



1 HashSet을 통한 실 세계 데이터 분석

# 학습<mark>목표</mark>

■ HashSet을 활용하여 실 세계 데이터 분석을 수행할 수 있다.

## 이메일 데이터에 존재하는 모든 ID 중복없이 유지하기

모든 ID를 HashSet(Integer) idSet에 삽입(자동 중복 제거)

모든 ID에 대해 ArrayList(Integer)에 !contains(ID)라면 삽입

모든 ID에 대해 LinkedList(Integer)에 !contains(ID)라면 삽입



실험1

[HashSet을 이용한 실 세계 데이터 분석]

### 이메일 데이터에 존재하는 각 ID에 대해, 보낸 사람에 존재하는지 확인



HashSet, ArrayList, LinkedList 사이에서 Key 탐색의 효율성 확인

#### 모든 ID를 갖는 HashSet(Integer) idSet 준비

## 보낸사람 ID

HashSet(Integer) senderSet 준비

(중복없이) 갖는 ArrayList(Integer)

senderArrayList (중복없이) 갖는 LinkedList〈Integer〉 senderLinkedList 준비

## idSet을 순회하며, 각 ID에 대해 다음의 속도 측정

senderSet.contains(id) 수행한 총 시간

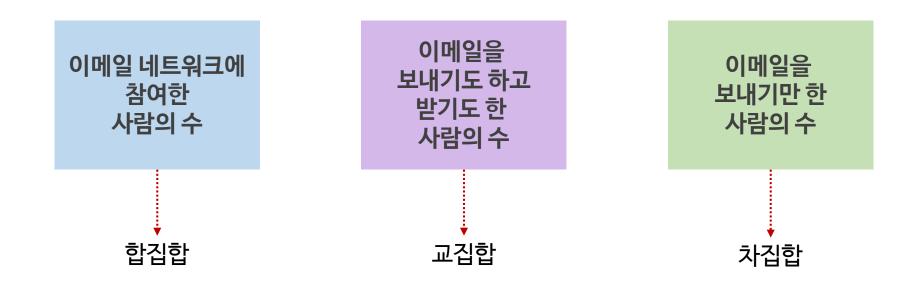
senderArrayList.contains(id) 수행한 총 시간

senderLinkedList.contains(id) 수행한 총시간

실험 2

[HashSet을 이용한 실 세계 데이터 분석]

## Set 집합 연산



[HashSet을 이용한 실 세계 데이터 분석]

실험 3



#### HashSet을 활용한 실 세계 데이터 분석

HashSet, ArrayList, LinkedList의 효율성 비교

key 탐색의 효율성

Set의 집합 연산을 이용한 문제