1.

(a) Note a fact: if there is at least $\frac{n}{2}$ numbers same, the median of the sorted list should be that number that repeat $\frac{n}{2}$

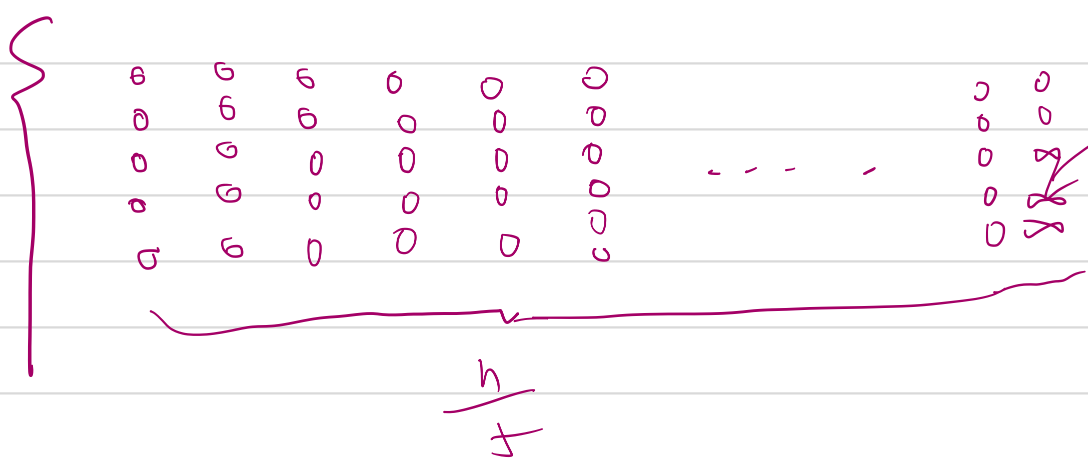
e.g. 1 1 1 1 1 2 3 8 7, 1 1 5 5 5 5 8 10,

1 2 3 4 $\boxed{9}$ 9 9 9

As a result, we aim for the middle one (or the middle two if the total number is even)

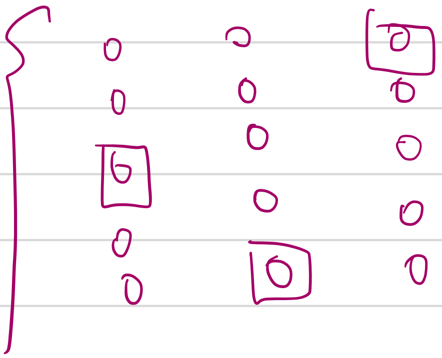
To find the $(\frac{n}{2})^{th}$ elem,

① Form $\frac{n}{5}$ groups of 5 elems ↗ $\boxed{\Theta(n)}$



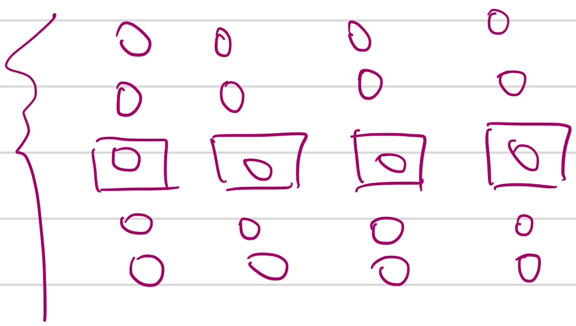
if not divided by 5, add 1 to 4 ∞

② Find median in each group

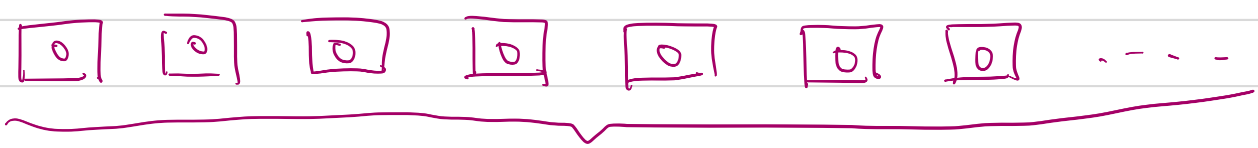


$$\frac{n}{5} \cdot \Theta(1) = \Theta(n)$$

$\Downarrow$ reorganize (affordable, constant time for each group)



③ take all the means together, which are:



$\frac{n}{5}$

unsorted !

