Evaluating the Effectiveness of HashMap, TreeMap, and ConcurrentSkipListMap for an Extractive Text Summarizer in Java

Juan Felix Kusnadi (2802536386)

Jonathan Mulyono (2802537054)

Iglesias Sidharta Handjo (2802530621)

Data Structures

BINUS University International

Introduction

A summary is a brief description that gives the main facts or ideas of something (Cambridge Dictionary, n.d.). Text summarization plays a significant role in a world with vast amounts of information, enabling humans to understand the importance of a lengthy text within a shorter period. Our interest in text summarization started from the desire to explore natural language processing and to deepen our understanding of object-oriented programming and the characteristics of different data structures. As such, we will develop an extractive text summarizer based on the Term Frequency-Inverse Document Frequency (TF-IDF) algorithm, a widely used statistical method for ranking the importance of sentences. In this study, we will implement this algorithm in Java to evaluate the three Map data structures, which are HashMap, TreeMap, and ConcurrentSkipListMap. Through this study, we aim to explore how the choice of different data structures impacts the performance of an extractive text summarizer. This project aims to produce a robust and modular extractive text summarizer and evaluate each Map's effectiveness in terms of runtime and memory behaviour. As large volumes of unstructured information continue to flood, this study offers insight to anyone willing to build an efficient summarization tool.

Problem Description

In the current modern world, vast amounts of textual data are generated every day. This makes it increasingly difficult for humans to process and extract key information efficiently. This overloading of information often creates a need for summarization tools that shorten texts into a meaningful, smaller piece without losing the essential content. Extractive summarization, which selects and compiles sentences based on their importance (Payong, 2024), offers a practical and effective solution. However, implementing this solution efficiently requires choosing the correct type of data structure to yield a balance between accuracy and performance, ensuring that the tool runs on varying text sizes (Ardash, 2024). As such, this study addresses the problem by investigating how different Java Map implementations affect the performance and behaviour of an extractive summarizer tool based on the TF-IDF approach.

Approach

Term Frequency-Inverse Document Frequency (TF-IDF) is a statistical approach used in natural language processing (NLP) to evaluate the importance of a word within a document relative to a collection of documents (corpus) (Jain, 2025). The key workings of TF-IDF state that:

1. Words that appear more often in a document are more important (higher Term Frequency) (Jain, 2025).
2. Words that are common across different documents in the corpus are less important as they are less useful for distinguishing different documents (lower Inverse Document Frequency) (Jain, 2025).

Unlike word frequency, TF-IDF balances common and rare terms to grasp the most important and meaningful words, hence helping to produce a relevant summary. In the context of this research, the algorithm will be modified to account for a single document or text input only, instead of a collection of documents. As such, calculating the Inverse Document Frequency will be based on the number of sentences from the input.

Methodology

Preparing the text input

1. Tokenize the input into distinct sentences, and tokenize each sentence into words.
2. Remove common stopwords such as “the”, “is”, “in”, “and”, etc., that are meaningless in identifying important terms.
3. Convert all words into lowercase for case-insensitive comparison.

4. Reduce words to their root form using stemming.

Calculate Term Frequency (TF)

1. Calculate w, the number of times word w appears.
2. Calculate S, the number of words in sentence S.
3. Calculate Term Frequency by $TF = \frac{w}{S}$.

Calculate Inverse Document Frequency (IDF)

1. Calculate N, the total number of sentences.
2. Calculate $SF(w)$, the sentence frequency for the number of sentences containing the word w.
3. Calculate Inverse Document Frequency by $IDF = log(\frac{N}{SF(w)})$.

Creating a summary

1. Calculate TF-IDF by $TF \times IDF$.
2. Store this key (word) and value (TF-IDF score) pair, and repeat this calculation for all words in all sentences.
3. Rank the importance of every sentence by finding the sum of the TF-IDF scores of all words in that sentence.
4. Select the top 20%-40% of sentences from the total number of sentences based on the user input (Demilie, 2022).

Choice of Data Structures

In this study, three implementations of the Map interface are compared based on their performance in executing the TF-IDF algorithm. Each implementation offers unique characteristics and ordering guarantees, which significantly affect both runtime and memory behaviour.

1. HashMap

    HashMap is an unsynchronized implementation of Map based on the hash table, storing the data in key-value pairs. HashMap uses a technique known as hashing, which converts a large String into a smaller one (GeeksforGeeks, 2025). The value stored in a HashMap can be accessed through its corresponding key. HashMap is chosen for this study as it is the default key-value storage in Java when order is not required.

    a. Advantages
        i. Fast operations for insertion and retrieval.
        ii. Lower overhead compared to tree-based maps when storing a large number of entries (GeeksforGeeks, 2025).
        iii. No need to maintain order.
    b. Disadvantages
        i. Values are returned in an arbitrary order, which can compromise the chronological order of a summary.
        ii. Poor distribution of hash codes can lead to more collisions, hence degrading performance.

2. TreeMap

    TreeMap is a Red-Black Tree implementation of NavigableMap that sorts entries based on their key order or through the use of a Comparator (GeeksforGeeks, 2025). TreeMap is an efficient way to sort key-value pairs, hence supporting the retrieval of sentence scores (GeeksforGeeks, 2025).

    a. Advantages
        i. TreeMap always returns entries in key order.
    b. Disadvantages

i. Higher memory overhead due to the presence of tree nodes and pointers.

3. ConcurrentSkipListMap

ConcurrentSkipListMap is based on the ConcurrentNavigableMap, which implements a concurrent variant of SkipList (Oracle, n.d.). A SkipList is a probabilistic data structure, where the average time complexity of common Map operations is determined by probabilistic analysis. A SkipList works by randomly skipping some elements in order to have faster searching (GeeksforGeeks, 2024).

a. Advantages

i. When the data size grows enormously, SkipList may provide a very efficient lookup due to its probabilistic nature.

b. Disadvantages

i. The efficient lookup holds only in expectation. In the worst-case scenario, performance degrades into O(n) time.

ii. More memory overhead due to the presence of multiple forward-pointer references.

Time and Space Complexity

Data Structure	**Best Time**	**Average Time**	**Worst Time**	**Space**
HashMap	O(1)	O(1)	O(n)	O(n)
TreeMap	O(log n)	O(log n)	O(log n)	O(n)
ConcurrentSkipListMap	O(log n)	O(log n)	O(n)	O(n log n)

Table 1. Time and space complexities for HashMap, TreeMap, and ConcurrentSkipListMap.

A HashMap uses an array of buckets with additional chaining (a LinkedList or a Red-Black Tree). A bucket is an index of the array where key-value pairs are stored (Pandey, 2023). HashMap stores values uniquely through generating a unique key (bucket index) using a hash function. In the event of a collision, chaining will occur on the index of the array. When the chain grows beyond the default 8 entries, the LinkedList will become a Red-Black Tree (Martinez, 2024). HashMap has an average time complexity of O(1) (similar to an ordinary array) when the hash function spreads data uniformly and each bucket index only holds a few entries. Worst time complexity of O(n) occurs when all hash function generates the same hash code, leading to a linear scan of the chain of length n, where n is the input size. Space complexity of O(n) is due to the proportional growth of memory size based on the input, as each key-value pair is stored in a LinkedList node or a Red-Black Tree node.

A TreeMap implements a Red-Black Tree, which ensures that every operation traverses from the root to a leaf node in roughly $\log_2 n$ steps, then does a number of rotations or colorings to maintain the balance and sorted order (Li et al., 2023). As such, it has a time complexity of O(log n). The O(n) space complexity is due to storing each entry in a single tree node.

Initially, ConcurrentSkipListMap assigns a new node with a random height value with a probability of getting level 1 of 0.5, getting level 2 of 0.25, getting level 3 of 0.125, etc (Tuladhar, 2015). Since each higher level occurs theoretically half as often, the overall height grows in $\log_2 n$. As such, the searching process takes roughly $\log_2 n$ forward movements and downward drops. This results in an average time complexity of O(log n). In the worst case scenario where every single node is assigned level 1, the SkipList transforms into a simple sorted LinkedList of size n, where n is the input size. In addition to storing n number of pointers on level 1, it also stores $\log_2 h$ number of pointers on level h. As such, this yields a space complexity of O(n log n).

Hypothesis

With an increase in input size, it is hypothesized that the memory behavior of HashMap, TreeMap, and ConcurrentSkipListMap will be similar, with a probability that it would be larger for ConcurrentSkipListMap on certain tests due to its probabilistic nature. The growth of runtime will be based on:

ConcurrentSkipListMap > TreeMap > HashMap

Preliminary Testing

A preliminary test was conducted on the command line version of the program for various types of input text, including informative articles, narrative texts, and news. While there was no significant problem found for informative and narrative texts, a major problem regarding the splitting of sentences was found for news articles. As news articles often contain honorifics, such as Mr, Ms, Dr, Prof, etc., splitting the input only using regular expressions is not accurate. The program recognizes the "." after the honorifics as the end of a sentence. As such, a customized sentence splitter was built based on three main strategies: skipping closing quotes, checking whether the next part after a full stop looks like a new sentence (either opening quotes or capital letters), and checking whether a word is an honorific (Appendix 1).

Application using Java

```java
import java.util.Arrays;
import java.util.HashSet;
import java.util.List;
import java.util.Set;
import java.util.stream.Collectors;

public class DefaultTextProcessor implements TextProcessor {
    private final Stemmer stemmer;
    private final Set<String> stopWords;
    private SentenceSplitter splitter = new SentenceSplitter();

    public DefaultTextProcessor(Stemmer stemmer) {
        this.stemmer = stemmer;
        this.stopWords = loadStopWords();
    }

    public static Set<String> loadStopWords() {
        return new HashSet<>(Arrays.asList(...a:"the", "is", "in", "and", "a", "to", "of",
        "for", "on", "by", "with", "at", "as", "from"));
    }

    public List<List<String>> preprocess(String text) {
        List<String> sentences = List.of(splitter.splitIntoSentences(text));
        return sentences.stream()
            .map(s -> Arrays.stream(
                    s.replaceAll(regex:"[^a-zA-Z\\s]", replacement:"") // Remove non alphabetical characters
                     .toLowerCase() // Convert the text to lowercase
                     .split(regex:"\\s+")) // Split the sentence into words
                .filter(tok -> !tok.isEmpty() && !stopWords.contains(tok)) // Filter out empty tokens and stop words
                .map(stemmer::stem) // Apply stemming
                .collect(Collectors.toList())) // Collect processed words as a list
            .collect(Collectors.toList()); // Collect all sentence lists into a list of sentences
    }
}
```

Figure 1. DefaultTextProcessor class.

Preparing the text input was done by the DefaultTextProcessor class, which implements the TextProcessor interface (Appendix 2). After the input text has been split by an instance of the SentenceSplitter class, and stopwords have been removed by filtering based on the set generated by the loadStopWords() method, stemming is handled by the SnowballStemmer class (Appendix 3) implementing Stemmer interface (Appendix 4). SnowballStemmer uses an external implementation of the Porter stemming method to reduce words into their root form. This stemming algorithm uses NLP to remove common endings of a word, such as from running to run and ran to run. This ensures that different forms of a word are treated as a single word.

After the input text is processed, sentences will be extracted to generate a summary based on the TF-IDF algorithm. This is done by the TFIDFSummarizer concrete class (Appendix 5), which extends the AbstractSummarizer abstract class (Appendix 6).

```
public TFIDFSummarizer(Supplier<Map<String, Double>> doubleMapSupplier, Supplier<Map<String, Integer>> intMapSupplier,
TextProcessor textProcessor) {
    super(doubleMapSupplier, intMapSupplier);
    this.textProcessor = textProcessor;
}
```

Figure 2. TFIDFSummarizer constructor.

As the aim of this study is to evaluate the effectiveness of HashMap, TreeMap, and ConcurrentSkipListMap, the constructor of the TDIDFSummarizer class uses the Supplier<T> functional interface, which returns type T when get() is called. The use of Supplier creates flexibility for this class, allowing it to seamlessly use the three different types of Map without any change of code. As shown by the figure below, the constructor initializes a new instance of TFIDFSummarizer that uses two HashMaps, making it more reusable and testable. In addition, it makes the comparison between the three Maps more reliable as no changes in the code are made.

```
TFIDFSummarizer summarizer = new TFIDFSummarizer(HashMap::new, HashMap::new, new DefaultTextProcessor(new SnowballStemmer()));
```

Figure 3. TFIDFSummarizer constructor.

