Analysis of Open Addressing: We'll look at the complexity of `INSERT` since, in open addressing, searching for a key k that is in the table takes exactly as long as it took to insert k in the first place. The time to search for an element k that does not appear in the table is the time it would take to insert that element in the table. You should check why these two statements are true.

It's not hard to come up with worst-case situations where the above types of open addressing require $\Theta(n)$ time for `INSERT`. On average, however, it can be very difficult to analyze a particular type of probing. Therefore, we will consider the following situation: there is a hash table with m locations that contains n elements and we want to insert a new key k. We will consider a random probe sequence for k—that is, it's probe sequence is equally likely to be any permutation of $(0, 1, ..., m-1)$. This is a realistic situation since, ideally, each key's probe sequence is as unrelated as possible to the probe sequence of any other key.

Let T denote the number of probes performed in the `INSERT`. Let A_i denote the event that every location up until the i-th probe is occupied. Then, $T >= i$ iff $A_1, A_2, \ldots, A_{i-1}$ all occur, so

$$\begin{aligned} \Pr(T \geq i) &= \Pr(A_1 \cap A_2 \cap \cdots \cap A_{i-1}) \\ &= \Pr(A_1)\Pr(A_2|A_1)\Pr(A_3|A_1 \cap A_2)\cdots\Pr(A_{i-1}|A_1 \cap \cdots \cap A_{i-2}) \end{aligned}$$

For $j \geq 1$,

$$Pr(A_j|A_1 \cap \cdots \cap A_{j-1}) = (n-j+1)/(m-j+1),$$

because there are $n-j+1$ elements that we haven't seen among the remaining $m-j+1$ slots that we haven't seen. Hence,

$$Pr(T \geq i) = n/m \cdot (n-1)/(m-1) \cdots (n-i+2)/(m-i+2) \leq (n/m)^{i-1} = a^{i-1}. \qquad (7)$$

Now we can calculate the expected value of T, or the average-case complexity of insert:

$$
\begin{aligned}
E(T) &= \sum_{i=0}^{m-1} i \Pr(T = i) \\
&\leq \sum_{i=1}^{\infty} i \Pr(T = i) \\
&= \sum_{i=1}^{\infty} i(\Pr(T \geq i) - \Pr(T \geq i + 1)) \\
&= \sum_{i=1}^{\infty} \Pr(T \geq i) && \text{by telescoping} \\
&\leq \sum_{i=1}^{\infty} a^{i-1} && \text{by (7)} \\
&= \sum_{i=0}^{\infty} a^{i} \\
&= \frac{1}{1-a}
\end{aligned}
$$

Remember that $a < 1$ since $n < m$. The bigger the load factor, however, the longer it takes to insert something. This is what we expect, intuitively.