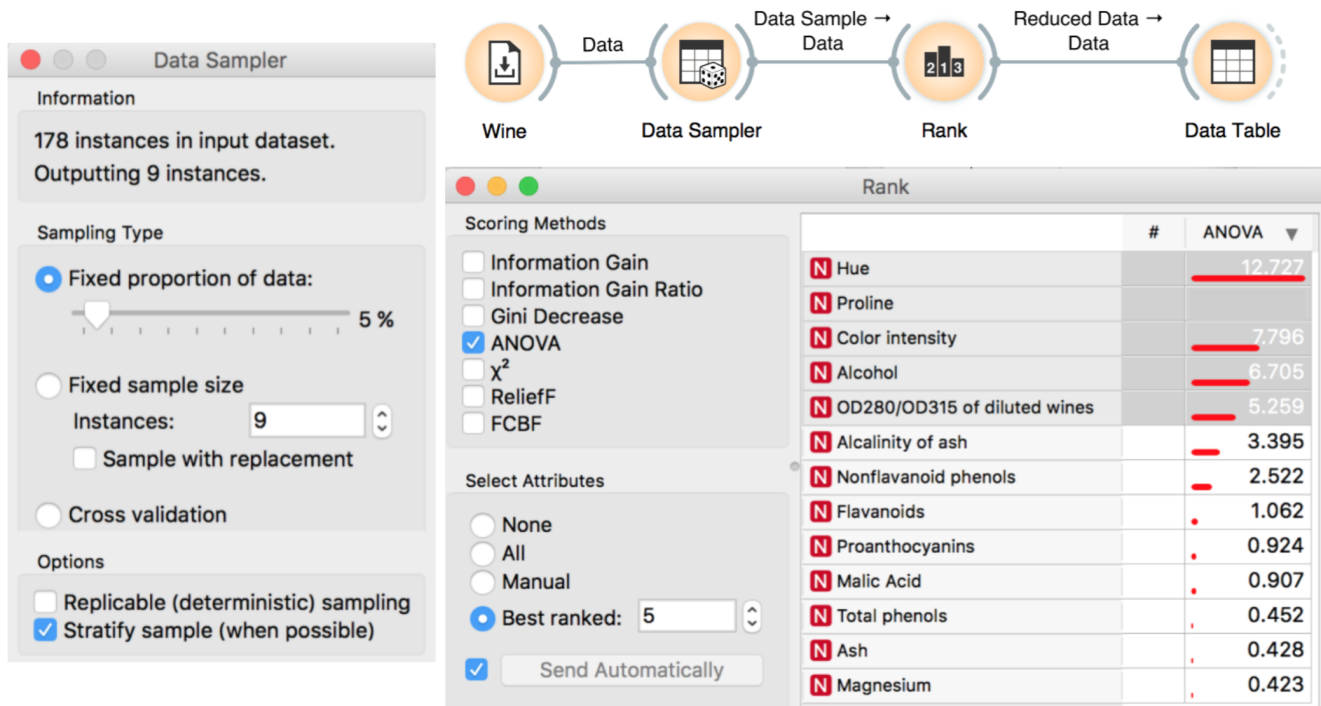
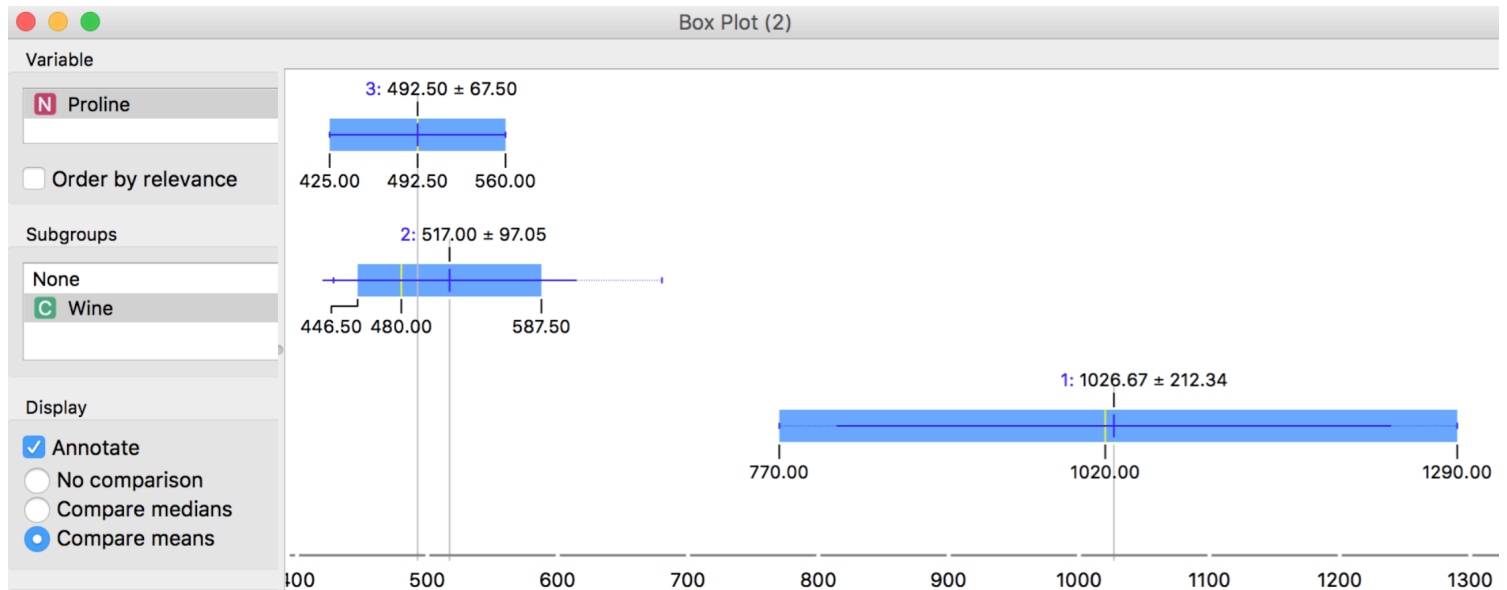


