

# Pueblo County (Two column)

**Notes:** This python scripts scrapes PDFs from a Delinquent tax payer report from different counties and creates a csv output file with the following items: **Account, Year, Date, Original, Interest, Fees, Total.** T The PDF is unique because the data is presented in two columns. It reads from top to bottom, left to right. The output filename is not unique and is static.

**PDF**

Pueblo County Held Properties		Date 07/08/2019	Page 1																														
<b>418000014</b> <b>Book: 998 Cert: 2019</b> <span style="float: right;"><b>10,803.44</b></span> SUBURBAN LAND CO 18-20-64 CONTG 2.1A IN THE NE4, SE4 SEC 18-20-64: BEG INTER OF NLY R/W LINE OF COLO HWY NO 47 (COLLEGE RD) AS PRESENTLY LOC ++ THE WLY R/W LINE OF COLO + S RR (NOW ABANDONED) TH N 12 DEG 02 MIN E ALG SD WLY LINE OF SD C-SRR R/W TO THE E-W C/L SD SEC 18 TH S 89 DEG 21 MIN E ALG SD C/L 38.08 FT TO A PT ON EXISTING FLOOD PLAIN DIST SD PT BEING TRU PT BEG TH SLY ALG FLOOD PLAIN DIST 2 COURSES: S 12 DEG 03 MIN W, 204.4 FT S 16 DEG 03 MIN W, 299.96 FT TO PT ON NLY R/W COLO HWY NO 47 (COLLEGE RD) TH N 84 DEG 18 MIN E, 190 FT ML TO INTER OF NLY R/W LINE SD HWY + IN BK 1941 PG 475 TH N 16 DEG 03 MIN E, 303.81 FT TH N 12 DEG 03 MIN E, 204.4 FT TO E-W C/L SEC 18 TH N 89 DEG 21 MIN E ALG SD C/L 180 FT ML TO TRUE PT BEG. FORMERLY #04181-00-014		TO E-W C/L SEC 18 TH N 89 DEG 21 MIN E ALG SD C/L 180 FT ML TO TRUE PT BEG. FORMERLY #04181-00-014 <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>1994 10/21/1998</td> <td>1,635.32</td> <td>4,977.64</td> <td>51.25</td> <td>6,684.21</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	1994 10/21/1998	1,635.32	4,977.64	51.25	6,684.21																		
Year	Date	Original	Interest	Fees	Total																												
Cert	1994 10/21/1998	1,635.32	4,977.64	51.25	6,684.21																												
<b>418000014</b> <b>Book: 998 Cert: 2023</b> <span style="float: right;"><b>6,332.88</b></span> SUBURBAN LAND CO 18-20-64 CONTG 2.1A IN THE NE4, SE4 SEC 18-20-64: BEG INTER OF NLY R/W LINE OF COLO HWY NO 47 (COLLEGE RD) AS PRESENTLY LOC ++ THE WLY R/W LINE OF COLO + S RR (NOW ABANDONED) TH N 12 DEG 02 MIN E ALG SD WLY LINE OF SD C-SRR R/W TO THE E-W C/L SD SEC 18 TH S 89 DEG 21 MIN E ALG SD C/L 38.08 FT TO A PT ON EXISTING FLOOD PLAIN DIST SD PT BEING TRU PT BEG TH SLY ALG FLOOD PLAIN DIST 2 COURSES: S 12 DEG 03 MIN W, 204.4 FT S 16 DEG 03 MIN W, 299.96 FT TO PT ON NLY R/W COLO HWY NO 47 (COLLEGE RD) TH N 84 DEG 18 MIN E, 190 FT ML TO INTER OF NLY R/W LINE SD HWY + IN BK 1941 PG 475 TH N 16 DEG 03 MIN E, 303.81 FT TH N 12 DEG 03 MIN E, 204.4 FT TO E-W C/L SEC 18 TH N 89 DEG 21 MIN E ALG SD C/L 180 FT ML TO TRUE PT BEG. FORMERLY #04181-00-014		<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>1995 10/21/1998</td> <td>1,564.60</td> <td>4,713.03</td> <td>51.25</td> <td>6,328.88</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	1995 10/21/1998	1,564.60	4,713.03	51.25	6,328.88																		
Year	Date	Original	Interest	Fees	Total																												
Cert	1995 10/21/1998	1,564.60	4,713.03	51.25	6,328.88																												
<b>418000014</b> <b>Book: 998 Cert: 2024</b> <span style="float: right;"><b>5,638.56</b></span> SUBURBAN LAND CO 18-20-64 CONTG 2.1A IN THE NE4, SE4 SEC 18-20-64: BEG INTER OF NLY R/W LINE OF COLO HWY NO 47 (COLLEGE RD) AS PRESENTLY LOC ++ THE WLY R/W LINE OF COLO + S RR (NOW ABANDONED) TH N 12 DEG 02 MIN E ALG SD WLY LINE OF SD C-SRR R/W TO THE E-W C/L SD SEC 18 TH S 89 DEG 21 MIN E ALG SD C/L 38.08 FT TO A PT ON EXISTING FLOOD PLAIN DIST SD PT BEING TRU PT BEG TH SLY ALG FLOOD PLAIN DIST 2 COURSES: S 12 DEG 03 MIN W, 204.4 FT S 16 DEG 03 MIN W, 299.96 FT TO PT ON NLY R/W COLO HWY NO 47 (COLLEGE RD) TH N 84 DEG 18 MIN E, 190 FT ML TO INTER OF NLY R/W LINE SD HWY + IN BK 1941 PG 475 TH N 16 DEG 03 MIN E, 303.81 FT TH N 12 DEG 03 MIN E, 204.4 FT TO E-W C/L SEC 18 TH N 89 DEG 21 MIN E ALG SD C/L 180 FT ML TO TRUE PT BEG. FORMERLY #04181-00-014		<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>1996 10/21/1998</td> <td>1,387.33</td> <td>4,195.98</td> <td>51.25</td> <td>5,634.56</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	1996 10/21/1998	1,387.33	4,195.98	51.25	5,634.56																		
