1. **System configuration:**

CPU used - Intel Core i7-6600U

Clock rate - 2.81 GHz

RAM - 16.0 GB

Cache size - 4MB Intel Smart Cache

1. **Quality:**
   1. Create a table that shows for each input string, the length of its corresponding fingerprint (task 1 output). As for the fingerprint themselves, please output them to a file.

|  |  |  |
| --- | --- | --- |
| **String ID** | **Strain Name** | **Fingerprint Length** |
| 1 | COVID-19 China |  |
| 2 | COVID-19 USA |  |
| 3 | COVID-19 Australia |  |
| 4 | COVID-19 India |  |
| 5 | COVID-19 Brazil |  |
| 6 | SARS\_2013 |  |
| 7 | SARS\_2017 |  |
| 8 | MERS\_2012\_Saudi |  |
| 9 | MERS\_2014\_Saudi |  |
| 10 | MERS\_2014\_USA |  |

* 1. Show the similarity matrix D as a table of values (for all pairs of sequences).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |  |  |

* 1. Discuss your results and observations.

More specifically,  
- what does the fingerprint lengths tell you about each strain and its uniqueness?  
- what does the similarity matrix D tell you about the relationship between these strains? Do you see any logical groupings visible through this matrix?

1. **Performance:**         
   Create tables and/or charts to show the runtime results of the code. Please break down the runtime into the different components:

Task 1 Fingerprinting performance: suffix tree construction time; time to identify fingerprints; *total*time for the entire fingerprinting task;

**Task 1 Performance Table**

|  |  |  |
| --- | --- | --- |
| **Task Name** | **Time Taken for Tasks in Seconds** | **Total Time for Tasks in Seconds** |
| Construction of Suffix Tree |  |  |
| Fingerprint Identification |  |

Task 2 Similarity matrix performance: time spent building suffix trees; time spent performing the alignments; *total*time for the computation of the matrix D.  
**Task 2 Performance Table**

|  |  |  |
| --- | --- | --- |
| **Task Name** | **Time Taken for Tasks in Seconds** | **Total Time for Matrix Computation in Seconds** |
| Building Suffix Trees |  |  |
| Time spent performing alignment |  |

In addition, for task 2 performance, report the length of the longest common substrings reported for every pair of strings. You can show that too in the form of a table.

**Longest Common Substring Length for Each Pair**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |  |  |

1. **Other comments:**    If you have any other observations or comments, you can add them here in this section.