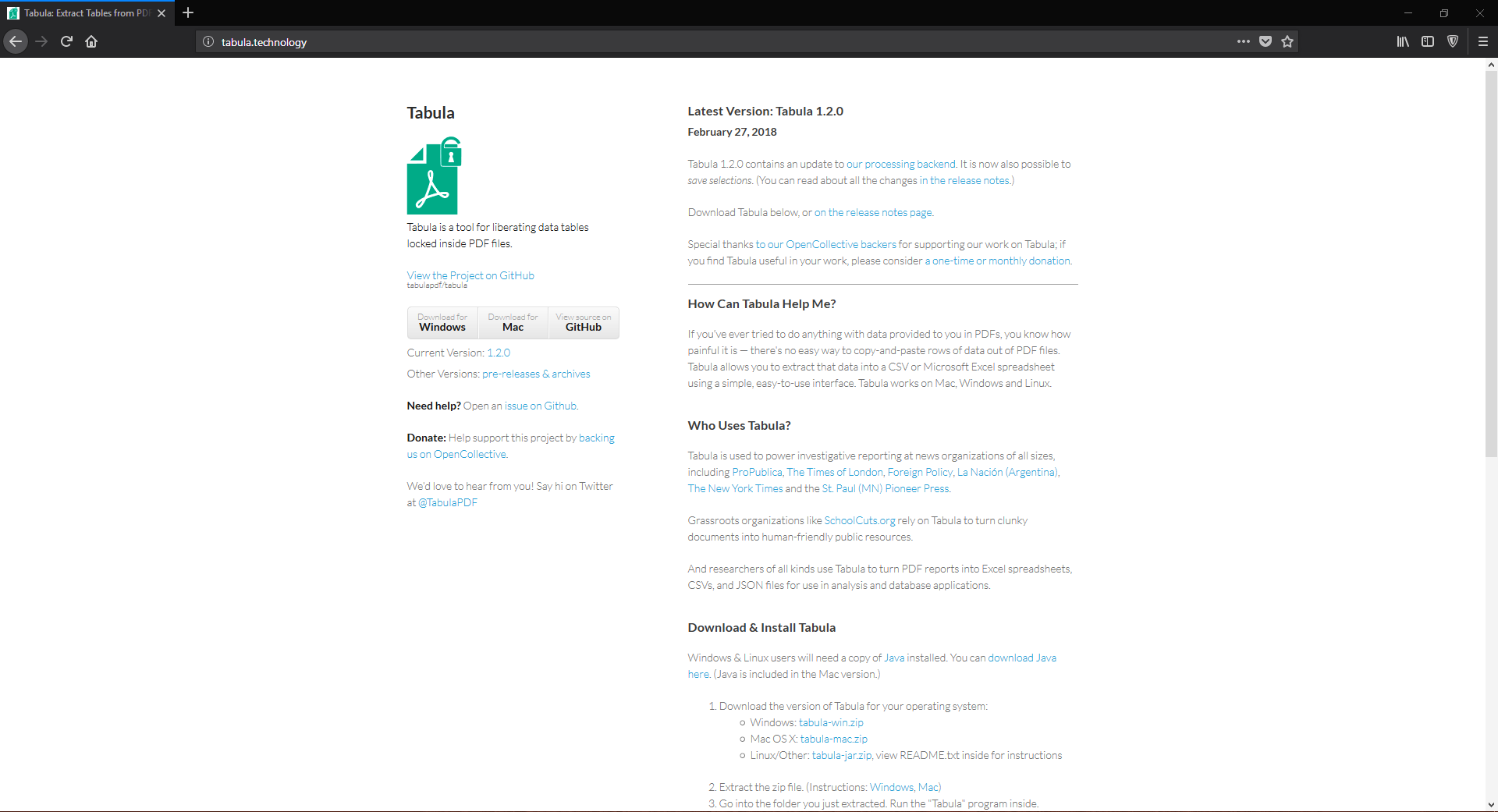
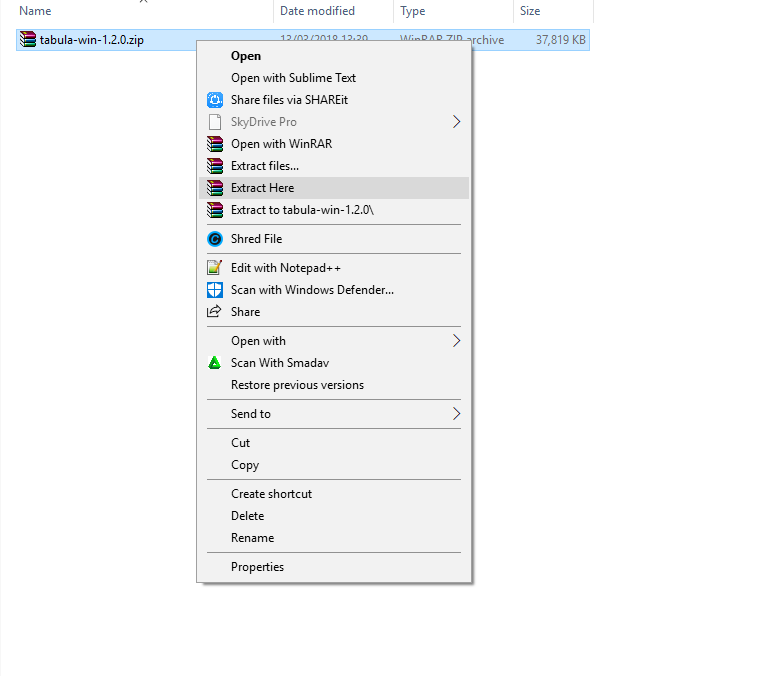
Laporan Scrape Data dari Pdf menjadi Excel(CSV) dengan aplikasi Tabula

