

Vaziyat. Quyida biror mahsulot bo'yicha mijozlarning sharhlari keltirilgan:

1. "Klipga gap yo'q. Mazza qildim" ☺
2. "Klipda haddan tashqari bachkanalik elementlari mavjud"
3. "OOOOOOOOO, gap yo'q"
4. "bir marta ko'rsa bo'ladi"
5. "Maslahat bermayman" ):

So'raladi: Yuqoridagi sharhlarni quyidagi dastlabki qayta ishlash jarayonlaridan o'tkazing.

Kichik harflarga keltirish;

Tokenizatsiya;

Tinish belgilari va maxsus belgilarni olib tashlash;

To'xtash so'zlarini olib tashlash;

Stemming yoki Lemmatization.