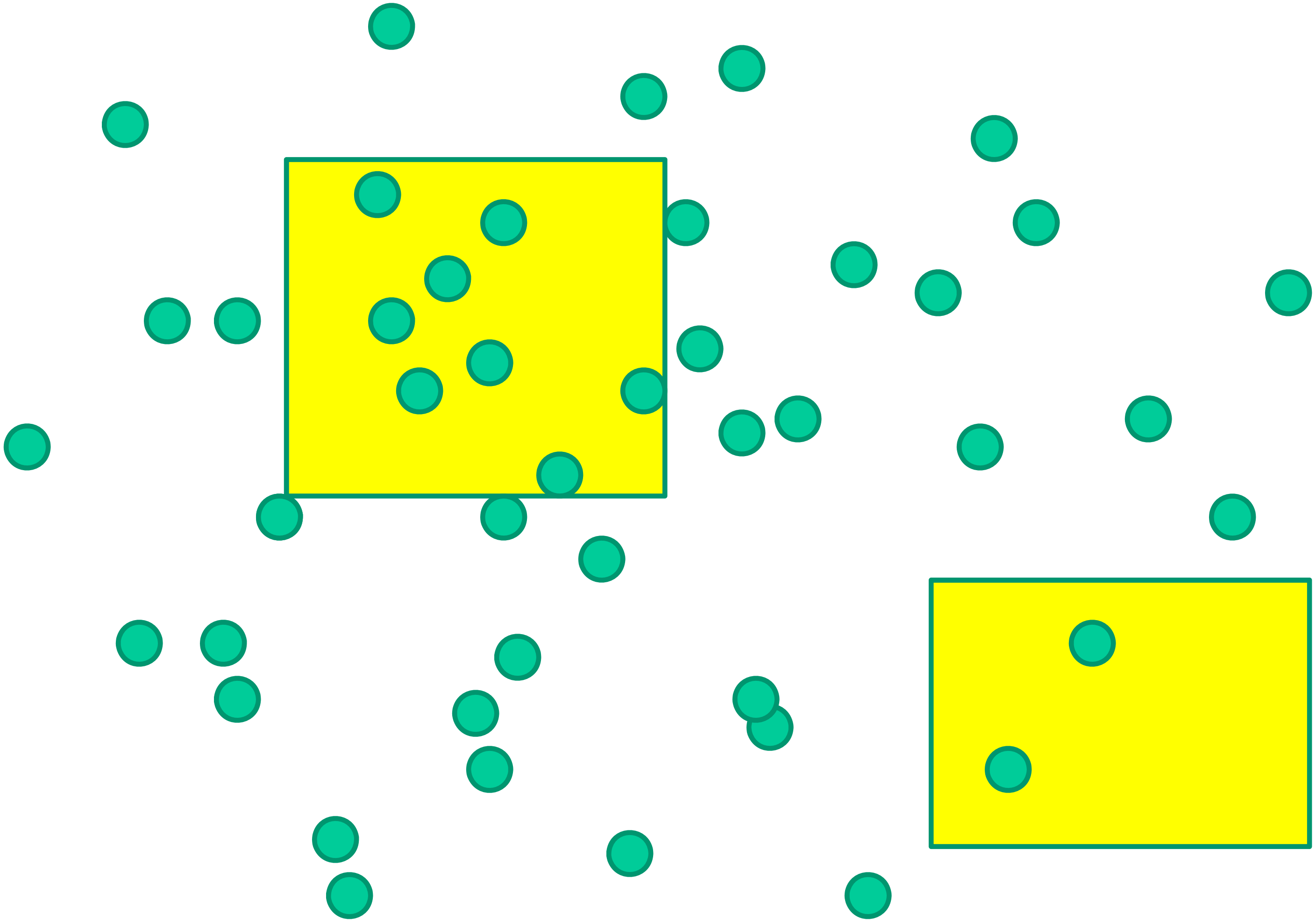
Range searching

- Given a faimly F of regions and a set S of n points in the d-dimensional space, construct a data structure D(S), such that we can do the following query efficiently.
- Reporting range query: Given any region R in F, report the set of points of S in R.
- Counting range query: answer the number of points of S in R

