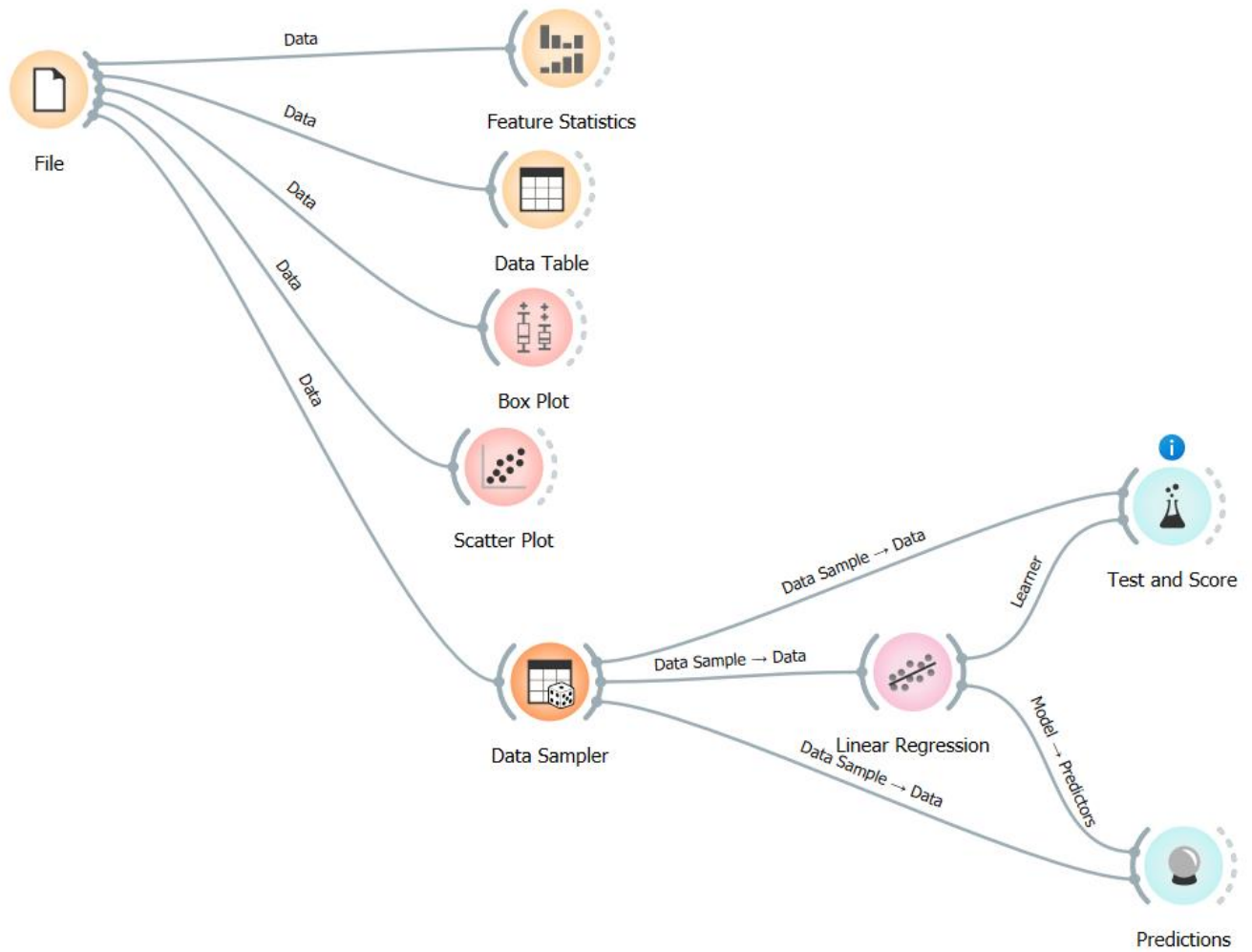
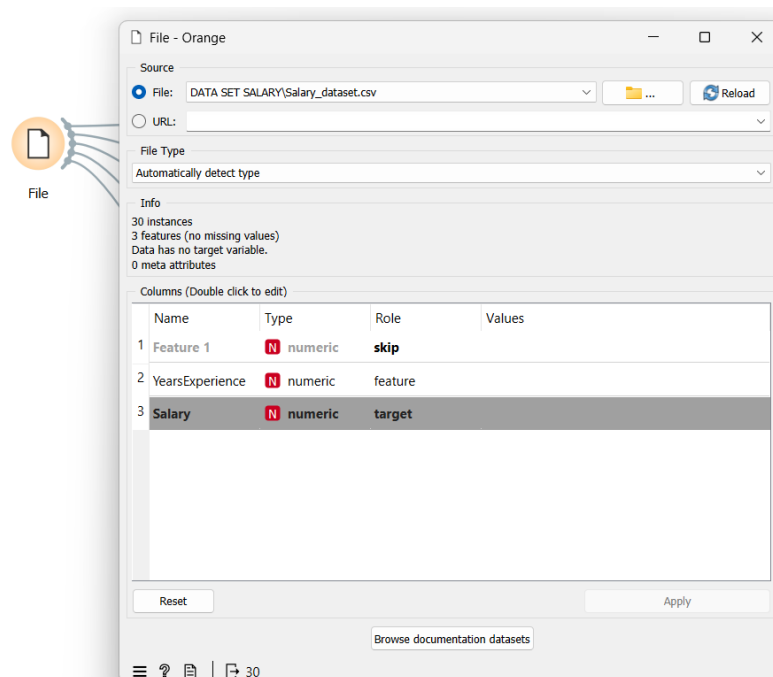


