Array and Hashing

1. Two Sum

Pattern: Arrays & Hashing

Problem Statement

> Given an array of integers `nums` and an integer `target`, return indices of the two numbers such that they add up to `target`.
>
> You may assume that each input would have exactly one solution, and you may not use the same element twice.
>
> You can return the answer in any order.

Sample Input & Output

```
Input: nums = [2,7,11,15], target = 9
Output: [0,1]
Explanation: nums[0] + nums[1] == 9, so we return [0, 1].
```

```
Input: nums = [3,3], target = 6
Output: [0,1]
Explanation: Two identical elements at different indices are valid.
```

```
Input: nums = [1,2,3], target = 7
Output: [] (or raises; but problem guarantees one solution)
```

LeetCode Editorial Solution + Inline Tests

```python
from typing import List

class Solution:
    def twoSum(self, nums: List[int], target: int) -> List[int]:
        # STEP 1: Initialize hash map to store value → index
        #   - Why? To check in O(1) if complement (target - num) exists
        seen = {}

        # STEP 2: Iterate through array with index
        #   - Why index? We need to return positions, not values
        for i, num in enumerate(nums):
            complement = target - num

            # STEP 3: Check if complement already seen
            #   - If yes, we found our pair: current index + stored index
            if complement in seen:
                return [seen[complement], i]
                # return [complement, num] -> returns the actual number

            # STEP 4: Store current number and index for future lookup
            #   - Why here? To avoid using same element twice
            seen[num] = i

        # STEP 5: Return empty if no solution (per constraints, won't happen)
        #   - Included for safety / clarity
        return []

# ------------------- INLINE TESTS -------------------
if __name__ == "__main__":
    sol = Solution()

    #  Test 1: Normal case
    result1 = sol.twoSum([2, 7, 11, 15], 9)
```

```
    print(f"Test 1: {result1} → Expected: [0, 1]")
    assert result1 == [0, 1], "Test 1 Failed"

    #  Test 2: Edge case - duplicate values
    result2 = sol.twoSum([3, 3], 6)
    print(f"Test 2: {result2} → Expected: [0, 1]")
    assert result2 == [0, 1], "Test 2 Failed"

    #  Test 3: Tricky - negative numbers
    result3 = sol.twoSum([-1, -2, -3, -4, -5], -8)
    print(f"Test 3: {result3} → Expected: [2, 4]")
    assert result3 == [2, 4], "Test 3 Failed"

    print(" All inline tests passed!")
```

How to use: Copy-paste this block into `.py` or Quarto cell → run directly → instant feedback.

Example Walkthrough

Let's walk through `nums = [2, 7, 11, 15], target = 9`.

Initial state:
`seen = {}` — empty hash map.
We'll iterate with index `i` and value `num`.

Step 1 — i=0, num=2:
- `complement = 9 - 2 = 7`
- Is 7 in `seen`? No → skip return
- Store `seen[2] = 0` → `seen = {2: 0}`
→ *Why store?* So if later we see 7, we know 2 was at index 0.

Step 2 — i=1, num=7:
- `complement = 9 - 7 = 2`
- Is `2` in `seen`? Yes → at index 0
- Return `[seen[2], 1]` → `[0, 1]`

→ *Why not store 7 first?*
Because we check *before* storing — this ensures we never use same index twice.

→ *Pattern insight*:
We trade space (hash map) for time — instead of nested loops $O(n^2)$, we do one pass $O(n)$. Hashing lets us "remember" what we've seen and instantly find complements.

Complexity Analysis

- **Time Complexity**: `O(n)`

  We traverse the list once. Each hash map lookup and insertion is O(1) average case.

- **Space Complexity**: `O(n)`

  In worst case, we store n-1 elements in the hash map before finding the solution.

---

2. Contains Duplicate

Pattern: Arrays & Hashing

Problem Statement

> Given an integer array `nums`, return `true` if any value appears at least twice in the array, and return `false` if every element is distinct.

