Michigan Foreclosure Sales

Notes: This python scripts scrapes a folder of PDFs from a Michigan Foreclosure Sales and creates a csv output file for each PDF. The following column headers are created and the associated data is collected: **County, Parcel and Purchase Price**. The PDF filenames are not unique and each CSV file is named based on the current number in a series. Multiple counties were listed in the PDFs and each county had a separate regular expression to search for the parcel number.

PDF

2017 Foreclosure Sales Disposition Details   2-20-18

County	Local Parcel Number	Disposition Date	Disposition Code	Sale Price	Deeded Party(s)
BRANCH	010-012-300-005-01	07/31/17	SOLD	$2,000.00	8392 S 29TH ST PROPERTIES LLC, 1230 DAY RD, UNION CITY, MI 49094
BRANCH	010-029-200-015-07	07/31/17	SOLD	$1,000.00	WENDY LAMARCHE, UNIT 1011 6223 31 AVE, CALGARY, AB T3B 3X2
BRANCH	010-C30-000-001-00	07/31/17	SOLD	$8,500.00	8392 S 29TH ST PROPERTIES LLC, 1230 DAY RD, UNION CITY, MI 49094
BRANCH	010-C30-000-175-00	07/31/17	SOLD	$900.00	WENDY LAMARCHE, UNIT 1011 6223 31 AVE, CALGARY, AB T3B 3X2
BRANCH	010-M21-000-068-00	11/02/17	SOLD	$650.00	CHRISTIAN E. RIEHL, 1535 BERGER ST. SOUTH, EAST CANTON, OH 44730
BRANCH	010-M21-000-069-00	07/31/17	SOLD	$1,400.00	SONG HUI KIM, 16876 HAWKS RD, JONES, MI 49061
BRANCH	010-O55-000-254-00	07/31/17	SOLD	$600.00	WARREN BUSSCHER PROPERTIES LLC, 6231 FILLMORE ST, ALLENDALE, MI 49401
BRANCH	011-000-009-016-00	07/31/17	SOLD	$2,000.00	WILLIAM MATHEW MAYHEW, 262 RAILROAD ST, SHERWOOD, MI 49089; VIRGINIA G MAYHEW, 43 ALLISON, BATTLE CREEK, MI 49037
BRANCH	020-002-200-015-01	07/31/17	SOLD	$19,000.00	8392 S 29TH ST PROPERTIES LLC, 1230 DAY RD, UNION CITY, MI 49094
BRANCH	020-005-300-016-00	12/30/17	Municipal Transfer	$0.00	Union Township
BRANCH	021-000-001-035-00	07/31/17	SOLD	$14,250.00	TOGDEN INVESTMENTS, LLC, 6183 6 MILE ROAD, EAST LEROY, MI 49051
BRANCH	021-000-020-180-00	07/31/17	SOLD	$3,100.00	WARREN BUSSCHER PROPERTIES LLC, 6231 FILLMORE ST, ALLENDALE, MI 49401
BRANCH	021-000-024-004-00	07/31/17	SOLD	$650.00	DAVID EVERETT JOHNSON BARBI SUE JOHNSON, 337 WOODRUFF ST, UNION CITY, MI 49094
BRANCH	021-000-029-010-01	12/30/17	Municipal Transfer	$0.00	Village of Union City
BRANCH	030-T25-000-038-99	07/31/17	SOLD	$30,500.00	MARK C LARDER TERESA I LARDER TRUST, 411 W MAIN ST, HOMER, MI 49245