Tugas Machine Learning

Membuat regression model dengan orange data mining



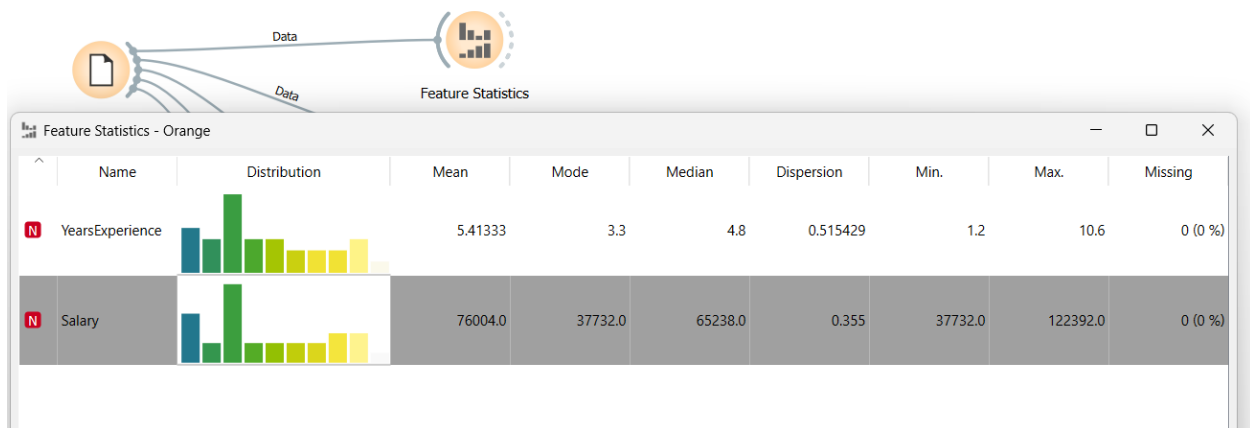
1. Mengupload File berisi Dataset yang akan digunakan



Columns (Double click to edit)				
	Name	Type	Role	Values
1	Feature 1	N numeric	skip	
2	YearsExperience	N numeric	feature	
3	Salary	N numeric	target	

