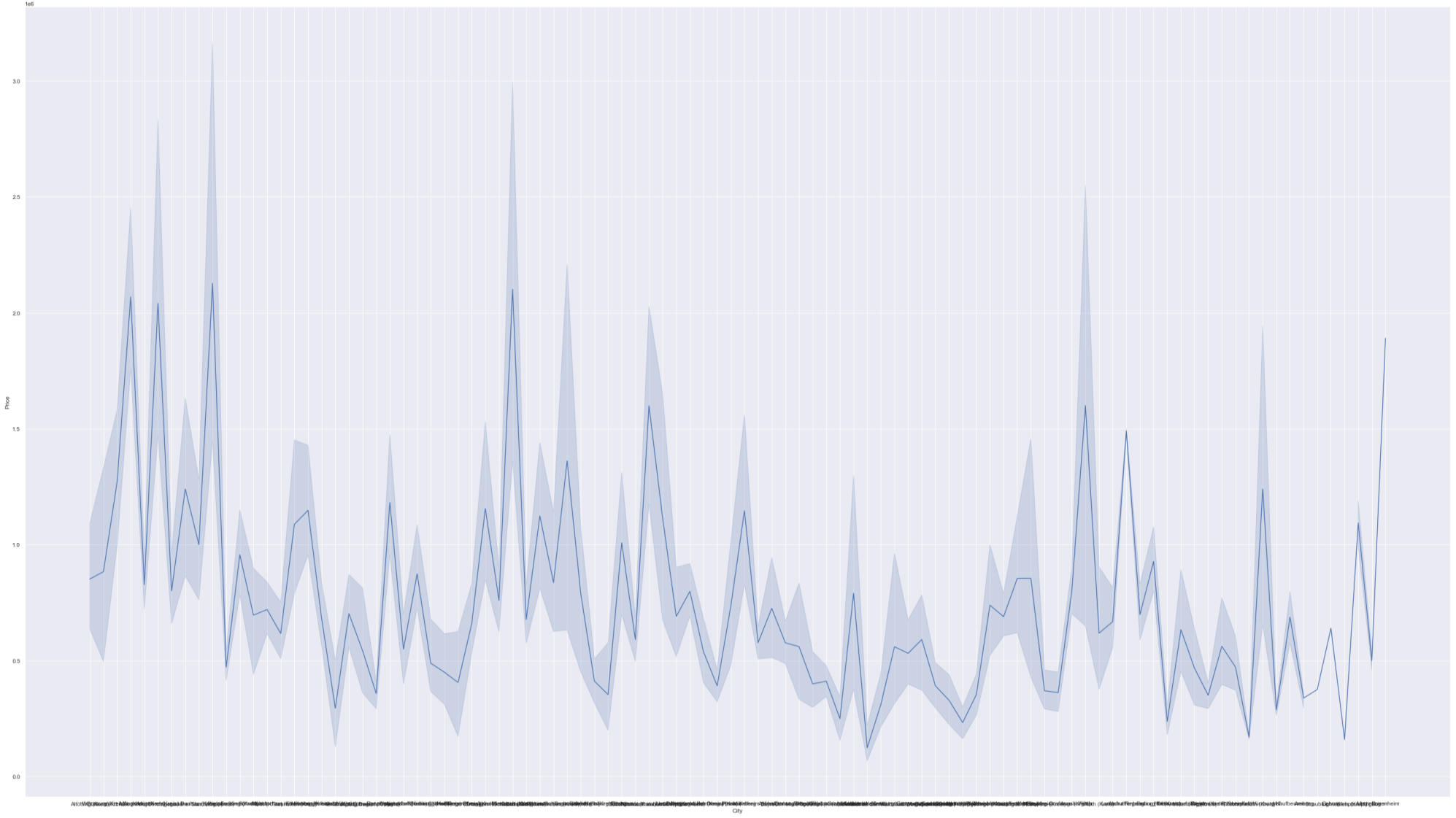
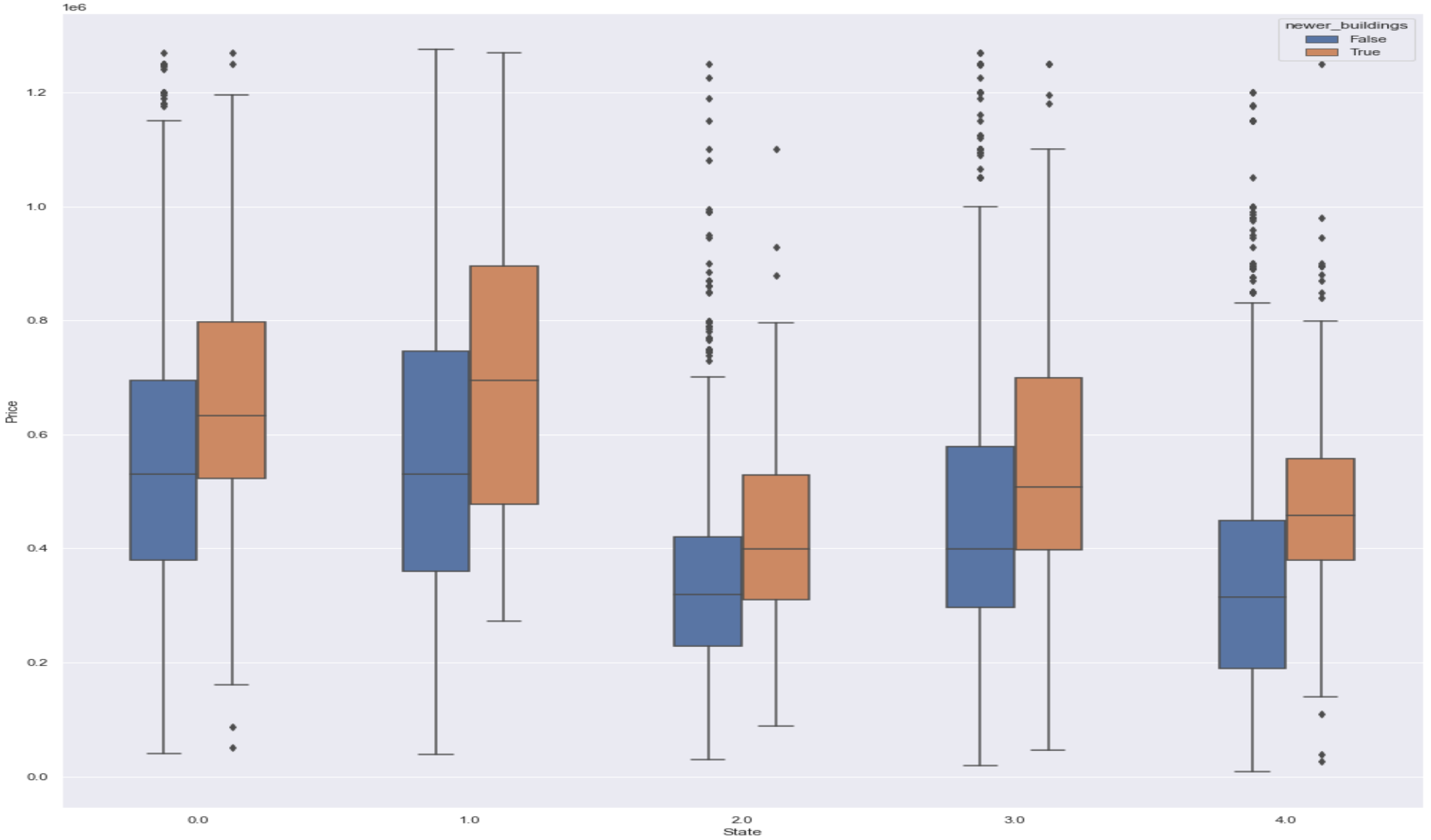
Appendix for the different figures in the analysis that has been scaled for the ease of visualisation:

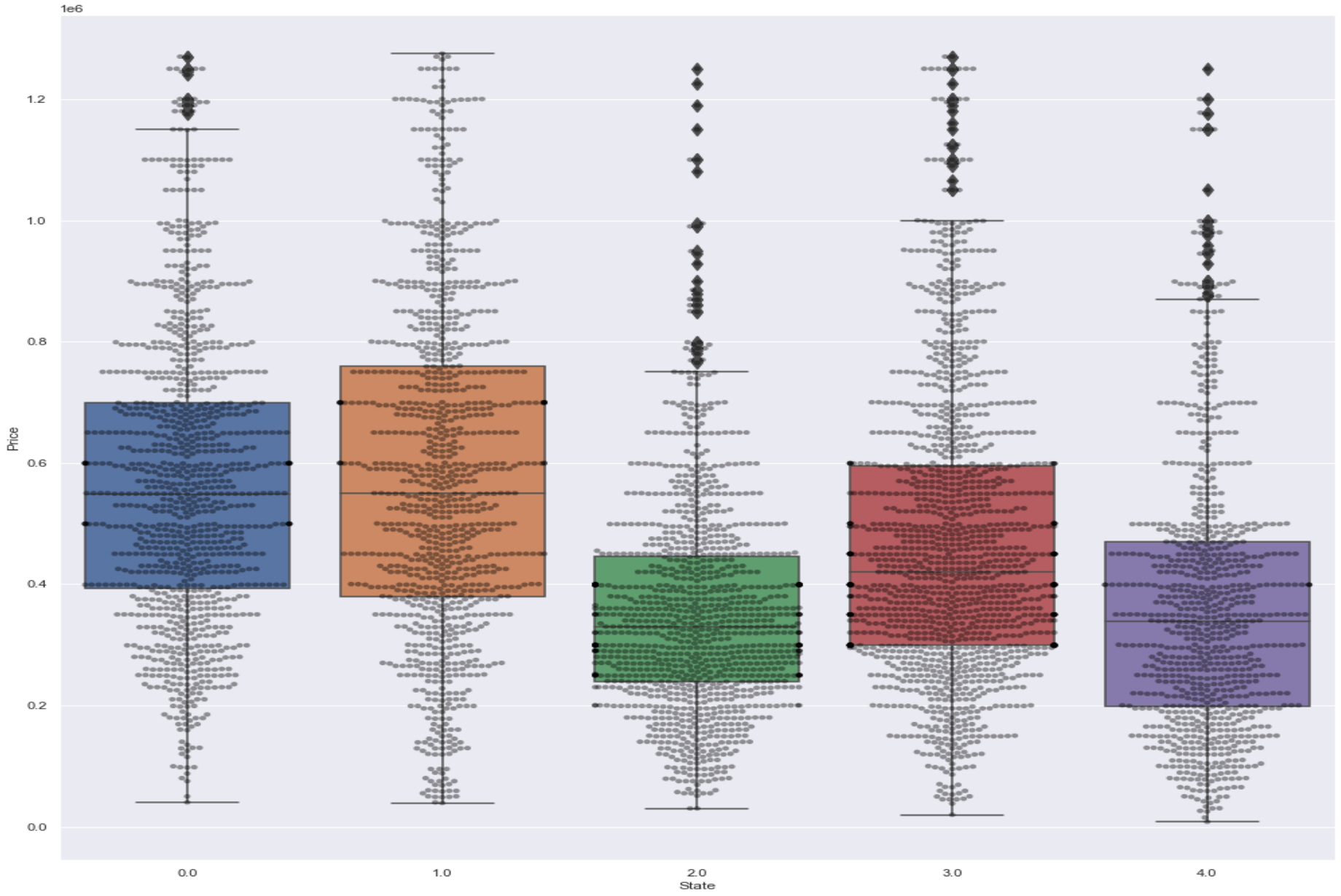
Figure(6.1) :



Figure(6.2) :


1e6
newer_buildings
False
True
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Price
0.0
1.0
2.0
3.0
4.0
State

Figure(6.3) :



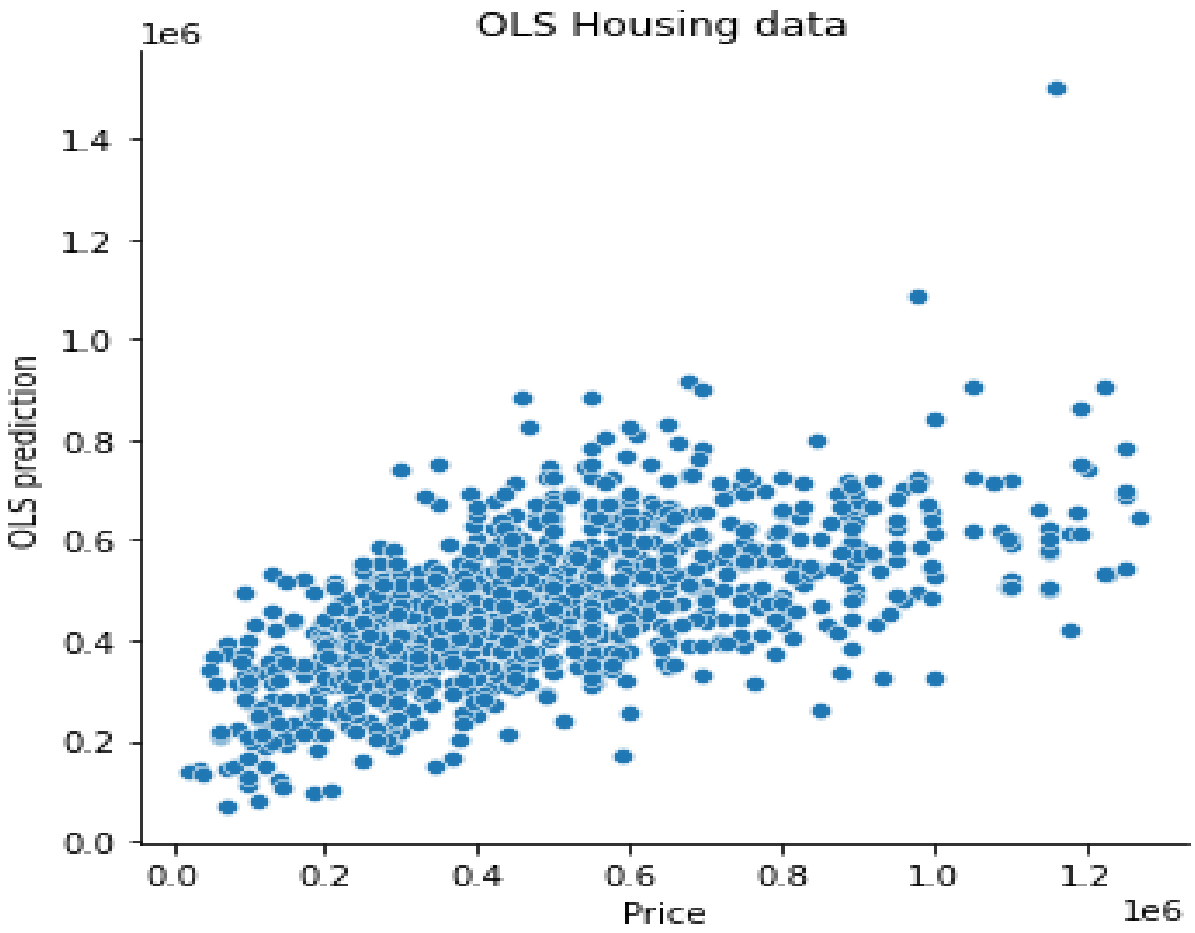
Figure(6.4):