BRANCH	040-B10-000-006-99	07/31/17	SOLD	$1,700.00	JEREMY JOSEPH KAY, 2911 SAUER CT, NORTH AURORA, IL 60542; RONALD WEISS AND MARY ANN WEISS, 1046 BROOKWOOD DRIVE, QUINCY, MI 49082; JILLIAN WEISS, 2911 SAUER CT, NORTH AURORA, IL 60542
BRANCH	081-N01-000-168-00	07/31/17	SOLD	$16,500.00	DANIELLE BADDERS, 408 N FOX RD, COLDWATER, MI 49036
BRANCH	081-501-000-268-00	07/31/17	SOLD	$18,500.00	DANIELLE BADDERS, 408 N FOX RD, COLDWATER, MI 49036
BRANCH	100-001-300-010-99	07/31/17	SOLD	$5,000.00	PAUL W. & G. ELAINE GREEN TRUST, 201 SOUTH BUTCHER ROAD, COLDWATER, MI 49036
BRANCH	100-031-200-005-01	07/31/17	SOLD	$10,250.00	DONALD LYNN MAYER CAROLYN JO MAYER, 652 S PARHAM RD, BRONSON, MI 49028
BRANCH	110-B40-000-078-00	12/30/17	MI Land Bank	$0.00	Michigan Land Bank Fast Track Authority
BRANCH	110-F20-000-143-00	07/31/17	SOLD	$750.00	JEFFERY THOMAS NOBLIT, 493 CHICKADEE DR, COLDWATER, MI 49036
BRANCH	140-014-200-030-00	07/31/17	SOLD	$15,750.00	KIRK A SKINNER AND DEANA L SKINNER, 539 BOOTH RD, BRONSON, MI 49028
BRANCH	150-004-200-030-06	07/31/17	SOLD	$13,750.00	FOUR KIDS EDUCATIONAL TRUST, 8268 ST RTE 15, BRYAN, OH 43506
BRANCH	150-011-300-015-01	07/31/17	SOLD	$2,500.00	LEO OSTERHOLT, 1102 GRAND AVE, CELINA, OH 45822
BRANCH	150-015-000-010-00	07/31/17	SOLD	$7,200.00	DIVINA ALEXAS CASTANEDA THRASHER, 3549 SCHOFIELD AVE, INDIANAPOLIS , IN 46218
BRANCH	200-001-000-244-00	07/31/17	SOLD	$18,000.00	CHRISTOPHER BRENT RUMSEY, 393 W CHICAGO RD, COLDWATER, MI 49036
BRANCH	200-002-000-169-00	07/31/17	SOLD	$29,000.00	ROBERTS RENTAL & REAL ESTATE LLC, 130 MATILDA ST, BRONSON, MI 49028
BRANCH	200-003-000-144-00	07/31/17	SOLD	$27,000.00	CHRISTOPHER BRENT RUMSEY, 393 W CHICAGO RD, COLDWATER, MI 49036
BRANCH	200-004-000-161-00	07/31/17	SOLD	$17,500.00	ROBERTS RENTAL & REAL ESTATE LLC, 130 MATILDA ST, BRONSON, MI 49028
BRANCH	301-000-000-983-00	10/06/17	Municipal Purchase	$1,011.76	CITY OF COLDWATER, ONE GRAND STREET, COLDWATER, MI 49036
BRANCH	304-000-000-395-00	11/02/17	SOLD	$4,600.00	RYAN JOHN SADDLER AND JESSICA MICHELLE SADDLER, 126 CUTTER AVE, COLDWATER, MI 49036
CLINTON	010-019-300-015-01	11/02/17	SOLD	$200.00	BZANY LLC, PO BOX 474, KEEGO HARBOR, MI 48320
CLINTON	010-034-400-031-00	08/28/17	SOLD	$16,500.00	MARK SANDERS LANSING #2, LLC, 8787 COLEMAN ROAD, HASLETT, MI 48840
CLINTON	010-100-005-003-00	12/30/17	Municipal Transfer	$0.00	Bath Township