Sample Input & Output

```
Input: nums = [1,2,3,1]
Output: true
Explanation: The number 1 appears twice.
```

```
Input: nums = [1,2,3,4]
Output: false
Explanation: All elements are unique.
```

```
Input: nums = [1]
Output: false
Explanation: Single element cannot have duplicates.
```

LeetCode Editorial Solution + Inline Tests

```python
from typing import List

class Solution:
    def containsDuplicate(self, nums: List[int]) -> bool:
        # STEP 1: Initialize hash set to track seen numbers
        #   - Set gives O(1) lookup; tracks uniqueness efficiently
        seen = set()

        # STEP 2: Iterate through each number
        #   - If num already in set → duplicate found → return True
        for num in nums:
            if num in seen:
                return True
            seen.add(num)  # Add to set for future checks

        # STEP 3: No duplicates found during iteration
        #   - Return False only after full traversal
        return False
        # -> below mentioned code also do the trick
        # return True if len(set(nums)) != len(num) else False
```

```python
# ------------------- INLINE TESTS -------------------
if __name__ == "__main__":
    sol = Solution()

    #  Test 1: Normal case - duplicate exists
    assert sol.containsDuplicate([1,2,3,1]) == True, "Test 1 Failed"
    print("  Test 1 passed: [1,2,3,1] → True")

    #  Test 2: Edge case - single element
    assert sol.containsDuplicate([1]) == False, "Test 2 Failed"
    print("  Test 2 passed: [1] → False")

    #  Test 3: Tricky/negative - all unique, large input
    assert sol.containsDuplicate([i for i in range(1000)]) == False, \
        "Test 3 Failed"
    print("  Test 3 passed: 0..999 (all unique) → False")
```

How to use: Copy-paste this block into `.py` or Quarto cell → run directly → instant feedback.

Example Walkthrough

Let's walk through `nums = [1, 2, 3, 1]` step by step:

Initial State:
- `seen = set()` — empty
- `nums = [1, 2, 3, 1]`

Step 1: Process `num = 1`
- Check: Is `1 in seen`? → No
- Action: Add `1` to `seen` → `seen = {1}`
- Why? We record that we've seen `1` so future `1`s will trigger duplicate.

Step 2: Process `num = 2`
- Check: Is `2 in seen`? → No
- Action: Add `2` → `seen = {1, 2}`
- Why? Still building our uniqueness tracker.

Step 3: Process `num = 3`
- Check: Is `3 in seen`? → No
- Action: Add `3` → `seen = {1, 2, 3}`
- Why? No duplicate yet — continue.

Step 4: Process `num = 1` (again)
- Check: Is `1 in seen`? → Yes!
- Action: Immediately `return True`
- Why? We found a duplicate — no need to process further.
- Pattern Insight: Hashing lets us detect reoccurrence in O(1), breaking early saves time.

Final Output: `True`

This exemplifies the **Arrays & Hashing** pattern:
→ Use a hash set to track what we've seen.
→ Leverage O(1) average lookup to detect duplicates on the fly.
→ Early termination optimizes best-case performance.

Complexity Analysis

- **Time Complexity**: `O(n)`

  We iterate through the array once. Each `in` check and `add` operation on the set is O(1) average case. Worst-case total = O(n).

- **Space Complexity**: `O(n)`

  In the worst case (all elements unique), we store all `n` elements in the set. Space scales linearly with input size.

3. Majority Element

Pattern: Arrays & Hashing

Problem Statement

Given an array `nums` of size `n`, return the majority element.

The majority element is the element that appears more than `n / 2` times. You may assume that the majority element always exists in the array.

