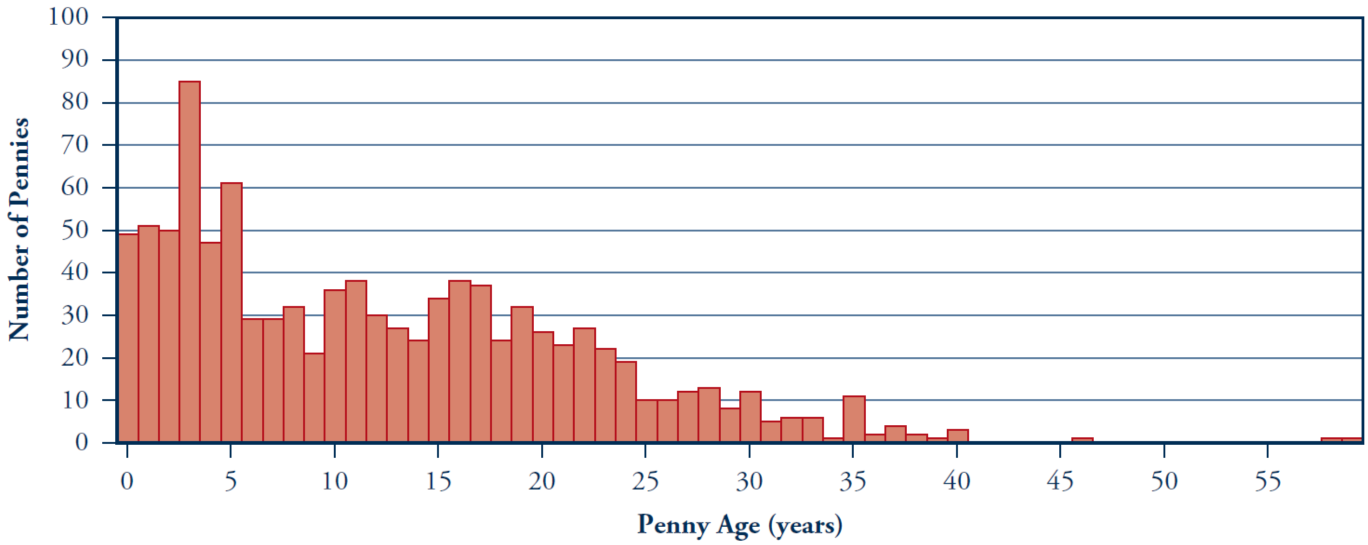
Introduction to Statistics	
Activity 7.3a	SOLUTIONS

Coin Ages

The following histogram displays the distribution of ages (in yrs) for a population of 1,000 pennies in circulation, collected in 1999. The following table has summary info for this distribution.

Number of Pennies	Mean	Standard Deviation	Min	Lower Quartile	Median	Upper Quartile	Max
1000	12.264	9.613	0	4	11	19	59



a) Identify the observational units and variable of interest here.
   Is this variable quantitative or qualitative?
   Observational units: **Pennies** Variable: **Age of penny**

   Quantitative or Qualitative? **Quantitative**

b) If you regard these 1,000 pennies as a population from which you can select samples, are the values in the previous table parameters or statistics? What symbols represent the mean and standard deviation?
   Parameters. μ, σ

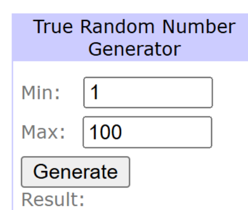
c) Does this population of coin ages roughly follow a normal distribution?
   If not, what shape does it have?
   No, skew right.

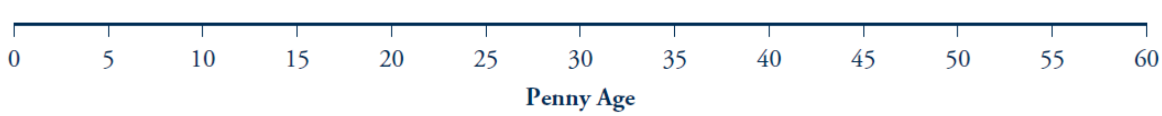
Rather than ask you to sample actual pennies from a container containing all 1,000 of these pennies, you will use a website to generate random samples of pennies from this population. To do this, we will assign a three-digit label to each of the 1000 pennies. The following table reports the number of pennies of each age and each penny's three-digit ID.

