

# real\_estate\_data\_2011\_vs\_2016

October 20, 2017

## 1 Compare real estate valuations and tax bills FY2011 vs FY2016

E.Quinn 10/17/2017

```
In [1]: import re
import numpy as np
import scipy as sc
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [109]: tot2016 = pd.read_csv("../EG_RE_Tax_Roll_assessed_12_31_2014_Totals.csv")
tot2016 = tot2016.rename(columns={'total_tax': 'total_tax2016'})
print(tot2016.shape)
print(tot2016.head())
```

```
(5375, 2)
   account  total_tax2016
0  100402392         7629.70
1  100003050        10073.47
2  100403672         5579.01
3  100009117         8538.27
4  100403456         1790.21
```

### 1.1 Read the FY2016 tax roll - assessed values on 12/31/2014

```
In [110]: tr2016a = pd.read_csv("../EG_RE_Tax_Roll_assessed_12_31_2014.csv")
tr2016a = tr2016a.rename(columns={'valuation': 'valuation2016', 'exemption': 'exemption2016',
                                   'taxbill': 'taxbill2016', 'address': 'address_2016'})
print(tr2016a.shape)
tr2016a.head()
```

```
(5569, 9)
```

```
Out[110]:
```

	account	parcel cat	state code	type \
0	100402392 079-012-101-0000	RE	1	ONE FAMI
1	100003050 044-011-418-0000	RE	1	ONE FAMI
2	100403672 054-007-187-0000	RE	1	ONE FAMI
3	100009117 013-019-006-0000	RE	33	FFOP
4	100403456 080-012-080-0000	RE	13	VACANT R

  

	address_2016	valuation2016	exemption2016	taxbill2016
0	1855 DIVISION ROAD	322200	0	7629.70
1	125 TAMARACK DRIVE	425400	0	10073.47
2	145 SOUTH PIERCE ROAD	235600	0	5579.01
3	145 SHIPPEE ROAD	360569	0	8538.27
4	DIVISION ROAD	75600	0	1790.21

```
In [111]: tot2016['total_tax2016'].sum()
```

```
Out[111]: 51480402.20999988
```

```
In [112]: tr2016 = tr2016a.merge(tot2016,left_on='account',right_on='account',how="right")
print(tr2016.shape)
tr2016.head()
```

```
(5941, 10)
```

```
Out[112]:
```

	account	parcel cat	state code	type \
0	100402392 079-012-101-0000	RE	1	ONE FAMI
1	100003050 044-011-418-0000	RE	1	ONE FAMI
2	100403672 054-007-187-0000	RE	1	ONE FAMI
3	100009117 013-019-006-0000	RE	33	FFOP
4	100403456 080-012-080-0000	RE	13	VACANT R

  

	address_2016	valuation2016	exemption2016	taxbill2016 \
0	1855 DIVISION ROAD	322200	0	7629.70
1	125 TAMARACK DRIVE	425400	0	10073.47
2	145 SOUTH PIERCE ROAD	235600	0	5579.01
3	145 SHIPPEE ROAD	360569	0	8538.27
4	DIVISION ROAD	75600	0	1790.21

  

	total_tax2016
0	7629.70
1	10073.47
2	5579.01
3	8538.27
4	1790.21

```
In [113]: trgrouped = tr2016.groupby(['account','parcel','state code','address_2016'])
print(len(trgrouped))
```

5569

```
In [114]: trg = trgrouped.agg({'taxbill2016' : np.sum, 'total_tax2016': np.max, 'exemption2016':  
trg.head()
```

```
Out[114]:
```

account	parcel	state	code	address_2016	taxbill2016 \
100000007	023-017-085-0000	1		15 CANONICUS TRAIL	4987.01
100000013	074-004-048-0000	1		137 KNOLLWOOD AVENUE	5642.94
100000026	034-015-160-0000	1		75 FALCON CIRCLE	11252.74
100000031	034-015-275-0000	1		8 RIVER RUN	13705.98
100000032	038-011-481-0000	1		60 JEFFERSON DRIVE	9670.91

  

account	parcel	state	code	address_2016	total_tax2016 \
100000007	023-017-085-0000	1		15 CANONICUS TRAIL	4987.01
100000013	074-004-048-0000	1		137 KNOLLWOOD AVENUE	5642.94
100000026	034-015-160-0000	1		75 FALCON CIRCLE	10401.74
100000031	034-015-275-0000	1		8 RIVER RUN	13705.98
100000032	038-011-481-0000	1		60 JEFFERSON DRIVE	9102.91