Sample Input & Output

```
Input: nums = [3,2,3]
Output: 3
Explanation: 3 appears 2 times > 3/2 = 1 → majority.
```

```
Input: nums = [2,2,1,1,1,2,2]
Output: 2
Explanation: 2 appears 4 times > 7/2 = 3 → majority.
```

```
Input: nums = [1]
Output: 1
Explanation: Single element is always majority by definition.
```

LeetCode Editorial Solution + Inline Tests

```python
from typing import List

class Solution:
    def majorityElement(self, nums: List[int]) -> int:
        # STEP 1: Initialize hash map to count occurrences
        #   - Tracks frequency of each element for quick lookup
        count = {}

        # STEP 2: Iterate and count each element
        #   - Invariant: after each step, count[x] = freq of x so far
        for num in nums:
            count[num] = count.get(num, 0) + 1

            # Early exit: if any count exceeds n//2, return immediately
            #   - Optimization: avoids full scan if found early
            if count[num] > len(nums) // 2:
                return num

        # STEP 3: Return fallback (problem guarantees existence)
        #   - Should never reach here per problem constraint
        return nums[0]

# ------------------- INLINE TESTS -------------------
if __name__ == "__main__":
    sol = Solution()

    #  Test 1: Normal case
    result1 = sol.majorityElement([3, 2, 3])
    print(f"Test 1 - Input: [3,2,3] → Output: {result1} (Expected: 3)")
    assert result1 == 3, "Test 1 Failed"

    #  Test 2: Edge case - single element
    result2 = sol.majorityElement([1])
    print(f"Test 2 - Input: [1] → Output: {result2} (Expected: 1)")
    assert result2 == 1, "Test 2 Failed"

    #  Test 3: Tricky case - long sequence, majority late
    result3 = sol.majorityElement([1, 1, 1, 2, 2, 2, 2])
    print(f"Test 3 - Input: [1,1,1,2,2,2,2] → Output: {result3} "
          f"(Expected: 2)")
    assert result3 == 2, "Test 3 Failed"
```

```python
    print("  All tests passed")
```

How to use: Copy-paste this block into `.py` or Quarto cell → run directly → instant feedback.

Example Walkthrough

Let's walk through `nums = [2,2,1,1,1,2,2]` step by step:

Initial State:
- `count = {}` (empty dictionary)
- `len(nums) = 7`, so majority threshold = `7//2 = 3`

Step 1: `num = 2`
→ `count.get(2, 0) + 1 = 0 + 1 = 1`
→ `count = {2: 1}`
→ `1 > 3?` → No → continue

Step 2: `num = 2`
→ `count[2] = 1 + 1 = 2`
→ `count = {2: 2}`
→ `2 > 3?` → No → continue

Step 3: `num = 1`
→ `count[1] = 0 + 1 = 1`
→ `count = {2:2, 1:1}`
→ `1 > 3?` → No → continue

Step 4: `num = 1`
→ `count[1] = 1 + 1 = 2`
→ `count = {2:2, 1:2}`
→ `2 > 3?` → No → continue

Step 5: `num = 1`
→ `count[1] = 2 + 1 = 3`
→ `count = {2:2, 1:3}`
→ `3 > 3?` → No → continue (note: must be >, not)

Step 6: `num = 2`
→ `count[2] = 2 + 1 = 3`

→ `count = {2:3, 1:3}`
→ `3 > 3?` → No → continue

Step 7: `num = 2`
→ `count[2] = 3 + 1 = 4`
→ `count = {2:4, 1:3}`
→ `4 > 3?` → YES → **return 2**

Why each step matters: - **Counting with dict**: Lets us track exact frequencies — core of hashing pattern. - **Early return**: Minor optimization — not required, but good practice. - **No fallback needed**: Problem guarantees majority exists, so we *will* hit return in loop.

Pattern Insight:
This is classic **Arrays & Hashing** — use a hash map to trade space for time, avoiding nested loops. Instead of $O(n^2)$, we get $O(n)$ by storing what we've seen.

Complexity Analysis

- **Time Complexity**: `O(n)`

  Single pass through array. Each dictionary `get` and assignment is O(1) average case.

- **Space Complexity**: `O(n)`

  In worst case, all elements are distinct until majority is found — hash map stores up to n keys.
  (Note: Boyer-Moore algorithm can solve this in O(1) space — but that's a different pattern.) -> stick with hashing for cleaner implementation.

