Data processing for PM data
GSA project

January 6, 2016

Contents

1 Introduction

The document records the process of plotting the new FY data

2 Checks

2.1 Sheets

sheet names:

```
# FY 13
[u'R1', u'R2', u'R3', u'R4', u'R5', u'R6', u'R7', u'R8', u'R9', u'R10', u'R11']
# FY 14
[u'r1 fy14', u'r2 fy14', u'r3 fy14', u'r4 fy14', u'r5 fy14', u'r6 fy14', u'r7 fy14', u'r8 fy14
# FY 15
[u'R1 FY15', u'R2 FY15', u'R3 FY15', u'R4 FY15', u'R5 FY15', u'R6 FY15', u'R7 FY15', u'R8 FY15
```

The names of sheets are not standard but the order of region is, so use index of sheets when reading. For faster processing, split excel files to single sheet csv files

2.2 Number of building

Table 1: Number of buildings in each file

Region	FY13	FY14	FY15
1	76	74	74
2	167	165	154
3	80	79	78
4	127	126	130
5	160	150	146
6	43	40	40
7	133	131	129
8	57	56	56
9	75	74	72
10	62	59	59
11	119	130	117

2.3 Common buildings

There are 1012 buildings that has records for all three years.

There are 1141 buildings covered in one of the three files.

Table 2: Building record and its year			
Building Number	2013	2014	2015
NJ4666ZZ	0	0	1
NY7182ZZ	1	1	1
OR0053ZZ	1	1	1
IA0121ZZ	1	1	1
OH2433ZZ	1	1	1
IL2540ZZ	1	1	1
AZ0000WW	1	1	1
MA0050ZZ	1	1	1
TX2747ZZ	1	1	1
ME0009ZZ	1	0	0
FL3323ZZ	0	1	1

note: "1" denotes the building has records in the corresponding year of the column, "0" denotes there are no records for this building in that year