Project 1

Manager: Brendan McDonnel
Masum Billah
Madhu Chencharapu
Gabriel Loos
Roshan Ravichandran

February 2026

1 Introduction

2 Auto MPG

2.1 EDA

2.2 Quality of Fit for each model

Table 1: Scalation - Auto MPG In-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	log1p
rSq	0.809255	0.776580	0.809163	0.835138	0.849102
rSqBar	0.806283	0.772507	0.806189	0.832569	0.846751
sst	23819.0	23819.0	23819.0	23819.0	23819.0
sse	4543.35	5321.63	4545.54	3926.85	3594.23
sde	3.40878	3.68883	3.40888	3.16757	3.02514
mse0	11.5902	13.5756	11.5958	10.0175	9.16895
rmse	3.40443	3.68451	3.40526	3.16504	3.02803
mae	2.61826	2.79509	2.61703	2.34675	2.18422
smape	12.0589	65.4181	11.9861	10.2046	9.32001
m	392.000	392.000	392.000	392.000	392.000
dfr	6.00000	7.00000	6.00000	6.00000	6.00000
df	385.000	384.000	385.000	385.000	385.000
fStat	272.234	190.677	272.072	325.048	361.067
aic	-1022.45	-1051.45	-1022.55	-993.873	-976.527
bic	-994.656	-1019.68	-994.750	-966.074	-948.728

Table 2: Scalation - Auto MPG Out-of-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	log1p
rSq	0.822842	0.797491	0.822903	0.846480	0.852864
rSqBar	0.820081	0.793799	0.820143	0.844088	0.850571
sst	4731.23	4731.23	4731.23	4731.23	4731.23
sse	838.174	958.118	837.889	726.337	696.133
sde	3.29026	3.51709	3.28969	3.05724	2.97969
mse0	10.7458	12.2836	10.7422	9.31202	8.92478
rmse	3.27808	3.50479	3.27752	3.05156	2.98744
mae	2.48735	2.62052	2.48643	2.11846	1.98665
smape	11.8858	61.9272	11.8808	9.00068	8.28831
m	78.0000	78.0000	78.0000	78.0000	78.0000
dfr	6.00000	7.00000	6.00000	6.00000	6.00000
df	385.000	384.000	385.000	385.000	385.000
fStat	298.034	216.030	298.158	353.804	371.939
aic	-189.284	-192.500	-189.271	-183.700	-182.048
bic	-172.787	-173.646	-172.774	-167.203	-165.551

Table 3: Statsmodels - Auto MPG In-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	Log1p
rSq	0.8093	0.8037	0.6416	0.8477	0.8725
rSqBar	0.8063	0.8012	0.6369	0.8453	0.8705
sst	23818.9935	23818.9935	23818.9935	252.1610	41.1728
sse	4543.3470	4675.2421	8537.0473	38.4071	5.2483
sde	3.4352	3.4802	4.7028	0.3158	0.1168
mse0	11.8009	11.9266	21.7782	0.0998	0.0136
rmse	3.4352	3.4535	4.6667	0.3158	0.1168
mae	2.6183	2.6269	3.6246	2.3467	2.1842
smape	12.0589	12.0433	16.2905	10.2046	9.3200
m	392.0000	392.0000	392.0000	392.0000	392.0000
dfr	6.0000	6.0000	6.0000	6.0000	6.0000
df	385.0000	386.0000	386.0000	385.0000	385.0000
fStat	272.2341	263.4262	115.1614	357.1178	439.2179
aic	2086.9095	983.6796	1219.7162	215.8245	-564.3874
bic	2114.7083	1007.5071	1243.5438	243.6234	-536.5886

2.3 Regression

2.4 Ridge

2.5 Lasso

2.6 Sqrt

2.7 log1p

3 Housing Prices

3.1 EDA

3.2 Quality of Fit for each model

3.3 Regression

3.4 Ridge

3.5 Lasso

3.6 Sqrt

3.7 log1p

Table 4: Statsmodels - Auto MPG Out-of-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	Log1p
rSq	0.8107	0.7854	0.5925	0.8482	0.8720
rSqBar	0.8070	0.7707	0.5646	0.8452	0.8695
sst	19750.2199	4032.2061	4032.2061	208.0819	33.8141
sse	3738.2664	865.2680	1643.1024	31.5951	4.3287
sde	3.4952	3.4428	4.7443	0.3213	0.1189
mse0	12.2166	10.9528	20.7988	0.1033	0.0141
rmse	3.4952	3.3095	4.5606	0.3213	0.1189
mae	2.5039	2.5877	3.7495	2.1509	1.9533
smape	12.3880	12.5974	17.6830	9.8913	8.7248
m	79.0000	79.0000	79.0000	79.0000	79.0000
dfr	6.0000	6.0000	6.0000	6.0000	6.0000
df	306.0000	73.0000	73.0000	306.0000	306.0000
fStat	218.4461	44.5308	17.6906	284.8805	347.3956
aic	1678.5500	201.0937	251.7566	184.4835	-437.6799
bic	1704.7734	215.3104	265.9733	210.7069	-411.4565

Table 5: Scalation - Auto MPG Linear Regression CV

Name	num	min	max	mean	stdev	interval
rSq	5	0.788	0.823	0.798	0.014	0.018
rSqBar	5	0.785	0.820	0.795	0.014	0.018
sst	5	3962.818	5671.580	4700.481	620.767	770.935
sse	5	824.554	1176.435	950.494	142.696	177.215
sde	5	3.177	3.738	3.431	0.226	0.281
mse0	5	10.571	15.083	12.186	1.829	2.272
rmse	5	3.251	3.884	3.483	0.256	0.318
mae	5	2.487	2.850	2.689	0.151	0.188
smape	5	11.886	12.905	12.372	0.427	0.530
m	5	78.000	78.000	78.000	0.000	0.000
dfr	5	6.000	6.000	6.000	0.000	0.000
df	5	385.000	385.000	385.000	0.000	0.000
fStat	5	239.054	298.034	254.430	24.517	30.448
aic	5	-205.110	-188.647	-194.539	6.676	8.291
bic	5	-188.613	-172.150	-178.042	6.676	8.291

Table 6: Statsmodels - Auto MPG Linear Regression CV

Name	In-num folds	min	max	mean	stdev
rSq	5	0.7654	0.8282	0.8010	0.0216
rSqBar	5	0.7458	0.8139	0.7845	0.0234
sst	5	4032.2061	5792.1365	4724.2751	617.8288
sse	5	745.1709	1359.0577	947.8346	213.7052
sde	5	3.2171	4.3446	3.5980	0.3912
mse0	5	51.0406	74.2582	60.2918	8.1328
rmse	5	3.2171	4.3446	3.5980	0.3912
mae	5	2.5039	3.1904	2.6786	0.2601
smape	5	11.2379	14.1325	12.3805	0.9795

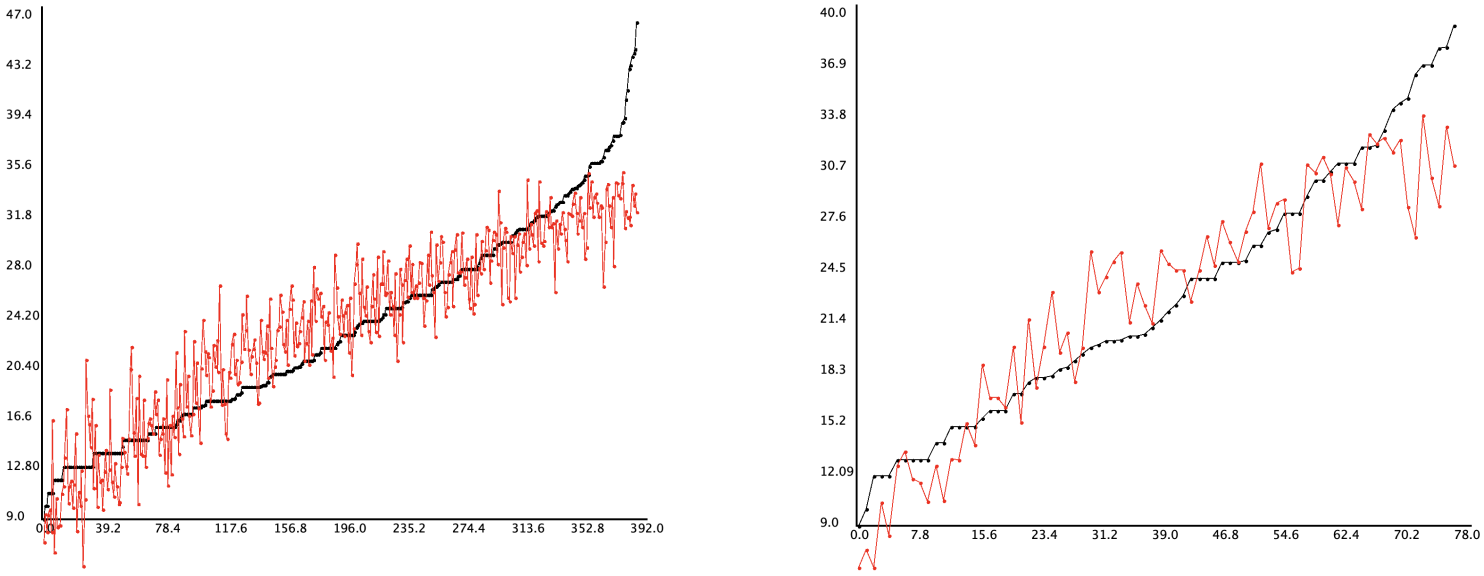


Figure 1: Scalation - Auto MPG Regression
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

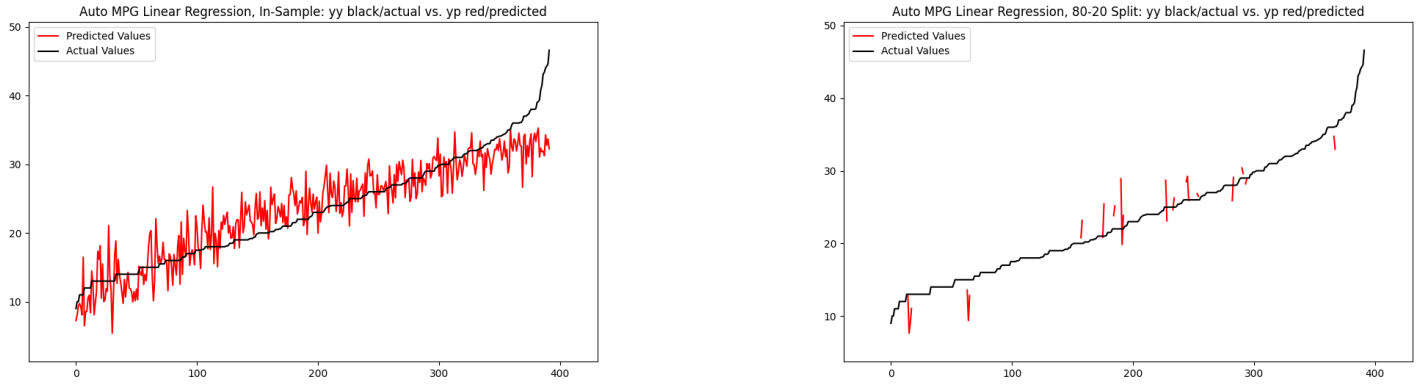


Figure 2: Statsmodels - Auto MPG Regression
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

