Relative Population

Start with a paper bag, of 10 candies of 3-5 different colors (have someone else fill the bag so that you don't know how many of each color are there).

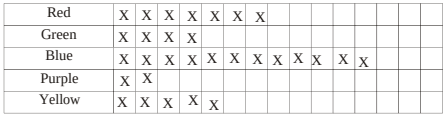
Procedure to gather statistics:

1) Shake the bag.
2) Pull out one candy.
3) Tally the candy color on the graph.
4) Replace the candy in the bag.

Repeat steps 1 through 4 10 times to find the number of occurrences of finding each color of candy.

After 10 samples, guess the number of different colors in your bag by completing Steps 5 - 7.

Sample Tally or Bar Graph of Number of Times Each Color of candy was selected after 30 samples:



Step 5: Totals after 30 Samples

Color	red	green	blue	purple	yellow
10 Samples	2	2	5		1
20 Samples	5	3	8	1	3
30 Samples	7	4	12	2	5

Step 6: Percentage of Times

Color	red	green	blue	purple	yellow
10 Samp.	20%	20%	50%		10%
20 Samp	25%	15%	40%	5%	15%
30 Samp	23.3%	13.3%	40%	6.7%	16.7%

Step 7: Best Estimate

Color	red	green	blue	purple	yellow
10 Samp.	2.0	2.0	5.0		1.0
20 Samp	2.5 (2 - 3)	1.5 (1-2)	4.0 4	.5 1	1.5 (1-2)
30 Samp	2.33 (2)	1.33 (1)	4.0 (4)	.67 (1)	1.67 (2)

The Solution...

There is no exact solution to a statistical population survey. As the number of samples increases, the relative populations (fractions of one color versus another color) become more accurately represented in the survey data.

Scientists frequently have no direct ways to make measurements. Surveys of populations are one example of this. Suppose you wanted to know the relative population of different kinds of squirrels (red, gray, black, etc.) in you neighborhood. One way might be to tabulate the number of times each type visited your local bird feeder. From this distribution, you might infer the relative numbers of squirrels.

To get the total number of each type of squirrel in your neighborhood, you might rely on other data to infer a total population, then use your distributions to find the number of each type. You might have different answers, based on different assumptions for your unknown quantity (the total population), and all the answers would be mathematically valid. Scientists contend with such messy data when direct measurements cannot be made.

Plots like the graph you have created are called histograms when the bins used for color are interpreted as aggregated data values. Look for more information about interpreting data from histograms in a book about statistics.

Lunchbox Math Bytes

easy to digest mathematics for your lunchbox

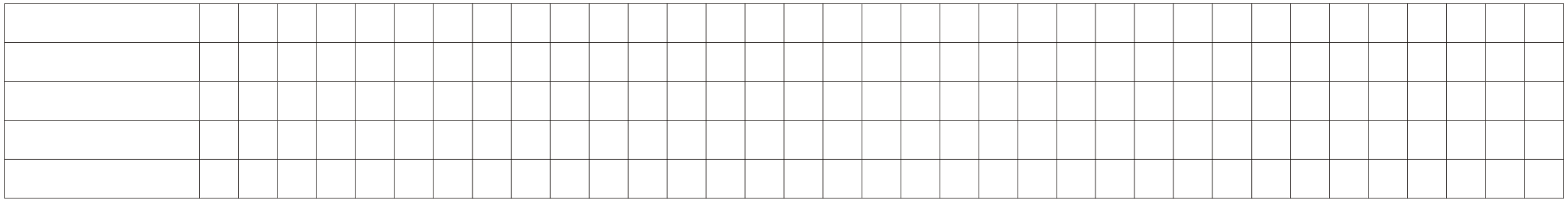
Messy Data

You will need to pack:

10 wrapped candies each of about 3-5 colors (ask someone to put them in a lunch bag, without you knowing how many of each color, but with knowing the different possible colors)

Before beginning, write the different color names in the spaces to the left of the bar graph and in shaded part of the table (write each color three times: one for each block portion of the table).
To gather the statistics: 1) Shake the bag with the 10 candies. 2) Without peeking, select one of the candies. 3) Mark a square with an "X" or other symbol on the bar chart to create a tally of the number of times you pulled that color candy out of the bag. 4) Replace the candy in the bag. It is important to always have all 10 candies in the bag when you sample.
Repeat steps 1-4 10 times, then stop and complete the first row, marked "10 samples" , in the table. 5) Write the number of times each color was pulled out of the bag in the first block. 6) Calculate the percentage of times each color was pulled. To calculate the percentage, divide the number of pulls by the total number of samples (which was 10), then multiply by 100. 7) To get the best guess at the number of a particular color in your bag, multiply the percentage of times the color was pulled by the total number of candies in the bag (which is 10) and divide that by 100.0 (because of the percentage). Round the value to the nearest integer. Since the number of candies is 10, you just need to divide the percentage by 10 (move the decimal to the left one place) then round the value off to an integer.
Gather more statistics by repeating steps 1-4 another 10 times. Then complete the table for the 20 samples row following steps 5-7. Did your guess of the number of each color candy in the bag change? Continue gathering statistics. Your guess should be better with more samples.

Color Count (Number of Times Each Color Candy Was Selected)



Total Number of Pulls	**Number of Times Pulled**					**Percentage of Times Pulled** 100 × (# Color Found) / (Total # Samples)					**Guess At Population** Guess = Percentage × (total in bag)/ 100) Guess = Percentage × 10 / 100 since 10 in bag Guess = Percentage/10 rounded to integer				
Color															
10 Samples															
20 Samples															
30 Samples															
40 Samples															
50 Samples															