Stats 506, F20, Final Project

Zhilin He, zhilinhe@umich.edu

December 14, 2020

Link to GitHub repository

https://github.com/zhilinheobv/506-Final-Project

Introduction

Fuel oil (including kerosene and diesel) is a common source of energy. Although it is not as environmental friendly as natural gas, it is used for heating in many homes and buildings, and other industries like fueling for trucks and ships. Some industries aim to reduce the use of polluting fuel oil since it might cause air pollution and harm people's health[1]. Therefore, proper usage of fuel oil is an interesting topic to consider.

Many buildings use fuel oil as an energy source for various reasons. It's possible that buildings' ownership affects the choice of energy sources. In this project, we are trying to find out whether government-owned commercial buildings use more fuel oil than not government-owned buildings, and what fuel oil is used for in different buildings.

We are going to use various different statistical methods including Pearson's chi-squared test and two sample t-test to identify the difference between goverment-owned buildings and not government-owned ones.

Data

This project uses the 2012 CBECS data. It contains 1119 variables and 6720 sampled observations. The data contains all aspects of information concerning energy consumption about commercial buildings all across the US.

We choose Detailed descriptions for the chosen variables are shown in Table 1 below.

Variable Name	Variable Description
PUBID (id)	Building ID
REGION	Census region
CENDIV (div)	Census division
GOVOWN (gov)	Government owned (1) or not (2)
FKUSED	Uses fuel oil/diesel/kerosene (1) or not (2)
FKTYPE	Fuel oil, diesel or kerosene used
FKHTBTU	Fuel oil^ used for heating (thousand BTU)
FKCLBTU	Fuel oil used for cooling (thousand BTU)
FKWTBTU	Fuel oil used for water heating (thousand BTU)
FKCKBTU	Fuel oil used for cooking (thousand BTU)
FKOTBTU	Fuel oil used for other reasons (thousand BTU)
FKBTU	Annual fuel oil consumption (thousand BTU)
FKEXP	Annual fuel oil expenditure ($)
FINALWT	Full sample building weight

Variable Name	Variable Description
FINALWT1-197	Final replicate weights

^ Including diesel and kerosene. Same with "Fuel oil" below.

Table 1. Variable descriptions

Methods

We use the jackknife method to estimate standard error. This method uses replicate weights to repeatedly estimate the statistics of interest and calculate the differences between these estimates and the full-sample estimate[2].

We also use Pearson's chi-squared test to verify whether the effect of building ownership is significant for fuel oil usage.

Analysis and Results

Data Cleaning

Firstly, we take a look at the data. Data summary is shown below.

id	div	gov	fo_used	fo_type	heating
Min. : 1	South Atlantic :1348	Goverment :1337	TRUE :1515	Fuel oil : 353	Min. : 0
1st Qu.:1681	Pacific :1113	Non-goverment:5383	FALSE:5205	Diesel : 782	1st Qu.: 0
Median :3360	East North Central: 900			Kerosene : 23	Median : 0
Mean :3360	West South Central: 852			More than one: 79	Mean : 394569
3rd Qu.:5040	Middle Atlantic : 787			Unknown : 278	3rd Qu.: 0
Max. :6720	West North Central: 559			NA's :5205	Max. :75555274
	(Other) :1161				NA's :5205

cooling	waterheating	cooking	other	total	expenditure
Min. : 0	Min. : 0	Min. : 0	Min. : 0	Min. : 0	Min. : 0
1st Qu.: 0	1st Qu.: 0	1st Qu.: 0	1st Qu.: 6805	1st Qu.: 33750	1st Qu.: 954
Median : 0	Median : 0	Median : 0	Median : 124553	Median : 208826	Median : 5374
Mean : 17896	Mean : 5680	Mean : 35505	Mean : 702369	Mean : 1156019	Mean : 27926
3rd Qu.: 0	3rd Qu.: 0	3rd Qu.: 0	3rd Qu.: 663206	3rd Qu.: 810147	3rd Qu.: 20295
Max. :10094242	Max. :796586	Max. :30923934	Max. :34166771	Max. :96306196	Max. :2041370
NA's :5205	NA's :5205	NA's :5205	NA's :5205	NA's :5205	NA's :5205

Table 2. Summary of variables

Fuel oil usage

We calculate the weighted frequency of fuel oil usage using the survey weights.