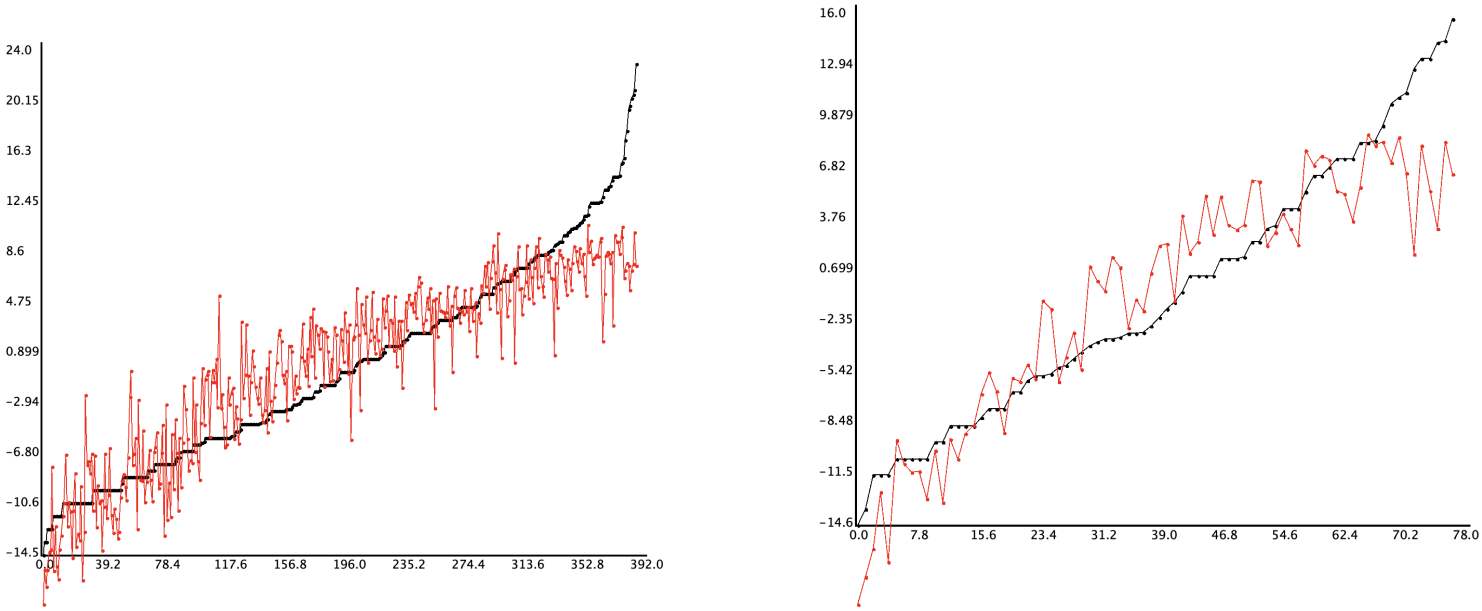


Figure 3: Scalation - Auto MPG Ridge
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

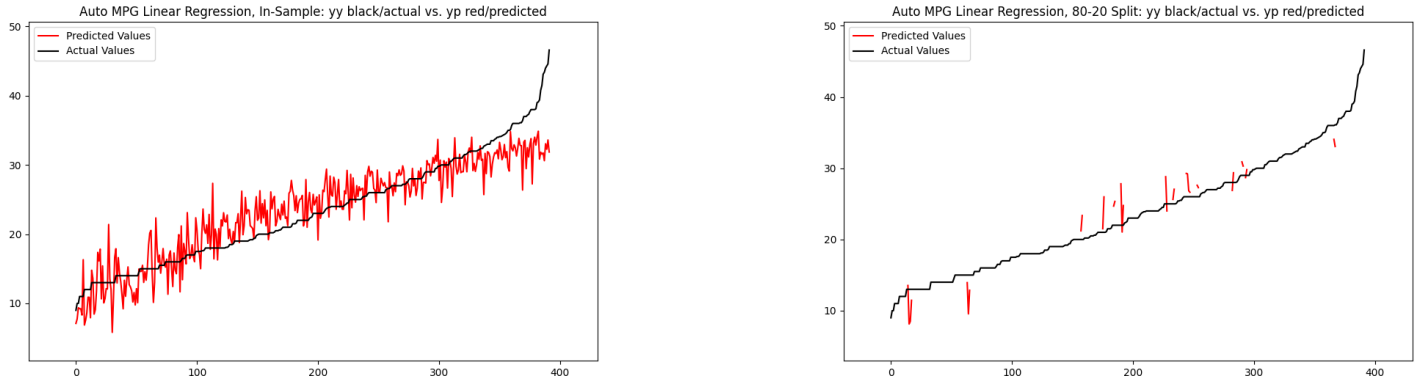


Figure 4: Statsmodels - Auto MPG Ridge
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

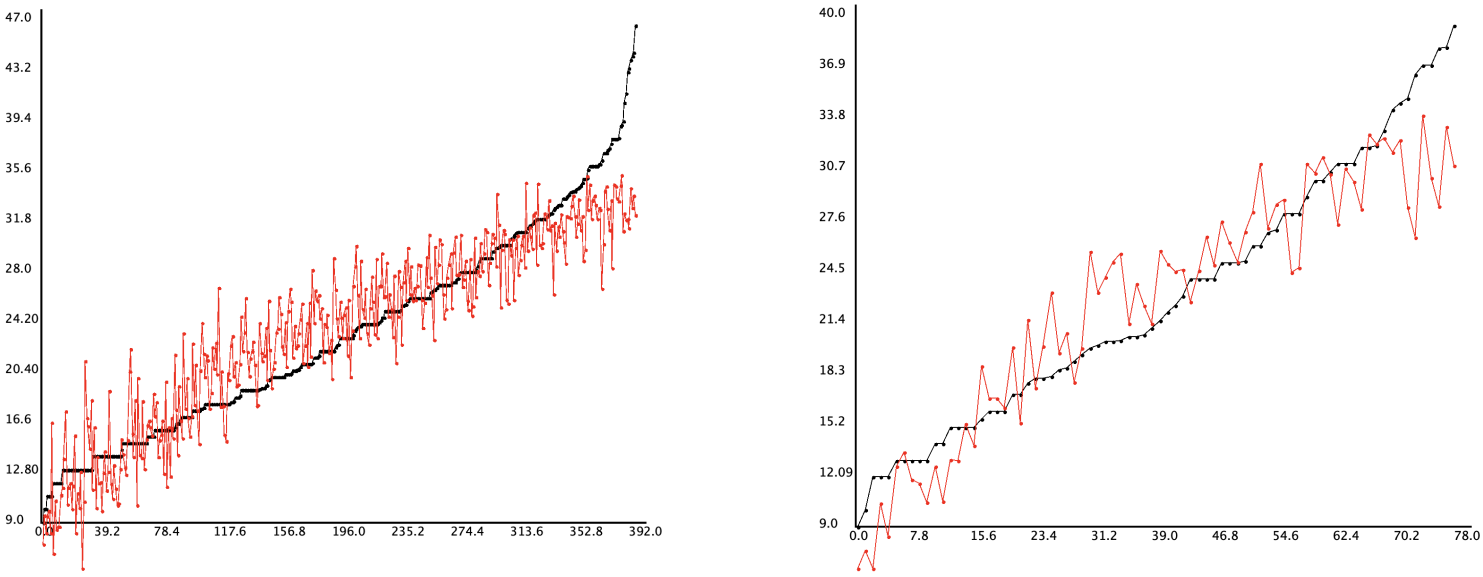


Figure 5: Scalation - Auto MPG Lasso
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

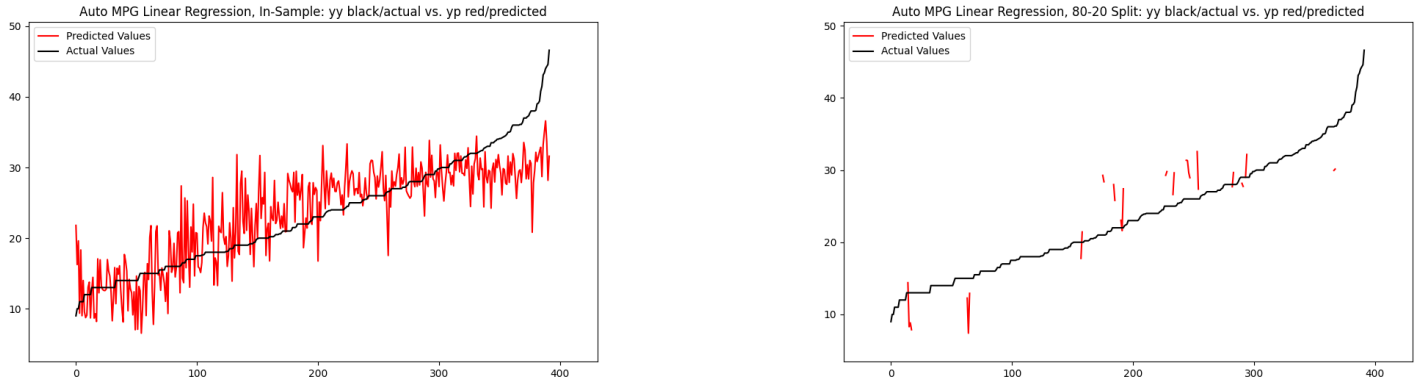


Figure 6: Statsmodels - Auto MPG Lasso
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

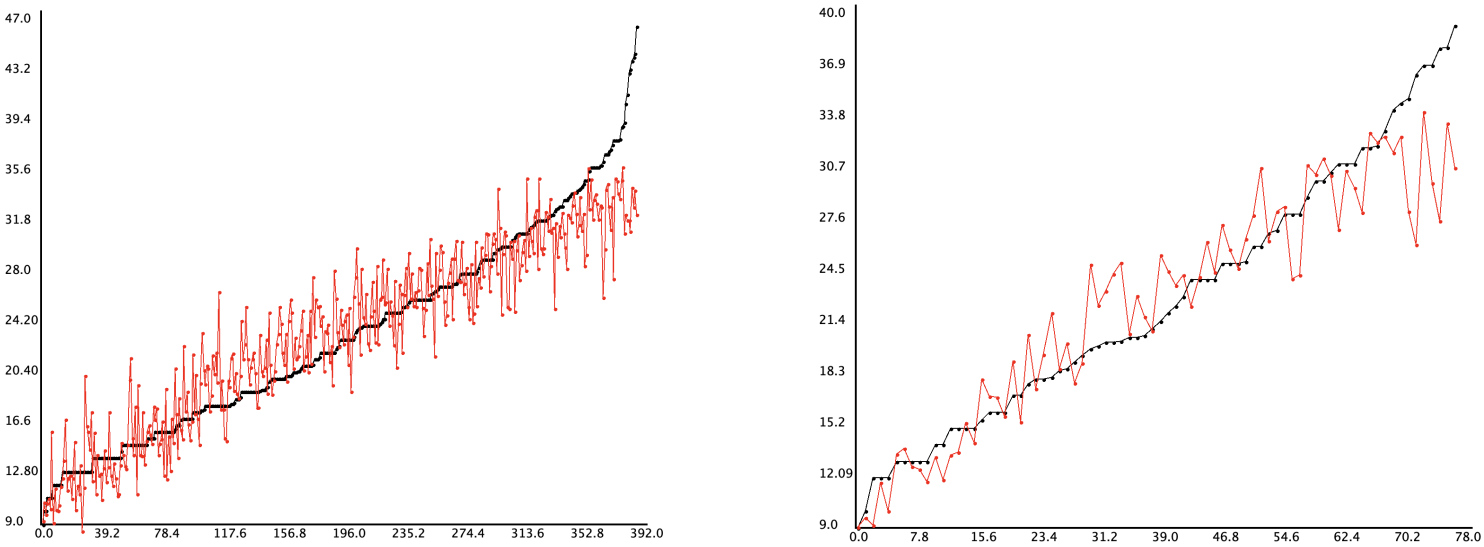


Figure 7: Scalation - Auto MPG Sqrt
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