	parameter_one	parameter_two	correlation	Do_correlate?
0	Price	Living_space	0.405045	yes
1	Price	HH_gross_income_mean	0.349621	yes
2	Price	HH_disposable_income_mean	0.345542	yes
3	Price	Phone_std	0.302892	yes
4	Price	Phone_mean	0.301412	yes
5	Price	Bathrooms	0.295349	yes
6	Price	Year_built	0.282400	yes
7	Price	Rooms	0.266642	yes
8	Price	Bedrooms	0.246157	yes
9	Price	HH_gross_income_std	0.245698	yes
10	Price	Dwelling_type_mean	0.241202	yes
11	Price	Floors	0.227775	yes
12	Price	HH_disposable_income_std	0.224764	yes
13	Price	Dwelling_type_std	0.208359	yes
14	Price	Garages	0.171165	yes
15	Price	Usable_area	0.163777	yes
16	Price	newer_buildings	0.161396	yes
17	Price	City	0.104831	yes
18	Price	Lot	0.081408	yes
19	Price	Place	0.070100	yes
20	Price	Energy_source	0.061736	yes
21	Price	Type	0.026038	no
22	Price	HH_size_mean	0.015123	no
23	Price	HH_social_service_dis_income_std	-0.006097	no
24	Price	Heating	-0.019852	no
25	Price	Garagetype	-0.059173	yes
26	Price	Energy_efficiency_class	-0.135330	yes
27	Price	HH_size_std	-0.149200	yes
28	Price	Condition	-0.164862	yes
29	Price	HH_social_service_dis_income_mean	-0.195904	yes
30	Price	crime_rate_mean	-0.254205	yes
31	Price	State	-0.265194	yes
32	Price	Relative_pov_cat_mean	-0.341566	yes
33	Price	Affording_holiday_std	-0.348558	yes
34	Price	Affording_holiday_mean	-0.351118	yes
35	Price	Relative_pov_cat_std	-0.360207	yes

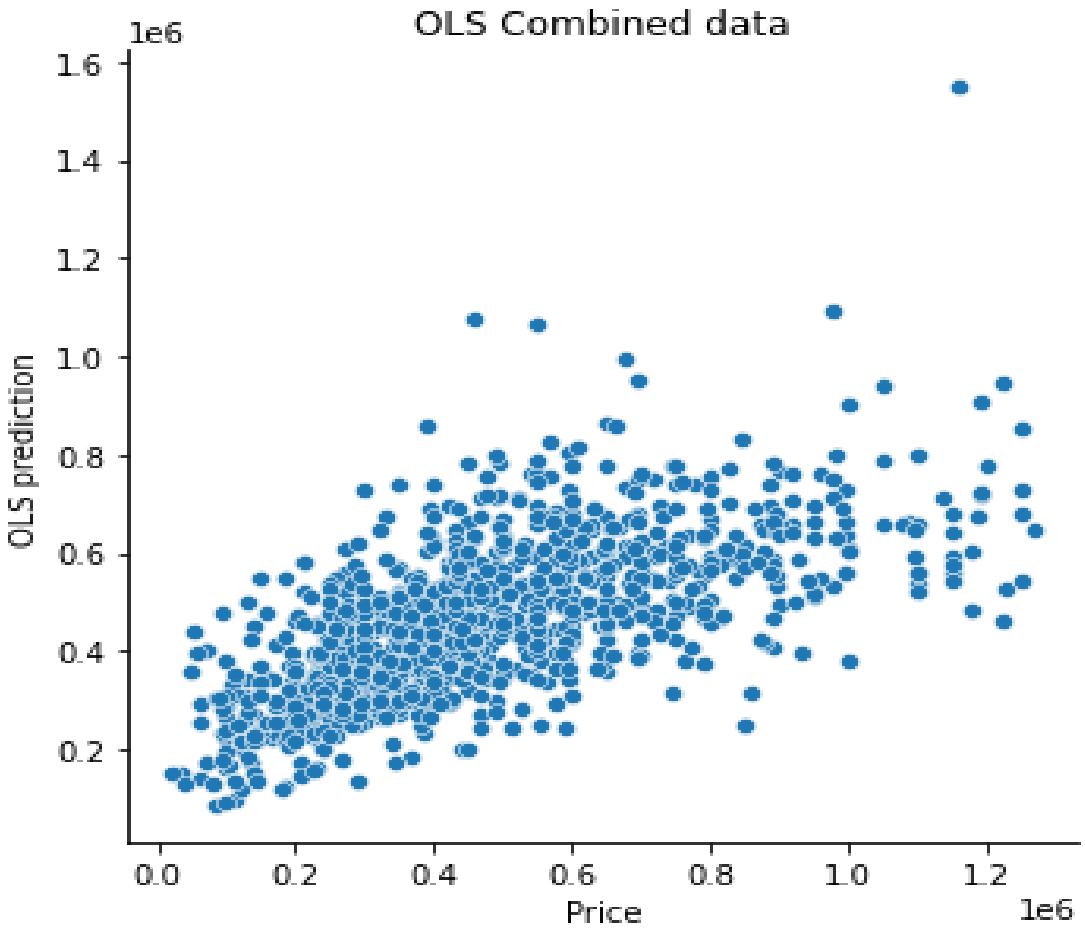
Figure(6.5) :

	PR(>F)	Do_correlate?
Garages	0.600569	no
Usable_area	0.527713	no
Bedrooms	0.312652	no
Place	0.180309	no
Heating	0.032318	yes
Energy_efficiency_class	0.003256	yes
Rooms	0.003077	yes
City	0.002767	yes
HH_social_service_dis_income_std	0.000273	yes
Type	0.000016	yes
HH_size_std	0.000012	yes
Garagetype	0.000012	yes
Bathrooms	0.000001	yes
Lot	0.000000	yes
Floors	0.000000	yes
HH_social_service_dis_income_mean	0.000000	yes
Energy_source	0.000000	yes
Condition	0.000000	yes
HH_gross_income_std	0.000000	yes
Phone_std	0.000000	yes
Dwelling_type_mean	0.000000	yes
HH_disposable_income_std	0.000000	yes
HH_gross_income_mean	0.000000	yes
Dwelling_type_std	0.000000	yes
Phone_mean	0.000000	yes
State	0.000000	yes
Relative_pov_cat_std	0.000000	yes
Affording_holiday_mean	0.000000	yes
Affording_holiday_std	0.000000	yes
Relative_pov_cat_mean	0.000000	yes
HH_disposable_income_mean	0.000000	yes
Year_built	0.000000	yes
Living_space	0.000000	yes
crime_rate_mean	0.000000	yes
HH_size_mean	0.000000	yes
Residual	nan	no

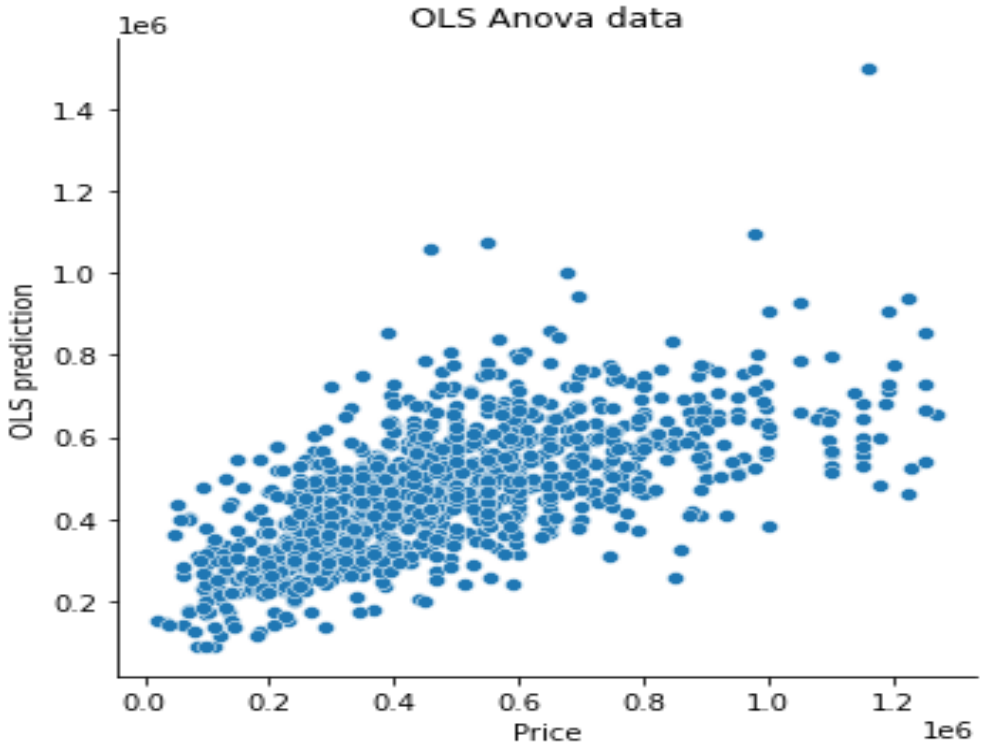
Figure(6.6)(a) :


OLS Housing data
1e6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
OLS prediction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Price
1e6

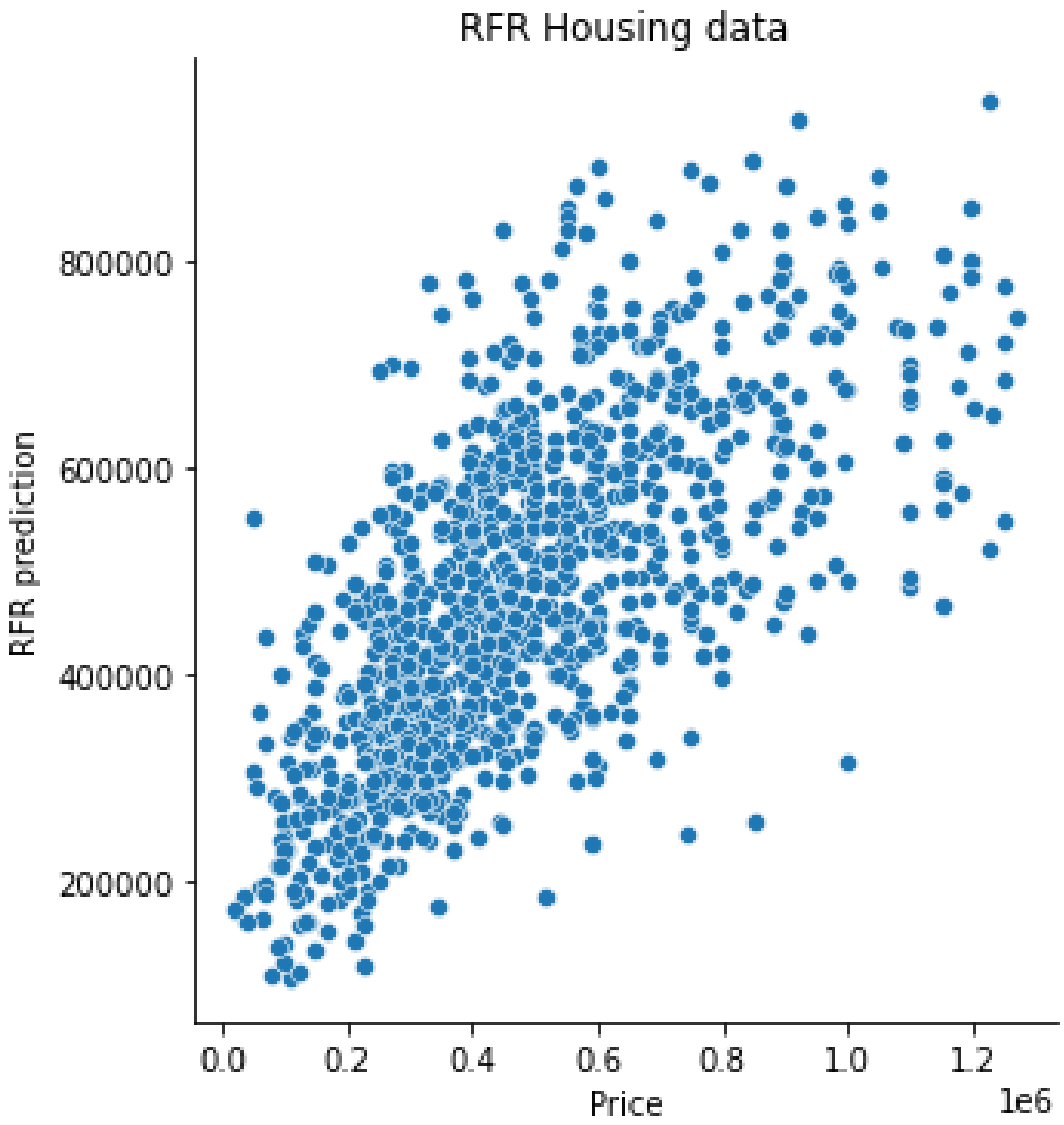
Figure(6.6)(b) :


OLS Combined data
1e6
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
OLS prediction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1e6
Price

