Nama: M. Rizky Fadillah

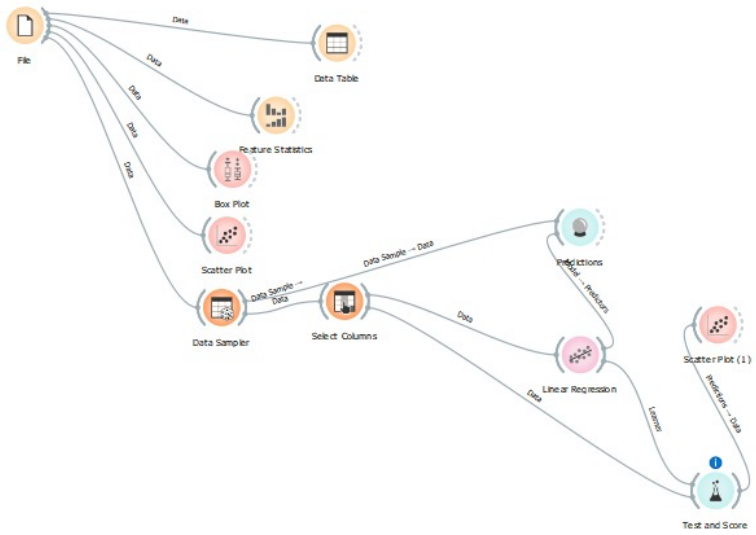
NIM: 1103210259

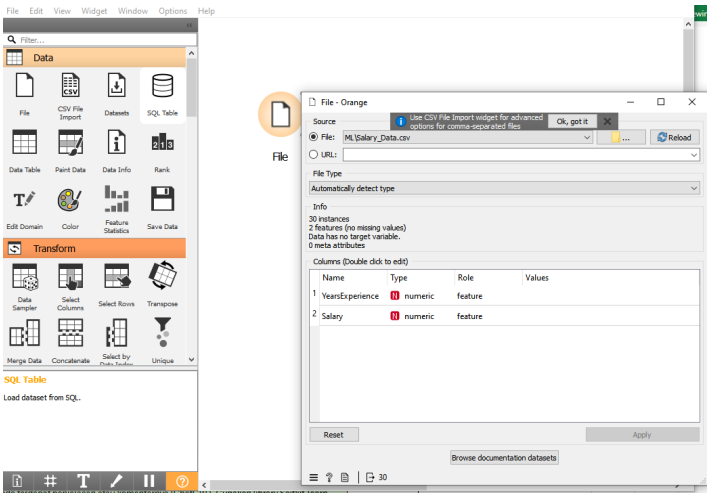
Kelas: TK-45-G04

TUGAS 1 PEMBELAJARAN MESIN

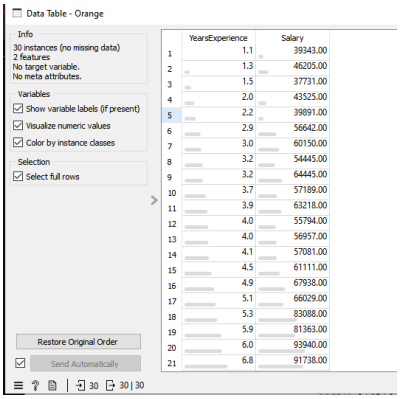
membuat 3 model regression (Simple Linear Regression, Multiple Linear Regression, dan Polynomial Regression)