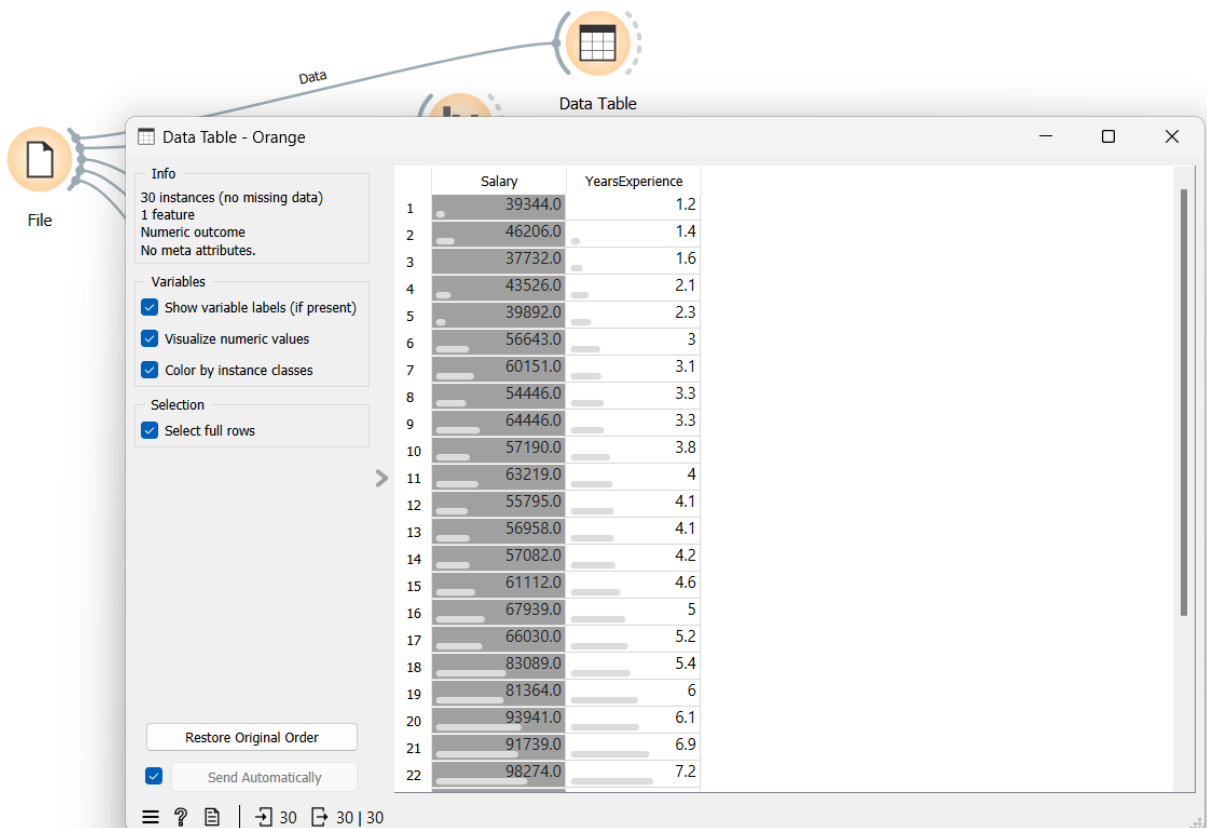
Kita dapat mengubah Role setiap Columns yang ada didalam Dataset seperti *feature* atau juga *target*, kita juga bisa mengubah rolenya *skip* untuk mengabaikan colum.