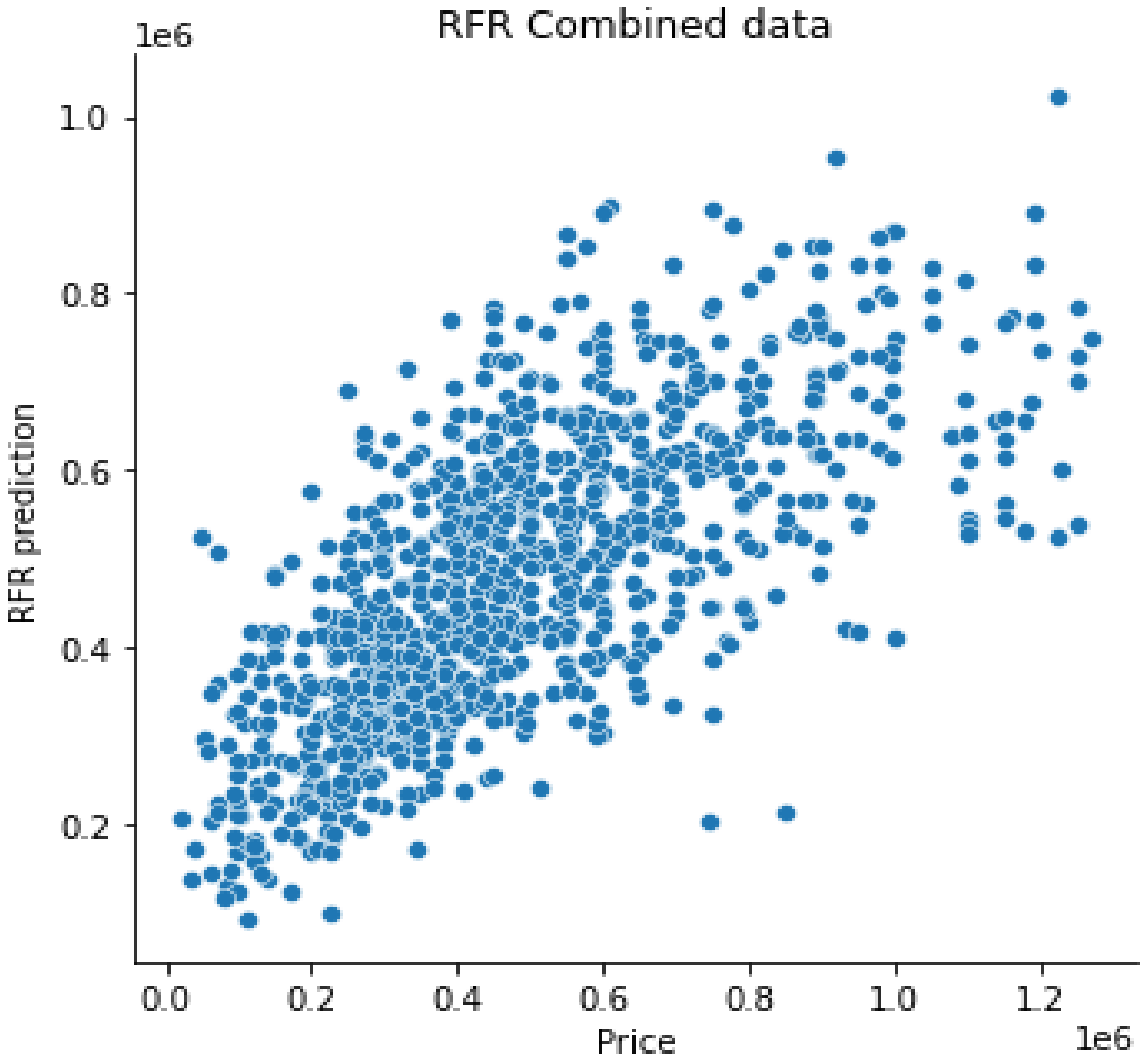
Figure(6.6)(c) :


OLS Anova data
1e6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
OLS prediction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1e6
Price

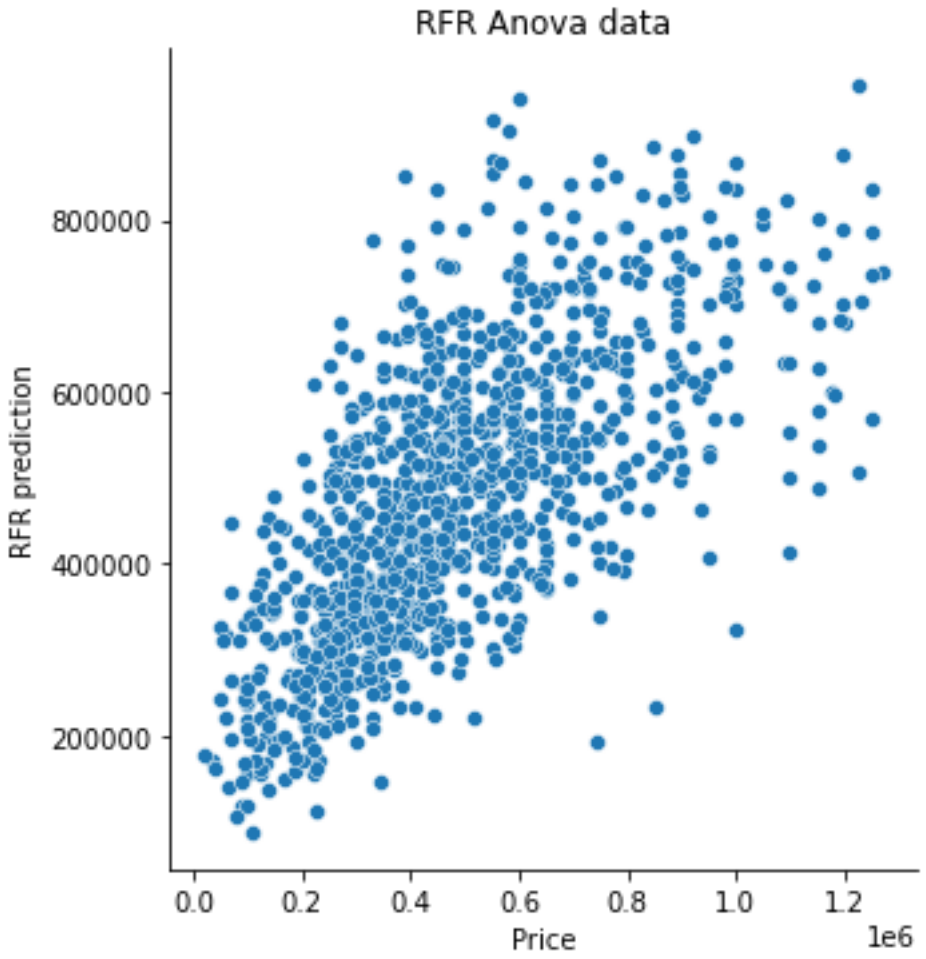
Figure(6.7)(a) :


RFR Housing data
RFR prediction
800000
600000
400000
200000
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Price
1e6

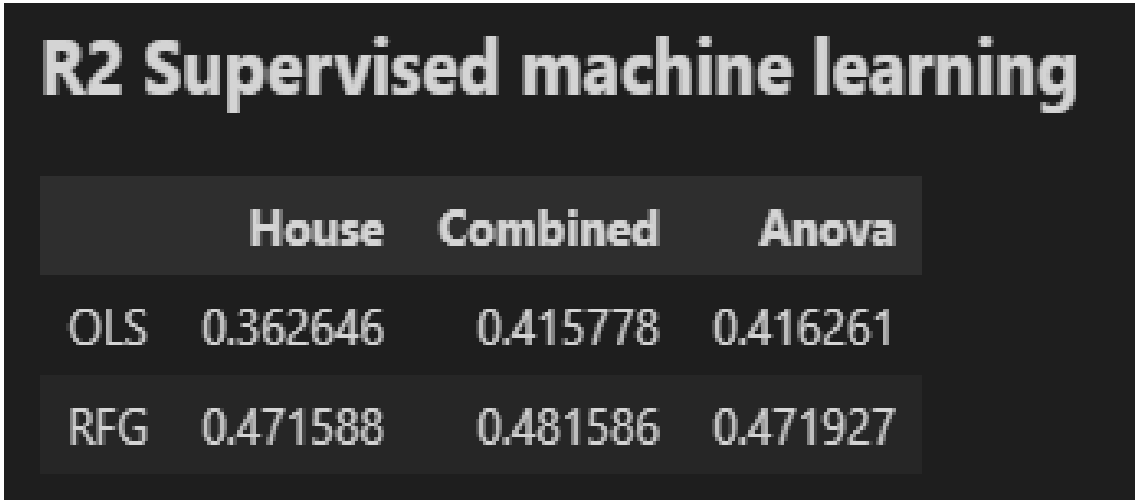
Figure(6.7)(b) :


RFR Combined data
1e6
1.0
0.8
0.6
0.4
0.2
RFR prediction
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Price
1e6

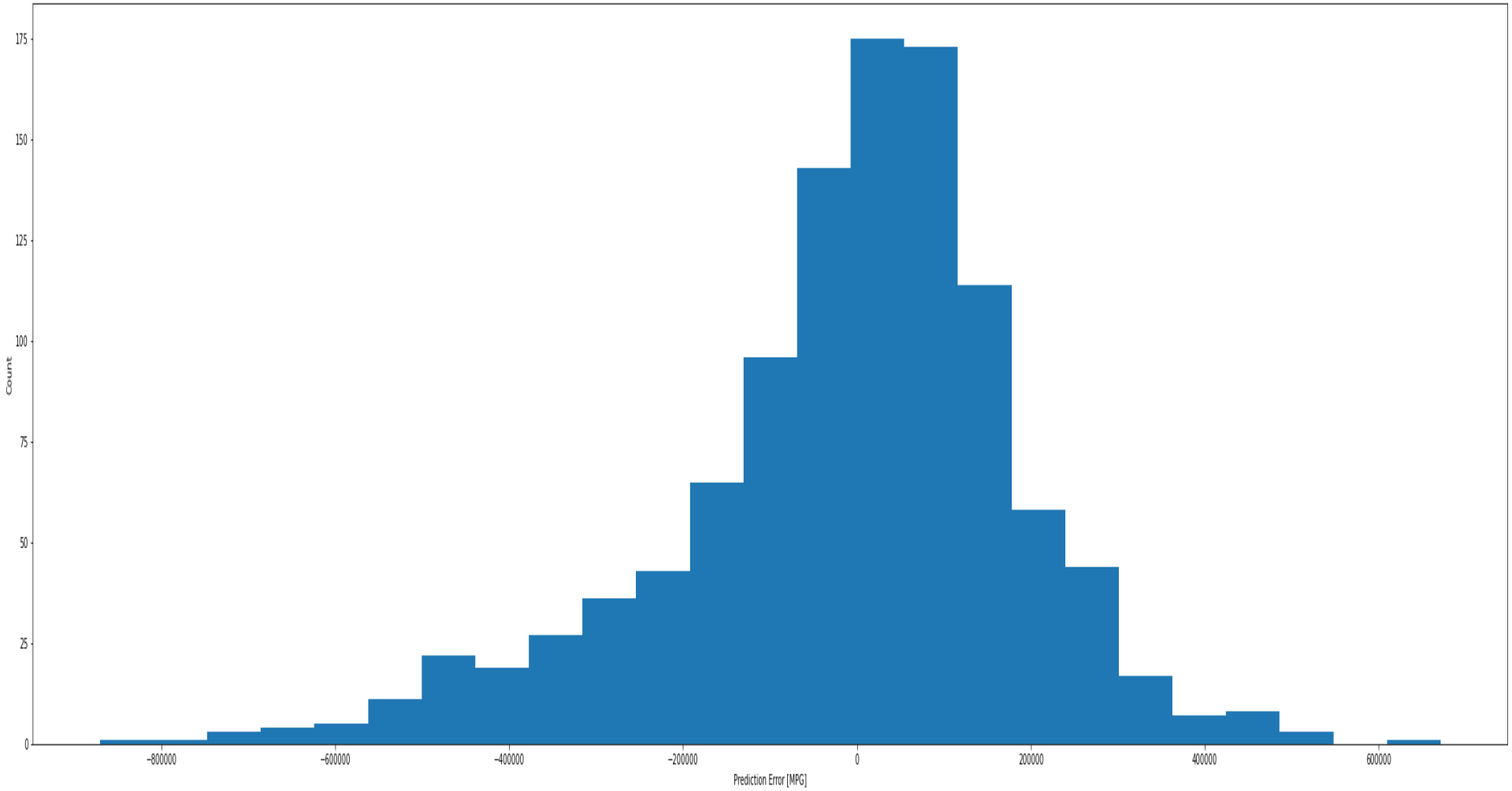
Figure(6.7)(c) :


RFR Anova data
RFR prediction
800000
600000
400000
200000
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Price
1e6

