The following notes are on the recurrence that arose when analyzing randomized selection in class on Wed., Sept. 28.

The recurrence was

$$T(n) = \sum_{x=1}^{n} \Pr(\text{partition is x smallest}) \cdot (\text{Running time when partition is x smallest}) \; .$$

Using $x$ and $n-x$ as an upper bound of the sizes of the two sides:

$$\begin{aligned}
T(n) &\leq \sum_{x=1}^{n} \frac{1}{n} \left( (T(x) \text{ or } T(n-x)) + O(n) \right) && (1)\\
&\leq \sum_{x=1}^{n} \frac{1}{n} \left( T(\max\{x, n-x\}) + O(n) \right) && (2)\\
&\leq \left(\frac{1}{n}\right) \sum_{x=1}^{n} \left( T(\max\{x, n-x\}) \right) + O(n) && (3)
\end{aligned}$$

We now rewrite the max term. Notice that as $x$ goes from 1 to $n$, the term $\max\{x, n-x\}$ takes on the values $n-1, n-2, n-3, \ldots, n/2, n/2, n/2+1, n/2+2, \ldots, n-1, n$. As an overestimate, we say that it takes all the values between $n/2$ and $n$ twice. Thus we substitute and obtain

$$\begin{aligned}
T(n) &\leq \left( \frac{2}{n} \sum_{x=0}^{n/2} T(n/2+x) \right) + O(n) && (4)\\
&= \frac{2}{n} T(n) + \left( \frac{2}{n} \sum_{x=0}^{n/2-1} T(n/2+x) \right) + O(n) && (5)
\end{aligned}$$

We pulled out the $T(n)$ terms to emphasized them. We might be worried about having $T(n)$ on the right side of the equation, so we will bring it over the left-hand side and obtain

$$\left(1 - \frac{2}{n}\right) T(n) \leq \left( \frac{2}{n} \sum_{x=0}^{n/2-1} T(n/2+x) \right) + O(n) \; .$$

We now multiply both sides of the inequality by $n/(n-2)$ to obtain:

$$T(n) \leq \left( \frac{2}{n-2} \sum_{x=0}^{n/2-1} T(n/2+x) \right) + kn^2/(n-2) \; .$$

We have replaced the $O(n)$ by $kn$ for some constant $k$ before multiplying by $n/(n-2)$. We do this because we will need to for the proof by induction below.

We now have a recurrence in a nice form. $T(n)$ is on the left, and the right has terms of the form $T(x)$ for $x < n$. We can therefore "guess" that $T(n) = O(n)$ and try to prove it. More precisely, we will prove by induction that $T(n) \leq cn$ for some $c$. Since the recurrence is in the stated form, we can substitute in on the right hand side and obtain

$$
\begin{aligned}
T(n) &\leq \left(\frac{2}{n-2}\sum_{x=0}^{n/2-1} T(n/2+x)\right) + kn^2/(n-2) && (6)\\
&\leq \left(\frac{2}{n-2}\sum_{x=0}^{n/2-1} c(n/2+x)\right) + kn^2/(n-2) && (7)\\
&= \left(\frac{2c}{n-2}\right)((n/2)(n/2) + (n/2-1)(n/2)/2) + kn^2/(n-2) && (8)\\
&= \left(\frac{2c}{n-2}\right)\left(3n^2/8 - n/4\right) + kn^2/(n-2) && (9)\\
&= \left(\frac{c}{n-2}\right)\left(3n^2/4 - n/8\right) + kn^2/(n-2) && (10)\\
&= \frac{1}{n-2}\left((3c/4+k)n^2 - (c/8)n\right) && (11)\\
&= \frac{n}{n-2}\left((3c/4+k)n - (c/8)\right) && (12)
\end{aligned}
$$

Looking at this last term, we see that the leading $n/(n-2)$ is slightly larger than 1, so we can upper bound it by, say 7/6 for $n \geq 14$ (there are many possible choices of upper bounds.) Our goal, remember, is to show that the term multiplying the $n$ is at most $c$, and as we will see, this suffices.

So we get

$$T(n) \leq (7/6)\left((3c/4+k)n - (c/8)\right)\ .$$

If the right hand side is at most $cn$ we are done. Whether it is will depend on the relative values of $c$ and $k$. Let's write the constraint we want

$$(7/6)\left((3c/4+k)n - (c/8)\right) \leq cn$$

and solve for $c$ in terms of $k$. We get

$$(7c/8 + 7k/6 - c)n \leq 7c/48$$

or

$(7k/6 - c/8)n \leq 7c/48$.

Clearly, if $7k/6 - c/8 < 0$ this will hold. So we just choose $c$ sufficiently larger than $k$, e.g. $c = 28k/3$ and we are done.