  

account	parcel	state	code	address_2016	valuation2016 \
100000007	023-017-085-0000	1		15 CANONICUS TRAIL	210600
100000013	074-004-048-0000	1		137 KNOLLWOOD AVENUE	238300
100000026	034-015-160-0000	1		75 FALCON CIRCLE	475200
100000031	034-015-275-0000	1		8 RIVER RUN	578800
100000032	038-011-481-0000	1		60 JEFFERSON DRIVE	408400

  

account	parcel	state	code	address_2016	exemption2016
100000007	023-017-085-0000	1		15 CANONICUS TRAIL	0
100000013	074-004-048-0000	1		137 KNOLLWOOD AVENUE	0
100000026	034-015-160-0000	1		75 FALCON CIRCLE	851
100000031	034-015-275-0000	1		8 RIVER RUN	0
100000032	038-011-481-0000	1		60 JEFFERSON DRIVE	568

```
In [115]: trg2 = trg.reset_index(level='account',drop=True)  
trg3 = trg2.reset_index()  
print(trg3.shape)  
trg3.head()
```

(5569, 7)

```
Out[115]:
```

	parcel	state	code	address_2016	taxbill2016 \
0	023-017-085-0000	1		15 CANONICUS TRAIL	4987.01
1	074-004-048-0000	1		137 KNOLLWOOD AVENUE	5642.94

2	034-015-160-0000	1	75 FALCON CIRCLE	11252.74
3	034-015-275-0000	1	8 RIVER RUN	13705.98
4	038-011-481-0000	1	60 JEFFERSON DRIVE	9670.91

	total_tax2016	valuation2016	exemption2016
0	4987.01	210600	0
1	5642.94	238300	0
2	10401.74	475200	851
3	13705.98	578800	0
4	9102.91	408400	568

```
In [116]: sf2016 = trg3.loc[trg3['state code']==1]
print(sf2016.shape)
sf2016.head()
```

(3901, 7)

```
Out[116]:
```

	parcel	state code	address_2016	taxbill2016	\
0	023-017-085-0000	1	15 CANONICUS TRAIL	4987.01	
1	074-004-048-0000	1	137 KNOLLWOOD AVENUE	5642.94	
2	034-015-160-0000	1	75 FALCON CIRCLE	11252.74	
3	034-015-275-0000	1	8 RIVER RUN	13705.98	
4	038-011-481-0000	1	60 JEFFERSON DRIVE	9670.91	

	total_tax2016	valuation2016	exemption2016
0	4987.01	210600	0
1	5642.94	238300	0
2	10401.74	475200	851
3	13705.98	578800	0
4	9102.91	408400	568

```
In [117]: sf2016.describe()
```

```
Out[117]:
```

	state code	taxbill2016	total_tax2016	valuation2016	exemption2016
count	3901.0	3901.000000	3901.000000	3.901000e+03	3901.000000
mean	1.0	10188.236462	10111.676855	4.310374e+05	1019.286593
std	0.0	4562.409181	4854.191530	1.916427e+05	11192.786816
min	1.0	1311.400000	0.000000	1.069000e+05	0.000000
25%	1.0	6857.730000	6642.240000	2.900000e+05	0.000000
50%	1.0	9412.800000	9233.720000	3.975000e+05	0.000000
75%	1.0	12223.620000	12361.200000	5.162000e+05	0.000000
max	1.0	53078.720000	53078.720000	2.241500e+06	262563.000000

## 1.2 Read FY2011 tax roll - assessed values on 12/31/2009

```
In [123]: tot2011 = pd.read_csv("../EG_RE_Tax_Roll_assessed_12_31_2009_Totals.csv")
tot2011 = tot2011.rename(columns={'total_tax': 'total_tax2011'})
print(tot2011.shape)
tot2011.head()
```

(5174, 2)

```
Out[123]:
```

	account	total_tax2011
0	100402392	6908.56
1	100003050	8678.43
2	100009117	8798.07
3	100039157	15342.03
4	100401774	94207.50

### 1.3 Read the 2011 tax roll - assessed values on 12/31/2010

```
In [124]: tr2011a = pd.read_csv("../EG_RE_Tax_Roll_assessed_12_31_2009.csv")
tr2011a = tr2011a.rename(columns={'valuation': 'valuation2011', 'exemption': 'exemption2011',
                                   'taxbill': 'taxbill2011', 'bill_no': 'bill_no2011', 'address': 'address2011'})
print(tr2011a.shape)
tr2011a.head()
```

