As Chapter 20 in your textbook points out, we begin to put science into Computer Science. Many tracks in STEM-based fields including Computer Science involve conducting research. The field of computer science involves research. Computer Science majors must utilize the scientific method. The scientific method is the logical scheme used by scientists searching for answers to the questions posed within science. The scientific method is used to produce scientific theories. Computer Science majors should be able to produce scientific theories and use empirical evidence to prove or disprove theories, a priori and a posteriori. A successful computer science major should be able to; pose a question in the context of existing knowledge, formulate a hypothesis as a tentative answer, deduce outcomes and make predictions, test the hypothesis in a specific experiment, produce research artifacts and documentation

In this assignment, you must utilize the scientific method. This assignment will represent your final project

You have studied chapters describing five sorting algorithms – bubble sort, selection sort, insertion sort, merge sort and quicksort. Chapter 20 provides information, illustrations, and example code on the merge sort and quicksort code. Chapter 19 provides information, illustrations, and example code on the bubble sort, selection sort, and insertions sort.

Create software testing tools for bubble sort, insertion sort, selection sort, merger sort, and quicksort. Your task for the Final Project is to conduct Bench Marking research to compare some of these algorithms you have studied in Computer Science I and II. You should design and implement a system to try to run the the sorting methods with random data sets of 10,000 20,000 100,000 200,000 1,000,000 2,000,000, 10,000,000 and 20,000,000 items.

You should run the program for each sorting algorithm 5 times for each data set size. Each time you run the program, you should record the time it takes to sort the data set and the size of the data. That is 5 readings for each data set size per sorting algorithm. If any sorting algorithm takes longer than 5 minutes to sort a particular data set size, you should stop the program for that trial and report this result in your report. If a sorting algorithm takes too long for anyone trial, such as for 100,000 items, then it is not necessary to complete the trials for larger data sets, simply report the size of the data set for which that algorithm started to take too long.

If you are creating separate projects for each type of sorting test, the program should have the following features

The program in the project should provide the capability to select or input the size of the array that will be populated with random integers.

It should also provide the capability to select the number of times that the program will rebuild the array and rerun the test.

If you are creating a single project that will perform all the tasks, it should have the following input features.

The ability to select or input the type of sorting algorithm to use

provide the capability to select or input the size of the array that will be populated with random integers

provide the capability to select the number of times that the program will rebuild the array and rerun the test.

Chapter 20 describes a way to create a timer, start the timer when the sort begins, stop the timer when the sort ends, and capture the value of the time that elapsed. Feel free to create your own way of recording the elapsed time. However, if you do not have the luxury of time, then the process described in chapter 20 will work fine.

You may separate the programs to test each sorting method into separate IntelliJ Projects. This approach would have one project per sorting algorithm. You can then run your varying data sets multiple times per project. The user should be able to select or enter the size of the data set to generate Run each data set size 5 times per project and record the results.

Alternately, you could incorporate each segment into a more comprehensive application. I would recommend making them individually first. If you have additional time, incorporate them into a single project. If you take this approach, the user should be able to select the type of sorting algorithm to use, select or enter the size of the data set, and select or enter the number of times to repeat the test.

You may also consider working on this project collaboratively. if so, you can separate the task of creating different programs with a group member and then share components. If you work on the programs collaboratively. Each student must conduct their own test on the data sets and produce their own data results, Workbook, report, and presentation.

Research Report - You should submit a Research Report with your results and conclusions about the experiment. Your report should describe how you conducted the experiment. Additionally, your report should include a table containing data from an aspect of the overall research framework and data that you collected. Your report should also contain a chart that visually explains your interpretation of a segment of your results you find interesting. Your report should also include a written description of what the chart represents. Your Research Report should summarize your research findings.

The Research Report is not the Lab Report. The Lab Report describes your software development activities. Your software project is just a tool that you will use to conduct your research. The Research Report should describe your thoughts and assumptions prior to the research. It should also contain the data that you accumulated. It should contain summaries of the data that you accumulated. It should contain at least one chart or graph illustrating something that you want to point out in your data. Finally, it should contain a conclusion on what you discovered from the data in your research.

Many times, what you thought prior to the research is not what you discovered after examining the results of the research. The research should have a format using the rules and guidelines that you have been following formatting your Lab Reports during this semester.

Questions like these should be in your mind as you think about the experiment and the report:

How does the time it takes for each algorithm to run compare to the size of the data set?

What are the best case, worst case, and average-case times for each method for each array length?

How does it fit with functions like T = f(n), T= f(n log(n) ) or T= f( n^2)?

What does our book tell us about the time efficiency of these algorithms?

Based on this, what did you expect the results to look like before running the time trials?

How did the results compare with this?

Which algorithm was the most efficient?

Why was this more efficient than the others?

You do not need to explicitly answer each of these questions, but they should guide you in considering what the experiment is all about. Your report should include:

a brief description of the topic

what is this experiment all about?