ownership	division	fuel oil used	fuel oil not used	percent used	LB used	UB used	LB not u
Goverment	New England	14031.772	21928.31	0.3902041	12590.067	15473.477	2476
Non-goverment	New England	100775.757	164899.36	0.3793195	89899.975	111651.539	17299

ownership	division	fuel oil used	fuel oil not used	percent used	LB used	UB used	LB not u
Goverment	Middle Atlantic	18537.168	28303.00	0.3957537	17703.126	19371.211	3001
Non-goverment	Middle Atlantic	83336.272	373671.66	0.1823519	77888.460	88784.083	38039
Goverment	East North Central	17556.040	82169.55	0.1760435	14774.439	20337.641	8524
Non-goverment	East North Central	28168.832	607353.11	0.0443239	27260.189	29077.474	62919
Goverment	West North Central	10038.200	61702.01	0.1399243	8978.305	11098.094	6548
Non-goverment	West North Central	21209.468	409282.07	0.0492680	19320.069	23098.868	42429
Goverment	South Atlantic	15893.248	104762.85	0.1317235	14989.815	16796.680	10941
Non-goverment	South Atlantic	53159.350	917568.99	0.0547623	50804.875	55513.825	94414
Goverment	East South Central	7662.804	59381.11	0.1142953	7251.888	8073.721	6180
Non-goverment	East South Central	21865.835	280725.32	0.0722620	18951.291	24780.380	28882
Goverment	West South Central	6747.545	94104.24	0.0669056	5938.566	7556.524	9987
Non-goverment	West South Central	25175.664	659756.56	0.0367564	22622.834	27728.495	69138
Goverment	Mountain	1919.560	58784.86	0.0316214	1746.692	2092.429	6402
Non-goverment	Mountain	6128.668	271611.53	0.0220662	4950.844	7306.491	28027
Goverment	Pacific	9037.605	163462.19	0.0523920	8135.849	9939.361	17236
Non-goverment	Pacific	25491.730	730936.21	0.0337001	23924.355	27059.105	77886

Table 3. Frequency table of buildings using oil by census division and ownership

From Table 3, we can see that the percentage of government buildings that use fuel oil is greater in all divisions, but the difference is more obvious in Middle Atlantic, East North Central, West North Central and South Atlantic divisions. We can use Pearson chi-squared test to verify whether this difference is statistically significant.

Division	Chi-squared statistic	p-value
New England	15.87111	$p < 0.001$
Middle Atlantic	11993.06659	$p < 0.001$
East North Central	25640.82270	$p < 0.001$
West North Central	8660.15322	$p < 0.001$
South Atlantic	10723.68072	$p < 0.001$
East South Central	1318.65061	$p < 0.001$
West South Central	2049.37892	$p < 0.001$
Mountain	195.51565	$p < 0.001$
Pacific	1370.74663	$p < 0.001$
Total	25577.63825	$p < 0.001$

Table 4. Result of chi-squared tests

It turns out the difference of fuel oil usage between government and non-government buildings is significant for all divisions. This is possibly due to the fact that the full sample size is large enough to identify this difference.