Finally, a Graphical User Interface (GUI) version of the program is created using JavaFX to allow an easier testing process, which involves large input and output sizes. Furthermore, several features are also added to enhance user experience, which include customizing the summary to be in paragraph or bullet points, showing word count for input and summarized text, a copy and paste button, a clear button, customizing the length of the summary, and statistics of the summary. Figure 4 shows the GUI window of this program.

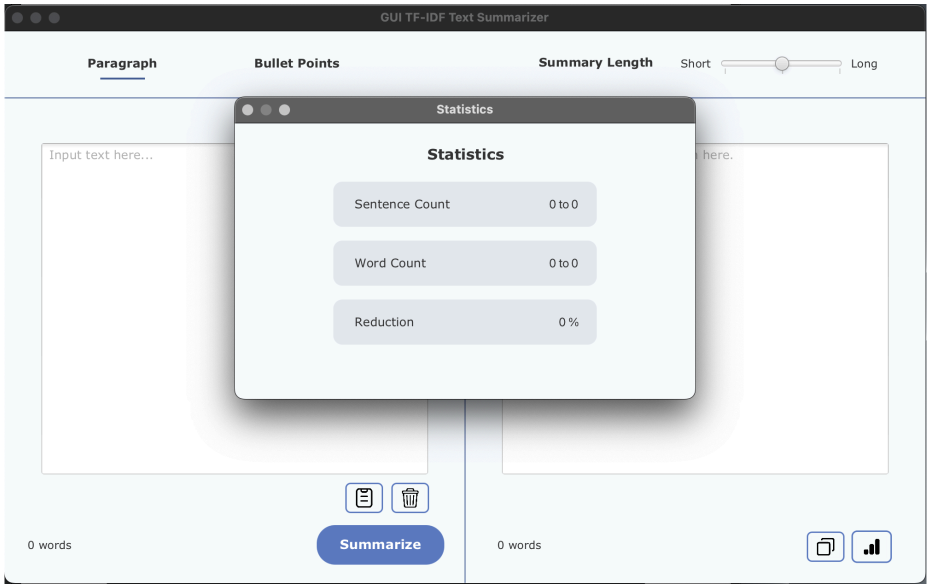


Figure 4. GUI program window.

Data Collection and Analysis

The performance of HashMap, TreeMap, and ConcurrentSkipListMap will be accurately measured using the lines of code shown in Figure 5. Although the runtime and memory behaviour are measured when summarize() is called, the differences in these two variables will be solely due to the difference in the performance of the get() and put() methods of the TFIDFSummarizer class. This is due to the fact that all lines of code remain the same (Appendix 5).

```java
summarizeBtn.setOnAction((ActionEvent ae) -> {
    Runtime runtime = Runtime.getRuntime();
    runtime.gc();
    long memoryBefore = runtime.totalMemory() - runtime.freeMemory();
    long startTime = System.nanoTime();
    String summary = summarizer.summarize(inputArea.getText (), summaryRatio);
    long endTime = System.nanoTime();
    long memoryAfter = runtime.totalMemory() - runtime.freeMemory();
    double duration = (double) (endTime - startTime) / 1_000_000;
    double memoryUsed = (double) (memoryAfter - memoryBefore) / 1024 / 1024;
    System.out.printf(format:"Runtime: %.3f ms, ", duration);
    System.out.printf(format:"Memory used: %.3f MB\n", memoryUsed);

    if (!paragraphFormat) {
        String[] sents = splitter.splitIntoSentences(summary);
        String formatted = IntStream.range(startInclusive:0, sents.length).mapToObj(i -> "• " + sents[i].trim())
            .collect(Collectors.joining(delimiter:"\n"));
        outputArea.setText(formatted);
    } else {
        outputArea.setText(summary);
    }
    outputCount.setText(countWord(summary) + " words");
});
```

Figure 5. Measuring runtime and memory behaviour.

Apple MacBook Pro (13-inch, 2020, Four Thunderbolt 3 ports)	
Processor	2 GHz Quad-Core Intel Core i5
Memory	16 GB 3733 MHz LPDDR4X
macOS	Sonoma 14.1

Table 2. Computer specifications used for testing.

#	Input (AI Generated)	Output (medium summary length)
1	Cybersecurity is essential in today's … keeping information secure. (100 words)	Cybersecurity is essential in today's … sensitive personal data. (20 words)
2	Artificial intelligence (AI) is revolutionizing the healthcare industry … determines their success in global health. (500 words)	Natural language processing (NLP) allows AI … success in global health. (51 words)
3	Renewable energy refers to sources … high-carbon infrastructure and transition directly. (1000 words)	Understanding how these sources work, their advantages … renewable portfolio standards. (112 words)
4	Social media has become an integral part of modern society ... of social media while minimizing harm. (2000 words)	For example, passive use, defined as scrolling … limits the ability to draw causal inferences. (124 words)
5	Coral reefs are among the most diverse and productive ... science to maintain credibility and trust. (5000 words)	Fertilization occurs externally, and resulting … emphasizing the economic logic of proactive investment. (192 words)

Table 3. Input and output of the program (Refer to Appendix 7-11 for full text).

The tables below show the runtime and memory behavior as raw data collected from running the above lines of code using medium summary length and paragraph format, retrieved from the terminal. The runtime data is collected after running the program 15 times to allow the Java Virtual Machine to warm up. It is noticeable that on the first several runs of the program, the runtime is significantly higher due to class loading and initialization, Just-in-Time (JIT) compilation, and garbage collection. As such, the data will be recorded for runs 16-20.

Runtime for input 1 (100 words) in ms			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	1.638	1.731	1.864
Trial 2	1.585	1.651	1.808
Trial 3	1.467	1.877	1.724
Trial 4	1.441	1.535	1.559
Trial 5	1.273	1.477	1.632
Runtime for input 2 (500 words) in ms			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	2.174	2.683	2.672
Trial 2	1.952	2.222	2.824
Trial 3	1.772	2.311	2.351
Trial 4	1.845	2.456	2.683
Trial 5	1.840	2.154	2.266
Runtime for input 3 (1000 words) in ms			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	2.567	3.456	3.498
Trial 2	2.814	3.416	3.616
Trial 3	2.723	3.375	3.380
Trial 4	2.777	3.251	3.580
Trial 5	2.939	3.514	3.363
Runtime for input 4 (2000 words) in ms			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	3.505	6.586	4.739
Trial 2	3.317	5.007	5.317
Trial 3	3.577	5.303	4.898
Trial 4	3.447	4.976	4.942

Trial 5	3.150	4.520	5.152
Runtime for input 5 (5000 words) in ms			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	5.211	9.666	10.686
Trial 2	6.605	9.236	10.282
Trial 3	6.625	9.221	10.258
Trial 4	6.659	9.697	10.440
Trial 5	6.318	9.109	10.195

Table 4. Runtime data collection.

Memory used for input 1 (100 words) in MB			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	0.892	0.668	0.893
Trial 2	0.892	0.588	0.893
Trial 3	0.646	0.834	0.647
Trial 4	0.590	0.914	0.837
Trial 5	0.590	0.834	0.591
Memory used input 2 (500 words) in MB			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	1.012	1.192	1.008
Trial 2	0.797	1.023	1.120
Trial 3	0.703	1.023	0.812
Trial 4	0.797	1.008	1.008
Trial 5	0.703	1.023	1.008
Memory used for input 3 (1000 words) in MB			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	0.990	0.779	1.216

Trial 2	1.374	1.274	1.524
Trial 3	1.130	1.170	1.123
Trial 4	1.234	1.170	1.524
Trial 5	0.990	1.274	1.123
Memory used for input 4 (2000 words) in MB			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	1.640	1.525	1.415
Trial 2	1.214	1.678	1.658
Trial 3	1.733	1.678	1.415
Trial 4	1.551	1.525	1.658
Trial 5	1.490	1.678	1.415
Memory used for input 5 (5000 words) in MB			
	HashMap	TreeMap	ConcurrentSkipListMap
Trial 1	2.839	2.907	2.986
Trial 2	2.839	2.868	2.769
Trial 3	2.924	2.907	2.769
Trial 4	2.839	2.907	2.769
Trial 5	2.924	2.907	2.769

Table 5. Memory behaviour data collection.

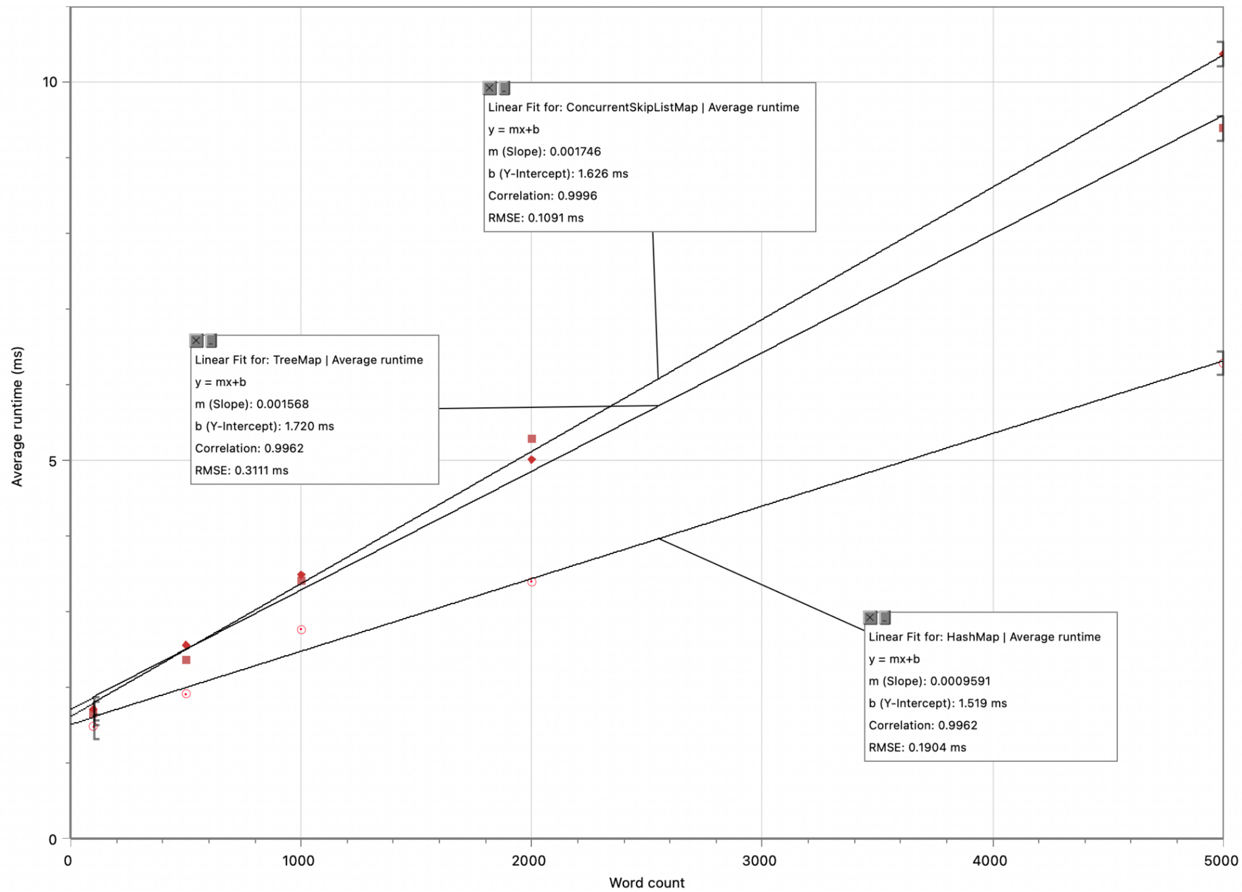
From the raw data collected, it is noticeable that the variability of runtime tends to be higher than memory behaviour, shown by the repeating values of memory used on different trials of the same input. Studies have shown that the object allocation patterns are highly repeatable for a given Java program and input (Li et al., 2021). The average memory used for different trials remains nearly constant. Another research focusing on heap usage also demonstrates that persistent objects are allocated in a predictable way, resulting in a stable heap occupancy (Briggs et al., 2016). Conversely, runtime data demonstrates high variability due to factors like timing and order of compilation, cache state, CPU scheduling, and garbage collection (Barrett et al., 2017).