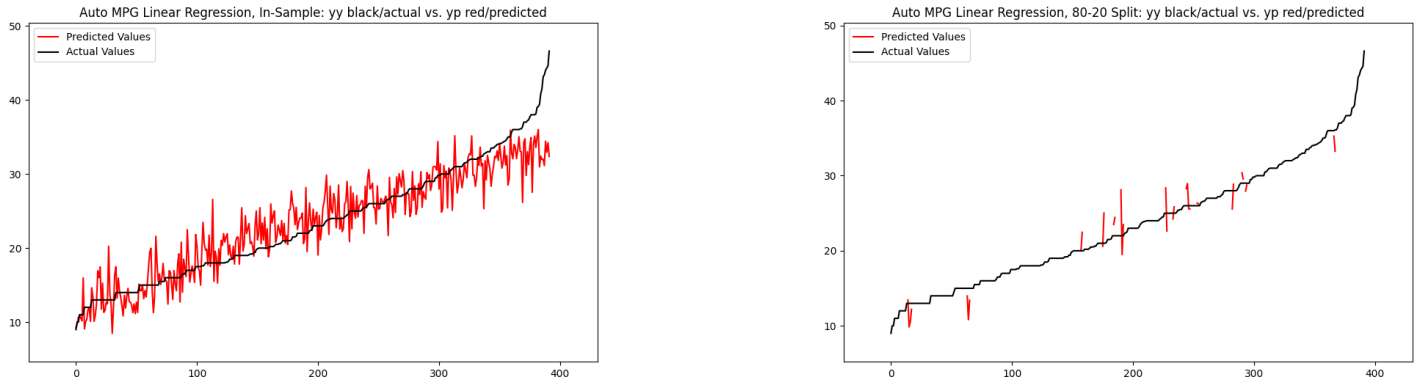


Figure 8: Statsmodels - Auto MPG Sqrt
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

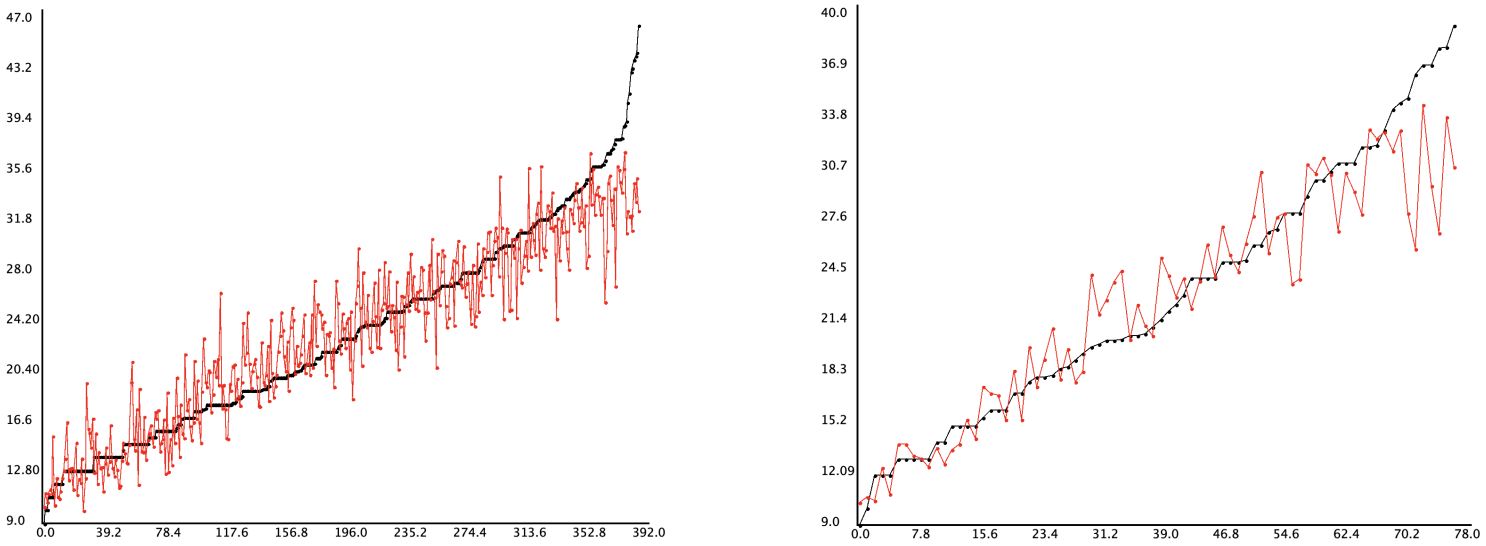


Figure 9: Scalation - Auto MPG log1p
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

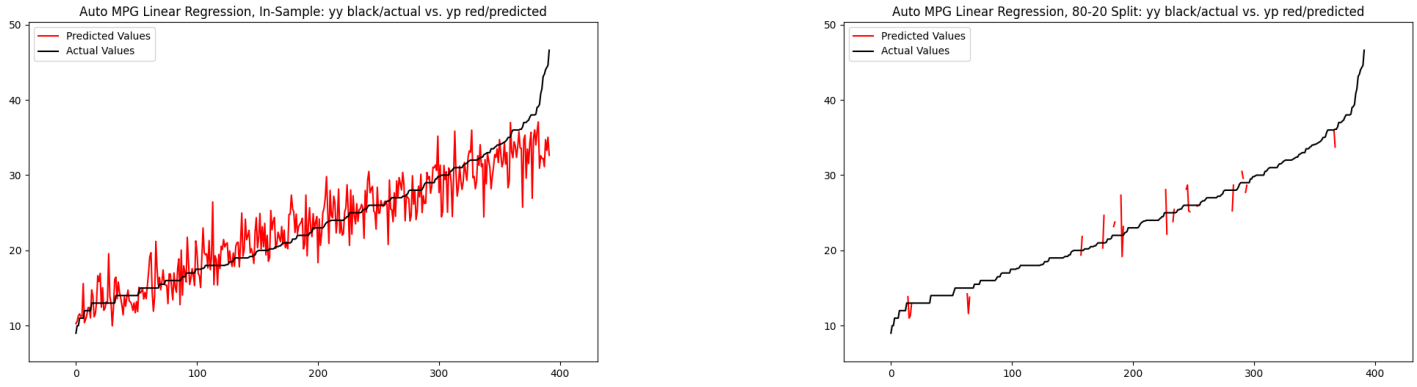


Figure 10: Statsmodels - Auto MPG log1p
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

Table 7: Scalation - Housing Prices In-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	log1p
rSq	0.998516	0.987398	0.998516	0.986152	0.922307
rSqBar	0.998507	0.987309	0.998507	0.986068	0.921838
sst	6.42325e+13	6.42325e+13	6.42325e+13	6.42325e+13	6.42325e+13
sse	9.53030e+10	8.09438e+11	9.53030e+10	8.89504e+11	4.99039e+12
sde	9767.21	28463.5	9767.21	29836.4	70604.5
mse0	9.53030e+07	8.09438e+08	9.53030e+07	8.89504e+08	4.99039e+09
rmse	9762.32	28450.6	9762.32	29824.5	70642.7
mae	7747.66	24147.7	7747.66	24309.3	53070.5
smape	1.57791	23.5576	1.57791	4.86164	9.22319
m	1000.00	1000.00	1000.00	1000.00	1000.00
dfr	6.00000	7.00000	6.00000	6.00000	6.00000
df	993.000	992.000	993.000	993.000	993.000
fStat	111378	11103.9	111378	11785.5	1964.69
aic	-10591.2	-11658.9	-10591.2	-11708.0	-12570.3
bic	-10556.9	-11619.6	-10556.9	-11673.7	-12536.0

Table 8: Scalation - Housing Prices Out-of-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	log1p
rSq	0.998649	0.989193	0.998649	0.984018	0.906860
rSqBar	0.998641	0.989117	0.998641	0.983921	0.906297
sst	1.33700e+13	1.33700e+13	1.33700e+13	1.33700e+13	1.33700e+13
sse	1.80638e+10	1.44485e+11	1.80638e+10	2.13678e+11	1.24528e+12
sde	9501.56	26880.4	9501.56	32757.9	78673.0
mse0	9.03190e+07	7.22424e+08	9.03189e+07	1.06839e+09	6.22641e+09
rmse	9503.63	26877.9	9503.63	32686.3	78907.6
mae	7484.48	22419.4	7484.48	26186.8	58148.8
smape	1.60847	19.6544	1.60847	5.19788	9.91629
m	200.000	200.000	200.000	200.000	200.000
dfr	6.00000	7.00000	6.00000	6.00000	6.00000
df	993.000	992.000	993.000	993.000	993.000
fStat	122330	12971.9	122330	10189.9	1611.39
aic	-2101.67	-2307.60	-2101.67	-2348.73	-2525.00
bic	-2078.59	-2281.21	-2078.59	-2325.64	-2501.91

Table 9: Statsmodels - House Price In-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	Log1p
rSq	0.9985	0.9906	0.9875	0.9855	0.9415
rSqBar	0.9985	0.9905	0.9875	0.9854	0.9411
sst	64232463468052.5469	64232463468052.5469	64232463468052.5469	29445342.0661	241.7492
sse	95249090298.3967	604344645977.1542	800169699033.4265	425598.9981	14.1461
sde	9798.8381	24669.9185	28386.7992	20.7131	0.1194
mse0	96017228.1234	604344645.9772	800169699.0334	429.0312	0.0143
rmse	9798.8381	24583.4222	28287.2710	20.7131	0.1194
mae	7740.4301	20205.2499	24002.5173	24296.4685	52989.2927
smape	1.5774	4.1035	4.9163	4.8579	9.2122
m	1000.0000	1000.0000	1000.0000	1000.0000	1000.0000
dfr	7.0000	7.0000	7.0000	7.0000	7.0000
df	992.0000	993.0000	993.0000	992.0000	992.0000
fStat	95425.1583	14935.3572	11245.5195	9662.8803	2280.1137
aic	21225.8831	20233.6552	20514.3344	8907.3747	-1404.4424
bic	21265.1451	20268.0095	20548.6887	8946.6367	-1365.1804