1. Simple Linear Regression dengan Sallary Dataset

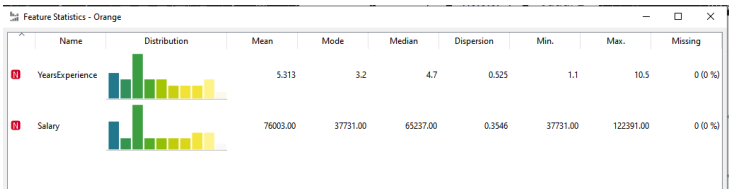




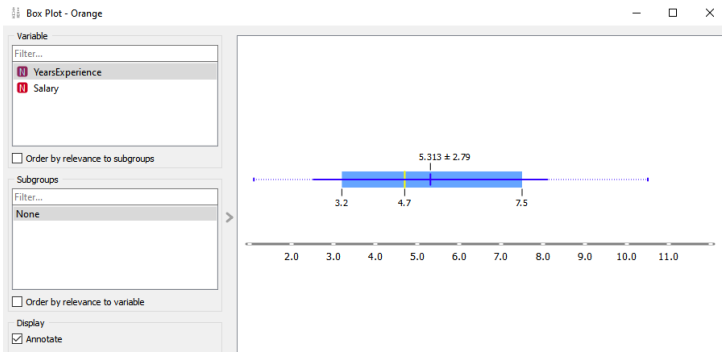
Data Table



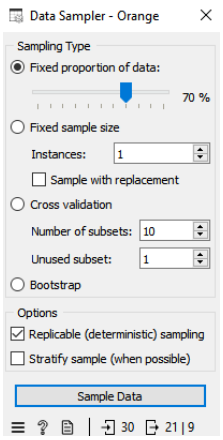
Feature Statistics



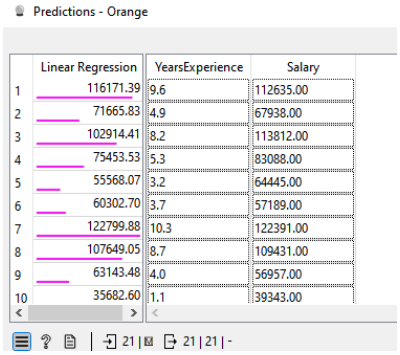