Year	Date	Original	Interest	Fees	Total																												
Cert	1996 10/21/1998	1,387.33	4,195.98	51.25	5,634.56																												
<b>418000014</b> <b>Book: 998 Cert: 2020</b> <span style="float: right;"><b>7,784.04</b></span> SUBURBAN LAND CO 18-20-64 CONTG 2.1A IN THE NE4, SE4 SEC 18-20-64: BEG INTER OF NLY R/W LINE OF COLO HWY NO 47 (COLLEGE RD) AS PRESENTLY LOC ++ THE WLY R/W LINE OF COLO + S RR (NOW ABANDONED) TH N 12 DEG 02 MIN E ALG SD WLY LINE OF SD C-SRR R/W TO THE E-W C/L SD SEC 18 TH S 89 DEG 21 MIN E ALG SD C/L 38.08 FT TO A PT ON EXISTING FLOOD PLAIN DIST SD PT BEING TRU PT BEG TH SLY ALG FLOOD PLAIN DIST 2 COURSES: S 12 DEG 03 MIN W, 204.4 FT S 16 DEG 03 MIN W, 299.96 FT TO PT ON NLY R/W COLO HWY NO 47 (COLLEGE RD) TH N 84 DEG 18 MIN E, 190 FT ML TO INTER OF NLY R/W LINE SD HWY + IN BK 1941 PG 475 TH N 16 DEG 03 MIN E, 303.81 FT TH N 12 DEG 03 MIN E, 204.4 FT TO E-W C/L SEC 18 TH N 89 DEG 21 MIN E ALG SD C/L 180 FT ML TO TRUE PT BEG. FORMERLY #04181-00-014		<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>1992 10/21/1998</td> <td>1,955.10</td> <td>5,793.69</td> <td>51.25</td> <td>7,780.04</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	1992 10/21/1998	1,955.10	5,793.69	51.25	7,780.04																		
Year	Date	Original	Interest	Fees	Total																												
Cert	1992 10/21/1998	1,955.10	5,793.69	51.25	7,780.04																												
<b>418000014</b> <b>Book: 998 Cert: 2021</b> <span style="float: right;"><b>7,236.12</b></span> SUBURBAN LAND CO 18-20-64 CONTG 2.1A IN THE NE4, SE4 SEC 18-20-64: BEG INTER OF NLY R/W LINE OF COLO HWY NO 47 (COLLEGE RD) AS PRESENTLY LOC ++ THE WLY R/W LINE OF COLO + S RR (NOW ABANDONED) TH N 12 DEG 02 MIN E ALG SD WLY LINE OF SD C-SRR R/W TO THE E-W C/L SD SEC 18 TH S 89 DEG 21 MIN E ALG SD C/L 38.08 FT TO A PT ON EXISTING FLOOD PLAIN DIST SD PT BEING TRU PT BEG TH SLY ALG FLOOD PLAIN DIST 2 COURSES: S 12 DEG 03 MIN W, 204.4 FT S 16 DEG 03 MIN W, 299.96 FT TO PT ON NLY R/W COLO HWY NO 47 (COLLEGE RD) TH N 84 DEG 18 MIN E, 190 FT ML TO INTER OF NLY R/W LINE SD HWY + IN BK 1941 PG 475 TH N 16 DEG 03 MIN E, 303.81 FT TH N 12 DEG 03 MIN E, 204.4 FT TO E-W C/L SEC 18 TH N 89 DEG 21 MIN E ALG SD C/L 180 FT ML TO TRUE PT BEG. FORMERLY #04181-00-014		<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>1993 10/21/1998</td> <td>1,795.21</td> <td>5,385.66</td> <td>51.25</td> <td>7,232.12</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	1993 10/21/1998	1,795.21	5,385.66	51.25	7,232.12																		
Year	Date	Original	Interest	Fees	Total																												
Cert	1993 10/21/1998	1,795.21	5,385.66	51.25	7,232.12																												
<b>418000014</b> <b>Book: 998 Cert: 2022</b> <span style="float: right;"><b>6,688.21</b></span> SUBURBAN LAND CO 18-20-64 CONTG 2.1A IN THE NE4, SE4 SEC 18-20-64: BEG INTER OF NLY R/W LINE OF COLO HWY NO 47 (COLLEGE RD) AS PRESENTLY LOC ++ THE WLY R/W LINE OF COLO + S RR (NOW ABANDONED) TH N 12 DEG 02 MIN E ALG SD WLY LINE OF SD C-SRR R/W TO THE E-W C/L SD SEC 18 TH S 89 DEG 21 MIN E ALG SD C/L 38.08 FT TO A PT ON EXISTING FLOOD PLAIN DIST SD PT BEING TRU PT BEG TH SLY ALG FLOOD PLAIN DIST 2 COURSES: S 12 DEG 03 MIN W, 204.4 FT S 16 DEG 03 MIN W, 299.96 FT TO PT ON NLY