https://voyant-tools.org/?corpus=e965de89827c9efc92b937597ebb709a

Find ## Replace Null

To remove the notations of pages from Hathi Trust plain text downloads.

Find <0x0c> Replace Null.

To remove Hathi Trust plain text formatting breaks.

To remove line breaks did the following:

* Find ([a-z]) ?\n(?=[a-z]) Replace \1space
* Find ([a-z]) ?\n(?=[A-Z]) Replace \1space
* Find ([a-z])- ?\n(?=[a-z]) Replace \1 Realized in my practice attempts before hitting replace all that this would also capture and misspell yellow-fever instances as well. Took entirely too many attempts but figured out that Find [w]-\n and Replace w- first would solve this problem. Only made one mistake to discover that Find [y]-\n and Replace y- would take care of the written numbers. Easily able then to fix scarlet-fever, enteric-fever, whopping-cough and small-pox.

<https://datastudio.google.com/u/0/reporting/d830d8c1-c6c7-480e-b7e1-fe0c3821bfbe/page/5QT2B>

<https://public.tableau.com/profile/david.ward5594#!/vizhome/Corpusv1/Sheet2>

<https://public.tableau.com/profile/will.hanley#!/vizhome/slc-extract/City>

<https://twitter.com/nassersaleh/status/1361840922825265153>

<https://www-cambridge-org.proxy.lib.fsu.edu/core/elements/network-turn/CC38F2EA9F51A6D1AFCB7E005218BBE5>

<https://www-cambridge-org.proxy.lib.fsu.edu/core/elements/network-turn/CC38F2EA9F51A6D1AFCB7E005218BBE5>