

ISC 1057 – Homework # 1

Fall 2016

Assigned: 9/13/16

Due: 9/22/16

In all exercises involving Google Ngram Viewer, search the English corpus of books. For all exercises involving Google Ngram Viewer submit (i) your answers and (ii) a copy of the results from Ngram Viewer (you can just take a screenshot).

1. In class we discussed the following Divide & Conquer algorithm for counting the number of people in a room.

Step 1 Each person in the room stands up and is assigned the number 1.

Step 2 Each standing person must pair up with another person standing in the room and add your respective numbers together.

Step 3 One of the two of you must sit down.

Step 4 If more than one person is standing, go to **Step 2** and repeat. If only one person is standing then their number is the total number of people in the room.

- (a) Complete the table below which compares the number of people remaining after Step 3 is completed for different size classrooms.

Loop #	# people standing after each loop			
	class: 17	class: 34	class: 68	class: 136
1				
2				
3				
4				
5				
6				
7				
8				

- (b) Using the trend you see from the table in (a) complete the table below and plot the data using PLOTLY.

Class Size	Number of loops required
17	
34	
68	
136	
272	
544	
1088	
2176	

-
2. Consider the sentence

It was the best of times, it was the worst of times.

List all the 1-grams, 2-grams, 3-grams, 4-grams in this sentence. Give the total number of each.

3. We can use Google Ngram Viewer to compare the rise and fall of businesses. Starting at 1900 compare the following businesses

Woolworth, Radio Shack, Best Buy, Amazon.com

- (a) For what period of time did Radio Shack surpass Woolworth?
 - (b) In what year was Woolworth the most popular?
 - (c) Does the first nonzero occurrence for Amazon.com seem reasonable? Why or why not?
 - (d) Make a new plot where the search is case insensitive. Which store's plot is most affected by this? Explain why you think this store was most affected.
-

4. Use Google Ngram Viewer with the "case-insensitive" box clicked to compare the following unigrams:

iPod, Walkman, iPad

starting at 1980 to as late as possible.

- (a) Why do you think that iPad hardly appears when it is so popular now?
 - (b) In what year is the number of references for iPod and Walkman the same?
 - (c) In 2004 the percentage of occurrences for iPod is approximately how many times greater than for Walkman? (Hint: click on the graph to get actual numbers.)
-
-