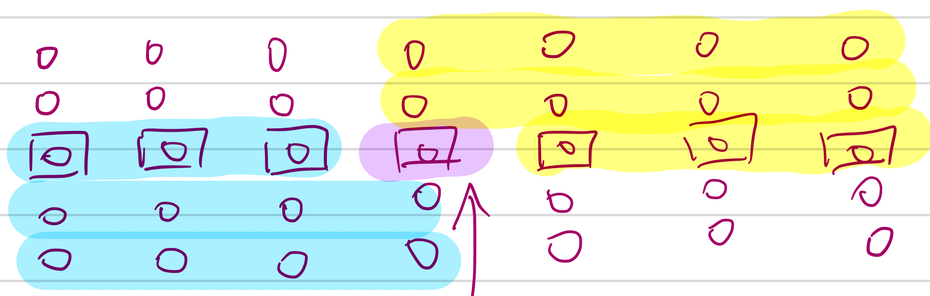
then do the same thing (repeat step ① ②, divide 5 and find median)

cost: $T(\frac{n}{5}) + 2\Theta(n) \Rightarrow T(\frac{n}{5}) + \Theta(n)$

finally

: the elem smaller than ultra median

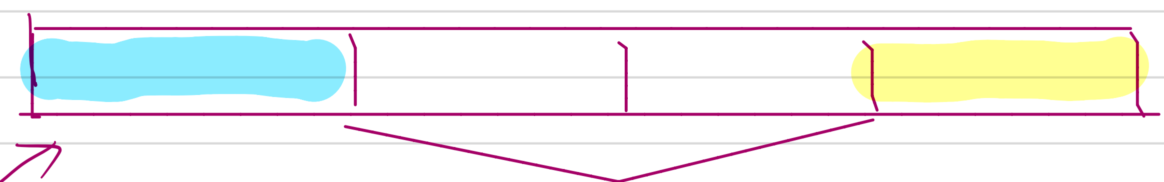
: the elem larger than ultra median



: the ultramedian is found

$\Rightarrow \quad \frac{1}{4}n \le \text{ultra median} \le \frac{3}{4}n$

definitely larger



definitely smaller

ultra-median range

Loop through the entire n, find the rank of x, if it is larger than $\frac{n}{2}$th, at least the yellow part (which even larger), shouldn't go into the recursive process again

if it is smaller, just recursivly do all the step before except blue part. if equal to $\frac{n}{2}$th, bingo! Got it!

After find the $\frac{n}{2}$th (or $\frac{n+1}{2}$th and $\frac{n-1}{2}$th if n is odd), linearly go through the list, set up the counter, see how many times that item appears. If it is not occuring $\frac{n}{2}$ time, the list then doesn't cousist a number that at least repeat $\frac{n}{2}$ times

The total time is:

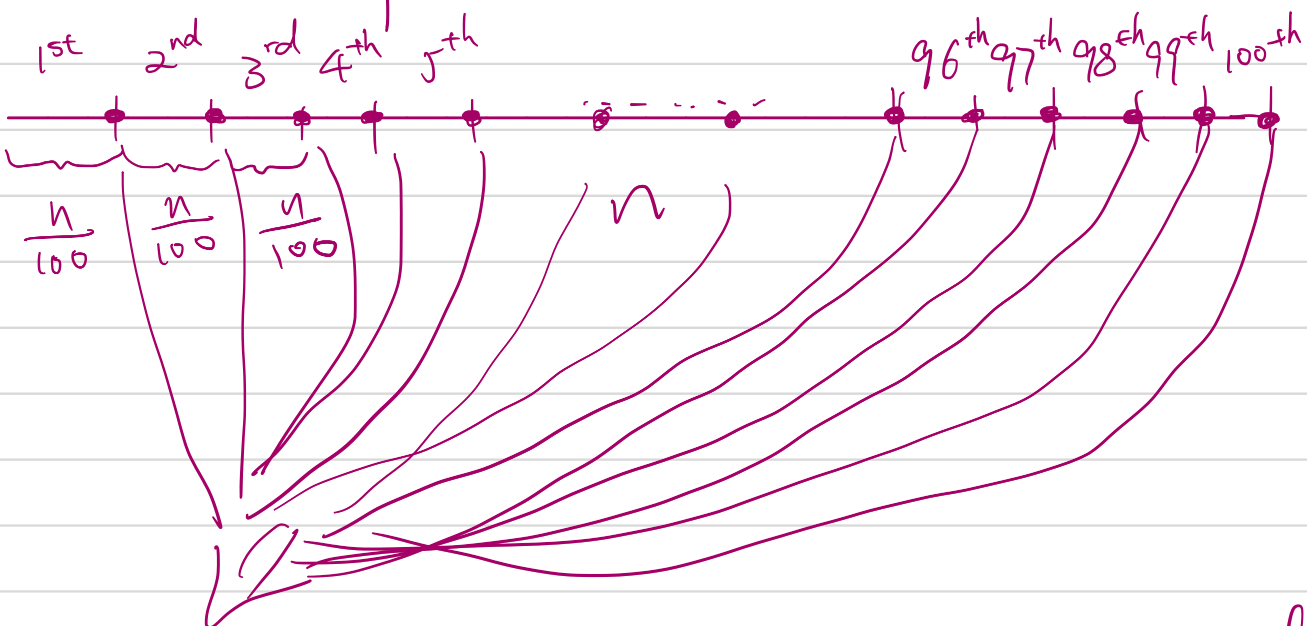
$$T(n) \le \underbrace{T\left(\frac{n}{5}\right)}_{\text{find ultramedian}} + \underbrace{T\left(\frac{3n}{4}\right)}_{\text{recursive if rank}(x)\text{ is not target}} + \underbrace{\theta(n)}_{}$$