Box Plot



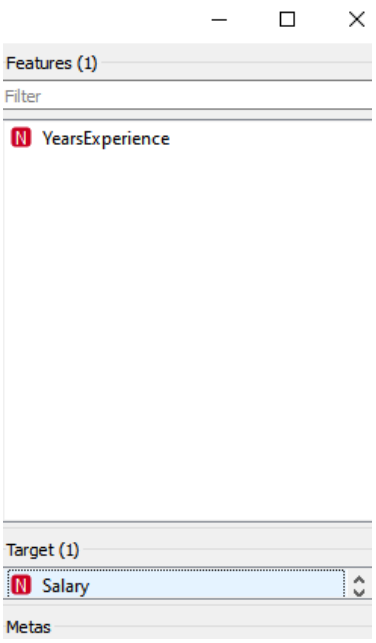
Data Sampler



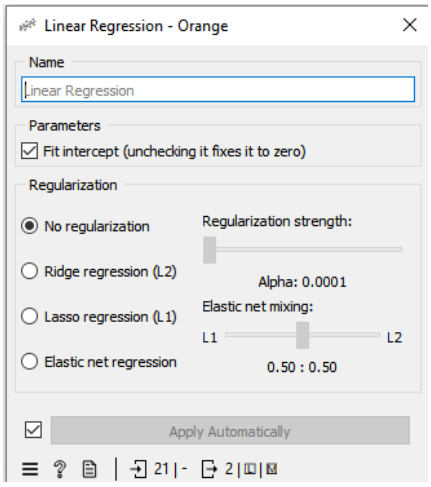
Predictions



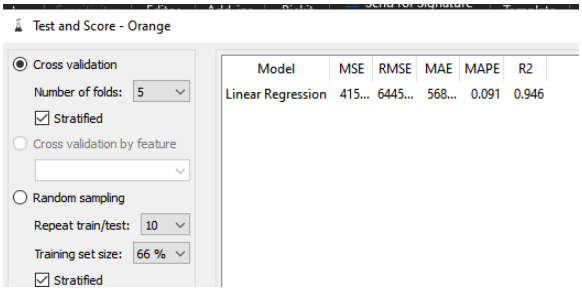
Select Columns



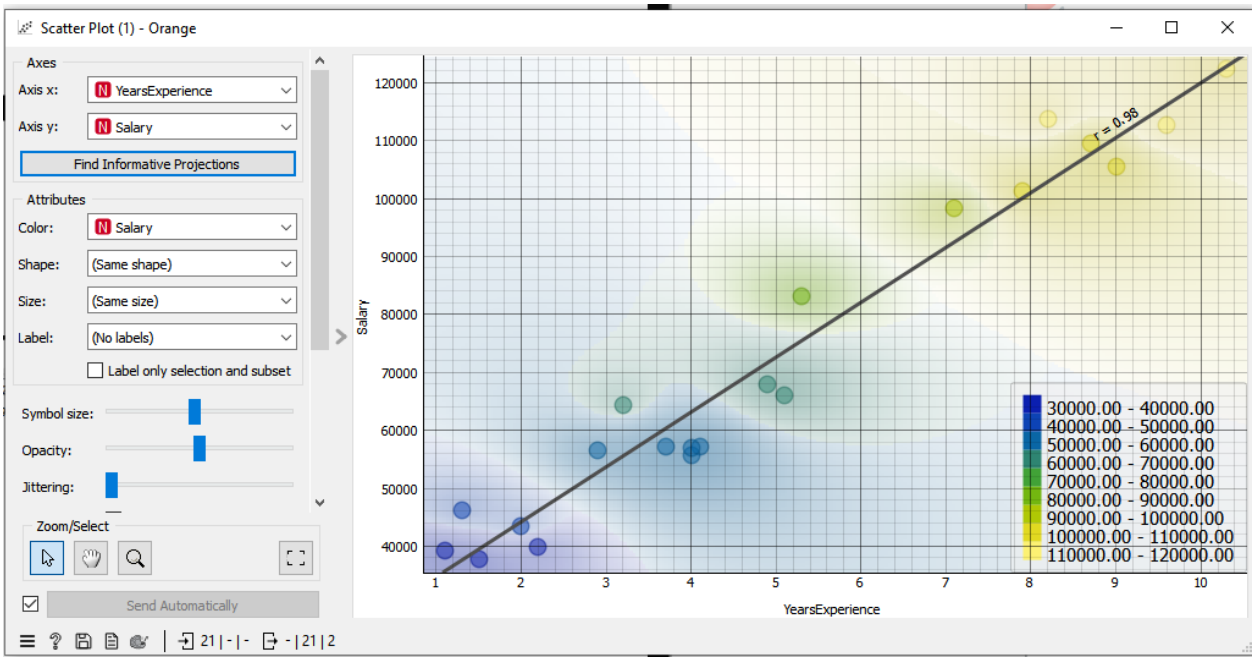
Linear Regression



