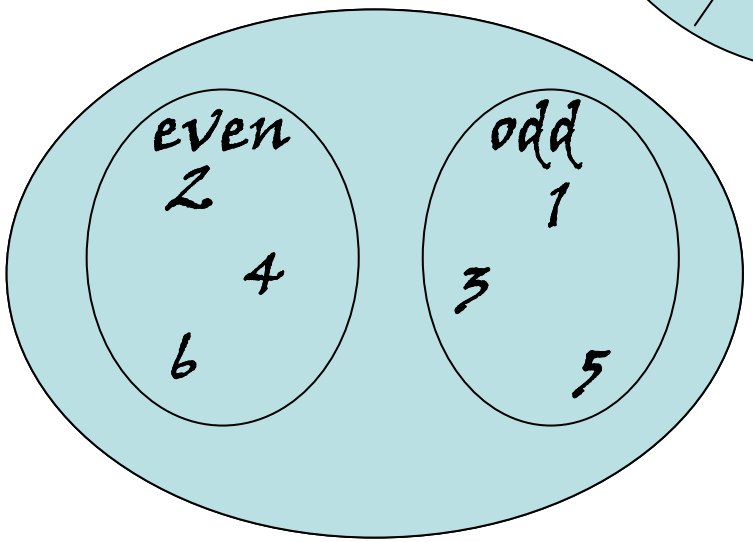
Probability as a Measure – as a number

0	1/6	2/6	3/6	4/6	5/6	1
0	1/6	1/3	1/2	2/3	5/6	1





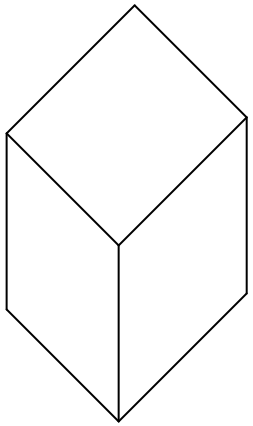
Roll the number cube once. How likely is it that it will land showing an even number up? Why?

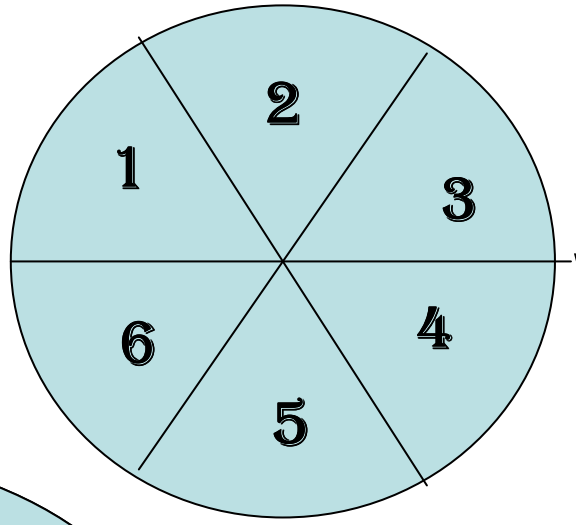


The probability is 3/6 or 1/2 of showing an even number because there are three even numbers out of the six possible numbers.

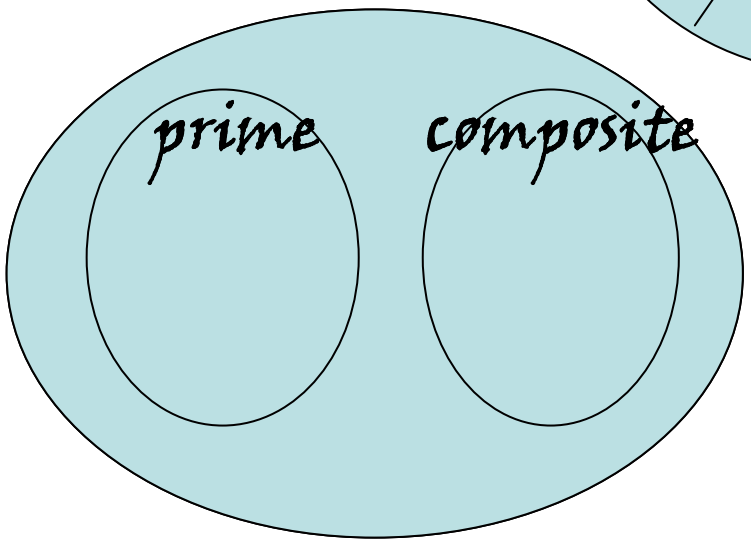
Probability as a Measure – as a number

0	1/6	2/6	3/6	4/6	5/6	1
0	1/6	1/3	1/2	2/3	5/6	1



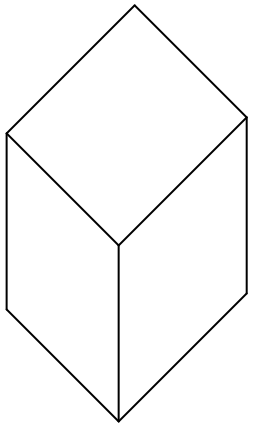


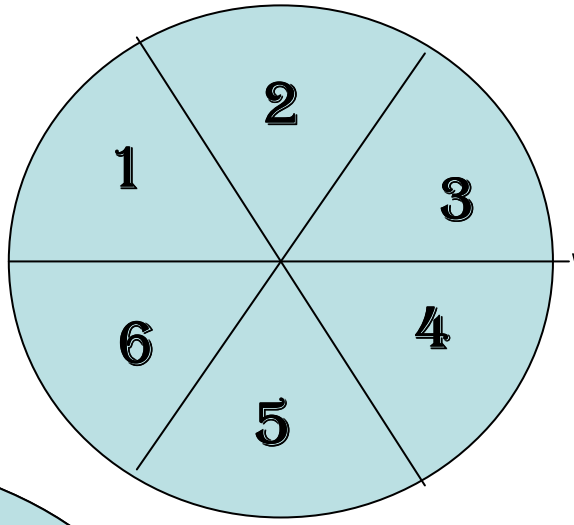
Roll the number cube once. How likely is it that it will land showing a prime number up? Why?



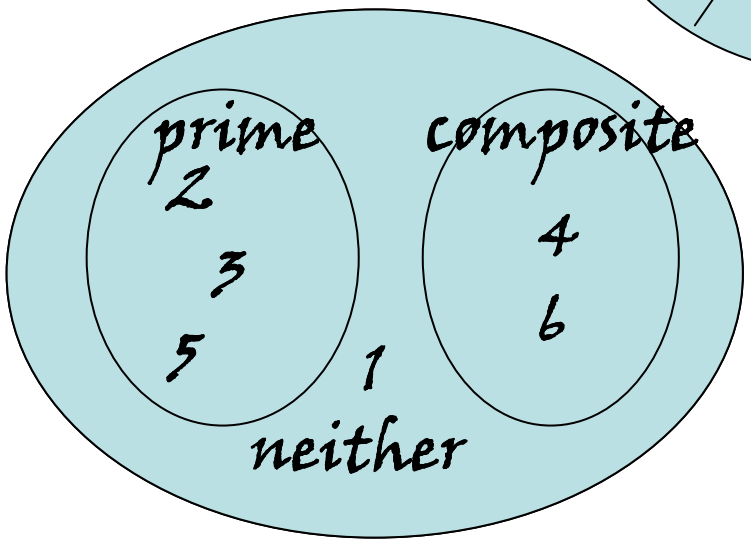
Probability as a Measure – as a number

0	1/6	2/6	3/6	4/6	5/6	1
0	1/6	1/3	1/2	2/3	5/6	1





Roll the number cube once. How likely is it that it will land showing a prime number up? Why?



The probability is 3/6 or 1/2 of showing a prime number because there are three prime numbers out of the six possible numbers : 2, 3, 5.

Estimating Probability using Empirical Data

The next level of probability development is to estimate probabilities from a distribution of data. Students should use experiments or simulations to collect data, build a frequency distribution for the observed outcomes, and then ask "how likely" (i.e., probability) questions based on the data set.

The following experiment involves *Austin ZOO Animal Crackers*, a Kellogg product.

The following experiment is based on a conversation I had with **Michael Kimmel**. *I am thankful for not only this suggestion but the many others he has given me over the past fifteen years.*

Estimating Probability using Empirical Data

There are 12 possible “zoo” animals:

bear, camel, elephant, lion, monkey, mountain goat, owl, penguin, rabbit, rhinoceros, tortoise, zebra.

One of the problems with using these animals as data is that they get “wounded” in transit. Discuss statistical protocol.

Estimating Probability using Empirical Data

Bear

Camel

Elephant

Lion

Monkey

Mountain Goat

Owl

Penguin

Rabbit

Rhinoceros

Tortoise

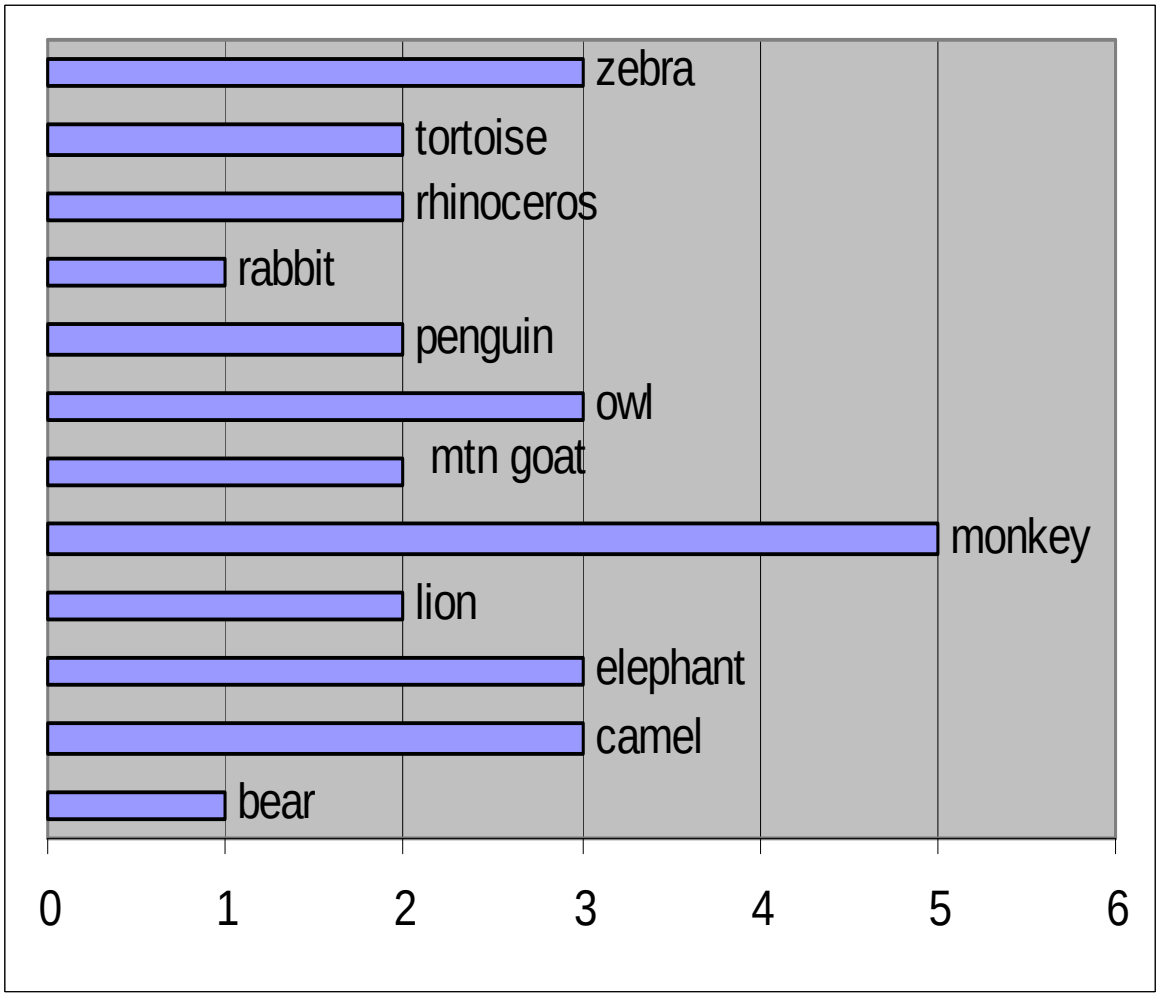
Zebra

Estimating Probability using Empirical Data

Build a frequency distribution for the animals in **your** sample. Tally your data first in a table and then draw an appropriate graph to illustrate your distribution.

Estimating Probability using Empirical Data Example

bear	1
camel	3
elephant	3
lion	2
monkey	5
mountain goat	2
owl	3
penguin	2
rabbit	1
rhinoceros	2
tortoise	2
zebra	3
Total	29



Estimating Probability using Empirical Data

If you were to draw one animal from your menagerie (i.e., choose one cracker without looking from your thoroughly mixed bag of animals), which animal in your data set is most likely to be drawn? Least likely?

Estimating Probability using Empirical Data

If you were to draw one animal from your menagerie (i.e., choose one cracker without looking from your thoroughly mixed bag of animals), which animal would you be most likely to draw? Least likely?

For the example menagerie, the most likely animal chosen would be a monkey since there are more of them, five, than any other animal, and the least likely would be either a bear or a rabbit since there is just one each of them.

B	C	E	L	M	MG	O	P	Ra	Rh	T	Z	
1	3	3	2	5	2	3	2	1	2	2	3	29

Estimating Probability using Empirical Data

If you were to draw one animal from your menagerie (i.e., choose one cracker without looking from your thoroughly mixed bag of animals), what is the probability of choosing your most likely animal? Least likely animal?

B	C	E	L	M	MG	O	P	Ra	Rh	T	Z	
1	3	3	2	5	2	3	2	1	2	2	3	29

Estimating Probability using Empirical Data

If you were to draw one animal from your menagerie (i.e., choose one cracker without looking from your thoroughly mixed bag of animals), what is the probability of choosing your most likely animal? Least likely animal?

For the example menagerie, the probability that a monkey will be chosen is 5/29 because there are 5 monkeys in the data set and 29 animals in the data set. The probability of choosing a bear is 1/29. The probability of choosing a rabbit is also 1/29.

B	C	E	L	M	MG	O	P	Ra	Rh	T	Z	
1	3	3	2	5	2	3	2	1	2	2	3	29

Estimating Probability using Empirical Data

If you were to draw one animal from your menagerie (i.e., choose one cracker without looking from your thoroughly mixed bag of animals), what is the probability of choosing:

a) **An animal with four feet?**
b) **An animal with claws? hooves?**
c) **An animal with knees?**
d) **An animal whose number of letters in its name is >= 5?**
e) **An animal whose species first letter is in the first half of the alphabet?**
f) **A carnivorous animal? herbivorous? omnivorous?**

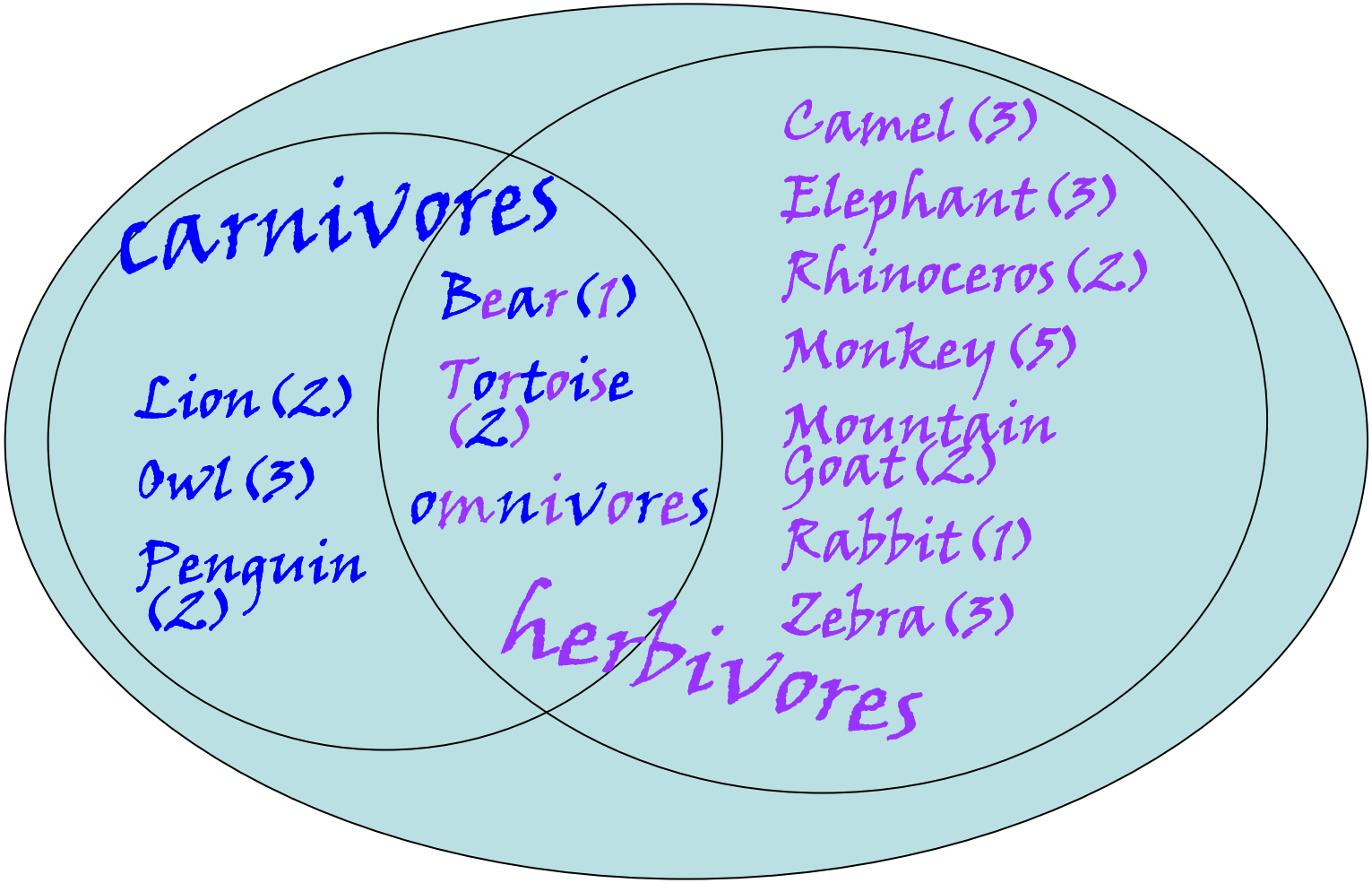
Estimating Probability using Empirical Data

From your menagerie, identify your carnivores, herbivores, and omnivores and draw an appropriate graph.

Suggestion: Draw a Venn Diagram.

Estimating Probability using Empirical Data

Here's a Venn Diagram identifying the carnivores, herbivores, and omnivores for the example menagerie.



