

**HYPERSENSITIVE IMAGE SEGMENTATION AND  
ANALYSIS TOWARDS AUTOMATED NON-  
DESTRUCTIVE MEASUREMENTS OF  
OLIGOSACCHARIDES AND OTHER TRAITS**

**HYSACS  
PROJECT**

WEDNESDAY, JUNE 3, 2020



```
graph TD; A[Segmentation] --> B[Dataset]; B --> C[Results];
```

Segmentation

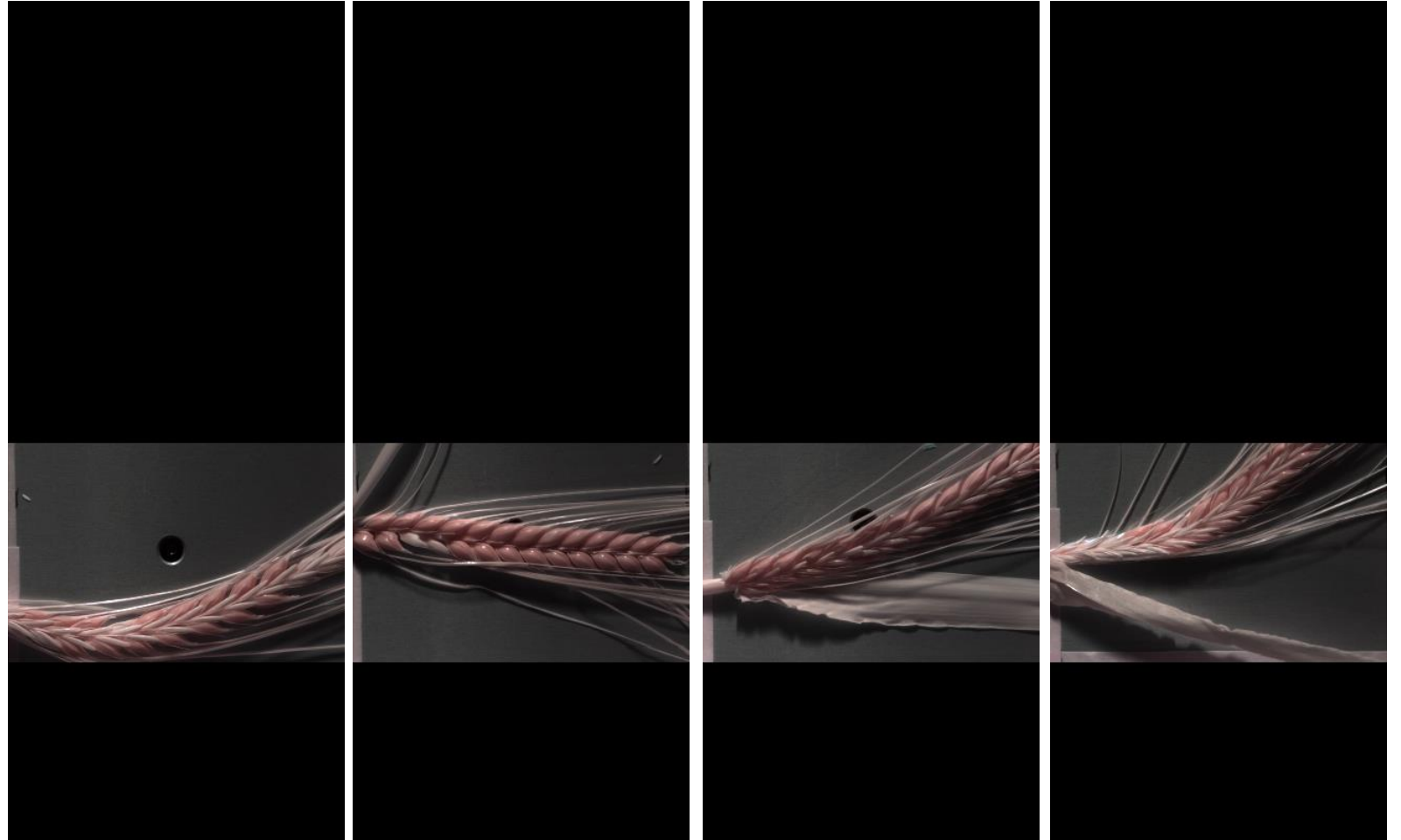
Dataset

Results

CONTENTS

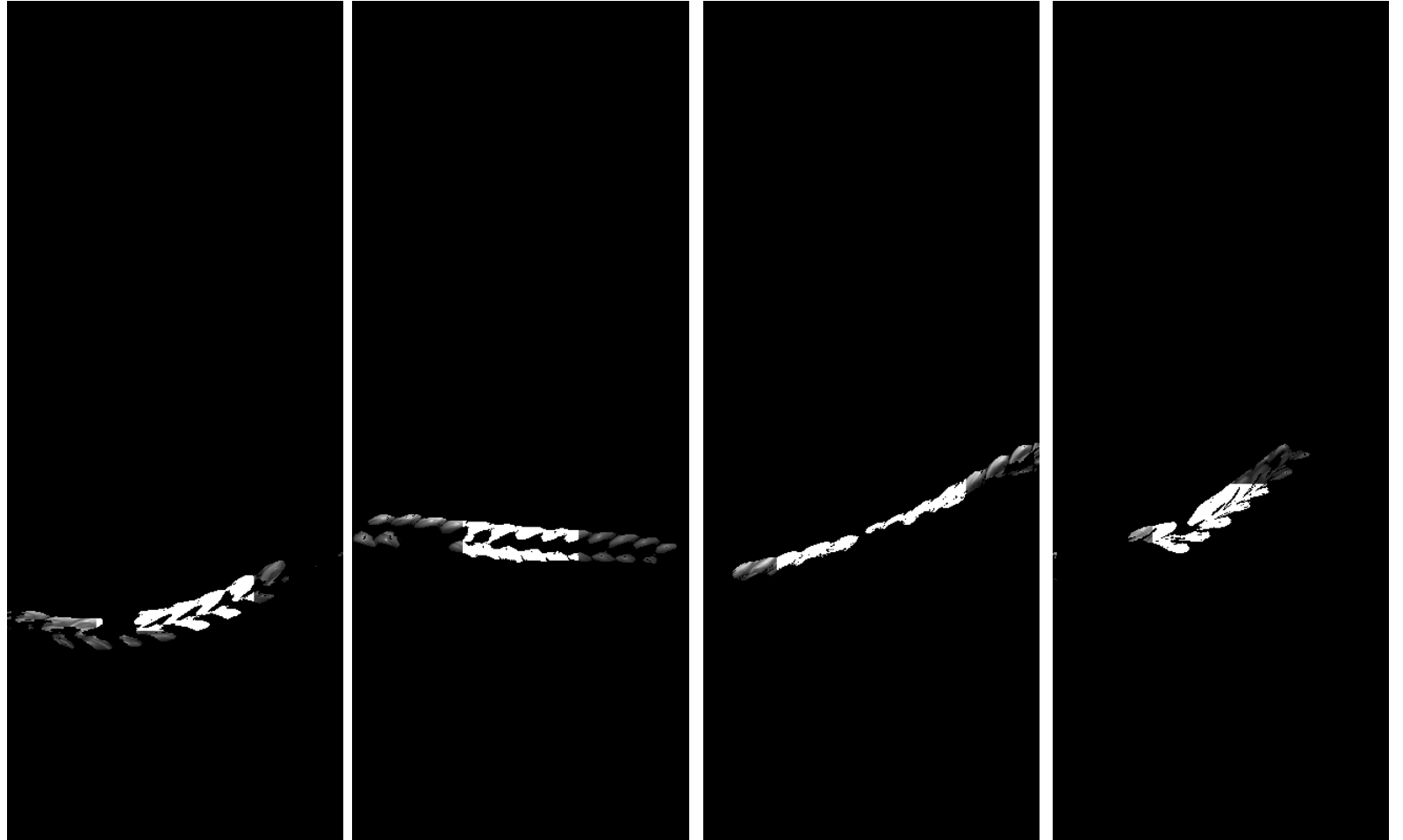
# SEGMENTATION

- 97% of images are properly segmented
- Five Images were damaged
- 14 images improperly segmented



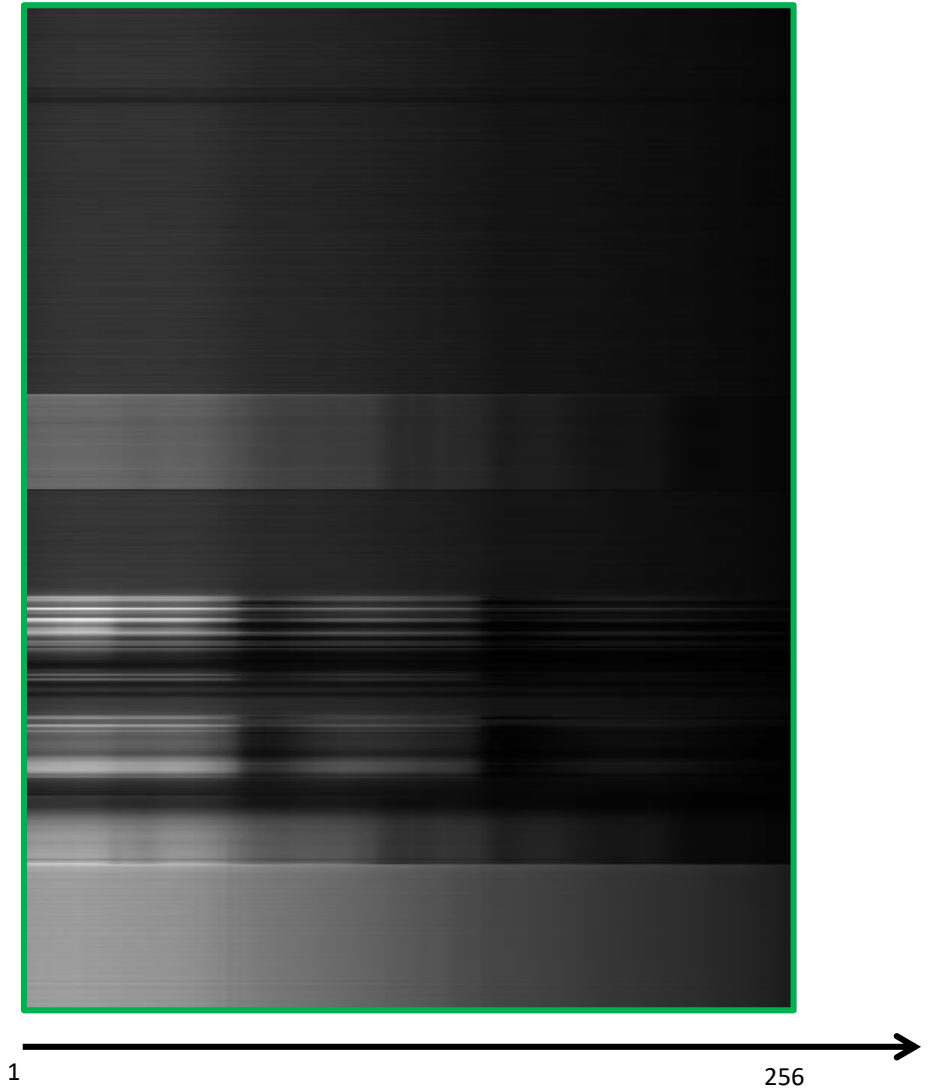
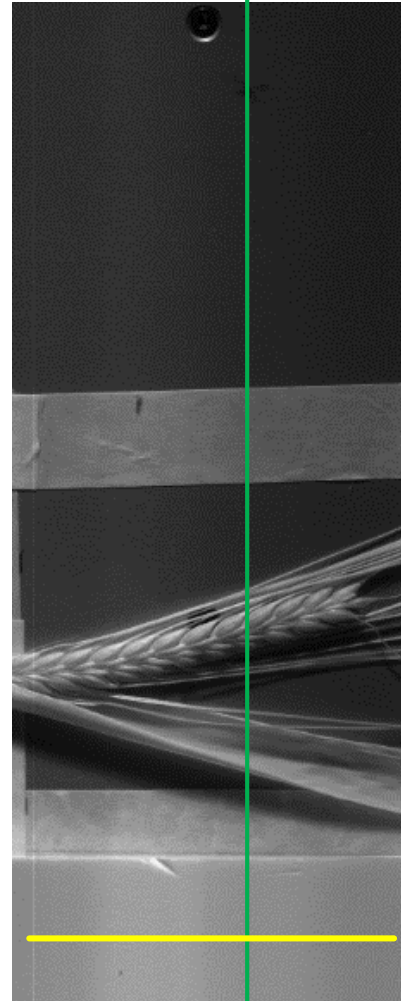
# SEGMENTATION

- 97% of images are properly segmented
- Five Images were damaged
- 14 images improperly segmented