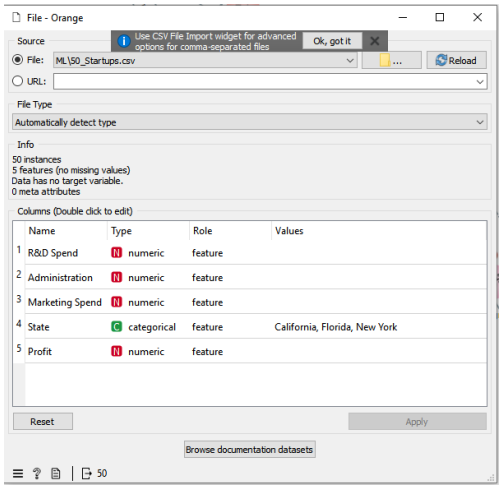
Test and Score

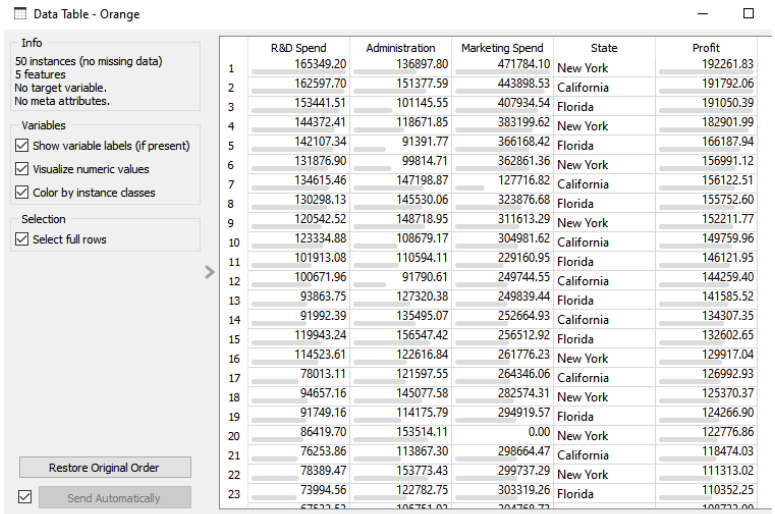


Scatter Plot

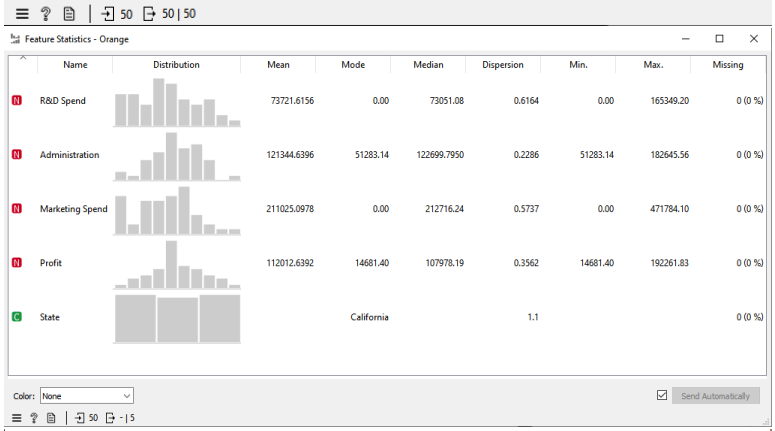


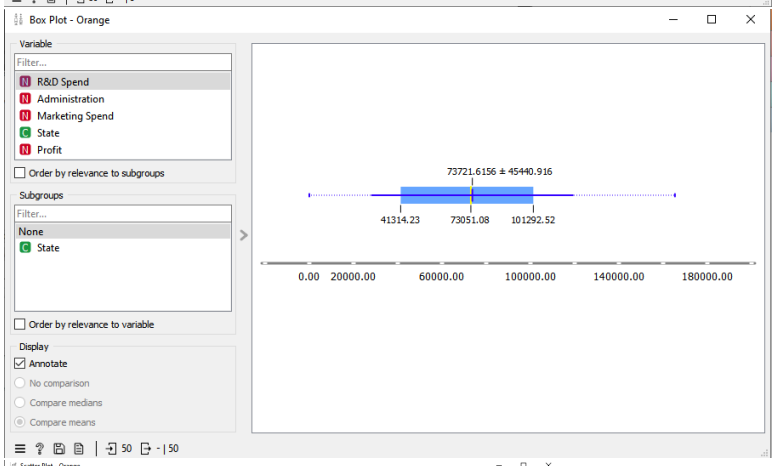
2. Multiple Linear Regression dengan 50 Start Up.csv