2. Feature Statistic

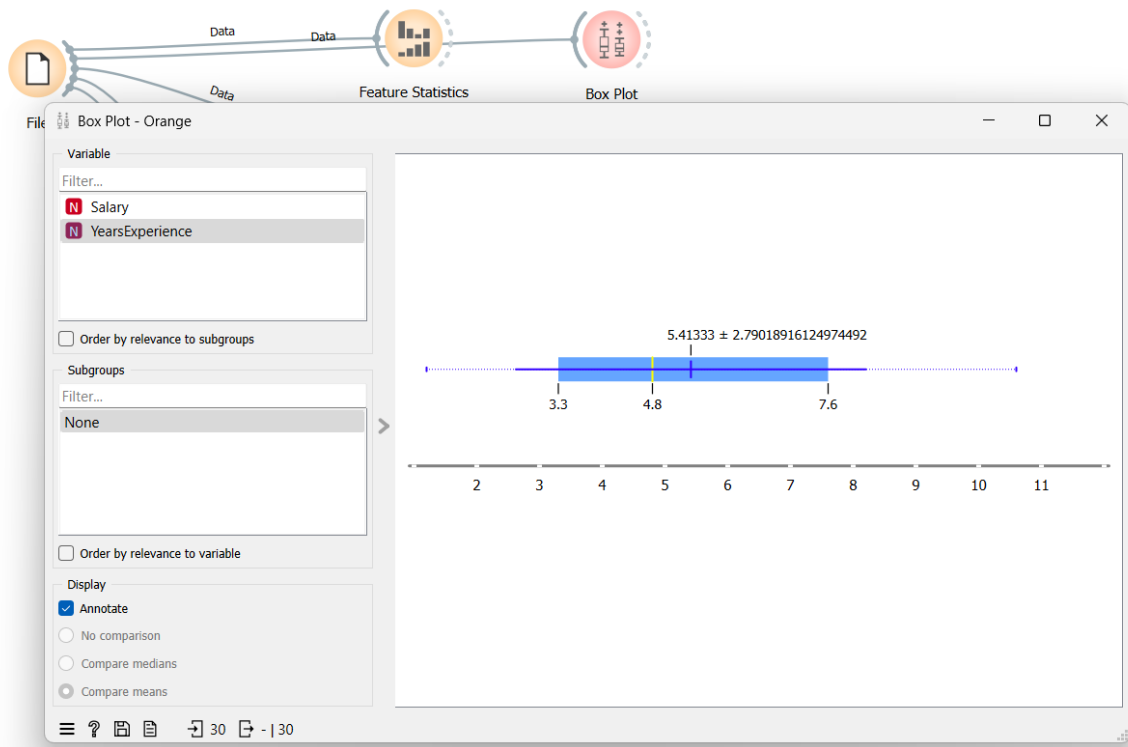


Untuk menampilkan data dalam dataset seperti *Mean*, *Mode*, *Median*, *Depression*, *Min.*, *Max.*, *Missing* dan juga distribusi datanya.