Nama : Mumuh Kustino Muharram

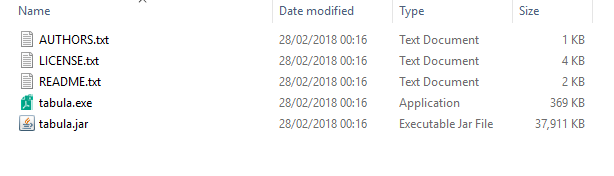
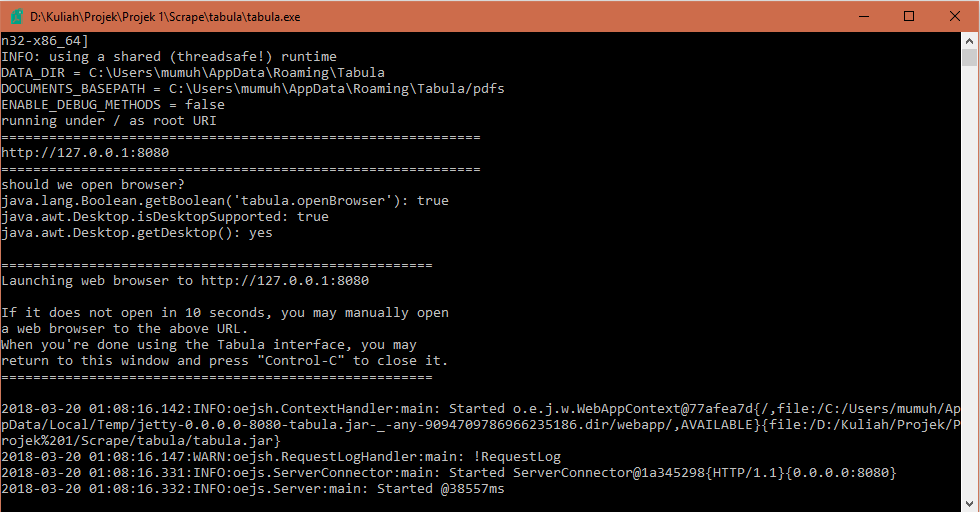
NIM : 171524025

