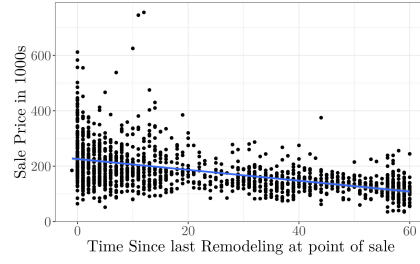
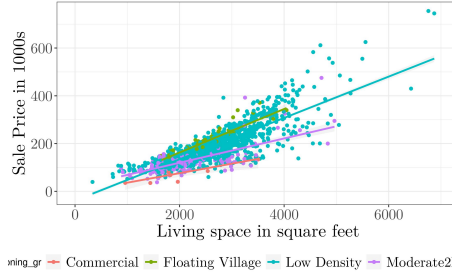


(A) Positive association of total Living Space & Sale Price & optimal line



(B) Positive association of total Living Space & Sale Price



(C) Positive association of total Living Space & Sale Price

FIGURE 1: Three Hypothesis Graphs displaying their respective association with the outcome variable

1 Data Prep, EDA, and Theory development

1.1 Variable Selection & Explanation

For the purpose of analyzing the determinants of prices of home sales in the US, the following variables were included in the analysis:

TABLE 1: DESCRIPTIVE STATISTICS OF NUMERIC VARIABLES

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
SalePrice	180.921	79.443	34.900	129.975	163.000	214.000	755.000
Lot Area	10,516.830	9,981.265	1,300	7,553.5	9,478.5	11,601.5	215,245
Quality	6.099	1.383	1	5	6	7	10
Condition	5.575	1.113	1	5	5	6	9
Total Living Space	2,572.893	823.598	334	2,014	2,479	3,008.5	11,752
Years Since Remodeling	22.950	20.641	-1	4	14	41	60

Notes: OLS estimates, robust standard errors in parentheses.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Plot 1: Causal relationship Scheme