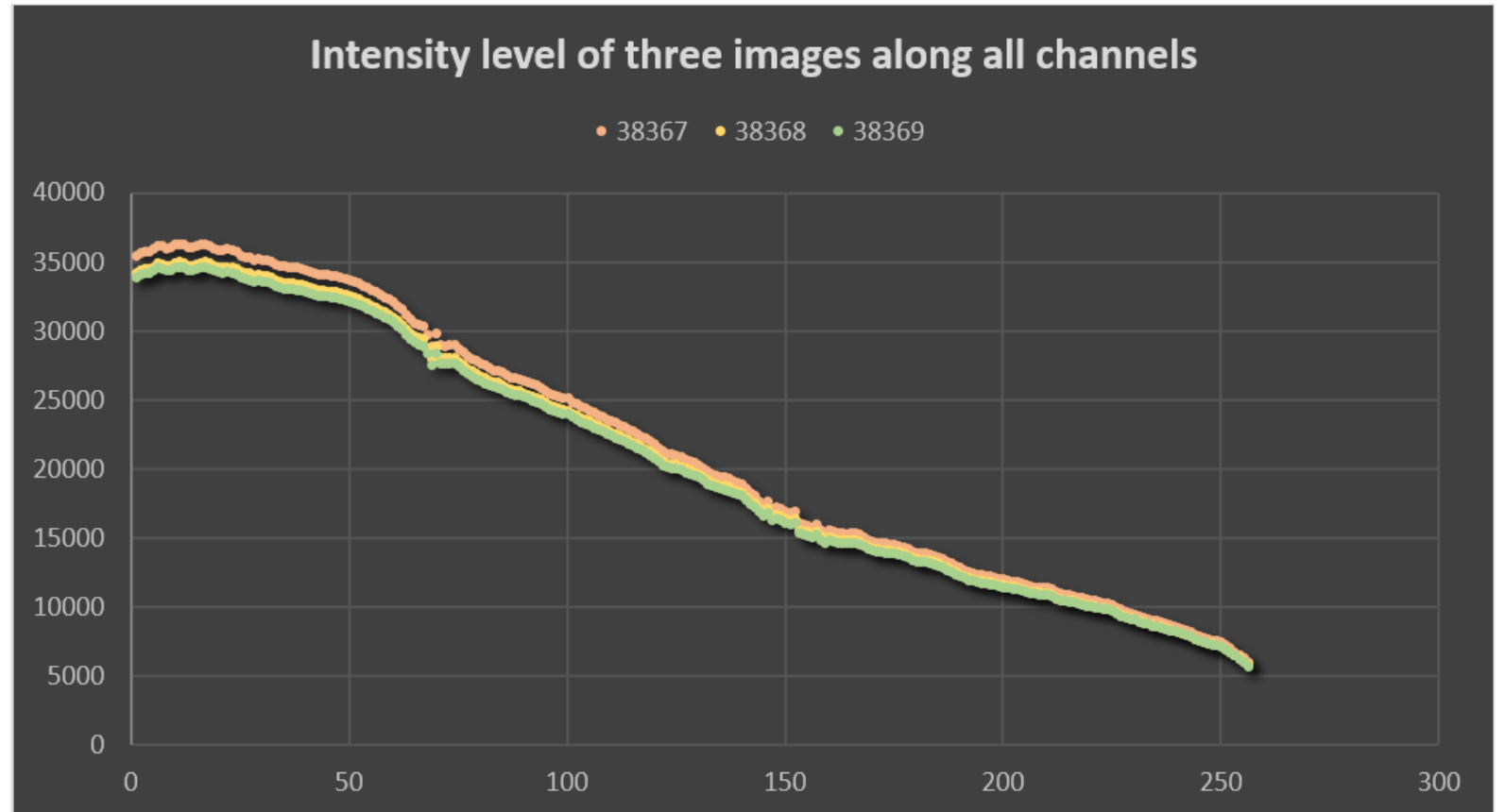
# DAYLIGHT INTENSITY

- Mean value of yellow highlighted area is taken as reference for daylight intensity value



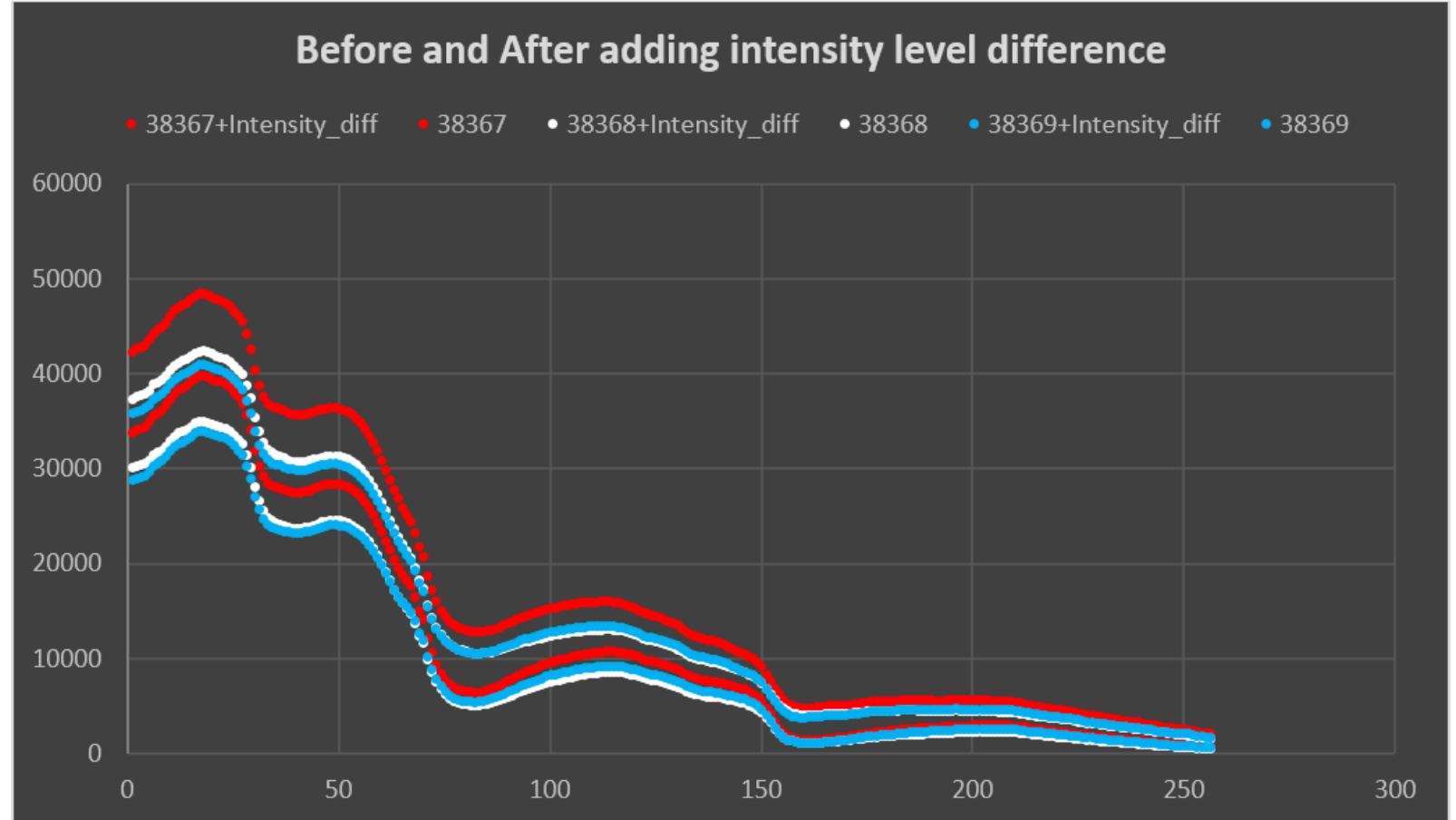
# DAYLIGHT INTENSITY

- Mean value of each channel is taken as reference for daylight intensity value



# DAYLIGHT INTENSITY

- Difference of intensity level of each image added on respective image though all channels



# DATASET

	Img_name	Band_001	Band_002	Band_003	...	Band_253	Band_254	Band_255	Band_256
0	38367	42411.773576	42731.784818	42856.157782	...	2340.470535	2268.815182	2193.448929	2151.578834
1	38368	37435.804808	37658.290769	37786.146346	...	1804.879038	1721.171346	1650.607115	1630.313462
2	38369	35867.342426	36086.528659	36225.673116	...	1931.900533	1871.970927	1819.157594	1791.436371
3	38370	41754.834217	42125.890133	42208.679896	...	2479.146990	2409.336211	2317.887066	2264.441994
4	38371	40995.739496	41192.064481	41291.218932	...	2805.463264	2697.943363	2626.431800	2521.017992
...	...	...	...	...	...	...	...	...	...
455	38866	34893.712285	35122.901273	35238.246404	...	2146.348202	2030.960300	1953.622472	1904.945393
456	38867	36965.963712	37215.449659	37320.237273	...	1968.034394	1907.351250	1815.729053	1779.045417
457	38868	41355.620445	41625.077755	41773.751780	...	1 2472.474916	2387.720544	2280.074461	2238.454322
458	38869	42324.013017	42554.751381	42647.630588	...	2053.766915	1940.010000	1905.009814	1823.892723
459	38870	33099.219507	33318.580099	33475.534631	...	1278.108276	1218.634975	1211.744778	1186.899951





# MODELS

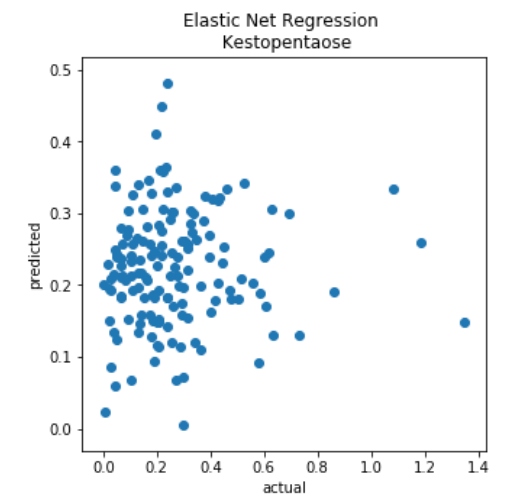
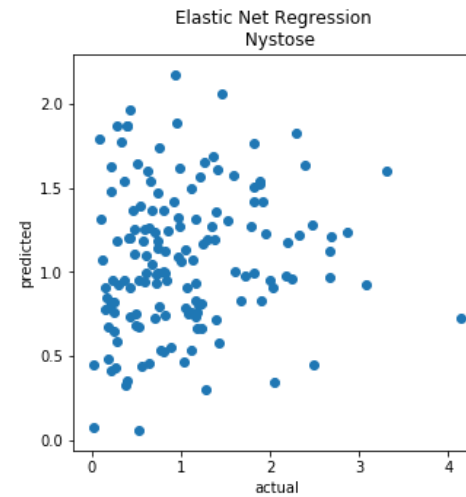
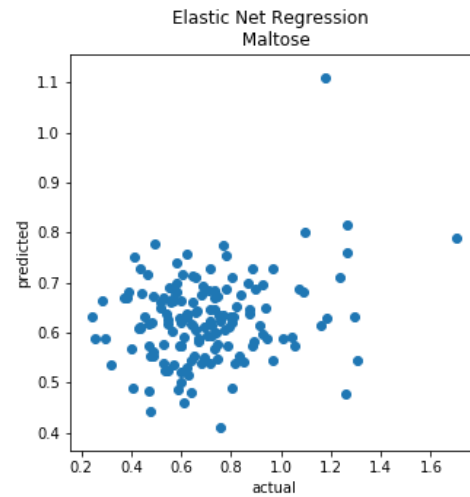
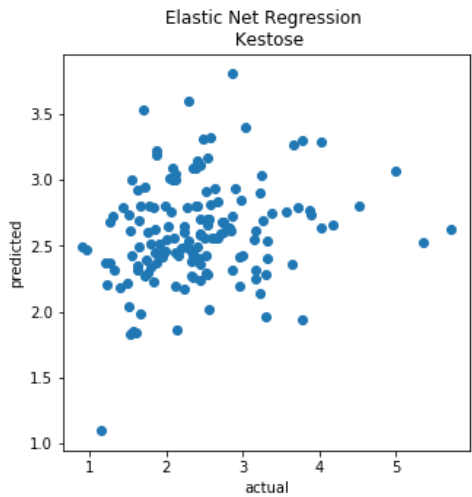
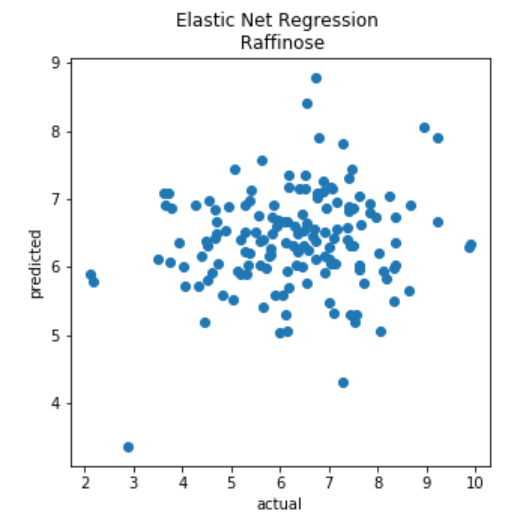
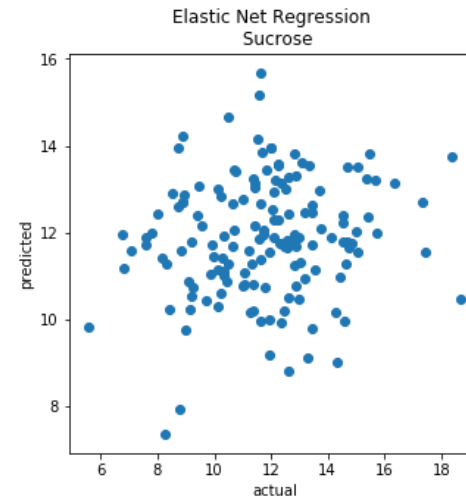
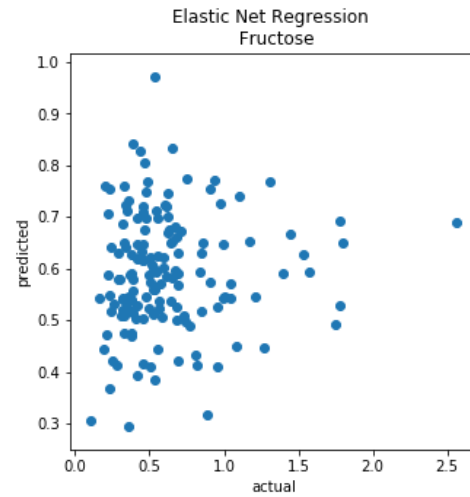
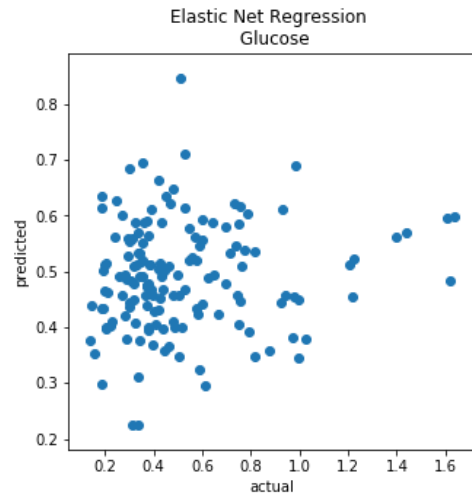
SPLIT THE DATASET BASED ON GENOTYPE