Estimating Probability using Empirical Data

For your menagerie, what is the probability of randomly choosing:

a) A meat-eating animal?

b) An animal that eats vegetation?

c) Do a) and b) add to one? Why or why not?

d) A pure carnivore?

e) A pure herbivore?

f) An omnivore?

g) Do d), e), and f) add to one? Why or why not?

Estimating Probability using Empirical Data

For the example menagerie, what is the probability of randomly choosing:

a) A meat-eating animal?

Estimating Probability using Empirical Data

MYSTERY SPINNER

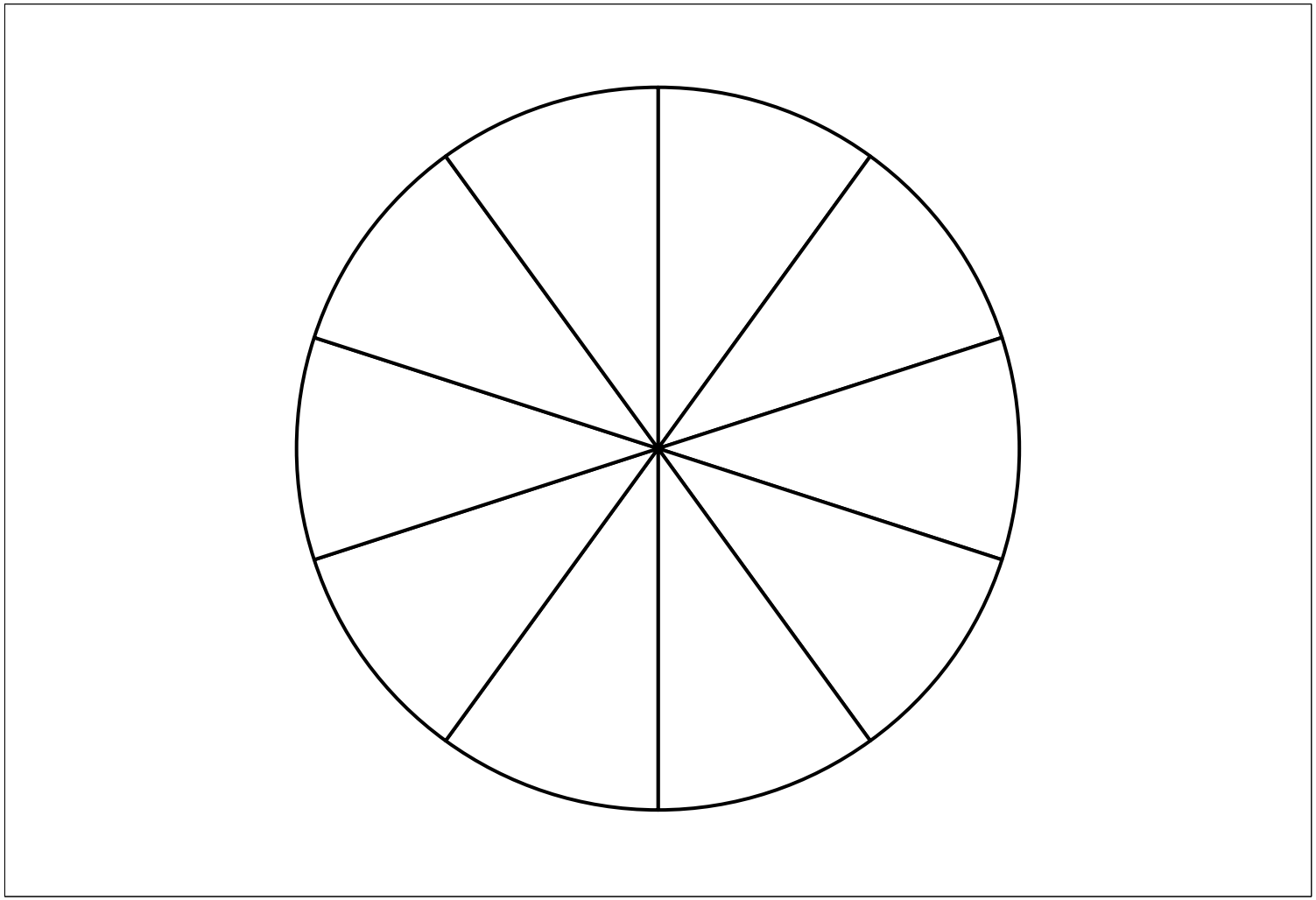
It's carnival time at your school and your students want to design a large spinner just like the one used last year, but it is missing and nobody can remember exactly what it looked like, except that it was divided into 10 equal parts.

Fortunately, you had written down data from last year's spins, namely, there were: 18 blue, 11 yellow, 11 red, 8 black, 8 purple, 13 pink, 22 green, and 9 white.

Using your data, can your students conjecture what last year's spinner looked like? If so, what might they come up with?

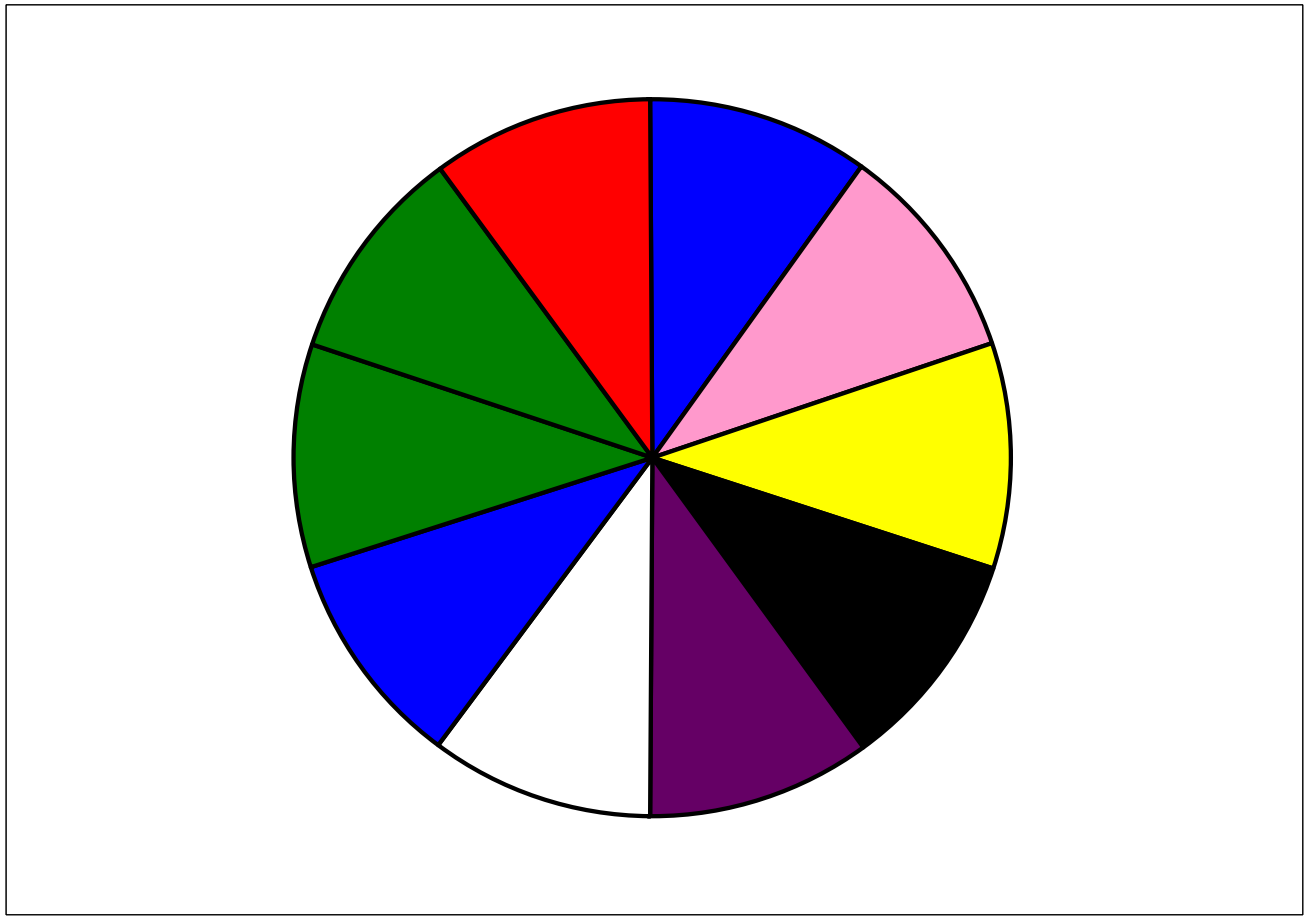
Estimating Probability using Empirical Data

MYSTERY SPINNER



Estimating Probability using Empirical Data

The students suggested the following spinner would satisfy the empirical data. Do you agree?



PROBABILITY REVIEW

Levels of development of probability in grades k-4:

1. Identifying when one event **compared** to another is **less likely**, **equally likely**, or **more likely** to happen.

2. Providing a **measure** of how likely an event can happen, **in words**, namely:

   Impossible Unlikely Equally Likely Likely Certain

3. Providing a **measure** of how likely an event can happen, **as a fraction**, for example:

Impossible	Unlikely	Equally Likely	Likely	Certain
0	**1/4**	**1/2**	**3/4**	**1**

4. Estimating probabilities using empirical data.

5. Long-run relative frequency; Law of Large Numbers.

OACS Data Analysis ...highlights

- K-2: Pose questions; sort and classify objects; draw picture graphs and bar graphs.
- 3-4: Do surveys and experiments; read, interpret, and construct tables, charts, graphs (bar, picture, line, line plot, Venn diagram). Describe data using mode, median, and range.
- 5-7: Read, create, and use line graphs, histograms, circle graphs, boxplots, stemplots. Interpret data by looking for patterns, relationships. Determine, use, and explain the range, mean, median, and mode for a set of data. Identify misuses of statistical data and displays.

State Standards

- **Mathematics - Most follow NCTM**
- **Science (e.g., OACS)**

  Use evidence and observations to explain and communicate the results of investigations.
- **Social Studies (e.g., OACS)**

  Read and interpret pictographs, bar graphs, line graphs, circle graphs, tables and flow charts.

GAISE Data Analysis

Data analysis is an investigative problem-solving process immersed in a context, and not a set of fancy tools and graphs and procedures for their own sake isolated from a context.

The GAISE Framework
Developmental Levels A, B, C
for evolving statistical concepts

- The three levels roughly parallel PSSM's preK-5, 6-8, and 9-12 grade bands.
- However, the Framework levels are based on experience, not age.
  - Level A…learning is more teacher driven
  - Level B…learning becomes more student centered
  - Level C…learning is highly student driven
- Hands-on, active learning is predominant throughout.

The Bottom Line of GAISE

Statistical analysis is an **investigative process** that turns often loosely formed ideas into scientific studies by:

- Formalizing a question that can be answered with data;
- Designing a plan to collect appropriate data;
- Analyzing the collected data by graphical and numerical methods;
- Interpreting the analysis so as to reflect light on the original question.

The Framework Model – two-dimensional

Process Component	Level A	Level B	Level C
Formulate Question	• See the GAISE document, pages 14-15 for definition and details. • See Statistics Teacher Network newsletter 68 for a complete example. *A Sequence of Activities for Developing Statistical Concepts* by Chris Franklin and Gary Kader		
Collect Data			
Analyze Data			
Interpret Results			

Distinction of Levels

- All four steps of the statistical process are used at all three levels A, B, C.
- The depth of understanding and sophistication of methods used increases across the levels.

For example,

a level A class may collect data to answer questions about their classroom;

a level B class may collect data to answer questions about their school; and,

a level C class may collect data to answer questions about their community and model the relationship between, say, housing prices and geographical variables such as the location of their schools.

Distinction of Levels

- Use of Probability:

- Level A
  - impossible…equally likely…certain
  - experimental and theoretical probability for coins/spinners

- Level B
  - concept of not equally likely; proportional reasoning;
  - Binomial distribution

- Level C
  - use of simulation for sampling distributions to examine p-value

Distinction of Levels

Mean:

- Level A

  mean as fair share

- Level B

  mean as a balancing point

- Level C

  mean as an estimate from a sample that will be used to make an inference about a population – concept of sampling distribution

Distinction of Levels

- What type of music is most popular among their peers in school? (rock, country, rap)

- Level A

  Summarize frequencies in table or bar graph

- Level B

  Transition to relative frequencies – proportional reasoning

- Level C

  Transition to sampling distributions for a sample proportion and role of probability in finding a margin of error (of using sample proportion to estimate population proportion)

GAISE SUMMARY

- Both conceptual understanding and procedural skills should be developed deliberately, but conceptual understanding should not be sacrificed for procedural proficiency.
- Active learning is key to the development of conceptual understanding.
- Real world data must be used wherever possible in statistics education.
- Appropriate technology is essential in order to emphasize concepts over calculations.

A Data Analysis Activity

based on *Hickory DIckory Dock: Navigating through Data Analysis*
D. Niezboda, P.Moyer-Packenham, NCTM Teaching Children Mathematics, Feb. 2005

London Bridge is falling down,
Falling down, falling down,
London Bridge is falling down,
My fair Lady.

Build it up with wood and clay,
Wood and clay, wood and clay,
Build it up with wood and clay,
My fair Lady.

Wood and clay will wash away,
Wash away, wash away,
Wood and clay will wash away,
My fair Lady.

- **Formulate a question:** Suppose that we are interested in the word structure of this nursery rhyme. For example, are there similar words in the London Bridge nursery rhyme? How often do they appear? Unique words? Short words? How short? Long words? How long? How long on average?
- **Collect data to answer the questions:**

- **Formulate a question:** Are there similar words in the London Bridge nursery rhyme? How often do they appear? Unique words? Short words? How short? Long words? How long? How many syllables per word?
- **Collect data to answer the questions:** Construct a tally and frequency table.

Word	Tally	Frequency

DATA

COLLECTION

Word	Tally	Frequency
London	//	2
Bridge	//	2
Is	//	2
falling	////	4
down	////	4
my	///	3
fair	///	3
Lady	///	3
build	//	2
it	//	2
up	//	2
with	//	2
wood	//////	6
and	//////	6
clay	//////	6
will	//	2
wash	////	4
away	////	4
18		59

Analyze the Data

Word	Tally	Frequency
London	//	2
Bridge	//	2
Is	//	2
falling	////	4
down	////	4
my	///	3
fair	///	3
Lady	///	3
build	//	2
it	//	2
up	//	2
with	//	2
wood	//////	6
and	//////	6
clay	//////	6
will	//	2
wash	////	4
away	////	4
18		59

Recall our questions; Which ones can be answered by our frequency table?

Formulate a question: Are there similar words in the London Bridge nursery rhyme? How often do they appear? Unique words? Short words? How short? Long words? How long? How long are the words on average? How many syllables per word?

Analyze the Data

Word	Tally	Frequency
London	//	2
Bridge	//	2
Is	//	2
falling	////	4
down	////	4
my	///	3
fair	///	3
Lady	///	3
build	//	2
it	//	2
up	//	2
with	//	2
wood	//////	6
and	//////	6
clay	//////	6
will	//	2
wash	////	4
away	////	4
18		59

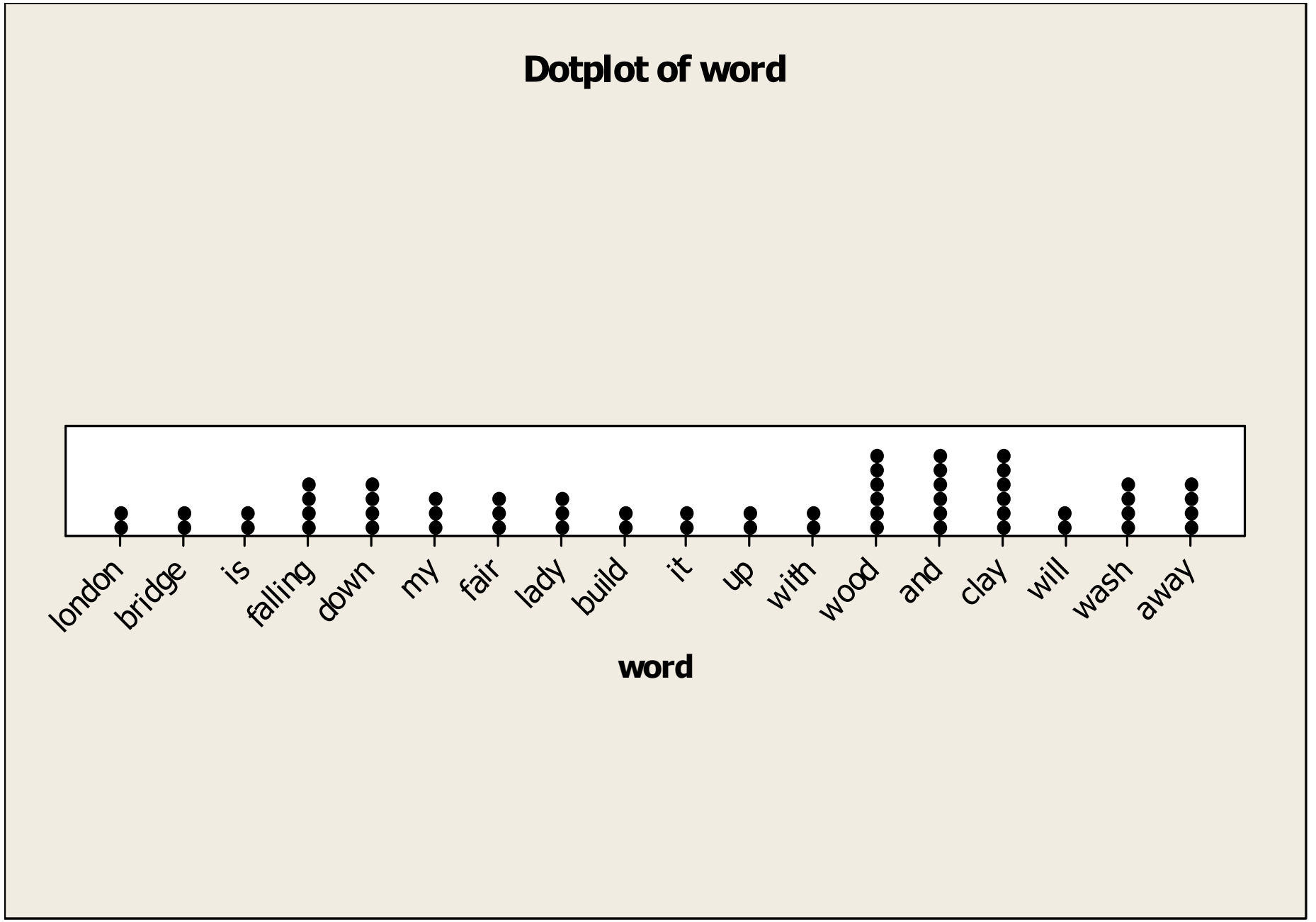
Answer the questions:

