Problem 4-1 (Humble Numbers)  

A number $k > 1$ is called humble if the only prime factors of $k$ are 3 and 5. Consider the task of, on input $n$, outputting the $n$ smallest humble numbers and the following algorithm to do it:

```python
Humble(n):
    count = 0, prevOutput = 0
    Heap.Insert(3)
    Heap.Insert(5)
    while (count < n)
        cur = Heap.ExtractMin
        if cur ≠ prevOutput then
            output cur
            Heap.Insert(3*cur)
            Heap.Insert(5*cur)
            count = count + 1
            prevOutput = cur
```

(a) (4 points) Argue that the algorithm above (1) outputs numbers in increasing order, (2) does not output any number twice, (3) only outputs humble numbers, and (4) outputs all of the first $n$ humble numbers.

(b) (2 points) Derive an exact (i.e., no $O$-notation) bound on the number of times `Heap.Insert` is called.

(c) (2 points) Bound exactly the number of times `Heap.ExtractMin` is called.  
   (Hint: Use (b).)

(d) (2 points) Use the answers to (b) and (c) above to argue that `Humble` runs in $O(n \log n)$ time. Assume that arithmetic can be performed in $O(1)$ time.

Problem 4-2 (Running Median)  

In this task you will design a data structure supporting the following operations:

- **Build($A[1 \ldots n]$):** Initializes the data structure with the elements of the array $A$.
- **Insert($x$):** Inserts the element $x$ into the data structure.
• **MEDIAN**: Returns the median\(^1\) of the currently stored elements.

In the following, you are allowed to use the standard heap operations from CLRS Chapter 6 (incurring the corresponding running times).

(a) (4 points) Describe a data structure such that you can perform the operations BUILD, INSERT, and MEDIAN with running times as required below.

(Hint: Use a min-heap and a max-heap.)

Describe in pseudo-code implementations of

(b) (3 points) BUILD running in time \(O(n)\) assuming that you can query an \(O(n)\) magic box for finding the median of an \(n\)-element array,\(^2\)

(c) (3 points) INSERT running in time \(O(\log n)\), where \(n\) is the number of elements in the data structure.

(d) (3 points) MEDIAN running in time \(O(1)\).

**Problem 4-3 (Three-way partitioning) 8 points**

Recall that quicksort selects an element as pivot, partitions an array around the pivot, and recurses on the left and on the right of the pivot. Consider an array that contains many duplicates and observe that for such an array, quicksort recurses on all duplicates of the pivot element. In this task you are to develop a new partitioning procedure that works well on arrays with many duplicates. The idea is to partition the array into elements less than the pivot, equal to the pivot and greater than the pivot.

(a) (4 points) Develop this idea into a partitioning algorithm and provide pseudocode. Make sure your algorithm is in-place (i.e., do not use more than a constant amount of extra space).

(b) (2 points) Use your partitioning algorithm to come up with a sorting algorithm. Analyze the worst-case running time of your algorithm.

(c) (2 points) Find an array on which the original quicksort runs in time \(\Theta(n^2)\) but your algorithm from (b) in \(\Theta(n)\).

**Problem 4-4 (Inversions) 14 points**

Recall that in the worst case the running time of (non-randomized) QUICKSORT and INSERTIONSORT are \(\Theta(n^2)\).

(a) (2 points) Give an example of array of length \(n\) where (non-randomized) QUICKSORT and INSERTIONSORT both take time \(\Omega(n^2)\). Justify your answer.

---

\(^1\)The median of \(n\) elements is the \(\lceil \frac{n}{2} \rceil\)-largest element.

\(^2\)Such an algorithms will be discussed in class in a few weeks’ time.
(b) (1 points) Give an example of array of length $n$ where (non-randomized) QUICKSORT runs in time $\Omega(n^2)$ but INSERTIONSORT runs in time $O(n)$. Justify your answer.

Let $A[1 \ldots n]$ be a random permutation of $\{1, 2, \ldots, n\}$. Recall that the pair $(i, j)$ is called an inversion if $i < j$ and $A[i] > A[j]$.

(c) (2 point) What is the expected running time of INSERTIONSORT on $A$? (Hint: Recall that the running time of INSERTIONSORT is $\Theta(n + I)$ where $I$ is the number of inversions in $A$. Thus, it suffices to compute the expected number of inversions in $A$.)

(d) (6 points) Assume we have a partition procedure that is stable – an algorithm is stable if it maintains the relative order of elements. Thus, a partition procedure is stable if it preserves the relative order of elements less than the pivot and the relative order of elements greater than pivot after the partition.

(i) Prove that a stable partition on $A$ where $A$ is a random permutation, results in subarrays that are random permutations themselves. (Hint: Consider the two distributions: (a) Running the stable partition on a random permutation $A$ and taking the left half. (b) a random permutation over $B$ where $B$ consists of all elements in $A$ less than the pivot.)

(ii) Use the above observation to write a recurrence for the expected running time of QUICKSORT on a random permutation $A$, and solve it. (Hint: Recall the recurrence for Randomized QUICKSORT from the lecture.)

(iii) Which of QUICKSORT and INSERTIONSORT has better expected running time on a random permutation?

(e) (3 points) [Extra credit] Prove that the expected running time of QUICKSORT computed in part (d) does not change if we use the standard (non-stable) Partition procedure from the book.