Age (in Yrs)	Count	ID #s	Age (in Yrs)	Count	ID #s	Age (in Yrs)	Count	ID #s
0	49	001–049	15	34	610–643	30	12	945–956
1	51	050–100	16	38	644–681	31	5	957–961
2	50	101–150	17	37	682–718	32	6	962–967
3	85	151–235	18	24	719–742	33	6	968–973
4	47	236–282	19	32	743–774	34	1	974
5	61	283–343	20	26	775–800	35	11	975–985
6	29	344–372	21	23	801–823	36	2	986–987
7	29	373–401	22	27	824–850	37	4	988–991
8	32	402–433	23	22	851–872	38	2	992–993
9	21	434–454	24	19	873–891	39	1	994
10	36	455–490	25	10	892–901	40	3	995–997
11	38	491–528	26	10	902–911	46	1	998
12	30	529–558	27	12	912–923	58	1	999
13	27	559–585	28	13	924–936	59	1	000
14	24	586–609	29	8	937–944	Total	1000	

Notice that each age has several ID labels assigned to it, equal to the number of pennies of that age in the population. Thus, for example, the age of 10 years has 36 ID labels because 36 of the 1000 pennies are 10 years old, whereas an age of 30 years has one-third as many ID labels because only 12 of the 1000 pennies are 30 years old.

d) Use the Random.org to generate a random sample of 5 penny ages from this population. Do this by changing "Max" to 1000 and "Generate"ing 5 random numbers between 1 and a 1000. Then, find the age assigned to each number in the table shown above. If you happen to get the same random number twice, ignore it and generate another number. Record the penny ages, and draw a dotplot of your sample distribution on the axis shown below:





e) Calculate the sample mean of your five penny ages.

f) Take four more random samples of five pennies each. Calculate the sample mean age for each sample, and record the results in the table:

	1	2	3	4	5
Sample Mean ($\bar{x}$)					

g) Did you get the same value for the sample mean all five times? What phenomenon that you studied earlier does this result again reveal?

No. Sample variability.

You are again encountering the notion of *sampling variability,* as it pertains to sample *means.* The sample mean varies from sample to sample not in a haphazard manner but according to a predictable long-term pattern known as a **sampling distribution.**

h) Use the applet on our class GoogleDoc called "One Variable with Sampling" to calculate the mean and standard deviation of your five sample means.
Mean ($\mu_{\bar{x}}$) of $\bar{x}$–values: SD ($\sigma_{\bar{x}}$) of $\bar{x}$–values:

i) Is this mean (of the $\bar{x}$–values) reasonably close to the population mean (μ = 12.264 years)? Is the standard deviation greater than, less than, or about equal to the population standard deviation (σ = 9.613 years)?
Yes. Less than.

The sample mean is an *unbiased* estimator of the population mean. In other words, the center of the distribution of sample means is the population mean. Also evident is that variability in the distribution of the sample mean is smaller than the variability in the original population (as long as the sample size is greater than one).

Now consider taking a random sample of 25 pennies from this same population. By taking five samples of 5 pennies each, you have essentially done this already. Consider all of your observations as a random sample of size 25. The sample mean of this sample of 25 pennies is exactly the mean of your five sample means recorded in part h.

j) Pool your sample mean from the sample of size 25 with those of your classmates. Produce a dotplot of these sample means. Label the axis clearly. What are the observational units in this plot?
[Plot]

Observational units: **The 25 penny samples**

k) Does this distribution appear to be centered at the population mean (μ = 12.264)?
Yes.

l) Do the values appear to be more or less spread out than either the population distribution or the distribution of your five sample means of size 5?
Less than both.

m) Does this distribution appear to be closer to a normal (bell/mound) shape than the distribution of ages in the original population? (Recall the histogram of the population distribution shown at the start of this activity.)
Yes, more normal shaped.