Consumption

We calculate the annual consumption of fuel oil among the buildings using fuel oil as an energy source. The consumptions for each specific use are also calculated. The graph for consumptions in different divisions is shown below, with estimated error bars:
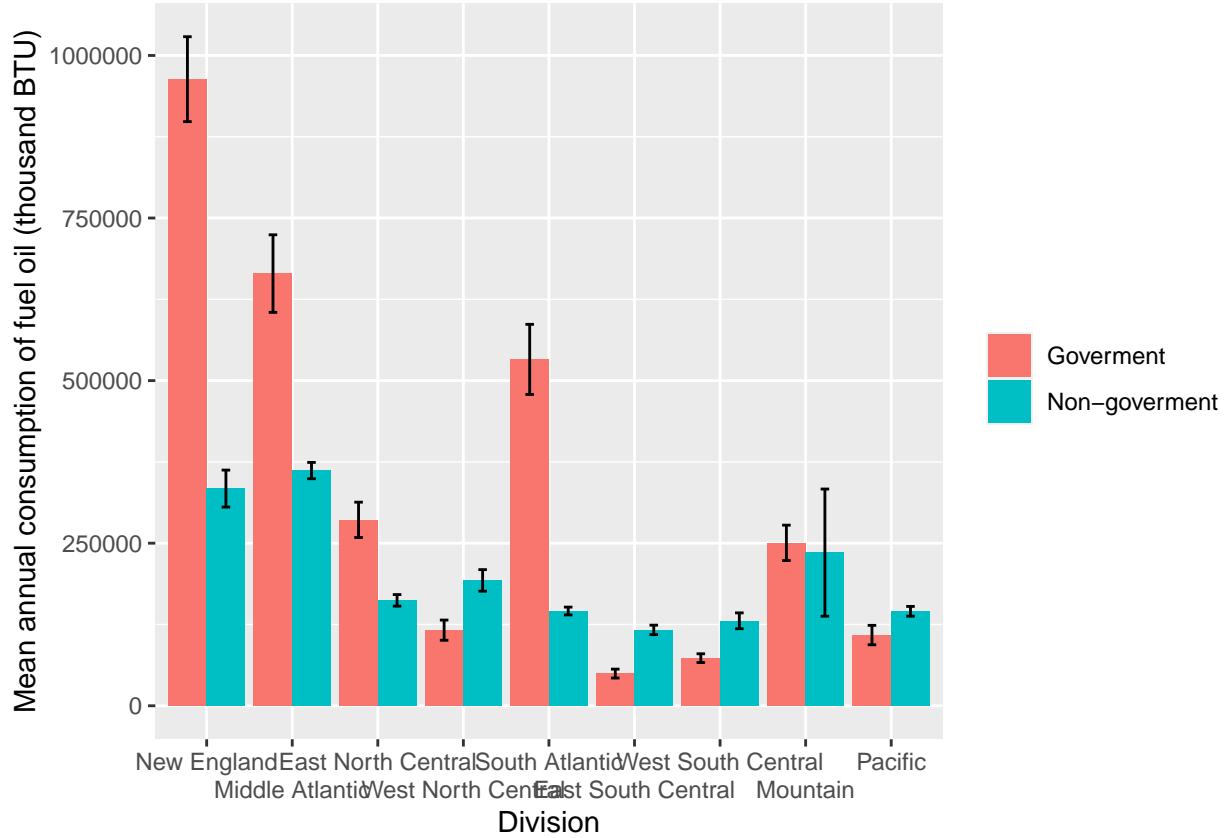
Figure 1. Mean annual fuel oil consumption for buildings by division and ownership

gov	div	mean	se	lower	upper
Goverment	New England	963511.69	33307.602	898229.99	1028793.39
Non-goverment	New England	333986.99	14512.312	305543.38	362430.60
Goverment	Middle Atlantic	664588.14	30426.772	604952.77	724223.52
Non-goverment	Middle Atlantic	361643.73	6389.122	349121.28	374166.18
Goverment	East North Central	285926.28	13854.153	258772.64	313079.92
Non-goverment	East North Central	162093.98	4539.531	153196.66	170991.29
Goverment	West North Central	116277.49	7922.396	100749.88	131805.10
Non-goverment	West North Central	192769.36	8448.660	176210.29	209328.43
Goverment	South Atlantic	532670.68	27502.782	478766.22	586575.14
Non-goverment	South Atlantic	145715.53	3064.887	139708.47	151722.60
Goverment	East South Central	49494.84	3491.005	42652.59	56337.08
Non-goverment	East South Central	116708.93	3711.652	109434.23	123983.64
Goverment	West South Central	73193.19	3447.682	66435.86	79950.52
Non-goverment	West South Central	130696.54	6207.395	118530.27	142862.81
Goverment	Mountain	250533.29	13886.984	223315.30	277751.28
Non-goverment	Mountain	235536.52	49878.258	137776.93	333296.11
Goverment	Pacific	108745.77	7627.133	93796.87	123694.68
Non-goverment	Pacific	145264.40	3842.659	137732.93	152795.88
Goverment	Total	422277.35	8729.799	405167.26	439387.45
Non-goverment	Total	249609.90	3144.955	243445.90	255773.90

Table 5. Mean annual fuel oil consumption for buildings

4

In New England, East North Central, Middle and South Atlantic divisions, government buildings consume significantly more fuel oil than non-government buildings on average. In Mountain Division, the difference is not significant. In other divisions, non-government buildings consume more fuel oil annually. Overall, because New England and Middle Atlantic divisions consume fuel oil the most in the US, government buildings consume significantly more fuel oil than other buildings annually on average.