Rectangular range searching(d=2)

- Given a set S of n points in the plane, construct a data structure D(S), such that we can do the following query efficiently.
- Rectangle range query (counting): Given an axis parallel rectangle R, report the set of points of S in R.
- Answer the number of points of S in R efficiently

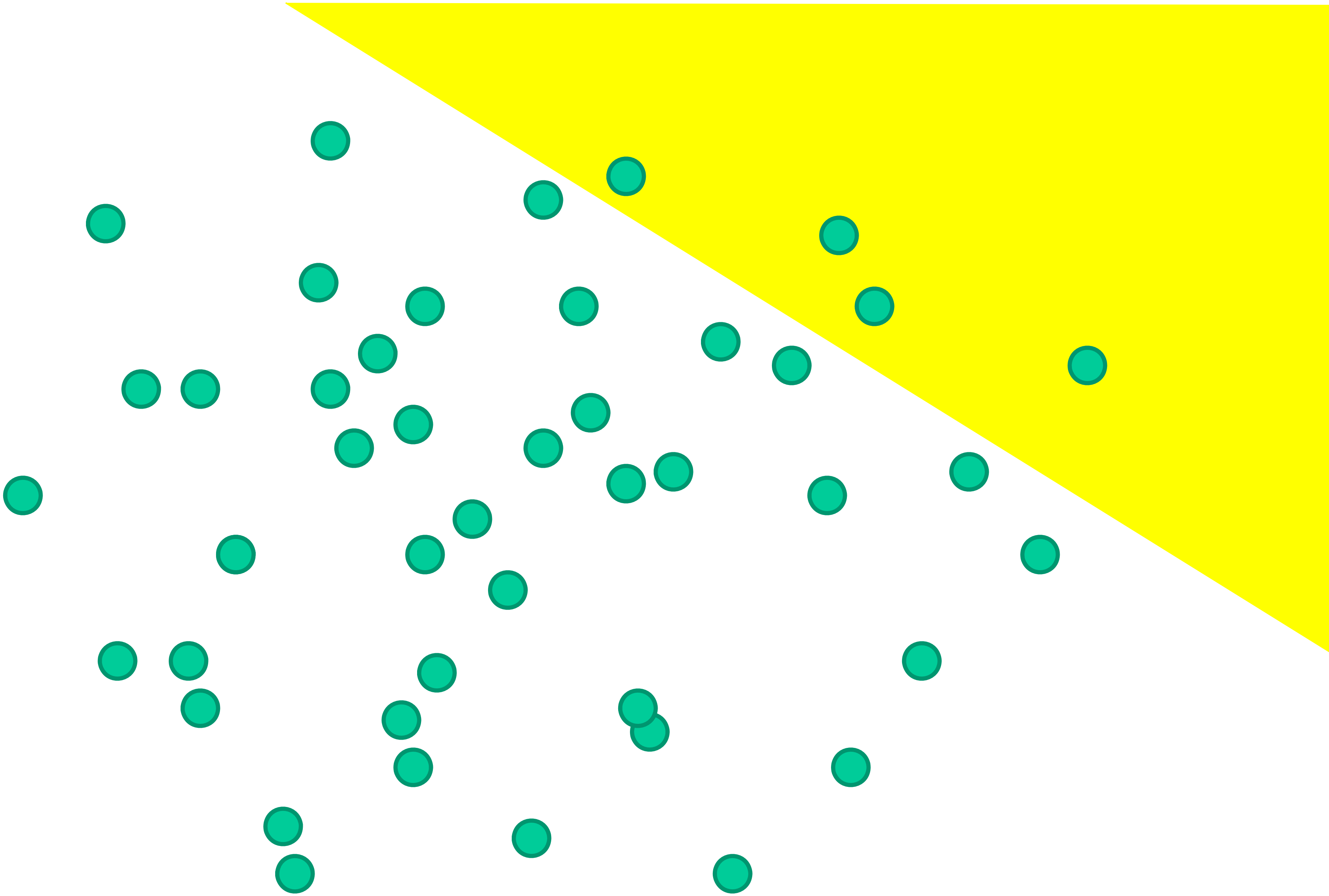


Example of application

- Given a database of customers with (income, sales), report the set of customers such that
  - 2M < income < 2.5M
  - 100K< sales< 300K