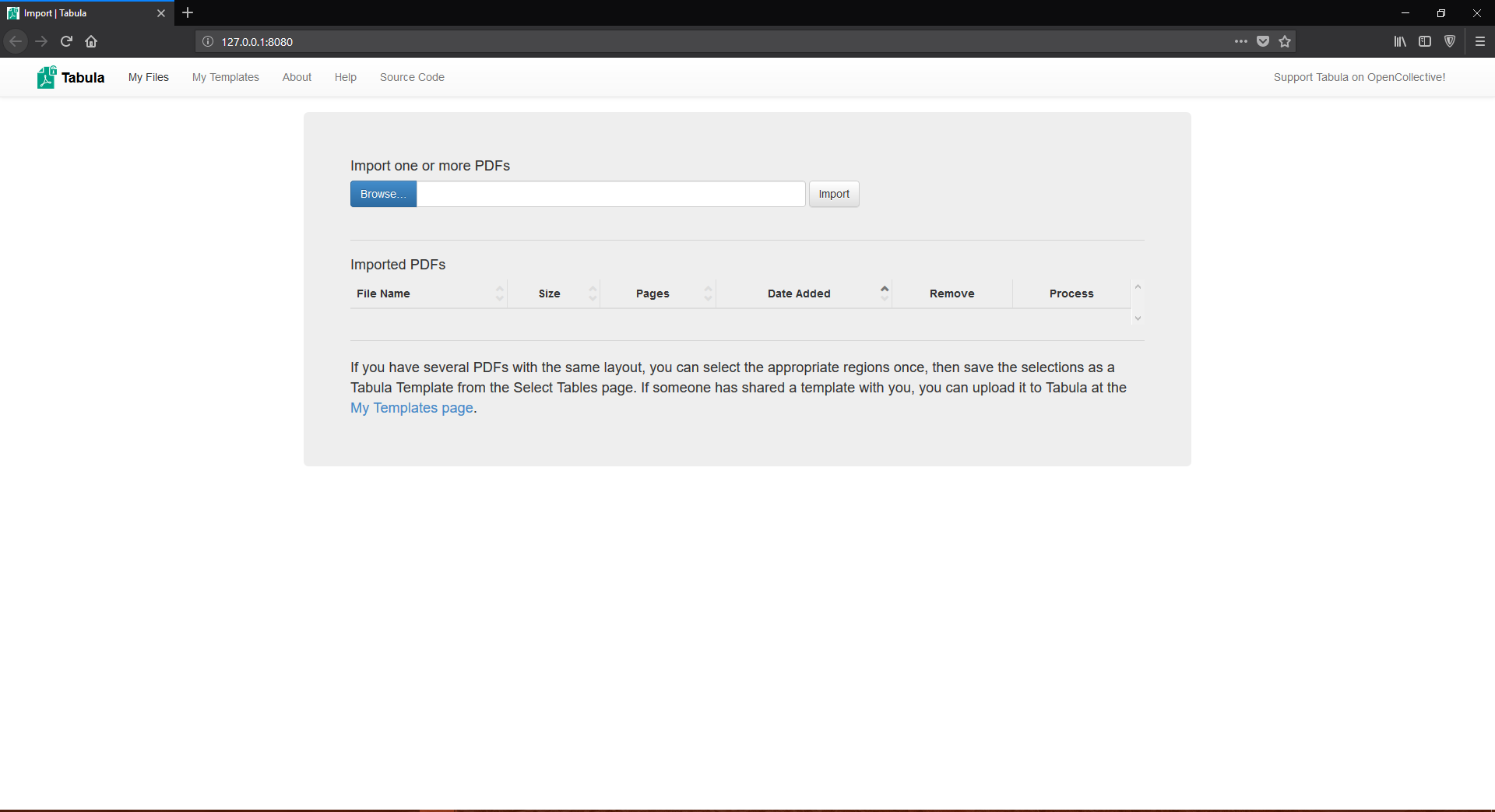
Kelas : 1A – D-4 Teknik Informatika

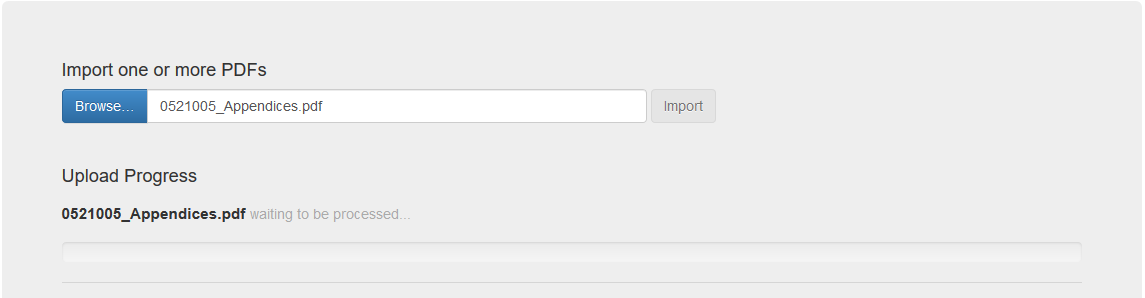
Pertama  
Download terlebih dahulu tabula di [http://tabula.technology/](http://tabula.technology/#)