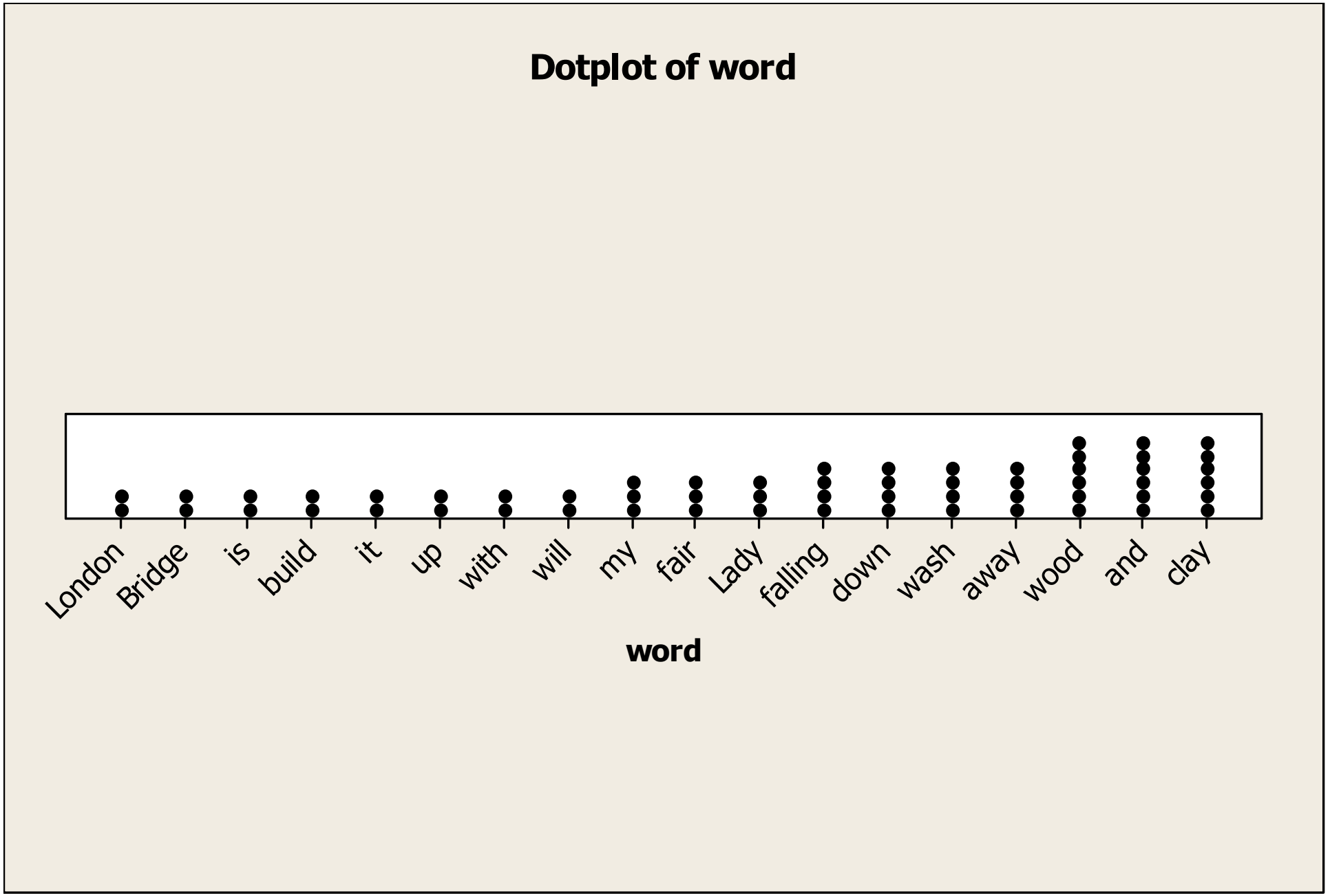
Are there similar words in the London Bridge nursery rhyme? Yes.

How often do they appear? London, Bridge, is, build, it, up, with, and will occur twice. My, fair, and Lady occur three times. Falling, down, wash, and away occur four times. Wood, and, clay occur six times each.

Unique words? There are no unique words.

The rest of the questions require more thought. Short words? How short? Long words? How long? How long are the words on average? How many syllables per word?


Dotplot of word
london
bridge
is
falling
down
my
fair
lady
build
it
up
with
wood
and
clay
will
wash
away
word


Dotplot of word
London
Bridge
is
build
it
up
with
will
my
fair
Lady
falling
down
wash
away
wood
and
clay
word

Extensions

- Analyze the letters in the words –
  - Frequency of letters in the whole rhyme
  - Frequency of the number of different letters per word

- Analyze the syllables in the words –
  - Frequency of syllables per word in the rhyme

Number of Syllables per Word	Frequency
1	14
2	4

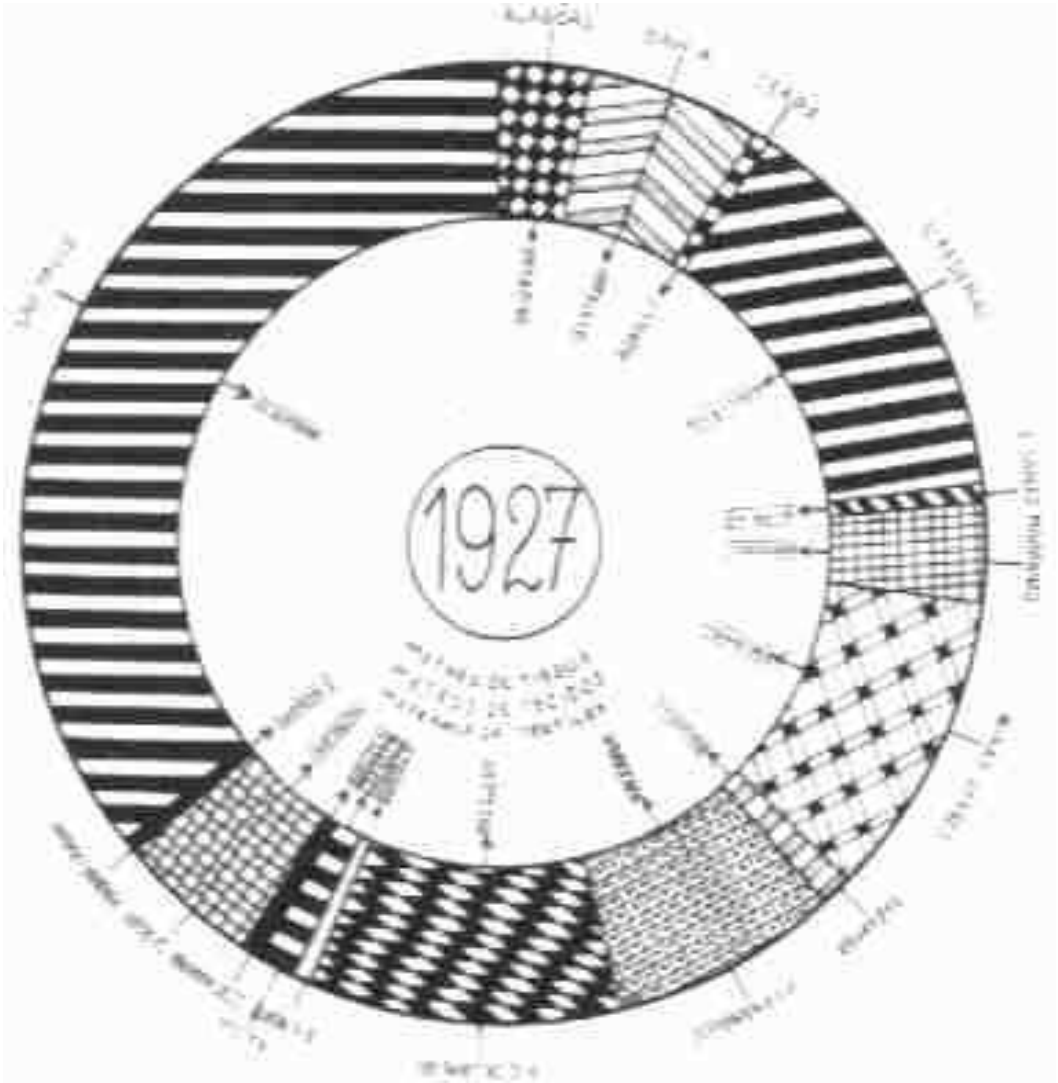
Scratching the Surface in Statistical Graphing

a few dos and don’ts

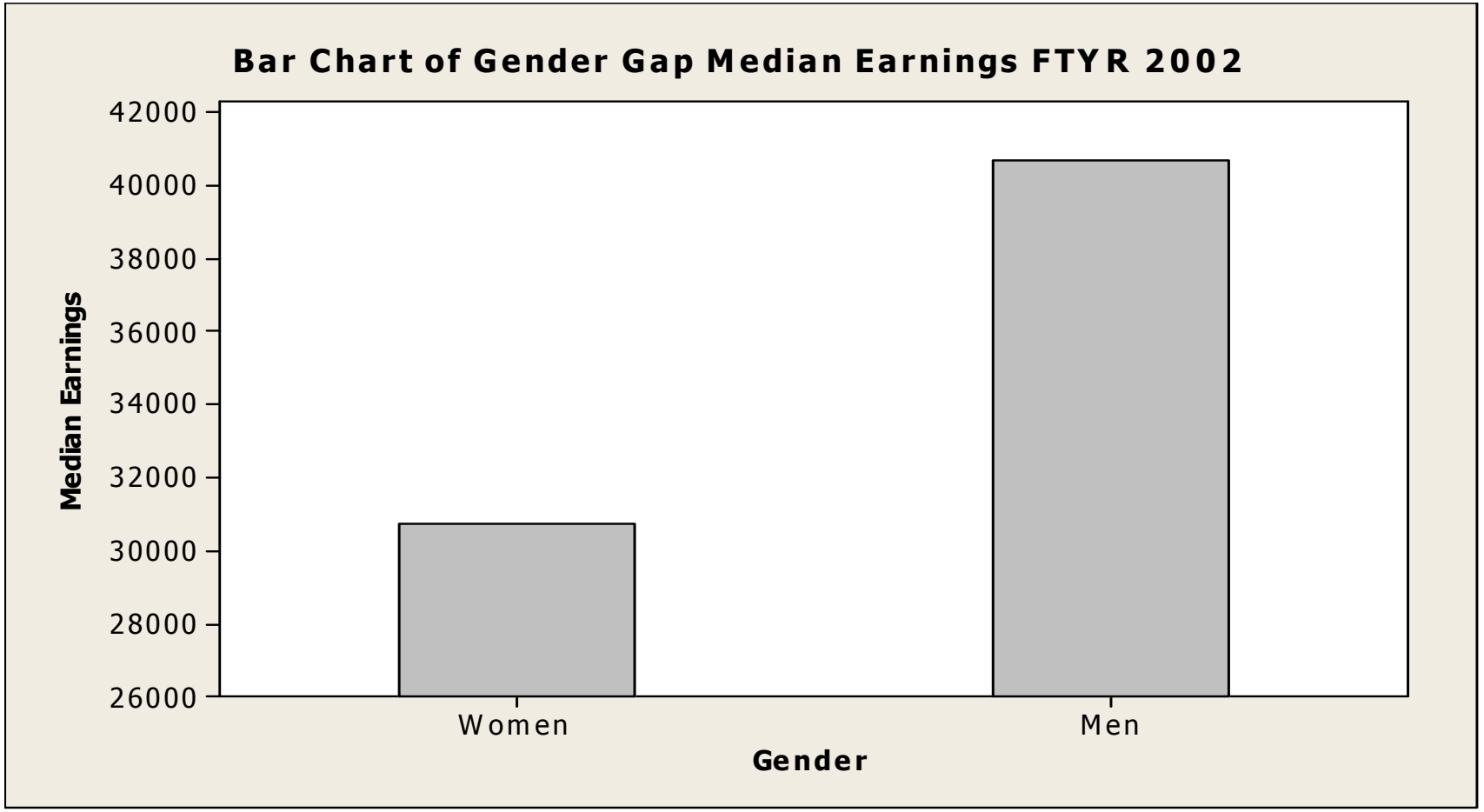
Keep In Mind

- A graph is used to provide a visualization of the information contained in data which enables the reader to summarize and to detect trends, relationships, and comparisons.
- This implies that data must be “clean.”
- The graph must be appropriate to the data and kept as simple as possible.

