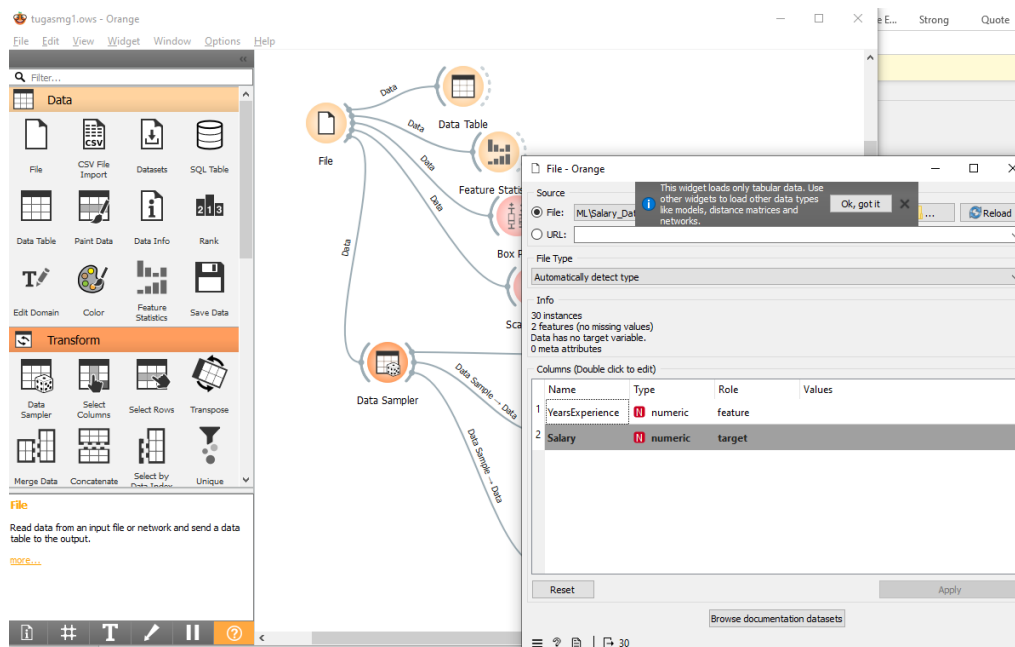
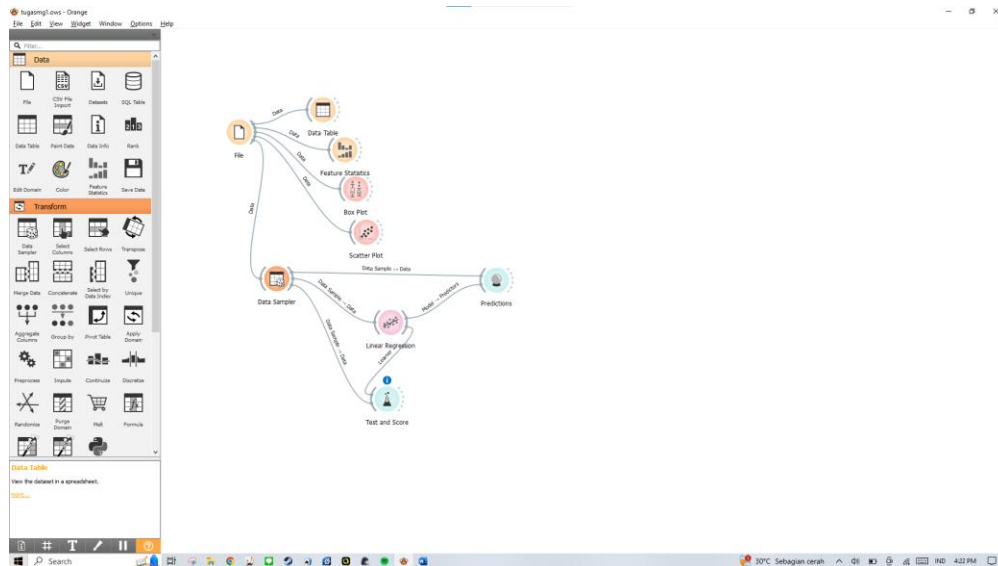