Average Runtime (ms)	HashMap	TreeMap	ConcurrentSkipListMap
Input 1	1.481	1.654	1.717
Input 2	1.917	2.365	2.559
Input 3	2.764	3.402	3.487
Input 4	3.399	5.278	5.010
Input 5	6.284	9.386	10.372

Table 6. Average runtime for all inputs.

Average Memory used (MB)	HashMap	TreeMap	ConcurrentSkipListMap
Input 1	0.722	0.768	0.772
Input 2	0.802	1.054	0.990
Input 3	1.144	1.133	1.302
Input 4	1.526	1.617	1.512
Input 5	2.873	2.899	2.812

Table 7. Average memory used for all inputs.



Graph 1. The relationship between average runtime (ms) and word count.

The graph above shows the relationship between average runtime in milliseconds and word count. Although the average time complexity of TreeMap and ConcurrentSkipListMap is O(log n), linear best fit is chosen to represent the relationship, as using a logarithmic function does not provide a good representation. The figure below shows the relationship represented by a logarithmic function, proving a bad representation due to not passing through any of the plotted points. Hence, using linear fit is still reasonable considering the small range of input size used (only 100 to 5000).

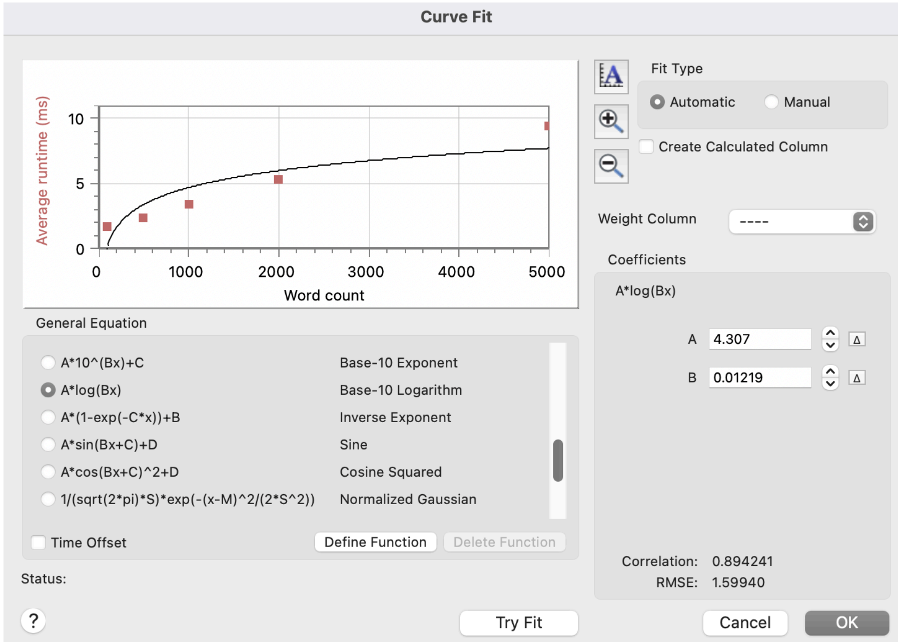
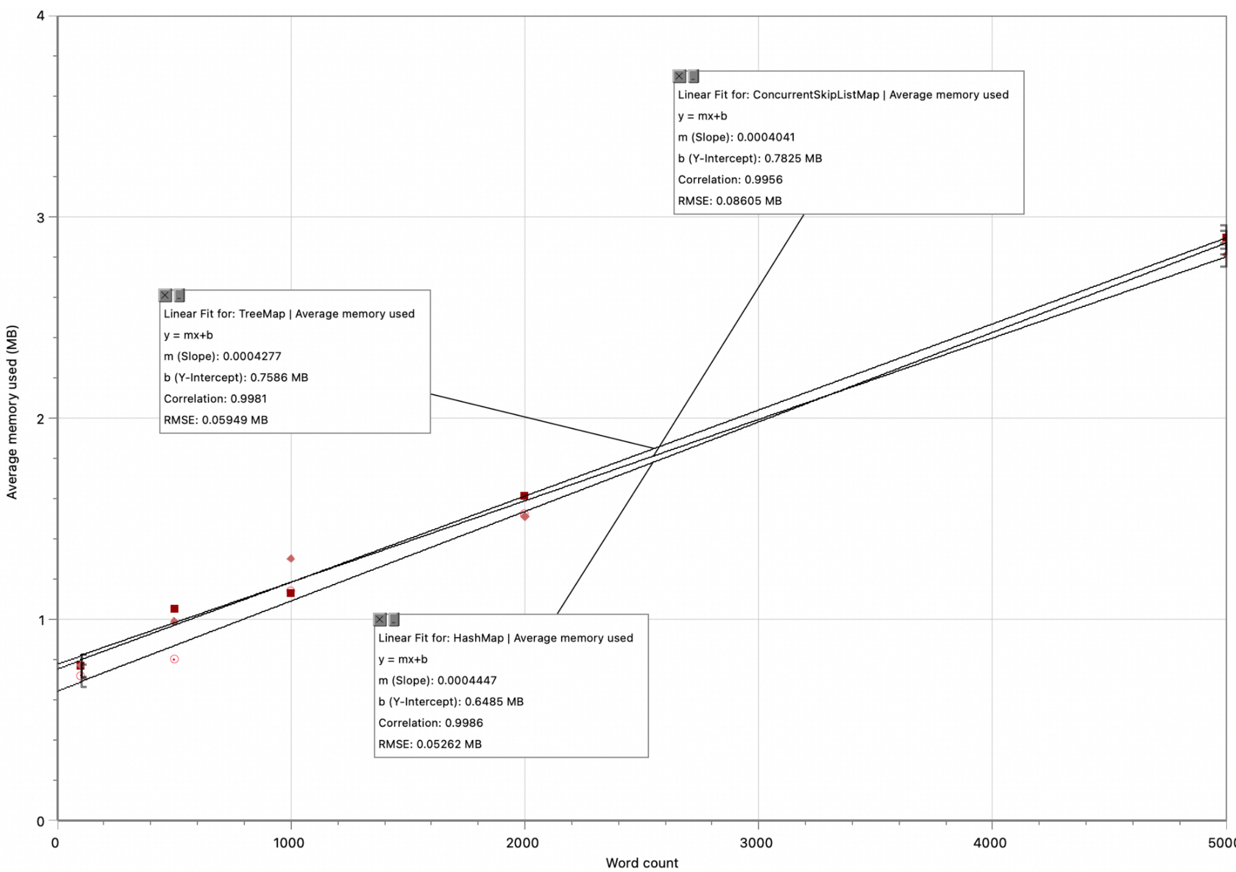


Figure 6. Logarithmic fit for TreeMap.

As word count increases, the increase in runtime for HashMap, TreeMap, and ConcurrentSkipListMap is different due to their unique nature. The increase in runtime due to an increase in word count for HashMap is significantly lower than for TreeMap and ConcurrentSkipListMap. This is proven by its lower gradient of 0.0009591, showing a larger gap between the two lines as word count increases further. The gradient of best fit for HashMap is about 39% lower than that of TreeMap, and 45% lower than that of ConcurrentSkipListMap. However, it is noticeable that the gradient of HashMap is not zero, as it is stated by the O(1) time complexity. The gradient for this line is roughly 0.0009591, which is close to the theoretical gradient of 0. The reason behind this discrepancy will be discussed in the evaluation section.

While the trend for TreeMap and ConcurrentSkipListMap is very similar and supports the O(n) time complexity through its considerably constant positive gradient, their gradients of best fit are different. The gradient of best fit for TreeMap is 0.001568, and 0.001746 for

ConcurrentSkipListMap, which is roughly 11% higher. This is due to the previously discussed probabilistic nature of ConcurrentSkipListMap, which involves giving each data node a level randomly and searching through it. This is also shown by Table 1, where the worst time complexity of ConcurrentSkipListMap is O(n), while O(log n) for TreeMap. This supports the initial hypothesis for the growth of runtime based on word count, that is ConcurrentSkipListMap > TreeMap > HashMap.



Graph 2. The relationship between average memory used (MB) and word count.

Graph 2 shows the relationship between average memory used in MB and word count for HashMap, TreeMap, and ConcurrentSkipListMap. The similar gradients of best fit of 0.0004447, 0.0004277, and 0.0004041 prove that the memory behaviour for the three types of

Map is relatively the same. While ConcurrentSkipListMap has a theoretical space complexity of O(n log n) due to the probable presence of additional forward pointers on different levels, its gradient is experimentally similar to HashMap and TreeMap. This shows that the test done resulted in an ideal scenario for ConcurrentSkipListMap, where the number of forward pointers on levels greater than 1 is small. However, further analysis on its memory behaviour cannot be made due to the limited range of input size. As such, this supports the hypothesis made, which stated that the memory behavior of HashMap, TreeMap, and ConcurrentSkipListMap is similar.

Conclusion

The result of this study supports the initial hypothesis, which shows that HashMap is the best in terms of time complexity, and all HashMap, TreeMap, and ConcurrentSkipListMap demonstrate a similar space complexity. All lines of best fit from Graph 1 and Graph 2 show a high correlation coefficient close to 1, proving a strong positive relationship between average runtime and memory used, to word count. Graph 1 shows that the runtime growth is greatest for ConcurrentSkipListMap and smallest for HashMap, proving HashMap to be the most efficient for an extractive summarizer tool. Additionally, Graph 2 shows that the memory behaviour of the three Maps is similar, as proven by the similar value of the gradient of best fit. As such, it can be concluded that HashMap is the most suitable implementation of Map with the lowest time complexity and reasonable space complexity.

Evaluation

Although the result supports the initial hypothesis, there are two main issues within this study that contradict the theoretical time complexity. With a constant time complexity for HashMap, the gradient should be theoretically zero. However, the gradient of best fit for HashMap in Graph 1 is non-zero (0.0009591). This non-zero gradient is caused by hash code computation and collisions. Each get() and put() method called for HashMap<K, V> must compute the key's hash code through hashCode(). If two keys hash to the same bucket, it will compare the keys using equals() as a key must be unique. Although the cost per operation is constant, the sum of these costs across input size n may grow proportionally to n.

Another issue is caused by the inability to represent the relationship for TreeMap and ConcurrentSkipListMap using a logarithmic function. This is due to the narrow input size range of only between 100 and 5000 words. On the test done, n only goes from 100 to 5000, so:

$$log(100) = 2,\ log(5000) \approx 3.70$$

This shows that Δy is only about 1.70 units, that is a 1.7 ms change in runtime, although the input size has increased by 50 times. The small value of Δy is dominated by the fixed-cost overheads, such as pointer traversal and tree rotation. This proves that $n = 5000$ is too small to see the asymptotical behavior of a logarithmic function, resulting in the poor statistical fit.

Extension

This study can be extended to evaluate more factors to better evaluate the true effectiveness of HashMap, TreeMap, and ConcurrentSkipListMap as the primary data structure for an extractive text summarizer. The test can be done using a much larger range of input sizes to be able to see the logarithmic behaviour of runtime for TreeMap and ConcurrentSkipListMap. In addition, it is also possible to implement the three Maps using other extractive summarizing algorithms to get a better picture of their effectiveness in different scenarios.