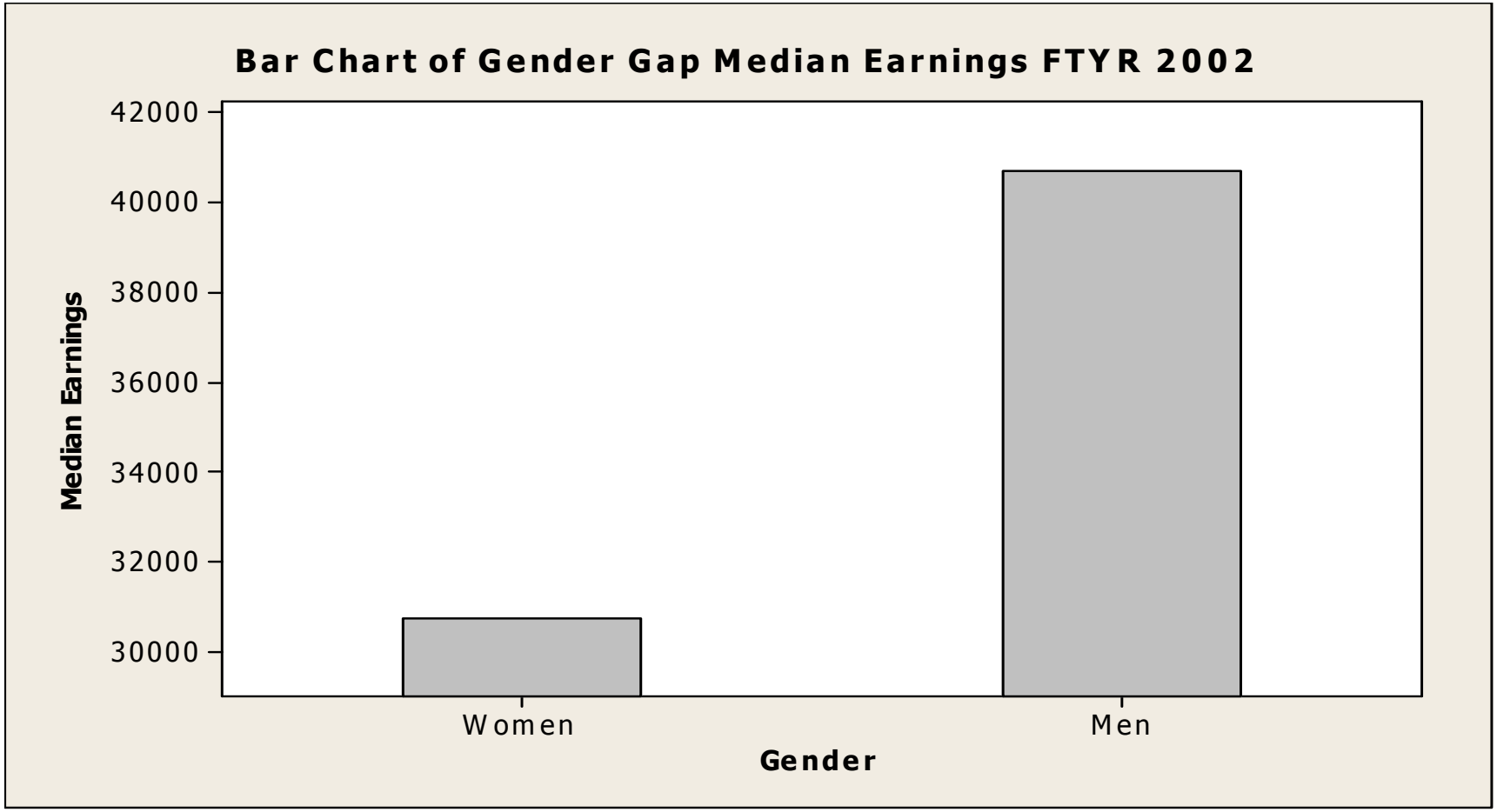
Avoid Chart Junk



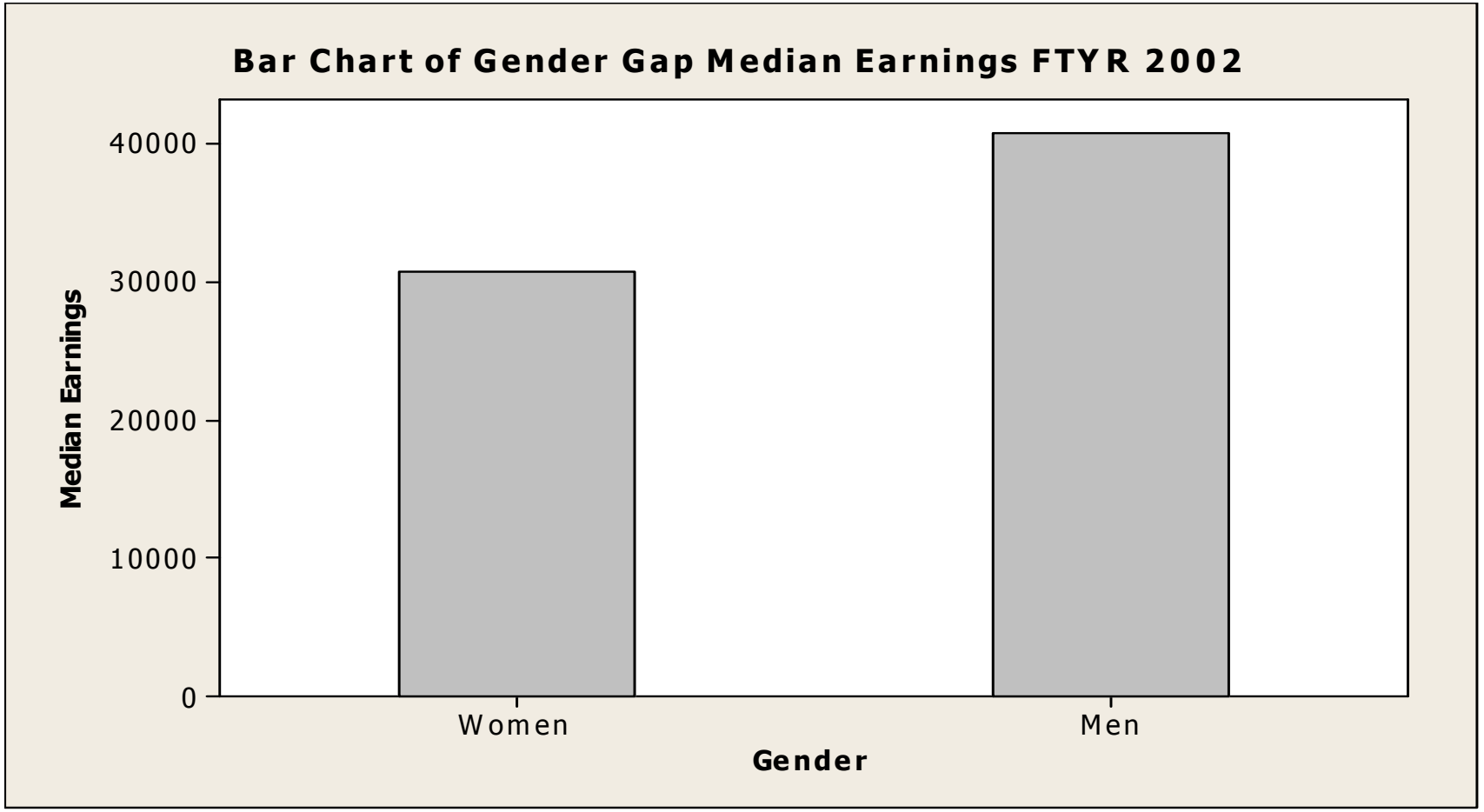
Beware of Missing Vertical 0



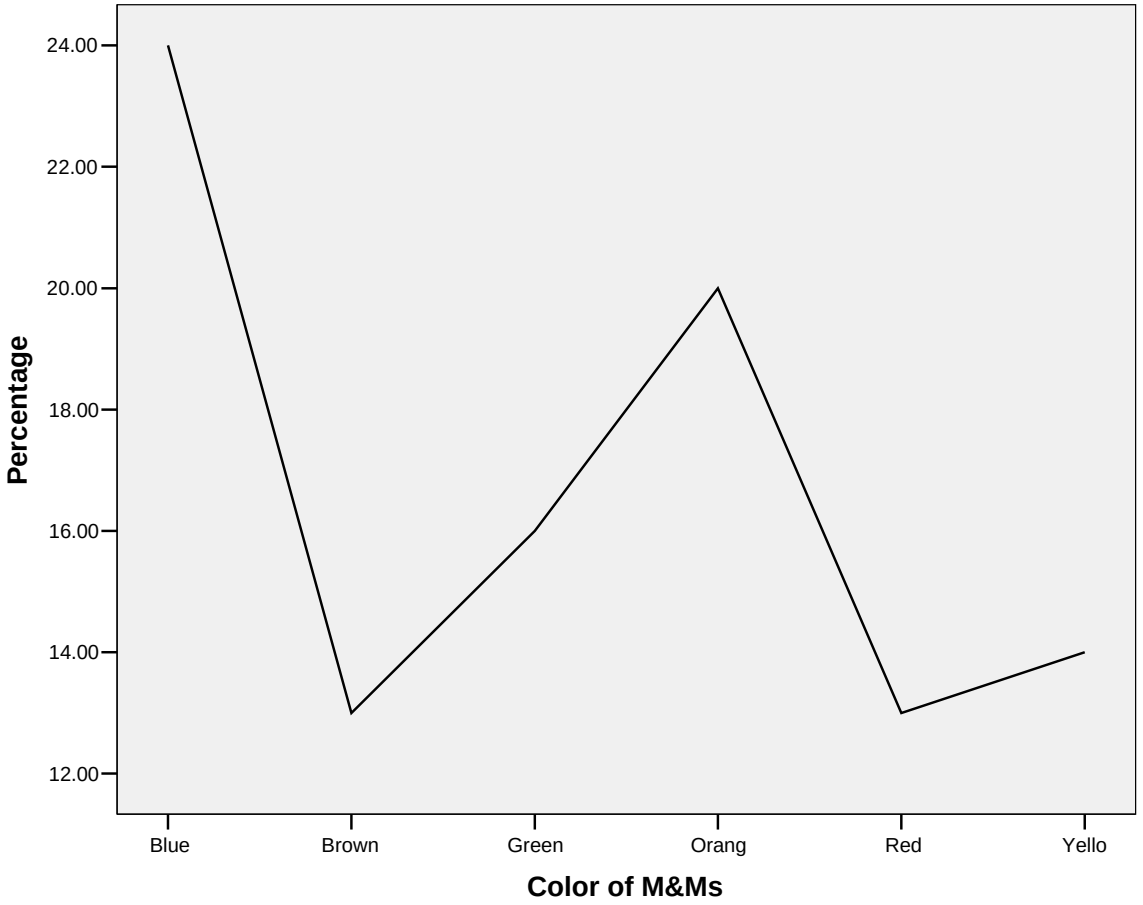
Worse Yet!


Bar Chart of Gender Gap Median Earnings FTYR 2002
Median Earnings
42000
40000
38000
36000
34000
32000
30000
Women
Men
Gender

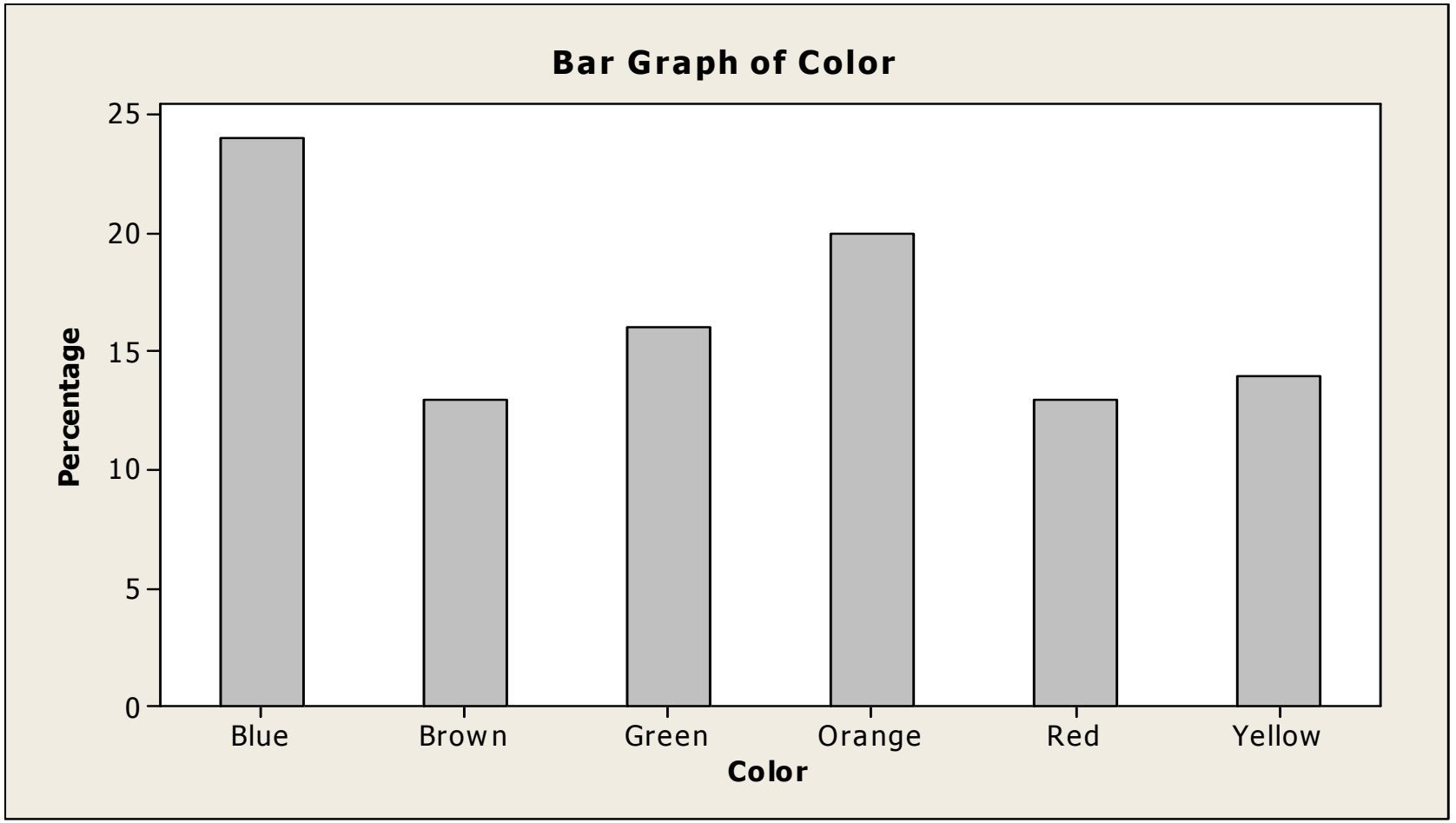
Correctly Done



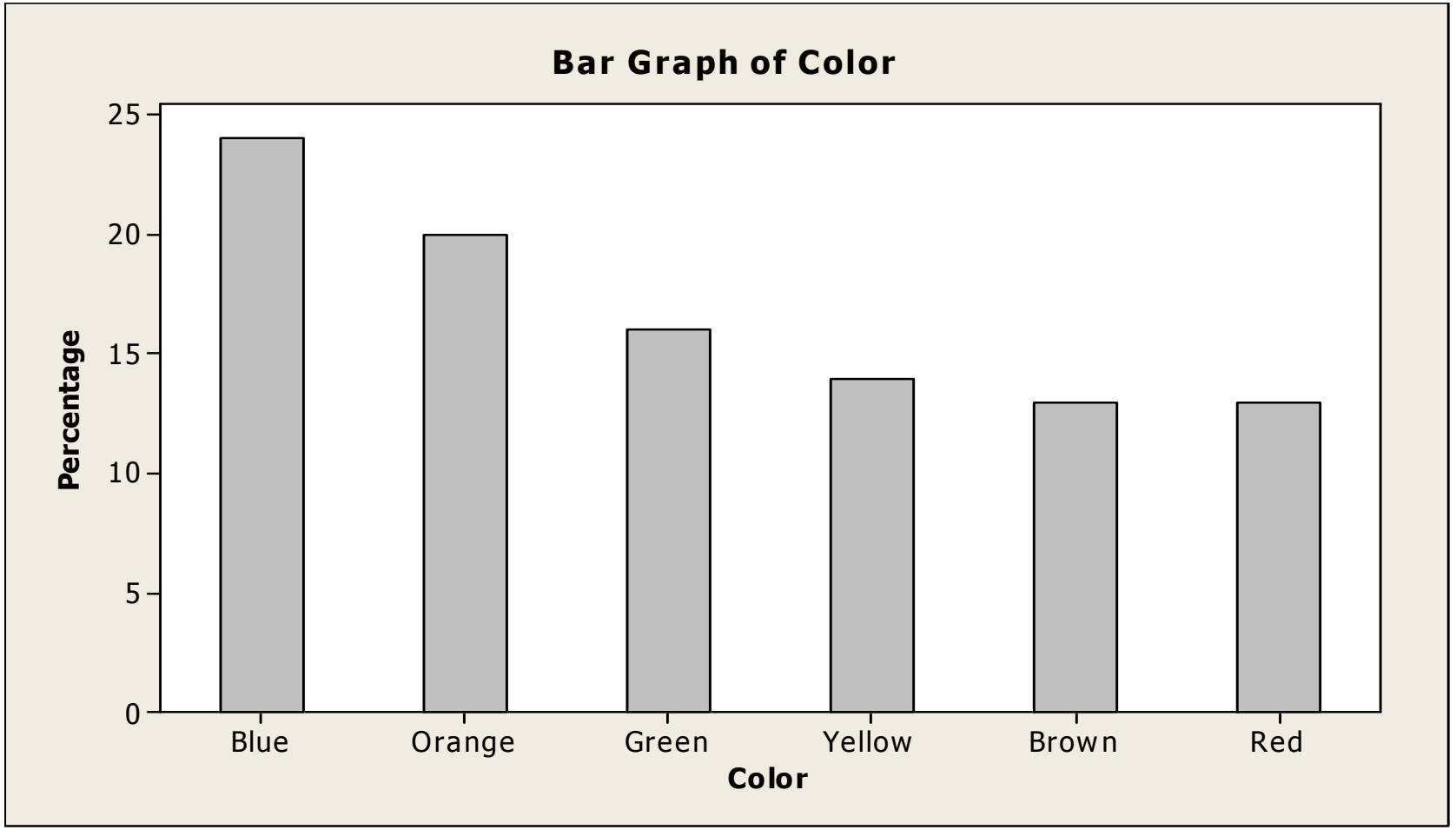
When Not to Connect the Dots



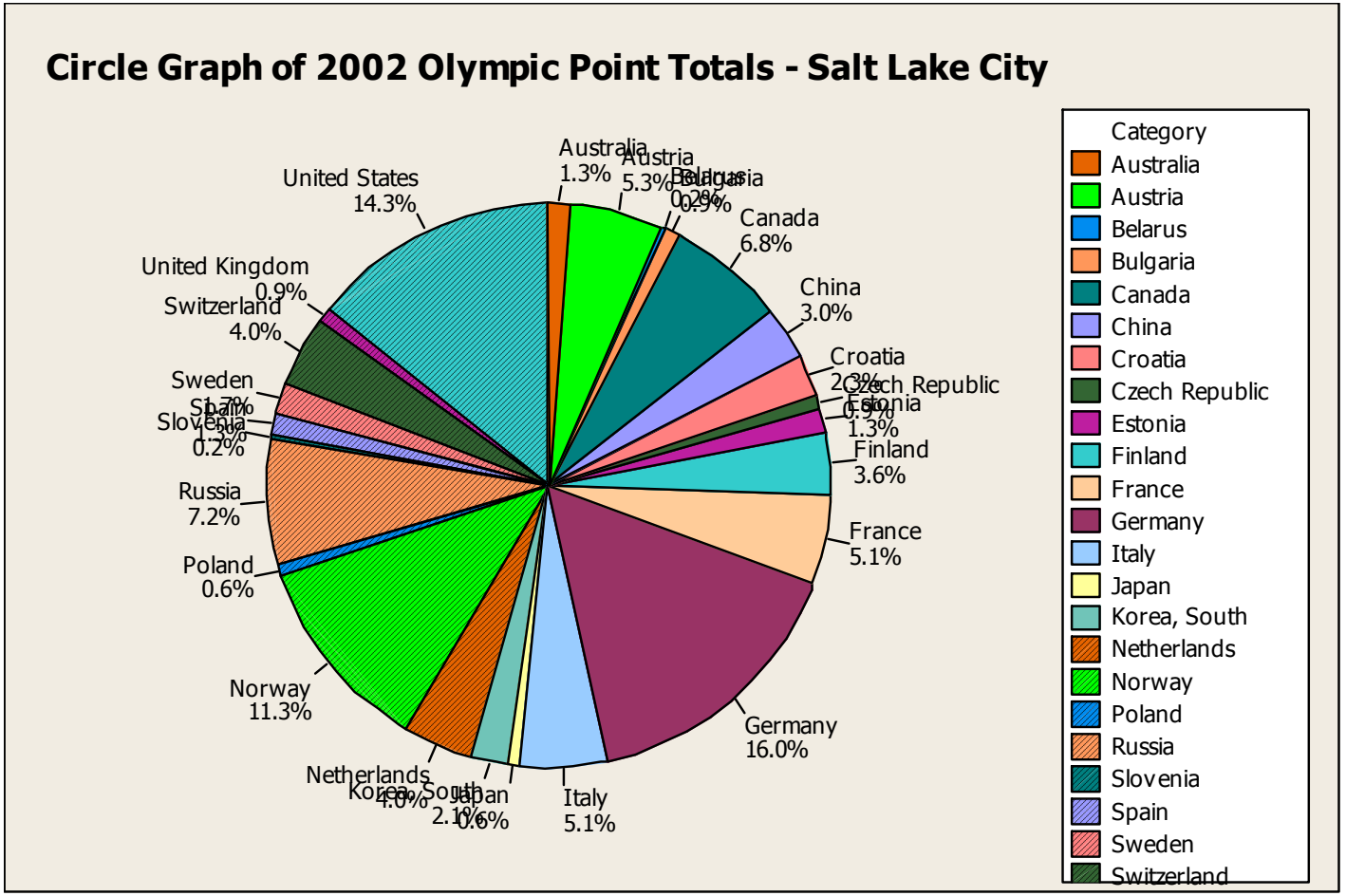
Correctly Shown



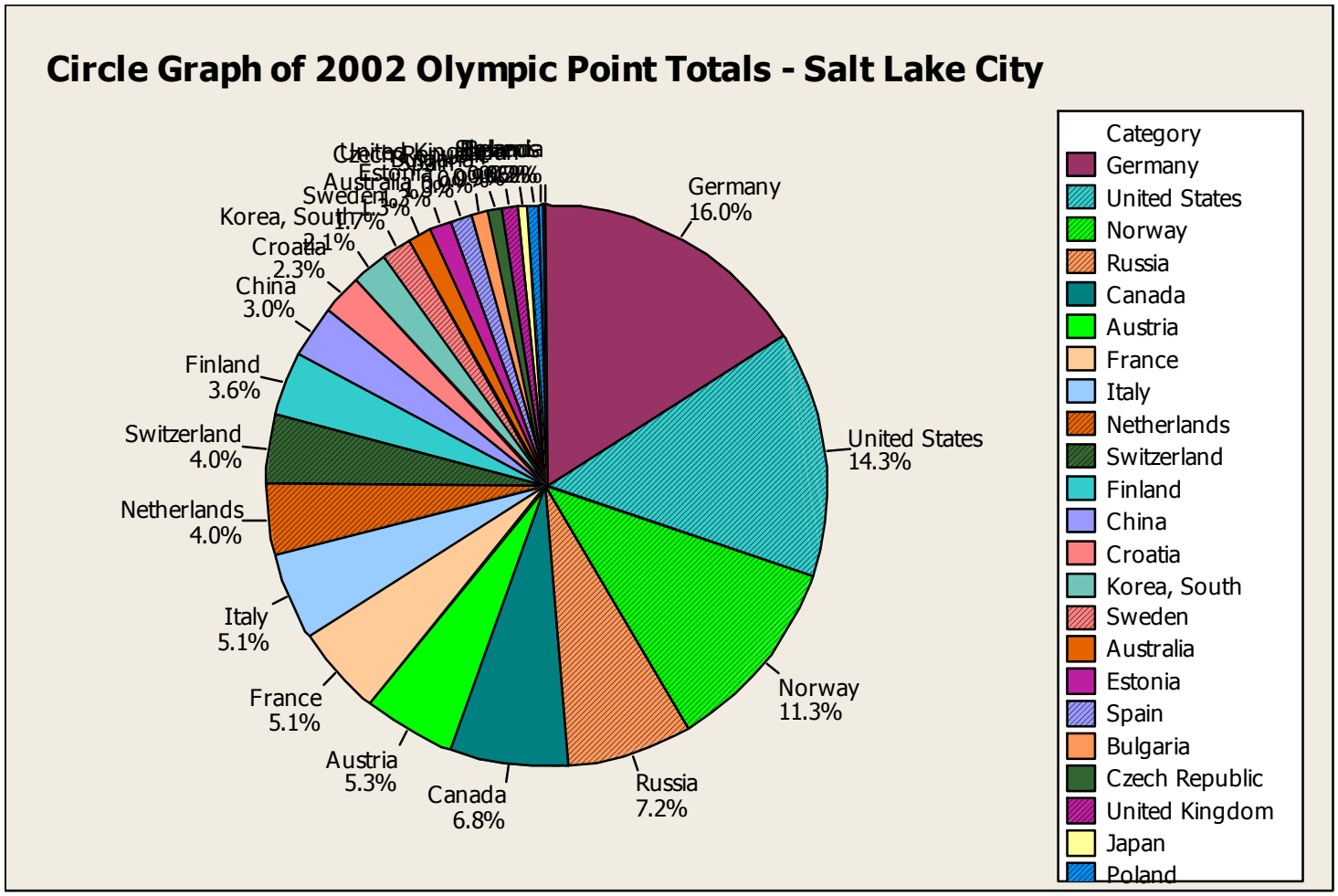
Better Yet!