Detailed uses

Fuel oil can be used in many ways in buildings. We consider the differences of uses between goverment buildings and non-government ones.

| gov	div	heating	cooling	waterheating	cooking	other
Goverment	New England	831720.04	0.0000	2832.7359868	0.00000	128958.91
Non-goverment	New England	299073.31	655.8435	1643.9363494	1726.87853	30887.02
Goverment	Middle Atlantic	551263.12	3611.3185	3050.3485278	0.00000	106663.35
Non-goverment	Middle Atlantic	254523.22	7640.7713	2090.4352639	5780.20675	91609.09
Goverment	East North Central	205466.86	1624.3409	14.2704173	0.00000	78820.81
Non-goverment	East North Central	81157.64	0.0000	6.9916443	0.00000	80929.35
Goverment	West North Central	53344.88	0.0000	2732.3831181	0.00000	60200.23
Non-goverment	West North Central	125394.44	0.0000	200.2502431	1078.22725	66096.44
Goverment	South Atlantic	241525.58	0.0000	2037.3908551	0.00000	289107.70
Non-goverment	South Atlantic	56405.43	1320.2462	5599.5459137	900.57979	81489.73
Goverment	East South Central	0.00	0.0000	12.1093298	0.00000	49482.73
Non-goverment	East South Central	74217.04	0.0000	9.8068558	0.00000	42482.09
Goverment	West South Central	1032.36	0.0000	0.0000000	0.00000	72160.83
Non-goverment	West South Central	12502.29	0.0000	18.8719399	0.00000	118175.37
Goverment	Mountain	0.00	0.0000	356.9030921	0.00000	250176.38
Non-goverment	Mountain	30745.62	271.2235	0.1660811	0.00000	204519.50
Goverment	Pacific	0.00	0.0000	818.3570389	1108.43920	106818.98
Non-goverment	Pacific	49790.53	0.0000	201.8280987	0.00000	95272.05
Goverment	Total	294581.59	941.2039	1622.1663039	98.76993	125033.63
Non-goverment	Total	171606.24	2120.6341	1773.3516725	1988.63459	72121.04

Table 6. Detailed uses of fuel oil
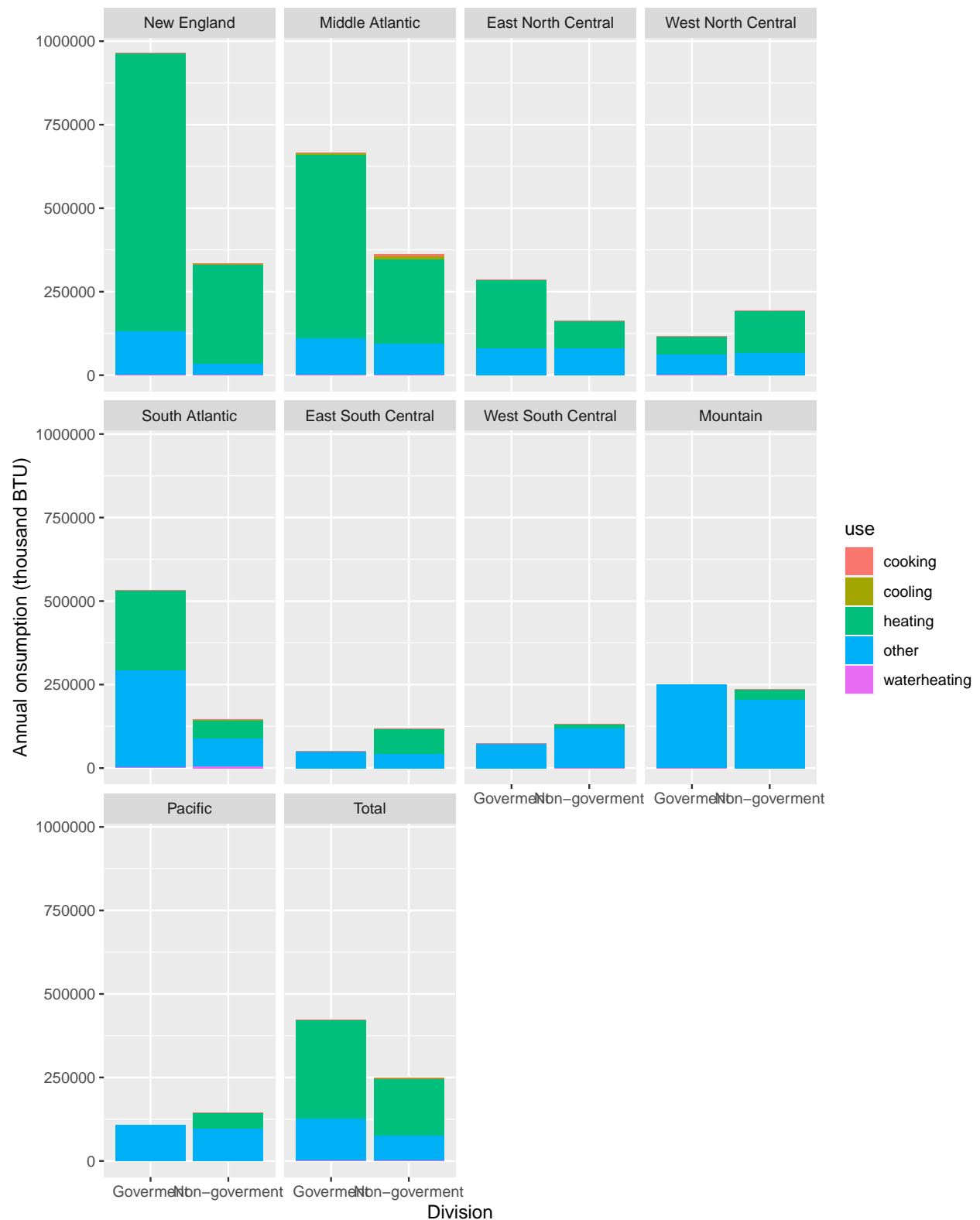
Figure 2. Stacked bar plots for different uses of fuel oil

