

Nama : Ameliani Kusmayadi

NIM : 1103213044

Kelas : TK-45-06

Tugas Regression Model dengan Orange Data Mining

The screenshot displays the Orange Data Mining interface with a workflow for building a Linear Regression model. The workflow consists of the following widgets and connections:

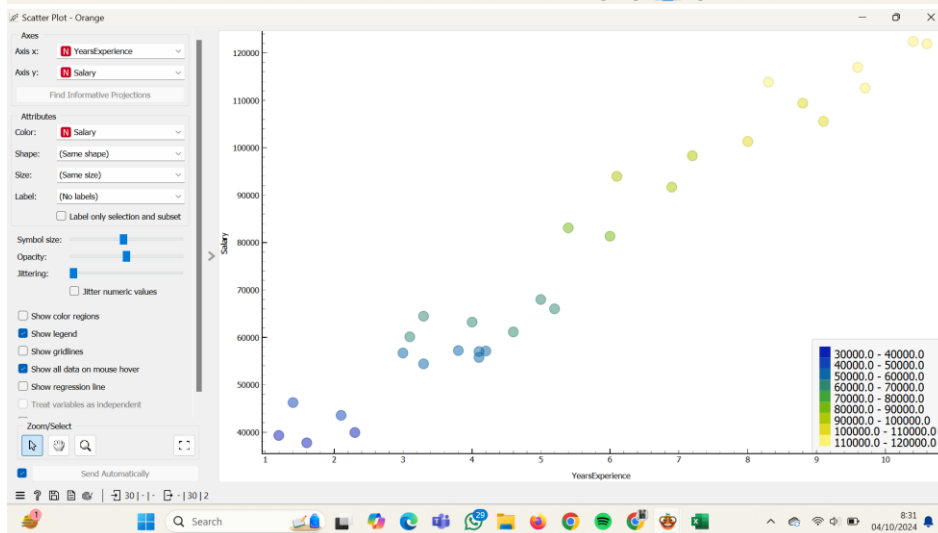
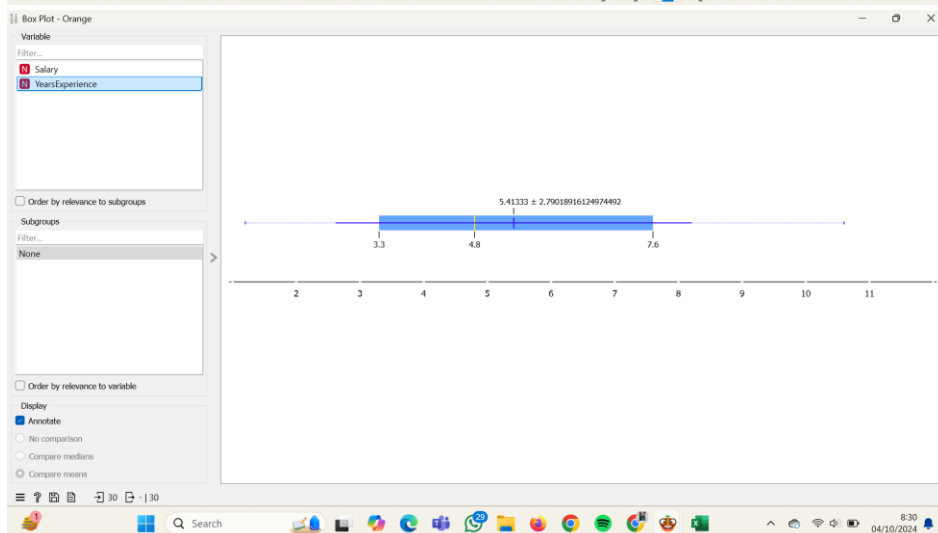
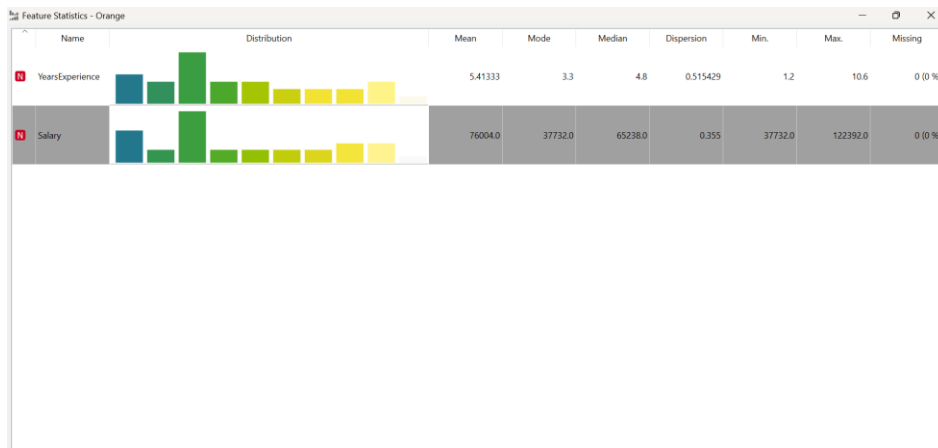
- File** widget connected to **Data Table**.
- Data Table** connected to **Feature Statistics**.
- Feature Statistics** connected to **Box Plot**.
- Box Plot** connected to **Scatter Plot**.
- Scatter Plot** connected to **Data Sampler**.
- Data Sampler** connected to **Linear Regression**.
- Linear Regression** connected to **Test and Score**.
- Test and Score** connected to **Predictions**.