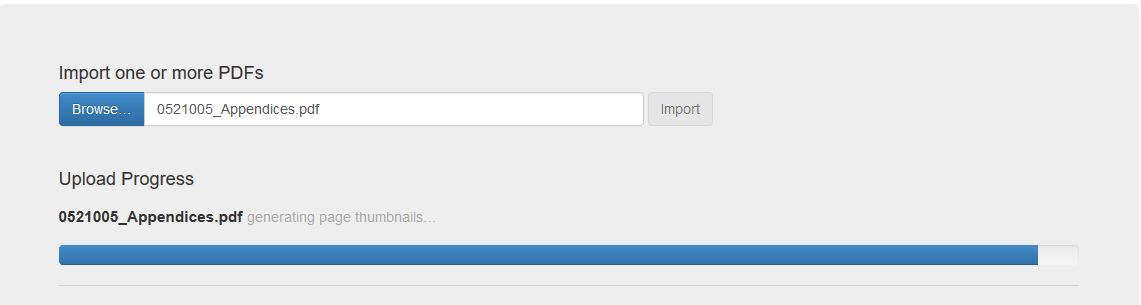
Kedua  
Ekstrak file yang di download tadi

Ketiga  
Jalankan aplikasi yang bernama tabula.exe

Maka akan keluar sebuah console yang akan membuka browser yang menuju sebuah aplikasi web 