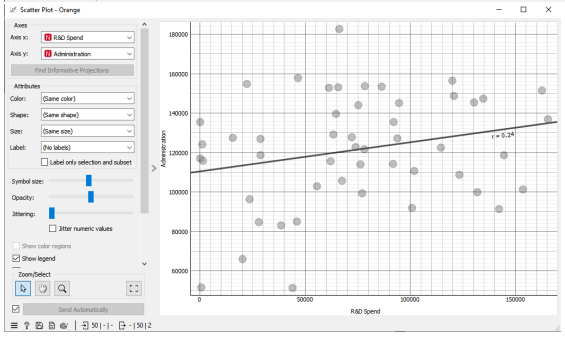


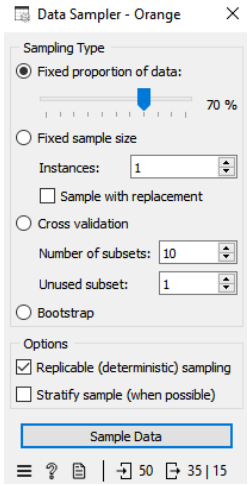


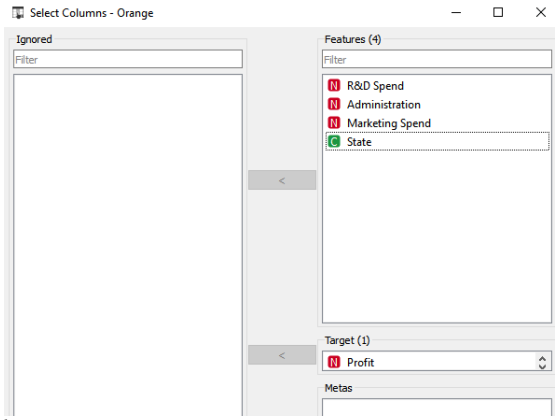
	R&D Spend	Administration	Marketing Spend	State	Profit
1	165349.20	136897.80	471784.10	New York	192261.83
2	162597.70	151377.59	443898.53	California	191792.06
3	153441.51	101145.55	407934.54	Florida	191050.39
4	144372.41	118671.85	383199.62	New York	182901.99
5	142107.34	91391.77	366168.42	Florida	166187.94
6	131876.90	99814.71	362861.36	New York	156991.12
7	134615.46	147198.87	127716.82	California	156122.51
8	130298.13	145530.06	323876.68	Florida	155752.60
9	120542.52	148718.95	311613.29	New York	152211.77
10	123334.88	108679.17	304981.62	California	149759.96
11	101913.08	110594.11	229160.95	Florida	146121.95
12	100671.96	91790.61	249744.55	California	144259.40
13	93863.75	127320.38	249839.44	Florida	141585.52
14	91992.39	135495.07	252664.93	California	134307.35
15	119943.24	156547.42	256512.92	Florida	132602.65
16	114523.61	122616.84	261776.23	New York	129917.04
17	78013.11	121597.55	264346.06	California	126992.93
18	94657.16	145077.58	282574.31	New York	125370.37
19	91749.16	114175.79	294919.57	Florida	124266.90
20	86419.70	153514.11	0.00	New York	122776.86
21	76253.86	113867.30	298664.47	California	118474.03
22	78389.47	153773.43	299737.29	New York	111313.02
23	73994.56	122782.75	303319.26	Florida	110352.25

