Query problems

Given \( n \) points on the plane:

\[ P_i : (x_i, y_i), \quad 1 \leq i \leq n \]

Build a data structure such that given a query rectangle \( R \), we want to:

(i) answer the \# points in \( R \): \( |P \cap R| \)

(ii) Report all the points contained in \( R \): \( P \cap R \)

Without any explicit data structure, query time: \( O(n) \) by testing each point against \( R \)
For query (i), the query complexity should be much smaller than n, preferably log n.

For query (ii), we should aim for output-sensitive performance, i.e., proportional to |P∩R|.

One-dimensional analogue

\[
\begin{array}{cccccccc}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
& & & & & & & \\
\end{array}
\]

Sort the points: Preprocessing time \(O(n \log n)\)

given an interval \([l, r]\), we can do a binary search on \(l\) and \(r\) to locate their position in the sorted set of points, say \(i\) and \(j\), \(j > i\).

\#Points in \(j - i\) \query Time : \(2 \log n\)

Cost of reporting : \(O(\log n + |P∩R|)\)
Can we extend it to 2 dim?

\[
\begin{align*}
K \cdot d \quad &\text{here} \\
L_1 &\quad \eta A \\
\eta_1 &\quad \eta_2 \\
\eta_3 &\quad \eta_4
\end{align*}
\]
Query time $\sim \# \text{ nodes visited}$

corresponds to some rectangle region