The **Data Table - Orange** window shows the dataset with the following columns:

Salary	Feature 1	YearsExperience
29344.0	0	1.2
46206.0	1	1.4
37732.0	2	1.6
43526.0	3	2.1
39692.0	4	2.3
56643.0	5	3
60151.0	6	3.1
54446.0	7	3.3
64446.0	8	3.3
57190.0	9	3.8
63219.0	10	4
55795.0	11	4.1
56958.0	12	4.1
57082.0	13	4.2
61112.0	14	4.6
67939.0	15	5
66030.0	16	5.2
83089.0	17	5.4
81364.0	18	6
93941.0	19	6.1
91739.0	20	6.9

The **Source** window shows the file path and column roles:

Name	Type	Role	Values
Feature 1	numeric	skip	
YearsExperience	numeric	feature	
Salary	numeric	target	



Data Table - Orange

Info

30 instances (no missing data)
1 feature
Numeric outcome
No meta attributes.

Variables

- Show variable labels (if present)
- Visualize numeric values
- Color by instance classes

Selection

- Select full rows

Restore Original Order

Send Automatically

	Salary	YearsExperience
1	39344.0	1.2
2	46206.0	1.4
3	37732.0	1.6
4	43526.0	2.1
5	39892.0	2.3
6	56643.0	3
7	60151.0	3.1
8	54446.0	3.3
9	64460.0	3.3
10	57190.0	3.8
11	63219.0	4
12	55795.0	4.1
13	56958.0	4.1
14	57082.0	4.2
15	61112.0	4.6
16	67939.0	5
17	66030.0	5.2
18	83089.0	5.4
19	81364.0	6
20	93941.0	6.1
21	91739.0	6.9
22	98274.0	7.2
23	101303.0	8
24	113813.0	8.3
25	109432.0	8.8
26	105583.0	9.1
27	116970.0	9.6
28	112636.0	9.7
29	122392.0	10.4
30	121871.0	10.6

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

Compare models by: Mean square error

Negligible diff.: 0.1

Linear Re...

Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

Stratification is ignored for regression

Predictions - Orange

Shown regression error: Difference

	Linear Regression error	Salary	YearsExperience
1	116172.4 2536.4	39344.0	1.2
2	71666.8 3727.8	46206.0	1.4
3	102015.4 -1099.	37732.0	1.6
4	75454.5 -7634.5	43526.0	2.1
5	55569.1 -8876.9	39892.0	2.3
6	60303.7 3113.7	56643.0	3
7	122800.9 408.9	60151.0	3.1
8	107650.0 -1782.0	54446.0	3.3
9	63144.5 6186.5	64460.0	3.3
10	35683.6 -3660.4	57190.0	3.8
11	46099.8 6207.8	63219.0	4
12	73560.7 7530.7	55795.0	4.1
13	52728.3 -3914.7	56958.0	4.1
14	64091.4 7009.4	57082.0	4.2
15	63144.5 7349.5	61112.0	4.6
16	100074.6 -1228.4	67939.0	5
17	37577.5 -8628.5	66030.0	5.2
18	39471.3 1739.3	83089.0	5.4
19	110490.8 4907.8	81364.0	6
20	44205.9 679.9	93941.0	6.1
21	92499.2 -5774.8	91739.0	6.9
		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