# RIDGE REGRESSION

Sugar Contents	Mean absolute error	R^2 error	Mean square error
Glucose	0.511074901	-3.88061	0.442623544
Fructose	0.626878935	-3.58233	0.658037805
Sucrose	4.516844231	-4.40465	30.37255988
Raffinose	2.893977891	-5.39806	13.55439056
Kestose	1.917278963	-7.69109	5.886601115
Maltose	0.578016671	-8.63727	0.5331861
Nystose	1.677731598	-7.77544	4.987306438
Kestopentaose	0.532623374	-8.84136	0.455078583
Total_Fructan	8.090380188	-6.3398	111.5696898

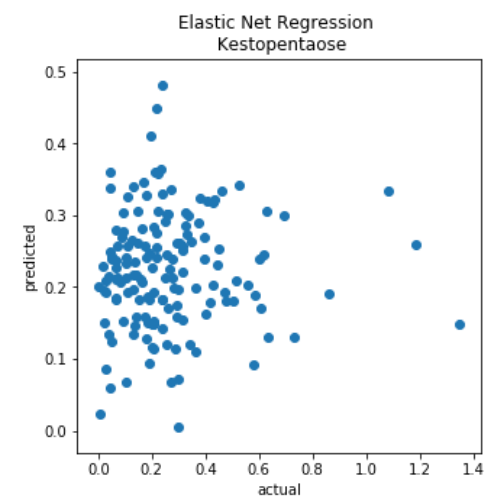
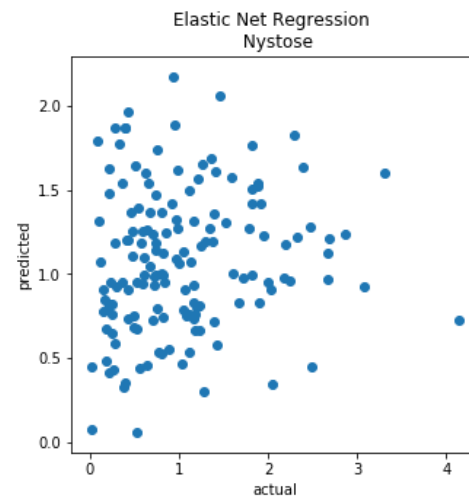
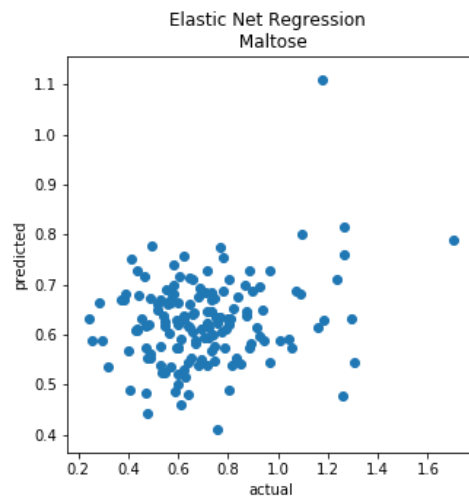
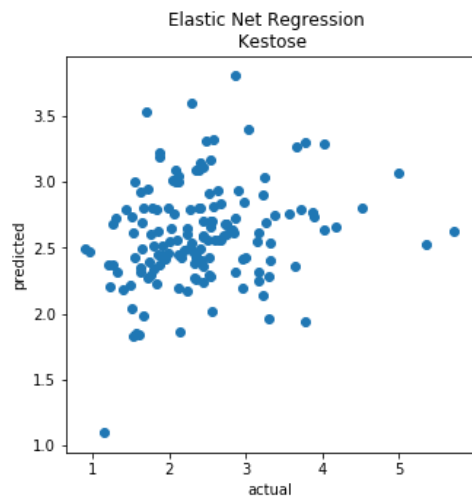
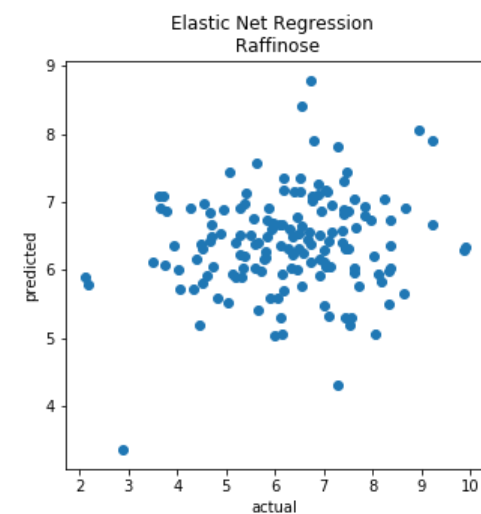
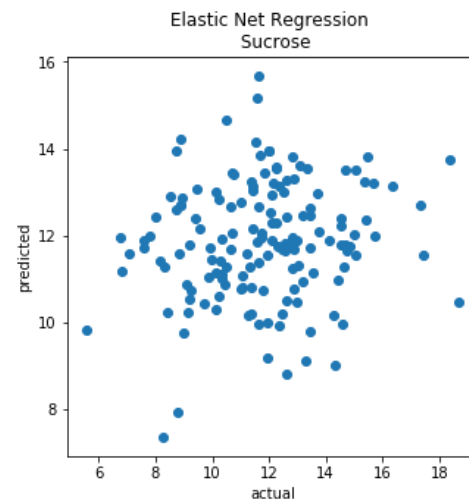
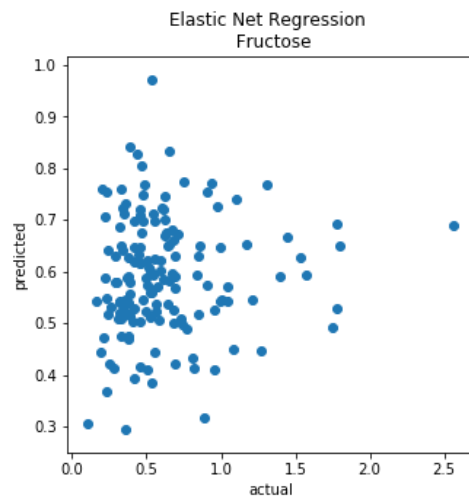
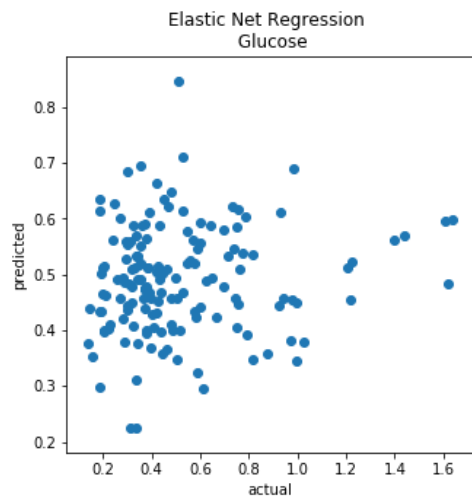
# RIDGE REGRESSION



# ELASTIC NET

Sugar Contents	Mean absolute error	R <sup>2</sup> error	Mean square error
Glucose	0.225395	-0.03672	0.09402
Fructose	0.272316	-0.04594	0.150201
Sucrose	2.072446	-0.17048	6.577736
Raffinose	1.223988	-0.10549	2.342003
Kestose	0.667266	-0.06532	0.721554
Maltose	0.179537	-0.05939	0.058611
Nystose	0.634522	-0.17936	0.670258
Kestopentaose	0.160034	-0.14356	0.05288
Total_Fructan	3.773329	-0.58032	24.02191

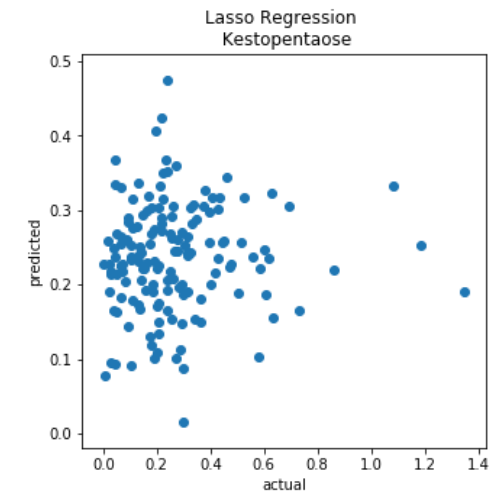
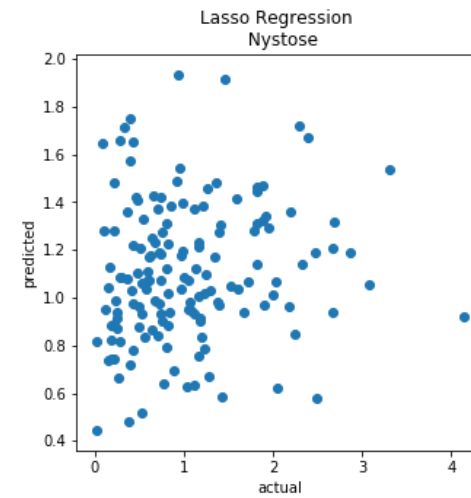
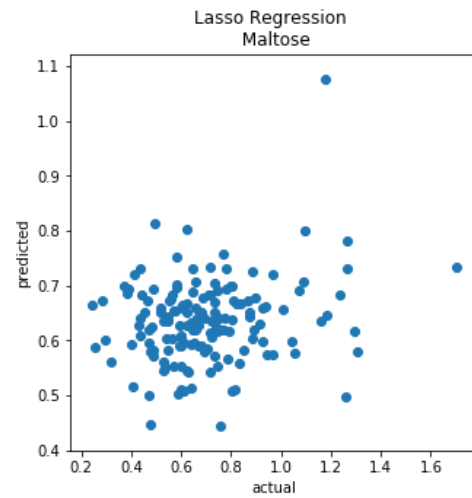
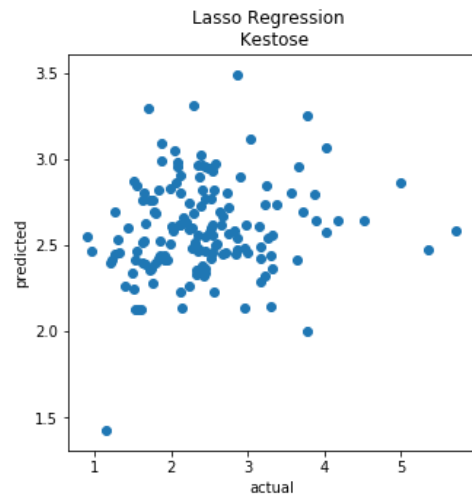
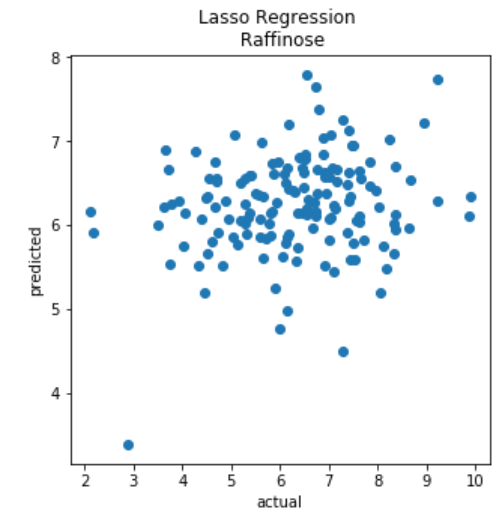
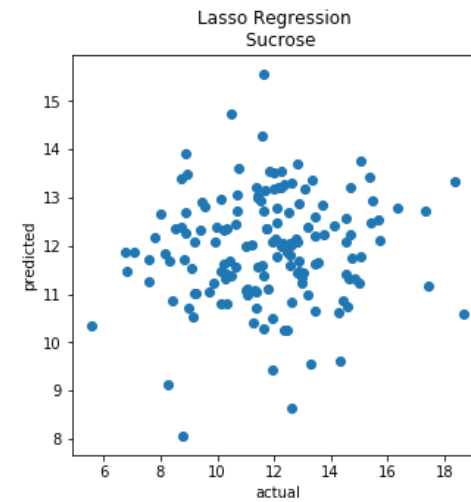
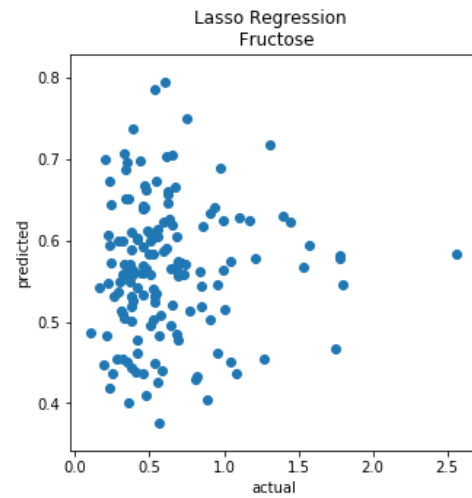
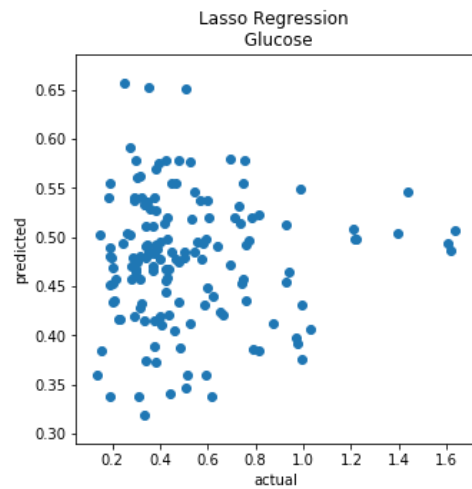
# ELASTIC NET



