

# 자바와 함께하는 자료구조의 이해

Chapter 08 정렬

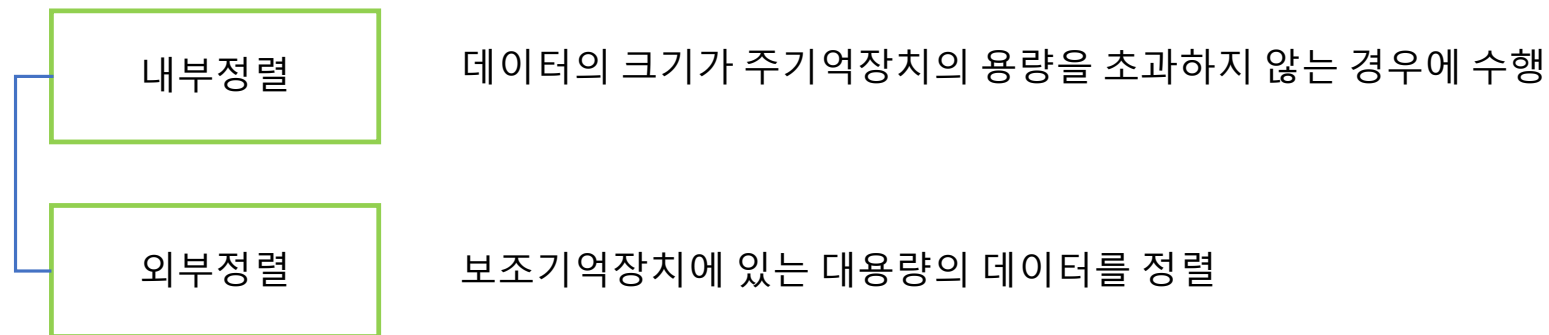
# 목차

1. Comparable vs Comparator
2. 선택정렬
3. 삽입정렬
4. 쉘정렬
5. 합병정렬
6. 퀵정렬
7. 버블정렬

# 정렬

## 정렬이란?

- **일반적:** 주어진 데이터를 오름차순으로 나열하는 것
- Cf) 오름차순: 1 -> 2 -> 3 -> 4 -> 5 ...
- **쉽고 간단 장점 정렬:** 선택정렬, 삽입정렬
- **빠른 수행시간 장점 정렬:** 쉘정렬, 힙정렬, 합병정렬, 퀵정렬



# 선택정렬

## 선택정렬이란?

- 배열에서 정렬되지 않은 부분의 원소들 중에서 최솟값을 선택하여 정렬된 부분의 바로 오른쪽 원소와 교환하는 것



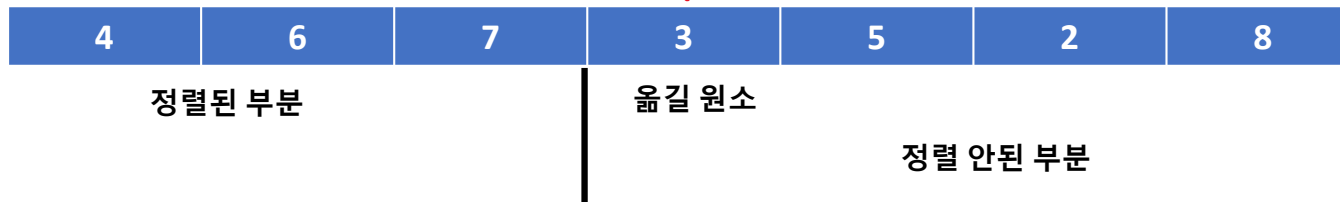
- 단, 전제로 정렬 안된 부분의 최소값은 정렬된 부분의 최대값보다 커야 한다.
- 선택정렬과 관련된 모든 시간 복잡도는  $O(N^2)$ 이다.
- 효율성 측면에서 뒤떨어져 잘 사용되지 않는다.

# 삽입정렬

## 삽입정렬이란?

- 정렬된 부분과 정렬되지 않은 부분으로 나뉜다
- 정렬되지 않은 부분의 가장 왼쪽 원소를 정렬된 부분에 '삽입'하는 방식의 알고리즘

← 방향으로 순서대로 비교하며 적정 위치 배치



- 입력에 민감한 알고리즘이다(역순으로 정렬되어 있다면 최악/매우 느리다)
- 시간 복잡도는  $O(N^2)$ 이다.
- 정렬된 파일에 뒤에 소량 데이터만 추가 정렬하는 경우나 입력 데이터가 적은 경우 효율적이다.

