* 2017-2020: no separator between author affiliations and text. Text contains Introduction, Methods, Results and Conclusions, instead of only abstract. Anyway, its been considered that the first occurrence of \n is separating authors and second occurrence affiliations and after that its text including intro, methods, results, conclusion. So, few authors could be found in affiliations column and few affiliations data could be found in text column need to be cleaned manually.
* 2021: Again, different pattern and lack of separators. Few of the Authors and all the affiliations appended in the end of Title column and some could be found at the start of Text column as well. Affiliations column contains random data based on second occurrence of \n.
* Where there are nans there you can find all data appended to the Title column