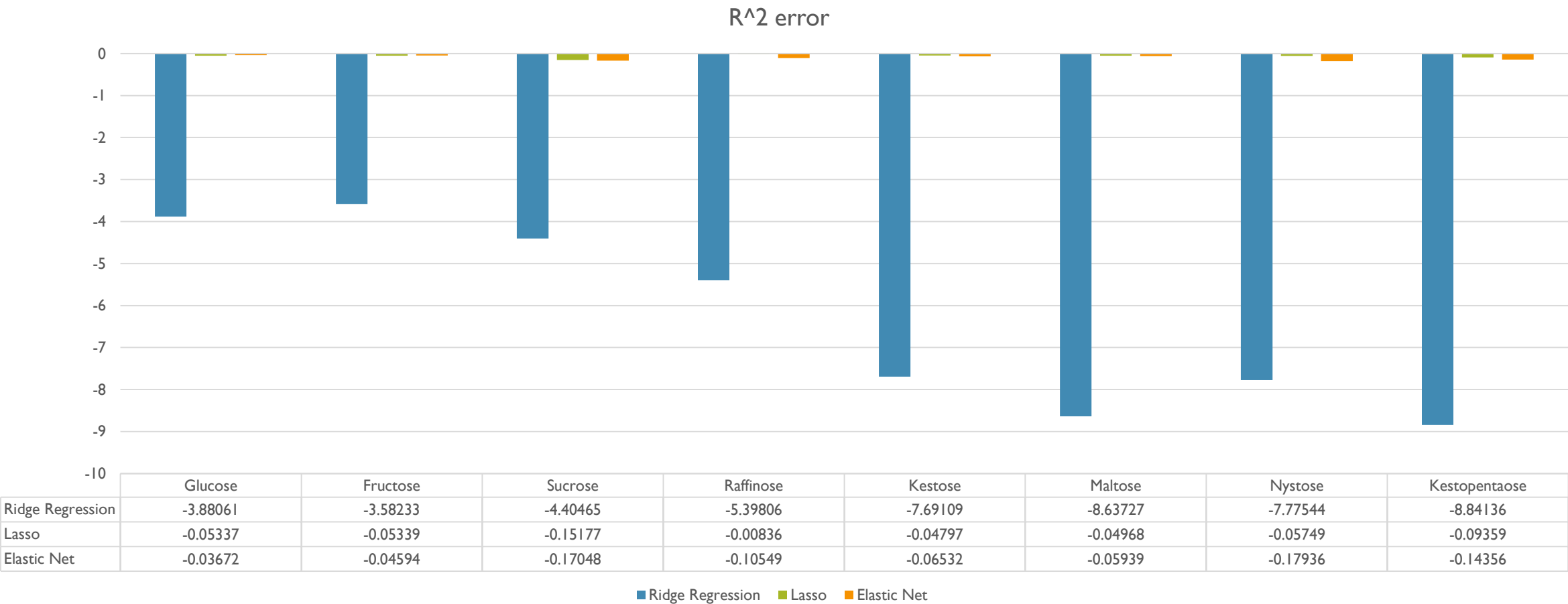
# LASSO

Sugar Contents	Mean absolute error	R <sup>2</sup> error	Mean square error
Glucose	0.22319	-0.05337	0.09553
Fructose	0.264749	-0.05339	0.15127
Sucrose	2.049952	-0.15177	6.472586
Raffinose	1.161771	-0.00836	2.136232
Kestose	0.662869	-0.04797	0.709804
Maltose	0.17603	-0.04968	0.058074
Nystose	0.612399	-0.05749	0.600999
Kestopentaose	0.158774	-0.09359	0.050569
Total_Fructan	3.405795	-0.30661	19.86133