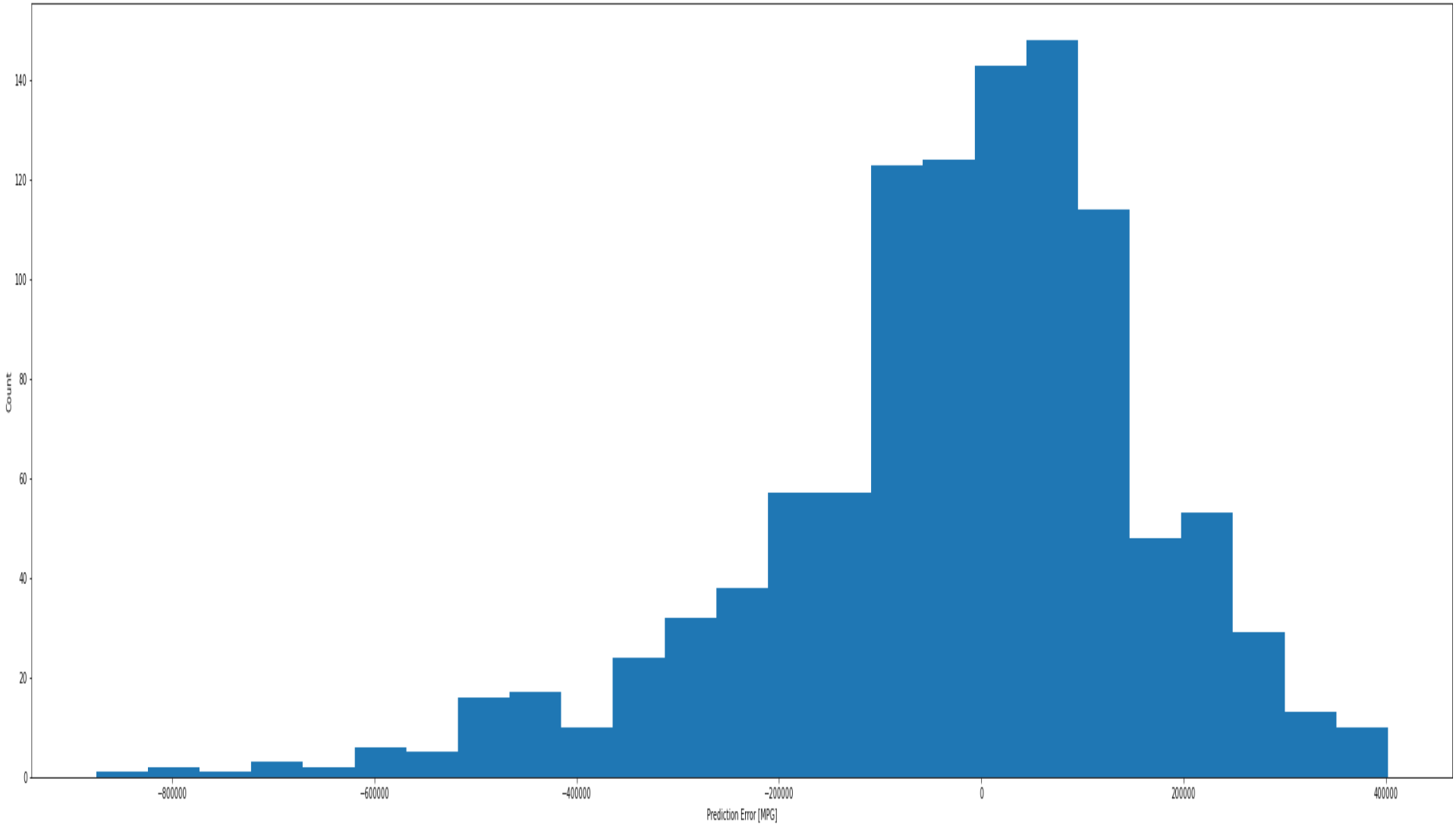
Figure(6.8) :


R2 Supervised machine learning
House
Combined
Anova
OLS
0.362646
0.415778
0.416261
RFG
0.471588
0.481586
0.471927

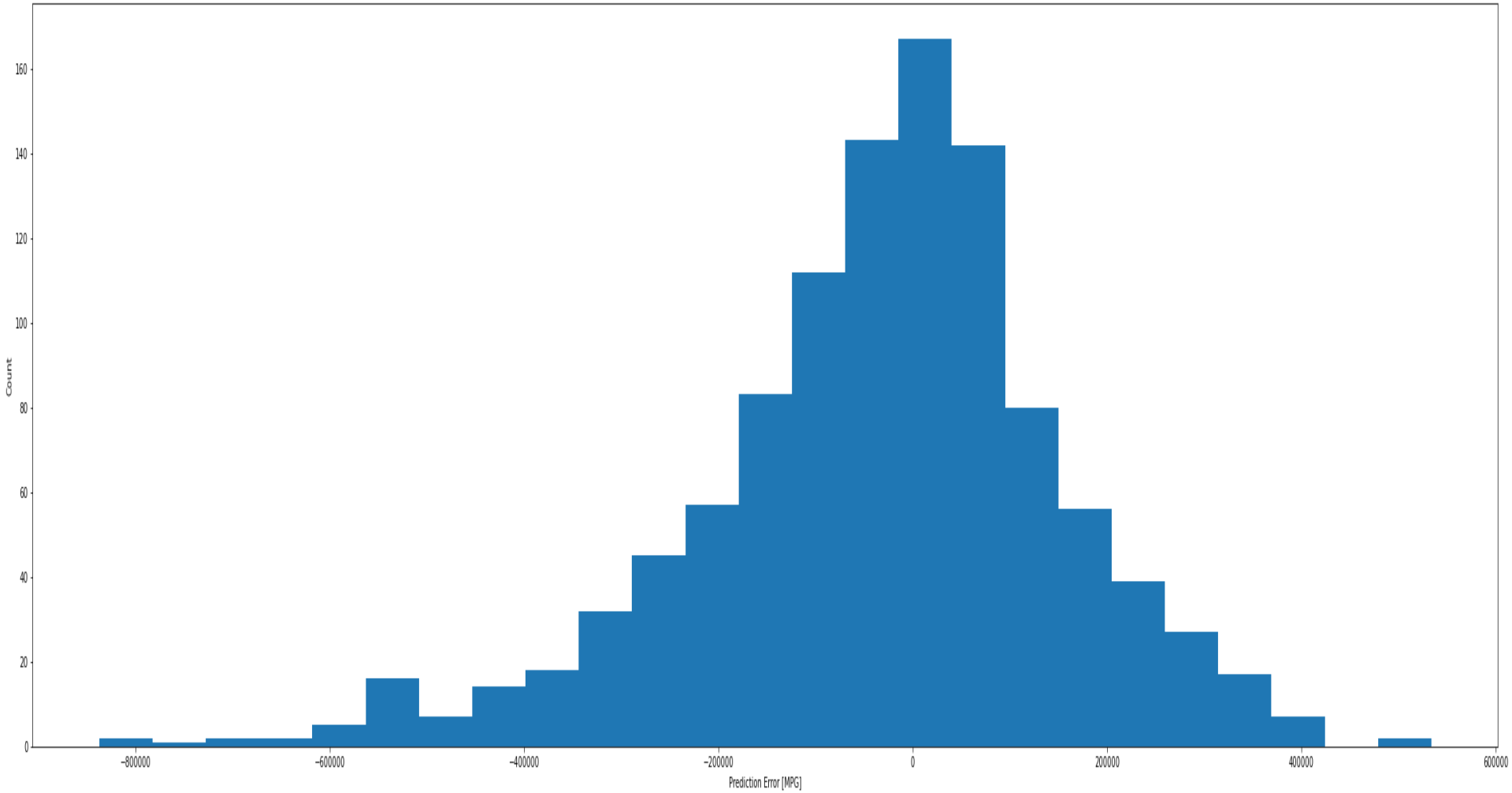
Figure(6.10) (a) :


175
150
125
100
75
50
25
0
Count
−800000
−600000
−400000
−200000
0
200000
400000
600000
Prediction Error [MPG]

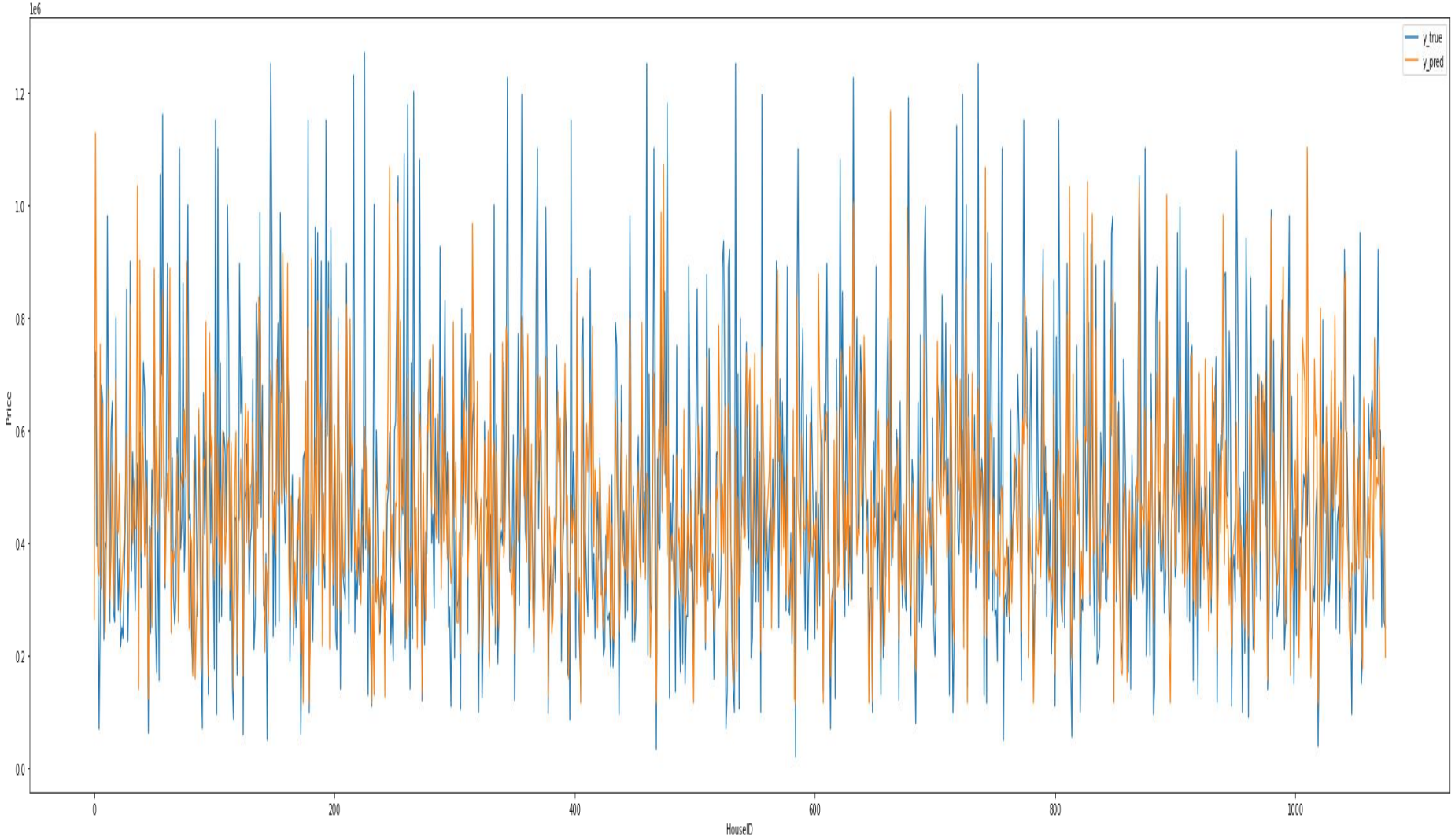
Figure(6.10) (b) :


140
120
100
80
60
40
20
0
Count
−800000
−600000
−400000
−200000
0
200000
400000
Prediction Error [MPG]

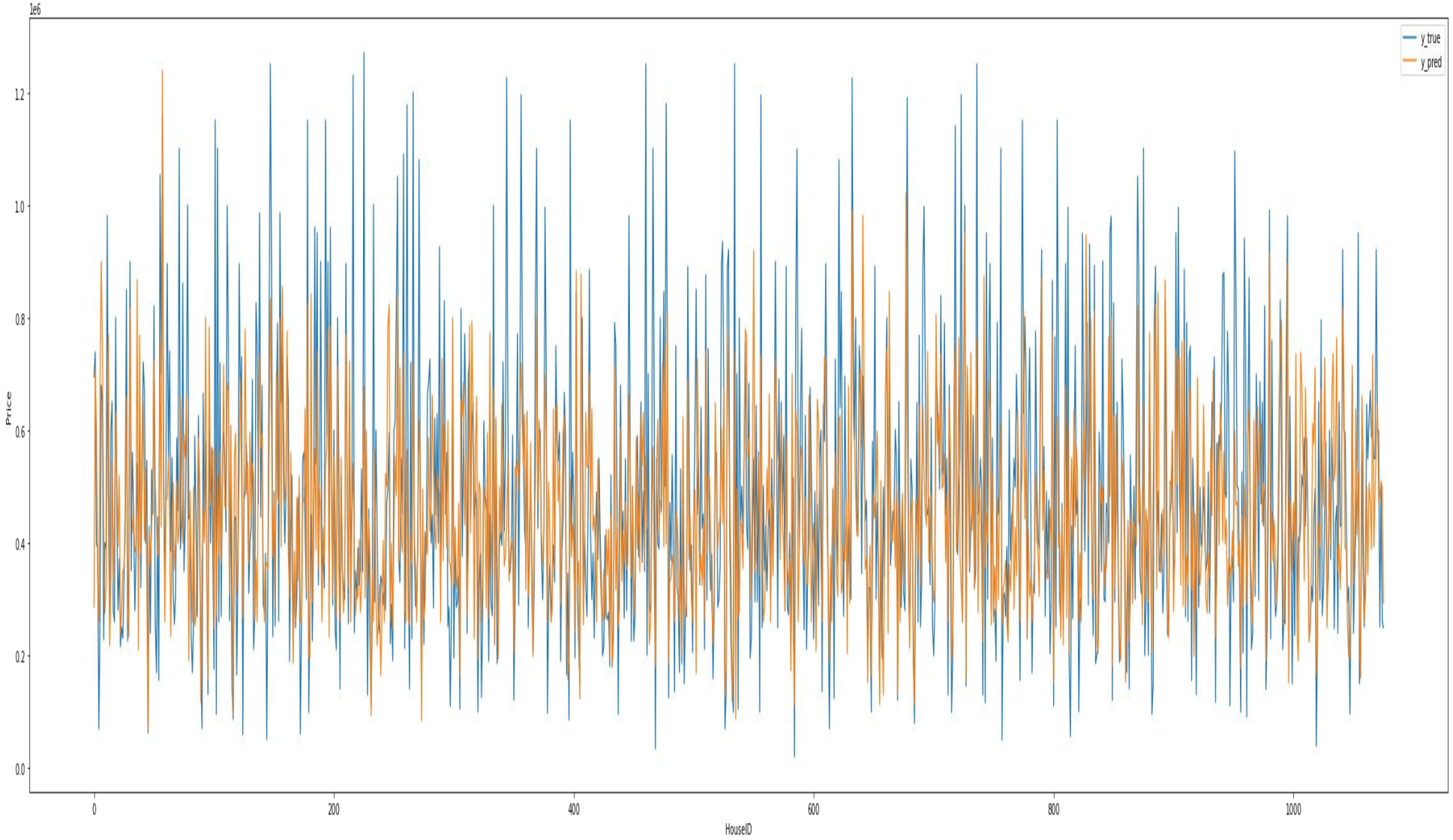
Figure(6.10) (c) :


