



NAYAK VINAYAK VINOD . <2021fc04135@wilp.bits-pilani.ac.in>

Clarification regarding non-utf-encoded data in retail-dataset file

Vinayak Nayak <2021fc04135@wilp.bits-pilani.ac.in>
To: "B. SUNIL" <sunilbhutada@wilp.bits-pilani.ac.in>
Bcc: SHREYSI KALRA <2021fc04586@wilp.bits-pilani.ac.in>

Thu, Mar 2, 2023 at 11:39 AM

Hi Sir,

Appreciate the prompt response:) Yes there is a way to modify the encoding and that is what I have done i.e. converted the file to utf-8 encoding and subsequently loaded it into HDFS.

However my question was that since we're not supposed to change the data in any way as mentioned in point 2, would this be considered acceptable?

Vinayak Vinod Nayak
2021FC04135
M.Tech (DSE)[2021 October Batch]
Cell: +91-8652380100

[Quoted text hidden]