Exercise 1

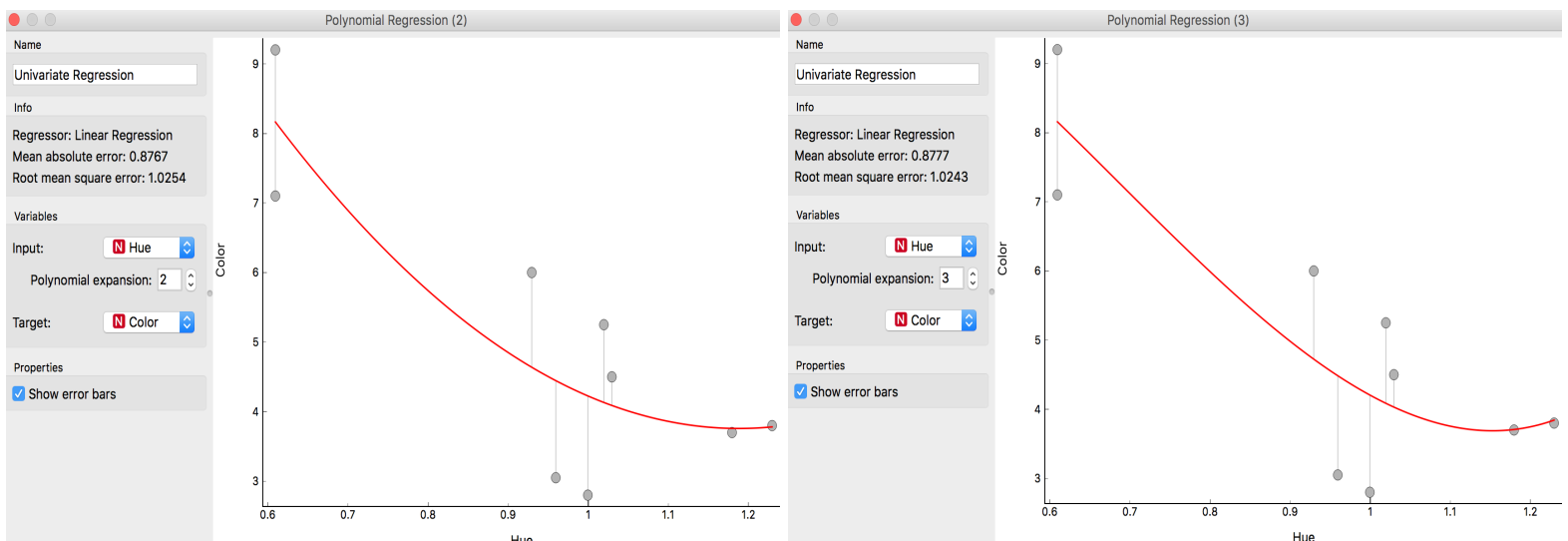


Exercise 2

2) *canvas*

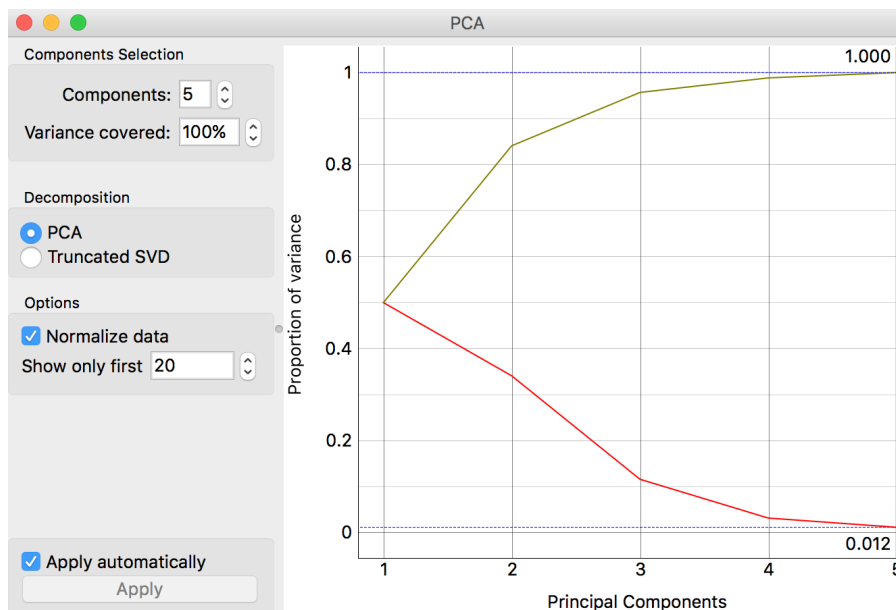


