We've talked about a lot of data structures this semester. So our final OLE of the semester is to perform, analyze and explain an Imperial Study to really demonstrate what the performance times we've talked about mean in more practical terms. Code to use is here --> DataType.java and NumberCruncher.java

DO NOT TEST/RUN THIS CODE ON SHELL!!! ONLY USE LAB COMPUTERS OR PERSONAL COMPUTERS

Phase 1 [10 points]: Points awarded based on completion of these 3 steps.

1. Compile and Run the pre-generated code 10 times.
2. Direct each run's output into a file.
3. **All 10 output files should be included in your OLE #4 submission.**

NOTE:: Running this code inside of an IDE (Dr. Java, eclipse, etc...) often results in the IDE running out of memory, so I HIGHLY encourage you to run the code from the command line.

An example of how to call/run the program and redirect it's output to a file is provided below. The "> *output.txt*" part is what does the output redirection, where **output.txt** is the name of the file your output will be redirected into. The other options tell the Java VM to reserve more memory than the JVM would reserve by default. This permits more RAM to be used by the JVM than would normally be allowed.

java -Xms64m -Xmx6g NumberCruncher > output.txt

Phase 2 [70 points]: Analyze your data from each run individually **AND** as averages for each data structure.

1. Does what you see make sense? Explain!
2. Are there outliers? (Are there any runs that are very odd) What do you think would have caused these outliers? (or...Why do you think your data didn't have any outliers?)
3. Which 3 data structures performed the best and which 3 data structures performed the worst for each of:
    - Adding
    - Removing
    - Searching
    - Overall (add time + search time + remove time)
4. Does the ranking you saw make sense? Why or Why not? (should be one section for each of the 4 categories)
5. Explain what you think you would see about the overal rankings if the number of searches was doubled. Test your theory and explain why the data supported or refuted your theory. **Include the output file your test run generated!**
6. Explain what you think you would see about the overal rankings if the number of searches was halved. Test your theory and explain why the data supported or refuted your theory. **Include the output file your test run generated!**
7. Include at least one other discussion about something you found interesting...**Each group member** should submit one of these discussion sections and the sections should note which group member wrote which section.

Phase 2 Point Breakdown:

- Parts 1 and 2 (10 Points)
- 5 points for each section discussion (Parts 3 and 4): add, remove, searching, overall (20 points total)
- Parts 5 and 6 (10 points):
  - theory (5 points)
  - Testing/data (5 points)
- Part 7 (10 points)
- Deep Thought points (10 points) -- These are given for thought provoking questions and/or comments related to your analysis of all 7 parts

---

Phase 3 [20 points]: Include in your report **AT LEAST 1 outside source per group member as this section is graded separately for each group member!!!!** Make sure to cite where in your report you used the source, and include enough citation information that I can find your source for myself. This outside source should be used to help explain some part of your data analysis. Points are awarded here for:

- Quality of Source (10 points)
- Discussion of of source and how you used it to help understand your data (10 points)

You may work in groups of up to size 6.