split in to groups of 5 and
find medians
check if median appears $\frac{n}{2}$ times

$T(n) = O(n)$, linear time

(b)

Same as the last question. if a number appears $\frac{n}{100}$ time, if we devide n to 100 parts:



these numbers, at least one should appear at least $\frac{n}{100}$ times, if not, then the list definitely don't have a number that appears $\frac{n}{100}$ times.

$\Rightarrow$

As a result, use the approach before, find $(\frac{n}{100})^{th}$, $(\frac{2n}{100})^{th}$, $(\frac{3n}{100})^{th}$ ...... number, then loop through the n linearly to see if any of them appears $\frac{n}{100}$ time.

Because the time complexity ~~to~~ find one rank in an unsorted list is $O(n)$ in part a, the worst case here is we test out $(\frac{n}{100})^{th}$ to $(\frac{99n}{100})^{th}$ still not finding out any of them repeat $\frac{n}{100}$ times, so we are going to find 100 ranks in total, make the case $T(n) = 100 \cdot O(n)$,

which by definition, $T(n) = O(n)$

2.

(a) let X = # bee visit each flower

Define r.v. $X_k = \begin{cases} 1, & \text{if a bee lands on flower} \\ 0, & \text{otherwise} \end{cases}$

$$E(X_k) = \frac{1}{n}$$

$$E(X) = E\left(\sum_{k=1}^{k} X_k\right) = E(X_1) + E(X_2) + E(X_3) + \cdots E(X_k) = k \cdot \frac{1}{n} = \frac{k}{n}$$

(b)

let X = # flowers that are being visited

Define r.v. $X_k = \begin{cases} 1, & \text{if a flower is visited by at least one bee} \\ 0, & \text{otherwise} \end{cases}$

for one bee visit one flower, probability is $\frac{1}{n}$, one bee NOT visit one flower, $\left(1 - \frac{1}{n}\right)$, which is $\frac{n-1}{n}$

for k bees, the probability is independent from each one, which is $\left(\frac{n-1}{n}\right)^k$

$$E(X_k) = 1 - \left(\frac{n-1}{n}\right)^k$$

$$E(X) = E\left(\sum_{n=1}^{n} X_n\right) = E(X_1) + E(X_2) + E(X_3) + \cdots + E(X_n)$$

$$= n \cdot \left(1 - \left(\frac{n-1}{n}\right)^k\right) = n - n \cdot \left(\frac{n-1}{n}\right)^k$$

(c) Yes. so for one flower, $n=1$, $E(X)=1$, 1 flower will be expected to be visited; one bee, $k=1$, $E(X) = n - \cancel{n} \cdot \frac{n-1}{\cancel{n}}$, $E(X)=1$, 1 flower will be expected to be visited; if $k=400$, $n=100$, the answer will be

$$100 - 100 \cdot \left(\frac{99}{100}\right)^{400}$$

$$= 100 - 1.795$$

$$= 98.205$$ flowers are expected to be visited