160
140
120
100
80
60
40
20
0
Count
−800000
−600000
−400000
−200000
0
200000
400000
600000
Prediction Error [MPG]

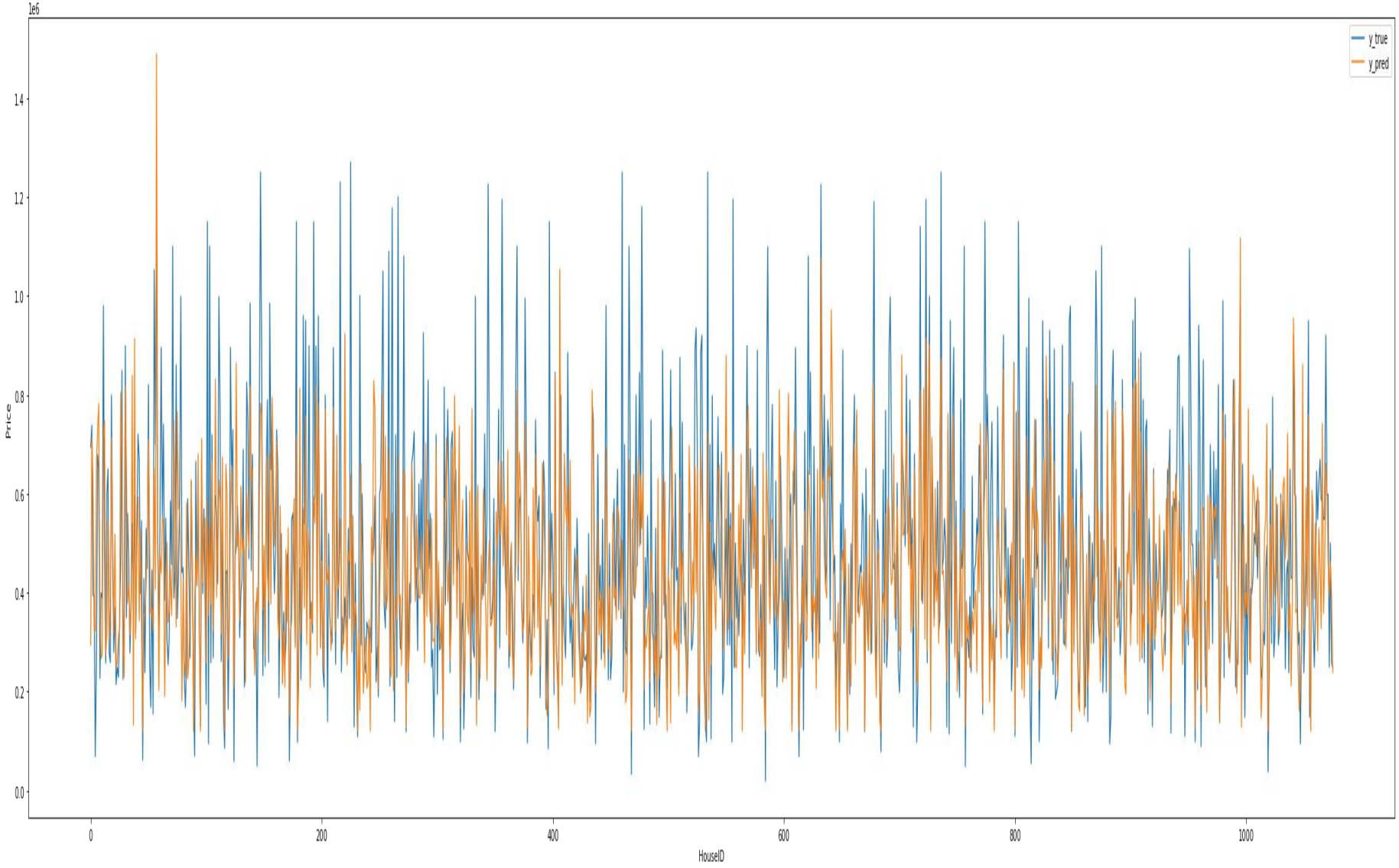
Figure(6.10) (a) :


1e6
y_true
y_pred
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Price
0
200
400
600
800
1000
HouseID

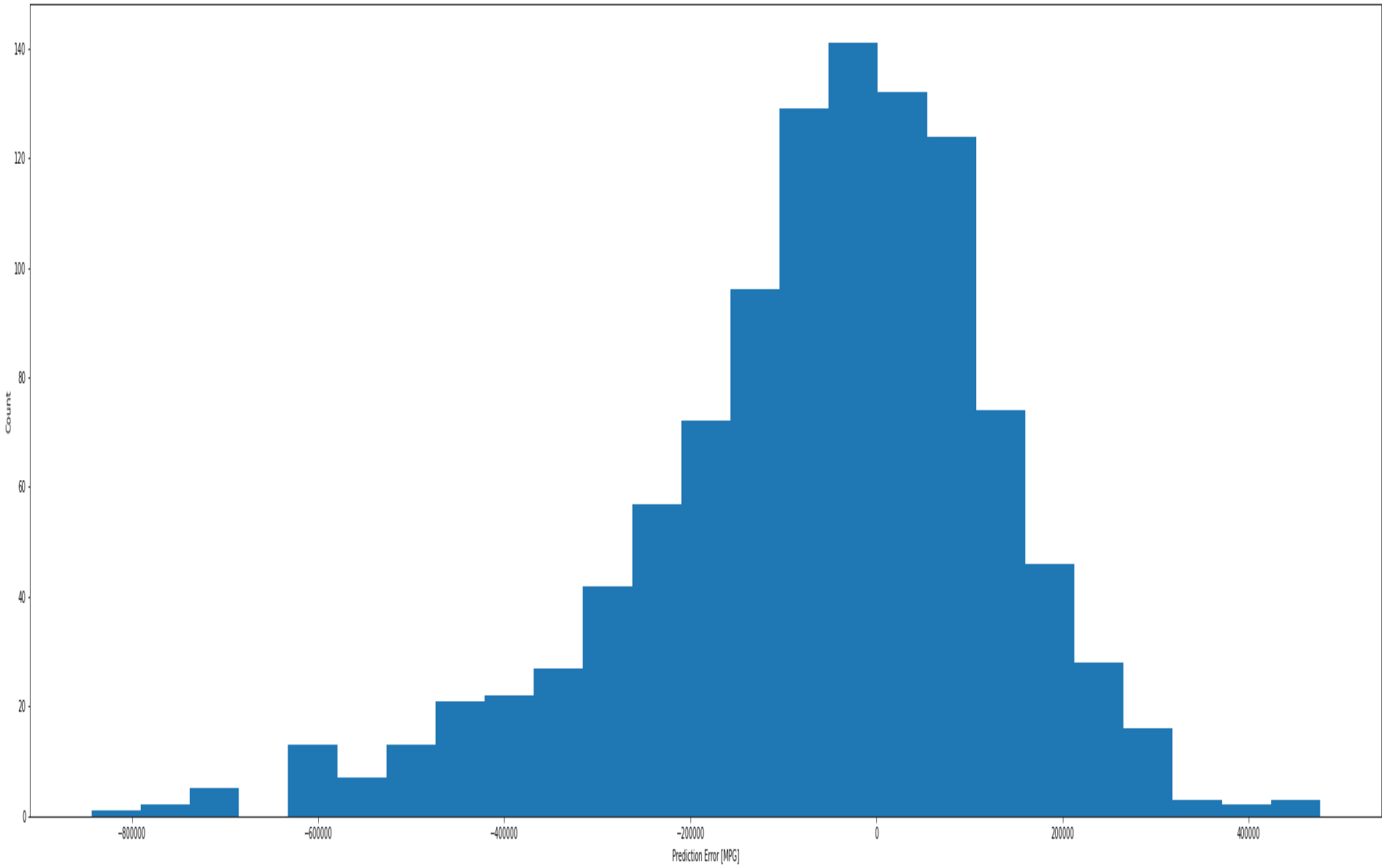
Figure(6.10) (b) :


1e6
y_true
y_pred
Price
HouseID
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
200
400
600
800
1000

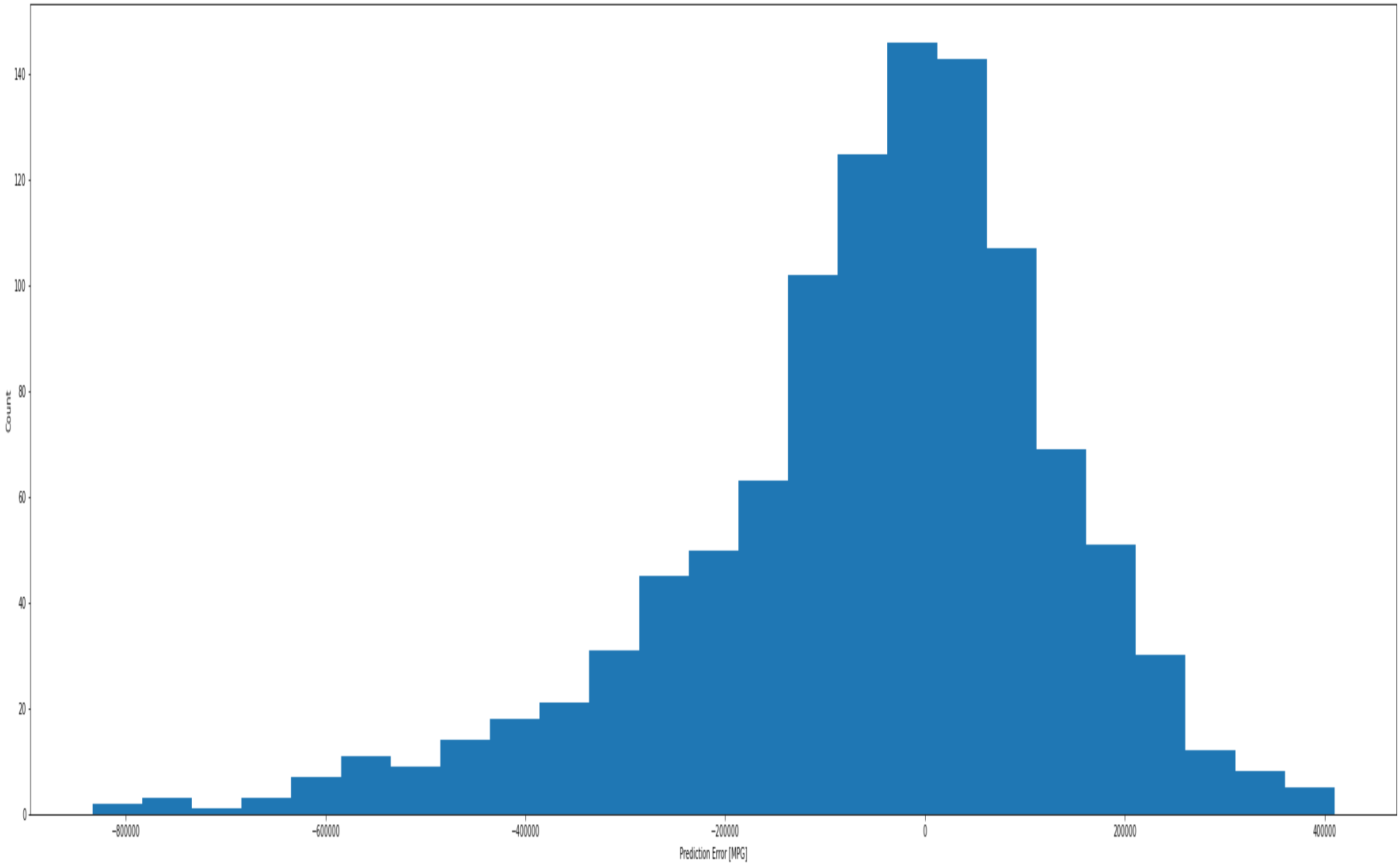
Figure(6.10) (c) :


1e6
y_true
y_pred
Price
HouseID
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0
200
400
600
800
1000

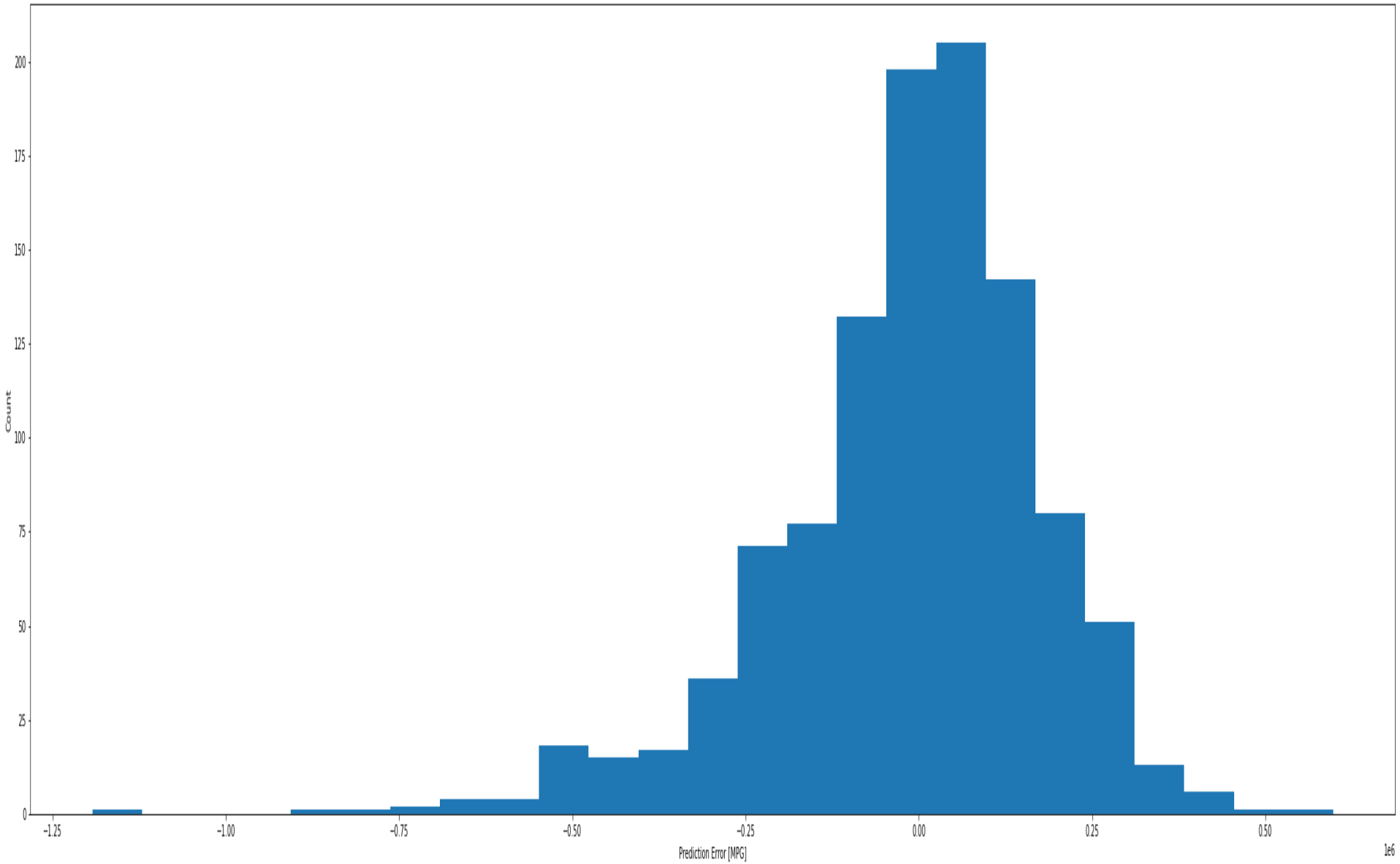
Figure(6.12) (a) :