From Figure 2 we can clearly see that uses of fuel oil mainly differs by census divisions, possibly due to climate reasons. For East South Central, West South Central, Mountain and Pacific divisions, heating is not the main use of fuel oil. In these divisions, government buildings almost only use fuel oil for other reasons

(including manufacturing and electricity generation). But non-government buildings still use some fuel oil for heating.

Type of fuel oil

In this project, "fuel oil" includes (heavy) fuel oil, kerosene and diesel. We consider the differences of type of fuel oil used in those buildings using fuel oil as an energy source.

gov	div	fo_typeFuel oil	fo_typeDiesel	fo_typeKerosene	fo_typeMore than one	fo_
Goverment	New England	10008.20926	1176.0464	1387.07659	1168.838533	
Non-goverment	New England	85696.44872	6175.2985	1387.07659	5148.016715	
Goverment	Middle Atlantic	11522.55998	1095.7556	1858.29285	457.483121	
Non-goverment	Middle Atlantic	39773.34814	15690.6690	11078.49751	3735.513553	
Goverment	East North Central	611.59269	4911.2823	4458.21006	317.358516	
Non-goverment	East North Central	12852.61609	6906.5392	178.39373	110.150013	
Goverment	West North Central	3231.51863	3756.5780	1484.51505	19.690263	
Non-goverment	West North Central	10082.69960	6438.3027	1807.46287	1755.892772	
Goverment	South Atlantic	1509.23496	9933.0040	0.00000	188.300041	
Non-goverment	South Atlantic	22302.43625	17449.3251	4913.02064	129.463685	
Goverment	East South Central	0.00000	5581.0438	0.00000	5.982684	
Non-goverment	East South Central	6978.71476	7145.5759	3614.92573	1733.152467	
Goverment	West South Central	0.00000	3321.9424	0.00000	3.657082	
Non-goverment	West South Central	1031.02549	12336.1951	1807.46287	897.685905	
Goverment	Mountain	385.44830	904.2397	0.00000	28.619654	
Non-goverment	Mountain	46.14503	5485.0594	49.81985	0.000000	
Goverment	Pacific	0.00000	4210.3745	0.00000	0.000000	
Non-goverment	Pacific	6309.18418	9735.6014	880.78346	147.401067	
Goverment	Total	27268.56382	34890.2667	9188.09454	2189.929894	
Non-goverment	Total	185072.61826	87362.5664	25717.44325	13657.276178	