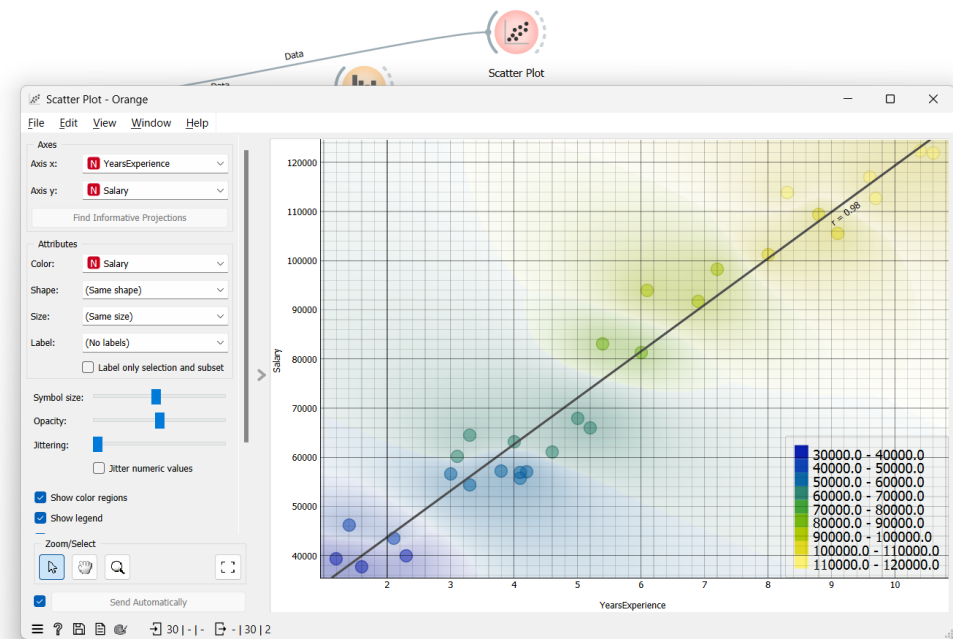
3. Data Table



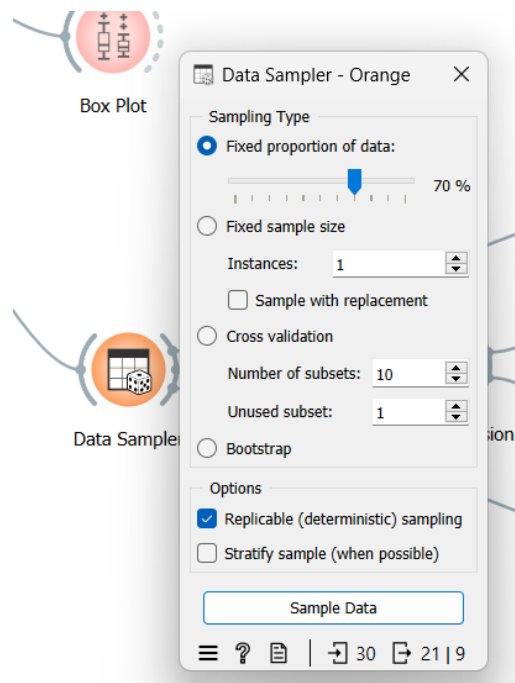
4. Box Plot



5. Scatter Plot

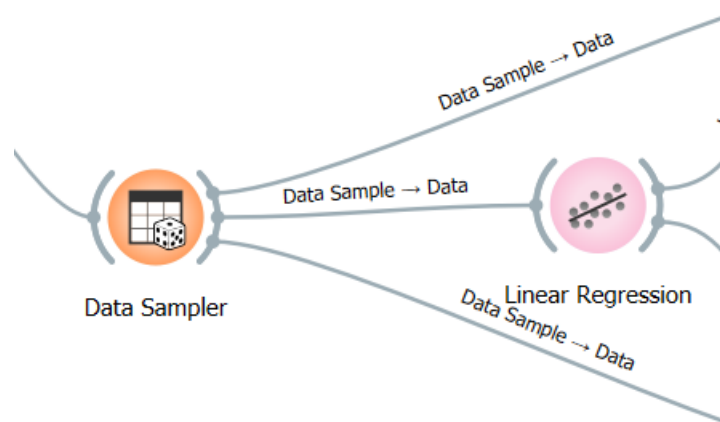


6. Data Sampler

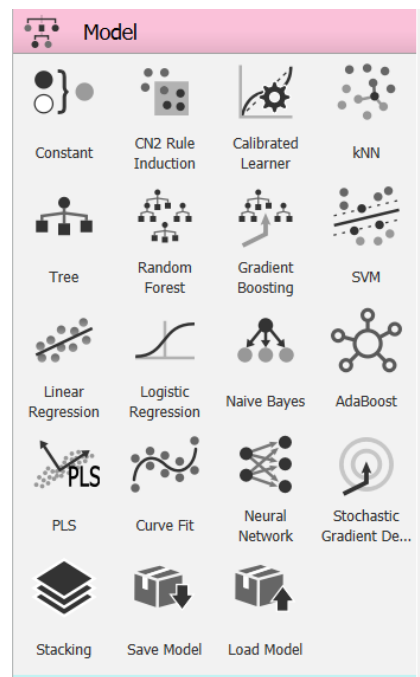


Kita bisa mengatur seberapa persen *Training set* dan juga *Test set*

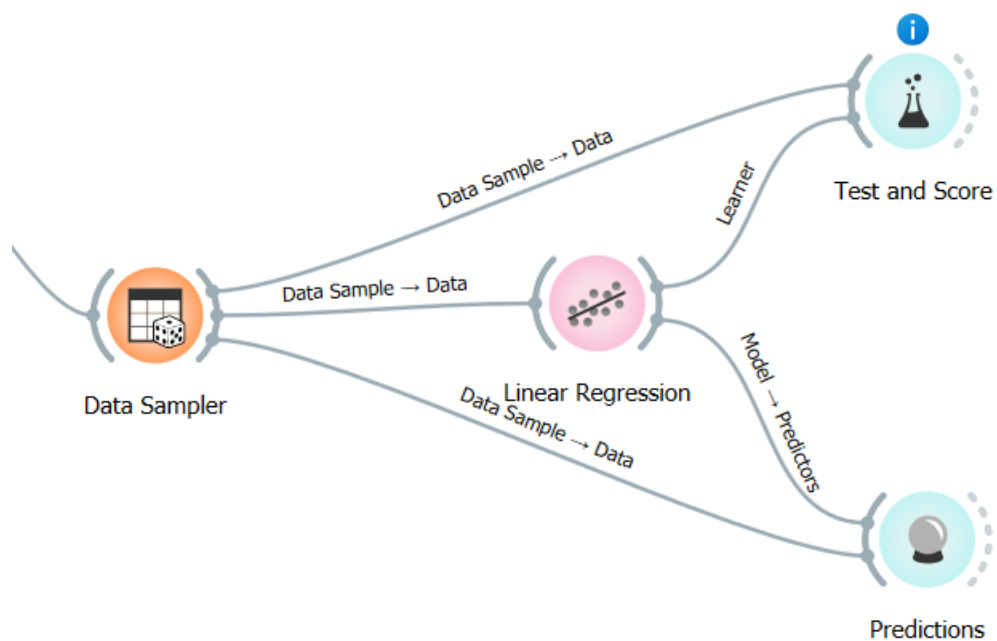
7. Model Liner Regression



Kita dapat memilih beberapa model yang akan digunakan salah satunya model ***Linear Regression***



8. Evaluate Data



Setelah memilih model Algoritma kita bisa melakukan ***Evaluate*** data salah satunya yaitu ***Test and Score*** dan juga ***Predictions***

a. Test and Score

Test and Score - Orange

☒ Cross validation

Number of folds: 5
 ☒ Stratified

☐ Cross validation by feature

☐ Random sampling

Repeat train/test: 10

Training set size: 66 %

☒ Stratified

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

Model	MSE	RMSE	MAE	MAPE	R2