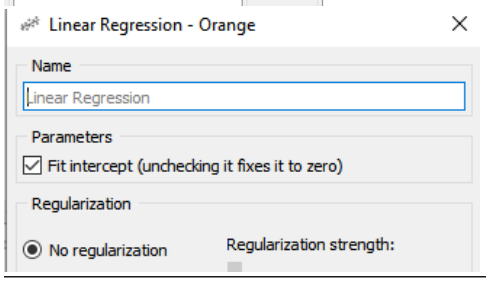


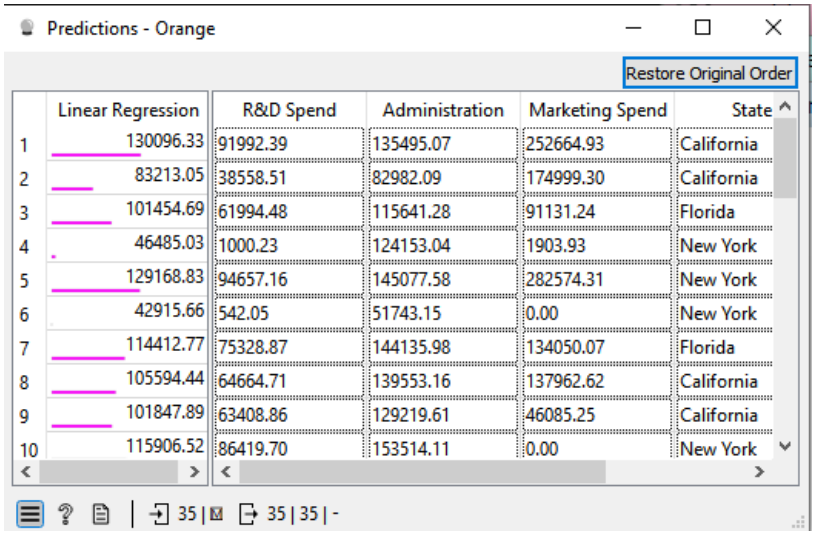




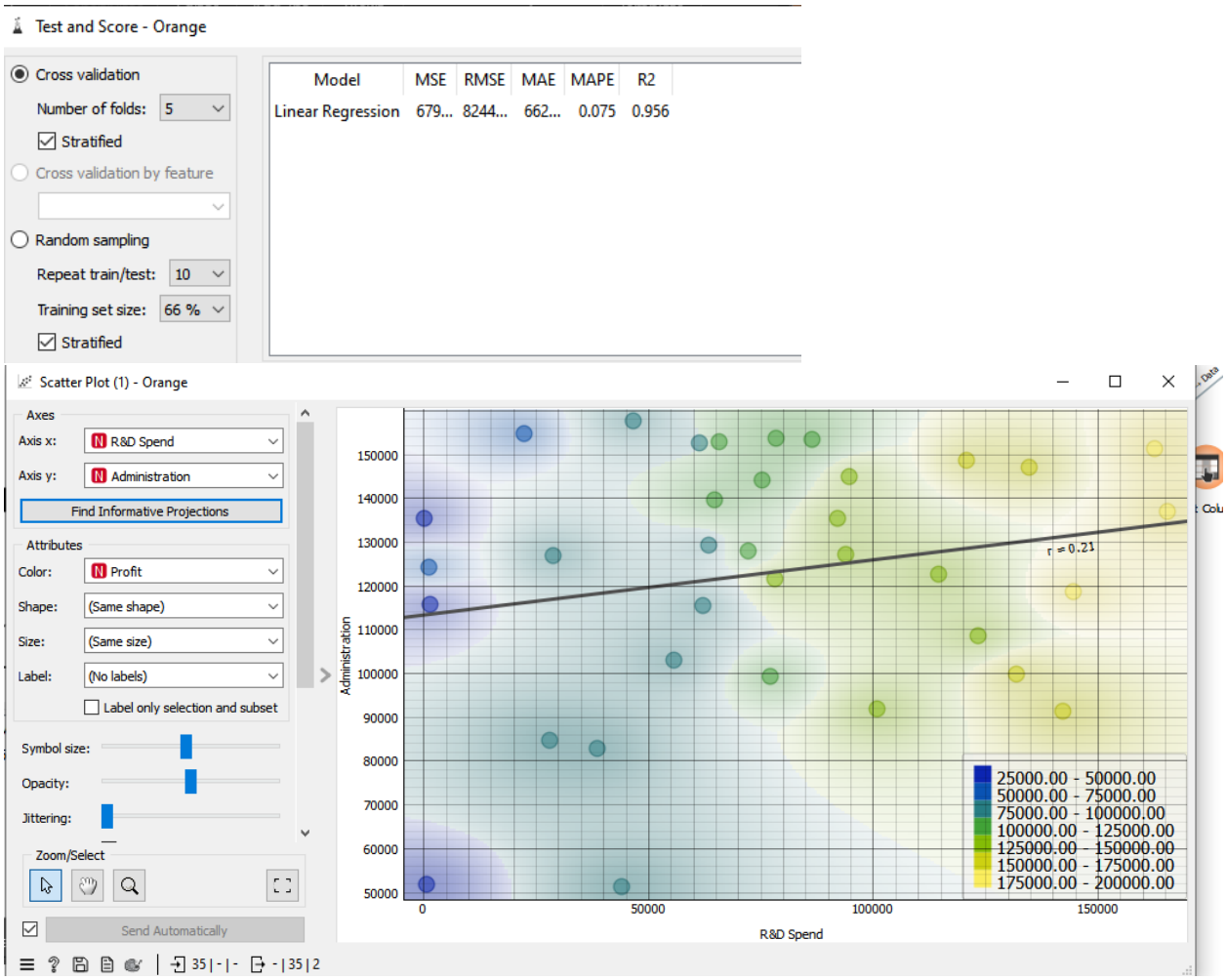




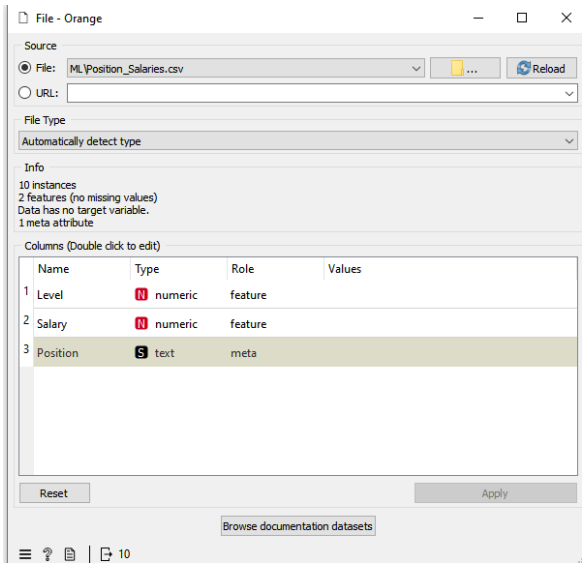




	Linear Regression	R&D Spend	Administration	Marketing Spend	State
1	130096.33	91992.39	135495.07	252664.93	California
2	83213.05	38558.51	82982.09	174999.30	California
3	101454.69	61994.48	115641.28	91131.24	Florida
4	46485.03	1000.23	124153.04	1903.93	New York
5	129168.83	94657.16	145077.58	282574.31	New York
6	42915.66	542.05	51743.15	0.00	New York
7	114412.77	75328.87	144135.98	134050.07	Florida
8	105594.44	64664.71	139553.16	137962.62	California
9	101847.89	63408.86	129219.61	46085.25	California
10	115906.52	86419.70	153514.11	0.00	New York



3. Polynomial Regression menggunakan Position Salarys.csv



Data Table - Orange

Info
10 instances (no missing data)
2 features
No target variable.
1 meta attribute

Variables
- [x] Show variable labels (if present)
- [x] Visualize numeric values
- [x] Color by instance classes

Selection
- [x] Select full rows

	Position	Level	Salary
1	Business Analyst	1	45000
2	Junior Consultant	2	50000
3	Senior Consulta...	3	60000
4	Manager	4	80000
5	Country Manager	5	110000
6	Region Manager	6	150000
7	Partner	7	200000
8	Senior Partner	8	300000
9	C-level	9	500000
10	CEO	10	1000000

Feature Statistics - Orange

Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
Level		5.50	1	5.50	0.52	1	10	0 (0 %)
Salary		249500	45000	130000	1.14	45000	1000000	0 (0 %)