# LASSO

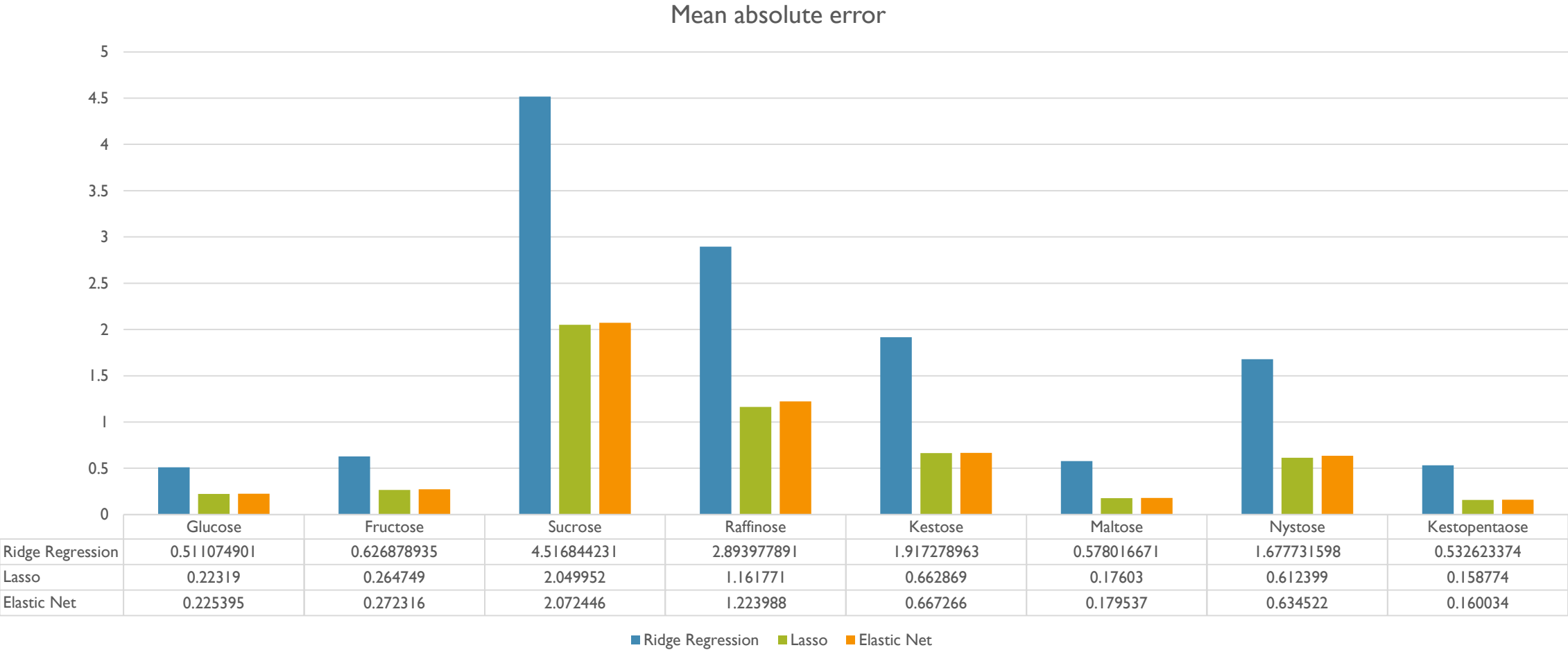


# R^2 ERROR



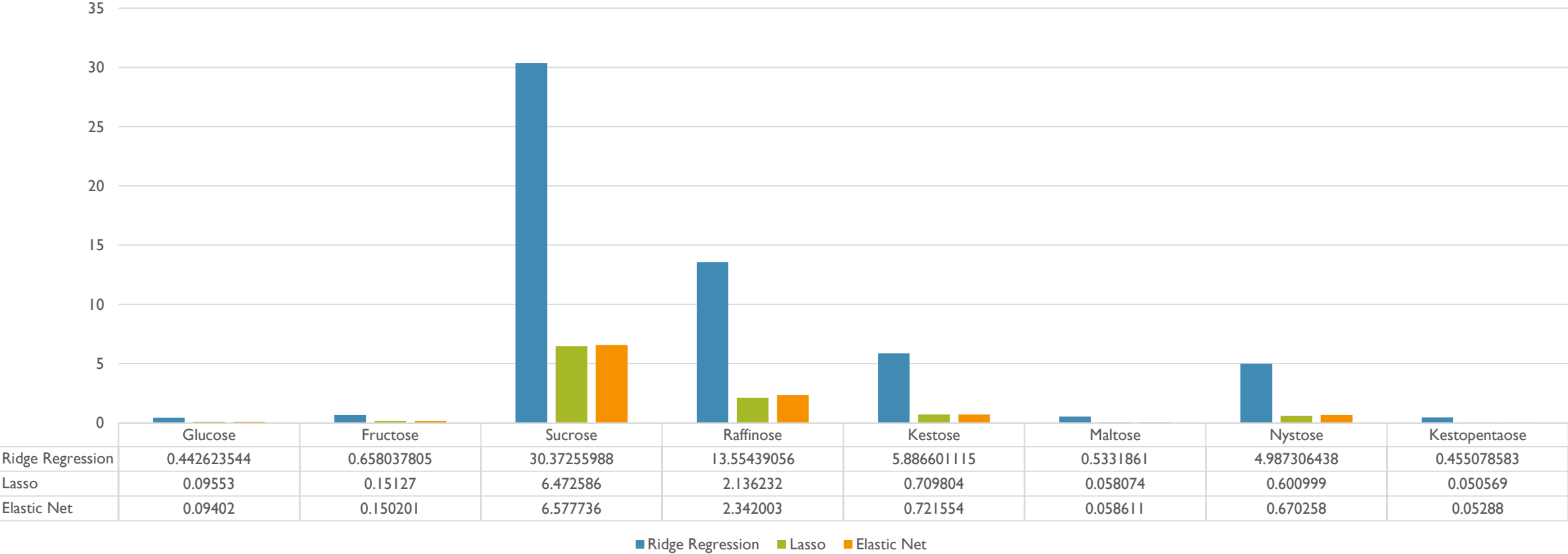


# MEAN ABSOLUTE ERROR



# MEAN SQUARE ERROR

Mean square error





# MODELS

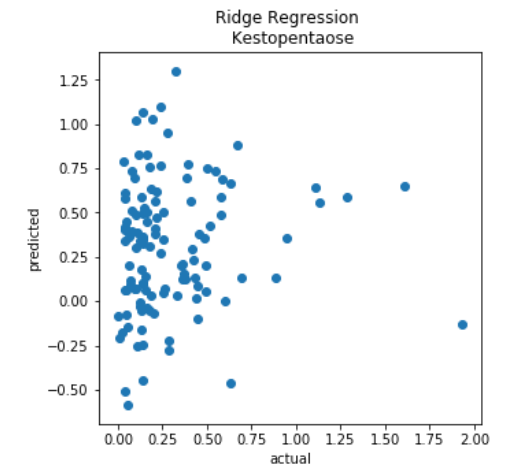
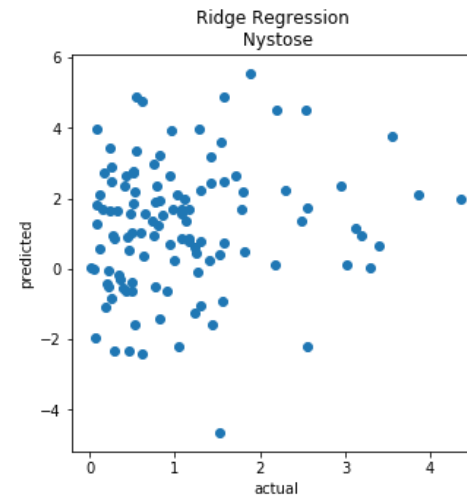
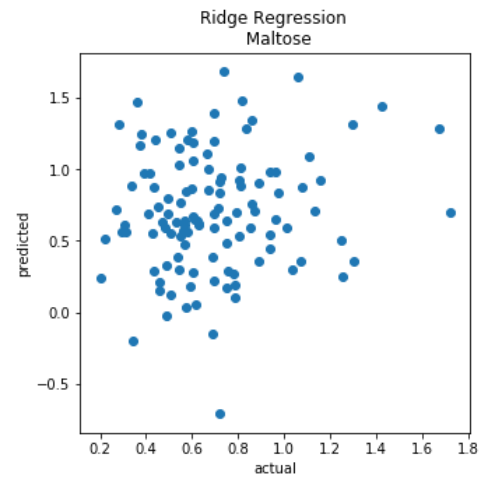
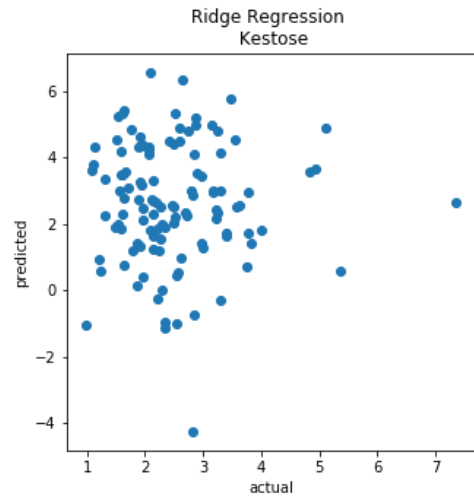
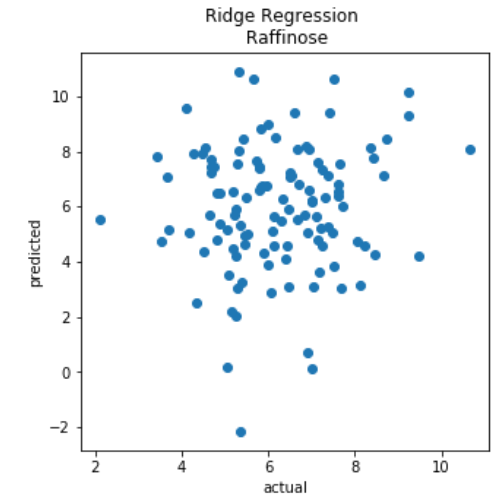
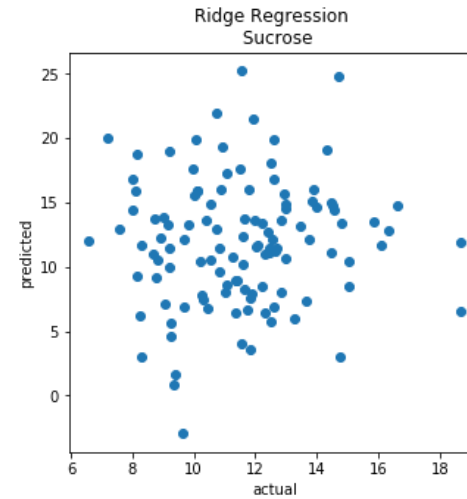
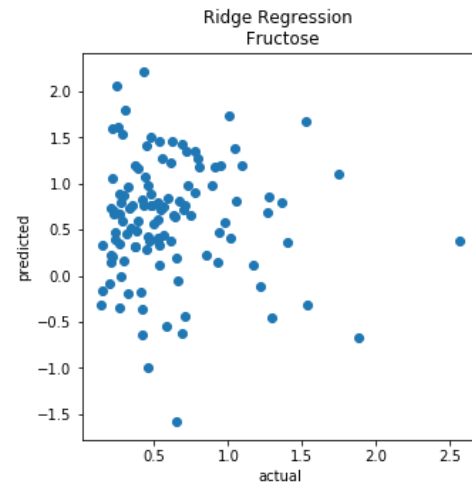
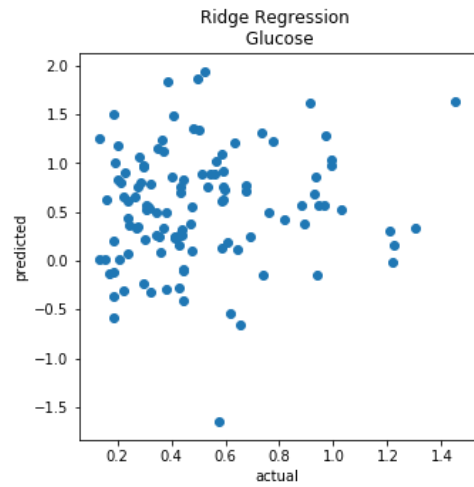
BASED ON RANDOMLY SPLITTING DATASET IN TO TRAIN AND TEST (TAKING MEANS OF 10 REPLICATES)



# RIDGE REGRESSION

Sugar Contents	Mean absolute error	R <sup>2</sup> error	Mean square error
Glucose	0.441288	-3.0415	0.324161
Fructose	0.579718	-3.48335	0.562616
Sucrose	3.550131	-2.64905	20.32791
Raffinose	2.058246	-2.58193	6.888539
Kestose	1.366227	-2.94252	2.942811
Maltose	0.392421	-2.48984	0.25688
Nystose	1.366642	-2.97526	2.933128
Kestopentaose	0.391586	-3.02246	0.249752
Total_Fructan	6.465502	-2.93823	67.35755

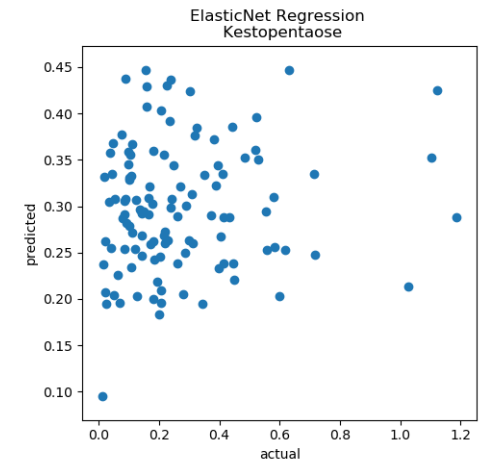
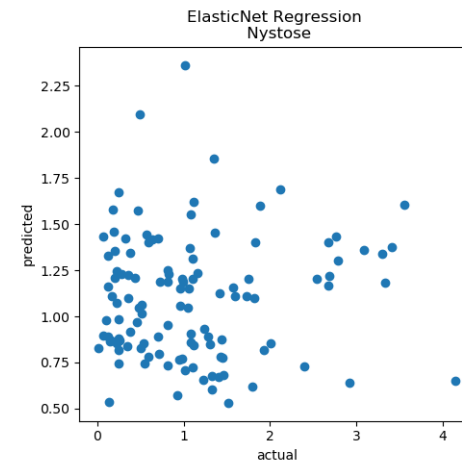
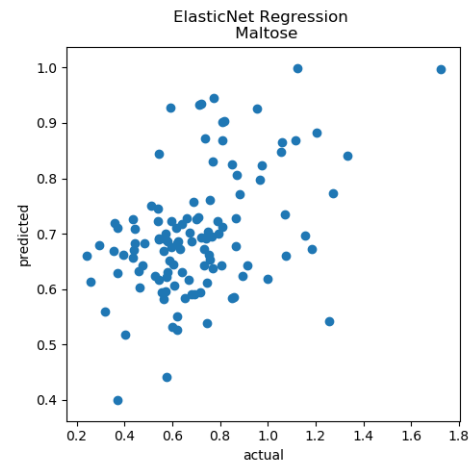
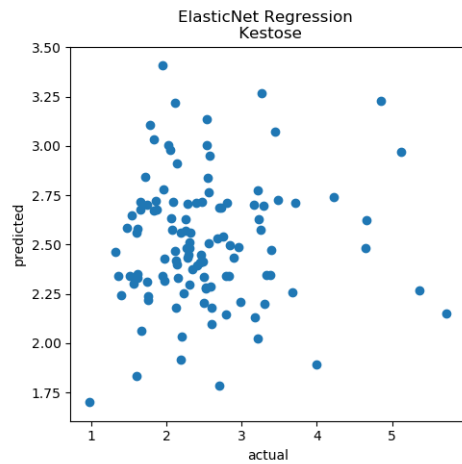
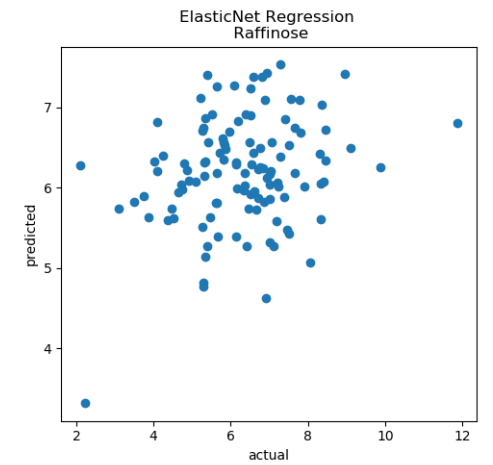
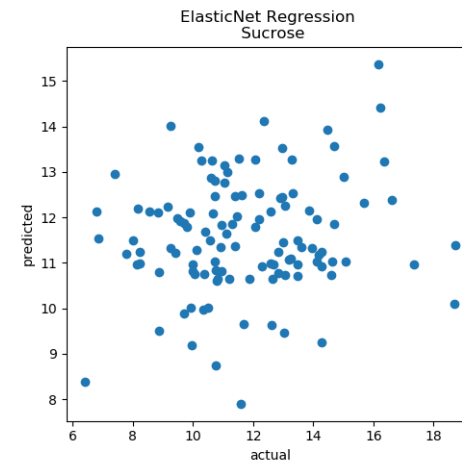
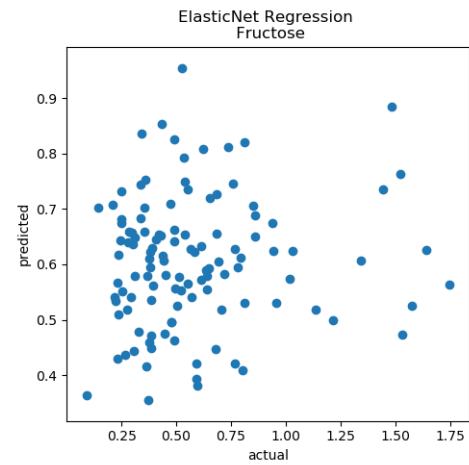
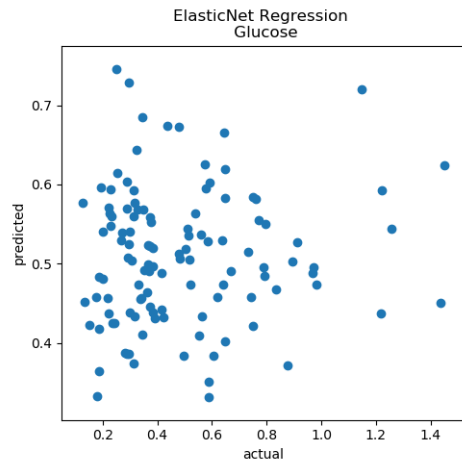
# RIDGE REGRESSION



# ELASTIC NET

Sugar Contents	Mean absolute error	R <sup>2</sup> error	Mean square error
Glucose	0.221439	0.056633	0.083809
Fructose	0.277698	0.04544	0.141775
Sucrose	1.960413	-0.08924	6.061645
Raffinose	1.106896	-0.01355	1.974434
Kestose	0.668826	0.007923	0.783563
Maltose	0.185754	0.132907	0.064505
Nystose	0.706149	0.02043	0.789722
Kestopentaose	0.196809	0.025075	0.073493
Total_Fructan	3.525995	-0.14442	21.76777

