1. You need to prepare the list of the dataset source names, volume in GBs, and total of articles in each source as a table **[5 pts]**.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sno.** | **Source name** | **Volume in GBs** | **Total no.of articles** |
| 1. | Kaggle  <https://www.kaggle.com/datasets/rushikeshdarge/hindi-text-corpus?select=hi_part_1.txt.gz> | 8.87 GB | 5 |
| 2. | Lindat Repository  <https://lindat.mff.cuni.cz/repository/xmlui/handle/11858/00-097C-0000-0023-6260-A> | 41.49 | 2 |
| 3. | Github  <https://github.com/cltk/hindi_text_ltrc> | 10 MB | 170 |
| 4. | Github  <https://github.com/gayatrivenugopal/Hindi-Aesthetics-Corpus> | 93 MB | 1047 |
| 5. | Cfilt  <https://www.cfilt.iitb.ac.in/iitb_parallel/> | 296 MB | 1 |
| 6. | Wortschatz  <https://wortschatz.uni-leipzig.de/en/download/Hindi> | 1.6 GB | 12 |
| 7. | Kaggle  <https://www.kaggle.com/datasets/abhishek/hindi-oscar-corpus/data> | 8.87GB | 1 |
| 8. |  |  |  |