References

Ardash. (2024, October 5). DSA for developers: Why data structures matter in real projects. *Medium*. https://medium.com/@adarshgs.909/dsa-for-developers-why-data-structures-matter-in-real-projects-9b373f706a17

Barrett, E., Tereick, C., Killick, R., Mount, S., & Tratt, L. (2017, October 6). Virtual machine warmup blows hot and cold. *Programming Languages.* https://doi.org/10.48550/arXiv.1602.00602

Briggs, K., Zhou, B., & Dueck, G. (2016, March 29). Cold object identification in the Java virtual machine. *Software: Practice and Experience, 47*(1), 79-95. https://doi.org/10.1002/spe.2396

Cambridge Dictionary. (n.d.). Meaning of summary in English. *Cambridge Dictionary*. https://dictionary.cambridge.org/dictionary/english/summary

Demilie, W. (2022, September 13). Comparative analysis of automated text summarization techniques: The case of Ethiopian languages. *Wireless Communication and Mobile Computing*, *2022*, 13-15. https://doi.org/10.1155/2022/3282127

GeeksforGeeks. (2024, September 4). Skip list - Efficient search, insert, and delete in linked list. *GeeksforGeeks*. https://www.geeksforgeeks.org/skip-list/

GeeksforGeeks. (2025, April 23). Map interface in Java. *GeeksforGeeks*. https://www.geeksforgeeks.org/map-interface-java-examples/

Jain, S. (2025, February 7). Understanding TF-IDF (Term Frequency-Inverse Document Frequency). *GeeksforGeeks*. https://www.geeksforgeeks.org/understanding-tf-idf-term-frequency-inverse-document-frequency/

Li, B., Su, P., Chabbi, M., Jiao, S., & Liu, X. (2021, April 7). DJXPerf: Identifying memory inefficiencies via object-centric profiling for Java. *Performance*. https://doi.org/10.48550/arXiv.2104.03388

Li, R., Grodin, H., & Harper, R. (2023, September 23). A verified cost analysis of joinable Red-Black Trees. *Programming Languages*, *2*, 7. https://arxiv.org/pdf/2309.11056

Martinez, D. (2024, January 8). Java HashMap load factor. *Baeldung*. https://www.baeldung.com/java-hashmap-load-factor

Oracle. (n.d.). Class ConcurrentSkipListMap<K, V>. *Oracle*. https://docs.oracle.com/javase/8/docs/api/java/util/concurrent/ConcurrentSkipListMap.html

Pandey, K. (2023, July 13). Understanding HashMap data structure with examples. *Masai School.* https://www.masaischool.com/blog/understanding-hashmap-data-structure-with-examples/

Payong, A. (2024, October 25). Introduction to extractive and abstractive summarization techniques. *DigitalOcean*. https://www.digitalocean.com/community/tutorials/extractive-and-abstractive-summarization-techniques

Tuladhar, A. (2015, May 1). Concurrent non-blocking skip list using multi-word compare and swap operation. *UNLV Theses, Dissertations, Professional Papers, and Capstones Libraries*, 51. http://dx.doi.org/10.34917/7646077

Appendices

Appendix 1

```java
import java.util.ArrayList;
import java.util.HashSet;
import java.util.List;
import java.util.Set;

public class SentenceSplitter {
    private static final Set<String> HONORIFICS = new HashSet<>(Set.of(
        e1:"Mr.", e2:"Mrs.", e3:"Dr.", e4:"Ms.", e5:"Jr.", e6:"Sr.", e7:"Prof."
    ));

    // ALGORITHM: (1) skipping closing quotes, (2) checking uppercase next, (3) checking a fixed list of abbreviations
    public String[] splitIntoSentences(String text) {
        List<String> sentences = new ArrayList<>();
        if (text == null || text.isBlank()) {
            return new String[0];
        }

        // Clean bullet points
        text = text.replace(target:"•", replacement:"");

        StringBuilder current = new StringBuilder();
        int length = text.length();

        for (int i = 0; i < length; i++) {
            char c = text.charAt(i);
            current.append(c);
            // Check for a potential sentence-ending punctuation
            if (c == '.' || c == '!' || c == '?') {
                // Look backward to capture the word or token ending at this punctuation
                int k = i;
                while (k >= 0 && (Character.isLetter(text.charAt(k)) || text.charAt(k) == '.')) {
                    k--;
                }
                String token = text.substring(k + 1, i + 1);
                // If token is a known honorific, do not split
                if (HONORIFICS.contains(token)) {
                    continue;
                }
                // Move forward past any spaces and any closing quotes
                int j = i + 1;
                while (j < length && (
                        Character.isWhitespace(text.charAt(j))
                        || text.charAt(j) == '”'
                        || text.charAt(j) == '“'
                    )) {
                    j++;
                }
                // If j is at end or the next character is uppercase or an opening quote, split the sentence
                if (j >= length
                        || Character.isUpperCase(text.charAt(j))
                        || text.charAt(j) == '”'
                        || text.charAt(j) == '“') {
                    sentences.add(current.toString().trim());
                    current.setLength(newLength:0);
                    // Advance i to j-1 (so the for loop's i++ lands at j)
                    i = j - 1;
                }
            }
        }

        // Add any remaining text as the final sentence (if not empty)
        if (current.length() > 0) {
            sentences.add(current.toString().trim());
        }

        return sentences.toArray(new String[0]);
    }
}
```

Appendix 1. SentenceSplitter class.

Appendix 2

```java
import java.util.List;

public interface TextProcessor {
    public List<List<String>> preprocess(String text);
}
```

Appendix 2. TextProcessor interface.

Appendix 3

```java
import org.tartarus.snowball.ext.PorterStemmer;

public class SnowballStemmer implements Stemmer {
    private final PorterStemmer stemmer = new PorterStemmer();

    public String stem(String word) {
        stemmer.setCurrent(word);
        stemmer.stem();
        return stemmer.getCurrent();
    }
}
```

Appendix 3. SnowballStemmer class.

Appendix 4

```java
public interface Stemmer {
    public String stem(String word);
}
```

Appendix 4. Stemmer interface.

Appendix 5

```java
import java.util.ArrayList;
import java.util.HashSet;
import java.util.List;
import java.util.Map;
import java.util.Set;
import java.util.function.Supplier;
import java.util.stream.Collectors;
import java.util.stream.IntStream;

public class TFIDFSummarizer extends AbstractSummarizer{
    private final TextProcessor textProcessor;
    private SentenceSplitter splitter = new SentenceSplitter();

    public TFIDFSummarizer(Supplier<Map<String, Double>> doubleMapSupplier, Supplier<Map<String, Integer>> intMapSupplier,
    TextProcessor textProcessor) {
        super(doubleMapSupplier, intMapSupplier);
        this.textProcessor = textProcessor;
    }

    @Override
    public String summarize(String text, double ratio) {
        List<List<String>> preprocessed = textProcessor.preprocess(text);
        List<String> sentences = List.of(splitter.splitIntoSentences(text));
        int sentenceSize = sentences.size();

        Map<String, Integer> sentenceFrequency = intMapSupplier.get();
        Map<String, Double> idf = doubleMapSupplier.get();

        // Calculate IDF
        for (List<String> sentence : preprocessed) {
            Set<String> uniqueWords = new HashSet<>(sentence);
            for (String word : uniqueWords) {
                sentenceFrequency.put(word, sentenceFrequency.getOrDefault(word, defaultValue:0) + 1);
            }
        }
        for (Map.Entry<String, Integer> entry : sentenceFrequency.entrySet()) {
            idf.put(entry.getKey(), Math.log((double)sentenceSize / entry.getValue()));
        }

        // Calculate TF and sentence score
        List<Double> scoreList = new ArrayList<>();
        for (List<String> sentence : preprocessed) {
            Map<String, Double> tf = doubleMapSupplier.get();
            for (String word : sentence) {
                tf.put(word, tf.getOrDefault(word, defaultValue:0.0) + 1.0);
            }
            double score = 0.0;
            for (String word : sentence) {
                double tfValue = tf.get(word) / sentence.size();
                double idfValue = idf.getOrDefault(word, defaultValue:0.0);
                score += tfValue * idfValue;
            }
            scoreList.add(score);
        }

        int summarySize = Math.max(a:1, (int) Math.ceil((sentenceSize * ratio) / Math.pow(Math.log10(sentenceSize + 1), b:2.5)));
        List<Integer> rank = IntStream.range(startInclusive:0, scoreList.size()) // Generates stream from index 0 to size of scoreList
        .boxed() // Box primitive int into wrapper class Integer
        .sorted((i, j) -> Double.compare(scoreList.get(j), scoreList.get(i))) // Sort sentence indices by their score descending
        .limit(summarySize) // Take top summarySize indices
        .sorted() // Restore original sentence order (ascending)
        .collect(Collectors.toList());

        return rank.stream().map(sentences::get).collect(Collectors.joining(delimiter:" "));
    }
}
```

Appendix 5. TFIDFSummarizer class.

Appendix 6

```java
import java.util.function.Supplier;
import java.util.Map;

public abstract class AbstractSummarizer {
    protected final Supplier<Map<String, Double>> doubleMapSupplier;
    protected final Supplier<Map<String, Integer>> intMapSupplier;

    protected AbstractSummarizer(Supplier<Map<String, Double>> doubleMapSupplier, Supplier<Map<String, Integer>> intMapSupplier) {
        this.doubleMapSupplier = doubleMapSupplier;
        this.intMapSupplier = intMapSupplier;
    }

    public abstract String summarize(String text, double ratio);
}
```

Appendix 6. AbstractSummarizer abstract class.

Appendix 7

Input 1 (AI Generated)
Cybersecurity is essential in today's digital world. It protects systems, networks, and data from unauthorized access. Strong passwords help prevent hacking. Two-factor authentication adds another layer of security. Regular software updates fix known vulnerabilities. Phishing emails try to trick users into giving information. Never click suspicious links or attachments. Use antivirus software to detect threats. Public Wi-Fi can be risky without a VPN. Data breaches can expose sensitive personal data. Companies must train employees to recognize cyber threats. Cyberattacks can cause financial loss and damage reputations. Staying informed helps reduce risk. Everyone plays an important role in keeping information secure.
Output 1
Cybersecurity is essential in today's digital world. Never click suspicious links or attachments. Data breaches can expose sensitive personal data.

Appendix 7. Full text for input and output 1.

Appendix 8

Input 2 (AI Generated)
Artificial intelligence (AI) is revolutionizing the healthcare industry by enhancing efficiency, accuracy, and decision-making. AI refers to computer systems that simulate human intelligence to perform tasks such as learning, reasoning, and problem-solving. In healthcare, AI applications range from diagnostics to treatment planning, patient monitoring, and administrative support. One of the most prominent uses of AI in healthcare is in medical imaging. AI algorithms can analyze X-rays, CT scans, and MRIs to detect abnormalities such as tumors, fractures, and infections. These tools assist radiologists by identifying patterns that may be too subtle for the human eye. For example, AI systems have been shown to detect early signs of breast cancer in mammograms with higher accuracy than traditional methods. This can lead to earlier treatment and improved patient outcomes. Predictive analytics is another critical application. AI models analyze historical patient data, genetic information, and lifestyle factors to predict the likelihood of diseases such as diabetes, heart conditions, or cancer. This enables healthcare providers to take preventative measures before symptoms even appear. Predictive tools also help hospitals forecast patient admissions, optimize staffing, and manage resources efficiently. Natural language processing (NLP) allows AI to understand and interpret human language. In healthcare, NLP is used to extract useful information from unstructured data such as doctors' notes, discharge summaries, and medical literature. This helps streamline documentation, reduce administrative burden, and ensure that vital information is easily accessible. AI chatbots equipped with NLP can also assist patients by answering common questions and scheduling appointments. Drug discovery and development is a traditionally lengthy and expensive process. AI accelerates this by predicting how different molecules will interact with biological targets. Machine learning models can identify potential drug candidates faster and at a lower cost than traditional methods. During the COVID-19 pandemic, AI helped researchers analyze existing drugs for potential repurposing, significantly speeding up research timelines. AI-powered robots are also entering operating rooms. These robotic systems, guided by surgeons and enhanced by AI, can perform precise, minimally invasive procedures. Benefits include reduced recovery time, lower risk of infection, and greater surgical consistency. In addition, service robots are being used in hospitals to deliver medications, disinfect rooms, and assist elderly or immobile patients. Despite its promise, AI in healthcare raises several concerns. Data privacy and security are major issues, as AI systems require access to sensitive medical records. Ensuring that data is anonymized and protected is essential. Another challenge is bias in AI algorithms. If training data lacks diversity, the AI may produce inaccurate or unfair outcomes. Regulatory oversight and rigorous testing are necessary to maintain ethical standards. In conclusion, artificial intelligence has the potential to transform healthcare in meaningful ways. From improving diagnostics to personalizing treatment, AI enhances patient care while reducing costs and workload. As technology continues to advance, careful integration of AI—with ethical safeguards—will be crucial for a smarter, more responsive healthcare system. Continued investment in research, collaboration between clinicians and technologists, and public education about AI are all essential. Building trust determines their success in global health.
Output 2
Natural language processing (NLP) allows AI to understand and interpret human language. During the COVID-19 pandemic, AI helped researchers analyze existing drugs for potential repurposing, significantly speeding up research timelines. Benefits include reduced recovery time, lower risk of infection, and greater surgical consistency. Building trust determines their success in global health.

