Practical 2
2CS501
19BCE248

AIM:- Convert tables in pdf file tables to different formats using camelot and tabula module

Camelot:
Camelot is an open source library provided to extract table information in a well defined format from a pdf file. Here each table is extracted into a pandas dataframe.

Tabula:
Tabula also is an open source library provided to extract table information from a pdf file. Various file format options are also made available like directly converting extracted data to json,csv file formats.

Camelot vs Tabula:
Both are handy tools for carrying out similar kinds of activities. According to my experience both are similar as I have not explored much of the options. But Camelot seems more user-friendly and accurate in my case.

Conclusion:
From these practicals we explored some built-in libraries for carrying out the above mentioned task in a smooth way. And also it teaches us to explore more and more options and not stick to one in which we feel more comfortable.