Tugas 1

Nama : muhammad makhluhi makbullah

Kelas : TK 45 01

NIM : 1103210171



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

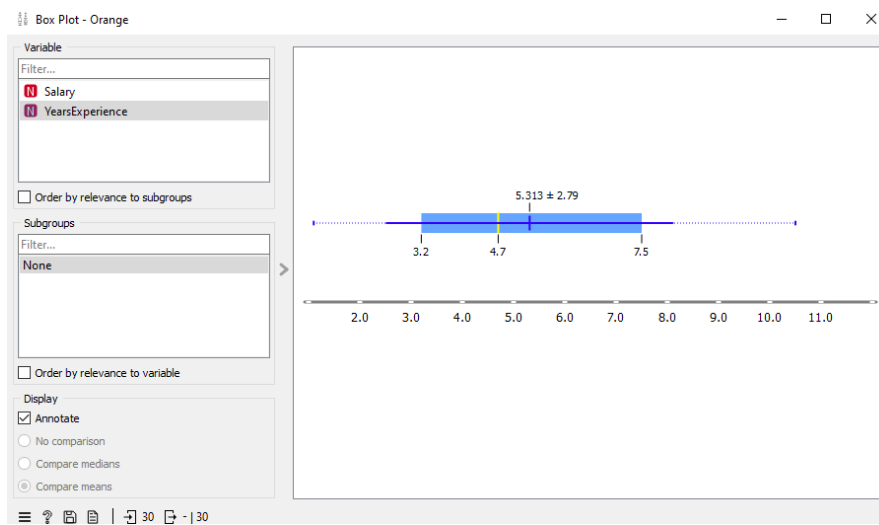
	Salary	YearsExperience
1	39343.00	1.1
2	46205.00	1.3
3	37731.00	1.5
4	43525.00	2.0
5	39891.00	2.2
6	56642.00	2.9
7	60150.00	3.0
8	54445.00	3.2
9	64445.00	3.2
10	57189.00	3.7
11	63218.00	3.9
12	55794.00	4.0
13	56957.00	4.0
14	57081.00	4.1
15	61111.00	4.5
16	67938.00	4.9
17	66029.00	5.1
18	83088.00	5.3
19	81363.00	5.9
20	93940.00	6.0
21	91738.00	6.8