# ELASTIC NET

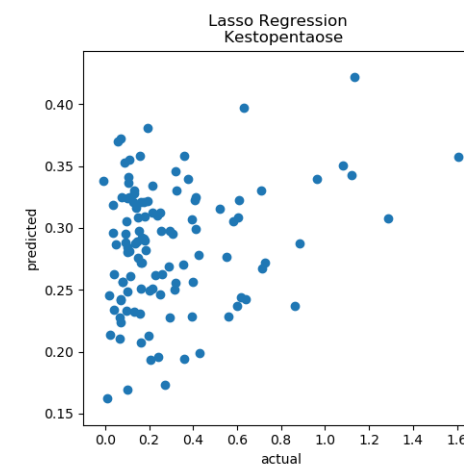
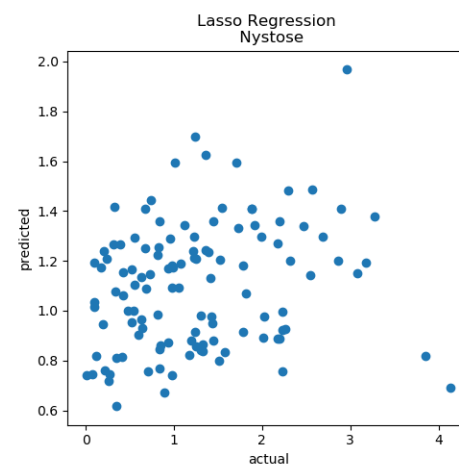
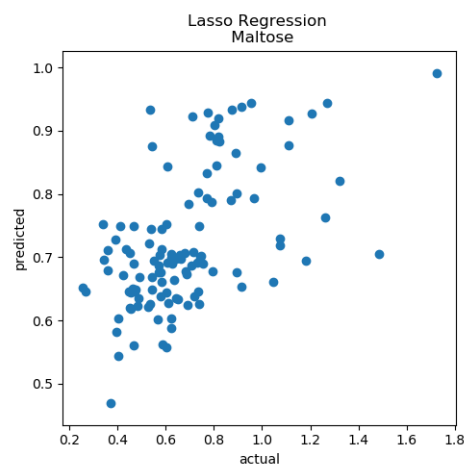
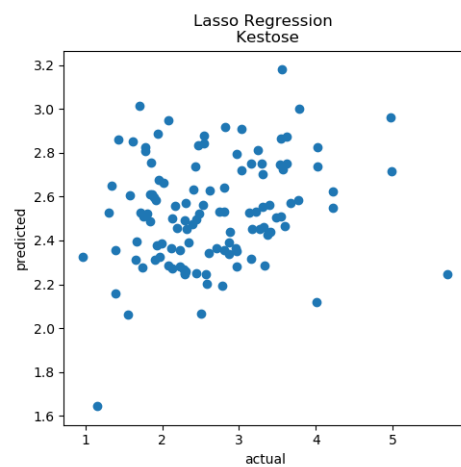
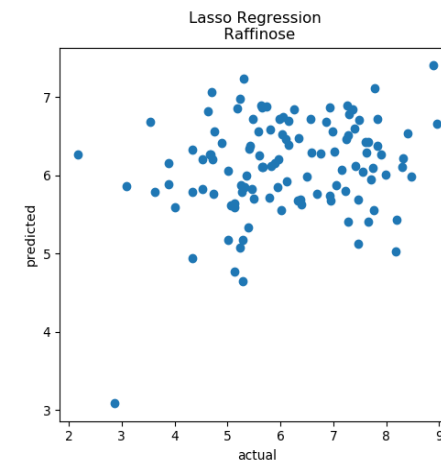
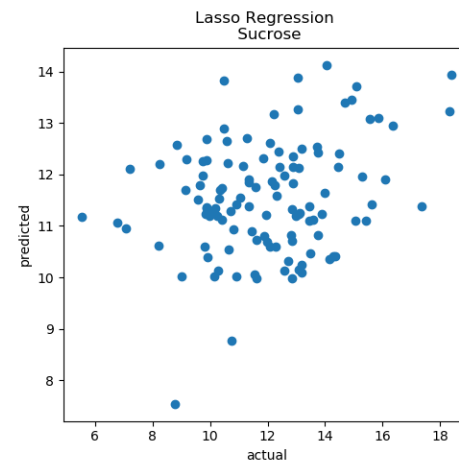
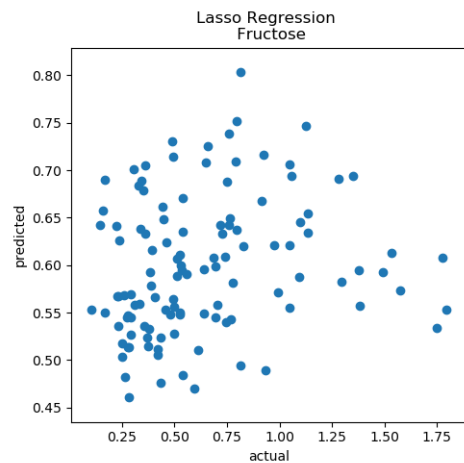
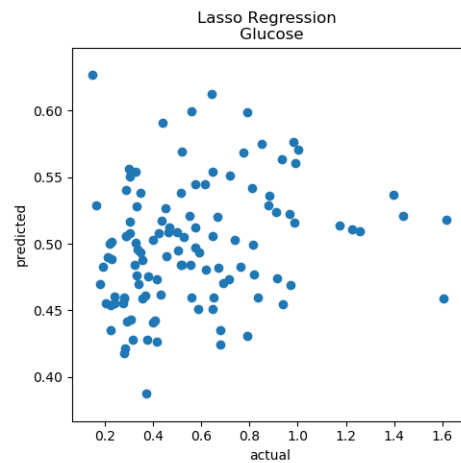


# LASSO

Sugar Contents	Mean absolute error	R <sup>2</sup> error	Mean square error
Glucose	0.220978	-0.01924	0.080412
Fructose	0.269796	-0.01317	0.126867
Sucrose	1.882335	0.014208	5.482845
Raffinose	1.104471	0.026226	1.918576
Kestose	0.652554	0.025089	0.73997
Maltose	0.187364	0.119368	0.070103
Nystose	0.678549	0.017393	0.732928
Kestopentaose	0.183541	-0.00063	0.058472
Total_Fructan	3.137438	-0.01234	17.19574