Table 10: Statsmodels - House Price Out-of-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	Log1p
rSq	0.9985	0.9913	0.9884	0.9852	0.9402
rSqBar	0.9985	0.9910	0.9880	0.9851	0.9397
sst	51340367232368.6250	12891771417242.6445	12891771417242.6445	23544096.3448	193.6629
sse	75080425564.8099	112341107192.6465	149469076702.9745	347277.4580	11.5792
sde	9736.4530	24126.2984	27828.9629	20.9400	0.1209
mse0	94798517.1273	561705535.9632	747345383.5149	438.4816	0.0146
rmse	9736.4530	23700.3278	27337.6185	20.9400	0.1209
mae	8174.5836	19629.7981	23012.0161	23553.0274	51521.6937
smape	1.6620	3.9691	4.6112	4.7594	9.0567
m	200.0000	200.0000	200.0000	200.0000	200.0000
dfr	7.0000	7.0000	7.0000	7.0000	7.0000
df	792.0000	193.0000	193.0000	792.0000	792.0000
fStat	77254.5038	3136.4045	2350.4760	7557.5143	1779.1786
aic	16972.0728	4043.2977	4100.4076	7144.9157	-1102.0191
bic	17009.5497	4066.3859	4123.4958	7182.3926	-1064.5422

Table 11: Scalation - House Price Linear Regression CV

Name	num folds	min	max	mean	stdev	interval
rSq	5	0.998	0.999	0.998	0.000	0.000
rSqBar	5	0.998	0.999	0.998	0.000	0.000
sst	5	11805100457351.200	13369989427686.710	12829460198984.865	606159535374.424	752793908917.133
sse	5	17790794374.705	22663619750.593	19394433560.564	2117948592.845	2630295668.134
sde	5	9365.942	10608.530	9818.348	533.068	662.021
mse0	5	88953971.874	113318098.753	96972167.803	10589742.964	13151478.341
rmse	5	9431.541	10645.097	9836.093	528.486	656.330
mae	5	7418.536	8609.308	7817.189	534.033	663.219
smape	5	1.408	1.804	1.592	0.141	0.176
m	5	200.000	200.000	200.000	0.000	0.000
dfr	5	6.000	6.000	6.000	0.000	0.000
df	5	993.000	993.000	993.000	0.000	0.000
fStat	5	96001.466	122329.999	110142.703	10843.190	13466.236
aic	5	-2127.278	-2100.155	-2109.082	11.789	14.641
bic	5	-2104.190	-2077.067	-2085.993	11.789	14.641

Table 12: Statsmodels - House Price Linear Regression CV

Name	In-num folds	min	max	mean	stdev
rSq	5	0.9984	0.9986	0.9985	0.0001
rSqBar	5	0.9984	0.9986	0.9984	0.0001
sst	5	12384264865084.2500	13564576490393.4668	12842056251568.3359	395911885307.9127
sse	5	17238729843.0851	20996608313.5931	19277864710.9665	1296174615.0581
sde	5	9450.9176	10430.2788	9988.5569	337.6978
mse0	5	61921324325.4212	67822882451.9673	64210281257.8417	1979559426.5396
rmse	5	9450.9176	10430.2788	9988.5569	337.6978
mae	5	7516.9137	8174.5836	7794.6342	224.6310
smape	5	1.5104	1.6620	1.5879	0.0540

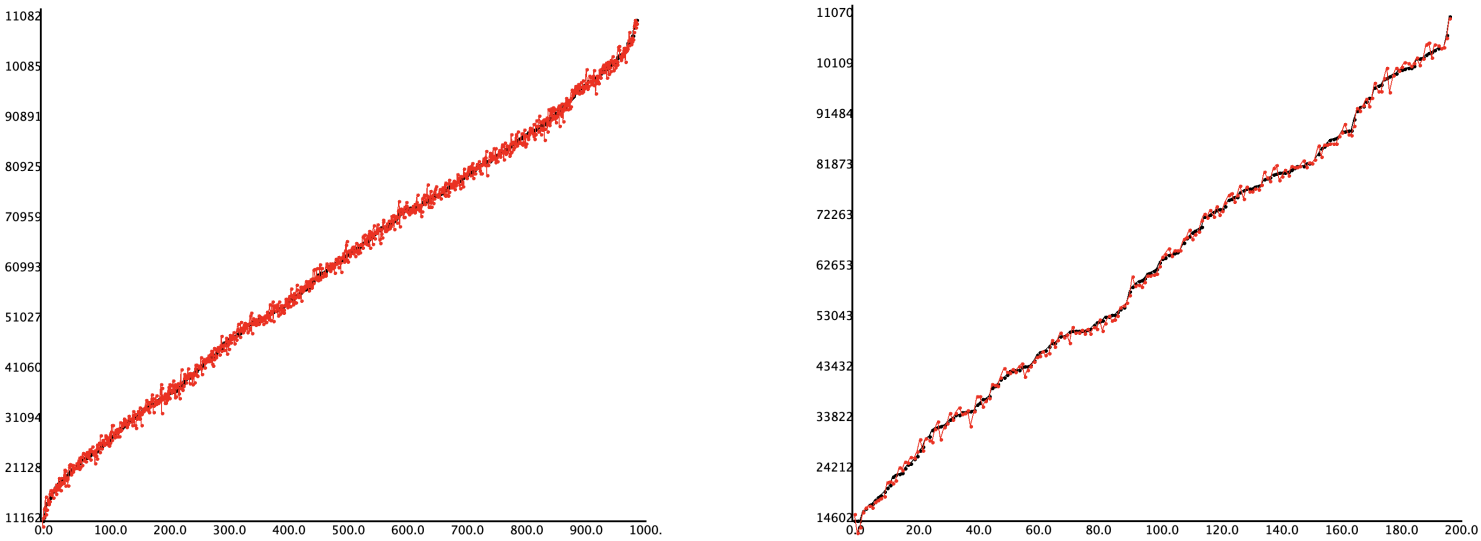


Figure 11: Scalation - House Price Regression
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

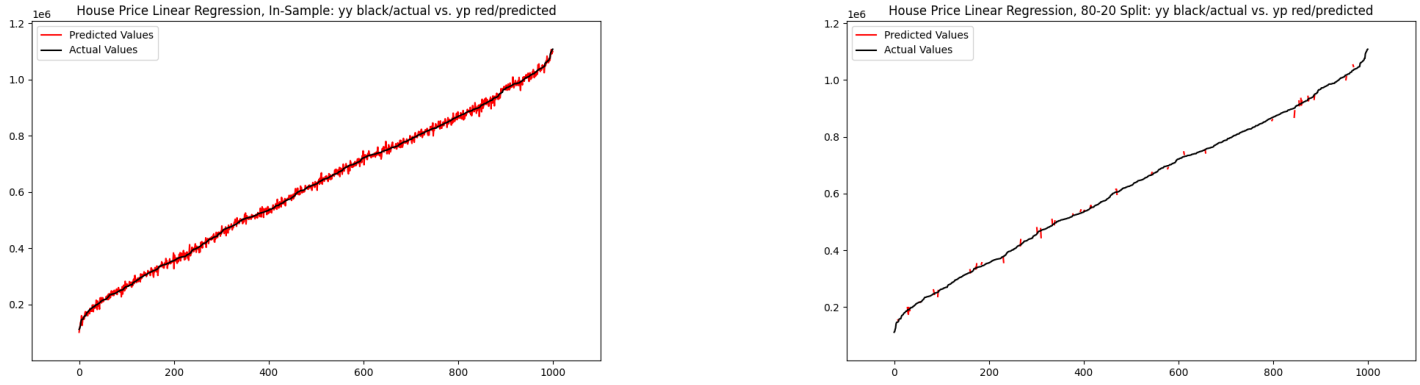


Figure 12: Statsmodels - House Price Regression
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

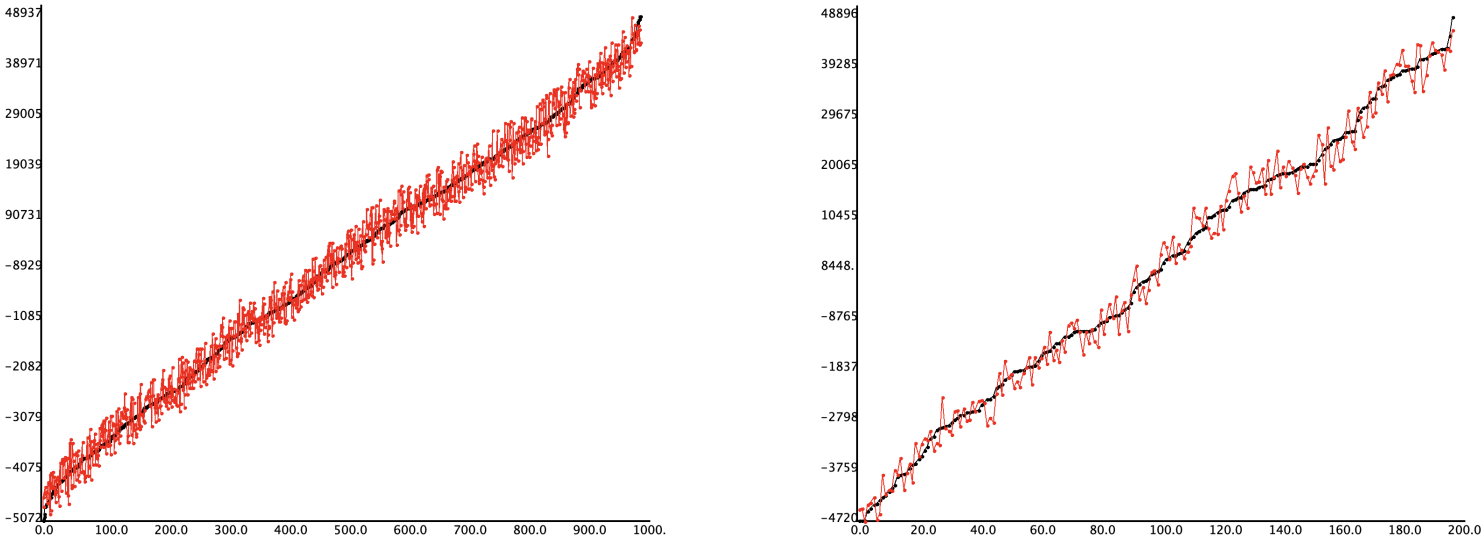


Figure 13: Scalation - House Price Ridge
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

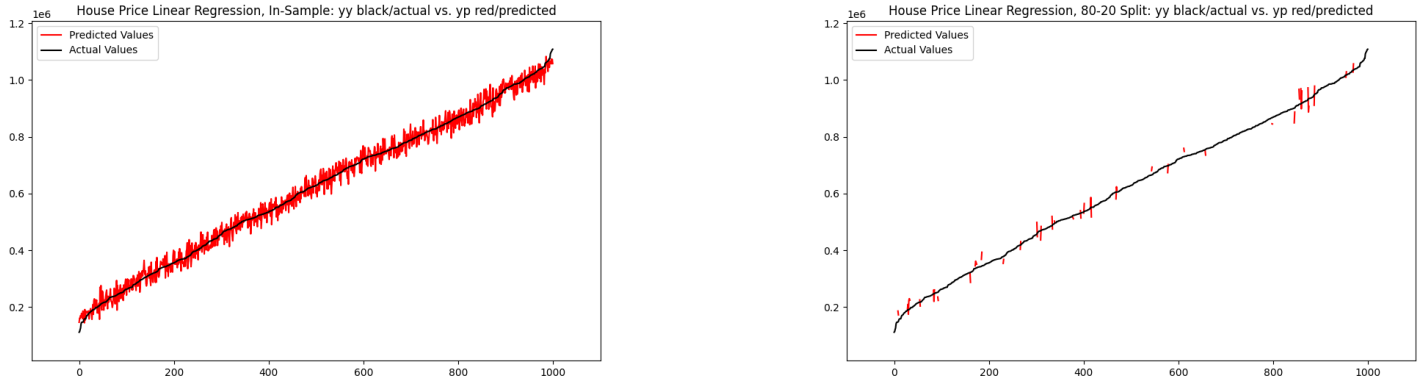


Figure 14: Statsmodels - House Price Ridge
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