Appendix 8. Full text for input and output 2.

Appendix 9

Input 3 (AI Generated)
Renewable energy refers to sources of power that are naturally replenished on a human timescale, such as sunlight, wind, water flow, geothermal heat, and biomass. Unlike fossil fuels, which take millions of years to form and release carbon dioxide and other pollutants when burned, renewable energy sources are less likely to run out and tend to produce fewer greenhouse gases. The use of renewable energy has grown rapidly over the past few decades as concerns about climate change, energy security, and pollution have intensified. Advances in technology, along with supportive policies and falling costs, have enabled renewable energy to move from niche applications into mainstream power generation, heating, and transportation. Understanding how these sources work, their advantages, and the challenges they present is essential for appreciating their role in shaping a sustainable future. Solar energy harnesses sunlight through photovoltaic cells or solar thermal collectors. Photovoltaic cells convert sunlight directly into electricity by allowing photons to knock electrons free from atoms, creating an electric current. Solar thermal systems, by contrast, concentrate sunlight to heat a fluid that drives a turbine connected to an electrical generator. Both approaches have seen significant improvements in efficiency and cost reduction in recent years. Panels and collectors can be installed on rooftops, in large utility-scale solar farms, or integrated into building materials such as windows and facades. Solar energy's modular nature means that installations can range from small portable chargers to sprawling desert solar fields. Despite variability—solar panels generate most electricity during sunny daytime hours—batteries and grid management strategies can store surplus energy for use during cloudy periods or nighttime. Wind energy captures the kinetic energy of moving air and converts it into electricity using turbines. Modern wind turbines consist of three large blades mounted on a tower; as wind pushes these blades, they spin a shaft connected to a generator. Wind farms can be located onshore, in areas with steady wind patterns, or offshore, where winds tend to be stronger and more reliable. Offshore turbines, though more expensive to install and maintain, often generate electricity at higher capacity factors than their land-based counterparts. Advances in blade design, materials, and control systems have allowed turbines to become larger and more efficient, reducing cost per kilowatt-hour. Wind energy is variable—generation dips when wind speeds are low and can exceed demand during storms—so integration with storage systems and grid balancing mechanisms remains crucial for ensuring reliable power supply. Hydroelectric power captures the energy of moving water, typically by building dams on rivers. Water released from the reservoir flows through turbines that generate electricity. Large hydroelectric dams can produce vast amounts of power, supplying electricity to millions of people. Small-scale run-of-river projects divert a portion of river flow through turbines without requiring large reservoirs. Hydropower is flexible and can ramp up generation quickly to meet fluctuating demand, making it an effective partner to variable renewable sources like solar and wind. However, constructing dams can have significant environmental and social impacts: flooding large areas can displace communities and wildlife, alter sediment transport, and affect downstream ecosystems. Ensuring that hydroelectric projects are designed and managed with ecological considerations in mind is critical to minimizing these adverse effects. Geothermal energy utilizes heat from beneath the Earth's surface. In regions where geothermal reservoirs are close to the surface, wells can be drilled to access hot water or steam that drives turbines for electricity generation. Geothermal heat can also be used directly for heating buildings, greenhouses, and industrial processes. Enhanced geothermal systems, an emerging technology, create artificial reservoirs by fracturing hot rock and circulating fluid, potentially expanding geothermal resources beyond naturally occurring sites. Geothermal energy provides a stable, baseload power output that does not depend on weather or time of day. Its environmental footprint is relatively small compared to fossil fuels, but geothermal plants may release trace amounts of greenhouse gases from underground sources and require careful management of water resources and subsurface pressure to avoid seismic risks. Biomass energy derives from organic materials such as wood, agricultural residues, and dedicated energy crops. Biomass can be burned directly for heat, processed into biofuels like ethanol and biodiesel for transportation, or converted into syngas through gasification. Bioenergy can play a role in waste management by utilizing agricultural or forestry byproducts. Unlike fossil fuels, biomass is part of the contemporary carbon cycle; the carbon dioxide released during combustion is roughly balanced by

the carbon dioxide absorbed during plant growth, making it potentially carbon-neutral if sourced sustainably. However, large-scale biomass production can compete with food crops for land and water, and unsustainable harvesting can lead to deforestation and biodiversity loss. Careful planning and certification systems are needed to ensure that biomass use contributes positively to climate goals. Each renewable energy source offers unique benefits and challenges. Solar and wind are abundant and widely distributed, making them suitable for a diverse range of climates and geographies. Both have zero fuel costs and produce no direct air pollution. Yet they face intermittency—solar panels do not generate power at night or during severe cloud cover, and wind turbines depend on wind patterns that may fluctuate. Hydropower provides firm, dispatchable electricity but is limited by geography and can have significant ecological consequences. Geothermal energy offers reliable baseload power in regions with accessible resources but is more geographically constrained to tectonically active areas. Biomass energy is flexible and dispatchable but requires sustainable management of organic feedstocks to avoid negative environmental impacts. Cost trends have increasingly favored renewables. Over the past decade, levelized costs for solar photovoltaic and onshore wind have fallen by more than 80 percent, making them competitive or even cheaper than new fossil fuel plants in many regions. Factors contributing to this decline include improvements in technology, economies of scale, supply chain optimizations, and supportive governmental policies such as tax incentives, feed-in tariffs, and renewable portfolio standards. Investment in research and development continues to drive efficiency gains and cost reductions. As renewable energy becomes more affordable, developing countries in particular have opportunities to bypass conventional high-carbon infrastructure and transition directly.
Output 3
Understanding how these sources work, their advantages, and the challenges they present is essential for appreciating their role in shaping a sustainable future. Modern wind turbines consist of three large blades mounted on a tower; as wind pushes these blades, they spin a shaft connected to a generator. Unlike fossil fuels, biomass is part of the contemporary carbon cycle; the carbon dioxide released during combustion is roughly balanced by the carbon dioxide absorbed during plant growth, making it potentially carbon-neutral if sourced sustainably. Factors contributing to this decline include improvements in technology, economies of scale, supply chain optimizations, and supportive governmental policies such as tax incentives, feed-in tariffs, and renewable portfolio standards.

Appendix 9. Full text for input and output 3.

Appendix 10

Input 4 (AI Generated)
Social media has become an integral part of modern society, influencing how individuals communicate, share information, and perceive themselves and the world around them. Since the early 2000s, platforms such as Facebook, Twitter, Instagram, and TikTok have seen explosive growth in user bases, connecting billions of people globally. The rise of social media has transformed communication norms and introduced new dynamics in interpersonal relationships. While these platforms offer benefits such as convenience, connectivity, and real-time updates, they also present unique challenges related to mental well-being. As users engage with online content and social networks, they encounter a mixture of supportive communities, misinformation, and pressures to present idealized versions of their lives. The constant flow of notifications and the drive for likes or comments can create an environment that heightens anxiety and fuels compulsive checking behaviors. In response, researchers and policymakers have sought to understand the nuanced ways in which social media impacts mental health, aiming to identify both the risks and the protective factors involved. One of the earliest popular social media platforms, Facebook, launched in 2004, quickly became synonymous with online social networking. Its success paved the way for subsequent platforms emphasizing different forms of content, such as Twitter for short text updates and Instagram for visual storytelling. Over time, social media evolved to incorporate diverse features, including live streaming, interactive stories, ephemeral content, and algorithm-driven feeds. These features are designed to increase user engagement by tailoring content to individual preferences, but they can also create filter bubbles that reinforce existing beliefs and limit exposure to diverse perspectives. Additionally, algorithms prioritize sensational or emotionally charged content, which can amplify negative emotions and contribute to polarization within online communities. As platforms compete for user attention, they continuously innovate with features that encourage longer usage sessions, raising concerns about excessive screen time and its implications for sleep, attention spans, and overall mental health. Research on social media and mental health has expanded significantly in recent years. Studies have investigated various outcomes, including symptoms of depression, anxiety, loneliness, self-esteem, and body image concerns. Findings suggest that the relationship between social media use and mental health is complex and influenced by multiple factors such as frequency of use, type of activity, personality traits, and offline social support systems. For example, passive use, defined as scrolling through feeds without actively engaging, has been associated with increased feelings of envy and decreased life satisfaction. In contrast, active use—such as messaging friends or posting original content—sometimes correlates with improved social connectivity and well-being. However, even active engagement can be detrimental if it leads to overexposure to negative feedback or trolling. Moreover, individuals with preexisting vulnerabilities, such as low self-esteem or social anxiety, may experience more negative outcomes from particular online interactions. A meta-analysis of studies conducted between 2010 and 2020 found that heavy social media users were more likely to report symptoms of depression and anxiety compared to casual users. Another study focusing on adolescents revealed that those spending more than three hours per day on social media were at a higher risk for mental health problems. However, these studies also highlight that causality is difficult to establish; individuals with preexisting mental health issues may be drawn to social media as a means of distraction or social connection. Longitudinal research has attempted to disentangle these effects, with some evidence suggesting that reducing time spent on social media can lead to modest improvements in mood and life satisfaction. Nevertheless, the directionality of these relationships remains under investigation, emphasizing the need for careful interpretation of cross-sectional data. Self-esteem and body image are particularly vulnerable to social media influences. Platforms that prioritize visual content, such as Instagram and TikTok, often feature images of idealized beauty standards, carefully curated lifestyles, and edited photographs. Exposure to such content can lead to social comparison, where individuals gauge their own worth against the seemingly perfect lives of others. This comparison process can exacerbate feelings of inadequacy and contribute to negative body image, eating disorders, and disordered patterns of social media use. Adolescents and young adults, who are in critical stages of identity formation, may be especially susceptible to these pressures. Moreover, the portrayal of lifestyle products, fitness routines, and dietary supplements by influencers can encourage harmful behaviors