# 셸정렬

## 셸정렬이란?

- 삽입정렬에 전처리과정을 추가한 것이다.
- **전처리 과정:** 작은 값을 가진 원소들을 배열의 앞부분으로 옮기며, 큰 값을 가진 원소들이 배열의 뒷부분에 자리잡도록 만드는 과정
- 1.  $h$ -정렬: 간격이  $h$ 인 원소끼리 정렬한다.
- 2.  $h$ 를 줄여나가며 1번과 동일한 작업을 진행한다
- 3.  $h=1$ 이 될 경우 삽입정렬을 진행한다.

# 셸정렬

## 셸정렬이란?

10	8	6	20	4	3	22	1	0	15	16
----	---	---	----	---	---	----	---	---	----	----

간격을 설정하는 방법에 따라 수행시간이 달라진다.

일반적으로 시간 복잡도는  $O(N^2)$ 과  $O(N\log N)$  사이에 있다고 볼 수 있다.

K(gap):  $10/2=5$

10					3					16
	8					22				
		6					1			
			20					0		
				4					15	

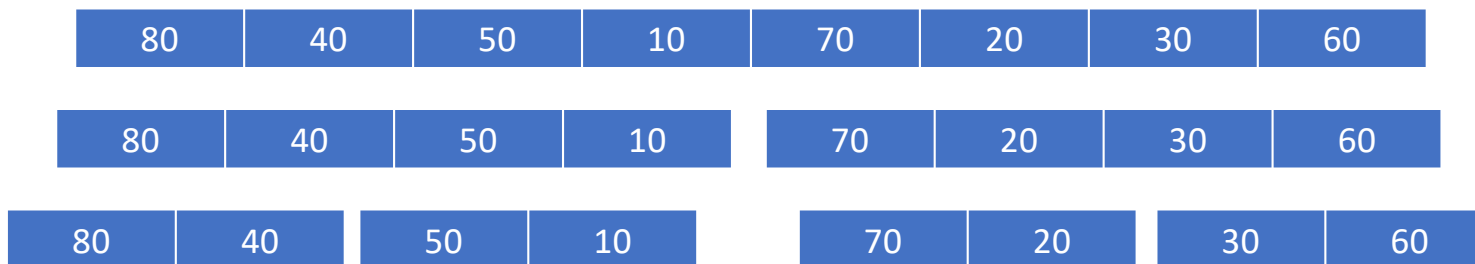
- 출처: 1인개발TV 유튜브

# 합병정렬

## 합병정렬이란?

- 입력된 데이터를 2개 단위까지 쪼개어 나눈 후 정렬하며 합병하는 방식

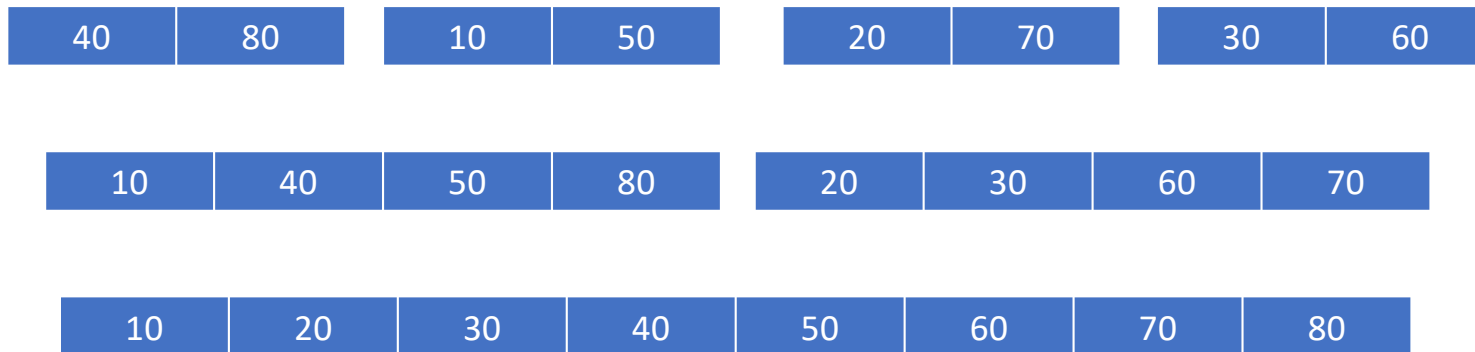
1. 데이터를 쪼갬다.