4. Valid Anagram

Pattern: Arrays & Hashing

Problem Statement

Given two strings `s` and `t`, return `true` if `t` is an anagram of `s`, and `false` otherwise.

An **anagram** is a word or phrase formed by rearranging the letters of a different word or phrase, typically using all the original letters exactly once.

Sample Input & Output

```
Input: s = "anagram", t = "nagaram"
Output: true
Explanation: All characters in 't' are a rearrangement of 's'.
```

```
Input: s = "rat", t = "car"
Output: false
Explanation: 't' has different characters than 's'.
```

```
Input: s = "", t = ""
Output: true
Explanation: Two empty strings are trivially anagrams.
```

LeetCode Editorial Solution + Inline Tests

```python
from typing import List

class Solution:
    def isAnagram(self, s: str, t: str) -> bool:
        # STEP 1: Initialize structures
        #   - Use two hash maps (dicts) to count char frequencies in s and t.
        #   - More flexible than array - handles any Unicode char.
        if len(s) != len(t):
            return False

        count_s, count_t = {}, {}
```

```python
        # STEP 2: Main loop - populate both hash maps
        #   - Traverse by index since strings are same length.
        #   - Use .get() to safely handle missing keys.
        for i in range(len(s)):
            count_s[s[i]] = 1 + count_s.get(s[i], 0)
            count_t[t[i]] = 1 + count_t.get(t[i], 0)

        # STEP 3: Update state / bookkeeping
        #   - No incremental update needed - we build full maps first.
        #   - Final comparison validates anagram property.

        # STEP 4: Return result
        #   - Direct dict equality check - Python compares keys & values.
        return count_s == count_t

# ------------------- INLINE TESTS -------------------
if __name__ == "__main__":
    sol = Solution()

    #  Test 1: Normal case
    assert sol.isAnagram("anagram", "nagaram") == True, "Normal case failed"
    print(" Test 1 passed: 'anagram' vs 'nagaram'")

    #  Test 2: Edge case - empty strings
    assert sol.isAnagram("", "") == True, "Empty strings should be anagrams"
    print(" Test 2 passed: empty strings")

    #  Test 3: Tricky/negative - different letters
    assert sol.isAnagram("rat", "car") == False, "Mismatched letters"
    print(" Test 3 passed: 'rat' vs 'car'")
```

How to use: Copy-paste this block into `.py` or Quarto cell → run directly → instant feedback.

Example Walkthrough

Let's walk through `s = "anagram"`, `t = "nagaram"`:

Initial State: `count_s = {}`, `count_t = {}`, and `len(s) == len(t)` → proceed.

After i=0 (`s[0]='a'`, `t[0]='n'`):
→ `count_s['a'] = 1 + 0 → {'a': 1}`
→ `count_t['n'] = 1 + 0 → {'n': 1}`

After i=1 (`s[1]='n'`, `t[1]='a'`):
→ `count_s['n'] = 1 + 0 → {'a':1, 'n':1}`
→ `count_t['a'] = 1 + 0 → {'n':1, 'a':1}`

Continue to end:
→ `'a'` appears 3x in both → count_s['a']=3, count_t['a']=3
→ `'g'`,`'r'`,`'m'`,`'n'` each appear 1x → all match

Final Comparison:
→ `count_s == {'a':3, 'n':1, 'g':1, 'r':1, 'm':1}`
→ `count_t == {'n':1, 'a':3, 'g':1, 'r':1, 'm':1}`
→ Dictionaries are equal → return `True`

Why necessary?
→ Length check avoids unnecessary work and index errors.
→ `.get(key, 0)` prevents KeyError — critical for robust hashing.
→ Direct `dict1 == dict2` leverages Python's deep equality — clean and readable.
→ This exemplifies **frequency counting with hash maps** — a foundational Arrays & Hashing technique.

Complexity Analysis

- **Time Complexity**: `O(n)`

  Single pass over both strings (n steps), then O(1) dict comparison (since alphabet size is bounded — even for Unicode, in practice it's limited per input).