Python Code (20190808_MichiganForeclosure.py)

```python
# Read a list of PDF files in folder, open, scan, and dump data individual text files

# Accumatch

# by Michael Keller

# Aug 7, 2019

# requires pdfplumber, re, datetime, os


# Libraries

import pdfplumber

import re

from datetime import datetime

import os

import sys


# Start Timer

start = datetime.now()


# Booleans

Parcel_Found = False

SalePrice_Found = False


# Variables

regSalePrice = "['$'][,0-9]+['.']\d{2}"

space = ' '

nospace = ''

comma = ','

_return = '\r'

count = 0

fileCount = 0

outString = []

outFilename = 'MichiganForeclosure_'

outputType = '.csv'

fileName = ''

##parcel = ''

##amount = ''

current_county = 'NotFound'
```

```python
expression = ''

dollarSign = '$'

restOfLine = ''

SalePrice = ''

colHeadings = 'County,Parcel,Sale Price'


# Regular Expression County Dictionary

counties = {

    'ALGER':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'ARENAC':'\d{3}\D{1}\d{1}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'BARAGA':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'BARRY':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'BAY':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'BENZIE':'\d{12}',

    'BRANCH':'\d{3}\D{1}\S{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'CALHOUN':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'CASS':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'CHARLEVOIX':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'CHIPPEWA':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'CLARE':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'CLINTON':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'CRAWFORD':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'DELTA':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}',

    'DICKINSON':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{0,4}',

    'EATON':'\d{2}.{1}\d{2,3}.{1}\d{2,3}.{1}\d{2,3}.{1}\d{3}.{1}\d{2,3}',

    'EMMET':'\d{2}\D{1}\d{2}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}',

    'GOGEBIC':'\d{2}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}',

    'GRAND':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'HOUGHTON':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'IOSCO':'\d{3}\D{1}.{3}\D\d{3}\D\d{3}\D\d{2}',

    'INGHAM':'\d{2}\D{1}\d{2}\D{1}\d{2}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}',

    'IRON':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{0,4}',

    'LEELANAU':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',

    'KALAMAZOO':'\d{2}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}',

    'KALKASKA':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{0,4}',

    'KEWEENAW':'\d{3}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}',
```

```python
    'LIVINGSTON':'\d{4}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}',
    'LUCE':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{0,4}',
    'MANISTEE':'\d{12}',
    'MARQUETTE':'\d{2}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'MASON':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'MECOSTA':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}',
    'MISSAUKEE':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'MONROE':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'MONTMORENCY':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'MUSKEGON':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{4}\D{1}\d{2}',
    'OCEANA':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'OGEMAW':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'ONTONAGON':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'OSCEOLA':'\d{12}',
    'OSCODA':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'OTSEGO':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'OTTAWA':'\d{2}\D{1}\d{2}\D{1}\d{2}\D{1}\d{3}\D{1}\d{3}',
    'SAINT CLAIR':'\d{14}',
    'SAINT JOSEPH':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'SAGINAW':'\d{1}\D{1}\d{2}\D{1}\d{1}\D{1}\d{5}\D{1}\d{3}\D{1}\d{2}',
    'SCHOOLCRAFT':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
    'SHIAWASSEE':'\d{3}\D{1}\d{2,3}\D{1}\d{3}\D{1}.{3,6}',
    'TUSCOLA':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{4}\D{1}\d{2}'
    }


##    'EATON':'\d{2}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{2}',
##    'IOSCO':'\d{3}\w\d{10,11}',
##    'SHIAWASSEE':'\d{3}\D{1}\d{3}\D{1}\d{3}\D{1}\d{3}',


# Messages
messageSaveBeg = 'Saving \''
messageSaveEnd = '\' file . . .'
messageEOL = '. . . end of line'


# Functions
def getCurrentDirectory():
```

```python
    os.chdir(os.path.dirname(__file__))

    return os.getcwd()


def getFiles(ext):

    for(dirpath,dirnames,filenames) in os.walk(getCurrentDirectory()):

        return (f for f in filenames if f.endswith(ext))


def printPDFfilename(fn):

    return 'Reading ' + fn + ' file . . .'


def openFile(fn):

    # Create output file

    try:

        file = open(fn,'w+')

        # Write out lines from array

        file.write(colHeadings + _return)

        file.close()

        print('Opening ' + fn + ' file . . .')

    except FileNotFoundError:

        sys.exit('File: ' + fn + ' does not exist')

    except PermissionError:

        sys.exit('Unable able to open output file for writing. Is the file currently open?')


def writeFile(fn,os):

    # Create and append to output file

    try:

        file = open(fn,'a+')

        for taxLine in os:

            file.write(taxLine + _return)

        file.close()

        print('Saving ' + fn + ' file . . .')

    except FileNotFoundError:

        sys.exit('File: ' + fn + ' does not exist')

    except PermissionError:

        sys.exit('Unable able to open output file for writing. Is the file currently open?')
```

```python
PDFs = getFiles('.pdf')


for x in PDFs:
    ## Message Run
    print(printPDFfilename(x))


    ## Open PDF file for reading
    pdf = pdfplumber.open(x)


    lastPage = len(pdf.pages)


    ## Reset counters
    count = 0
    fileCount = fileCount + 1


    print('Number of Pages: ' + str(lastPage))


    fileName = outFilename + str(fileCount) + outputType
    openFile(fileName)


    for page in pdf.pages:
        p1 = pdf.pages[count]
        count = count + 1
        pltext = p1.extract_text()
        text = pltext.splitlines()


##     Break for Testing
##          if(count == 6):
##              break


        for item in text:
            Item_Found = False
            if (current_county == item[0:len(current_county)]):
                Item_Found = True
            else:
                for x in counties:
```

```python
                    if(x == item[0:len(x)]):
                        current_county = x
                        Item_Found = True
                        break


            if(Item_Found == True):
                expression = counties.get(current_county)
                Parcel_Found = re.search(expression,item)
                if not (Parcel_Found is None):
                    Parcel = item[Parcel_Found.start():Parcel_Found.end()]
                    Parcel = Parcel.replace('-','-')
                    if(space in Parcel):
                        blob = Parcel.split()
                        Parcel = blob[0]
                    restOfLine = item[Parcel_Found.end():len(item)]
                    SalePrice_Found = re.search(regSalePrice,restOfLine)
                    SalePrice = restOfLine[SalePrice_Found.start():SalePrice_Found.end()]
                    SalePrice = SalePrice.replace(comma,nospace)
                    SalePrice = SalePrice.replace(dollarSign,nospace)
                    outString.append(current_county + comma + Parcel + comma + SalePrice)


    ## Close PDF
    pdf.close()


    ## Write CSV file
    writeFile(fileName,outString)


    ## Clear variables
    pl = ''
    pltext = ''
    text = ''
    _TownshipFound = False
    ## Make sure outString is clear for next file's data
    outString.clear()


#end timer
```

```python
print('Runtime: ' + str(datetime.now() - start))


print(messageEOL)


# EOF
```

