Web Scrapping Assignment

Foreclosure is a legal process in which a lender attempts to recover the balance of a loan from a borrower who has stopped making payments. Homebuyers who want a good deal in real estate often think about buying a foreclosure. Our job is to identify and collect details of such properties which are being foreclosed. Information about this properties can be found at district court or bank's website.

Your task is to collect such data for **Martin County** of **Florida** state. Scrawl through the county's website and collect following details for each such property identified:

- 1) CaseDate date on which the case was filed
- 2) CaseId unique ID given to the cases being filed in the court
- 3) CaseType foreclosed properties can be homestead, non-homestead or commercial
- 4) DefendantName is the name of the borrower
- 5) BankName usually referred to as plaintiff, is the lender
- 6) ClaimValue is the amount borrower failed to repay

URL: https://court.martinclerk.com/Home.aspx/Search

Sequence of screenshots are provided to give you walkthrough as to how and where you will get the information.

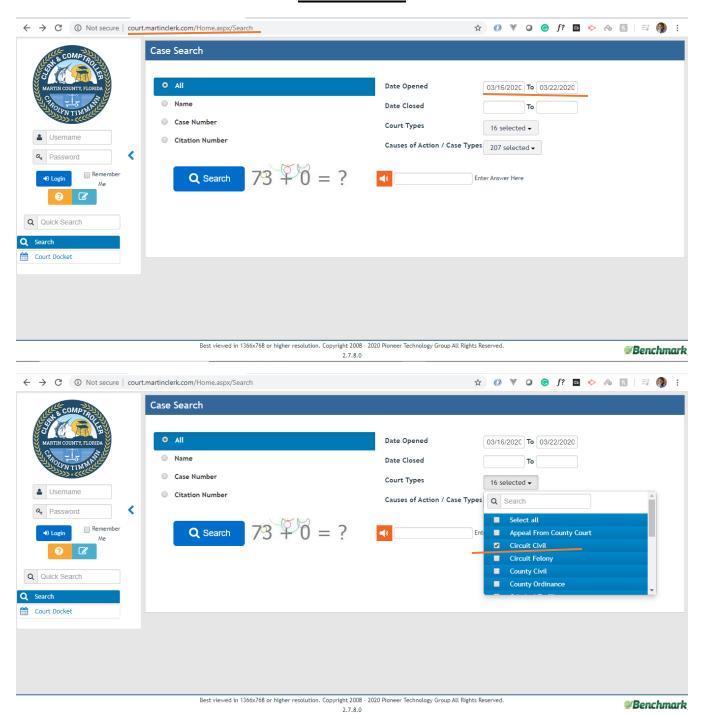
You need to automate this workflow using python, selenium and pytessaract (for pdf->image/image to text processing).

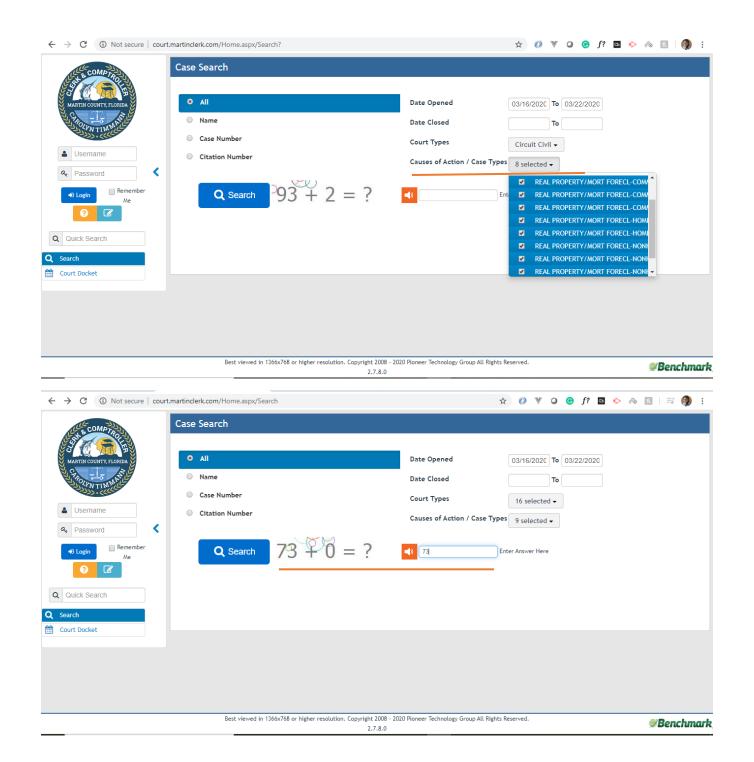
Save the output in a csv file with name martin_county.csv