Count
140
120
100
80
60
40
20
0
−800000
−600000
−400000
−200000
0
200000
400000
Prediction Error [MPG]

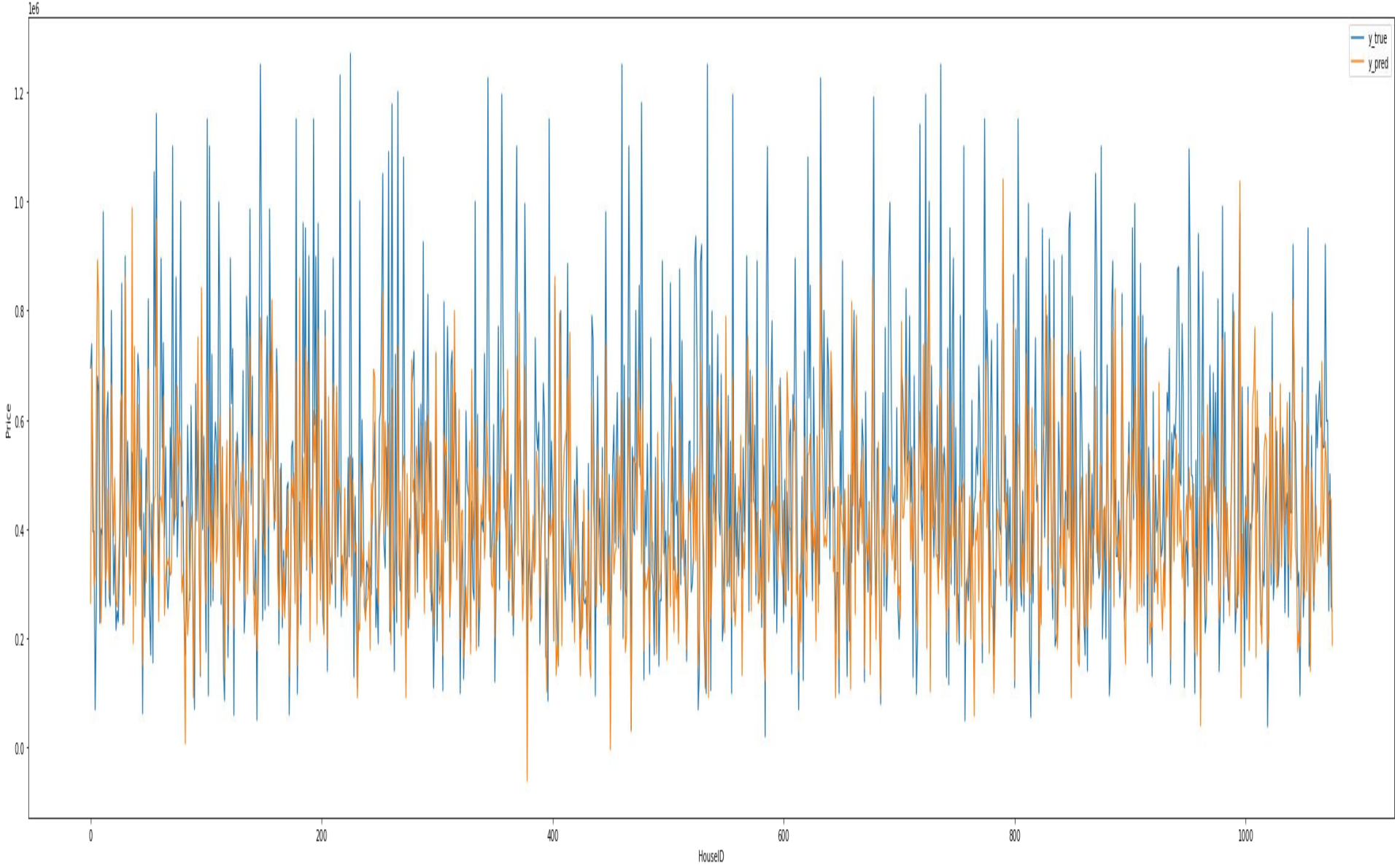
Figure(6.12) (b) :


140
120
100
80
60
40
20
0
Count
−800000
−600000
−400000
−200000
0
200000
400000
Prediction Error [MPG]

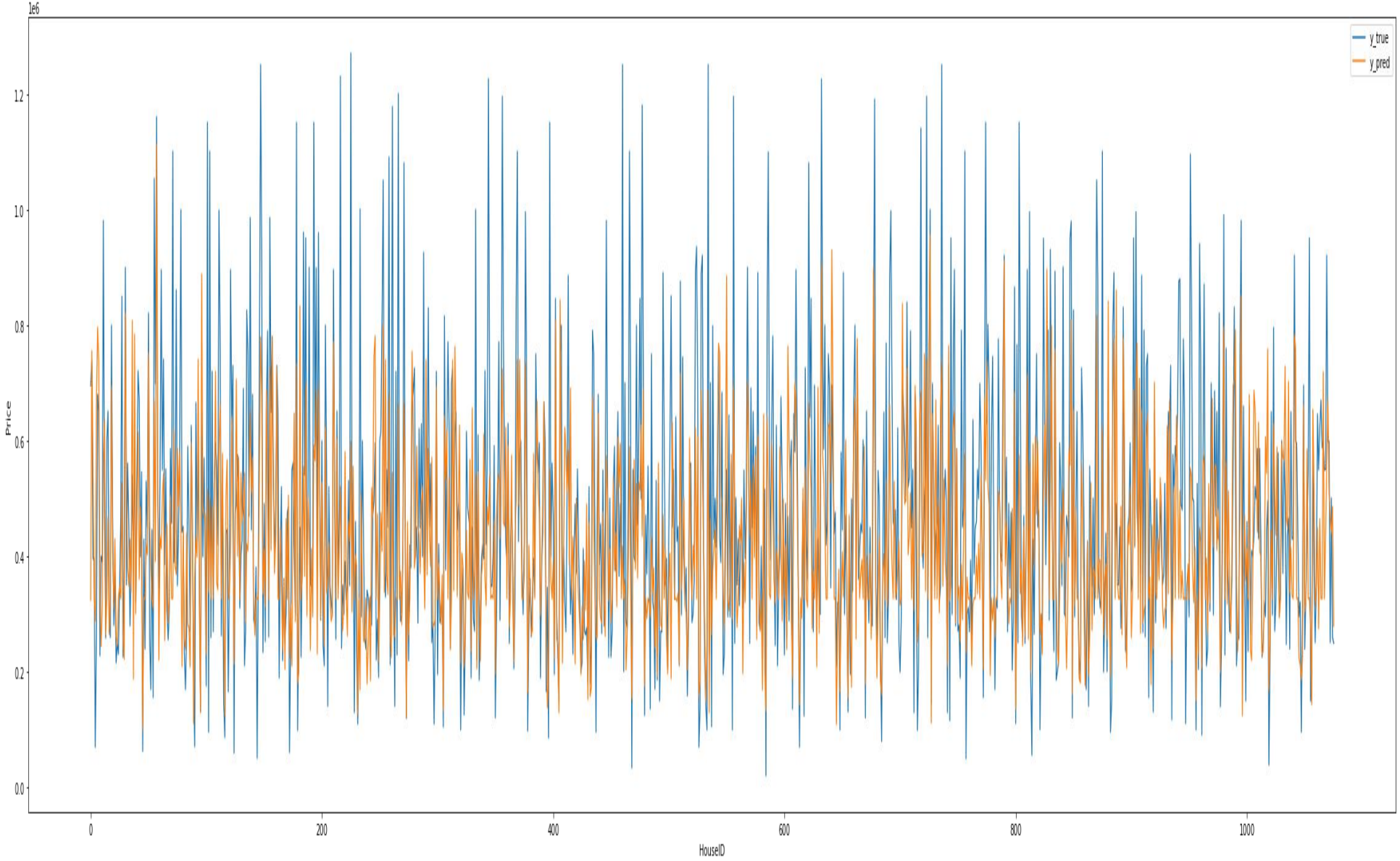
Figure(6.12) (c) :


Count
200
175
150
125
100
75
50
25
0
−1.25
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
1e6
Prediction Error [MPG]

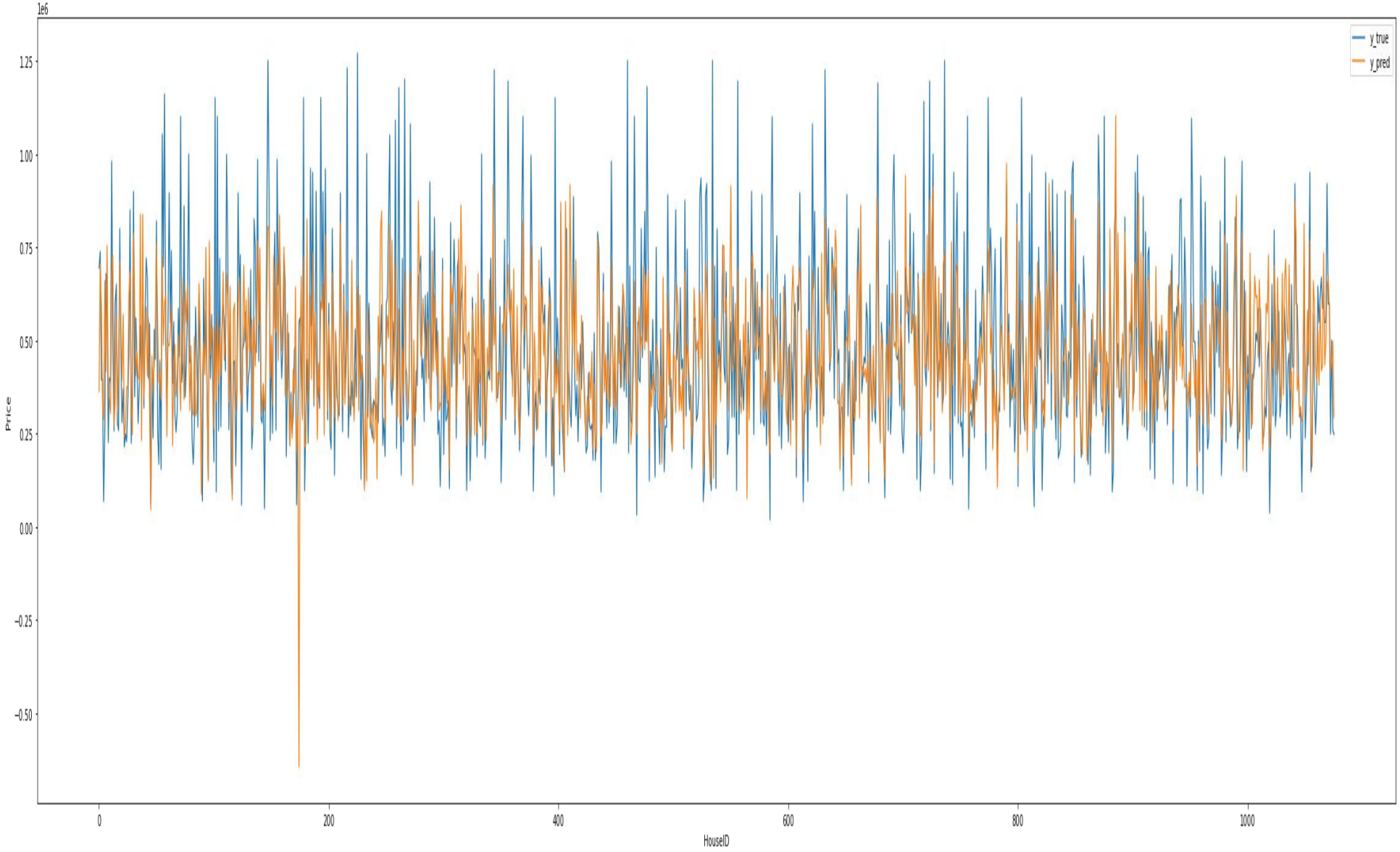
Figure(6.12) (a) :


1e6
y_true
y_pred
Price
HouseID
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
200
400
600
800
1000