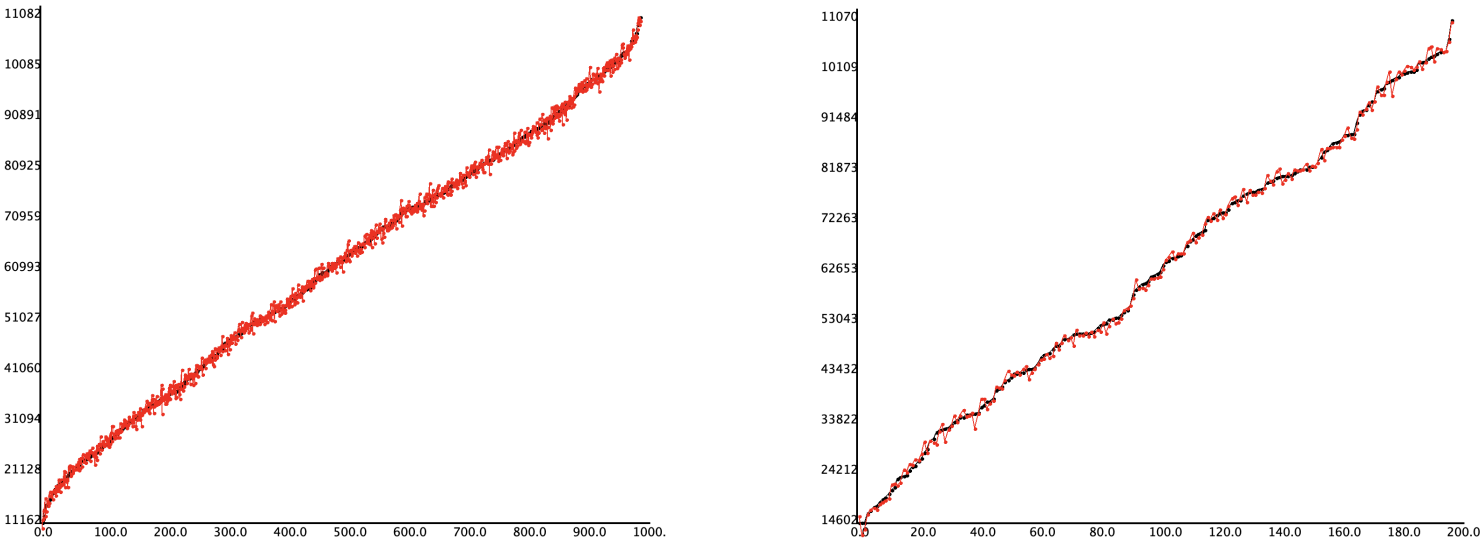
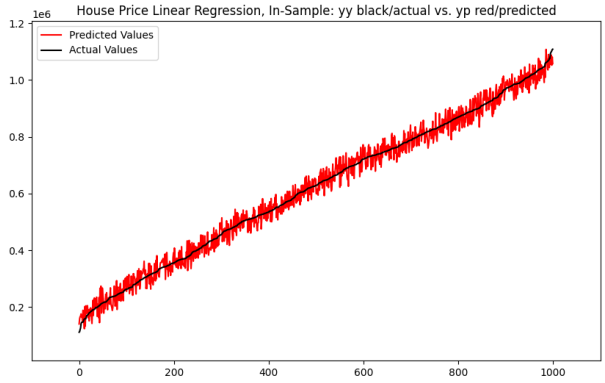


Figure 15: Scalation - House Price Lasso
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted



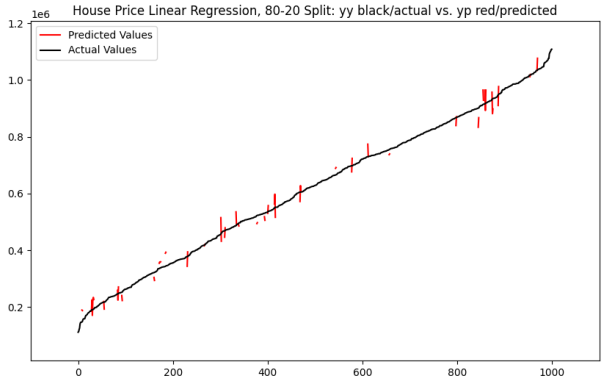
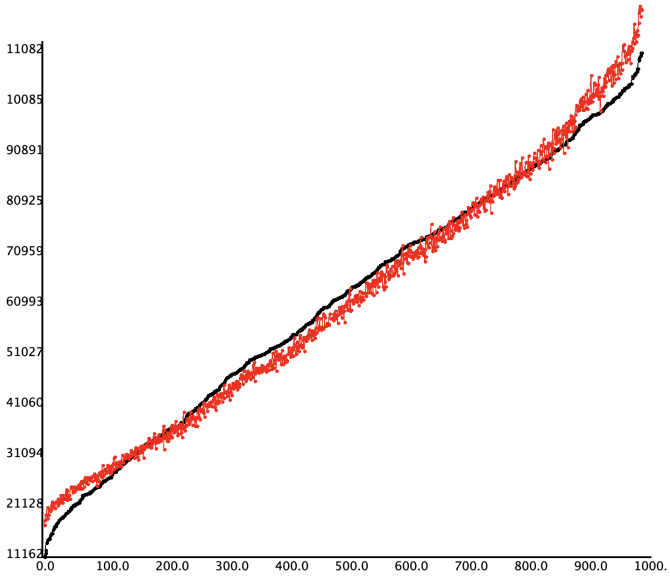


Figure 16: Statsmodels - House Price Lasso
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted



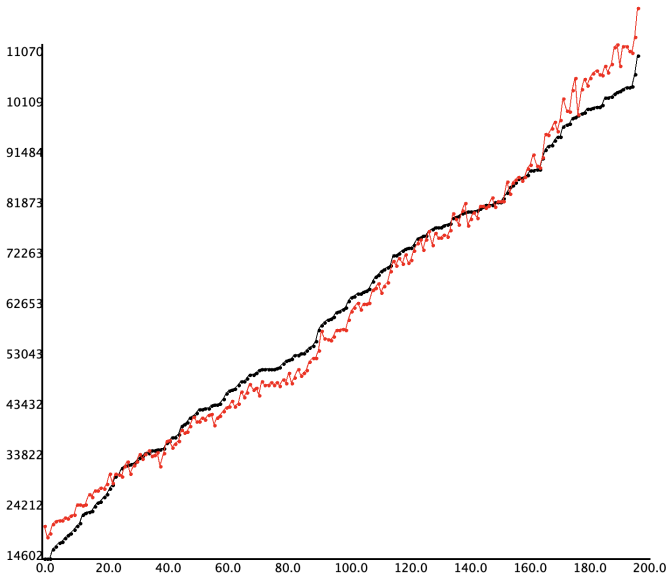


Figure 17: Scalation - House Price Sqrt
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

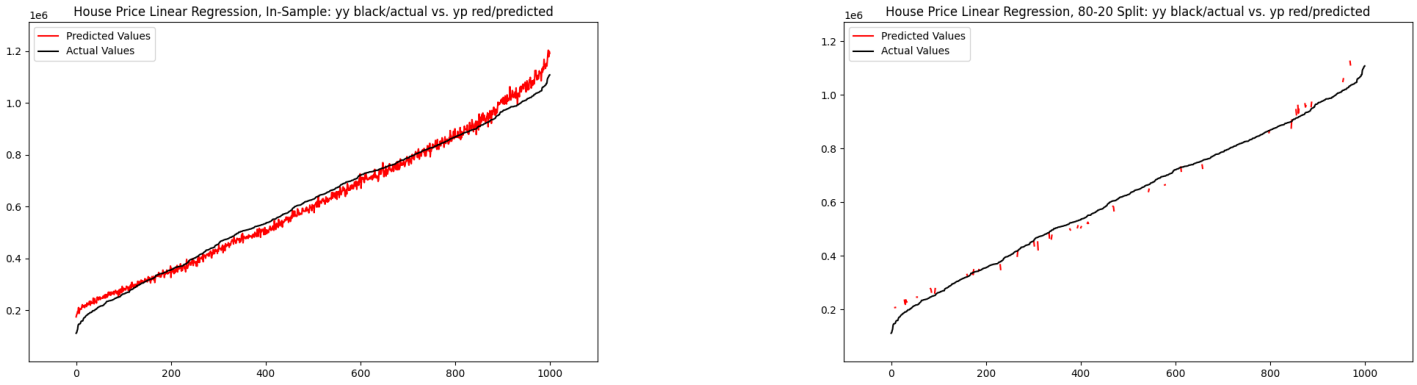


Figure 18: Statsmodels - House Price Sqrt
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

