

Homework 2

rbulkunde3 Bulkunde

January 2023

1 Introduction

2 Sample mean table

See table 1.

2.1 Python version

	Mean (s.d.)	(s.d.)	Diff (p value)
Electricity (kWh)	1181.33 (454.31)	1086.75 (423.96)	-3.403304 0.000692
Home (sqft)	1633.05 (682.90)	1657.55 (686.27)	0.565839 0.571630
Temperature	79.89 (2.16)	79.89 (1.97)	0.016128 0.987135
Observations	501	499	1000.000000

Table 1: Standard deviation are in paranthesis

2.2 Stata version

	(1)		(2)		(3)	
	mean	sd	mean	sd	b	t
electricity	1181.329	454.308	1086.745	423.960	94.584***	(3.404)
sqft	1633.052	682.904	1657.551	686.271	-24.499	(-0.566)
temp	79.891	2.163	79.893	1.968	-0.002	(-0.016)
Observations	501		499		1000	

Table 2: Summary statistics produced using Stata

(1)	
VARIABLES	Ordinary least squares
sqft	0.62** (0.01)
retrofit	-109.67** (7.94)
temp	3.26 (1.93)
Constant	-83.60 (154.69)
Observations	1,000
R-squared	0.92
Robust standard errors in parentheses	
** p<0.01, * p<0.05	

Table 3: Regression produced using Stata

3 STATA Regression

4 Graphs

4.1 Python version: Kernel Density Plot

4.2 Stata version: two-way scatterplot

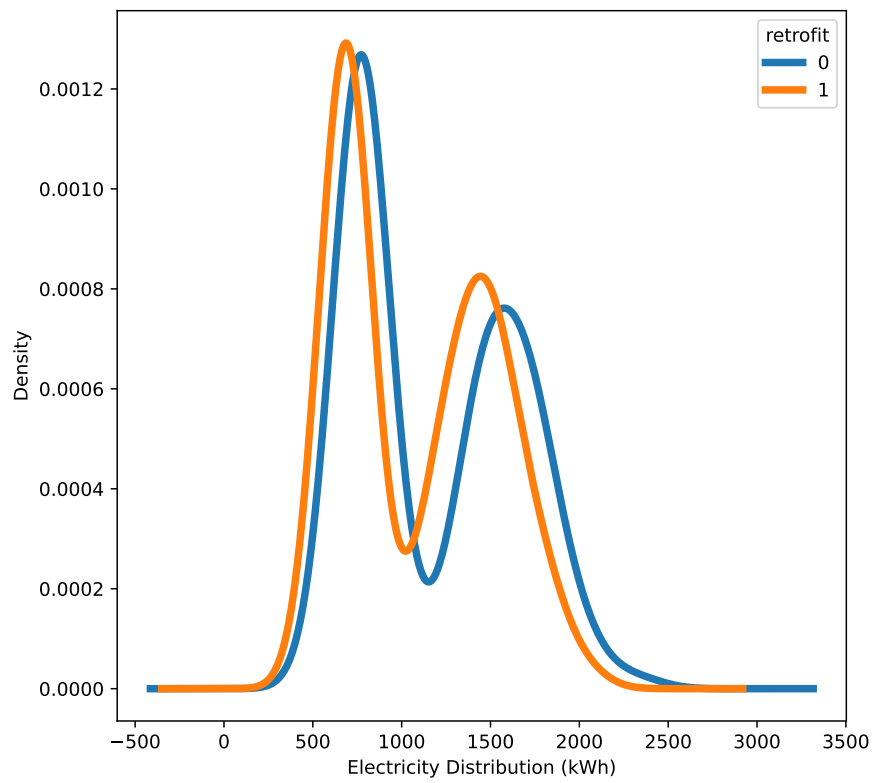


Figure 1: Sample kernel density plot of the outcome variable.

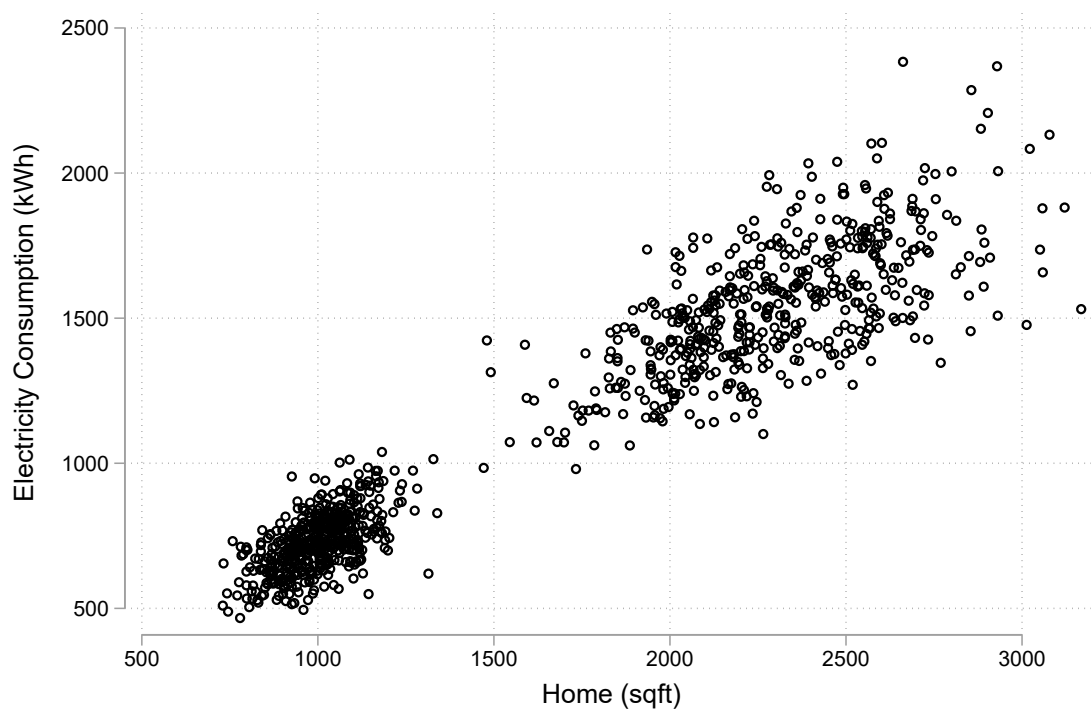


Figure 2: two-way scatterplot