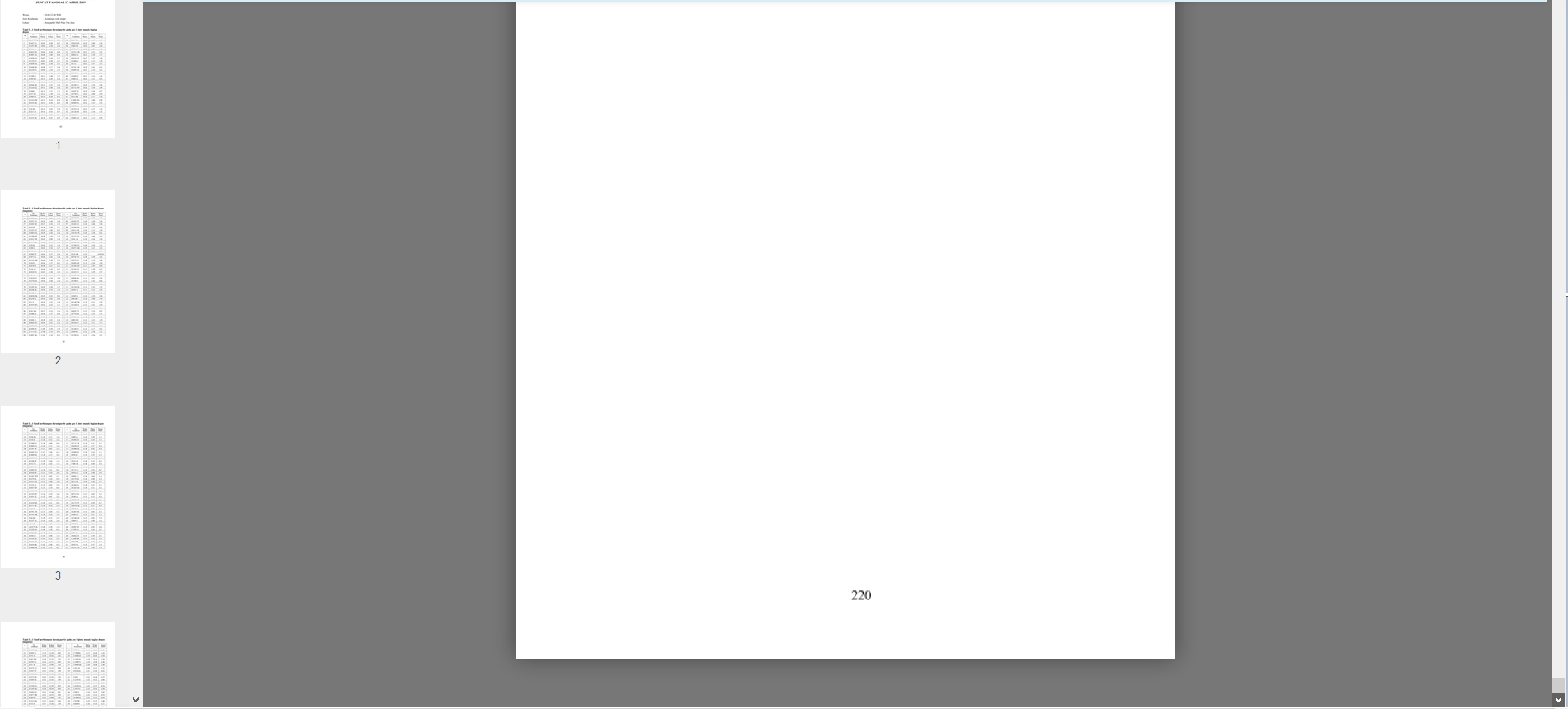
Berikut tampilan dari aplikasi web Tabul

