Sampling Distributions of the Sample Mean—Pocket Pennies

You will need
- 25 pennies collected from recent day-to-day change

Some of the distributions of data that you have studied have had a roughly normal shape, but many others were not normal at all. What kind of distribution tends to emerge when you create sampling distributions of the mean from these non-normal populations? Here you'll explore that question.

COLLECT DATA

1. Enter the dates on your random sample of 25 pennies into a Fathom case table, using the attribute *Year*. Create a second attribute, *Age*, with a formula that calculates the difference between the current year and *Year*.

Pocket Pennies

	Year	Age
=		2005 - Year
1	1996	9
2	1985	20
3	1985	20
4	1993	12

Q1 If you were to make a histogram of the ages of all the pennies from all the students in your class, what do you think the shape of the distribution would look like? Sketch your prediction.

You can copy cases from one case table to another by choosing **Select All Cases, Copy,** and **Paste** from the **Edit** menu.

2. Combine everyone's data into one collection and make sure everyone has a copy of that Fathom document.
3. Using the complete collection, make a histogram of the ages of all the pennies in the class.

Q2 How does the actual distribution compare with your prediction in Q1?

Q3 Estimate the mean and standard deviation of the distribution. Confirm these estimates by computing the mean and standard deviation in Fathom. Either plot the values on the histogram or use a summary table.

INVESTIGATE

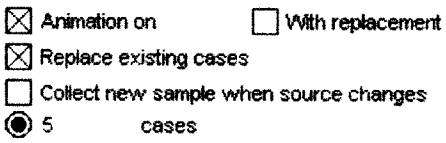
Building a Sampling Distribution

Next you'll take a random sample of size 5 from the ages of your class's pennies.

4. Select the collection, and choose **Collection | Sample Cases.** By default, Fathom takes a sample of ten cases with replacement and places them in a new collection named Sample of Pocket Pennies. You'll change this to five cases without replacement.

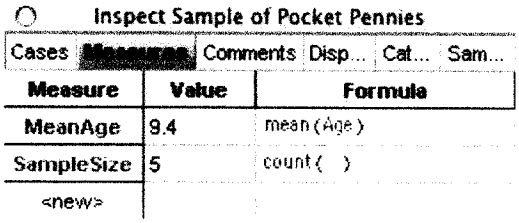
Notice that animation is on by default. You may want to change this later.

5. Double-click the Sample of Pocket Pennies collection to show its inspector. On the **Sample** panel, change the settings to match these. Click **Sample More Cases** to re-collect your sample.



The *SampleSize* measure will allow you to compare different sample sizes later on.

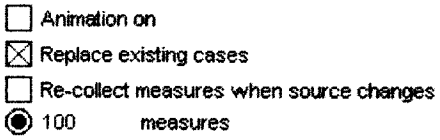
6. Go to the **Measures** panel of the inspector and define these measures.



Q4 If you were to make a histogram of the mean ages from several samples, do you think the mean of the values in this histogram would be larger than, smaller than, or the same as the mean of the population of the ages of all pennies? Regardless of your choice, try to make an argument to support each choice. Estimate what the standard deviation of the distribution of the mean ages will be.

You may want to turn off animation in the sample and measures collections.

7. Collect the mean ages from several samples by selecting the sample collection, then choosing **Collection | Collect Measures.** Show the inspector for the measures collection and change to these settings. Click **Collect More Measures.**



8. Make a histogram of *MeanAge.* Compute the mean and standard deviation of *MeanAge* by plotting values on the graph or using a summary table.

Q5 Which of the three choices in Q4 appears to be correct?

Changing the Sample Size

You'll now collect measures for samples of size 10 and size 25. You'll be able to create a split histogram to compare the effect of sample size.

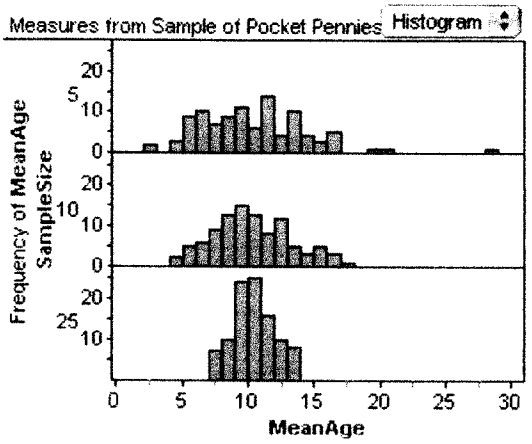
9. Show the inspector for the sample collection. On the **Sample** panel, change the sample size to 10.

10. Show the inspector for the measures collection. On the **Collect Measures** panel, uncheck Replace existing cases. This allows you to put the measures from all the samples into one collection. Then click **Collect More Measures.**

11. Repeat steps 9 and 10 to change the sample size to 25 and collect 100 more measures.

Holding down the Shift key tells Fathom to use the numerical values of the attribute as categories.

12. Drag the attribute *SampleSize* of the measures collection and drop it on the vertical axis of the histogram for *MeanAge* while holding down the Shift key. Your histogram should split three ways, showing distributions for each sample size (5, 10, and 25).



13. Compute the mean and standard deviation for each of the three sampling distributions, using a summary table. Again, hold down the Shift key when you drop *SampleSize* in the summary table.

Measures from Sample of Pocket Pennies

		MeanAge
SampleSize	5	
	10	
	25	

Column Summary

Q6 Look at the four histograms you constructed. As the sample size increases, what can you say about the shape of the histogram of sample means? About the center? About the spread?

Q7 Compare the values you got in step 13 for the mean and SD for the three sampling distributions with the values you got in Q3 for the whole population. Then figure out formulas for the mean and SD of a sampling distribution that relate them to the population mean and SD.

14. On your histogram, plot the values mean, mean *plus* 2SD, and mean *minus* 2SD.

MeanAge

mean( ) = 10.2621
mean( ) + 2 stdDev( ) = 16.3756
mean( ) − 2 stdDev( ) = 4.14756

Q8 What percentage of sample means are within 2 SD's of the population mean for each sample size?

Q9 For which sample sizes would it be reasonable to use the rule stating that 95% of all sample means lie within approximately 2 SD's of the population mean?

EXPLORE MORE

Open the Fathom document **LifeExp.ftm.** In this file you will find data on the life expectancy for females in Asia and Africa. Discuss the shapes of the original population. Take 200 samples of size 5 from each population. Do your conclusions from Q6–Q9 still hold up?