and unrealistic expectations regarding appearance. Another significant concern is the role of social media in fostering cyberbullying and harassment. Unlike traditional forms of bullying that occur in physical spaces, cyberbullying can take place at any time and reach a wide audience with minimal effort. Victims may experience emotional distress, social isolation, and in severe cases, suicidal ideation. Social media platforms have implemented guidelines and reporting mechanisms to address harassment, but the sheer volume of content and the anonymity afforded by online environments complicate moderation efforts. Moreover, hate speech and coordinated harassment campaigns can create hostile digital spaces, particularly for marginalized communities. These negative experiences can have lasting psychological effects, prompting users to withdraw from online engagement or develop heightened distrust toward digital communication channels. While much attention is given to potential harms, it is also important to recognize the positive aspects of social media for mental health. For individuals who are geographically isolated, have mobility limitations, or experience social anxiety, online platforms can offer opportunities to connect with like-minded people and form supportive communities. For example, online forums and groups dedicated to specific mental health conditions can provide a sense of belonging and validation. Some platforms facilitate live peer support sessions, virtual mental health workshops, and psychoeducational content created by licensed professionals. These resources can serve as valuable supplements to traditional therapy, particularly in areas where mental health services are scarce or difficult to access. In crisis situations, social media has also been used to mobilize rapid support, such as fundraising for medical expenses or organizing local volunteer efforts, demonstrating its potential as a tool for collective resilience. Social media also plays a role in shaping cultural conversations around mental health. High-profile individuals and celebrities sharing their own experiences with mental illness have helped normalize discussions and reduce stigma. Hashtags such as #MentalHealthAwareness and #EndTheStigma trend regularly, encouraging users to share their stories and resources. Campaigns during World Mental Health Day and Mental Health Awareness Month leverage social media's reach to disseminate information and promote preventive measures. Additionally, digital influencers focused on wellness, mindfulness, and self-care have gained significant followings, often collaborating with mental health professionals to develop content that educates audiences about coping strategies, stress management, and resilience-building. However, the quality and accuracy of mental health advice vary widely, underscoring the importance of critical evaluation by users. Privacy and data security are additional factors influencing the mental health impact of social media. Platforms collect vast amounts of personal data to tailor content and advertisements. While targeted ads can enhance user experience, they also raise concerns about surveillance and manipulation. Instances where personal information is leaked or misused can lead to anxiety and distrust. Understanding the algorithms that curate content and recognizing that user behavior is monitored can contribute to feelings of vulnerability and loss of autonomy. Data breaches at major social media companies have exposed sensitive information such as private messages, location data, and personal preferences, heightening concerns about digital safety. Educational efforts aimed at improving digital literacy and encouraging users to review privacy settings can mitigate some risks, but systemic changes in platform policies are also necessary to enhance transparency. Demographic factors play a significant role in how social media affects mental health. Adolescents and young adults are among the most active users and also the most susceptible to negative outcomes. Their developmental stage involves identity formation and social validation, making them more likely to engage in social comparison and be influenced by peer feedback. Women, particularly teenage girls, often report higher levels of social media-related anxiety and body dissatisfaction. Men may experience pressure to conform to ideals of masculinity, leading to different forms of social comparison and potential concealment of emotional vulnerabilities. Cultural background also influences how users interpret online interactions; individuals from collectivist societies may be more attuned to community feedback, whereas those from individualistic cultures might focus on personal achievement and self-expression. To mitigate the negative impacts of social media on mental health, several approaches can be implemented. First, individuals can adopt healthier usage patterns, such as setting time limits, disabling notifications, and engaging in regular digital detoxes. Mindful use involves being conscious of emotional responses while using platforms and pausing when feelings of distress arise. Second, education programs in schools can teach digital literacy, empathy, and coping skills. These programs should emphasize the importance of critical thinking, guiding young users to discern between trustworthy information and clickbait, and encouraging respectful online communication. Third, parents

and caregivers can model balanced media consumption and maintain open dialogues with children about their online experiences. Establishing family media plans that include agreed-upon screen time limits and designated device-free zones can foster healthier habits. Policy interventions also have a role to play. Governments and regulatory bodies can enforce stricter guidelines regarding data privacy, algorithmic transparency, and age-appropriate content. Legislation such as the General Data Protection Regulation (GDPR) in Europe has set a precedent for empowering users to control their data. Similar regulations in other regions can help ensure that platforms prioritize user well-being over profit. Additionally, funding for mental health research focused on digital environments can facilitate evidence-based policy decisions. Collaboration between regulatory agencies, academic institutions, and private sector stakeholders is essential to develop standards for responsible data collection and algorithmic accountability. Longitudinal studies are essential to gain a deeper understanding of long-term effects. Most existing research relies on cross-sectional data, which limits the ability to draw causal inferences. Tracking individuals over time can reveal how patterns of social media use influence mental health trajectories and identify critical periods where interventions may be most effective. Collaboration between academia, industry, and healthcare institutions can facilitate access to larger datasets while ensuring ethical standards are maintained. Technological advancements in real-time data analytics and wearable devices may enable researchers to monitor physiological indicators of stress and well-being, providing richer insights into how online interactions manifest in physical health outcomes. Future directions in social media and mental health research should also explore emerging technologies such as virtual reality (VR) and augmented reality (AR). While current platforms primarily involve two-dimensional content, immersive environments could intensify the emotional impact of online interactions. VR-based social spaces may offer novel therapeutic opportunities, such as exposure therapy for phobias or simulated social skills training. However, they also raise new ethical and psychological concerns, such as heightened susceptibility to addiction, blurred boundaries between virtual and real experiences, and increased vulnerability to manipulation. Researchers must examine how prolonged immersion in VR or AR social settings affects cognitive processes, emotional regulation, and identity development. Platform designers and developers are increasingly recognizing their responsibility to prioritize user well-being in the architecture of social media environments. By integrating mental health resources directly into the user interface, platforms can offer immediate support for users experiencing emotional distress. Features such as mood-tracking prompts, direct links to crisis helplines, and educational pop-ups about healthy online behavior can cultivate a safer environment. Community-driven moderation tools, where trusted volunteers help identify harmful content, can complement algorithmic filtering to reduce exposure to toxic material. Encouraging diverse representation among developers and content moderators can help identify and mitigate cultural biases in community guidelines. Building partnerships with mental health professionals ensures that resources are accurate and easily beneficial. In conclusion, social media's impact on mental health is multifaceted, encompassing both risks and benefits. The widespread adoption of digital platforms has transformed social dynamics, offering unprecedented opportunities for connection, support, and information sharing. At the same time, the design features that drive engagement can lead to negative outcomes such as depression, anxiety, and body image concerns. Addressing these challenges requires coordinated efforts from individuals, educators, policymakers, and platform providers. By promoting digital literacy, implementing mindful usage strategies, enforcing data transparency regulations, and investing in longitudinal research, society can harness the positive potential of social media while minimizing harm.
Output 4
For example, passive use, defined as scrolling through feeds without actively engaging, has been associated with increased feelings of envy and decreased life satisfaction. Nevertheless, the directionality of these relationships remains under investigation, emphasizing the need for careful interpretation of cross-sectional data. This comparison process can exacerbate feelings of inadequacy and contribute to negative body image, eating disorders, and disordered patterns of social media use. Moreover, the portrayal of lifestyle products, fitness routines, and dietary supplements by influencers can encourage harmful behaviors and unrealistic expectations regarding appearance. Legislation such as the General Data Protection Regulation (GDPR) in Europe has set a precedent for empowering

users to control their data. Most existing research relies on cross-sectional data, which limits the ability to draw causal inferences.

Appendix 10. Full text for input and output 4.

Appendix 11

Input 5 (AI Generated)
Coral reefs are among the most diverse and productive ecosystems on Earth. Located in shallow tropical and subtropical waters, they harbor a rich array of species adapted to live in complex three-dimensional structures created by reef-building corals. These structures form a living framework that supports not only colorful invertebrates and fish, but also numerous algae, sponges, crustaceans, mollusks, echinoderms, and microorganisms. Reefs provide essential ecosystem services including coastal protection, nutrient cycling, and fisheries that support food security for millions of people. Their beauty and biodiversity attract tourism, fueling local economies and cultural identities. Despite occupying less than one percent of the ocean floor, coral reefs support about twenty-five percent of marine species. Understanding the ecology and behavior of coral reefs is vital for addressing the threats they face from climate change, pollution, and overexploitation. Coral reefs are primarily constructed by scleractinian corals, sometimes called stony corals. These marine invertebrates belong to the phylum Cnidaria and are related to jellyfish and sea anemones. Reef-building corals possess specialized cells called cnidocytes that contain nematocysts—stinging organelles used to capture small plankton and deter predators. Each coral polyp secretes calcium carbonate to form a protective exoskeleton. Over time, colonies of polyps divide asexually to expand the skeleton and build massive reef structures that can persist for centuries. Coral colonies grow slowly, often only a few centimeters per year in favorable conditions, yet by accreting skeletons over millennia, they create expansive reef platforms that can stretch for hundreds of kilometers. The relationship between corals and zooxanthellae is mutualistic and essential for reef health. Through photosynthesis, zooxanthellae convert sunlight, carbon dioxide, and nutrients into energy, providing up to ninety percent of the energy requirements of the coral host. This energy supports coral growth, reproduction, and calcification processes. In turn, corals supply zooxanthellae with carbon dioxide, nitrogen, and phosphorus derived from their metabolic waste. This symbiosis allows corals to thrive in nutrient-poor tropical waters, where photosynthetic productivity is otherwise limited. During daylight hours, coral polyps extend their tentacles to capture plankton, complementing energy gains from zooxanthellae. Coral reef ecosystems exhibit remarkable biodiversity that rivals tropical rainforests in their complexity. Fish communities on reefs include colorful angelfish, butterflyfish, parrotfish, surgeonfish, wrasses, and groupers among many others. Many fish species rely on the reef for food, shelter, and breeding grounds. Herbivorous fish such as parrotfish and surgeonfish graze on algae, preventing algal overgrowth that can outcompete corals for space and light. Predatory fish, such as groupers and sharks, regulate lower trophic levels, maintaining ecological balance. Invertebrates such as spiny lobsters, shrimps, sea cucumbers, and starfish perform important ecological roles, including detritus consumption and bioerosion, which helps recycle calcium carbonate and shape reef structure. Reef-building corals form complex three-dimensional habitats that house cryptic species and juvenile fish that find refuge within crevices and under ledges. The structural complexity of reefs influences species diversity, abundance, and interactions. Some fish species are obligate reef dwellers, relying on specific coral species for camouflage or spawning. For example, the long-spined sea urchin often inhabits crevices within branching corals to avoid predation and to graze on nearby algae and detritus. Meanwhile, small crustaceans, such as cleaner shrimps and gobies, establish symbiotic relationships with larger fish, providing cleaning services by removing parasites in exchange for shelter. Reproductive strategies of corals and associated reef organisms are equally diverse. Many coral species reproduce sexually through mass spawning events synchronized by environmental cues such as lunar cycles, water temperature, and photoperiod. During these events, corals release gametes—eggs and sperm—into

the water column simultaneously. Fertilization occurs externally, and resulting planula larvae drift as plankton before settling onto suitable substrate to form new colonies. Genetic mixing during spawning events increases genetic diversity within coral populations, potentially enhancing resilience to stressors. Some coral species are hermaphroditic, producing both eggs and sperm, while others have separate male and female colonies. Fish and invertebrate species frequently exhibit their own specialized reproductive behaviors. Many tropical fish engage in spawning aggregations where large numbers gather in specific locations for synchronized breeding. These aggregations may occur seasonally and are often targeted by fisheries, making them vulnerable to overexploitation. Certain invertebrates such as sea turtles and coconut crabs undertake long migrations to reach breeding sites—sea turtles return to the same beaches where they hatched decades earlier to lay eggs, illustrating natal homing behavior. Broadcast-spawning invertebrates like sea urchins release gametes into the water, relying on currents for fertilization. Nutrient cycles on coral reefs are intricate, involving exchanges among organisms and between the reef and surrounding waters. Primary production by photosynthetic organisms—zooxanthellae within corals, microalgae, cyanobacteria, and macroalgae—is the foundation of energy flow in reef ecosystems. Through photosynthesis, these primary producers convert inorganic carbon into organic matter. A portion is consumed by herbivores, while excess organic matter fuels detrital pathways. Detritivores such as sea cucumbers and certain crustaceans break down organic debris, releasing nutrients that support further primary production. Nutrients can also be exported from reefs to adjacent ecosystems during high tides and wave action. Coral reefs also interact with surrounding ecosystems, including mangroves, seagrass beds, and open ocean waters. Many juvenile fish species use mangrove nurseries before moving to coral reefs as they mature. Mangroves and seagrasses filter sediments and pollutants from terrestrial runoff, protecting reefs from turbidity and nutrient overload. Larval connectivity between habitats supports genetic exchange and resilience across reef networks. Conversely, coral reefs provide habitat and food sources for a variety of migratory species such as rays and groupers that visit seagrass meadows and mangrove lagoons at different life stages. Coral reefs face numerous threats from human activities and environmental change. Climate change poses the most significant long-term threat through elevated sea surface temperatures causing widespread bleaching events. When water temperatures rise by as little as one to two degrees Celsius above local seasonal maxima, corals may expel their symbiotic zooxanthellae. While some corals can recover if temperatures return to normal quickly, prolonged thermal stress leads to mortality. In 1998, a global bleaching event killed nearly sixteen percent of the world's reefs. Subsequent bleaching incidents in 2010, 2016, and 2017 further depleted coral cover in many regions. Coral species differ in bleaching sensitivity; some massive corals show greater tolerance, while branching corals often bleach more readily. Local stressors such as overfishing, destructive fishing practices, and coastal development degrade reef habitats. Overfishing of herbivorous fish like parrotfish and surgeonfish removes critical grazers that control algal growth, allowing macroalgae to overgrow and outcompete corals. Destructive practices such as cyanide fishing and blast fishing physically damage reef structures and kill non-target species. Coastal development that leads to dredging, seawall construction, and land reclamation can increase sedimentation, reducing light available for photosynthesis. Pollutants from agriculture and industry introduce excess nutrients and toxins that impair coral health and contribute to disease outbreaks. Marine pollution also includes plastic debris, which accumulates on reefs and entangles or is ingested by marine life. Microplastics can adhere to coral surfaces, reducing light penetration and impairing feeding. Coral disease outbreaks, such as white band disease, black band disease, and stony coral tissue loss disease, have increased in frequency and severity, further stressing reef communities. Disease transmission is often facilitated by waterborne pathogens and vectors such as butterflyfish that feed on infected tissues. Crown-of-thorns starfish are native predators of corals, but population outbreaks driven by elevated nutrient levels can decimate reef areas in the Pacific. Conservation efforts to protect coral reefs operate at local, regional, and global scales. Marine protected areas (MPAs) that restrict fishing and other damaging activities can help maintain fish populations, restore ecological balance, and enhance reef resilience. Well-enforced MPAs have demonstrated increased coral cover, higher abundances of herbivorous fish, and improved recovery from disturbances. Community-based management, where local stakeholders are involved in decision-making and enforcement, often leads to greater compliance and long-term success. In places like Fiji and the Philippines, locally managed marine areas combine traditional practices with modern science to sustain livelihoods and biodiversity.