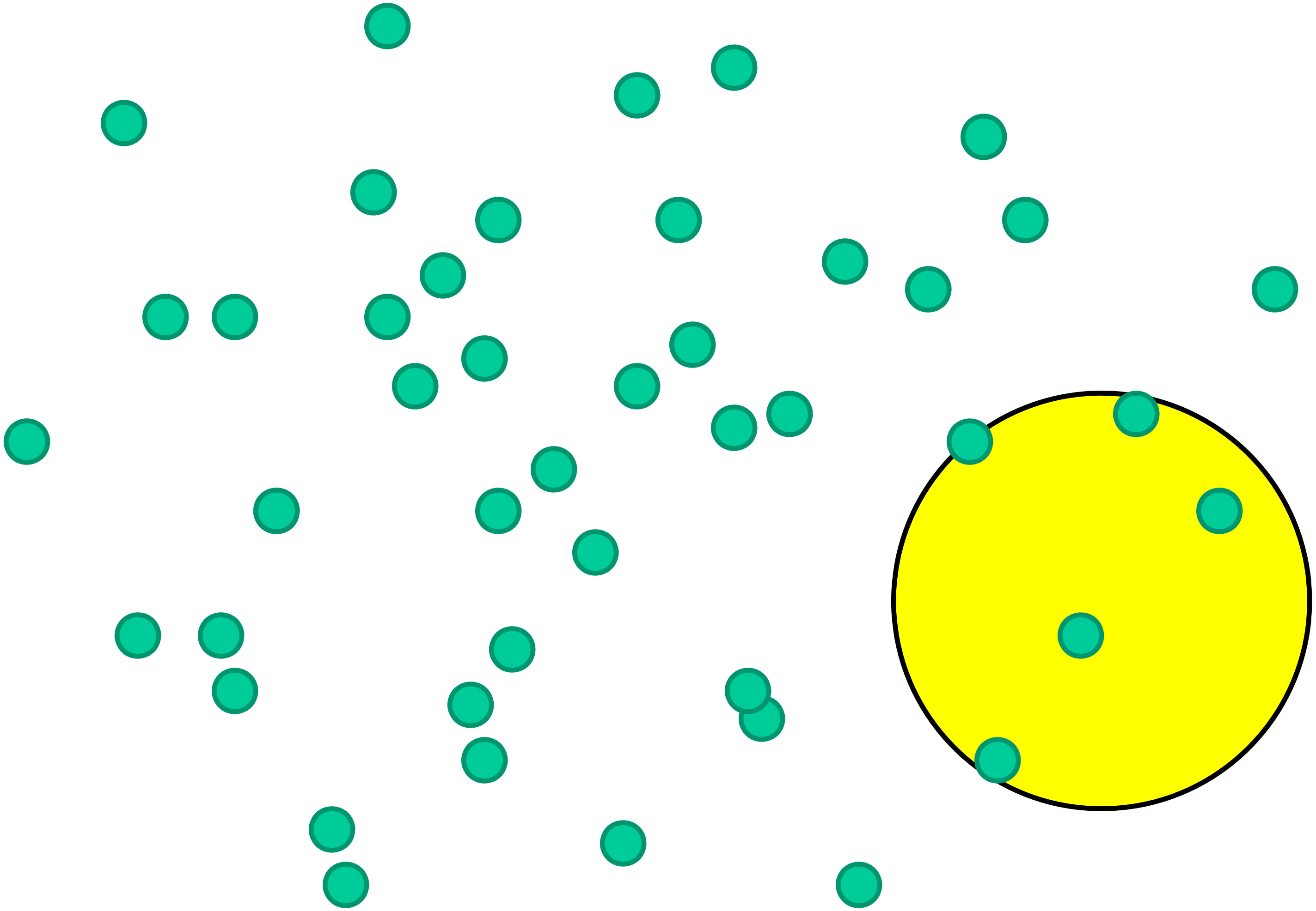
Halfplane range searching(d=2)

