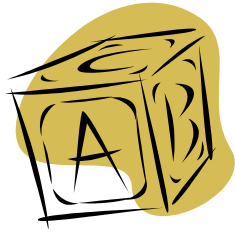# Likely

Letters

Likely Letters

_____
_____
_____

Statistics and Probability Project

Whether you are watching the game show Wheel of Fortune or playing Hangman with a friend, it helps to know which letters have the highest probability of use in words. During this project you will conduct an experiment and collect statistical data to discover which letters are the most commonly used in written English language.

Probability, or the likelihood that a specific event will occur, can be determined theoretically or experimentally. Theoretical probability is a ratio that compares the number of specific outcomes to the total number of outcomes possible. For example, to calculate the probability of rolling the number 2 on a number cube, divide 1 (the number of specific outcomes that are a 2) by 6 (the number of total possible outcomes) to determine the theoretical probability of 1/6 or about 16.7%. Theoretical probability can only be determined by mathematical calculation.

Experimental probability is based on performing an actual experiment to collect data. To determine the experimental probability simply divide the number of times that an actual event occurs by the number of times that the experiment is done. For instance, if you flip a coin and it lands on tails 27 out of 50 times, then the experimental probability of getting tails is 27/50 or 54%.

If we took the 26 letters of the alphabet and placed them in a hat and then asked you to choose one without looking, the theoretical probability of choosing each letter would be 1 out of 26 or about 3.8%. However, if we choose a letter at random out of a book or other written material, does each letter still have an equal probability of being chosen? Explain.

_____

_____

To find the probability of choosing any letter in the alphabet (at random) out of a book or other written source, would you use theoretical or experimental probability? Why?

_____

_____

To find the experimental probability of each letter in the alphabet being used, you will collect 100 letters randomly by selecting a small reading passage and then recording your data on the following page. Use 100 consecutive letters. Do not skip around on the page.

Before beginning, **predict** what you think the **top five most-used letters** will be.

1. ☐   2. ☐   3. ☐   4. ☐   5. ☐

Likely Letters

Using the 100 letters from your randomly selected reading passage, fill in the individual data in the chart below. Then enter the data for your group. Finally, enter the data for the entire class.

Letter	Individual Number In Sample	Individual % of __100__ Letters	Individual Letter Rank	Group Number In Sample	Group % of _____ Letters	Group Letter Rank	Class Number In Sample	Class % of _____ Letters	Class Letter Rank
A									
B									
C									
D									
E									
F									
G									
H									
I									
J									
K									
L									
M									

Likely Letters

Lesson Description: Likely Letters is a statistics and probability project that requires students to use experimental probability to determine the letters with the highest frequency of use in written English. The students differentiate between theoretical and experimental probability, make predictions, collect and organize data, and analyze their results to discover which letters of the alphabet are used the most.

Math Content: Statistics; Probability; Predicting; Collecting, Organizing, and Analyzing Data; Drawing Conclusions based on their data

Time Required: 1-2 Class Periods

Likely Letters includes:
* 2 Likely Letters assignment sheets
* 2 Likely Letters Data Entry Sheets
* 3 Likely Letters Teacher Tips pages
* 1 Likely Letters Cover Page

Materials Needed: Reading materials (books, magazines, etc.) for data collection

Suggested Grade Level: 5th - 8th

Teacher Testimonial:
Likely Letters is a lesson that makes statistics and data collection relevant to the students. Most students have watched Wheel of Fortune or played the game Hangman. Both games require a knowledge of letters that are likely to be used in the unknown words. This gives purpose to the data collection. You can also add suspense to the assignment by posting a covered list of the most frequently used letters that will only be revealed after the students complete their investigation.

Teacher Tips:
* On the Data Entry Sheet the students will write in the number of each letter found in their written English sample of 100 letters.
* Make sure that students understand how to fill in the total number of letters in their collected data where it says **"Percent of _____ Letters"** in the middle column of their data sheet. The number **100** has already been filled in under the "Individual" column since each student is recording the data for 100 letters. If there are three other students in their group then **400** would be written under the "Group" column. Finally, if the number of students in the class is 32, then **3200** would be written in under the "Class" column.
* Students will calculate the "Percent of _____ Letters" column by taking the number of each letter that they have recorded (from the first column under each heading) and dividing it by the total number of letters recorded. Have students convert the decimal to a percent and **round the percent probability on their data entry sheets to the nearest tenth** of a percent.