### Task 1

Consider the population of 269 words in the passage below. Let’s say we want to take a sample of words from the passage that will allow us to draw a conclusion about the **average length of all the words** in the passage. ***Please read carefully through that passage and use your judgment to select a sample of 10 words that you believe to be representative of the population. Circle those words in the passage so you can easily refer back to them.***

"Four score and seven years ago our fathers brought forth, on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we cannot dedicate—we cannot consecrate—we cannot hallow—this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us—that from these honored dead we take increased devotion to that cause for which they here gave the last full measure of devotion—that we here highly resolve that these dead shall not have died in vain—that this nation, under God, shall have a new birth of freedom—and that government of the people, by the people, for the people, shall not perish from the earth."

Write the average number of letters in the words you circled both here \_\_\_\_\_\_\_\_\_\_\_\_\_\_ and enter your data into the app using the section number given to you by your instructor. Choose “Judgment” as your sample type.

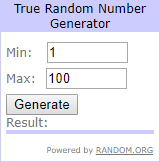
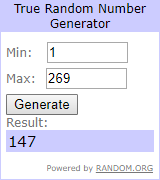
1. Explain what you did in order to choose your sample. How did you decide which ten words to choose?
2. Why do you think the ten words in your sample are representative of the 269 words in the population? Explain briefly.

When the sampling method produces characteristics of the sample that systematically differ from those characteristics of the population, we say that the sampling method is **biased**. In a moment, we’ll come back to the sample you took based on your own judgment to consider whether that method of sampling was biased. To try to eliminate potential biases, however, we like to take simple **random** samples. This should create representative samples, *no matter what variable we decide to focus on*. Humans are not very good “random samplers,” so it is important to use other techniques to do the sampling for us.

### Task 2

Let’s now use technology to take a **simple random sample** of 10 words from this population. Go to the website <http://random.org>

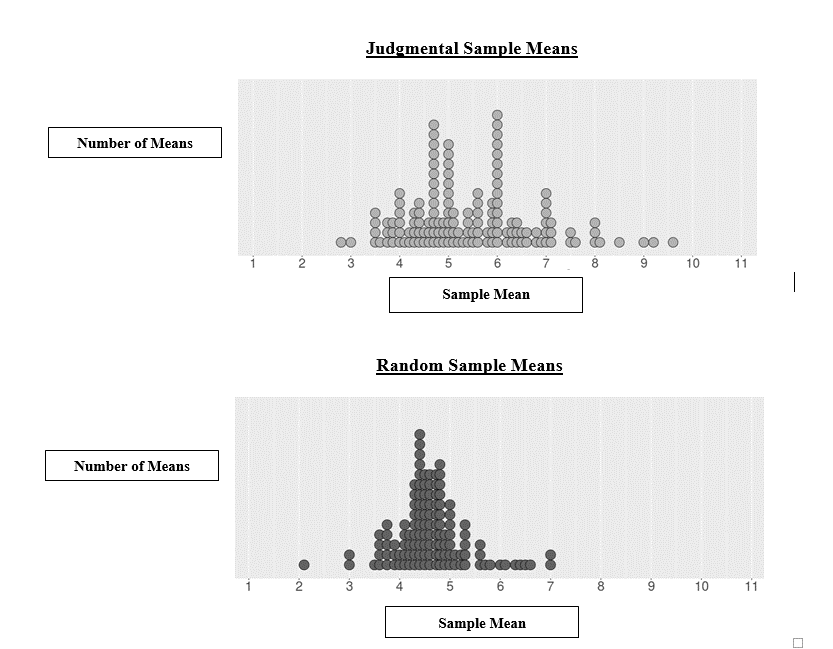
Near the top of the page you will see the "True Random Number Generator" widget. Set the values in the widget with a Min of 1 and a Max of 269. (There are 269 words in the Gettysburg Address.) Press the "Generate" button 10 times and record the result in the Number column of the table below. If you get the same number twice, just hit "Generate" an extra time so you have 10 unique numbers between 1 and 269.

Record the numbers you generate in the Number column in the table below, then use the table you’ve been given to find the words associated with the random numbers you generated. Record the word and count the number of letters in each word.

| Number | Word | # of letters |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Write your average both here \_\_\_\_\_\_\_\_\_\_\_\_\_\_ and enter your data into the app using the section number given to you by your instructor. Choose “Random” as your sample type.

We actually asked several students to share the means they got in Task 1 (where judgment was used to take a sample) and Task 2 (where the random.org applet was used to select a simple random sample). The plots of these means are shown to the right. The plots are called dotplots, and each dot (in each plot) represents one student’s sample mean. Within each plot, the x-axis (or horizontal axis) shows the value of the sample mean, and the y-axis (or vertical axis) shows how many means had particular values (e.g., in the Judgment Sample means plot, we can see that exactly three means were at 8). We can see that in the Judgment Sample means, the smallest mean was about 2.7 and the largest was about 9.7. In the Random Sample means, the smallest mean was about 2.2 and the largest mean was 7. Look carefully at these plots since you will need to answer the following questions based on what you are seeing.

1. It turns out, that the population average, or the mean length of all words in the Gettysburg Address, is 4.295 letters. Examine the “Judgment Sample Means” and “Random Sample Means” plots. The average of all the means in the Judgment Sample Means plot is about 5.70; the average of all the means in the Random Sample Means plot is very close to 4.295**. We would say that the judgment sampling method is biased.** What do you see in the Judgment Sample Means plot that indicates bias, when compared to the Random Sample Means plot? Please be as specific as you can.

**IMPORTANT:** Remember that we say a sampling method is biased if it tends to give results that, on average, too low or too high (or that underrepresent or over-represent a characteristic of the population). This can happen if the method of selecting the sample or the method of getting a response is flawed in some way. We can never know for sure if our sample is truly representative of the population, but we can know if we have used a method that will lead to bias. A sampling method is biased if it tends to lead to non-representative samples.

1. Why would you expect asking someone to use judgement to take a “representative” sample from the Gettysburg Address, in order to estimate average word length, to lead to bias?