(10400, 10)

```
Out[124]:
```

	account	bill_no2011	parcel cat	state code	type \
0	100402392	4739	079-012-101-0000	RE	1 ONE FAMI
1	100402392	4739	079-012-101-0000	RE	1 ONE FAMI
2	100003050	1093	044-011-418-0000	RE	1 ONE FAMI
3	100003050	1093	044-011-418-0000	RE	1 ONE FAMI
4	100009117	3114	013-019-006-0000	RE	33 FFOP

  

	address_2011	valuation2011	exemption2011	taxbill2011
0	1855 DIVISION ROA	363800	0	633.01
1	1855 DIVISION ROA	363800	0	6275.55
2	125 TAMARACK DRIV	457000	0	795.18
3	125 TAMARACK DRIV	457000	0	7883.25
4	145 SHIPPEE ROAD	463300	0	806.14

```
In [125]: tr2011 = tr2011a.merge(tot2011, left_on='account', right_on='account', how="right")
print(tr2011.shape)
tr2011.head()
```

(10638, 11)

```
Out[125]:
```

	account	bill_no2011	parcel cat	state code	type \
0	100402392	4739	079-012-101-0000	RE	1 ONE FAMI
1	100402392	4739	079-012-101-0000	RE	1 ONE FAMI
2	100003050	1093	044-011-418-0000	RE	1 ONE FAMI
3	100003050	1093	044-011-418-0000	RE	1 ONE FAMI
4	100009117	3114	013-019-006-0000	RE	33 FFOP

	address_2011	valuation2011	exemption2011	taxbill2011	total_tax2011
0	1855 DIVISION ROA	363800	0	633.01	6908.56
1	1855 DIVISION ROA	363800	0	6275.55	6908.56
2	125 TAMARACK DRIV	457000	0	795.18	8678.43
3	125 TAMARACK DRIV	457000	0	7883.25	8678.43
4	145 SHIPPEE ROAD	463300	0	806.14	8798.07

```
In [126]: trgrouped = tr2011.groupby(['account','parcel','state code'])
          print(len(trgrouped))
```

5165

```
In [131]: trg = trgrouped.agg({'taxbill2011' : np.sum, 'total_tax2011': np.max, 'exemption2011':
                             trg.head())
```

```
Out[131]:
```

account	parcel	state code	exemption2011	valuation2011	\
100000007	023-017-085-0000	1	0	207800	
100000011	074-004-007-0000	1	65186	296000	
100000013	074-004-048-0000	1	29788	291200	
100000026	034-015-160-0000	1	0	517200	
100000031	034-015-275-0000	1	0	610200	

  

account	parcel	state code	total_tax2011	taxbill2011
100000007	023-017-085-0000	1	3946.12	3946.12
100000011	074-004-007-0000	1	4383.16	4383.16
100000013	074-004-048-0000	1	4964.22	4964.22
100000026	034-015-160-0000	1	9821.63	9821.63
100000031	034-015-275-0000	1	11587.70	11587.70

```
In [132]: trg2 = trg.reset_index(level='account',drop=True)
          trg3 = trg2.reset_index()
          print(trg3.shape)
          trg3.head()
```

(5165, 6)

```
Out[132]:
```

	parcel	state code	exemption2011	valuation2011	total_tax2011	\
0	023-017-085-0000	1	0	207800	3946.12	
1	074-004-007-0000	1	65186	296000	4383.16	
2	074-004-048-0000	1	29788	291200	4964.22	
3	034-015-160-0000	1	0	517200	9821.63	
4	034-015-275-0000	1	0	610200	11587.70	

taxbill2011

```

0      3946.12
1      4383.16
2      4964.22
3      9821.63
4     11587.70

```

```

In [133]: sf2011 = trg3.loc[trg3['state code']==1]
          print(sf2011.shape)
          sf2011.head()

```

(3865, 6)

```

Out[133]:
           parcel  state code  exemption2011  valuation2011  total_tax2011  \
0  023-017-085-0000           1             0         207800         3946.12
1  074-004-007-0000           1         65186         296000         4383.16
2  074-004-048-0000           1         29788         291200         4964.22
3  034-015-160-0000           1             0         517200         9821.63
4  034-015-275-0000           1             0         610200        11587.70

           taxbill2011
0           3946.12
1           4383.16
2           4964.22
3           9821.63
4          11587.70

```

```

In [134]: sf2011.describe()

```

```

Out[134]:
           state code  exemption2011  valuation2011  total_tax2011  taxbill2011
count           3865.0    3.865000e+03    3.865000e+03    3865.000000    3865.000000
mean              1.0    1.310182e+04    4.674164e+05    8748.063806    8914.540877
std              0.0    7.978753e+04    2.045106e+05    4074.067578    13982.296689
min              1.0    0.000000e+00    1.120000e+05         0.000000         0.000000
25%              1.0    0.000000e+00    3.199000e+05    5810.750000    5791.950000
50%              1.0    0.000000e+00    4.314000e+05    8061.260000    8019.700000
75%              1.0    0.000000e+00    5.620000e+05   10714.160000   10649.590000
max              1.0    4.617140e+06    2.110200e+06   40072.700000   840978.850000