Bar Graph of Color
Percentage
25
20
15
10
5
0
Blue
Orange
Green
Yellow
Brown
Red
Color

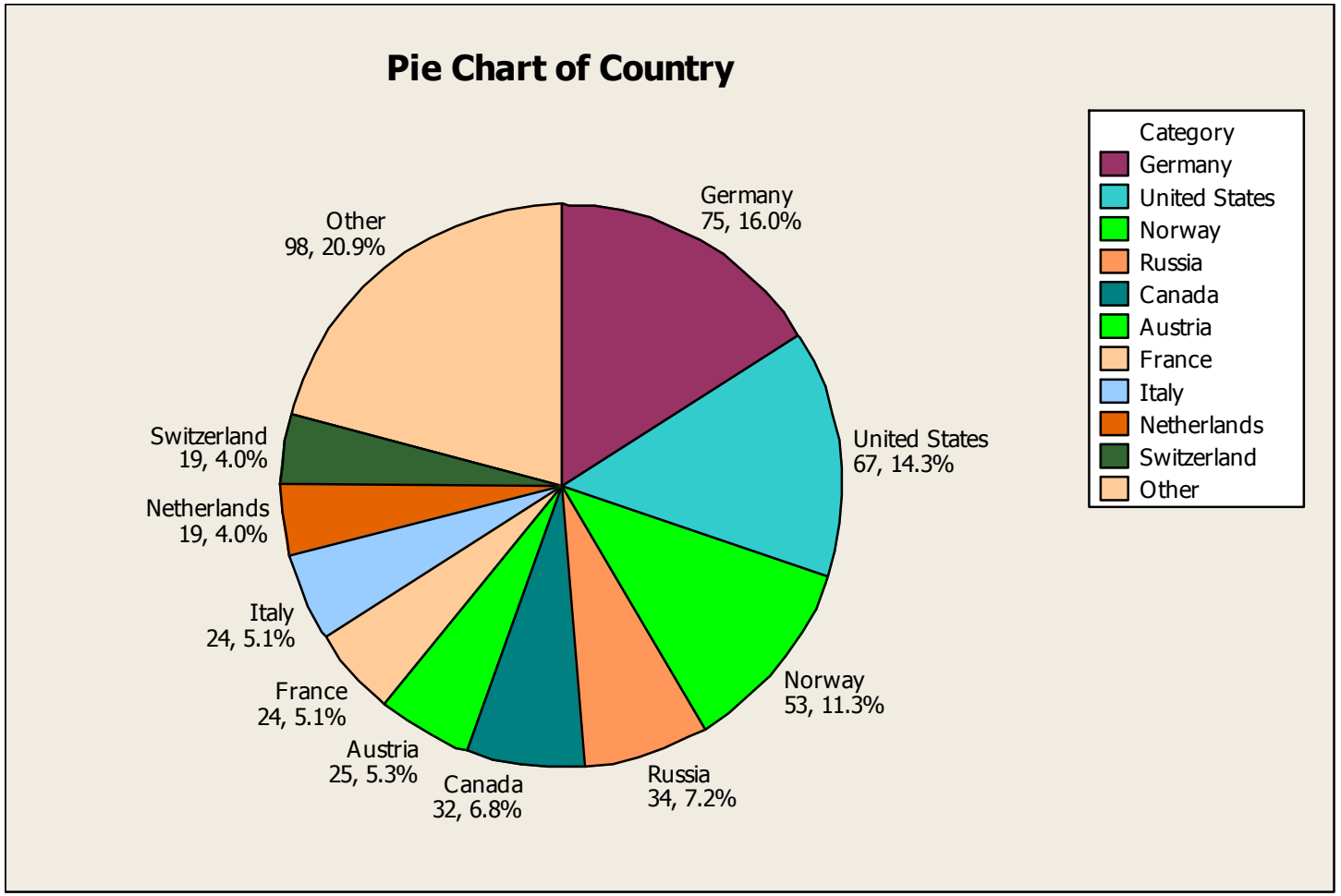
Too Many Pieces of the Pie!



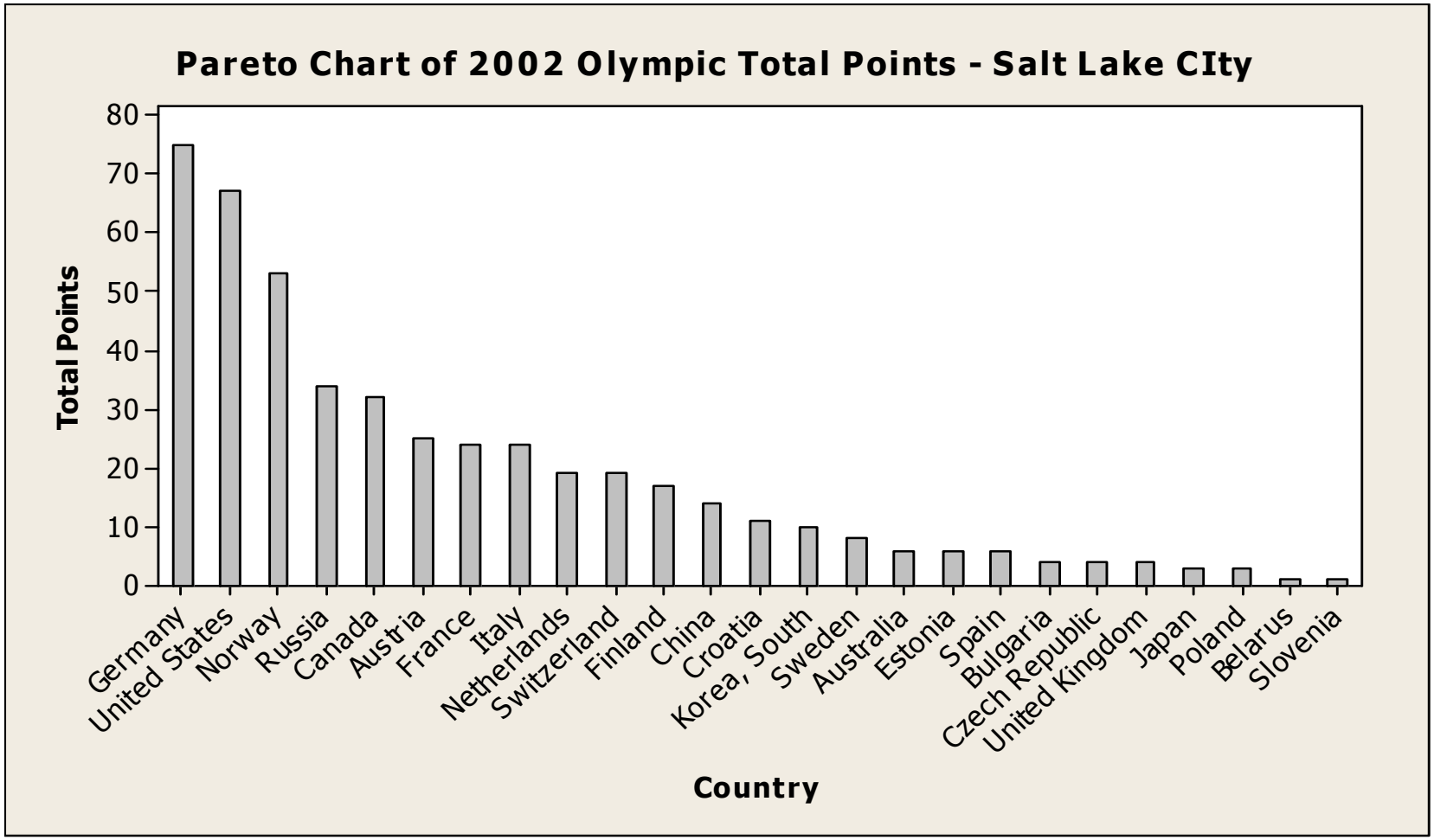
A Better Ordering, But Still “Busy”



The Top Ten Point-getters



Pareto to the Rescue

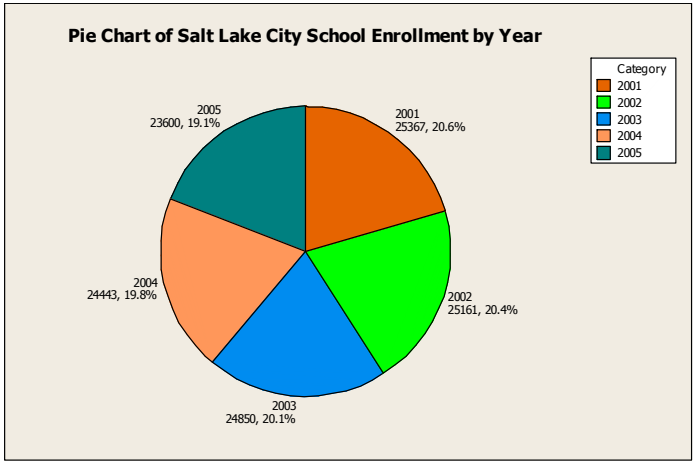


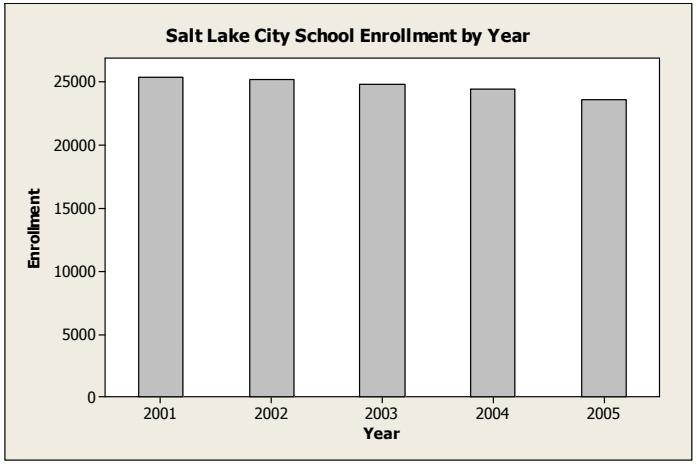
Use the Graph Most Appropriate for the Data

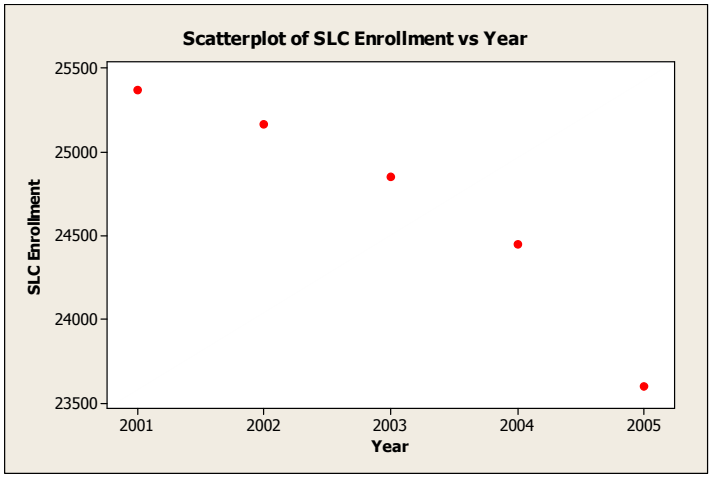
Salt Lake City School Enrollment

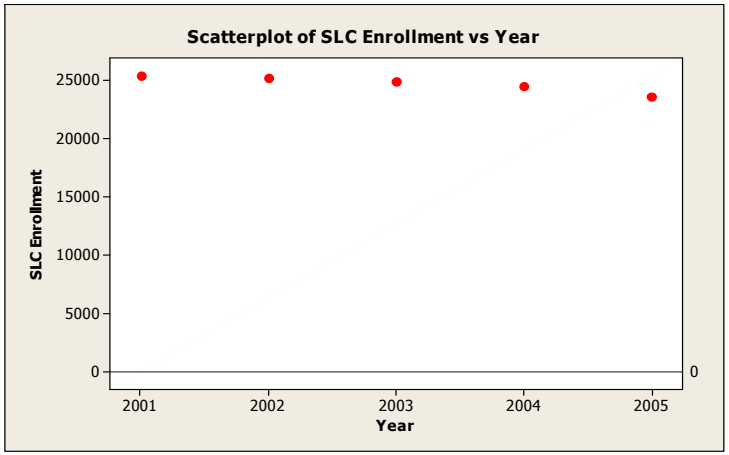
YEAR	ENROLLMENT
2001	25,367
2002	25,161
2003	24,850
2004	24,443
2005	23,600

Some Graphs to Critique









3-D Graphs – no! no!no!no!no!no!no!

KEEP AMERICA BEAUTIFUL

Focus: **Facts on municipal solid waste**

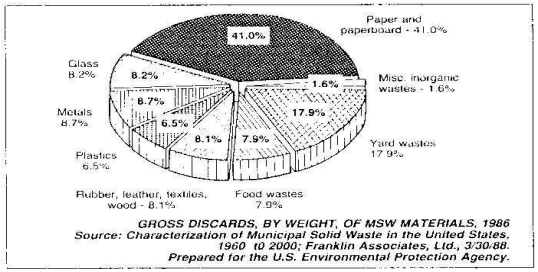
No. 1
December 1989

AN INTRODUCTION TO MUNICIPAL SOLID WASTE MANAGEMENT

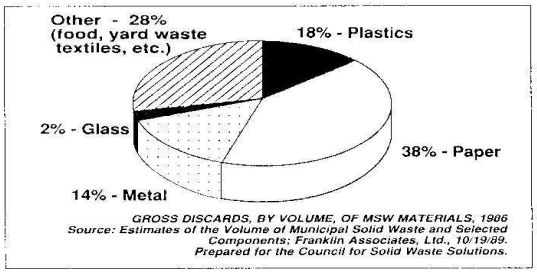
This issue of Focus is the first in a bi-monthly series on current topics relating to municipal solid waste management. Focus is published by the Keep America Beautiful, Inc. Solid Waste Task Force.

On average, every American "throws away" 3 1/2 pounds of municipal solid waste (MSW) a day. This figure is determined by dividing 160 million tons, the estimated amount of waste generated on an annual basis in the United States, by the total population. As the population increases, this amount is expected to expand to 193 million tons by the year 2000. It is important to remember that these quantities are estimates. While the true amount of waste may be more or less, estimates are useful tools in developing strategies to handle solid waste.

MSW consists of wastes from households, commercial establishments, institutions, light industry, and small quantities of special wastes such as from hospitals and laboratories.



GROSS DISCARDS, BY WEIGHT, OF MSW MATERIALS, 1986
Source: Characterization of Municipal Solid Waste in the United States, 1960 to 2000; Franklin Associates, Ltd., 3/30/88. Prepared for the U.S. Environmental Protection Agency.



GROSS DISCARDS, BY VOLUME, OF MSW MATERIALS, 1986
Source: Estimates of the Volume of Municipal Solid Waste and Selected Components; Franklin Associates, Ltd., 10/19/89. Prepared for the Council for Solid Waste Solutions.

COMPOSITION OF WASTE

The components of the waste stream have been analyzed by both their weight and volume as the charts above demonstrate. These charts serve as guides for the composition of waste. However, waste composition varies with type and size of each community, climate, and time of year.

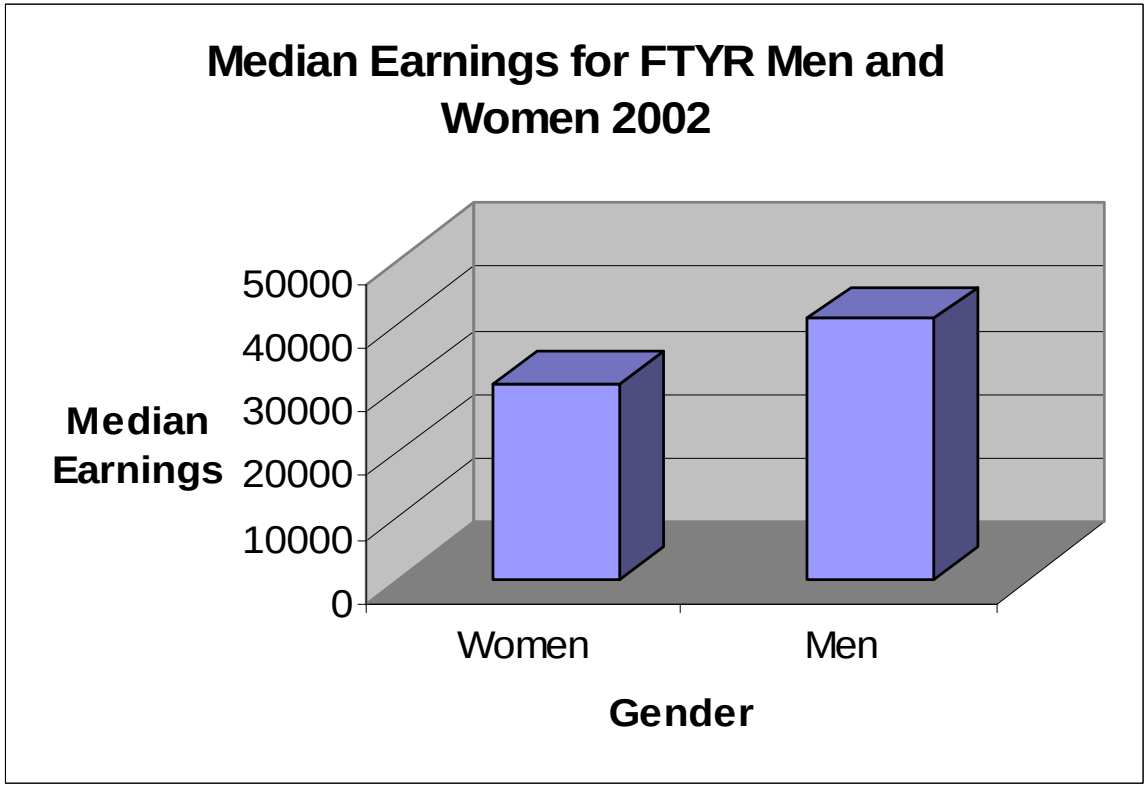
MEANS OF SOLID WASTE DISPOSAL

Historically, MSW has been disposed of in landfills; it was inexpensive, and large parcels of land were relatively easy to acquire. Today, landfills are closing at a rapid pace because they have filled up or were not constructed to meet current stringent environmental regulations. In 1986, there were approximately 6,000 operating landfills. By 1991, the U.S. EPA says that 45% will be closed leaving just 3,300 sanitary landfills remaining in operation.