- Given a set S of n points in the plane, construct a data structure D(S), such that we can do the following query efficiently.
- Halfplane range query: Given a halfplane H, report the set of points of S in H efficiently
- Answer the number of points of S in H.



Circle range searching(d=2)

- Given a set S of n points in the plane, construct a data structure D(S), such that we can do the following query efficiently.
- Circle range query: Given a circle C, report the set of points of S inside C.
- Answer the number of points of S in C.

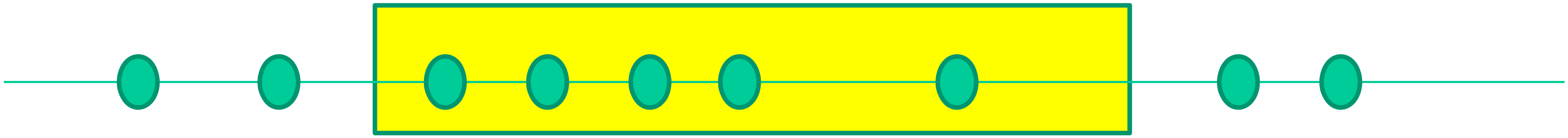


Example

- Answer the set of Italian restaurants within distance of 300 M from Sendai station.

Interval Range searching (d= 1)

- Given a set S of n data with real key values, construct a data structure D(S), such that we can do the following query efficiently.
- Interval range query (reporting): Given an interval I, report the set of data of S such that the key values are in I.
- Interval range query (counting): Answer the set of data of S whose key values are in I.



Theorem 1

There is an O(n) size data structure D(S) to answer the reporting interval range query in O(k+log n) time, where k is the number of reported elements. Also, the counting interval query can be compted in O(log n) time using O(n) size data structure.