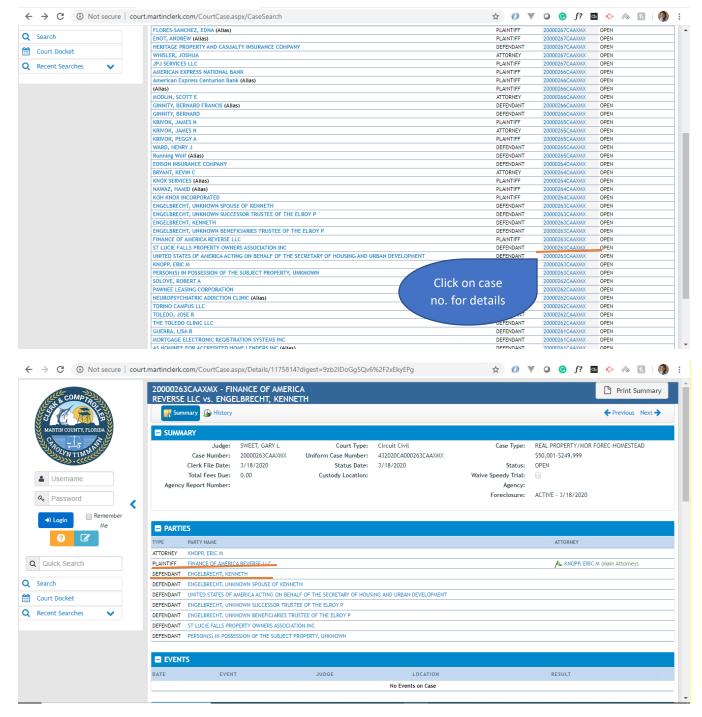
Things that are expected from code:

- Clean and reusable modules
- Give variables and functions meaningful names
- Proper comments should be added to make code self understood
- Exceptions need to be taken care and with user readable messages
- DO NOT HARD CODE anything (like path, or reference files)
- Also, make requirements.txt file containing all the required packages needed for execution

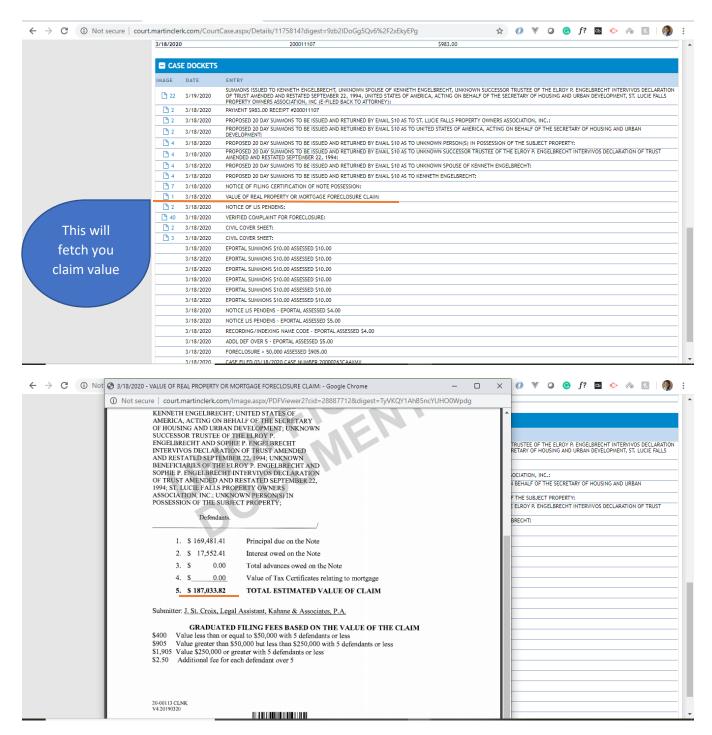
Work Flow







Always consider first defendant in the list



For this case, claim value = \$187,033.82