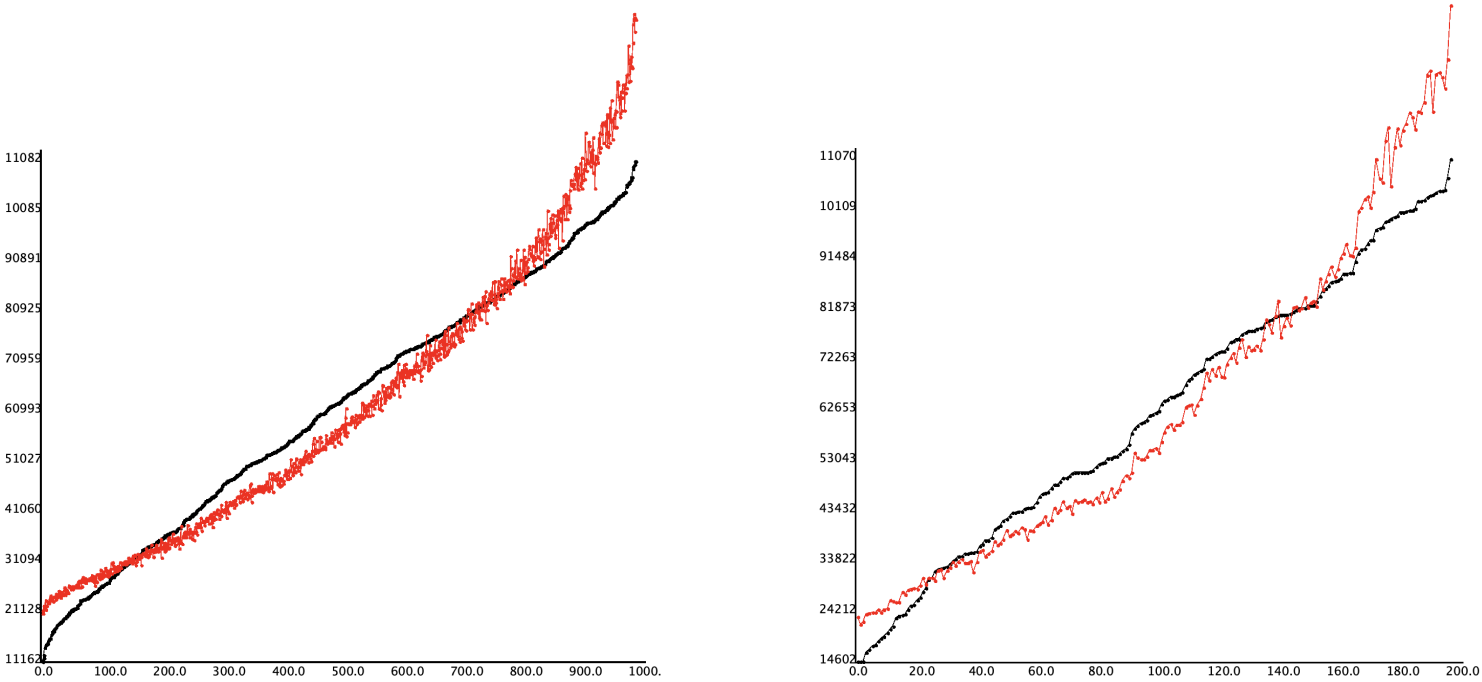
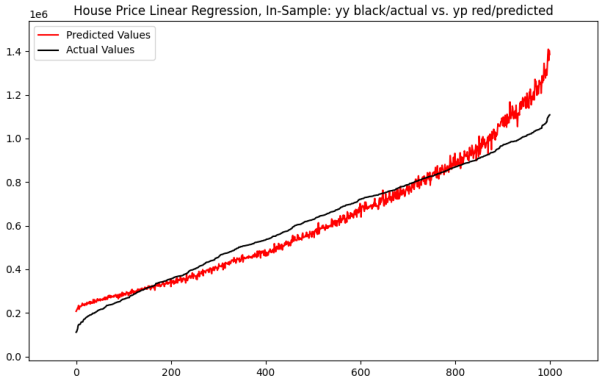


Figure 19: Scalation - House Price log1p
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted



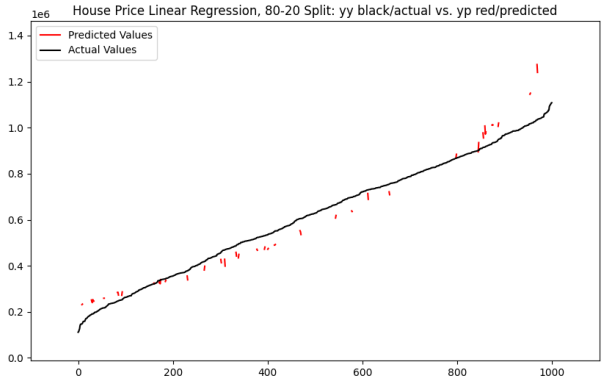


Figure 20: Statsmodels - House Price log1p
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

Table 13: Scalation - Insurance Charges In-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	log1p
rSq	0.750157	0.749962	0.750157	0.752656	0.527431
rSqBar	0.748842	0.748457	0.748842	0.751354	0.524943
sst	1.96074e+11	1.96074e+11	1.96074e+11	1.96074e+11	1.96074e+11
sse	4.89878e+10	4.90260e+10	4.89878e+10	4.84977e+10	9.26587e+10
sde	6053.11	6055.42	6053.11	6001.45	8316.89
mse0	3.66127e+07	3.66413e+07	3.66127e+07	3.62464e+07	6.92516e+07
rmse	6050.84	6053.20	6050.84	6020.50	8321.76
mae	4179.54	4150.46	4179.54	3623.20	4215.04
smape	37.9722	67.8472	37.9722	27.9341	26.5034
m	1338.00	1338.00	1338.00	1338.00	1338.00
dfr	7.00000	8.00000	7.00000	7.00000	7.00000
df	1330.00	1329.00	1330.00	1330.00	1330.00
fStat	570.477	498.274	570.477	578.161	212.057
aic	-13533.8	-13532.3	-13533.8	-13527.1	-13960.2
bic	-13492.2	-13485.5	-13492.2	-13485.5	-13918.6

Table 14: Scalation - Insurance Charges Out-of-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	log1p
rSq	0.720005	0.719714	0.720005	0.730926	0.583107
rSqBar	0.718531	0.718027	0.718531	0.729510	0.580913
sst	4.06432e+10	4.06432e+10	4.06432e+10	4.06432e+10	4.06432e+10
sse	1.13799e+10	1.13917e+10	1.13799e+10	1.09360e+10	1.69438e+10
sde	6540.46	6543.82	6540.46	6396.98	7972.86
mse0	4.26213e+07	4.26655e+07	4.26213e+07	4.09588e+07	6.34601e+07
rmse	6528.50	6531.89	6528.50	6399.91	7966.18
mae	4430.66	4429.38	4430.66	3868.65	4087.97
smape	40.0602	62.9178	40.0602	30.5895	27.5909
m	267.000	267.000	267.000	267.000	267.000
dfr	7.00000	8.00000	7.00000	7.00000	7.00000
df	1330.00	1329.00	1330.00	1330.00	1330.00
fStat	488.584	426.574	488.584	516.126	265.753
aic	-2708.17	-2706.31	-2708.17	-2702.86	-2761.31
bic	-2679.47	-2674.02	-2679.47	-2674.16	-2732.61

Table 15: Statsmodels - Insurance Charges In-Sample QoF Comparison

Metric	Regression	Ridge	Lasso	Sqrt	Log1p
rSq	0.7509	0.2718	0.7469	0.7795	0.7680
rSqBar	0.7494	0.2680	0.7456	0.7782	0.7666
sst	196074221568.3671	196074221568.3671	196074221568.3671	3051091.5131	1130.1100
sse	48839532843.9219	142780871945.3538	49623455821.6671	672635.9707	262.2315
sde	6062.1023	10361.1794	6108.2624	22.4972	0.4442
mse0	36749084.1564	106712161.3941	37087784.6201	506.1219	0.1973
rmse	6062.1023	10330.1579	6089.9741	22.4972	0.4442
mae	4170.8869	8134.7714	4115.6884	3613.8958	4219.5115
smape	37.8059	66.1576	34.1997	27.6903	26.2889
m	1338.0000	1338.0000	1338.0000	1338.0000	1338.0000
dfr	8.0000	8.0000	8.0000	8.0000	8.0000
df	1329.0000	1330.0000	1330.0000	1329.0000	1329.0000
fStat	500.8107	62.0533	490.6438	587.4216	549.8054
aic	27113.5058	24749.7939	23335.7320	12137.4774	1634.5362
bic	27160.2962	24791.3854	23377.3235	12184.2678	1681.3266

Table 16: Statsmodels - Insurance Charges Out-of-Sample QoF Comparison

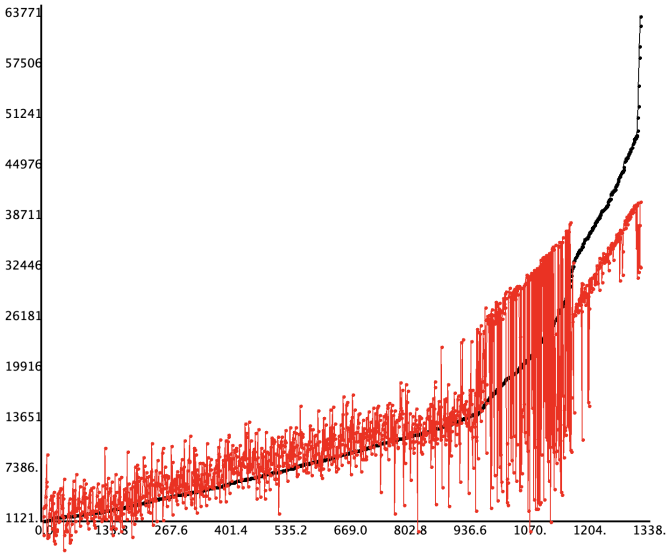
Metric	Regression	Ridge	Lasso	Sqrt	Log1p
rSq	0.7417	0.2951	0.7797	0.7701	0.7572
rSqBar	0.7398	0.2761	0.7738	0.7684	0.7554
sst	154436975468.4681	41606660039.7953	41606660039.7953	2399231.6774	888.0691
sse	39887119421.1600	29328394351.2134	9164643207.2381	551569.9183	215.5882
sde	6131.3858	10620.8058	5937.0555	22.8004	0.4508
mse0	37593892.0086	109434307.2806	34196429.8778	519.8585	0.2032
rmse	6131.3858	10461.0854	5847.7714	22.8004	0.4508
mae	4181.1945	8328.6585	4069.5309	3556.9640	3888.4432
smape	40.0220	69.4557	35.5690	29.1245	25.7136
m	268.0000	268.0000	268.0000	268.0000	268.0000
dfr	8.0000	8.0000	8.0000	8.0000	8.0000
df	1061.0000	260.0000	260.0000	1061.0000	1061.0000
fStat	380.8792	13.6061	115.0471	444.2703	413.6950
aic	21708.8072	4976.9038	4665.1653	9736.7961	1340.3418
bic	21753.5859	5005.6317	4693.8932	9781.5748	1385.1205

Table 17: Scalation - Insurance Charges Linear Regression CV

Name	num	min	max	mean	stdev	interval
rSq	5	0.701	0.814	0.743	0.046	0.057
rSqBar	5	0.699	0.813	0.742	0.046	0.057
sst	5	31902777173.848	43430486230.657	38949749086.422	4343029725.651	5393640012.108
sse	5	7539480548.420	11454966761.392	9918152392.401	1670500799.560	2074607019.047
sde	5	5320.957	6562.018	6084.654	526.821	654.263
mse0	5	28237754.863	42902497.234	37146638.174	6256557.302	7770063.742
rmse	5	5313.921	6550.000	6076.688	525.006	652.009
mae	5	3810.754	4511.498	4216.703	292.688	363.491
smape	5	36.128	40.060	38.143	1.833	2.277
m	5	267.000	267.000	267.000	0.000	0.000
dfr	5	7.000	7.000	7.000	0.000	0.000
df	5	1330.000	1330.000	1330.000	0.000	0.000
fStat	5	444.882	830.519	573.015	157.534	195.643
aic	5	-2709.047	-2663.110	-2691.017	19.599	24.340
bic	5	-2680.349	-2634.412	-2662.319	19.599	24.340

Table 18: Statsmodels - Insurance Charges Linear Regression CV

Name	In-num folds	min	max	mean	stdev
rSq	5	0.6324	0.7956	0.7402	0.0578
rSqBar	5	0.6211	0.7893	0.7322	0.0596
sst	5	30189024179.7055	43857198758.4016	39154186092.5516	4823002859.8713
sse	5	8965018845.2774	11096336332.7613	9899546225.2180	821555419.0107
sde	5	5872.0390	6545.4562	6170.1628	260.9758
mse0	5	113067506.2910	163646264.0239	146298959.6759	17905426.2621
rmse	5	5872.0390	6545.4562	6170.1628	260.9758
mae	5	4054.1099	4427.9335	4203.4121	129.0554
smape	5	35.6194	40.0220	38.1279	1.5723