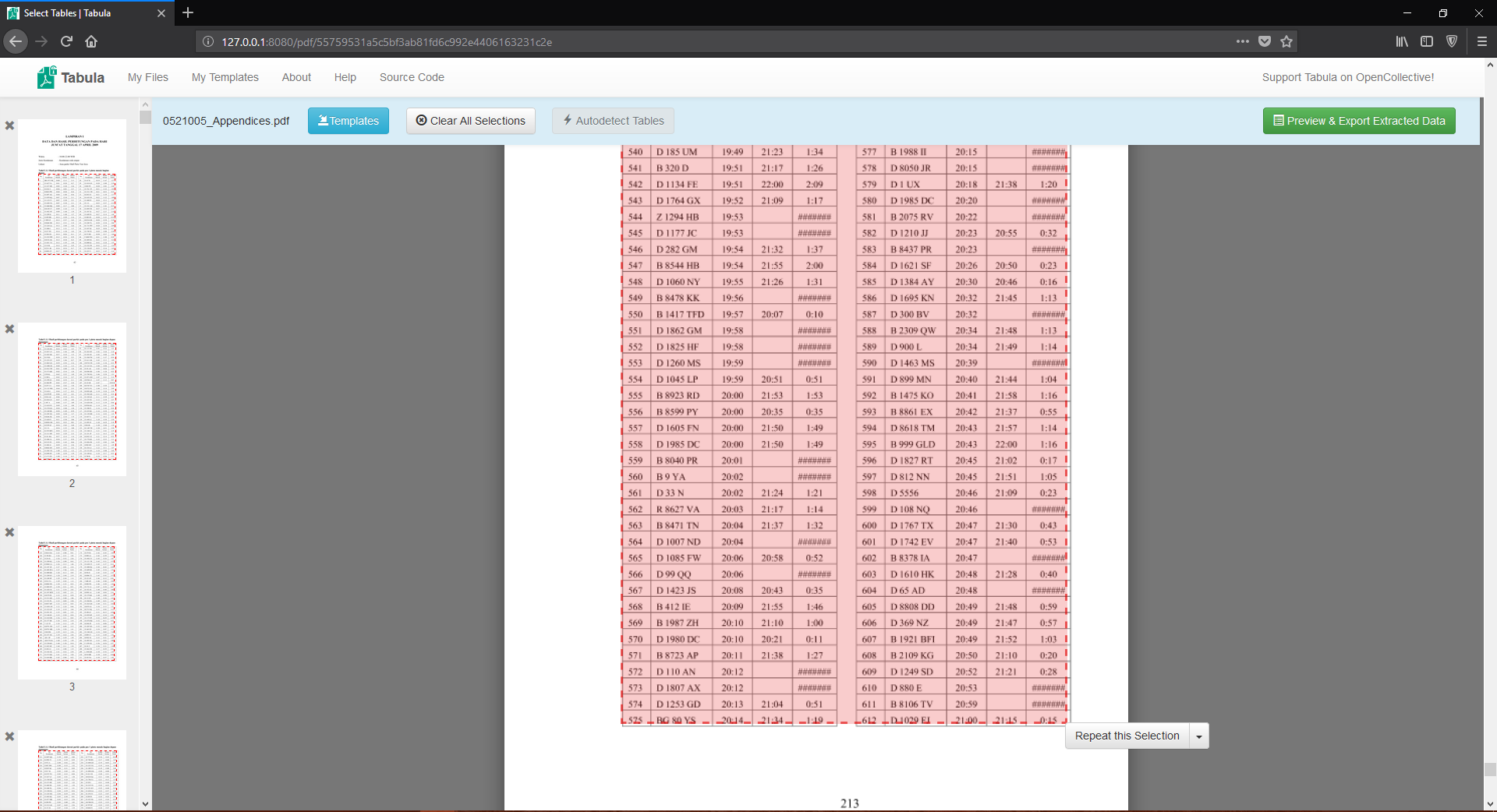
Keempat  
Tekan browse dan cari file yang ingin di ekstrak datanya, lalu tekan import. Tunggu loading. 