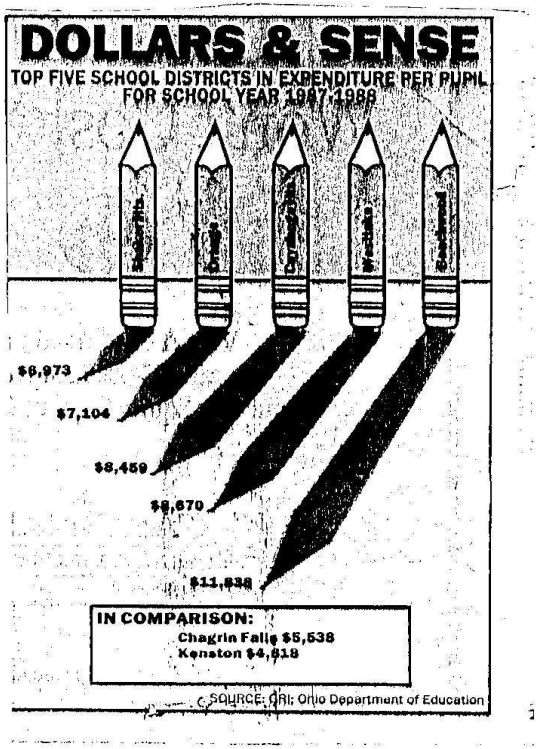
A nonprofit public education organization dedicated to improving waste handling practices in American communities.

3-D Graphs – no! no!no!no!no!no!

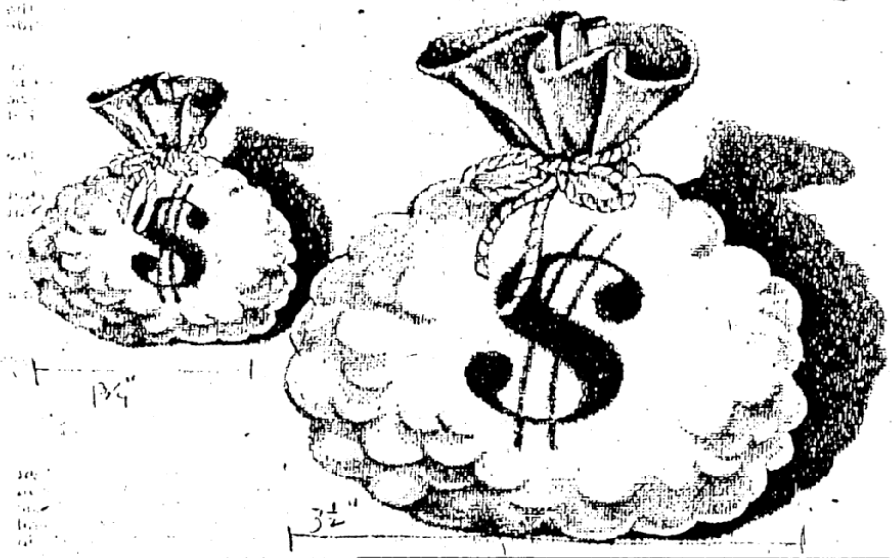
Recall the data: Men $40,668; Women $30,724



Pictographs –
An Easy Way to Lie with Statistics



Now you can double your money at Cleveland Federal Savings



WITH A CLEVELAND FEDERAL SAVINGS 7¾% per annum/$1,000 minimum certificate of deposit you can more than double your money in just 10 years.

There's no need to invest at an unknown rate (such as treasury notes or common stock) when you can be sure of a high yield. And you can choose from many options.

Our certificates earn from 5¼% to 7¾% per annum, depending on the maturity date.

At Cleveland Federal Savings our interest is in your future.

SAVINGS PLANS TO MEET YOUR NEEDS

5% per annum — DAILY INTEREST ON PASSBOOK AND CLUB ACCOUNTS

5¼% per annum — DAILY INTEREST ON GOLD KEY ACCOUNTS

ESLIC

CERTIFICATES OF DEPOSIT
($1,000 Minimum)

5¼% per annum	5¾% per annum	6½% per annum	6¾% per annum	7½% per annum	7¾% per annum
3-5 month maturity	6-11 month maturity	12-29 month maturity	30-47 month maturity	48-71 month maturity	72-120 month maturity

Regarding certificates of deposit, Federal regulations require a substantial interest penalty for early withdrawal.

Cleveland Federal Savings

AND LOAN ASSOCIATION

Downtown at 614 Euclid Avenue • Clark-West 25th • Parma • Westgate • Southland • University Heights
Euclid-Green • Memphis-Fulton • Westlake • Brooklyn • Richmond Mall • Maple Heights • Wellington
Aurora Commons • Ridge-Ridgewood • Mentor • Vermilion • Eton Square • Lakewood • Strongsville
Satellite Offices: Hamen's Rocky River, Cedar-Green, Cedar-Taylor, Aurora Commons, and Mentor Town Square Stores.

2/7/77

The Ohio Statistical Poster Competition

www.bio.ri.ccf.org/ASA/

Click on K-12 Programs
Click on Northeast Ohio Statistical Poster Competition
All rules and entry form are there. Deadline March 20.

There's a link to the national competition and pictures of past winners.
Ohio is a region. Other regions include Pennsylvania, Michigan, Connecticut, Utah, Colorado/Wyoming, Washington DC, Nevada, New York. Regional winners are submitted to the national. Entries that are not part of a region are sent to Alexandria Virginia composing the "Other" region.

Some Reasons to Have a Statistics Poster Competition

- **Statistics and statistical concepts are an integral part of the school curriculum.**
- **A statistics poster can be across the curriculum incorporating topics from science, social sciences, health, art, English, mathematics.**
- **A poster gives students a viable creative way to display their statistical knowledge.**

Categories and Prizes

- **The competition has four grade categories: K-3, 4-6, 7-9, and 10-12.**
- **For grades 4-12, entries may be done by teams of one to four students.**
- **For K-3, an entire class may work on an entry.**
- **There are cash prizes at the Ohio and national levels.**

What is a Statistics Poster?

- **Posters tell stories!**
- **A statistics poster states a question that can be answered by data, and displays two or more related graphs that summarize the data, show different points of view of the data, while visually answering the stated question(s).**

The Poster

- **A poster must be able to stand alone – without a narrator to tell the story, or a report to discuss the data.**
- **Viewers must be able to understand individual graphs, the relationships among the graphics, and how the graphics address the central question.**

Judging Criteria

The formal criteria are:

1) Clarity of Message

2) Appropriateness of the Graphics

3) Details of the Graphs

4) Creativity

5) Overall Impact

1) Clarity of the Message

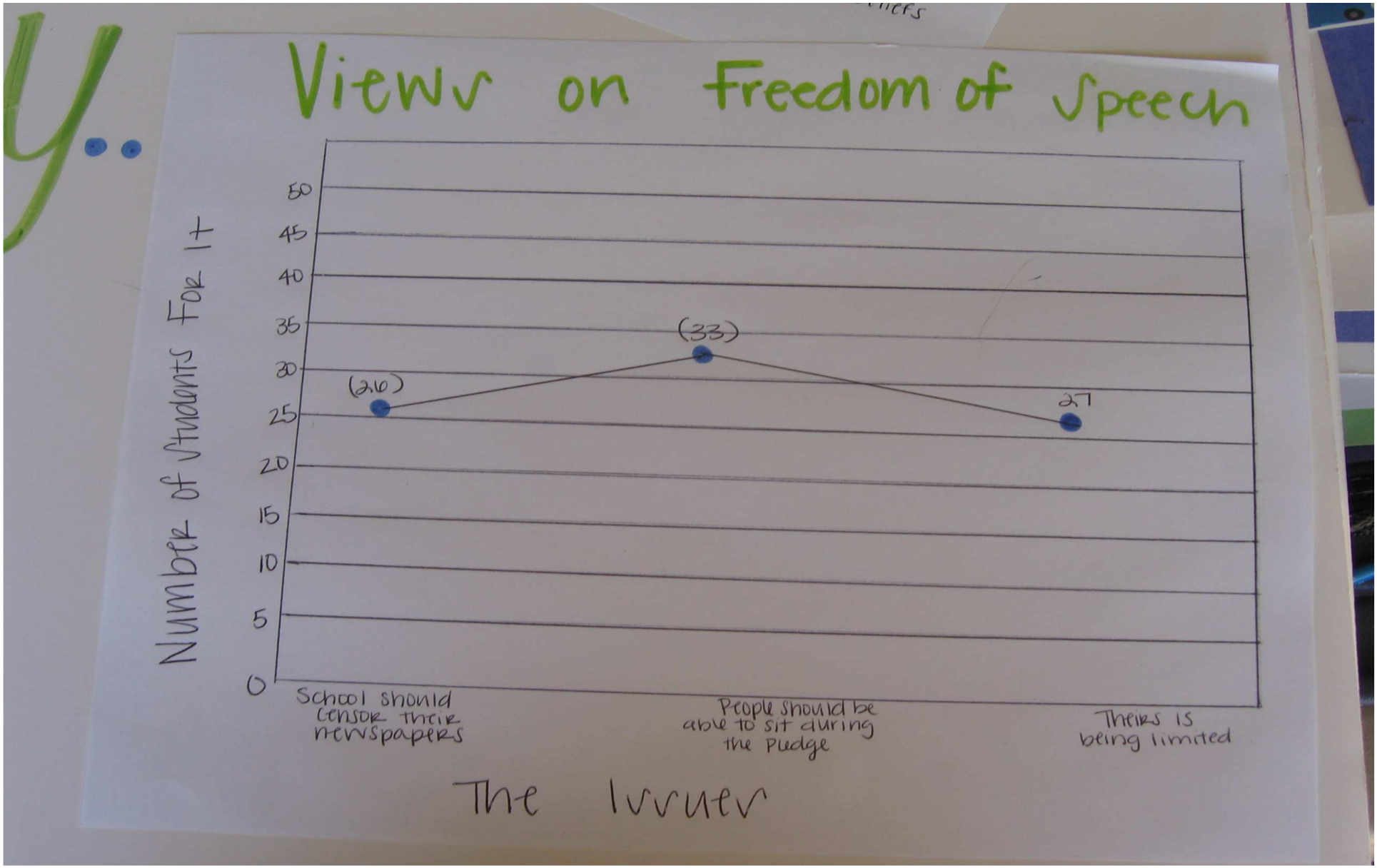
- **What is the story behind the poster?**
- **Do the graphs interconnect? How does each graph contribute to the story?**
- **Is the central message prominently and clearly presented?**
- **Are the conclusions obvious from the graphs?**

The circle graph is hardly visible. Also there is too much chart junk!

2) Appropriateness of the Graphics

- **Were the correct data collected to answer the question?**
- **Are the graphs appropriate for the type of data used, continuous or categorical?**

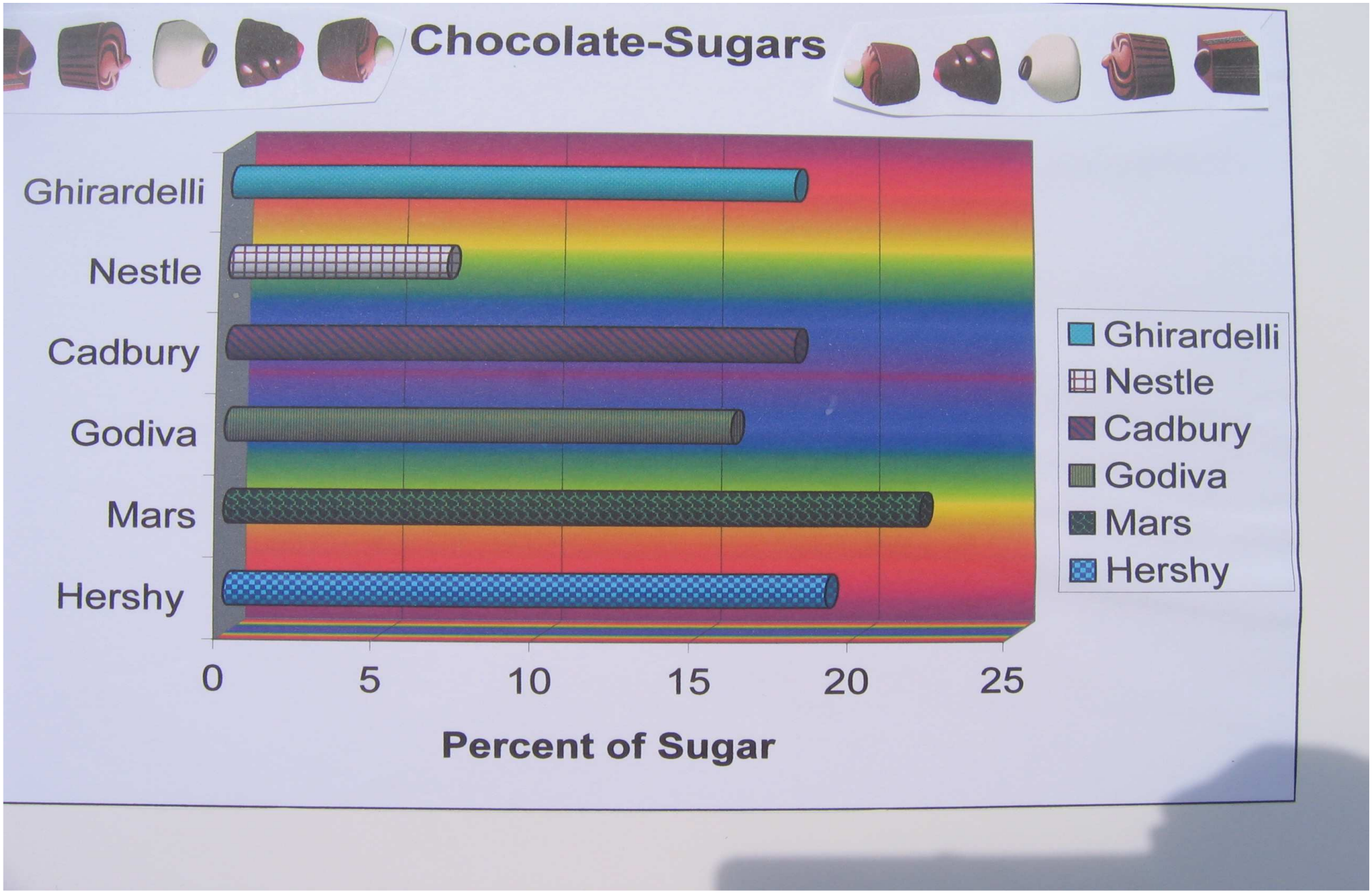
Line graphs are not appropriate for categorical data



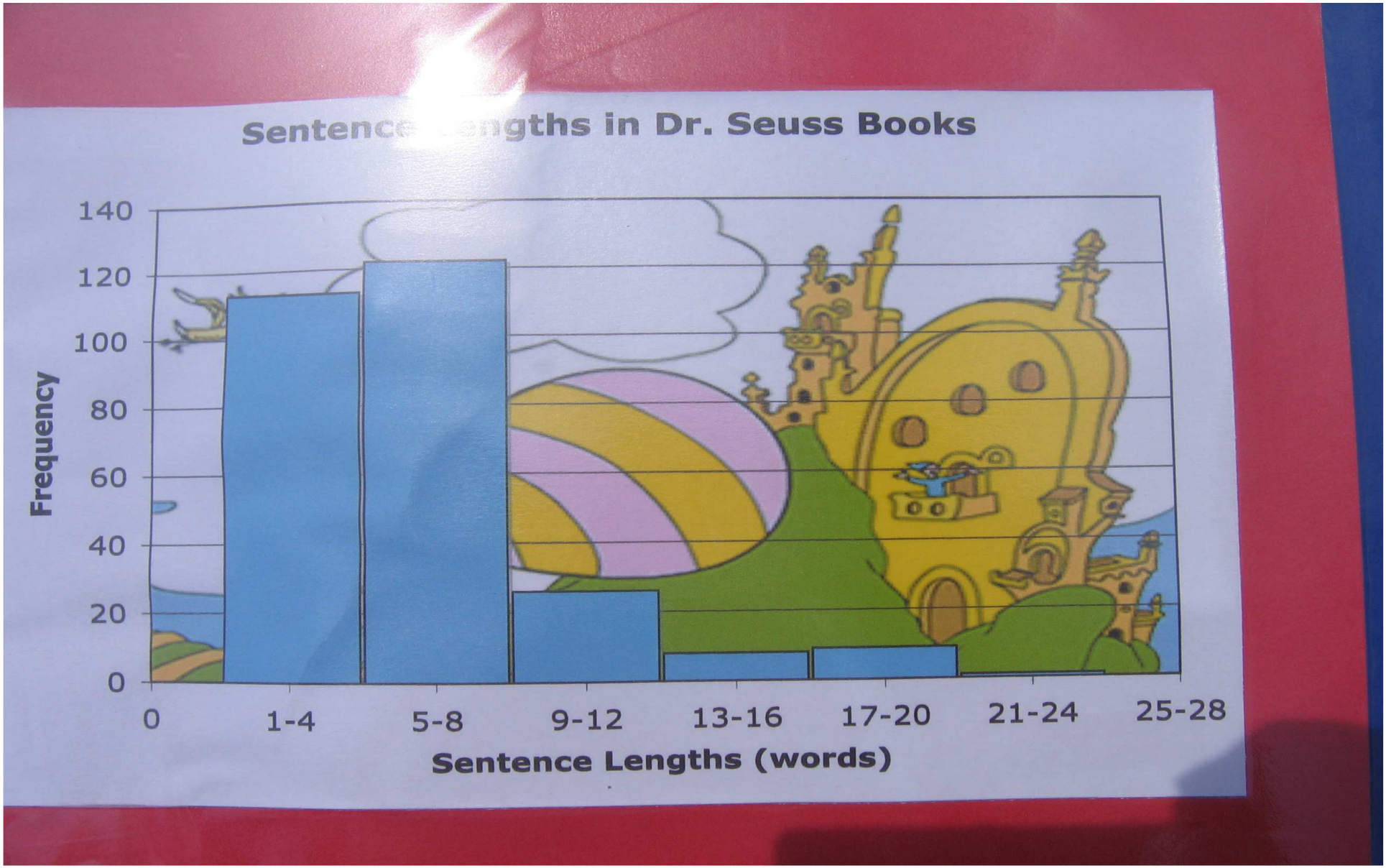
3) Details of the Graphs

- **Does each graph have its own title, labels, and legend?**
- **Are axes scaled correctly and identified?**
- **Are colors (if used), patterns, and scales used consistently and effectively?**

Hardly a good use of color



Good topic, clear, correct scales, distracting color background ruins what is otherwise a very nice poster