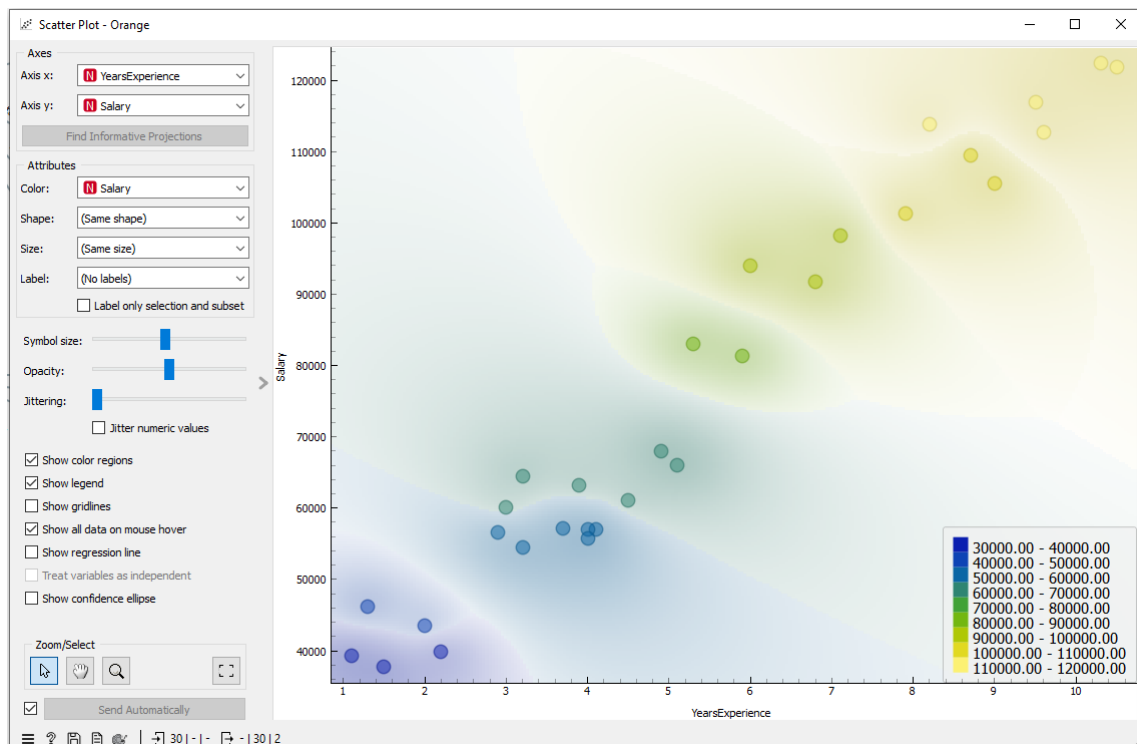
Feature Statistics - Orange

	Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
YearsExperience			5.313	3.2	4.7	0.525	1.1	10.5	0 (0 %)
Salary			76003.00	37731.00	65237.00	0.3546	37731.00	122391.00	0 (0 %)

Colors: ☒ Salary

☒ Send Automatically





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

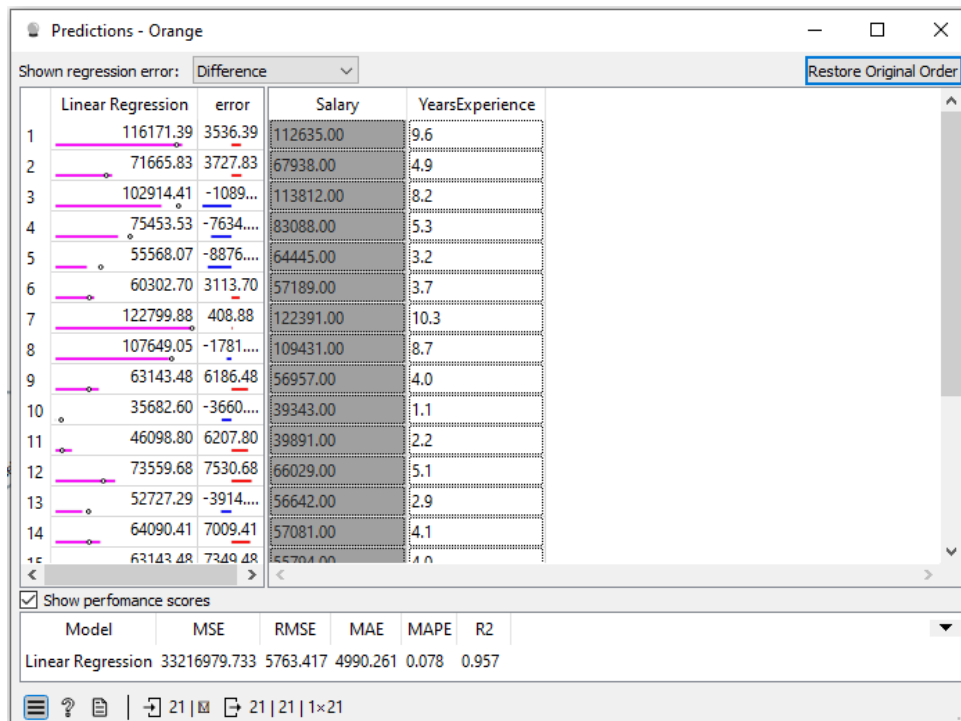
Model	MSE	RMSE	MAE	MAPE	R2
Linear Regression	415...	6445...	568...	0.091	0.946

Compare models by: Mean square error
☐ Negligible diff.: 0.1

Linear Regression	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



```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score
from sklearn.neighbors import KNeighborsRegressor
from sklearn.metrics import mean_squared_error
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay

pip install scikit-learn

Requirement already satisfied: scikit-learn in /usr/local/lib/python3.10/dist-packages (1.5.2)
Requirement already satisfied: numpy>=1.19.5 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.26.4)
Requirement already satisfied: scipy>=1.6.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.13.1)
Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.4.2)
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (3.5.0)

data = pd.read_csv('Salary_Data.csv')
print(data.head())

YearsExperience  Salary
0               1.1   39343.0
1               1.3   46205.0
2               1.5   37731.0
3               2.0   43525.0
4               2.2   39891.0

print(data.dtypes)

YearsExperience    float64
Salary             float64
dtype: object

plt.scatter(data['YearsExperience'], data['Salary'], color='blue')
plt.title("Years of Experience vs Salary")
plt.xlabel("Years of Experience")
plt.ylabel("Salary")
plt.grid(True)
plt.show()
```