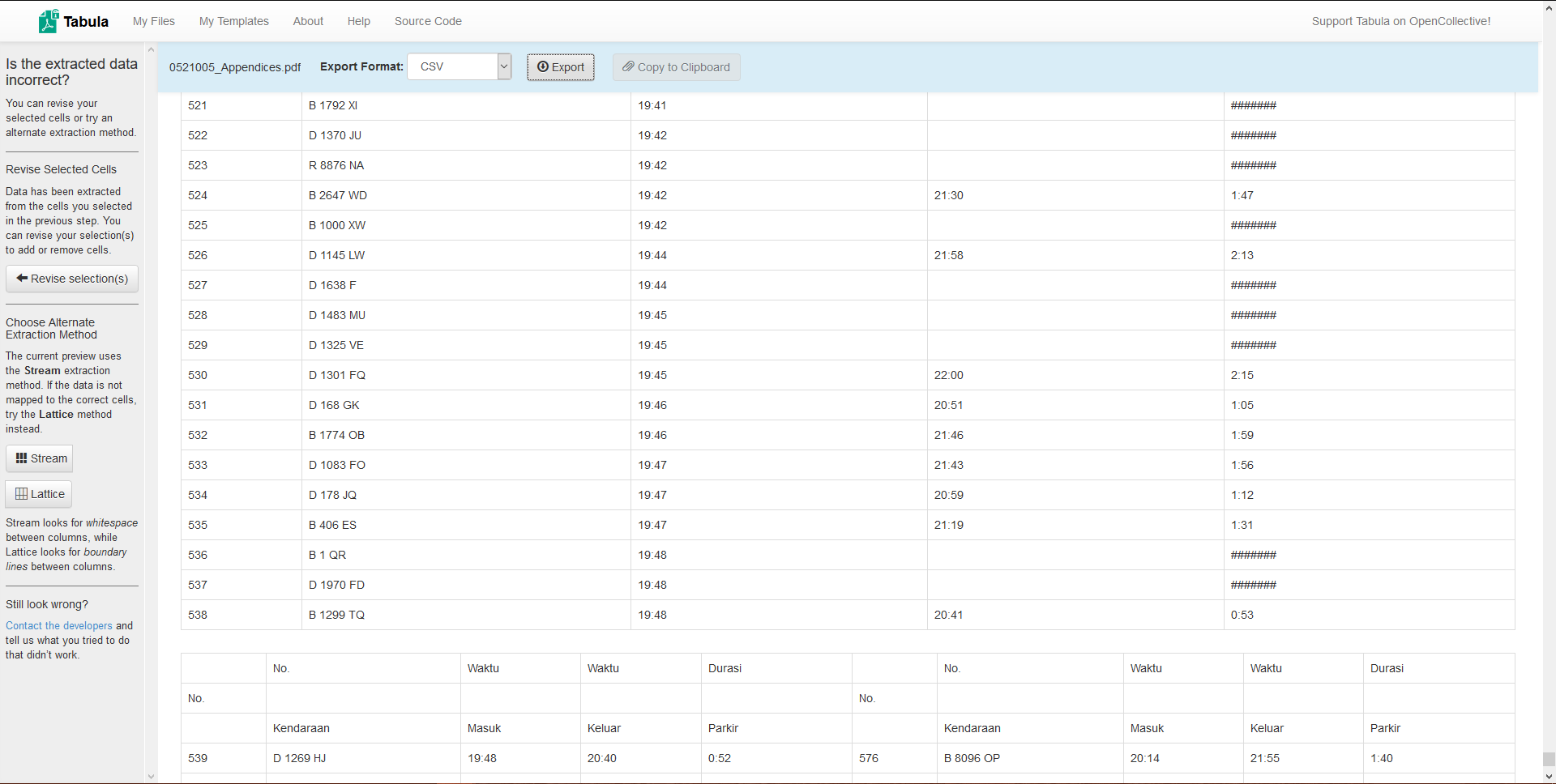
Kelima  
Ketika selesai tampilan pada web akan berubah, lalu scroll ke bawah sampai semua isi file tampil.