# 합병정렬

## 2. 정렬하며 붙여 나간다

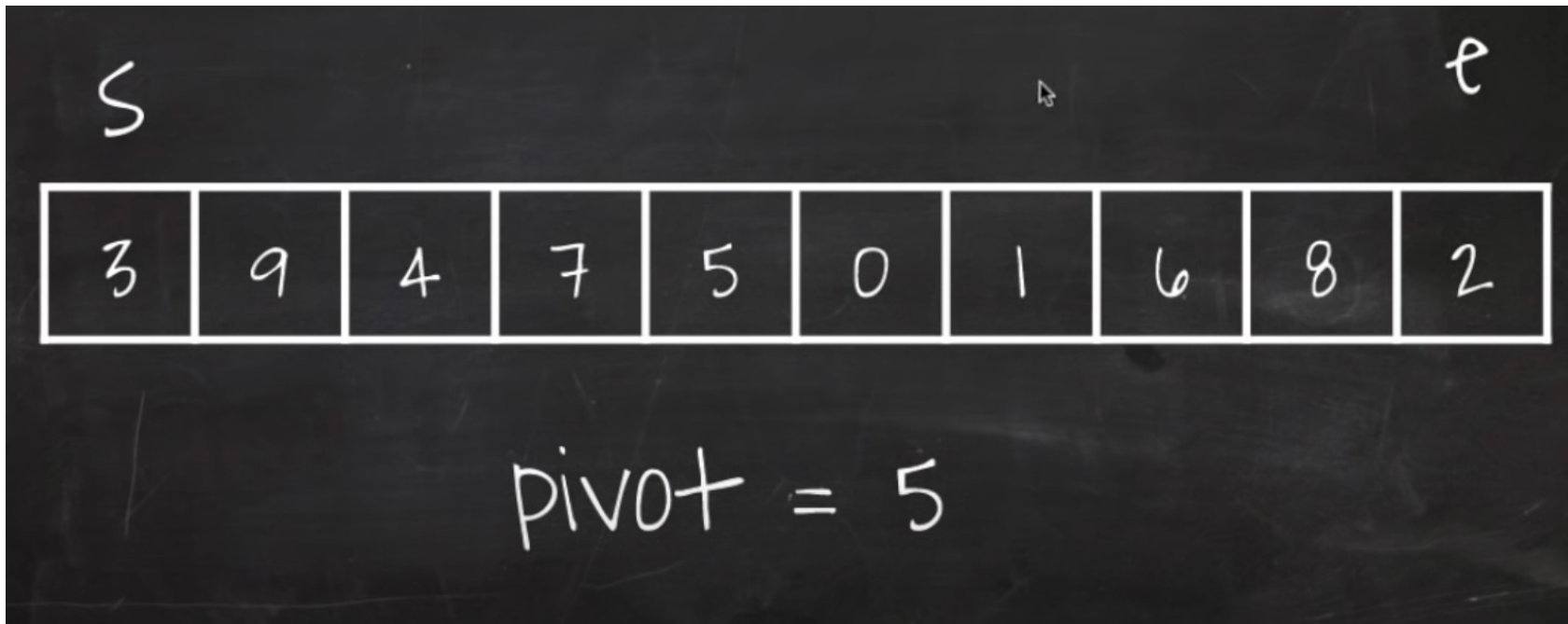


- 데이터 수( $N$ ) \* 분할 수( $\log N$ ) = 시간 복잡도는  $O(N \log N)$ 이다.
- 분할하는 과정에서 메모리에 저장 해둬야 하므로 메모리가 부족한 경우 비효율적이다

# 퀵정렬

## 퀵정렬이란?

- 임의의 피벗값(기준값)을 잡고 start, end점부터 pivot값과 크고 작은 여부를 비교해온다.
- $start > pivot$  &  $pivot > end$ 의 경우 swap(교환) 해준다



- 출처: 엔지니어대한민국 유튜브

# 퀵정렬

퀵정렬이란?

- 평균/최선 경우 시간 복잡도:  $O(N\log N)$
- 최악 경우 시간 복잡도:  $O(N^2)$
- 성능 향상 방법으로 Median-of-Three 방식이 있다.

# 버블정렬

## 버블정렬이란?

- 배열을 앞에서부터 2개를 묶는다.
- 묶은 값을 비교하여 정렬한다
- 이때 비교 순서가 넘어갈 때마다 값 하나는 중복이 된다.

80	40	50	10	70	20	30	60
----	----	----	----	----	----	----	----

1. (80, 40) -> (40, 80)

40	80	50	10	70	20	30	60
----	----	----	----	----	----	----	----

2. (80, 50) -> (50, 80)

40	50	80	10	70	20	30	60
----	----	----	----	----	----	----	----

- 모두 정렬이 될 때까지 루프를 돈다
- 시간 복잡도는  $O(N^2)$ 이다.