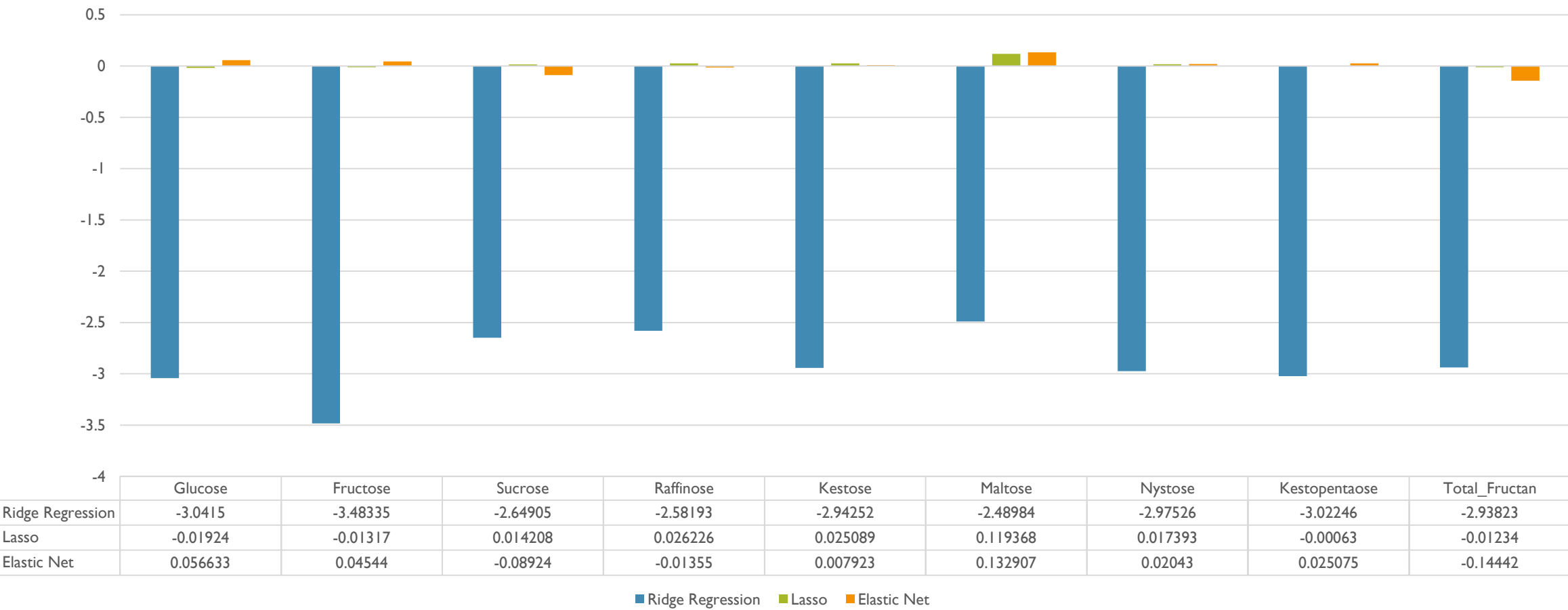


# LASSO



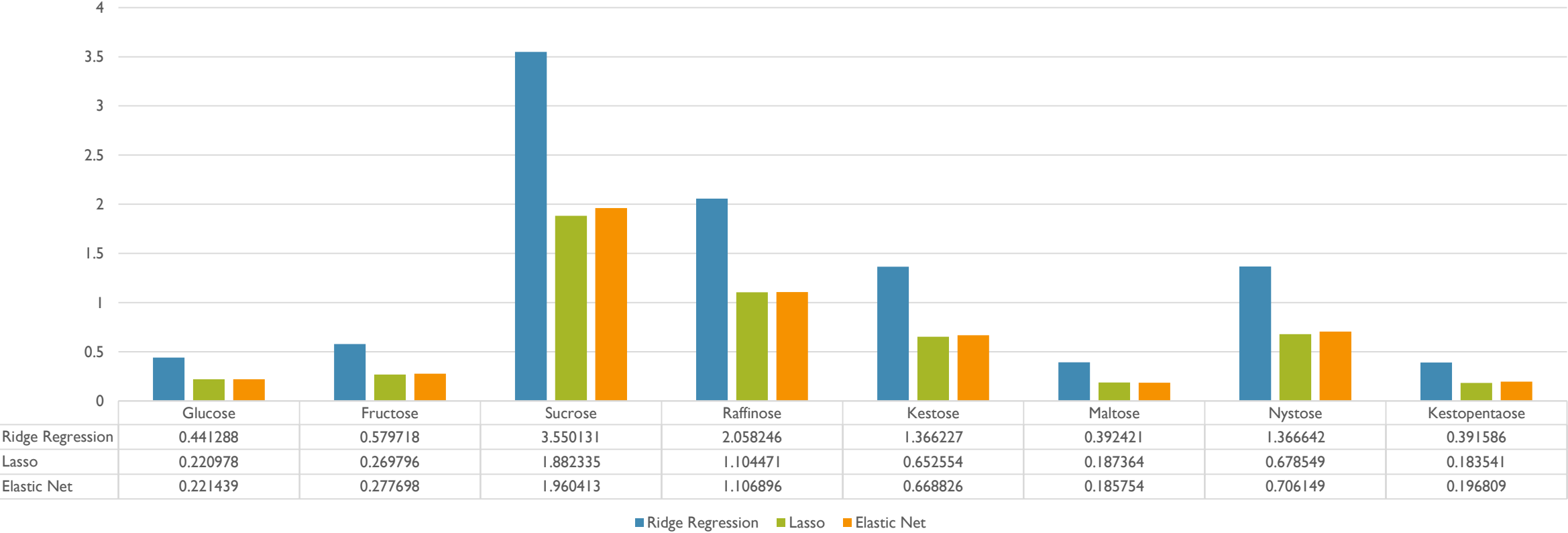
# R^2 ERROR

R^2 error



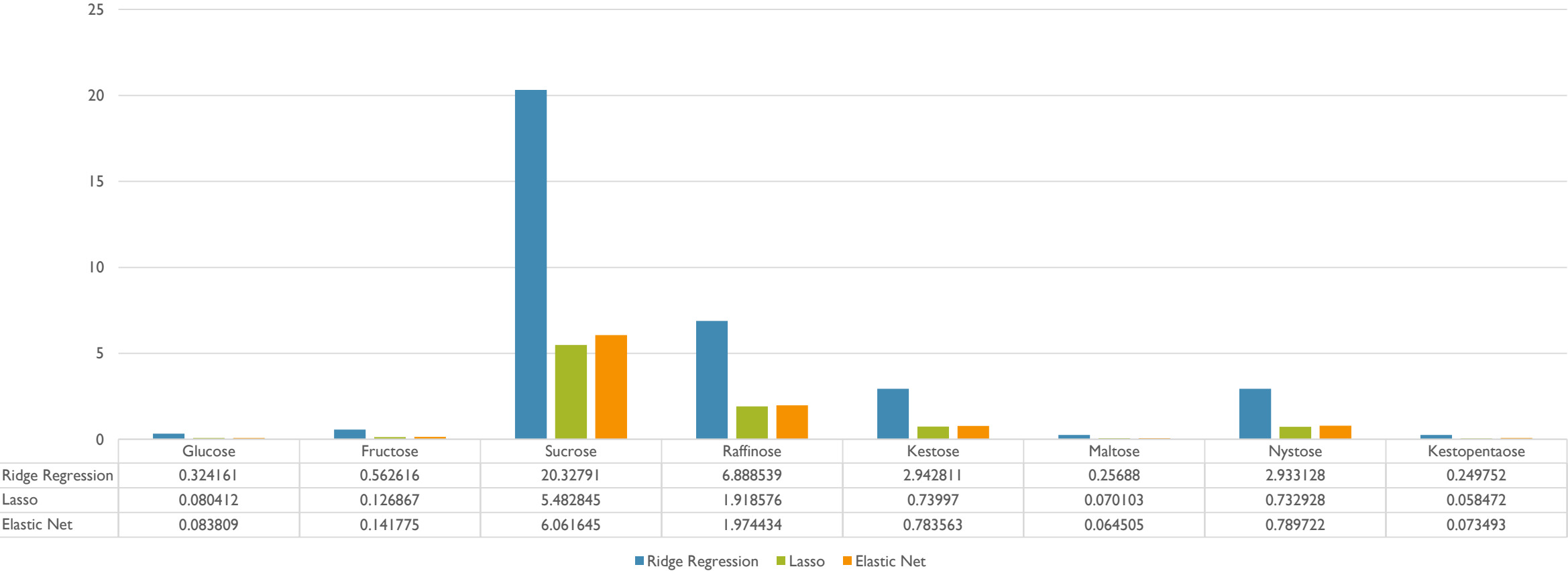
# MEAN ABSOLUTE ERROR

Mean absolute error



# MEAN SQUARE ERROR

Mean square error





# DATASET AGAINST SUGAR VALUES

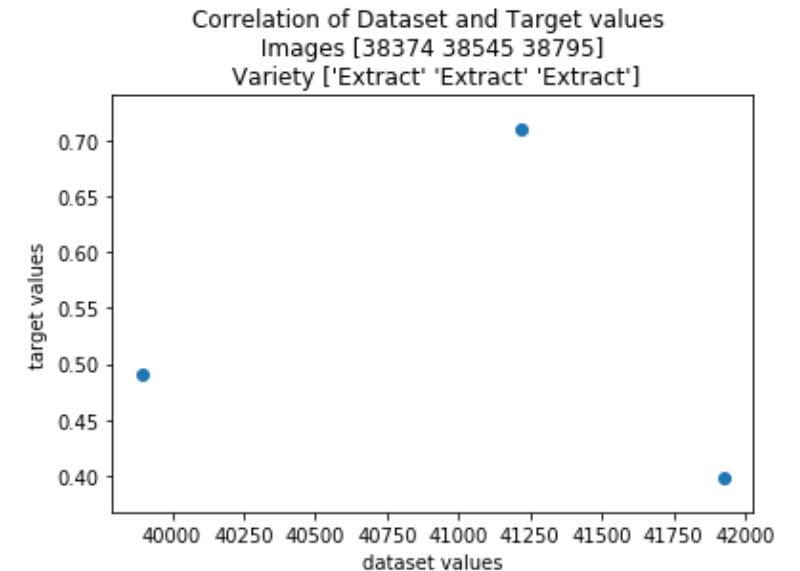
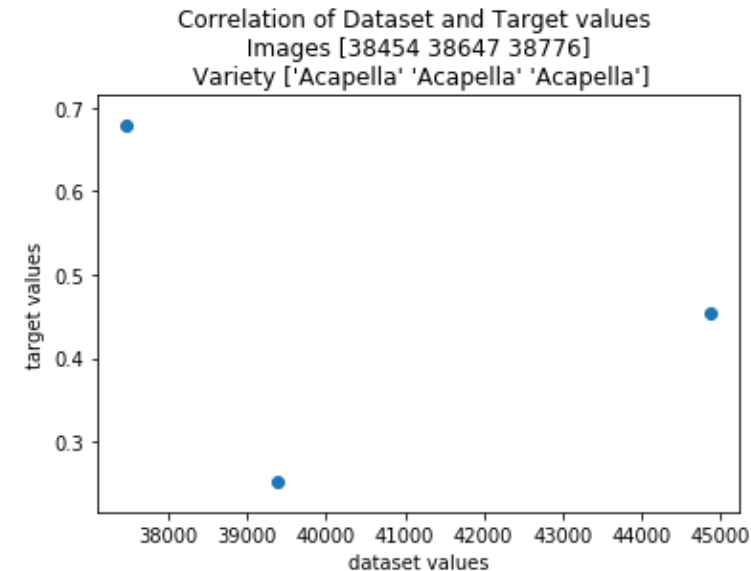
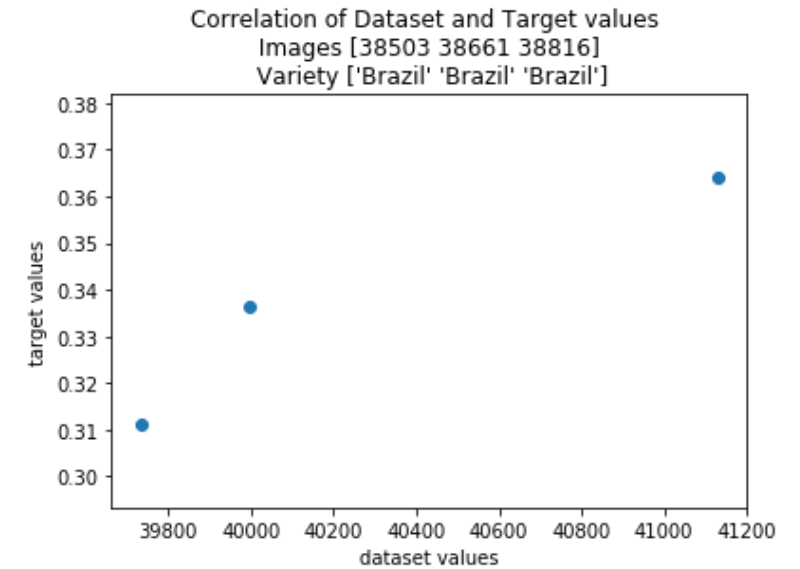
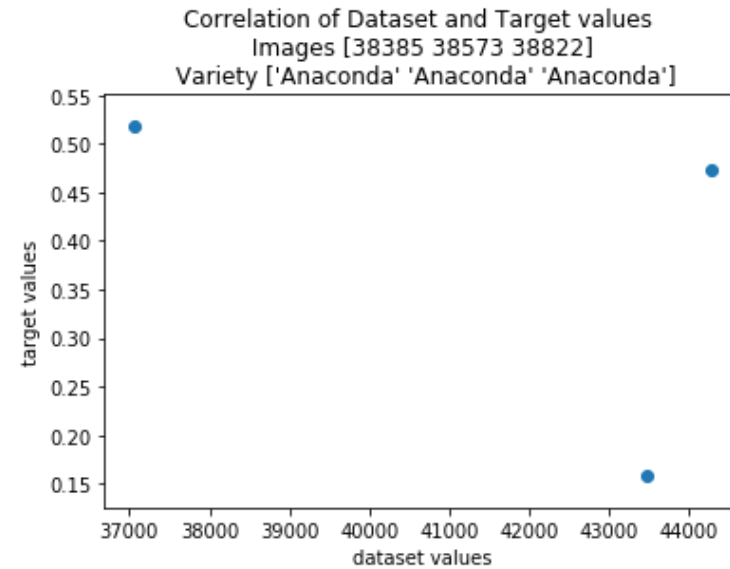
CORRELATION?



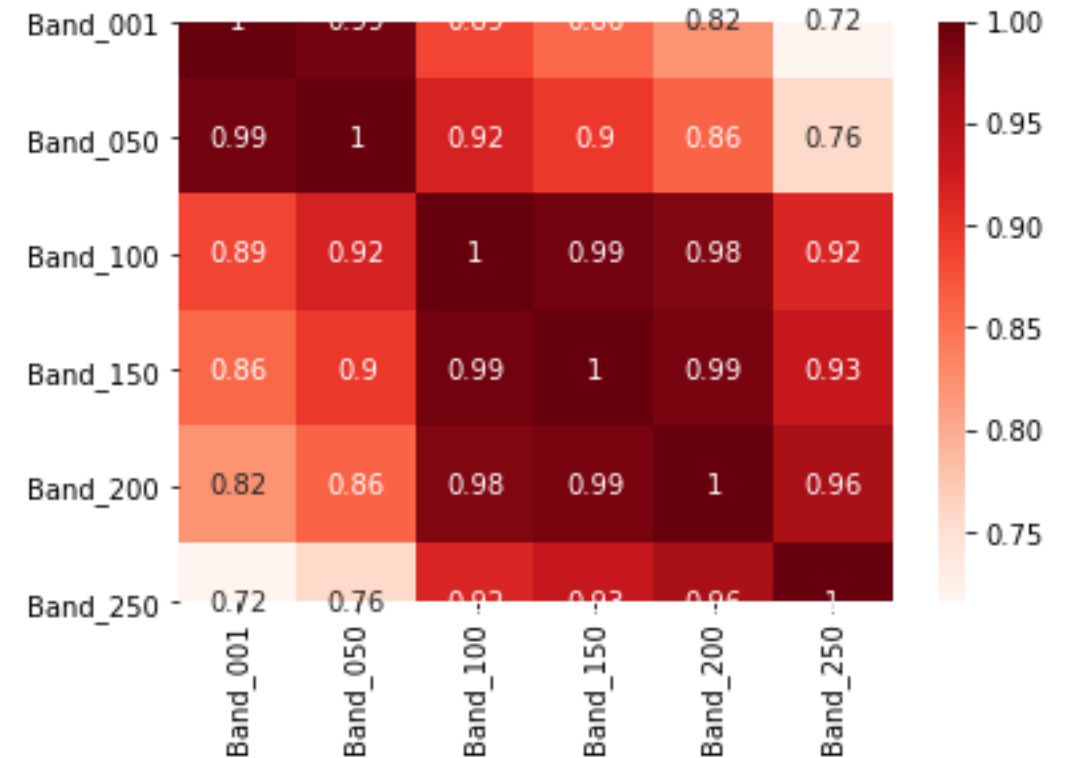
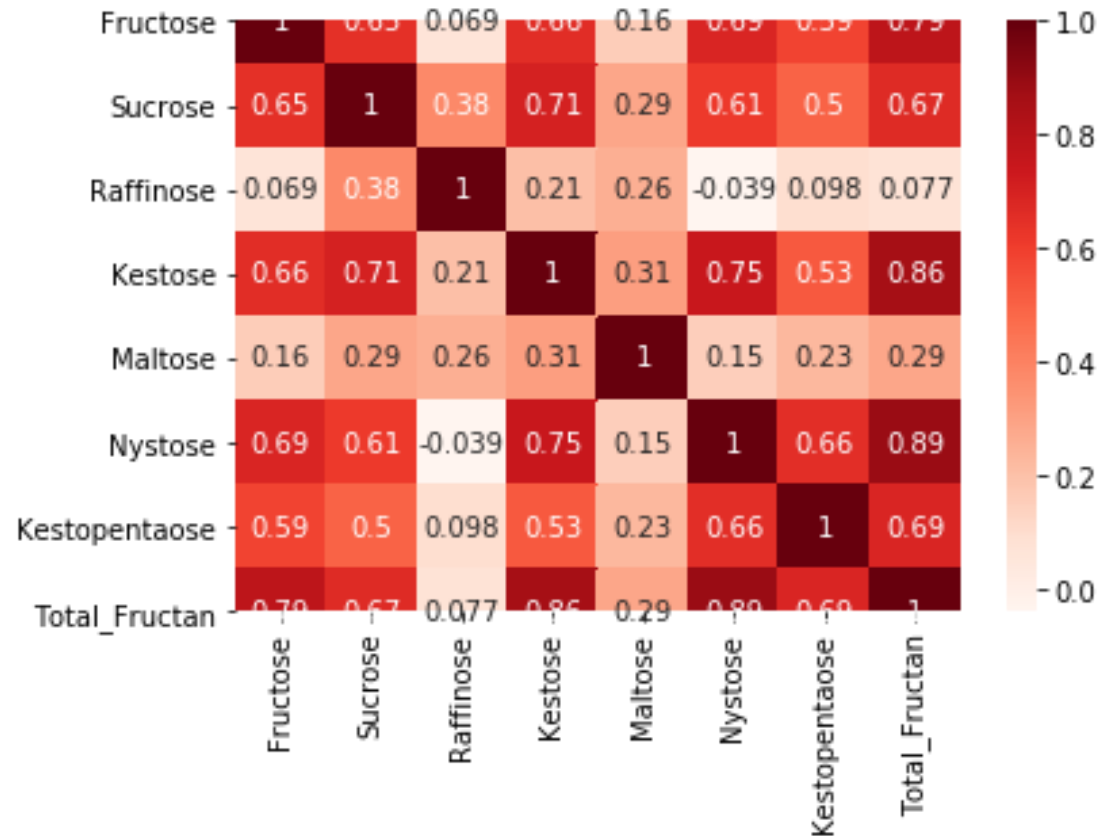
# GENOTYPE AGAINST DATASET

- Normally we have three images for each Variety ID

Correlation?  
Noise?