Box Plot - Orange

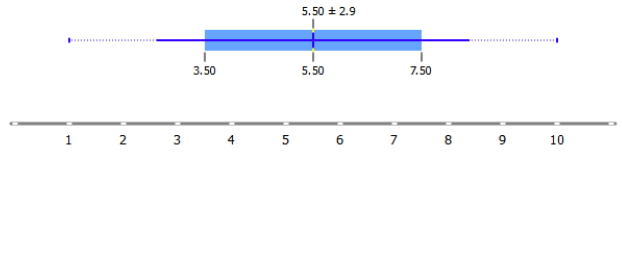
Variable
Filter...
- Level
- Salary

Order by relevance to subgroups

Subgroups
Filter...
- None

Order by relevance to variable

Display
- [x] Annotate



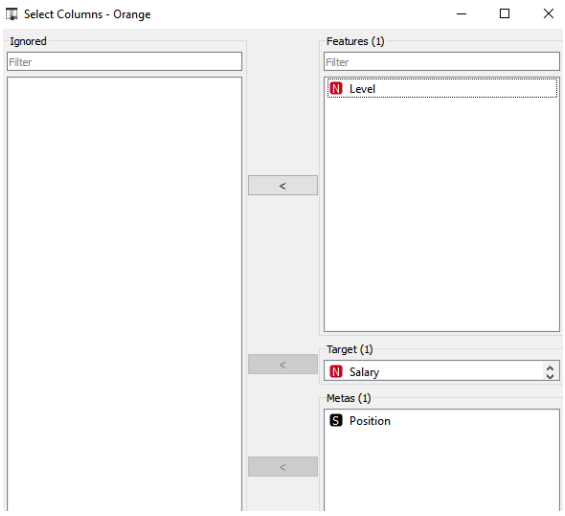
Data Sampler - Orange

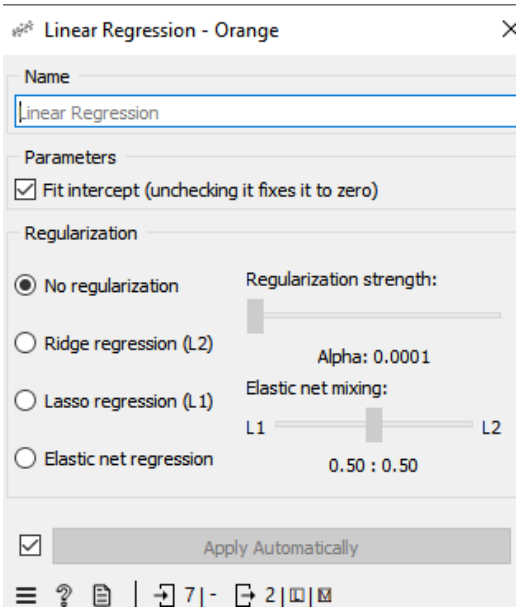
Sampling Type
- (•) Fixed proportion of data: 70 %
- ( ) Fixed sample size
  - Instances: 1
  - [ ] Sample with replacement
- ( ) Cross validation
  - Number of subsets: 10
  - Unused subset: 1
- ( ) Bootstrap

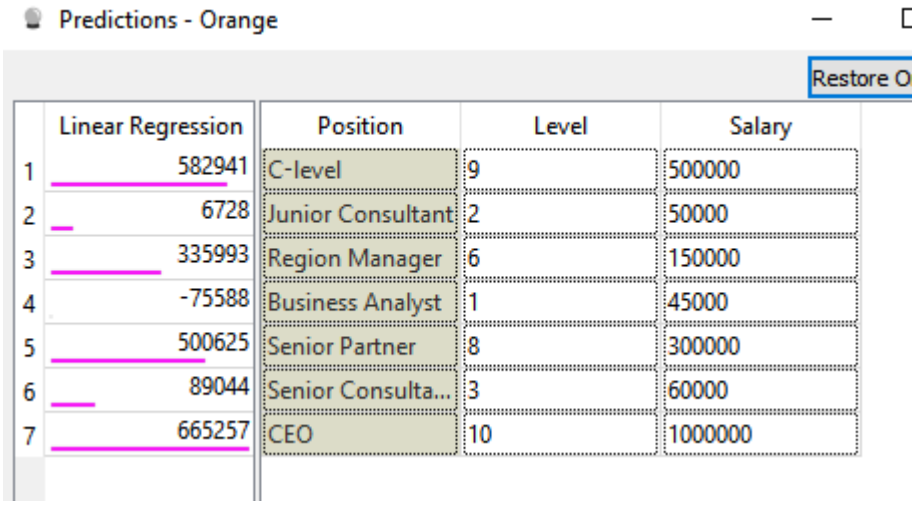
Options
- [x] Replicable (deterministic) sampling
- [ ] Stratify sample (when possible)

Sample Data

10 | 7 | 3

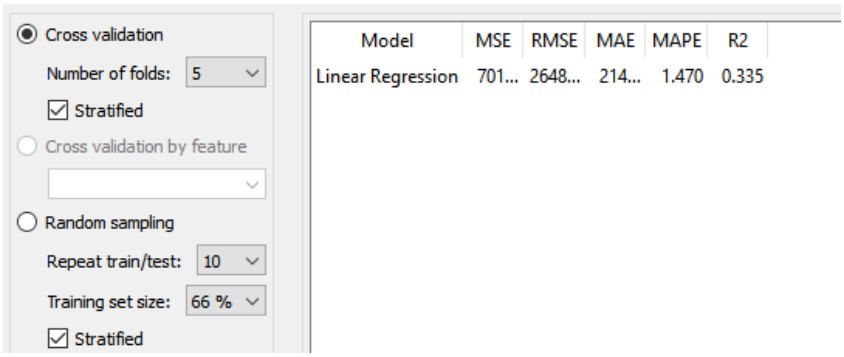






	Linear Regression	Position	Level	Salary
1	582941	C-level	9	500000
2	6728	Junior Consultant	2	50000
3	335993	Region Manager	6	150000
4	-75588	Business Analyst	1	45000
5	500625	Senior Partner	8	300000
6	89044	Senior Consulta...	3	60000
7	665257	CEO	10	1000000

Test and Score - Orange



Model	MSE	RMSE	MAE	MAPE	R2
Linear Regression	701...	2648...	214...	1.470	0.335