**Workflow (UNIX commands)**

How to find all non-ASCII characters in UNIX feed file ?

Use below command:

**grep --color='auto' -P -n "[\x80-\xFF]" <filename>**

How to remove non ASCII characters with one space only

* **cat 1.test | LANG=iso-8859-1 sed 's/[\x80-\xFF][\x80-\xFF]/\x20/g' > 2.test**
* **LANG=iso-8859-1 sed -i 's/[\x80-\xFF][\x80-\xFF]/\x20/g' 1.test**
* **LANG=iso-8859-1 sed -i 's/[\x80-\xFF][\x80-\xFF][\x80-\xFF]/\x20/g;s/[\x80-\xFF][\x80-\xFF]/\x20/g' 1.test**

