

Problem Statement A:

#1. Write a program to read the content of any of the below website and all its sub pages and perform following actions:

1. Parse all the pages and sub pages of News, Sports and Business section
2. Extract the content, Image and Links
3. Dump the Content, Image and Links into the respective mongo collections

Websites

1. <https://timesofindia.indiatimes.com/>
2. <https://www.firstpost.com/>
3. <https://indianexpress.com/>
4. <https://www.moneycontrol.com/> (Use Market, news and Mutual Funds section)

Notes:

1. Submit all the code and instruction to execute the code
2. You can use only java/node.js programming language to perform following tasks