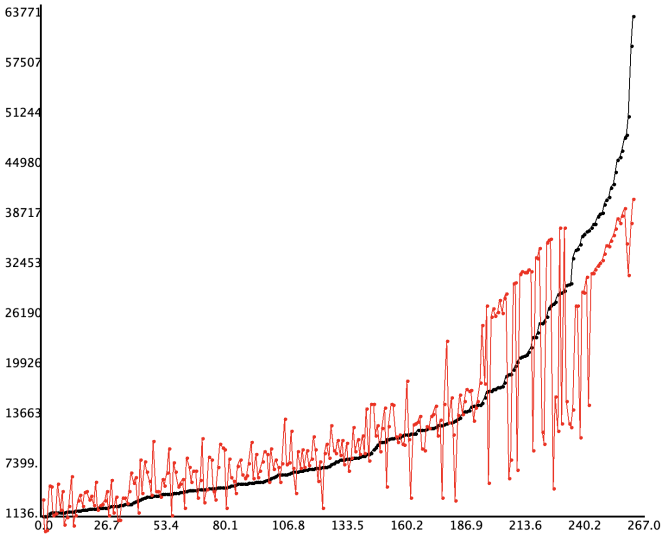
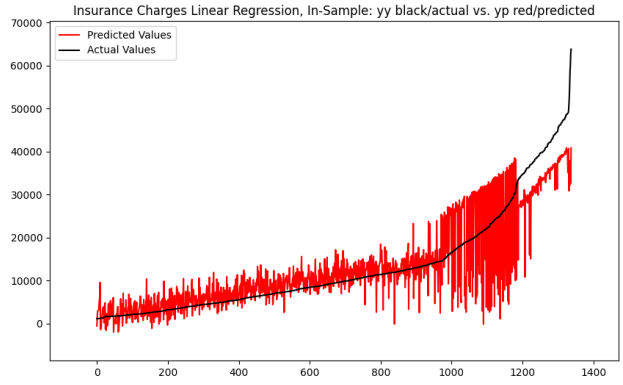


Figure 21: Scalation - Insurance Charges Regression
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted



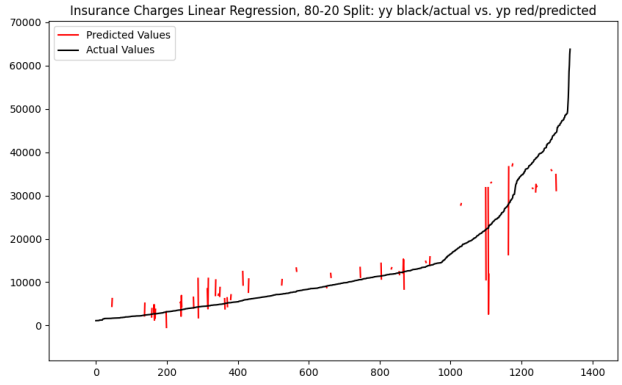


Figure 22: Statsmodels - Insurance Charges Regression
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

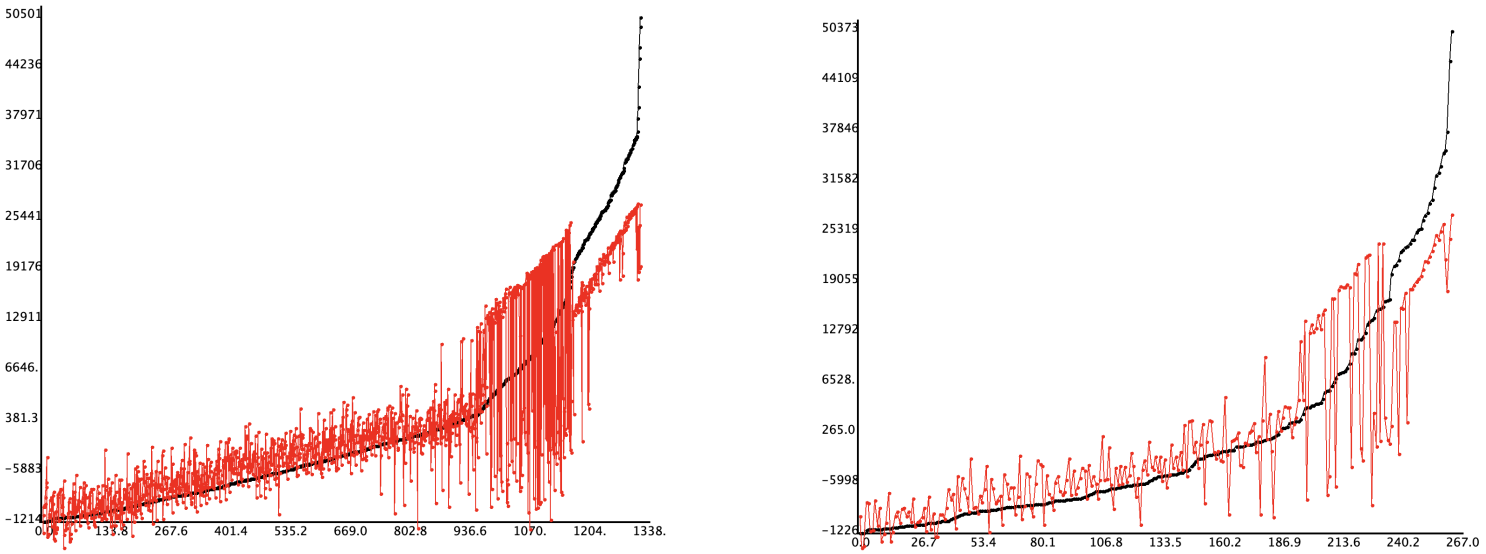


Figure 23: Scalation - Insurance Charges Ridge
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

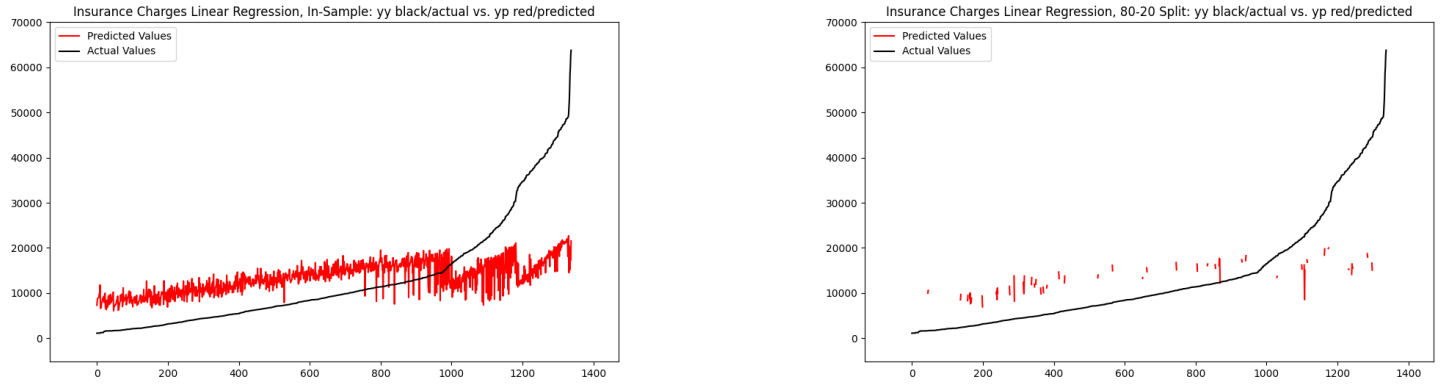


Figure 24: Statsmodels - Insurance Charges Ridge
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

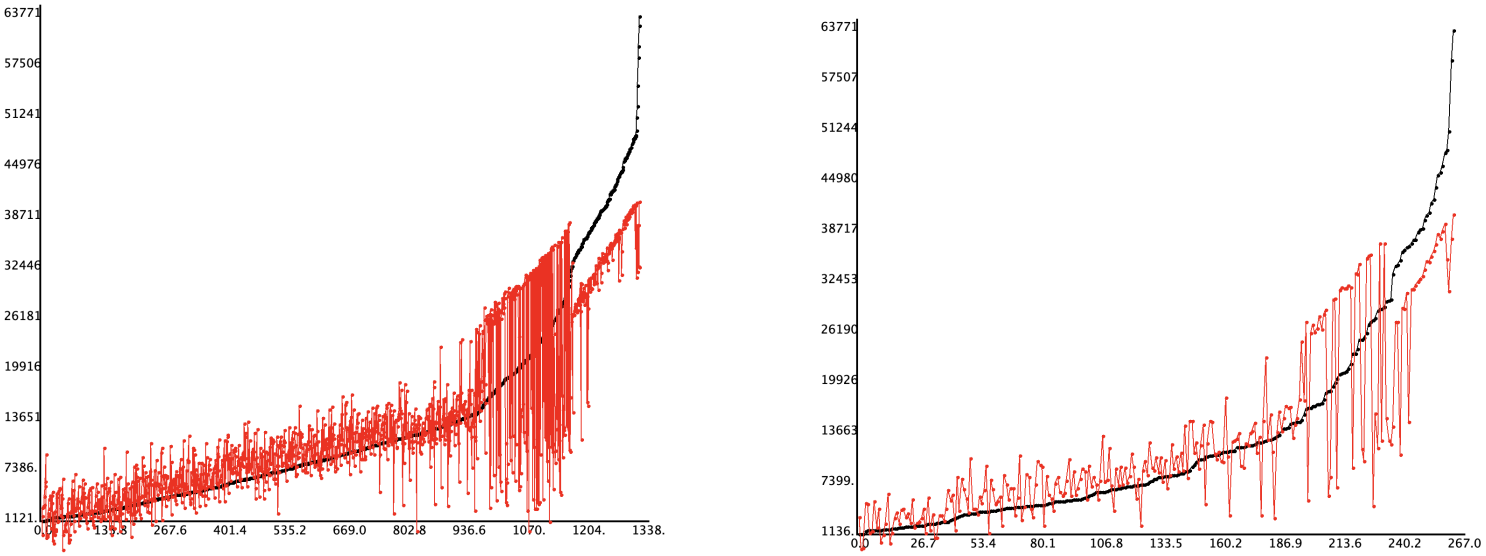


Figure 25: Scalation - Insurance Charges Lasso
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