Global efforts to address climate change, such as reducing greenhouse gas emissions, are essential for the long-term survival of coral reefs. International agreements and policies targeting carbon reduction, coupled with local efforts to improve water quality and reduce overfishing, can synergistically enhance reef resilience. Initiatives like the International Coral Reef Initiative, the Coral Reef Alliance, and the Global Coral Reef Monitoring Network facilitate collaboration among scientists, governments, NGOs, and local communities. Through sharing data, best practices, and resources, stakeholders can implement adaptive management strategies that respond to emerging threats and prioritize high-value reef areas. Future research on coral reefs focuses on improving understanding of reef resilience mechanisms, particularly how some corals adapt to thermal stress or recover from bleaching. Studies on coral microbiomes explore how associated bacteria and viruses influence coral health, disease resistance, and stress tolerance. Innovative monitoring technologies such as autonomous underwater vehicles, remote sensing, and molecular tools enable rapid assessment of reef conditions at broader spatial and temporal scales. For instance, environmental DNA (eDNA) sampling can detect the presence of rare or cryptic species without direct observation, expanding our knowledge of biodiversity on reefs. Coral reefs have existed for millions of years, with fossil evidence indicating that reef ecosystems flourished even during periods of significant climatic fluctuations. Ancient reef deposits built by long-extinct organisms like rudist bivalves and stromatoporoids formed massive limestone structures during the Cretaceous and Paleozoic eras. Modern reef-building corals began proliferating during the Eocene epoch, about fifty million years ago. Over geological time, sea-level changes influenced the location and extent of reef habitats, with ice ages and interglacial periods causing reefs to grow or die back. Studying fossilized reefs provides insights into how reef communities responded to past environmental changes, informing predictions for future resilience. The complex three-dimensional architecture of reefs is a product of both biological and physical processes. Wave action, tides, and currents interact with coral skeletons to shape reef crests, fore reefs, and back reefs into distinct geomorphological zones. Fringing reefs grow directly from continental landmasses, forming narrow bands along coastlines. Barrier reefs develop parallel to shore but are separated from land by deeper lagoons. Coral atolls form circular or horseshoe-shaped reefs surrounding central lagoons, often atop submerged volcanic islands. The structure and orientation of these different reef types influence their susceptibility to wave damage and bleaching. Reef restoration initiatives have gained momentum in recent decades to combat the decline of degraded reefs. One common approach is coral gardening, which involves the collection of coral fragments from healthy donor colonies, their cultivation in underwater nurseries, and eventual transplantation onto damaged reefs. Nurseries may be constructed using floating rope structures or fixed frames, where fragments grow until reaching a suitable size for outplanting. Successful restoration projects require careful selection of resilient coral species, appropriate site selection, and ongoing monitoring to evaluate survival rates and ecological integration. In addition to coral outplanting, some initiatives are testing sediment stabilization techniques to reduce smothering of juvenile corals. Technological advances have transformed coral reef research and monitoring. Autonomous underwater vehicles equipped with cameras and sensors can map reef terrain at high resolution, capturing three-dimensional models of reef complexity. Drones, flown over narrow reef lagoons or shallow coastal areas, provide aerial imagery that helps estimate coral cover, detect bleaching events, and monitor human impacts. Molecular tools such as environmental DNA (eDNA) sampling allow researchers to detect the presence of thousands of species in water samples, including rare or cryptic organisms. Coupling eDNA with traditional survey methods enhances our understanding of reef biodiversity and ecosystem health. Economic valuation of coral reef ecosystem services highlights their importance for coastal communities and national economies. Reefs generate revenue through tourism, fisheries, and coastal protection. In many tropical nations, diving and snorkeling tourism centered on reef experiences contribute significantly to gross domestic product. Hotels, dive operators, and tour guides rely on healthy reefs to attract visitors. Recreational fishing for species like snapper, grouper, and barracuda depends on productive reef habitats that support adult populations. Commercial fisheries targeting lobster, crab, and ornamental species also derive benefits from reefs. Estimating the market value of these services can incentivize governments to invest in reef conservation and recognize reefs as natural capital assets. Social and cultural dimensions of coral reefs are profound, with many indigenous and coastal communities maintaining spiritual and traditional connections to reef environments. Folklore, rituals, and customary marine tenure systems have governed reef resource use for generations. Traditional ecological knowledge, accumulated through centuries of interaction with reefs, offers valuable insights into sustainable harvesting practices, seasonal patterns of fish migrations, and the interdependence of reef species. Incorporating this knowledge into management plans respects cultural heritage and enhances conservation outcomes. In the Pacific Islands, for instance, customary reef closure periods during spawning seasons help maintain fish stocks and food security for communities. Tourism can generate both positive and negative impacts on reefs. While ecotourism can provide economic incentives for conservation, unchecked visitation can cause physical damage through anchor drops, trampling, and coral breakage by snorkelers and divers. Education programs that train dive operators to follow best practices, such as using mooring buoys instead of anchors, conducting briefings on reef etiquette, and limiting group sizes, help reduce tourist-related stress. Certification programs highlighting operators that adhere to reef-friendly practices encourage responsible tourism and promote reef awareness among visitors. Coral reefs also face emerging threats from ocean acidification, a consequence of elevated atmospheric carbon dioxide dissolving into seawater and lowering pH levels. Acidified waters reduce the availability of carbonate ions, which corals need to build calcium carbonate skeletons. Over long periods, acidification can slow coral growth, weaken existing structures, and impair reproduction. Combined with warming seas, acidification poses a “double jeopardy” for coral calcifiers. Experiments on coral recruitment under high-CO_2 conditions indicate reduced settlement success and skeletal density. Scientists are investigating whether local management actions such as reducing pollution can moderate acidification impacts at community scales. Another layer of reef stress emerges from pollutant runoff, particularly from agricultural fertilizers and pesticides. Nutrient enrichment can fuel algal blooms, which degrade water quality and smother corals by blocking sunlight. Pesticides such as diuron and atrazine, frequently used in agriculture, can impair coral photosynthesis and increase susceptibility to disease. Urban runoff, containing sewage and industrial effluents, introduces pathogens and heavy metals into reef waters. Wastewater treatment and sustainable agricultural practices are essential to reduce pollutant loads. In regions like Southeast Asia, where rapid coastal development has increased pollution, implementing buffer zones and constructed wetlands can filter contaminants before they reach reefs. Of particular concern are invasive species that alter reef community composition. The Indo-Pacific lionfish, introduced into the Atlantic, has rapidly spread along Caribbean reefs, preying on native fish and outcompeting local predators. Lionfish can decimate juvenile fish populations, reducing ecological balance and impairing coral recovery. Removal efforts, often led by volunteers, involve organized derbies and spearfishing campaigns to reduce lionfish densities. Similar threats include invasive algae such as caulerpa and kudzu-like species that can overgrow coral surfaces and inhibit settlement of coral larvae. Early detection and rapid response are critical for managing invasive species on reefs. Coral reef connectivity across geographical regions influences recovery potential after disturbances. Larval dispersal, mediated by ocean currents, shapes metapopulation dynamics and genetic exchange. Reefs that serve as larval sources for down-current reefs are critical nodes in network models. Protecting source reefs enhances resilience at landscape scales. Studies using oceanographic modeling combined with larval behavior experiments help identify corridors and barriers to connectivity. For instance, the coral larvae of some Acropora species exhibit vertical migration behavior that influences settlement patterns. Understanding these dynamics informs the design of networks of MPAs that support connectivity corridors. Restoration of reef-associated fisheries often incorporates community-based initiatives such as fishery cooperatives and no-take zones. By regulating fishing gear, seasonal closures, and catch limits, these programs aim to rebuild fish stocks and enhance food security for local livelihoods. Successes in reef fish recovery depend on enforcement of regulations, monitoring of critical fish populations, and adaptive management that responds to ecological feedback. In some Southeast Asian countries, fishers use traditional taboos to prevent fishing during spawning seasons, demonstrating how cultural norms can complement scientific approaches to reef management. Global collaborative research efforts, such as the International Year of the Reef and the Global Coral Reef Monitoring Network, aggregate observations from thousands of sites worldwide. Standardized monitoring protocols, covering metrics such as coral cover, fish abundance, and disease prevalence, enable comparisons across regions and time periods. Data from citizen science platforms, where divers and snorkelers contribute sightings through mobile applications, expand monitoring capacity. These datasets help detect early signs of reef decline and inform rapid response strategies. For example, data showing declines in herbivorous fish led to the establishment of targeted closures in parts of the Caribbean. Educational programs targeting youth, such as reef camps, school curricula, and hands-on field experiences, foster the next generation of marine stewards. By engaging students in reef surveys, species identification, and hands-on restoration activities, educators build awareness and cultivate conservation ethics. Outreach materials, including documentaries, interactive exhibits, and virtual reef tours, bring the underwater world to audiences far removed from tropical coastlines. In an increasingly connected world, these tools help forge emotional connections between people and reefs, inspiring broader support for reef conservation. Innovative financing mechanisms support reef conservation and restoration. Blue carbon initiatives that value coastal ecosystems' carbon sequestration potential often include mangroves, seagrasses, and salt marshes. Although coral reefs themselves do not sequester carbon in the same way, they rely on adjacent habitats that do. By bundling conservation of interconnected coastal ecosystems, blue carbon projects can unlock funding for integrated management. Additionally, biodiversity offsets and environmental impact bonds channel private investment toward reef protection, linking investor returns to ecological outcomes such as increased coral cover or fish biomass. Emerging coral farming techniques leverage genetic insights to cultivate thermally tolerant genotypes. Assisted evolution experiments involve exposing corals to elevated temperatures to select for resilient individuals. Selective breeding among diverse coral strains may produce offspring with enhanced stress tolerance, potentially seeding future populations better adapted to warming seas. Microbiome manipulation, where corals are inoculated with beneficial bacteria that boost immune function or mitigate bleaching, represents another frontier. While field trials show promise, scaling these interventions requires careful consideration of ecological risks and genetic diversity to avoid unintended consequences. Indigenous and local communities often hold generations of knowledge regarding reef dynamics, seasonal cues, and species interactions. Incorporating traditional ecological knowledge into scientific research enriches understanding of long-term ecological changes. For example, ancient fishing taboos that restrict harvest of certain species during spawning seasons provide insight into sustainable management. Community mapping exercises, participatory workshops, and oral history documentation integrate local perspectives into conservation planning. When communities take ownership of reef protection efforts, outcomes are more durable, as local stakeholders perceive the benefits of healthy reefs for their livelihoods and cultural heritage. Regional case studies illustrate varied trajectories of reef decline and recovery. The Great Barrier Reef off Australia has experienced multiple mass bleaching events over recent decades. Combining monitoring data with remote sensing, scientists track bleaching intensity, coral mortality, and recovery rates. Research shows that reefs with higher connectivity and cooler upwelling zones exhibit faster recovery, guiding targeted conservation efforts. In contrast, Caribbean reefs have faced compounded challenges from overfishing, pollution, and disease outbreaks, resulting in declines in stony coral cover from fifty percent in the 1970s to less than fifteen percent today. Restoration projects in Jamaica and Bonaire demonstrate that localized interventions—such as creating artificial reef structures and reintroducing key herbivorous species—can catalyze small-scale recoveries even when broader conditions remain challenging. Innovative policies such as reef insurance, where governments and businesses purchase insurance policies against coral bleaching events, represent novel approaches to risk management. When triggered by predefined thresholds—such as extended sea surface temperature anomalies—insurance payouts support emergency relief, rapid response restoration, and community adaptation programs. For example, Caribbean nations have explored parametric insurance instruments that provide swift funding after severe bleaching, allowing for immediate deployment of restoration teams, purchase of food and water during tourism declines, and support for affected fishers. While premiums can be high, pilot programs demonstrate the value of rapid response in mitigating socioeconomic losses. Coral reef eco-entrepreneurship fosters economic opportunities that align with conservation. Local enterprises focusing on sustainable seafood production, eco-friendly diving tours, and handcrafts made from sustainably sourced materials diversify income streams. By providing economic incentives that depend on healthy reefs, communities gain vested interests in long-term conservation. Certification programs, such as the Marine Stewardship Council, validate sustainable fishing practices, allowing products to fetch premium prices in global markets. Partnerships between businesses and conservation organizations create platforms for public-private collaborations that support reef-friendly innovations. Interdisciplinary research that bridges ecology, economics, sociology, and technology is crucial for holistic reef management. Social scientists analyze governance structures, stakeholder incentives, and cultural values to develop policies that accommodate diverse perspectives. Economists quantify ecosystem services and investigate market-based conservation approaches. Engineers design pollution control systems and reef-friendly coastal infrastructure. Biologists and ecologists track species interactions, assess community resilience, and refine restoration protocols. By integrating insights from multiple disciplines, managers can craft adaptive strategies that address the root causes of reef decline while supporting sustainable development. Education and outreach remain fundamental to building global reef champions. Documentaries like Chasing Coral and Coral Reef Adventures capture stunning footage of reef beauty while highlighting urgent conservation needs. Virtual reality experiences, such as 360-degree underwater tours, allow viewers to immerse themselves in reef environments, fostering empathy and appreciation. Citizen science platforms, including Reef Check and iNaturalist, enable reef enthusiasts to contribute observational data to scientists, bridging the gap between research and public engagement. Public lectures, museum exhibits, and art installations that feature reef themes broaden reach and inspire action beyond scientific circles. Climate change mitigation and adaptation strategies are essential at multiple scales. Locally, managers can implement early warning systems that use satellite data to predict bleaching events, triggering temporary closures of reef areas to minimize additional stress from tourism. Water quality improvement projects that reduce agricultural runoff and wastewater discharge enhance coral resilience by preventing compounding stressors. Restoration efforts focusing on hybrid coral genotypes and resilient symbionts provide short-term gains but cannot replace long-term climate stability. As climate impacts intensify, planning for potential relocations of human communities away from low-lying coastal zones may become necessary. In summary, coral reefs represent dynamic, complex ecosystems that support extraordinary biodiversity and provide critical services to human societies. Their ecological success is built upon intricate symbiotic relationships, efficient nutrient cycling, and structural complexity. Yet these same attributes render reefs vulnerable to a suite of global and local stressors. Understanding the ecology and behavior of coral reef organisms—from the cellular mechanisms of coral–algal symbiosis to the population-level dynamics of fish communities—provides the foundation for effective conservation strategies. Innovative restoration techniques, integrated management approaches, and community engagement offer hope for sustaining reefs in the face of change. However, the ultimate determinant of coral reef futures will remain the global trajectory of greenhouse gas emissions and humanity's collective willingness to safeguard these irreplaceable ecosystems. To complement these recovery and protection strategies, international cooperation is crucial. Coral reef conservation cannot succeed in isolation, as ocean currents carry coral larvae across political boundaries. Regional alliances such as the Coral Triangle Initiative bring together nations to share resources, scientific data, and management practices. By harmonizing policies on coastal development, fishing regulations, and pollution control, neighboring countries can reduce transboundary threats and enhance overall reef resilience. Innovative financing mechanisms are equally essential. Payment for ecosystem services programs can provide incentives for protecting reefs. For instance, coastal landowners might receive compensation for maintaining mangrove buffers that filter sediments and enhance reef water quality. Conservation agreements between governments and private stakeholders may include performance-based payments tied to measurable improvements such as increased coral cover or fish biomass. Ecotourism levies, where a small fee is collected from visitors who dive or snorkel, can be directed to local reef management funds, supporting patrols, monitoring, and education campaigns. Microfinance schemes that offer low-interest loans for reef-friendly businesses—such as eco-lodges or sustainable fisheries—enable communities to develop livelihoods without degrading reef ecosystems. Education remains a cornerstone for long-term resilience. Incorporating reef ecology into school curricula fosters appreciation from an early age. Interactive modules that include snorkeling field trips, water testing, and coral identification engage students in hands-on learning. University courses that integrate reef science with policy, economics, and social dimensions prepare future leaders to tackle multifaceted challenges. Public awareness campaigns through social media platforms, television documentaries, and community events highlight the interconnectedness of reefs and human well-being, encouraging individuals to adopt reef-friendly practices. Climate change education tailored for coastal populations provides guidance on adaptation strategies. Workshops that demonstrate how to construct raised garden beds, rainwater harvesting systems, and safe evacuation routes enhance community preparedness for extreme weather events like cyclones. Coastal zoning planning based on sea-level rise projections safeguards critical infrastructure and minimizes habitat loss. In the long term, integrated coastal zone management plans that consider ecosystem interactions—such as how wetlands, mangroves, and reefs function collectively to buffer storm surge—can guide sustainable development. Technological innovations offer promising avenues for reef conservation. Satellite remote sensing tools provide real-time data on sea surface temperatures, allowing early detection of thermal anomalies. Combined with in situ temperature loggers and mobile apps for citizen reporting, managers can anticipate bleaching alerts and close recreational activities to reduce stress. Acoustical monitoring using hydrophones records reef soundscapes; healthy reefs exhibit complex acoustic signatures from fish and invertebrate spawning calls, while degraded reefs tend to be quieter. By correlating sound levels with fish abundance and coral cover, researchers develop non-invasive monitoring methods that cover vast reef areas efficiently. Genomic and bioinformatics tools are rapidly advancing our understanding of coral biology. Whole-genome sequencing of multiple coral species reveals genetic markers associated with resilience traits, such as heat shock proteins and antioxidant pathways. Researchers can use these data to develop genomic selection protocols for breeding programs. CRISPR gene editing has been demonstrated in coral larvae under laboratory conditions, opening possibilities for enhancing stress tolerance. Ethical considerations and potential ecological risks must be carefully evaluated before field application, but the promise of directed evolution to produce robust coral strains drives ongoing research. The concept of assisted migration also garners attention. Translocating corals from thermally resilient populations to more vulnerable reefs may bolster resistance to warming events. Field trials in the Red Sea, where corals experience naturally high temperatures, aim to identify candidate genotypes for transplant. However, success depends on factors such as local water quality, disease prevalence, and compatibility with native reef communities. Careful experimental designs with control plots and long-term monitoring assess survival rates, genetic diversity, and ecological impacts, informing best practices for assisted migration. Urban coastal areas present unique challenges and opportunities for reef conservation. Coastal cities often generate pollution but also concentrate human resources and funding. Green infrastructure initiatives such as living shorelines, artificial wetlands, and seagrass restoration can filter urban runoff before it reaches reefs. In situ bioremediation projects explore the use of specific microbial consortia to degrade pollutants and reduce nutrient loads. Urban reef projects sometimes partner with aquariums, universities, and volunteer organizations to develop micro-reef installations in harbors. These mini-reefs, constructed from bio-enhanced concrete modules, support biodiversity and serve as educational platforms that connect urban populations to marine ecosystems. Coral reef restoration's scalability remains a critical issue. While small-scale nurseries can produce hundreds or thousands of coral fragments annually, global reef loss occurs at rates that far exceed these efforts. To address scale, researchers are exploring automated microfragmentation techniques where robots cut coral tissue into uniform pieces that grow rapidly. Modular floating nurseries that optimize light exposure and water flow can accelerate coral growth. Partnerships with the private sector, such as technology companies and shipping firms, enable the development of automated cartilaginous reef fabrication—large structures pre-seeded with coral fragments that can be towed to degraded sites and deployed en masse. Addressing knowledge gaps in coral taxonomy helps refine conservation strategies. Advances in taxonomy using morphological and genetic data clarify species boundaries and cryptic diversity within coral genera. Accurate species identification underpins assessments of biodiversity, resilience potential, and restoration priorities. Databases such as the World Register of Marine Species (WoRMS) and the Global Coral Trait Database provide standardized information on species distributions, life history traits, and ecological roles. Citizen science platforms that allow divers to photograph and submit coral sightings help update distribution records, flag invasive species, and monitor changes over time. Legal frameworks at international, national, and local levels shape reef governance. The Convention on Biological Diversity (CBD) and the United Nations Sustainable Development Goals (SDGs) include targets relevant to coral reef conservation, such as protecting at least ten percent of coastal and marine areas by 2020 under Aichi Biodiversity Target 11. While some countries meet or exceed this target, others struggle due to enforcement challenges and competing economic demands. Strengthening compliance mechanisms—such as requiring Environmental Impact Assessments for coastal developments and mandating regular monitoring—improves accountability. Supporting community-led co-management agreements, where local fishers and government agencies share authority, enhances stewardship and reduces conflicts. The economic valuation of ecosystem services underscores reefs' importance in global markets. Recent studies estimate that coral reefs contribute hundreds of billions of dollars annually through tourism, fisheries, and coastline protection. Cost-benefit analyses demonstrate that every dollar