Linear Regression	41541929.995	6445.303	5688.801	0.091	0.946

b. Prediction

Predictions - Orange

Shown regression error: Difference

	Linear Regression	error	Salary	YearsExperience
1	116172.4	3536.4	112636.0	9.7
2	71666.8	3727.8	67939.0	5
3	102915.4	-1089...	113813.0	8.3
4	75454.5	-7634.5	83089.0	5.4
5	55569.1	-8876.9	64446.0	3.3
6	60303.7	3113.7	57190.0	3.8
7	122800.9	408.9	122392.0	10.4
8	107650.0	-1782.0	109432.0	8.8
9	63144.5	6186.5	56958.0	4.1
10	35683.6	-3660.4	39344.0	1.2
11	46099.8	6207.8	39892.0	2.3
12	73560.7	7530.7	66030.0	5.2
13	52728.3	-3914.7	56643.0	3
14	64091.4	7009.4	57082.0	4.2
15	63144.5	7349.5	55795.0	4.1
16	100074.6	-1228.4	101303.0	8
17	37577.5	-8628.5	46206.0	1.4
18	39471.3	1739.3	37732.0	1.6
19	110490.8	4907.8	105583.0	9.1
20	44205.9	679.9	43526.0	2.1
21	92499.2	-5774.8	98274.0	7.2

☒ Show performance scores

Model	MSE	RMSE	MAE	MAPE	R2
Linear Regression	33216979.733	5763.417	4990.261	0.078	0.957