R/W COLO HWY NO 47 (COLLEGE RD) TH N 84 DEG 18 MIN E, 190 FT ML TO INTER OF NLY R/W LINE SD HWY + IN BK 1941 PG 475 TH N 16 DEG 03 MIN E, 303.81 FT TH N 12 DEG 03 MIN E, 204.4 FT TO E-W C/L SEC 18 TH N 89 DEG 21 MIN E ALG SD C/L 180 FT ML TO TRUE PT BEG. FORMERLY #04181-00-014		<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>1997 10/21/1998</td> <td>54.87</td> <td>309.53</td> <td>51.25</td> <td>415.65</td> </tr> <tr> <td>SubTax</td> <td>1998 08/24/1999</td> <td>53.29</td> <td>163.22</td> <td>5.00</td> <td>221.51</td> </tr> <tr> <td>SubTax</td> <td>1999 08/22/2000</td> <td>51.11</td> <td>149.26</td> <td>5.00</td> <td>205.37</td> </tr> <tr> <td>SubTax</td> <td>2000 08/10/2001</td> <td>52.02</td> <td>143.69</td> <td>5.00</td> <td>200.71</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	1997 10/21/1998	54.87	309.53	51.25	415.65	SubTax	1998 08/24/1999	53.29	163.22	5.00	221.51	SubTax	1999 08/22/2000	51.11	149.26	5.00	205.37	SubTax	2000 08/10/2001	52.02	143.69	5.00	200.71
Year	Date	Original	Interest	Fees	Total																												
Cert	1997 10/21/1998	54.87	309.53	51.25	415.65																												
SubTax	1998 08/24/1999	53.29	163.22	5.00	221.51																												
SubTax	1999 08/22/2000	51.11	149.26	5.00	205.37																												
SubTax	2000 08/10/2001	52.02	143.69	5.00	200.71																												
<b>419309003</b> <b>Book: 2015 Cert: 1670</b> <span style="float: right;"><b>1,992.72</b></span> LYNCH HOLDCO LLC 107 E 24TH ST THAT PART VAC SANTA FE AVE ADDJ ON W LOTS 13 TO 16 INC BLK 19 GARFIELD PL + ADJ ON W FRAC BLK 80 BARNOLLAR + LOWTHERS ADD (LIES BETW SD LOTS + HWY R/W) ALSO W 110.112 FT OF W 1/2 OF BLK 80 BARNOLLAR + LOWTHERS ADD W 110.112 FT OF LOTS 9 TO 16 INC: S 24.27 FT OF W 110.12 FT OF LOT 8 + ALSO ALL THAT POR LOT 8 LYING N OF W 24.47 FT THEREOF + W OF W R/W D-RG + ALL LOTS 1 TO 7 INC LYING W OF W R/W D-RG ALL IN BLK 19 GARFIELD PL FORMERLY #04193-01-003		<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>2017 10/16/2018</td> <td>1,785.22</td> <td>180.79</td> <td>22.71</td> <td>1,988.72</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	2017 10/16/2018	1,785.22	180.79	22.71	1,988.72																		
Year	Date	Original	Interest	Fees	Total																												
Cert	2017 10/16/2018	1,785.22	180.79	22.71	1,988.72																												
<b>423080006</b> <b>Book: 2010 Cert: 2192</b> <span style="float: right;"><b>2,500.58</b></span> SANTISTEVAN LUCY B FRAC LOTS 29 THRU 32 INC LYING NELY OF HWY BLK 26 EASTWOOD RESUB		<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Year</th> <th>Date</th> <th>Original</th> <th>Interest</th> <th>Fees</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>Cert</td> <td>2009 10/21/2010</td> <td>82.11</td> <td>84.97</td> <td>14.00</td> <td>181.08</td> </tr> <tr> <td>SubTax</td> <td>2010 08/18/2011</td> <td>83.39</td> <td>70.77</td> <td>5.00</td> <td>159.16</td> </tr> <tr> <td>SubTax</td> <td>2011 08/20/2012</td> <td>78.56</td> <td>58.54</td> <td>5.00</td> <td>142.10</td> </tr> <tr> <td>SubTax</td> <td>2012 08/22/2013</td> <td>80.17</td> <td>51.14</td> <td>5.29</td> <td>136.60</td> </tr> </tbody> </table>		Year	Date	Original	Interest	Fees	Total	Cert	2009 10/21/2010	82.11	84.97	14.00	181.08	SubTax	2010 08/18/2011	83.39	70.77	5.00	159.16	SubTax	2011 08/20/2012	78.56	58.54	5.00	142.10	SubTax	2012 08/22/2013	80.17	51.14	5.29	136.60
Year	Date	Original	Interest	Fees	Total																												
Cert	2009 10/21/2010	82.11	84.97	14.00	181.08																												
SubTax	2010 08/18/2011	83.39	70.77	5.00	159.16																												
SubTax	2011 08/20/2012	78.56	58.54	5.00	142.10																												
SubTax	2012 08/22/2013	80.17	51.14	5.29	136.60																												