- **Space Complexity**: `O(k)`

  Where k = number of unique characters. In worst case, k = n (all chars distinct). For lowercase English, k 26 → effectively O(1). For general Unicode, O(k) where k scales with input diversity.

5. Group Anagrams

Pattern: Arrays & Hashing

Problem Statement

Given an array of strings `strs`, group the anagrams together. You can return the answer in **any order**.

An Anagram is a word or phrase formed by rearranging the letters of a different word or phrase, typically using all the original letters exactly once.

Constraints: - `1 <= strs.length <= 10^4` - `0 <= strs[i].length <= 100` - `strs[i]` consists of lowercase English letters.

Sample Input & Output

```
Input: strs = ["eat","tea","tan","ate","nat","bat"]
Output: [["eat","tea","ate"],["tan","nat"],["bat"]]
Explanation: Words with same frequency signature
grouped together.

Input: strs = [""]
Output: [[""]]
Explanation: Single empty string forms its own group.

Input: strs = ["a"]
Output: [["a"]]
Explanation: Single letter trivially grouped.
```

LeetCode Editorial Solution + Inline Tests

```python
from typing import List
from collections import defaultdict

class Solution:
    def groupAnagrams(self, strs: List[str]) -> List[List[str]]:
        # STEP 1: Create hashmap to group by freq signature
        anagram_map = defaultdict(list) # -> defaultdict(list, {})
```

```python
        for word in strs:
            # STEP 2: Count frequency of 26 letters
            freq = [0] * 26
            for ch in word:
                freq[ord(ch) - ord('a')] += 1

            # STEP 3: Use tuple of freq as key
            key = tuple(freq)
            anagram_map[key].append(word)

        # STEP 4: Return grouped anagrams
        return list(anagram_map.values())

# ------------------- INLINE TESTS -------------------
if __name__ == "__main__":
    sol = Solution()

    #   Test 1: Normal case
    strs1 = ["eat","tea","tan","ate","nat","bat"]
    print(sol.groupAnagrams(strs1))
    # Expected: groups like [["eat","tea","ate"],["tan","nat"],["bat"]]

    #   Test 2: Edge case - single empty string
    strs2 = [""]
    print(sol.groupAnagrams(strs2))
    # Expected: [[""]]

    #   Test 3: Tricky case - single letter
    strs3 = ["a"]
    print(sol.groupAnagrams(strs3))
    # Expected: [["a"]]
```

How to use: Copy-paste this block into `.py` or Quarto cell → run directly → instant feedback.

Detailed Example Walkthrough

Pick input: `["eat", "tea", "tan", "ate", "nat", "bat"]`.

1. **Initialize**: `anagram_map = {}` (empty defaultdict). Tracks groups keyed by letter-frequency tuple.
2. **Word = "eat"**
   - freq = `[1,0,0,0,1,0,...,1,...]`
   - key = `(1,0,0,0,1,0,...,1,...)`
   - Add → `anagram_map[key] = ["eat"]`.
3. **Word = "tea"**
   - Same frequency tuple as `"eat"`.
   - Append → `["eat","tea"]`.
4. **Word = "tan"**
   - freq different → new key.
   - Add → `["tan"]`.
5. **Word = "ate"**
   - Same key as `"eat"`.
   - Append → `["eat","tea","ate"]`.
6. **Word = "nat"**
   - Same key as `"tan"`.
   - Append → `["tan","nat"]`.
7. **Word = "bat"**
   - New frequency key → `["bat"]`.
8. **Final Output**: `[["eat","tea","ate"],["tan","nat"],["bat"]]`.

Pattern Insight: We reduce each string into a **hashable signature** (frequency tuple). Hashing ensures `O(1)` group lookup, which is the core of **Arrays & Hashing** mastery.

Complexity Analysis

- **Time Complexity**:
  - For each word (n words), count letters (k length).
  - Total: `O(n * k)`.
- **Space Complexity**:
  - Hashmap stores up to `n` groups.
  - Each key is a 26-length tuple.
  - Total: `O(n * k)` for output + `O(26)` per key.

"'