What do we hope to learn from it?

Why is this important?

a brief discussion of what you would expect to see in the experiment based on what you know about the sorting techniques before the experiment.

A description of your experiment –what did you do?

What equipment and software did you use?

Include the clock speed of your computer, and the specific compiler you used.

A presentation of your results –such as a table from the second spreadsheet with the summary results. If you can, create an Excel graph or set of graphs (one for each sorting technique) that illustrates your results.

A commentary or analysis of results.

What does the data show?

How does this compare to what you expected?

You should be able to cut and paste the table with your summary results from Excel into Word for inclusion in your report. You should submit your software project and your report. Remember, your report does not need to belong, but it should include the parts described in the list above.

Files and Items to Include in your Final Project Submission

Spreadsheet Workbook - Your Spreadsheet Workbook should contain the detailed data generated by your software testing tool. The spreadsheet should include aggregation functions or formulas that summarize your data. Examples of aggregation functions are count(), sum(), average(), min(), max(). The Workbook should also contain charts or graphs illustrating something that you found interesting in the data that you analyzed

PowerPoint Presentation - You must create and present a brief PowerPoint Presentation that outlines your research, findings, and conclusions. The PowerPoint Presentation should not be more than 7 slides. I have included a document describing the outline of a small research presentation.

Video Presentation - Since this is an online section of CSCI 112, you cannot do a face-to-face presentation using your PowerPoint slides as a guide. You are going to substitute this by using a multimedia desktop recording tool such as the service named Screencast. This tool will permit you to capture the activity on your computer screen while recording audio. I provided a link to register for a free subscription to Screencast. The link will also guide you to do an online tutorial on using Screencast. The service tool is very intuitive. You should limit your video presentation to a 5-minute presentation.

If you are familiar with an alternate software tool to create your Video Presentation, you are open to using the tool that you are familiar with. The Video Presentation should not be longer than 5 minutes.

Upload the following

Your Research Report as a Word Processing Document

Your Spreadsheet Workbook used to analyze your data and create the chart

The Code is used to Create Data for the Report in a Zipped Container File. This will be either several different programming projects or one comprehensive project in a zipped file container

A Lab Report describing your software development processes

PowerPoint Presentation

link or file to play a multimedia presentation.

In Conclusion

Details on the Artifacts to Submit for the Final Project

You should submit a report with your results and conclusions about the experiment. Your report should describe how you conducted the experiment. Additionally, your report should include a table containing data from an aspect of the overall research framework and data that you collected. Your report should also contain a chart that visually explains your interpretation of the results that the chart represents. Your report should summarize your findings

The research Report should describe your research and the development of the tools used to conduct your research. The Research Report is not the Lab Report. The Lab Report describes your software development activities. Your software project is just a tool that you will use to conduct your research. The Research Report should describe your thoughts and assumptions prior to the research. It should also contain the data that you accumulated. It should contain summaries of the data that you accumulated. It should contain at least one chart or graph illustrating something that you want to point out in your data. Finally, it should contain a conclusion on what you discovered from the data in your research. Many times what you thought prior to the research is not what you discovered after examining the results in the research. The research should be formatted using the rules and guidelines that you have been following formatting your Lab Reports during this semester.

You must create and present a brief PowerPoint Presentation that outlines your research, findings, and conclusions. The PowerPoint Presentation should not be more than 7 slides. I have included a document describing the outline of a small research presentation.

Zip File containing the Project or Projects used to accumulate your research data

A Lab Report describing your Software Development Activities. This includes software design specifications (pseudo\_code, flowcharts). The Lab Report is a separate document from the Research Report.

A Spreadsheet Workbook containing your empirical data. This should contain aggregation functions to summarize your results and support your conclusions. The Spreadsheet Workbook should also contain 1 or two charts and or graphs visually illustrating your findings. The charts and or graphs should be accompanied by a brief paragraph or two describing what the charts and or graphs represent.

Screencast or equivalent Video Presentation

Here is an outline for your final presentation

PowerPoint Assignment Final Project.docx

PowerPoint Assignment Final Project.pdf

Create a Video Presentation of Your PowerPoint and Report. Use the Screencast Video Recording Service to create and post your presentation

<https://www.screencast.com/> (Links to an external site.) (Links to an external site.)

You can also google it

The Final Research Report should contain the same minimum requirements as a Lab Report.

The Minimum Requirements are Listed Below

Document Formatting

A document header - The Header section should contain your name

A document footer - The footer should contain a page number

Paragraphs Formatting should be formatted

1.5 line spacing

12 points before and after the paragraph

Character Formatting

Body Text should be 12 points

Segment Headers should be boldface

Here is an Outline of the Framework of a Scientific Research Report

Basic Framework of a Research Report-1.pdf

All Java Programming Assignments must be submitted in a Zipped IntelliJ Project Folder Format.