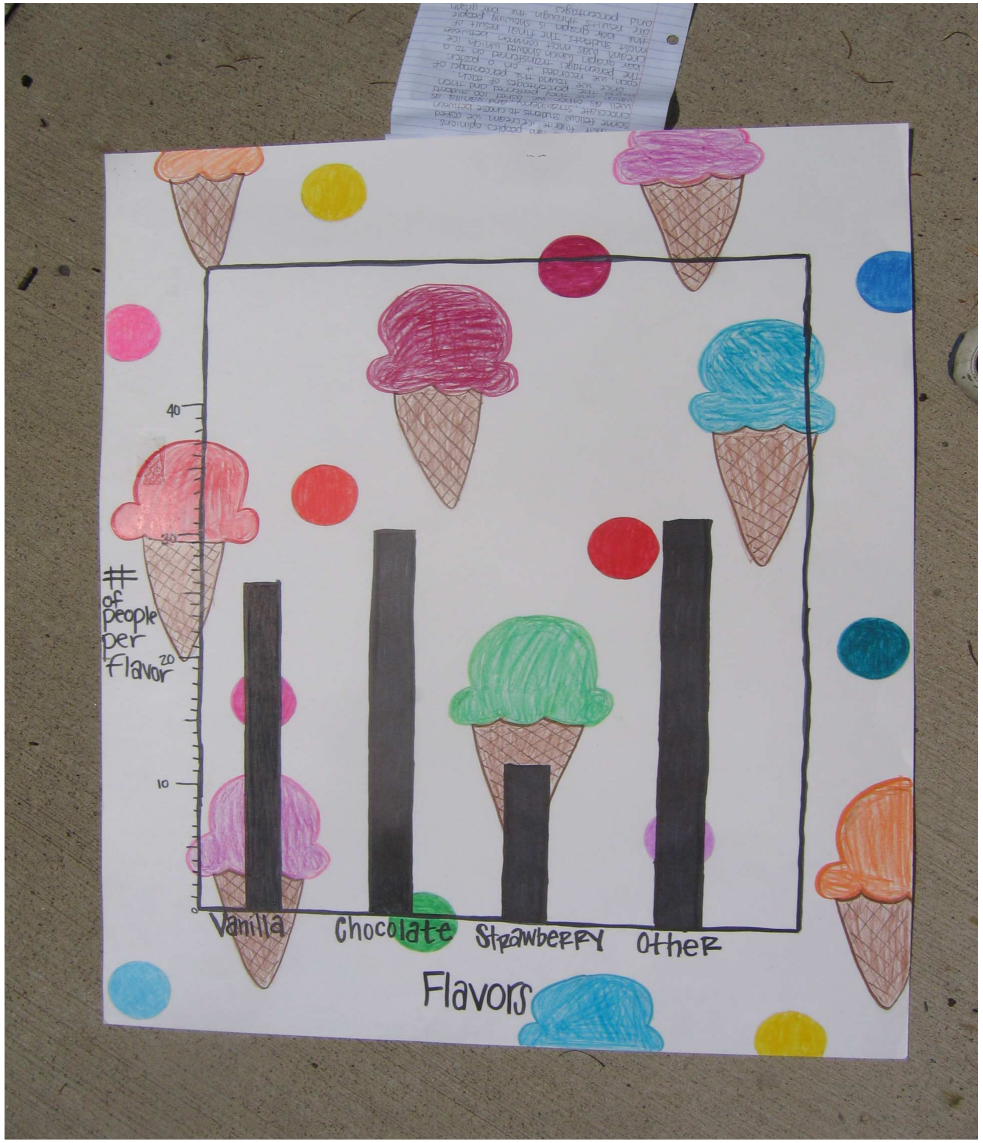


4) Creativity

- **Is there any evidence of creative thought or imagination?**
  - **In the choice of topic**
  - **In how the data are collected**
  - **In how the data are displayed**

  - **Is the topic appropriate to grade level? (have to be careful here)**

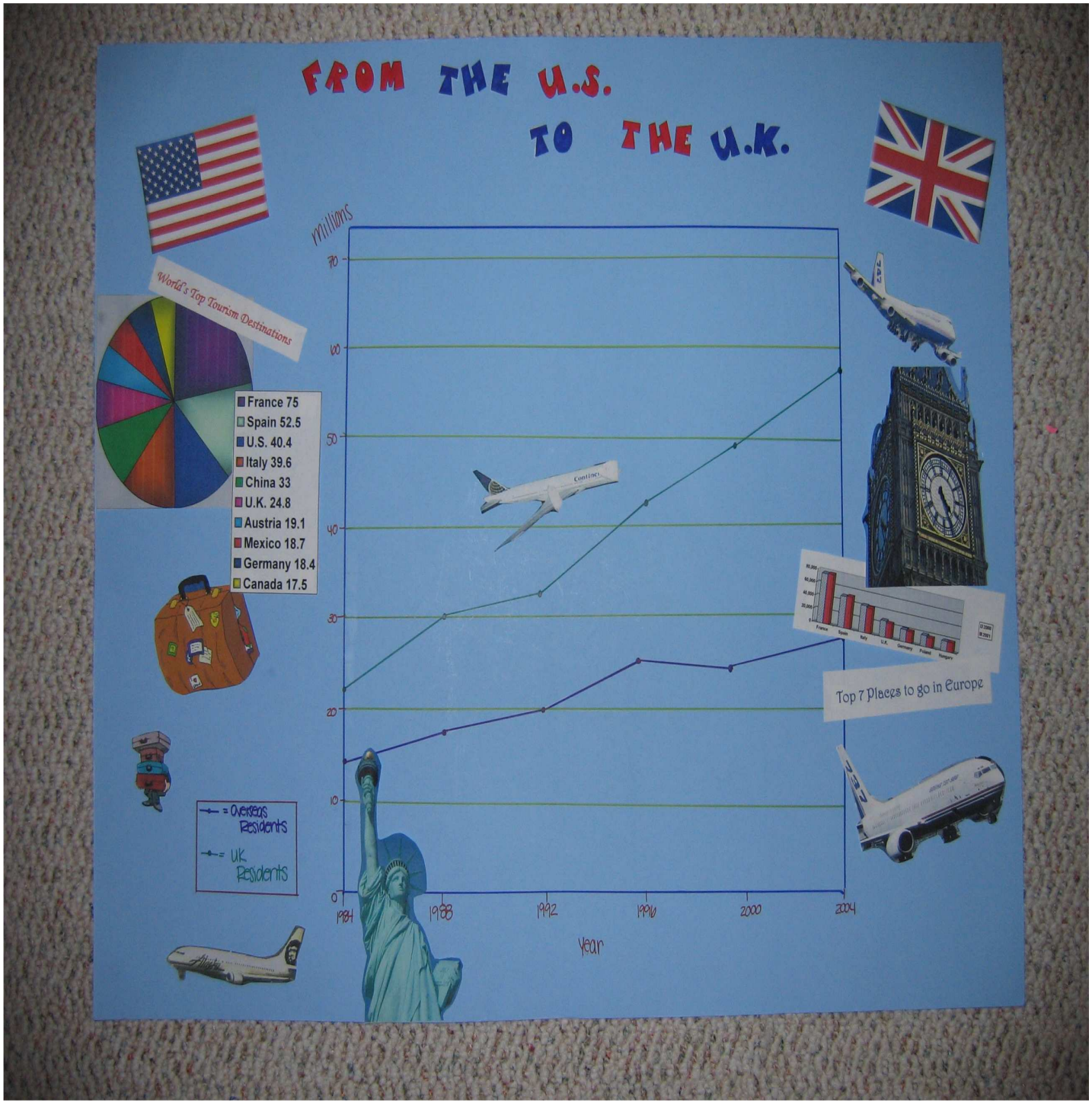
High school entry



First grade entry

5) Overall Impact

- **Is the poster eye-catching and visually attractive?**
- **Is the poster neatly done? Is it uncluttered and well-organized?**
- **Do the graphs take up the majority of the poster space? (about 75%)**
- **Is it readable from six feet away?**
- **Is the poster interesting?**



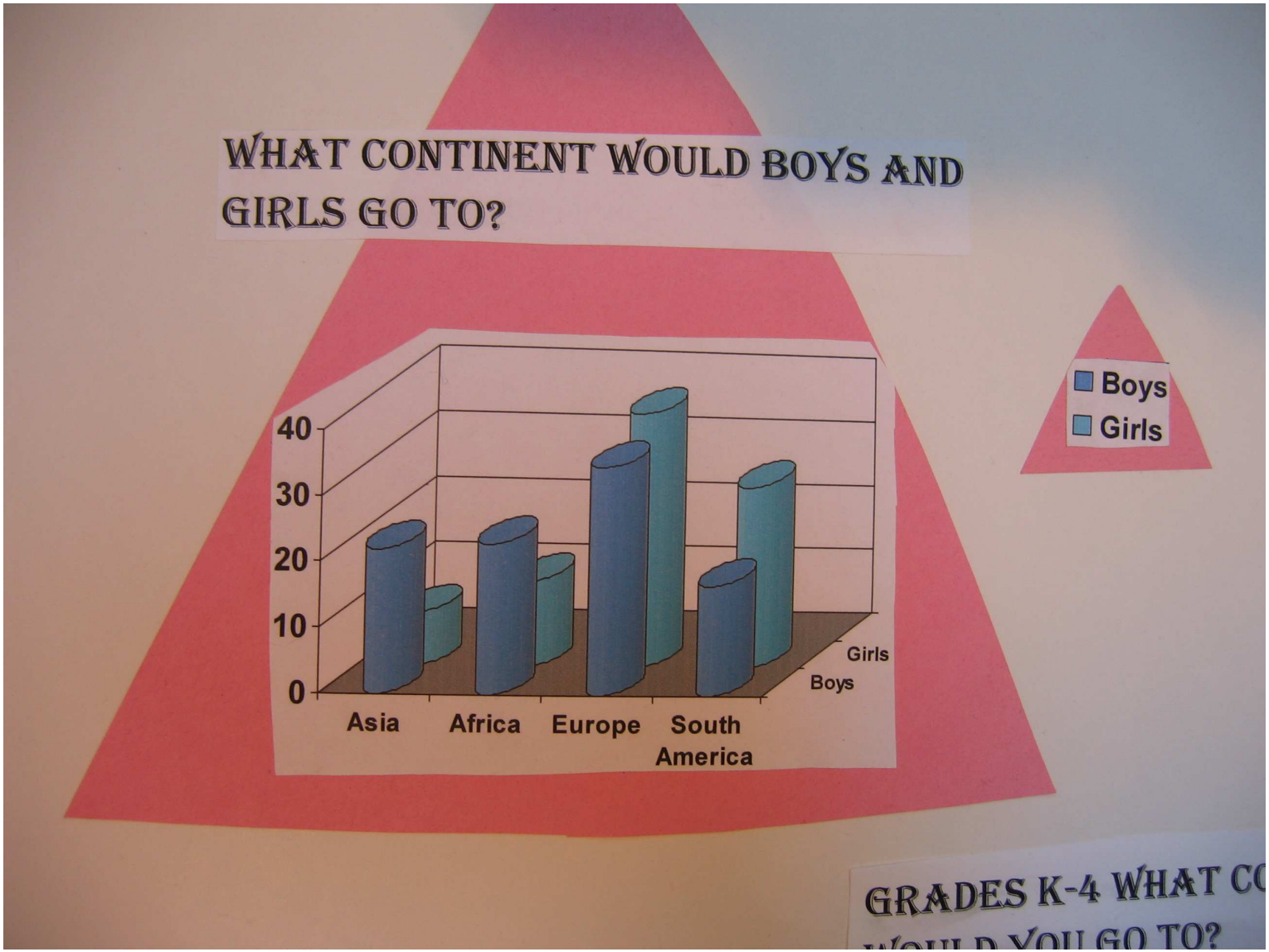
V
E
R
Y
C
O
N
F
U
S
I
N
G

OTHER ISSUES

3-DIMENSIONAL GRAPHICS

- **While 3-D may produce an interesting picture to look at, they are often misleading statistically.**
- **If the third dimension is meaningless, don’t use it. Use a 2-D graph.**
- **3-D graphs with a meaningful third dimension are often still difficult to figure out. So, be sure they are clear.**

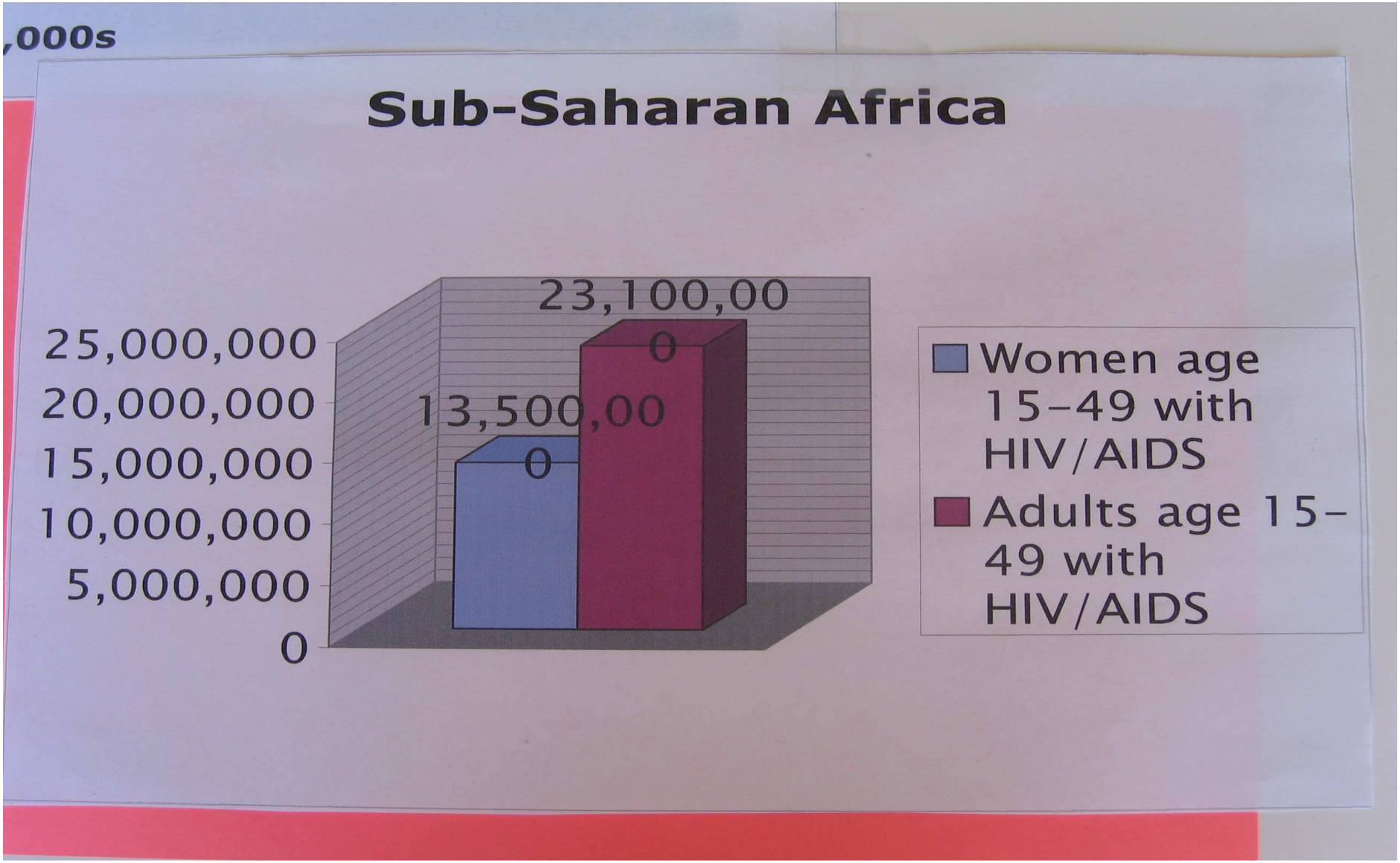
3-D Graphs should have a third variable dimension. They should NOT be used just for effect! 2-D side-by-side bar graph, better.



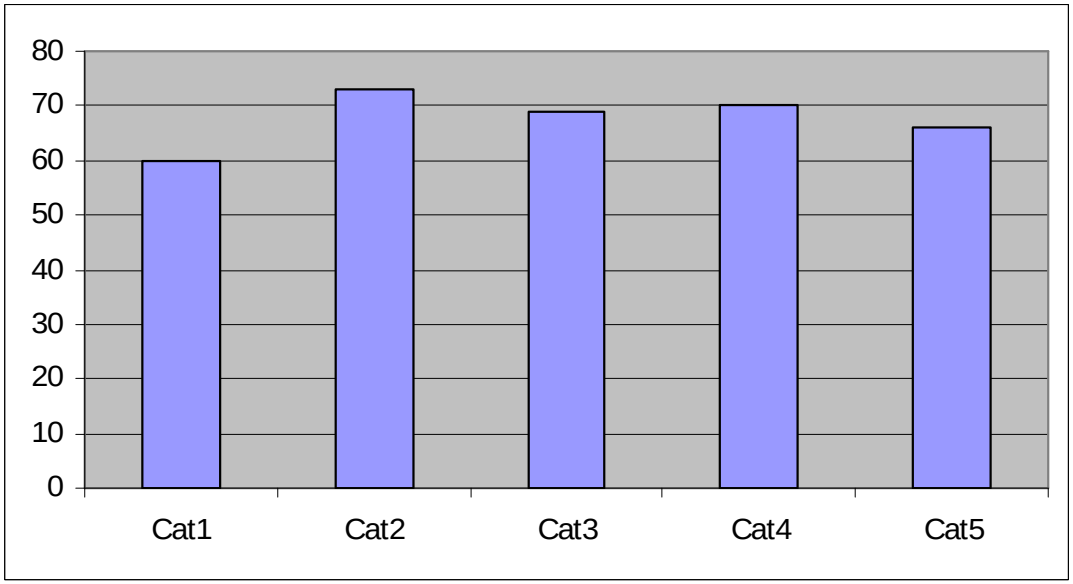
COMPUTER vs. HAND DRAWN

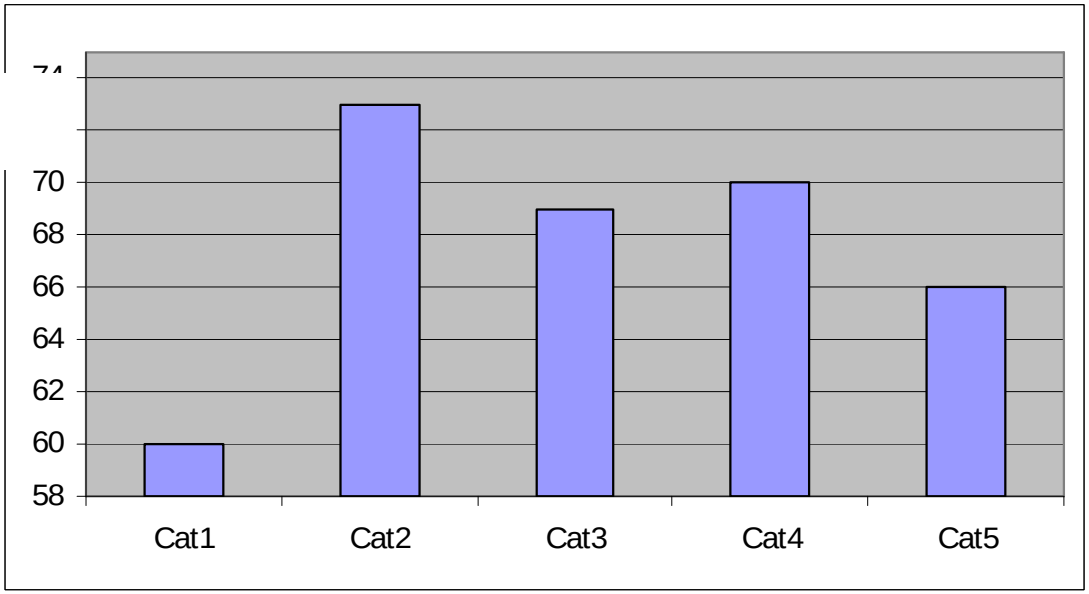
- **The default labeling in computer-created graphs is often too small to read.**
- **Computer software can sometimes default scale axes with units that make no sense.**
- **Hand drawn should be neat and scaled properly. Straight lines should be done with a ruler and circles with a compass.**

Here's unnecessary three dimensions and the label is poorly formatted.