4) modèle (2) et modèle (3)



Exercise 3

1) ACP

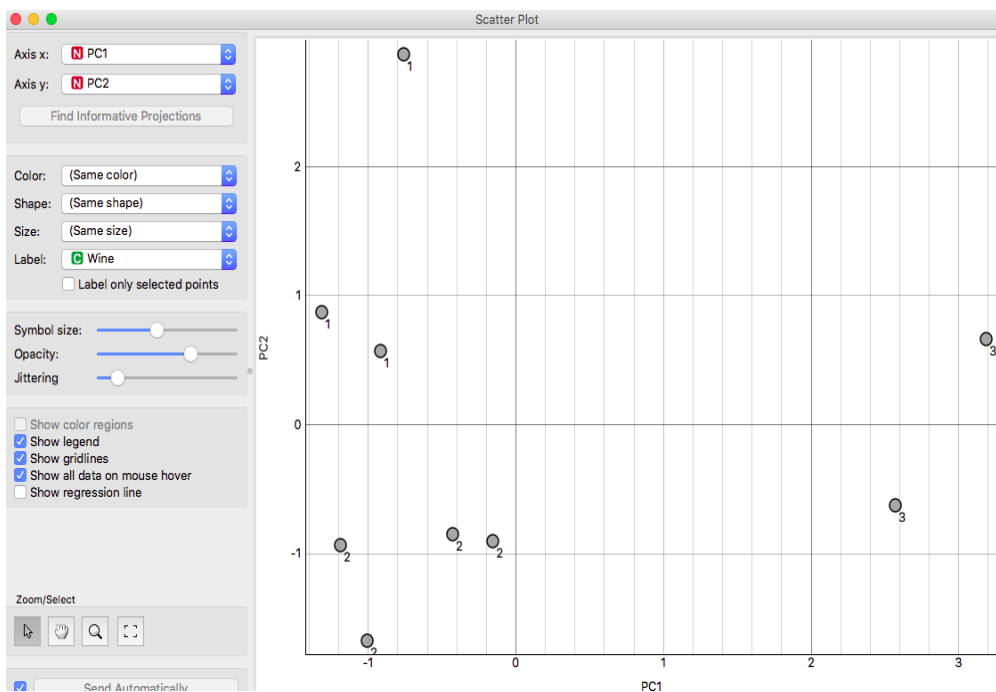


Correlations (1)