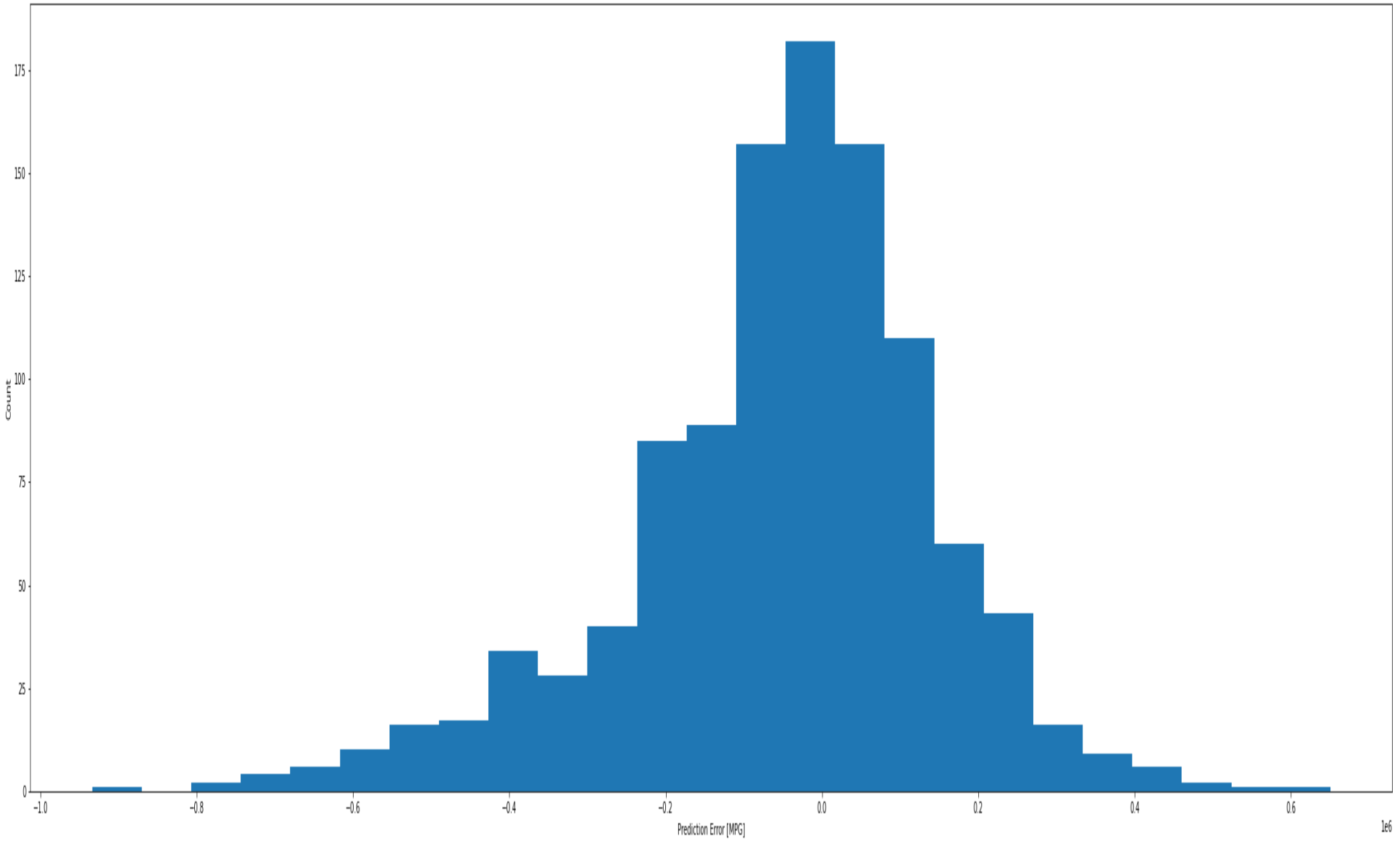
Figure(6.12) (b) :


1e6
y_true
y_pred
Price
HouseID
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
200
400
600
800
1000

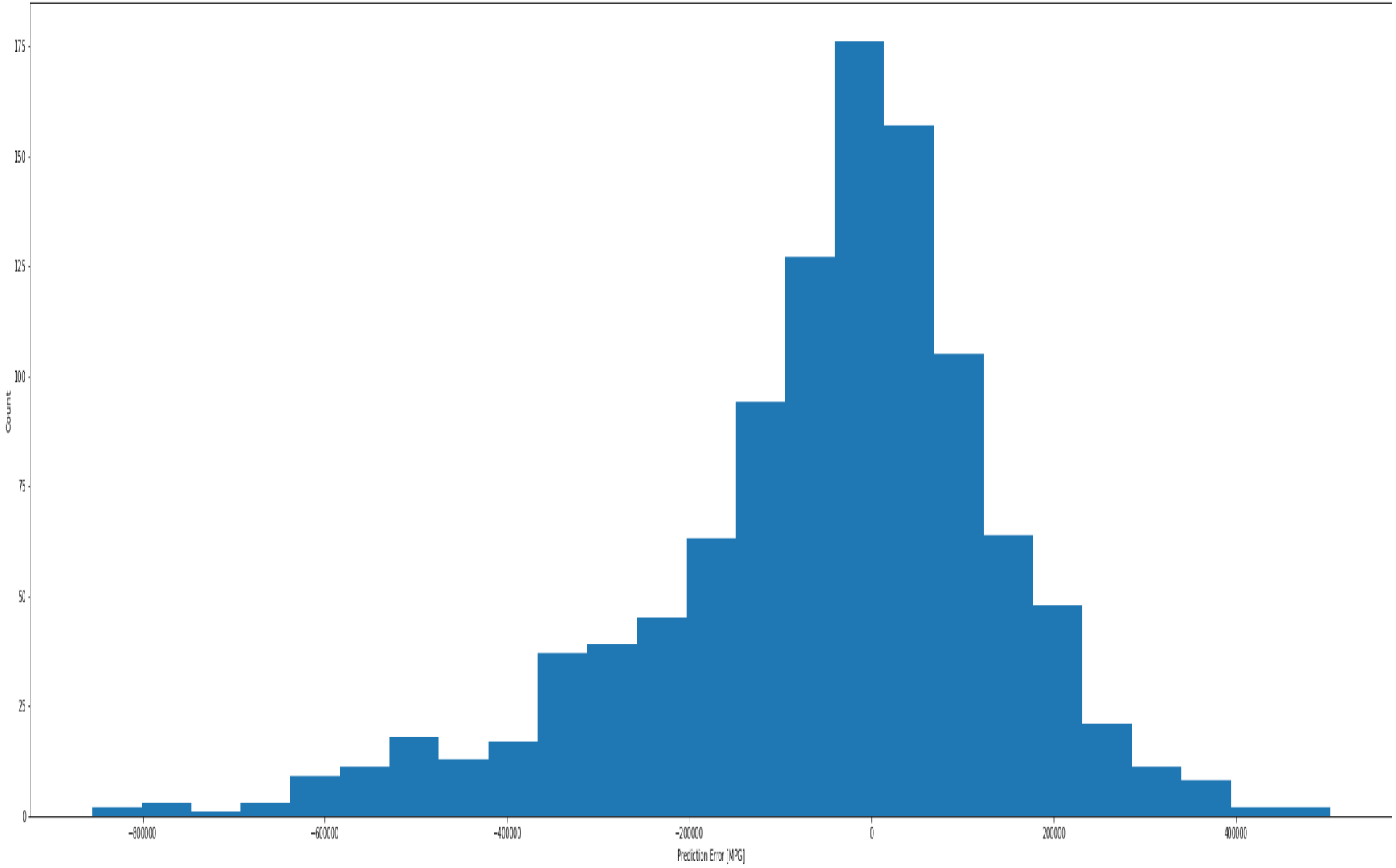
Figure(6.12) (c) :


1e6
1.25
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
Price
0
200
400
600
800
1000
HouseID
y_true
y_pred

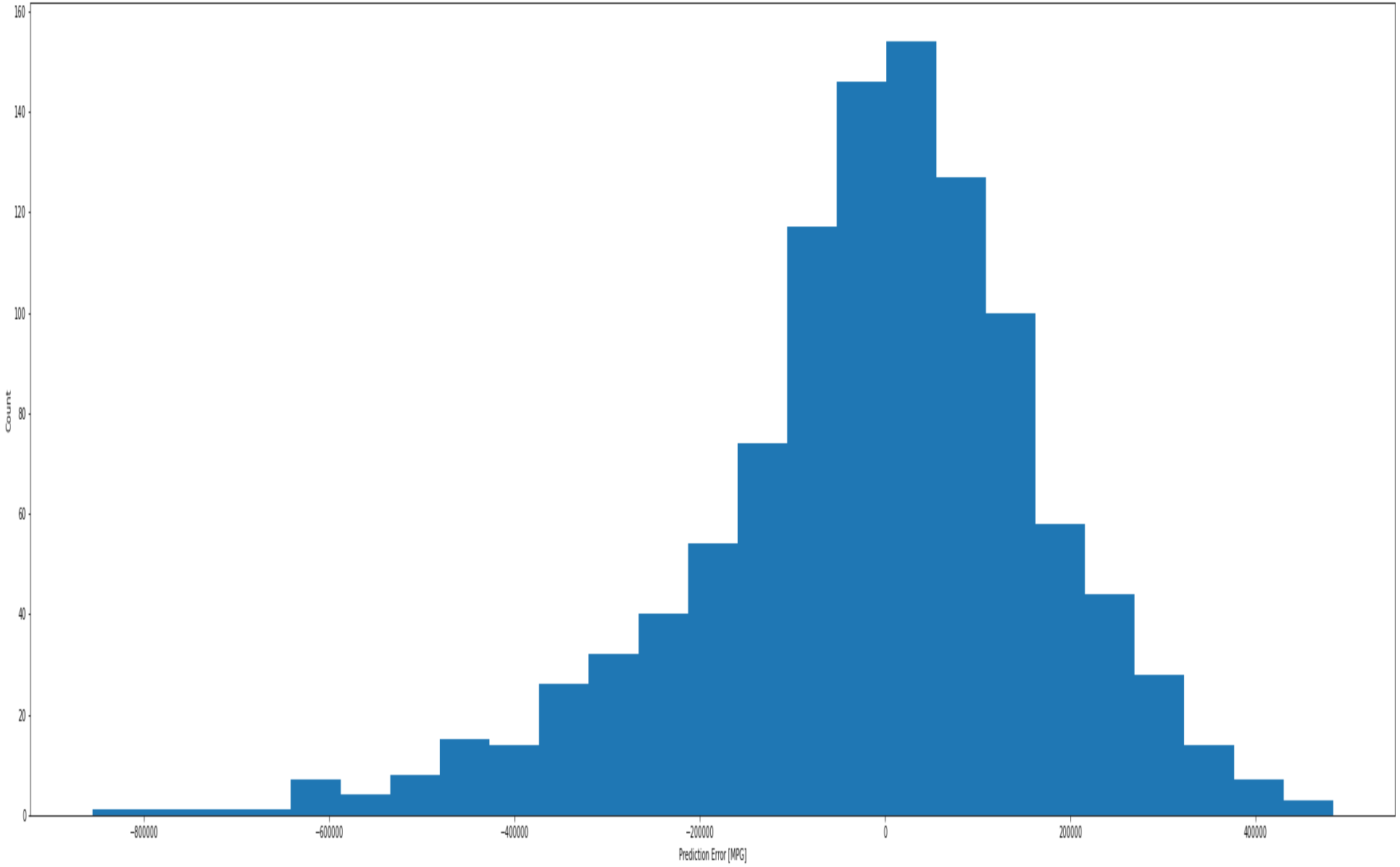
Figure(6.14) (a) :


Count
175
150
125
100
75
50
25
0
−1.0
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
1e6
Prediction Error [MPG]

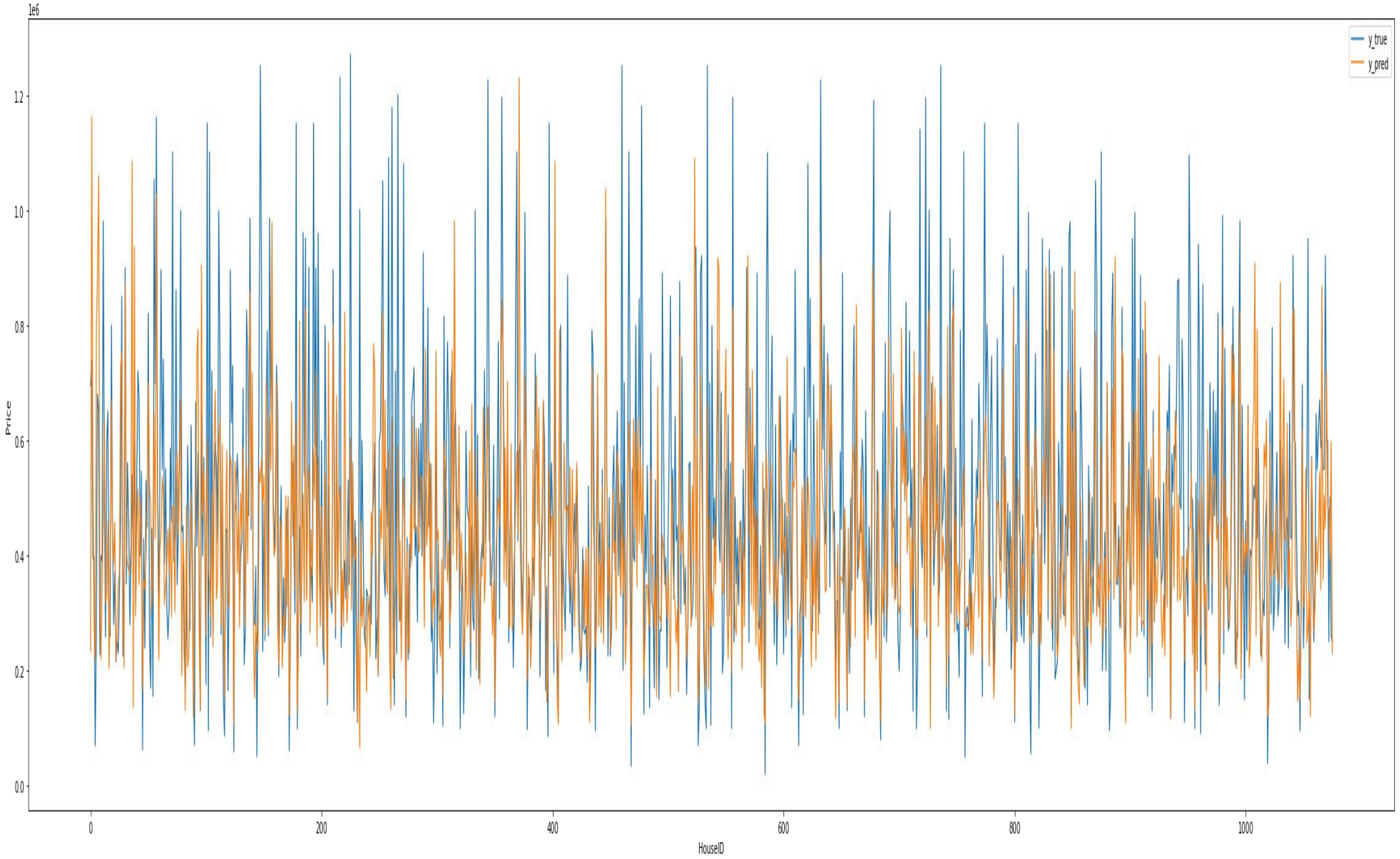
Figure(6.14) (b) :