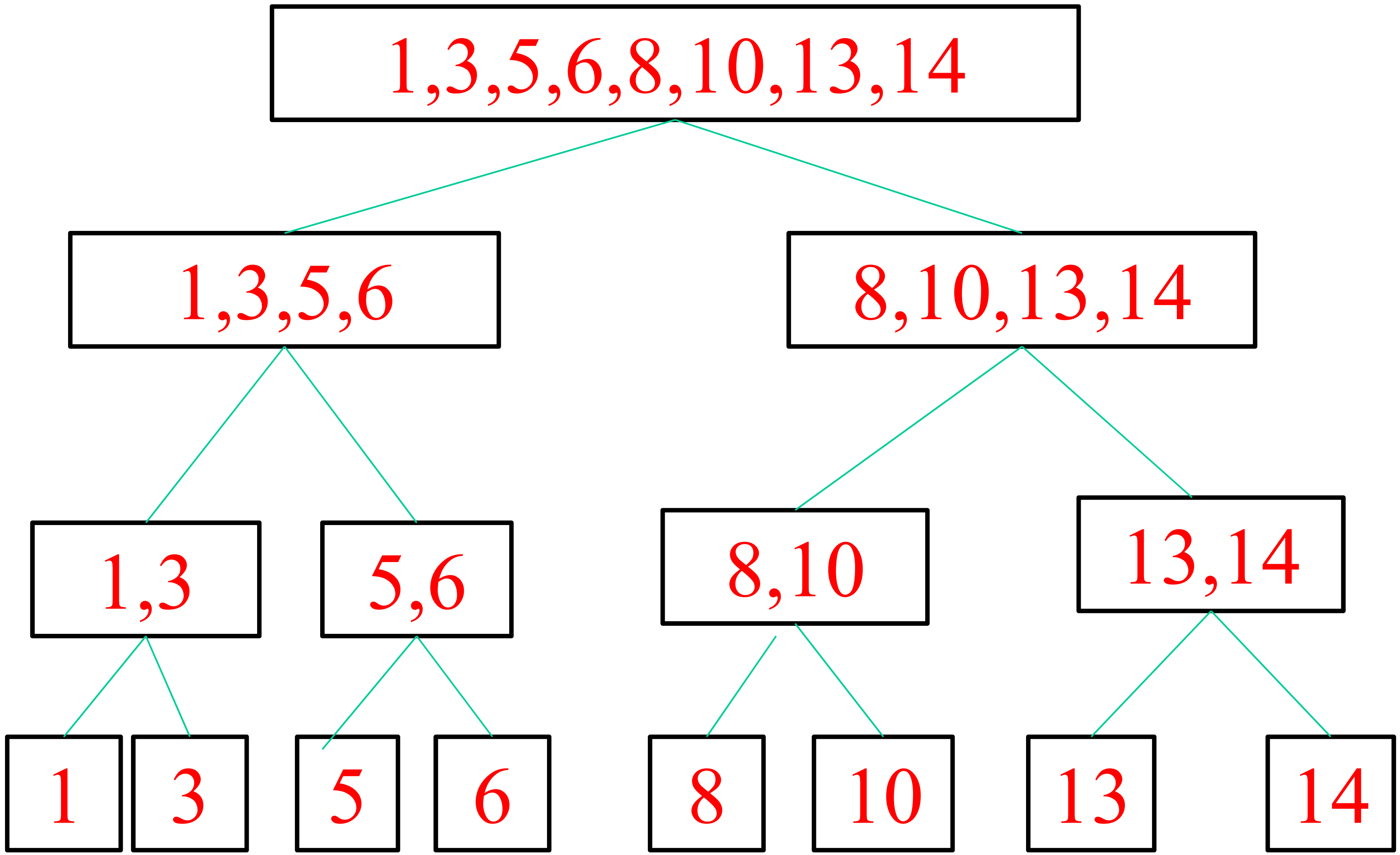
Interval Range searching (d= 1)

- Given a set S of n data with real key values key(x) and real data value data(x) for each x of S, construct a data structure D(S), such that we can do the following query efficiently.
- Range minimum query : Given an interval I, report x of S such that the key(x) values are in I and minimizing data(x).
  - Very important in data compression and data mining.

Interval tree

- Store n data into a binary tree T(S)
- The root of the tree contains S
- The left child is $T(S_1)$, where S_1 is the set of the n/2 data with small key values
- The right child is $T(S_2)$, where $S_2 = S - S_1$

A set stored at a vertex of the interval tree is called a primary set.


1,3,5,6,8,10,13,14
1,3,5,6
8,10,13,14
1,3
5,6
8,10
13,14
1
3
5
6
8
10
13
14

Interval query using T(S)

- Lemma 1.

For any interval I, $S \cap I$ is represented as a union of O(log n) primary sets.

- Lemma 2.

The primary sets considered in Theorem 1 can be computed in O(log n) time.

- Theorem 2. The range minimum query can be computed in O(log n) time

Rectangular range searching(d=2)

- Given a set S of n points in the plane, construct a data structure D(S), such that we can do the following query efficiently.
- Rectangle range query (counting): Given an axis parallel rectangle R, report the set of points of S in R.
- Answer the number of points of S in R efficiently

Range searching

- Theorem 3

There is a data structure of size O(n log n) to answer the reporting range query in O(k + log n) time , where k is the number of reported elements.

- Theorem 4.

There is a data structure of size O(n log n) to answer the counting range query in $O(\log^2 n)$ time.

- Theorem 5. The counting range query can be done in O(log n) time using an O(n log n) size data structure.