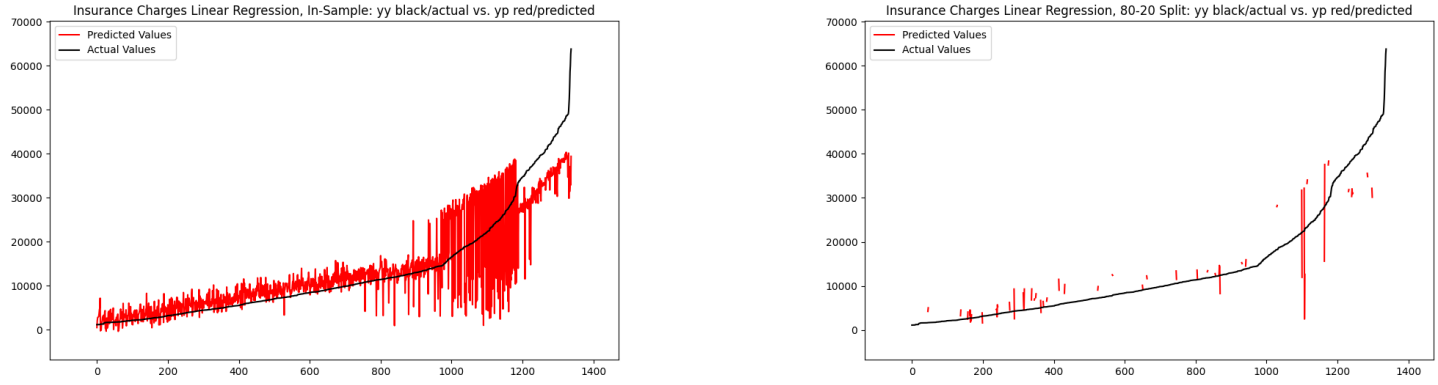


Figure 26: Statsmodels - Insurance Charges Lasso
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

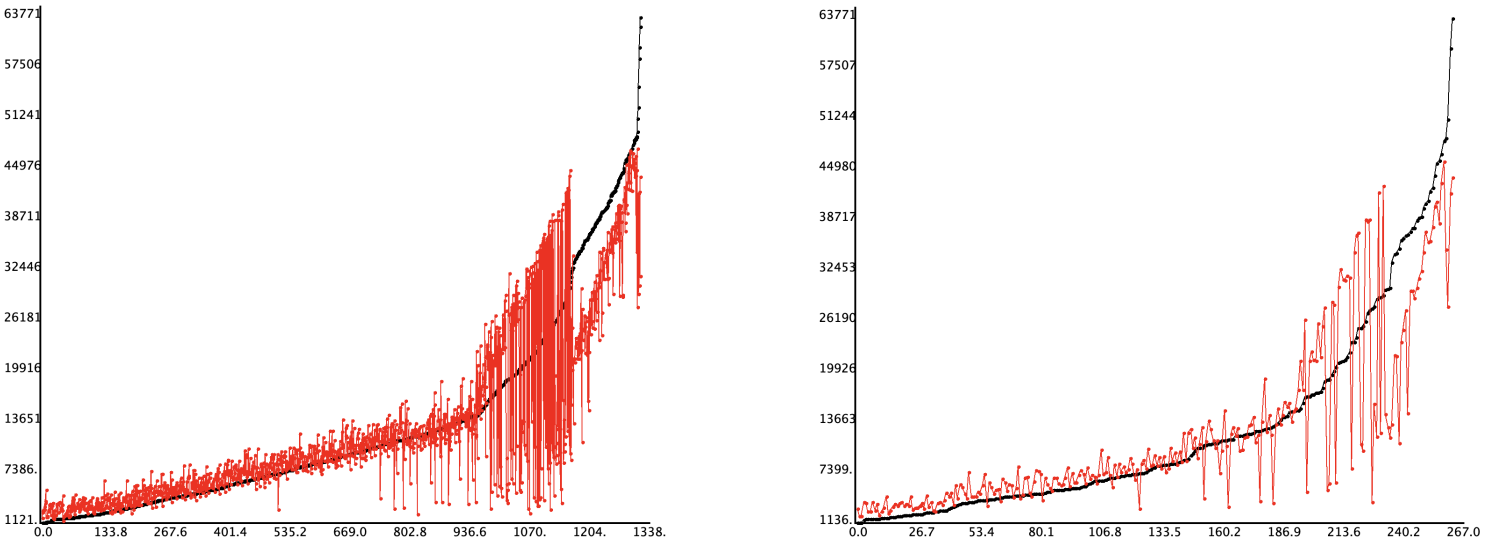


Figure 27: Scalation - Insurance Charges Sqrt
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

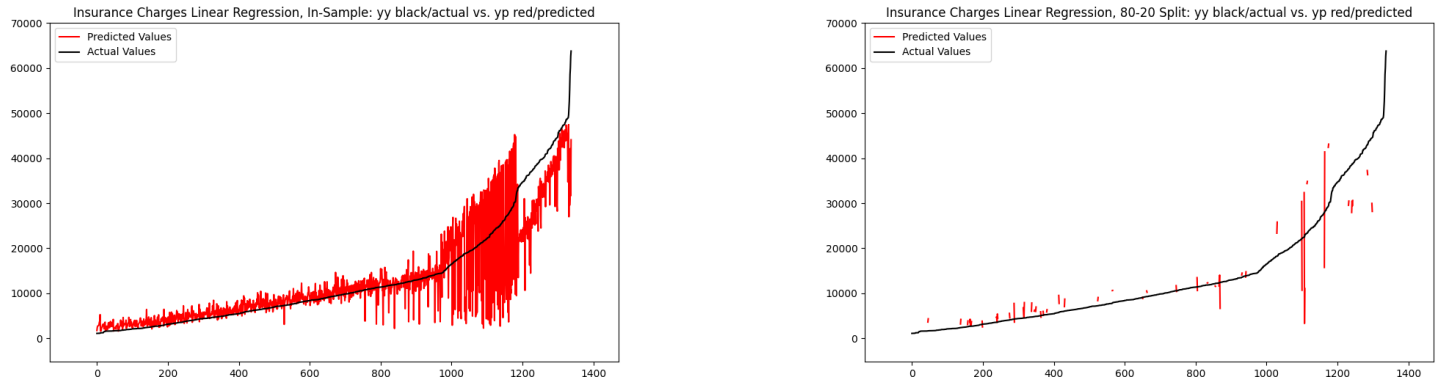


Figure 28: Statsmodels - Insurance Charges Sqrt
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

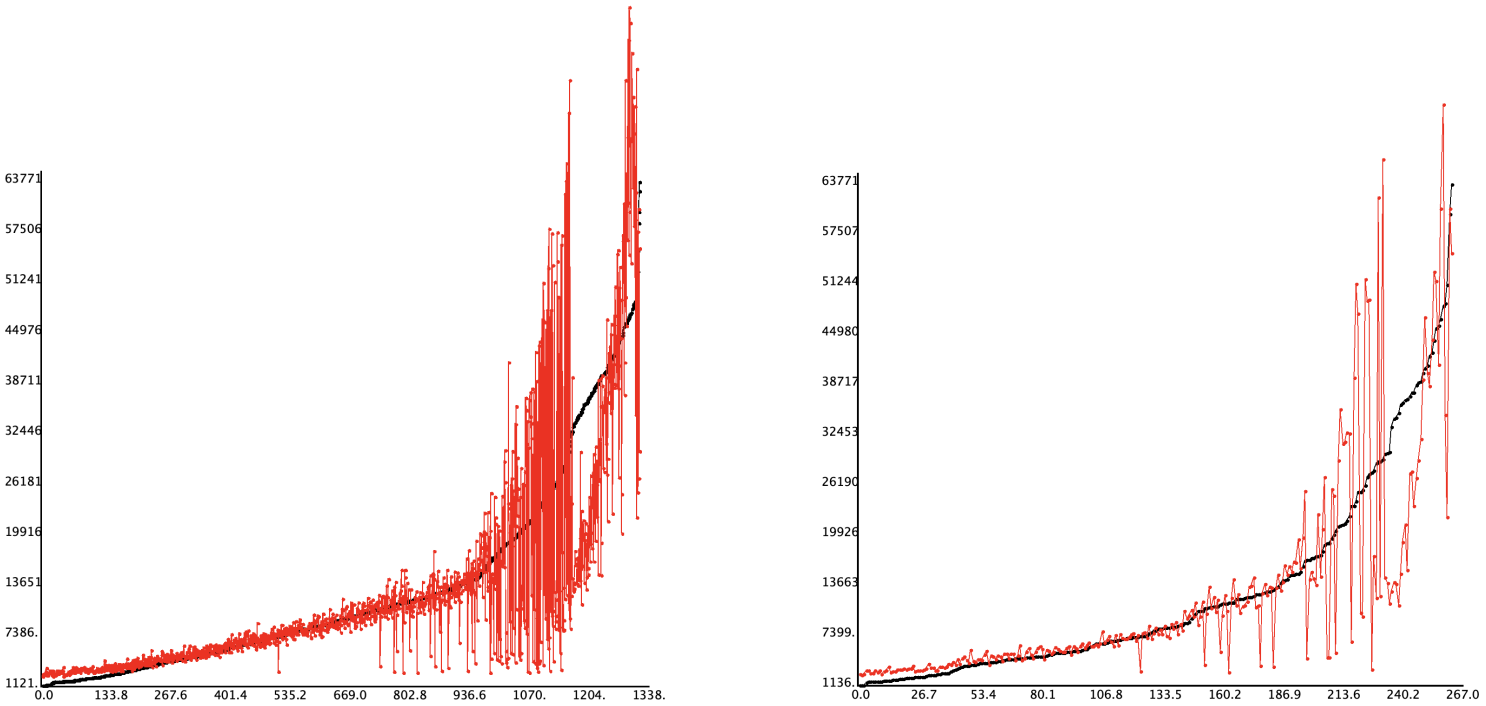


Figure 29: Scalation - Insurance Charges log1p
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted

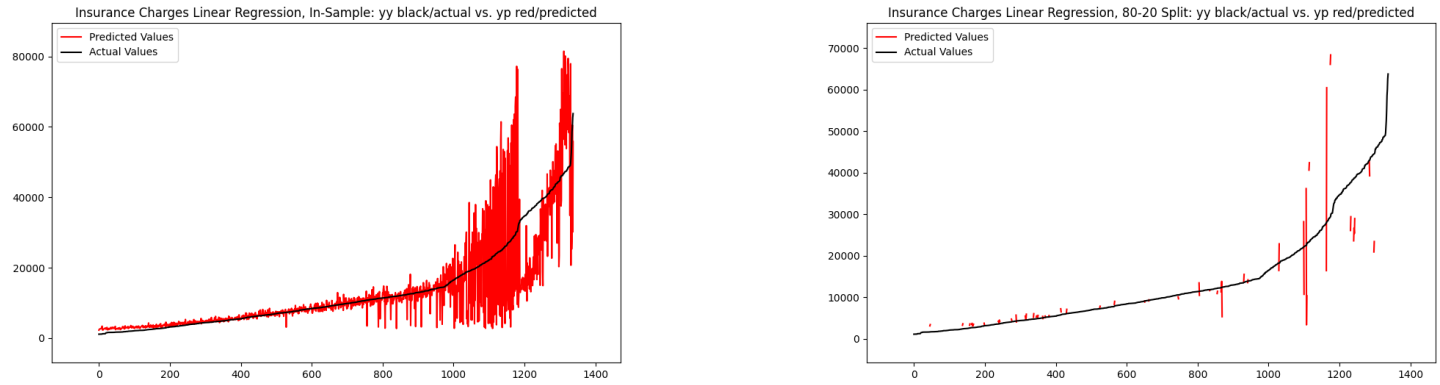


Figure 30: Statsmodels - Insurance Charges log1p
Left: In Sample Predictions
Right: 80-20 Out of Sample Predictions
yy black/actual vs. yp red/predicted