```

```

In [169]: sf_2016_2011 = sf2011.merge(sf2016,how="inner")
          print(sf_2016_2011.shape)
          sf_2016_2011.head()

```

(3818, 11)

```

Out[169]:
           parcel  state code  exemption2011  valuation2011  total_tax2011  \
0  023-017-085-0000           1             0         207800         3946.12
1  074-004-007-0000           1         65186         296000         4383.16

```

2	074-004-048-0000	1	29788	291200	4964.22
3	034-015-160-0000	1	0	517200	9821.63
4	034-015-275-0000	1	0	610200	11587.70

	taxbill2011	address_2016	taxbill2016	total_tax2016	\
0	3946.12	15 CANONICUS TRAIL	4987.01	4987.01	
1	4383.16	196 KNOLLWOOD AVENUE	6552.26	6552.26	
2	4964.22	137 KNOLLWOOD AVENUE	5642.94	5642.94	
3	9821.63	75 FALCON CIRCLE	11252.74	10401.74	
4	11587.70	8 RIVER RUN	13705.98	13705.98	

	valuation2016	exemption2016
0	210600	0
1	276700	0
2	238300	0
3	475200	851
4	578800	0

```
In [170]: sf_2016_2011['tax_bill_pct_chg'] = 100*((sf_2016_2011['total_tax2016']/sf_2016_2011['t
sf_2016_2011.head()
```

```
Out[170]:
```

	parcel	state code	exemption2011	valuation2011	total_tax2011	\
0	023-017-085-0000	1	0	207800	3946.12	
1	074-004-007-0000	1	65186	296000	4383.16	
2	074-004-048-0000	1	29788	291200	4964.22	
3	034-015-160-0000	1	0	517200	9821.63	
4	034-015-275-0000	1	0	610200	11587.70	

	taxbill2011	address_2016	taxbill2016	total_tax2016	\
0	3946.12	15 CANONICUS TRAIL	4987.01	4987.01	
1	4383.16	196 KNOLLWOOD AVENUE	6552.26	6552.26	
2	4964.22	137 KNOLLWOOD AVENUE	5642.94	5642.94	
3	9821.63	75 FALCON CIRCLE	11252.74	10401.74	
4	11587.70	8 RIVER RUN	13705.98	13705.98	

	valuation2016	exemption2016	tax_bill_pct_chg
0	210600	0	26.377556
1	276700	0	49.487128
2	238300	0	13.672239
3	475200	851	5.906453
4	578800	0	18.280418

```
In [171]: sf_2016_2011.describe()
```

```
Out[171]:
```

	state code	exemption2011	valuation2011	total_tax2011	taxbill2011	\
count	3818.0	3.818000e+03	3.818000e+03	3818.000000	3818.000000	
mean	1.0	1.312747e+04	4.698532e+05	8738.020592	8915.308683	
std	0.0	8.014756e+04	2.039789e+05	4058.619862	14057.518808	
min	1.0	0.000000e+00	1.403000e+05	0.000000	0.000000	



25%	1.0	0.000000e+00	3.222000e+05	5802.400000	5788.740000
50%	1.0	0.000000e+00	4.329500e+05	8048.915000	8017.580000
75%	1.0	0.000000e+00	5.632750e+05	10714.160000	10666.212500
max	1.0	4.617140e+06	2.110200e+06	40072.700000	840978.850000

	taxbill2016	total_tax2016	valuation2016	exemption2016	\
count	3818.000000	3818.000000	3.818000e+03	3818.000000	
mean	10136.006234	10039.499793	4.288489e+05	1039.542431	
std	4458.657937	4741.961085	1.872239e+05	11312.877472	
min	1311.400000	0.000000	1.069000e+05	0.000000	
25%	6855.952500	6628.030000	2.899250e+05	0.000000	
50%	9386.750000	9205.600000	3.964000e+05	0.000000	
75%	12169.150000	12230.127500	5.139000e+05	0.000000	
max	51830.780000	51830.780000	2.188800e+06	262563.000000	