Pearson correlation

PC1

1	-0.934	Hue, PC1
3	+0.869	Color, PC1
6	-0.701	OD, PC1
12	+0.460	Alcohol, PC1
13	-0.411	PC1, Proline



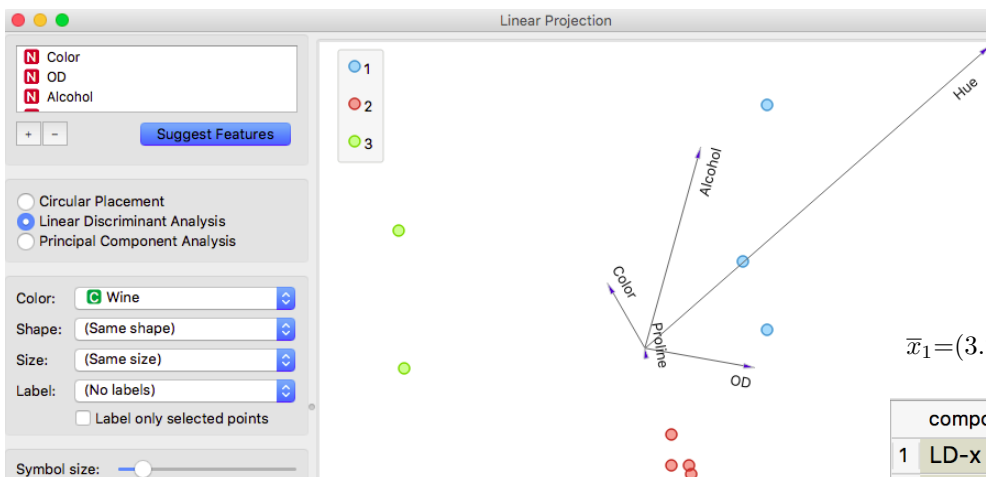
Correlations (1)

Pearson correlation

PC2

2	+0.887	PC2, Proline
4	+0.841	Alcohol, PC2
15	+0.384	OD, PC2
20	+0.249	Color, PC2
33	-0.032	Hue, PC2

2)3)4) AFD



	Wine	LD-x	LD-y
1	2	2.98	8.78
2	2	3.04	8.78
3	1	3.29	11.83
4	2	2.98	9.05
5	2	3.05	8.71
6	1	3.21	10.51
7	1	3.29	9.93
8	3	2.11	9.6
9	3	2.1	10.77

$$\bar{x}_1 = (3.26, 10.76) \quad \bar{x}_2 = (3.01, 8.83) \quad \bar{x}_3 = (2.1, 10.18)$$

	component	Hue	Proline	Color	Alcohol	OD
1	LD-x	0.94	-0.0	-0.10	0.15	0.30
2	LD-y	0.82	0.0	0.18	0.54	-0.05

Exercise 4

Contingency Table

Rows: Wine

Columns: Cluster

Scores:
ARI: 0.416
AMI: 0.448
 $\chi^2=9.00$, $p=0.061$

		Cluster			Σ
		1	2	3	
Wine	1	2	0	1	3
	2	0	4	0	4
	3	0	2	0	2
	Σ	2	6	1	9

Exercise 5

Test & Score

Sampling

☐ Cross validation
Number of folds: 2
☒ Stratified

☐ Cross validation by feature

☐ Random sampling
Repeat train/test: 2
Training set size: 50 %
☒ Stratified

☒ Leave one out
☐ Test on train data
☐ Test on test data

Target Class
(Average over classes)

Evaluation Results

Method	AUC	CA	F1	Precision	Recall
kNN	0.510	0.444	0.420		0.444
Random Forest	0.683	0.556	0.524	0.500	0.556

kNN

Name
kNN

Neighbors
Number of neighbors: 2
Metric: Euclidean
Weight: Distance

Confusion Matrix

Show: Number of instances

kNN
Random Forest

Output
☒ Predictions ☐ Probabilities
☒ Send Automatically

		Predicted			Σ
		1	2	3	
Actual	1	2	1	0	3
	2	1	2	1	4
	3	0	2	0	2
	Σ	3	5	1	9

Select Correct Select Misclassified Clear Selection