CSV Output

	A	B	C
1	County	Parcel	Sale Price
2	BRANCH	010-012-300-005-01	2000
3	BRANCH	010-029-200-015-07	1000
4	BRANCH	010-C30-000-001-00	8500
5	BRANCH	010-C30-000-175-00	900
6	BRANCH	010-M21-000-068-00	650
7	BRANCH	010-M21-000-069-00	1400
8	BRANCH	010-O55-000-254-00	600
9	BRANCH	011-000-009-016-00	2000
10	BRANCH	020-002-200-015-01	19000
11	BRANCH	020-005-300-016-00	0
12	BRANCH	021-000-001-035-00	14250
13	BRANCH	021-000-020-180-00	3100
14	BRANCH	021-000-024-004-00	650
15	BRANCH	021-000-029-010-01	0
16	BRANCH	030-T25-000-038-99	30500
17	BRANCH	040-B10-000-006-99	1700
18	BRANCH	081-N01-000-168-00	16500
19	BRANCH	081-S01-000-268-00	18500
20	BRANCH	100-001-300-010-99	5000
21	BRANCH	100-031-200-005-01	10250
22	BRANCH	110-B40-000-078-00	0
23	BRANCH	110-F20-000-143-00	750
24	BRANCH	140-014-200-030-00	15750
25	BRANCH	150-004-200-030-06	13750
26	BRANCH	150-011-300-015-01	2500
27	BRANCH	150-O15-000-010-00	7200
28	BRANCH	200-001-000-244-00	18000

Michael Keller
Accumatch
20190808_MichiganForeclosure
Aug 9, 2019