	tax_bill_pct_chg
count	3817.000000
mean	inf
std	NaN
min	-100.000000
25%	6.279868
50%	13.319830
75%	20.947909
max	inf

```
In [172]: sf_2016_2011.to_csv("../EG_Real_Estate_Tax_Bill_Percent_Changes_2011_to_2016.csv",columns=['valuation2011','valuation2016','total_tax2011','total_tax2016','exemption2011','exemption2016','tax_bill_pct_chg'])
```

```
In [173]: sf_2016_2011.quantile([.01, .05, .10, .15, .20, .25, .30, .35, .40, .45, .50, .55, .60, .85, .90, .95, .955, .957, .958, .959])
```

```
Out[173]:
```

	state	code	exemption2011	valuation2011	total_tax2011	taxbill2011	\
0.010	1.0		0.00	191436.0	2666.89420	2666.89420	
0.050	1.0		0.00	234600.0	3944.23000	3933.13150	
0.100	1.0		0.00	262350.0	4625.66000	4612.73300	
0.150	1.0		0.00	281210.0	5045.83000	5038.39800	
0.200	1.0		0.00	299940.0	5394.68000	5384.42600	
0.250	1.0		0.00	322200.0	5802.40000	5788.74000	
0.300	1.0		0.00	342510.0	6243.93300	6234.46900	
0.350	1.0		0.00	365000.0	6648.40000	6627.32000	
0.400	1.0		0.00	387360.0	7132.26200	7088.59000	
0.450	1.0		0.00	411700.0	7602.27000	7560.21900	
0.500	1.0		0.00	432950.0	8048.91500	8017.58000	
0.550	1.0		0.00	454835.0	8503.16000	8468.57350	
0.600	1.0		0.00	474900.0	8906.69000	8872.13000	
0.650	1.0		0.00	498900.0	9404.04000	9363.97000	
0.700	1.0		0.00	526790.0	10005.64900	9956.27000	

0.750	1.0	0.00	563275.0	10714.16000	10666.21250
0.800	1.0	5610.00	597380.0	11352.49000	11321.46000
0.850	1.0	29788.00	642080.0	12188.76600	12133.48500
0.900	1.0	50292.00	740020.0	14097.98700	14039.12000
0.950	1.0	89364.00	874420.0	16665.72350	16593.27850
0.955	1.0	90682.35	888500.0	16876.74250	16860.99965
0.957	1.0	94974.00	890460.7	16933.04252	16903.90440
0.958	1.0	94974.00	892511.6	16961.98020	16939.34112
0.959	1.0	94974.00	893550.9	16999.89057	16964.72570

	taxbill2016	total_tax2016	valuation2016	exemption2016	\
0.010	4048.51990	3272.20140	186351.0	0.0	
0.050	5152.77000	4636.45800	219710.0	0.0	
0.100	5637.49900	5265.01100	238870.0	0.0	
0.150	6013.65350	5706.47950	254655.0	0.0	
0.200	6458.48400	6129.32800	273500.0	0.0	
0.250	6855.95250	6628.03000	289925.0	0.0	
0.300	7279.46700	7066.17400	307600.0	0.0	
0.350	7811.79350	7561.02000	329995.0	0.0	
0.400	8345.78000	8153.02000	352580.0	0.0	
0.450	8898.11400	8671.26650	375765.0	0.0	
0.500	9386.75000	9205.60000	396400.0	0.0	
0.550	9863.54950	9739.58000	416535.0	0.0	
0.600	10364.74000	10323.71600	437700.0	0.0	
0.650	10883.44850	10871.60850	459605.0	0.0	
0.700	11543.76300	11543.76300	487490.0	0.0	
0.750	12169.15000	12230.12750	513900.0	0.0	
0.800	12999.37200	13123.92800	548960.0	300.0	
0.850	14005.41650	14095.99450	591445.0	678.0	
0.900	15740.57100	16077.27200	664720.0	961.0	
0.950	18469.45050	18734.31500	779960.0	1712.0	
0.955	18734.91850	19124.38240	791170.5	1813.0	
0.957	18905.49037	19201.37792	798373.8	1813.0	
0.958	18947.99582	19267.06224	800168.6	1813.0	
0.959	18991.37422	19351.30919	802000.6	1813.0	

	tax_bill_pct_chg
0.010	-18.298043
0.050	-5.231021
0.100	-0.182993
0.150	2.791477
0.200	4.805443
0.250	6.279868
0.300	7.670248
0.350	9.376854
0.400	10.676640
0.450	12.009017
0.500	13.319830

0.550	14.593974
0.600	15.969958
0.650	17.357235
0.700	18.990851
0.750	20.947909
0.800	22.847154
0.850	25.749952
0.900	30.459224
0.950	42.454060
0.955	45.909354
0.957	48.511371
0.958	49.591663
0.959	50.525023

In [ ]: