|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Earth (431 Kb)** | **Mystery (444 Kb)** | **Myths (750 Kb)** | **Simak (311 Kb)** | **Wodehouse(403 Kb)** |
| **Canonical Collection 1** | | | | | |
| **words1** | 1459 | 1439 | 2276 | 1011 | 1322 |
| **Canonical Collection 2** | | | | | |
| **Short Text 1** | 1889 | 1915 | 3015 | 1328 | 1721 |
| **Short Text 2** | 1852 | 1857 | 3031 | 1323 | 1736 |
| **Short Text 3** | 1875 | 1865 | 3060 | 1323 | 1717 |
| **Short Text 4** | 1859 | 1828 | 2998 | 1300 | 1697 |
| **Short Text 5** | 1846 | 1839 | 3026 | 1299 | 1672 |
| **Short Text 6** | 1849 | 1830 | 3119 | 1302 | 1687 |
| **Short Text 7** | 1846 | 1834 | 2961 | 1296 | 1718 |
| **Short Text 8** | 1839 | 1843 | 2967 | 1302 | 1724 |
| **Short Text 9** | 1889 | 1925 | 3100 | 1333 | 1770 |
| **Short Text 10** | 1837 | 1859 | 3029 | 1311 | 1695 |
| **Canonical Collection 3** | | | | | |
| **Chesterton** | 1959 | 2022 | 3492 | 1394 | 1812 |
| **Dickens** | 1955 | 1970 | 3651 | 1374 | 1782 |

**Results:**

**Earth:** Canonical Collection 1 is best for earth file.

**Mystery:** Canonical Collection 1 is best for mystery file.

**Myths:** Canonical Collection 1 is best for myths file.

**Simak:** Canonical Collection 1 is best for simak file.

**Wodehouse:** Canonical Collection 1 is best for Wodehouse file.

SoCanonical Collection 1 is best for part 2.