Please note that these parcels include Real Property, Mobile Homes, and Oil, Gas & Mineral Rights

## Python Code (20190712\_PuebloCountyPDF.py)

```
# Extract back years from Pueblo txt file
# Accumatch
# by Michael Keller
# July 9, 2019
# requires pdfplumber, re, datetime

# libraries
import pdfplumber
import re
from datetime import datetime

# start timer
start = datetime.now()

# booleans
#buyerFound = False
accountFound = False

# Messages
messageRun = 'Reading input file . . .'
messageSaveBeg = 'Saving file . . .'
messageEOL = '. . . end of line'

# Variables
count = 0
regAccounts = '\\d{9,10}'
oldAccount = ''
currentAccount = ''
startLineCert = 'Cert'
startLineSubTax = 'SubTax'
comma = ','
spaces = '\\s{2,}'
space = ' '
nospace = ''
outString = []
```

```

outputFile = 'PuebloCountyHeldProperties.csv'

_return = '\r'

colHeadingsFull = 'Account,Year,Date,Original,Interest,Fees,Total'

pdf = pdfplumber.open("countyheldadvert.pdf")
fileOut = open(outputFile,"w")

# Add column headings to first line in file
fileOut.write(colHeadingsFull + _return)

print(messageRun)

for page in pdf.pages:
    p1 = pdf.pages[count]
    count = count + 1
    pltext = p1.extract_tables()
    text = pltext[0][1]
    body = text[0].split('\n') + text[1].split('\n')

    for x in body:
        # Check if line starts with 'Cert' or 'SubTax'
        if((x[0:4] == startLineCert) or (x[0:6] == startLineSubTax)):
            carp = re.sub(spaces, space, x)
            # Take out commas in amounts
            carp = carp.replace(comma,nospace)
            #print('taxline found...')
            carp2 = carp.rsplit(' ',7)
            # Make sure year is 2015 or newer
            if(int(carp2[1]) >= 2015):
                # Add only the last three items or amounts found from the split
                outString.append(str(currentAccount) + comma + str(carp2[1]) + comma +
str(carp2[2]) + comma + str(carp2[3]) + comma + str(carp2[4]) + comma + str(carp2[5]) + comma +
str(carp2[6]))

            # Search for Account Number in line read from PDF
            accountFound = re.search(regAccounts,x)