175
150
125
100
75
50
25
0
Count
-800000
-600000
-400000
-200000
0
200000
400000
Prediction Error [MPG]

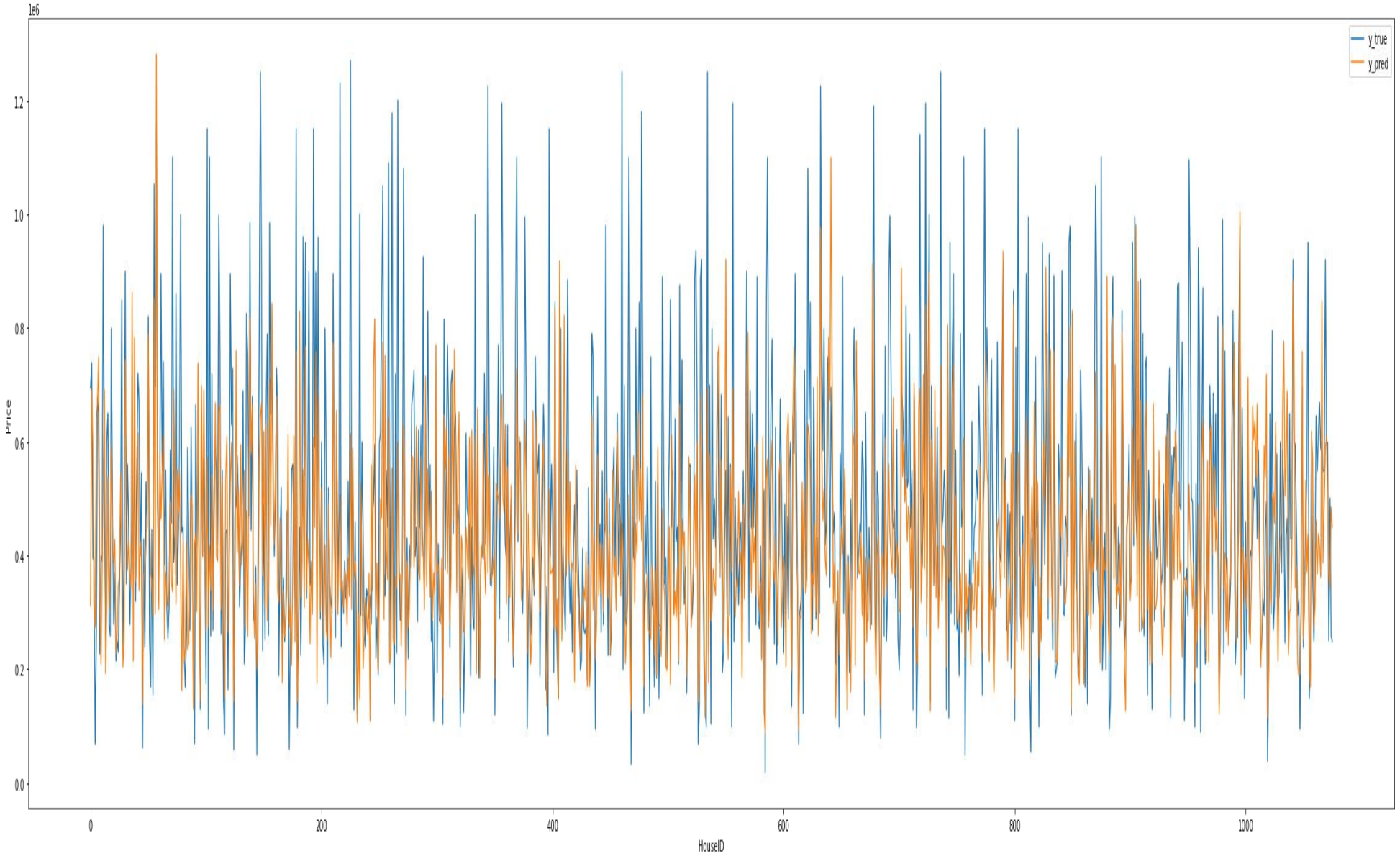
Figure(6.14) (c) :


160
140
120
100
80
60
40
20
0
Count
−800000
−600000
−400000
−200000
0
200000
400000
Prediction Error [MPG]

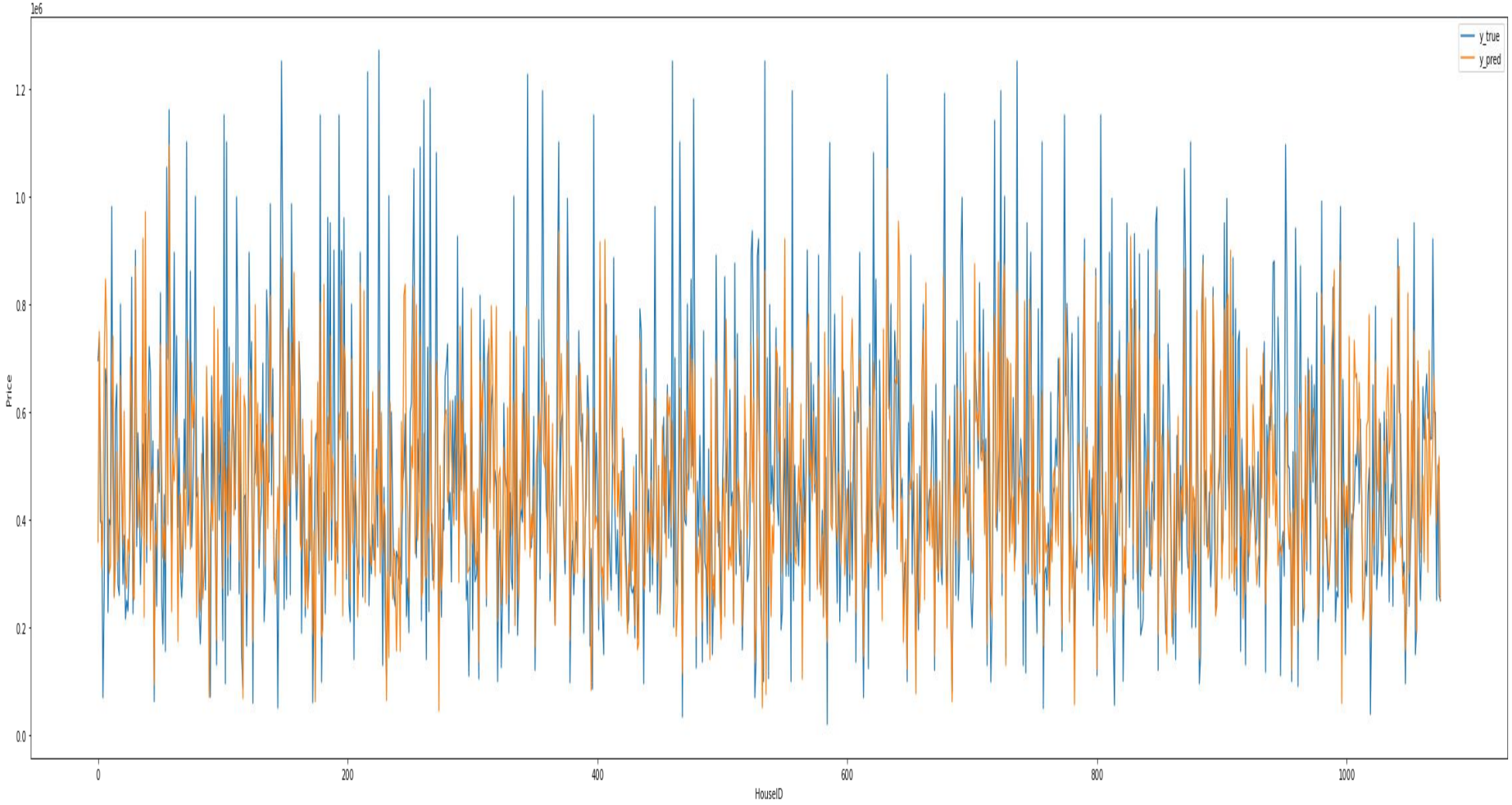
Figure(6.14) (a) :


1e6
y_true
y_pred
Price
HouseID
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
200
400
600
800
1000

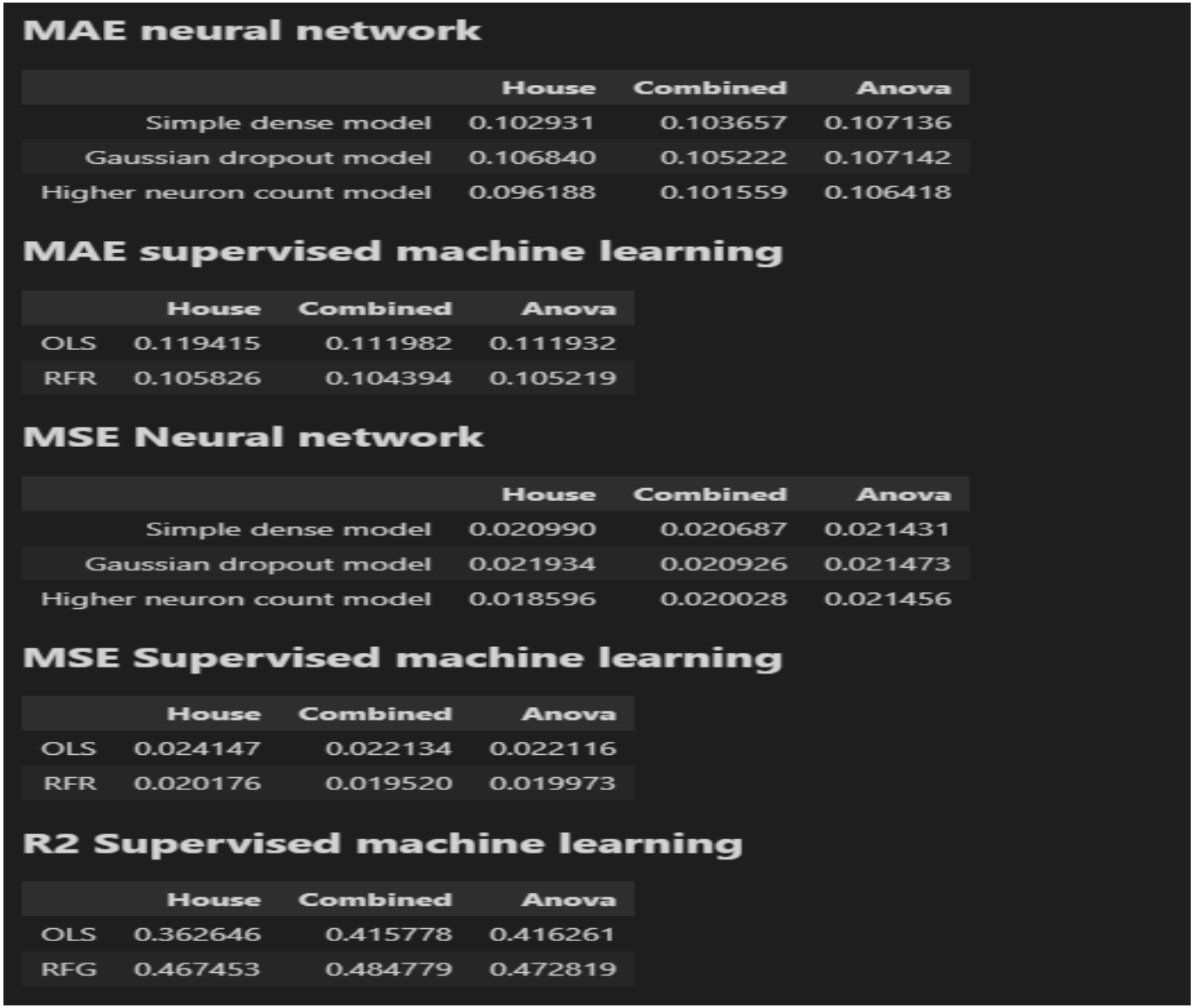
Figure(6.14) (b) :


1e6
y_true
y_pred
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Price
0
200
400
600
800
1000
HouseID

Figure(6.14) (c) :


1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Price
0
200
400
600
800
1000
HouseID
y_true
y_pred

Figure(6.15) :



MAE neural network

	House	Combined	Anova
Simple dense model	0.102931	0.103657	0.107136
Gaussian dropout model	0.106840	0.105222	0.107142
Higher neuron count model	0.096188	0.101559	0.106418

MAE supervised machine learning

	House	Combined	Anova
OLS	0.119415	0.111982	0.111932
RFR	0.105826	0.104394	0.105219

MSE Neural network

	House	Combined	Anova
Simple dense model	0.020990	0.020687	0.021431
Gaussian dropout model	0.021934	0.020926	0.021473
Higher neuron count model	0.018596	0.020028	0.021456

MSE Supervised machine learning

	House	Combined	Anova
OLS	0.024147	0.022134	0.022116
RFR	0.020176	0.019520	0.019973

R2 Supervised machine learning

	House	Combined	Anova
OLS	0.362646	0.415778	0.416261
RFG	0.467453	0.484779	0.472819