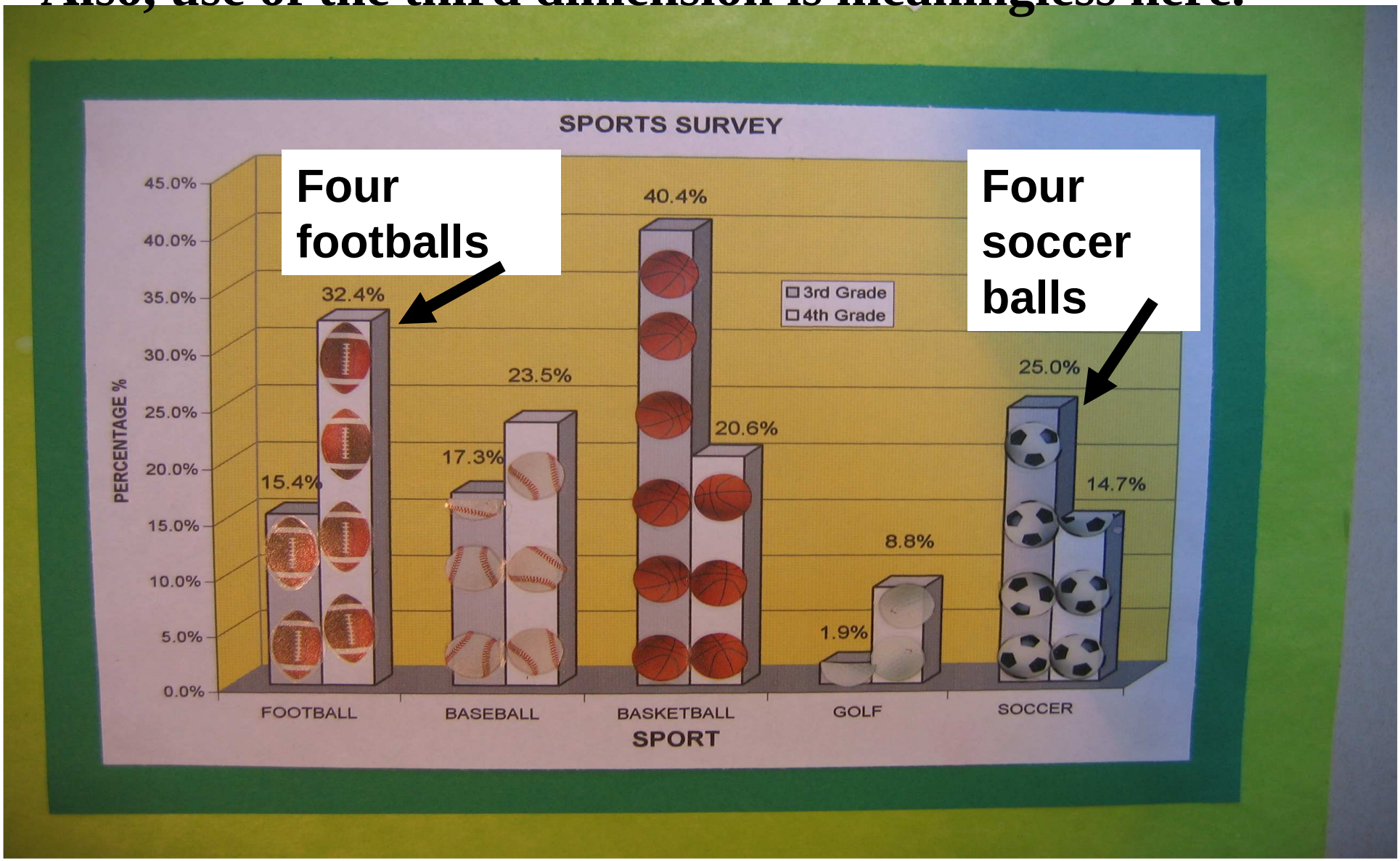
Bar charts should include a meaningful vertical 0.



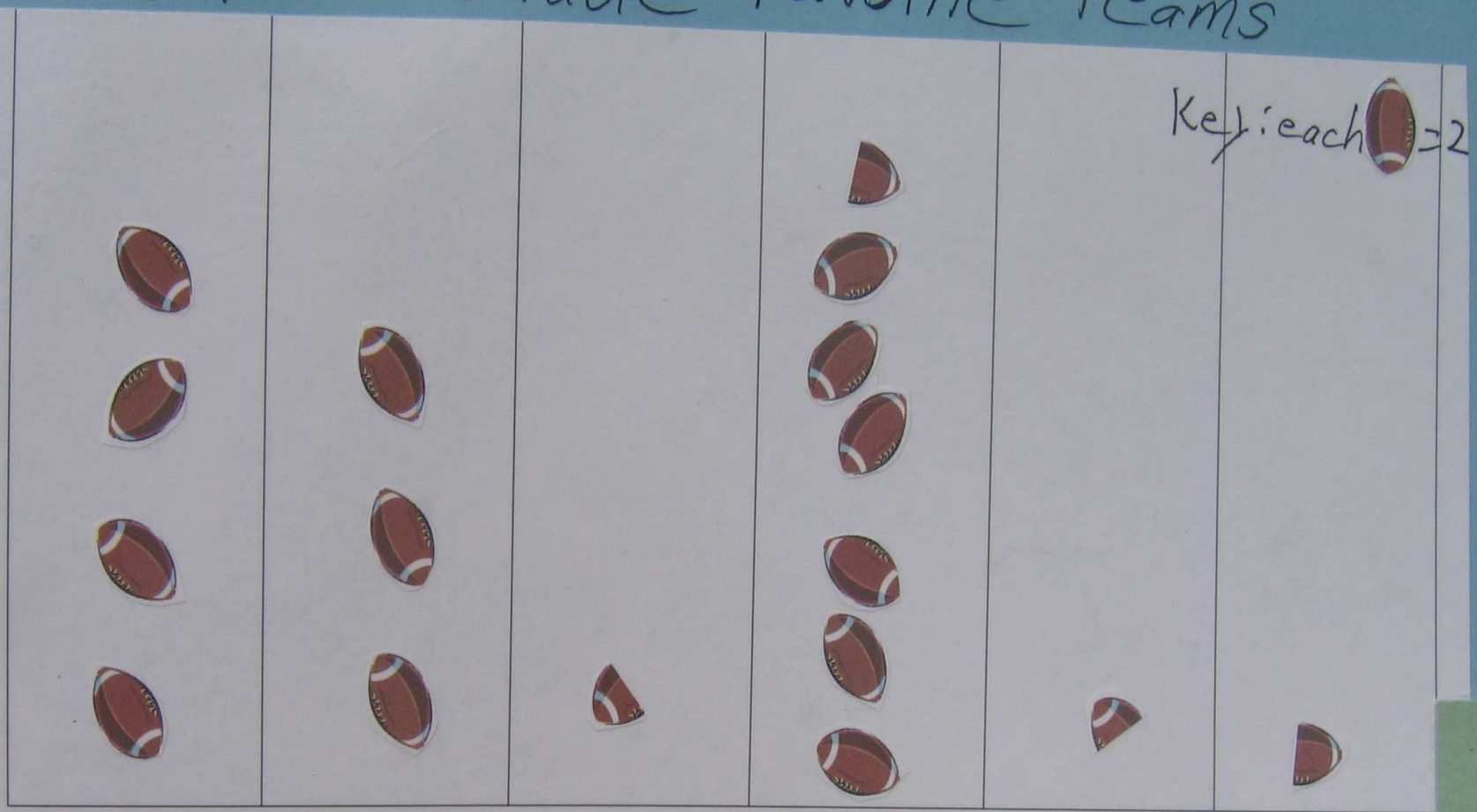


Pictures in a pictogram need to be the same size.

Also, use of the third dimension is meaningless here.

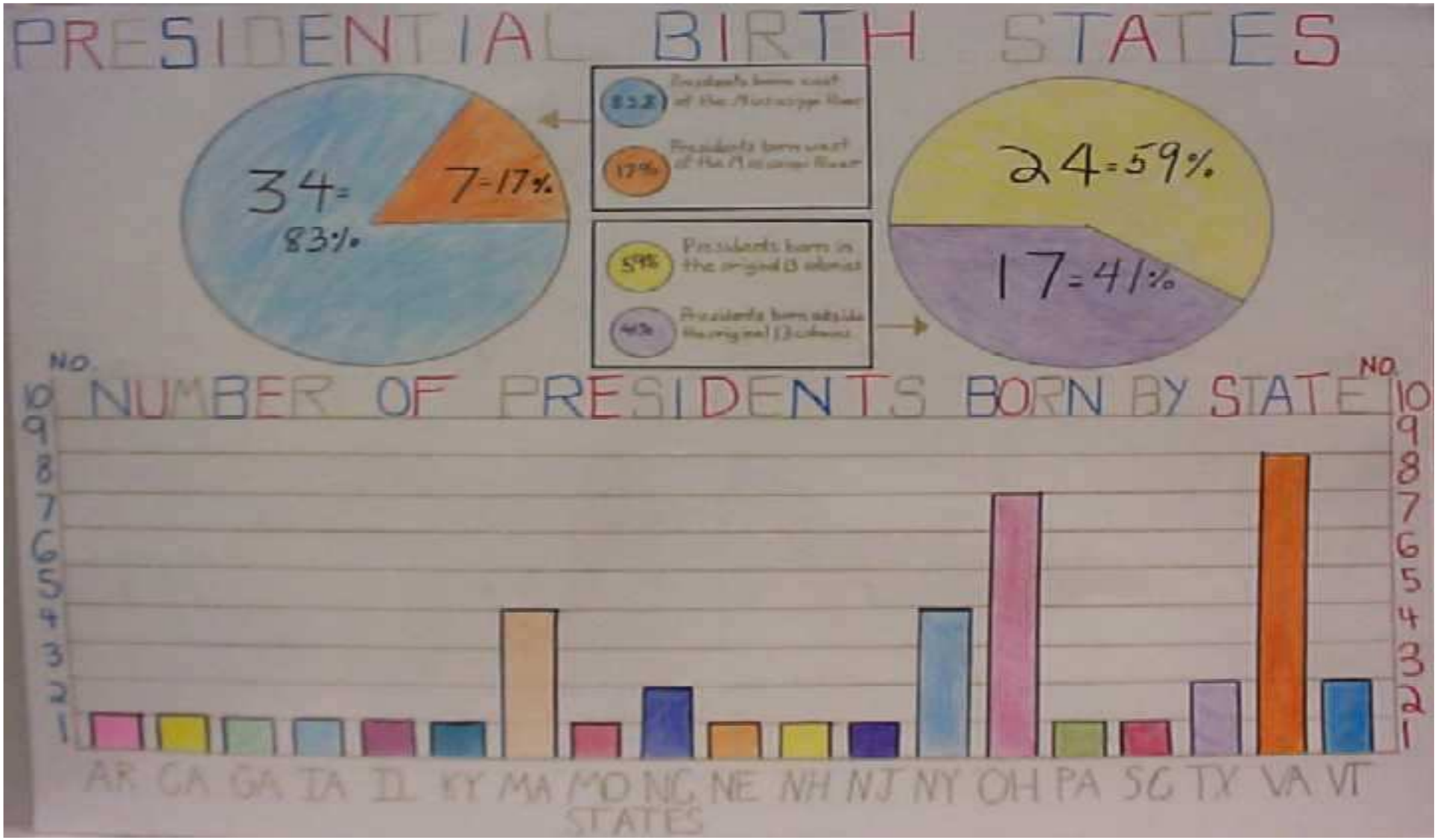


Third Grade Favorite Teams



Can’t compare products: Cans are different sizes, are not lined up, do not have a common starting position.

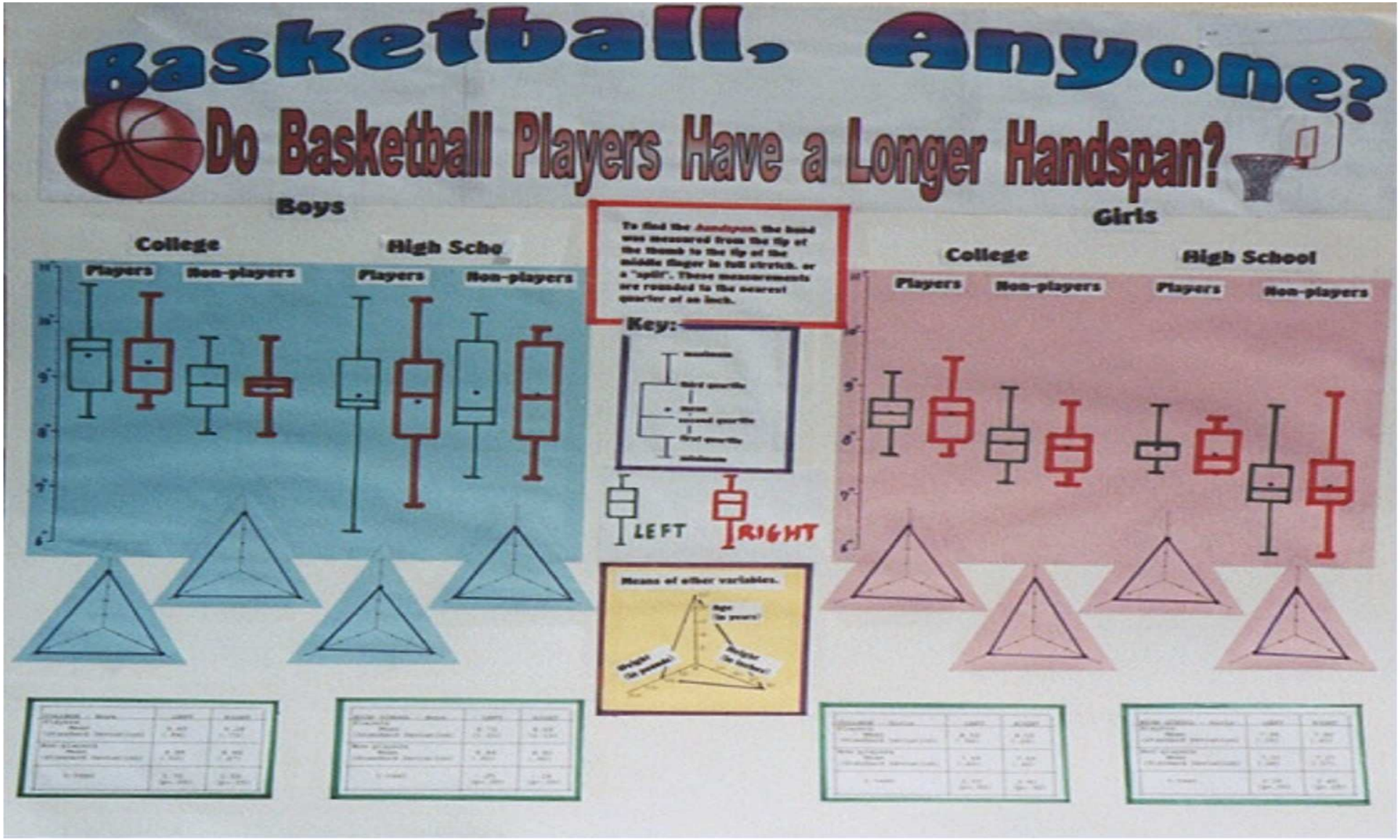
No scale – what does a half of a can mean? Or a whole can?



K-3 A winning poster.
There were multiple questions answered about where our Presidents have been born: What state? East or west of the Mississippi? Original 13 colonies or not? Very neatly drawn and well labeled. Good choice of Graphs. States arranged alphabetically. Nice use of color. Vertical axis numbering could be improved. Pareto ordering might be better.

4-6 A winning poster.

This poster is about how large people's hands are. There are several related questions here. Are basketball players different than non-athletes? Are boys different than girls? Are college different than high school? Hand-drawn is very neat and legible. The triangular graphs are creative but slightly difficult to



Resources

Exploring Statistics in the Elementary Grades Book One and Book Two
www.pearsonschool.com/index.cfm
Click on **curriculum,** then **elementary,** then **mathematics.**
Under **"Find a Product,"** type in **Exploring Statistics**

Navigating through Data Analysis and Probability in PreK – 2
Navigating through Data Analysis and Probability in Grades 3-5
National Council of Teachers of Mathematics publications

An Oldie but Goodie**: *Used Numbers*** (primarily data analysis) See Pearson as above.

Coming Attraction **– *STEW, Statistics Education on the Web*,**
an ASA peer-reviewed electronic journal, first edition spring 2009 (www.amstat.org)

Coming Attraction **- *GAP*, *GAISE Activities Project* ,**
a product of the ASA/NCTM Joint Committee on Curriculum in Statistics and Probability for Grades k-12, watch for a publication this fall **(www.amstat.org)**

The *GAISE Report for Grades Prek-12* (free download, or $15 purchase of bound volume at www.amstat.org)

www.ixl.com/math/practice/ (practice in many math skills; free 30 day trial; subscription)

QuizBus www.pbs4549.org/quizbus (Dealing with Data for grades 4 – 8; connects to OACS)

Don't hesitate to contact me if you are interested in workshops in mathematics, data analysis, probability through CMSETT: the Center for Mathematics and Science Education, Teaching, and Technology at John Carroll University. Linda Gojak, Director; Pam Keiper, Associate Director.

Also, if you would like for me to visit your classroom regarding the poster competition, call!

THANKS!