```

```
# Only run if Account is found in line
if(accountFound is not None):
    oldAccount = x[accountFound.start():accountFound.end()]
    if(currentAccount != oldAccount):
        currentAccount = oldAccount

##     if(count == 2):
##         break

#PDF File Close
pdf.close()

print(messageSaveBeg)

# Now dump string array into file
for lines in outString:
    fileOut.write(lines + _return)

print(messageEOL)

#fileIn.close()
fileOut.close()

#end timer
print('Runtime: ' + str(datetime.now() - start))

# EOF
```

CSV Output

	A	B	C	D	E	F	G
1	Account	Year	Date	Original	Interest	Fees	Total
2	418000014	2015	8/31/2016	56.57	25.86	5.2	87.63
3	418000014	2016	8/31/2017	56.77	17.3	5.2	79.27
4	418000014	2017	8/29/2018	56.81	8.65	5.28	70.74
5	419309003	2017	10/16/2018	1785.22	180.79	22.71	1988.72
6	428308006	2015	8/31/2016	79.22	25.29	5.3	109.81
7	428308006	2016	8/31/2017	122.02	25.42	1068.02	1215.46
8	428308006	2017	8/29/2018	79.58	8.46	5.4	93.44
9	428311008	2017	10/16/2018	0.95	1.5	14	16.45
10	428311009	2015	8/31/2016	0.93	1.78	5	7.71
11	428311009	2016	8/31/2017	0.94	1.19	5	7.13
12	428311009	2017	8/29/2018	0.94	0.59	5	6.53
13	428313015	2015	8/31/2016	47.18	25.05	5.18	77.41
14	428313015	2016	8/31/2017	47.37	16.76	5.18	69.31
15	428313015	2017	8/29/2018	47.39	8.38	5.24	61.01
16	428407010	2015	8/31/2016	94.25	35.73	5.34	135.32
17	428407010	2016	8/31/2017	309.91	75.58	5387.76	5773.25
18	428407010	2017	8/29/2018	94.68	11.96	5.46	112.1
19	428412011	2017	10/16/2018	57.35	7.16	14.28	78.79
20	429118002	2016	10/17/2017	0.86	3	14	17.86
21	429118002	2017	8/29/2018	0.84	0.64	5	6.48
22	429122040	2015	8/31/2016	7.14	5.1	5	17.24
23	429122040	2016	8/31/2017	7.18	3.41	5	15.59
24	429122040	2017	8/29/2018	7.18	1.71	5	13.89
25	429409007	2015	8/31/2016	141.44	43.97	5.52	190.93
26	429409007	2016	8/31/2017	165.48	34.12	592.9	792.5
27	429409007	2017	8/29/2018	198.73	20.39	1422.66	1641.78
28	429415007	2015	8/31/2016	53.86	28.25	5.2	87.31

Michael Keller  
Accumatch  
20190712\_PuebloCountyPDF  
July 17, 2019