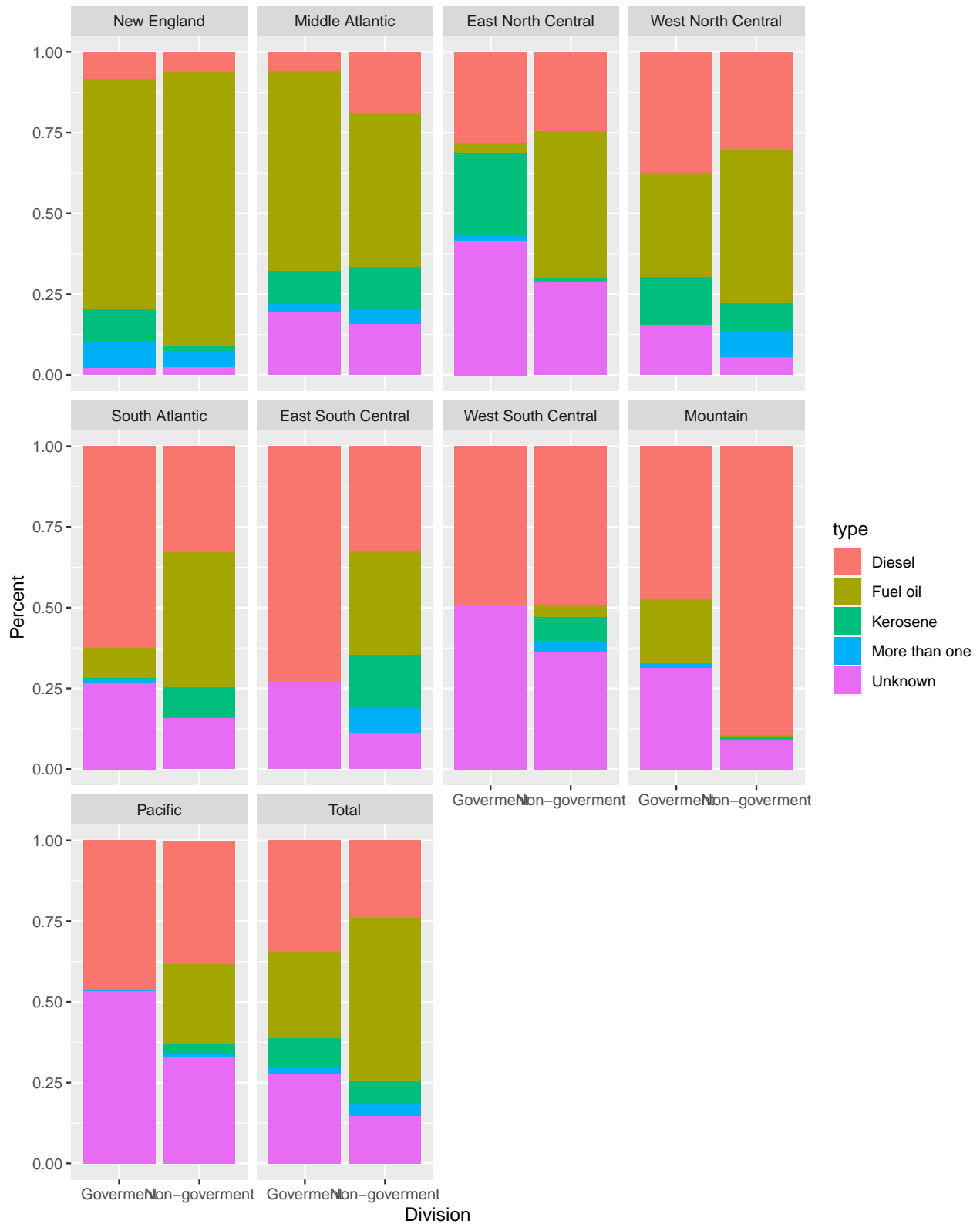**Table 7**. Types of fuel oil

Figure 3. Stacked bar plots for type of fuel oil

In general, government buildings use less heavy fuel oil than non-government buildings. However, in Mountain Division, government buildings use significantly more heavy fuel oil. In New England, East and West North

Central divisions, government buildings use kerosene more, but in other divisions, the difference is not obvious or non-government buildings use kerosene more. The type of fuel oil that buildings use also differs by census division.

Conclusion

The analysis about CBECS 2012 data provides some evidence for the assumption that government-owned buildings are more likely to use fuel oil as a source of energy, and they consume more fuel oil. However, generally non-government buildings consume more heavy fuel oil which is more polluting than kerosene and diesel.

Different census divisions use fuel oil for different reasons. In divisions where heating is not the main use of fuel oil, government buildings don't use fuel oil for heating entirely, while non-government buildings still use it for heating.

References

[1] Fuel oil. Wikipedia.

[2] Jackknife resampling. Wikipedia.