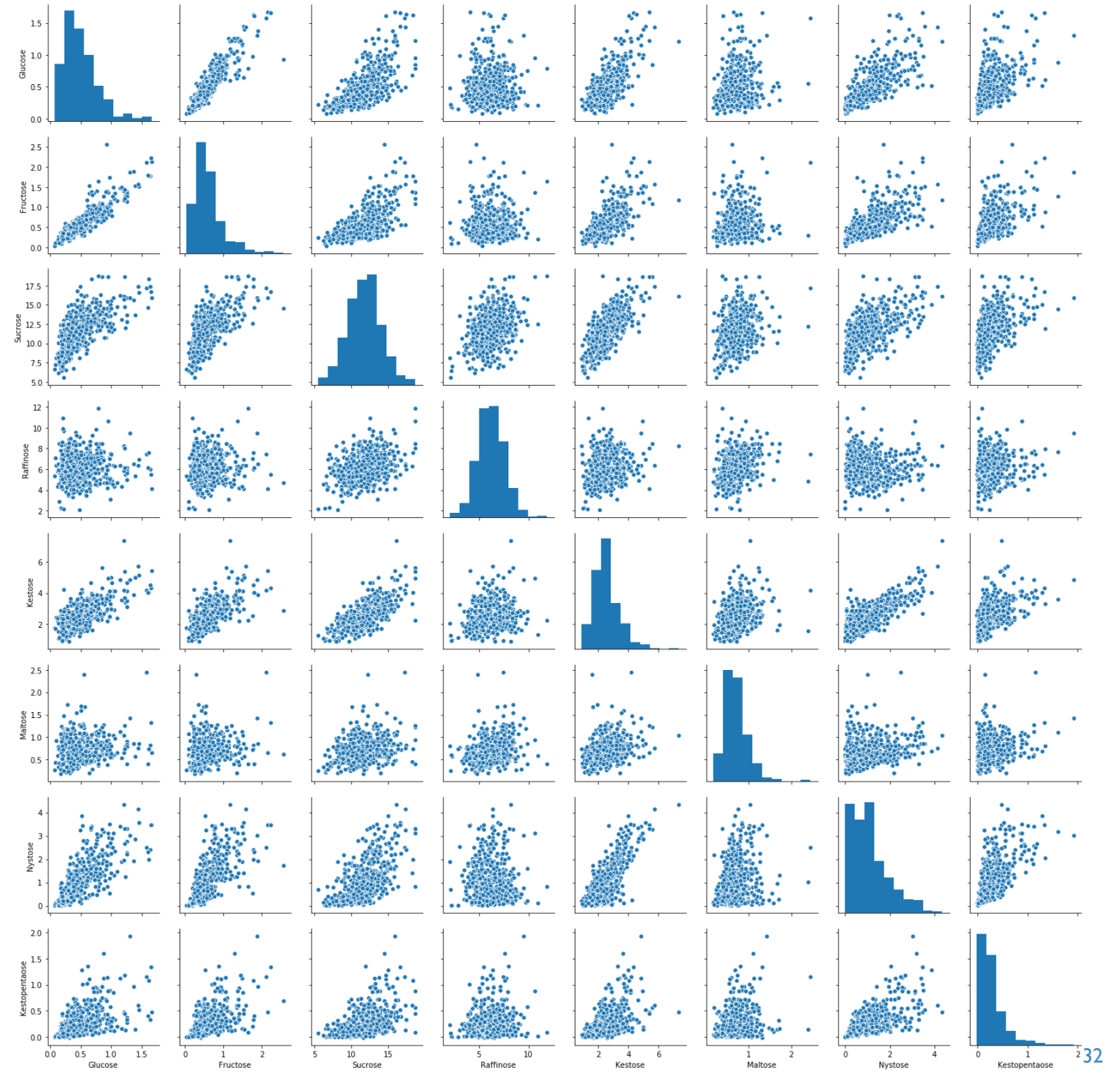
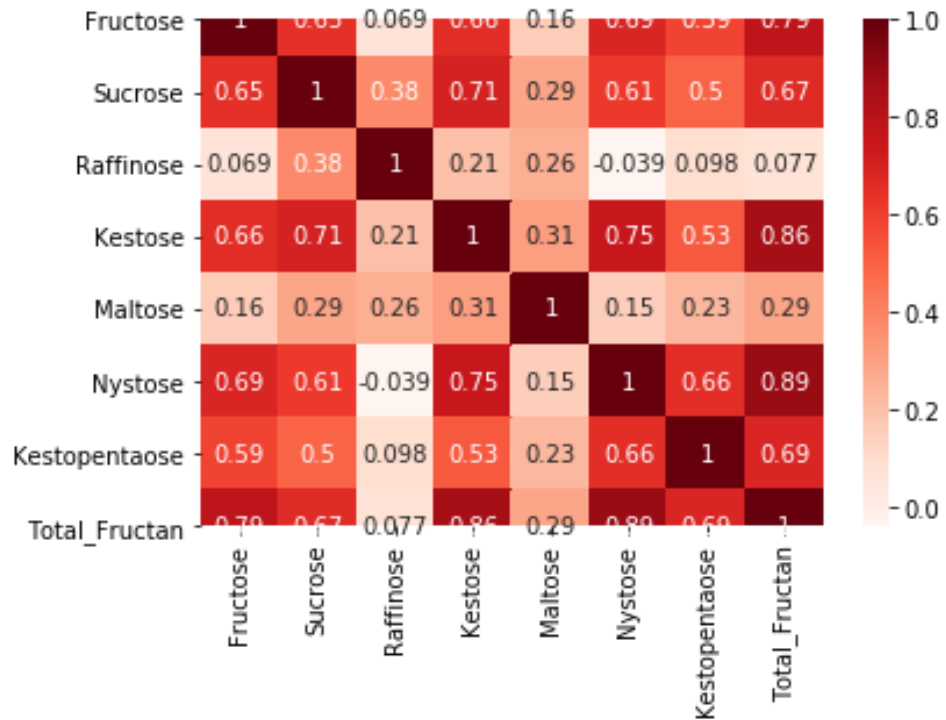


# CORRELATION OF DATASET AND SUGAR CONTENTS INDIVIDUALLY



# SUGAR CONTENTS

- Mean value of each channel for all the images build our dataset





# REMOVING NOISY IMAGES

- Taking Sucrose in to consideration
- Relative Error > 0.4

---

## List of 28 images having relative error larger than 0.4

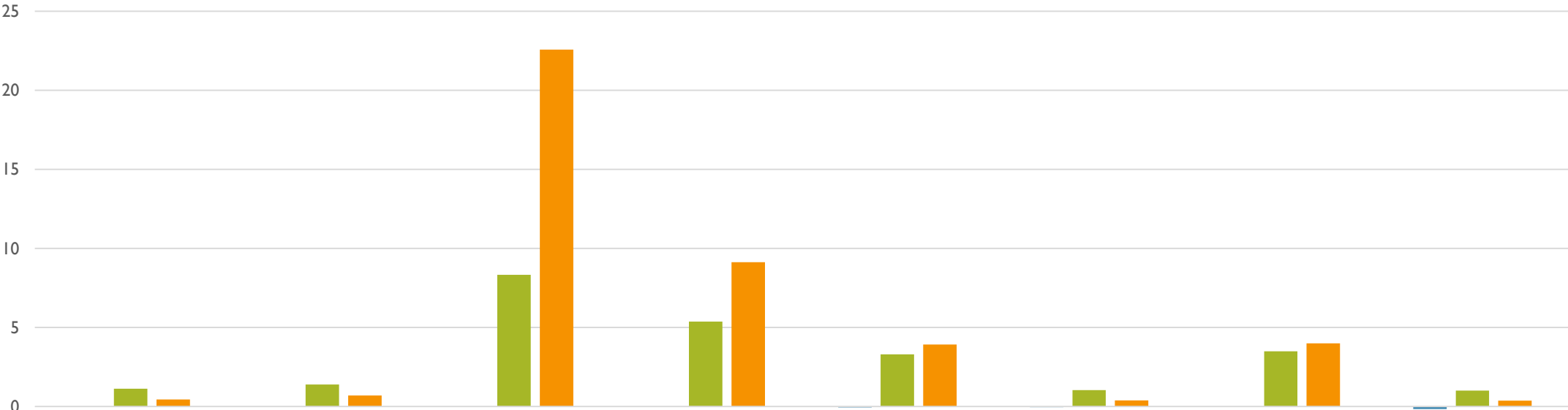
---

38371	38571	38629	38805
38390	38583	38699	38806
38405	38588	38727	38824
38473	38589	38728	38828
38475	38608	38750	38858
38530	38622	38791	38861
38568	38627	38793	38865

---

# REMOVING NOISY IMAGES

Elastic net



	Glucose	Fructose	Sucrose	Raffinose	Kestose	Maltose	Nystose	Kestopentaose
r2_score_error	-0.022733179	-0.040082646	-0.026646006	-0.018498523	-0.063862966	-0.056756575	-0.033618863	-0.176293146
mean_absolute_error	1.111945009	1.390164268	8.321508529	5.36557718	3.29616158	1.026738812	3.482611302	0.991954344
mean_squared_error	0.439913399	0.684157559	22.56793305	9.116560128	3.909442054	0.384637954	3.991911993	0.358894982

■ r2\_score\_error ■ mean\_absolute\_error ■ mean\_squared\_error

# ALL ALGORITHMS TRIED

**Ridge Regression**

**Lasso**

**Elastic Net**

**SVR**

**Random Forest Regression**

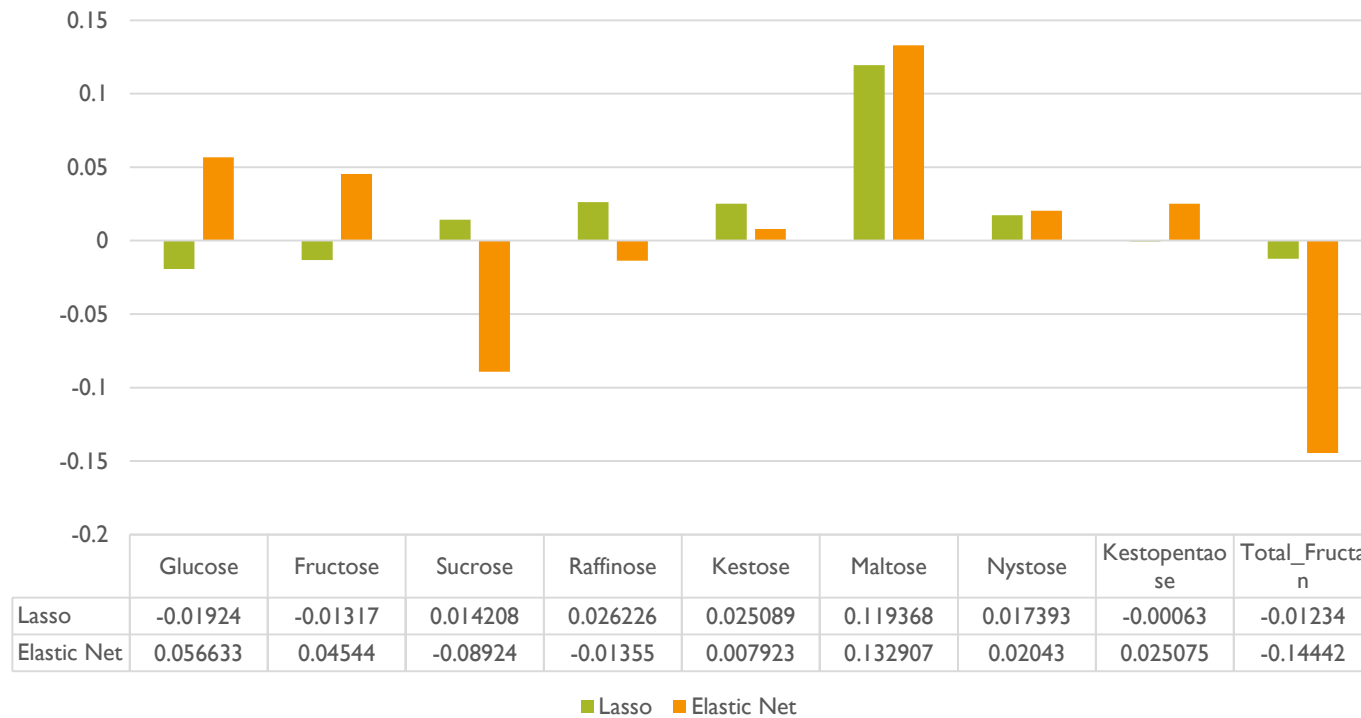
**PLS Regression**

**Stepwise OLS regression**

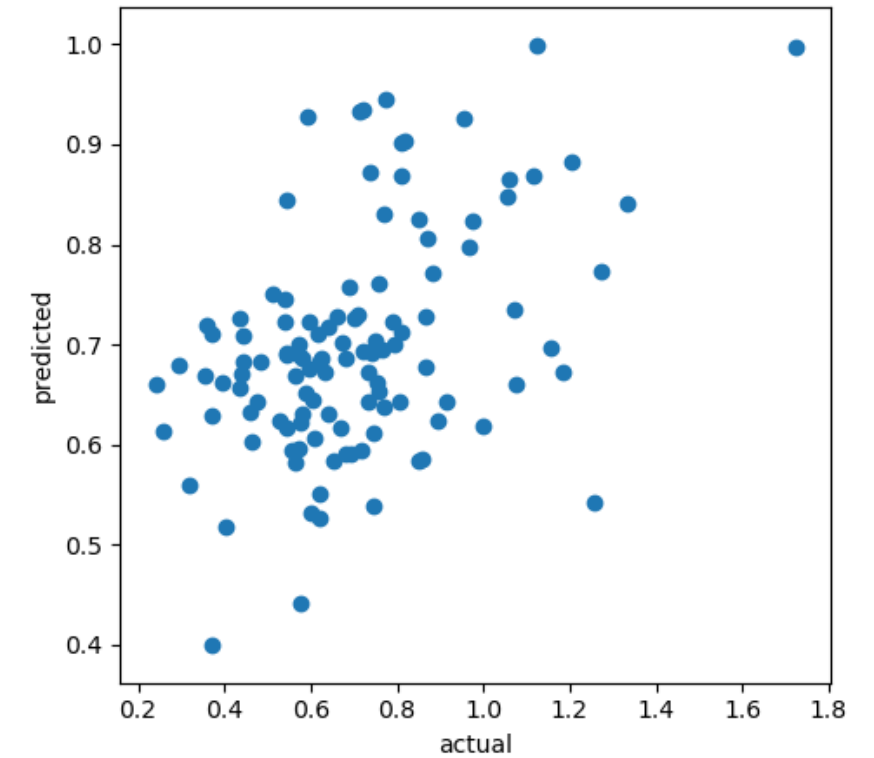
**Deep Neural Networks**

# ELASTIC NET

R<sup>2</sup> error



ElasticNet Regression  
Maltose



THANK YOU!