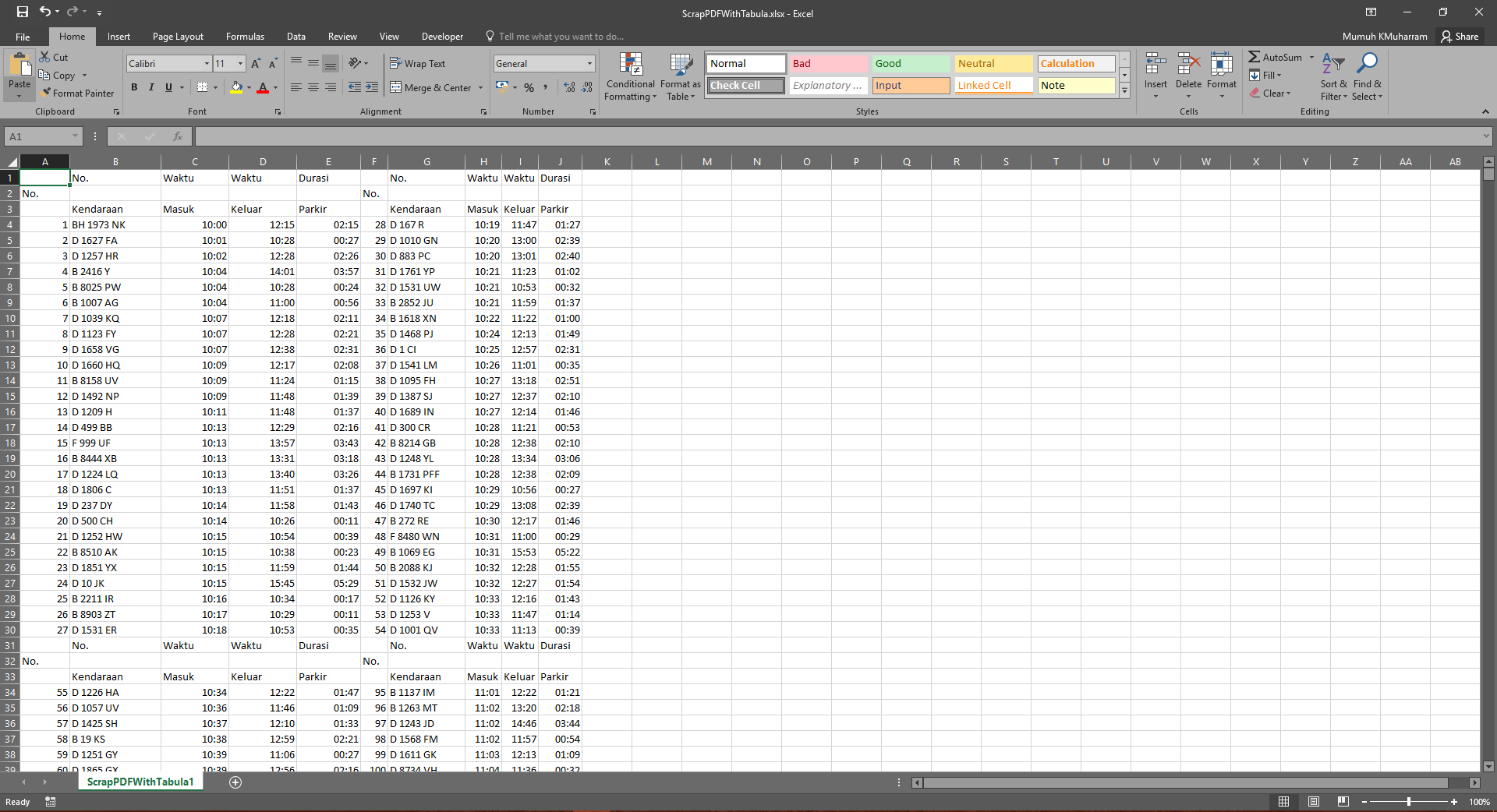
Lalu tekan button “Autodetect Tables” dan cek hingga semua table terdeteksi dengan benar.



Keenam  
Tekan tombol berwarna hijau yaitu “Preview & Export Extracted Data”. 

Ketujuh  
Cek kembali data yang sudah di scrape, lalu tekan “Export”, tombol “Export Format” opsional.



Kedelapan  
File akan didownload sesuai dengan format, dan untuk .csv dapat dibuka di Excel. 

Kelebihan

* Data yang ingin di ekstrak dapat di pilih secara manual, apakah berupa tabel saja atau dengan tulisan atau tulisan saja.
* Format dari data yang di scrape dapat disimpan dan digunakan lagi untuk file yang berbeda tetapi format sama.
* Ketika akan mengekstrak data dengan format yang sama dengan sebelumnya kita hanya perlu menggunakan button “Template”, dan mengekstraknya dengan button “Export”.

Kekurangan

* Export dari hasil scrape tidak bisa dengan format Excel (\*.xlsx).
* Di beberapa computer yang format dari systemnya tidak English(UK) maka ketika file.csv dibuka tidak akan tampil dengan rapih, tetapi harus di setting dahulu dengan button “Text to Columns” untuk mengubah dari char “ , “ comma supaya menjadi cell-cell yang berbeda, dan terlihat rapih.
* Waktu yang dibutuhkan lebih lama.