|  |  |  |  |
| --- | --- | --- | --- |
| Book1.txt | Vie.xlsx | combined\_data.xlsx | combined\_data.txt |
| Đổi file xlsx thành DataFrame, rồi lưu dưới dạng txt, sau đó bỏ đi dòng header (Tiếng việt, tiếng anh) cho đúng format với file “vie.txt” | Đổi file txt thành Dataframe, rồi lưu dưới dạng xlsx, sau đó thay các header (0,1,2 thành Tiếng\_Việt, Tiếng\_Anh), xóa cột thứ 3 cho đúng format với file “Book.xlsx” | Với 2 file “Book1.xlsx” và “vie.xlsx” cùng format, chuyển cả 2 thành DataFrame, sau đó hợp nhất thành 1 file “combined\_data.xlsx” có hình dạng (263518x2), bỏ qua header | Tương tự cách làm với file “combined\_data.xlsx”, tạo ra file “combined\_data.txt” có hình dạng (263518x2) |