invested in reef conservation can yield multiple dollars in ecosystem services, emphasizing the economic logic of proactive investment. Financial institutions and insurers increasingly incorporate nature-based solutions in their risk assessments. For instance, companies may invest in reef restoration to reduce insurance premiums for coastal infrastructure by lowering the risk of storm damage. Cross-disciplinary partnerships are essential to foster innovation. Collaboration among marine biologists, economists, social scientists, and engineers yields integrated solutions. For example, developing solar-powered desalination plants that use reef algae biomass as feedstock for biofuel production can address freshwater scarcity and provide clean energy, reducing pressure on reef ecosystems. Eco-friendly aquaculture systems that culture bivalves on reef-adjacent structures improve water quality by filtering particulates, supporting both restoration and local livelihoods. Architects and urban planners incorporate reef data into coastal zoning to minimize disruption to key ecological corridors. Communicating uncertainty is paramount in reef science to maintain credibility and trust.
Output 5
Fertilization occurs externally, and resulting planula larvae drift as plankton before settling onto suitable substrate to form new colonies. Certain invertebrates such as sea turtles and coconut crabs undertake long migrations to reach breeding sites—sea turtles return to the same beaches where they hatched decades earlier to lay eggs, illustrating natal homing behavior. A portion is consumed by herbivores, while excess organic matter fuels detrital pathways. Coral disease outbreaks, such as white band disease, black band disease, and stony coral tissue loss disease, have increased in frequency and severity, further stressing reef communities. Although coral reefs themselves do not sequester carbon in the same way, they rely on adjacent habitats that do. Local enterprises focusing on sustainable seafood production, eco-friendly diving tours, and handcrafts made from sustainably sourced materials diversify income streams. The concept of assisted migration also garners attention. Databases such as the World Register of Marine Species (WoRMS) and the Global Coral Trait Database provide standardized information on species distributions, life history traits, and ecological roles. Cost-benefit analyses demonstrate that every dollar invested in reef conservation can yield multiple dollars in ecosystem services, emphasizing the economic logic of proactive investment.

Appendix 11. Full text for input and output 5.

Appendix 12 - Program Manual

1. Download Java SDK 24 (if not yet downloaded).
2. Change the path of Java SDK in launch.json and settings.json.
3. Open Main.java and run main().

Appendix 13

The GitHub repository for this program can be found here.

Appendix 14

The presentation for this project can be found here.