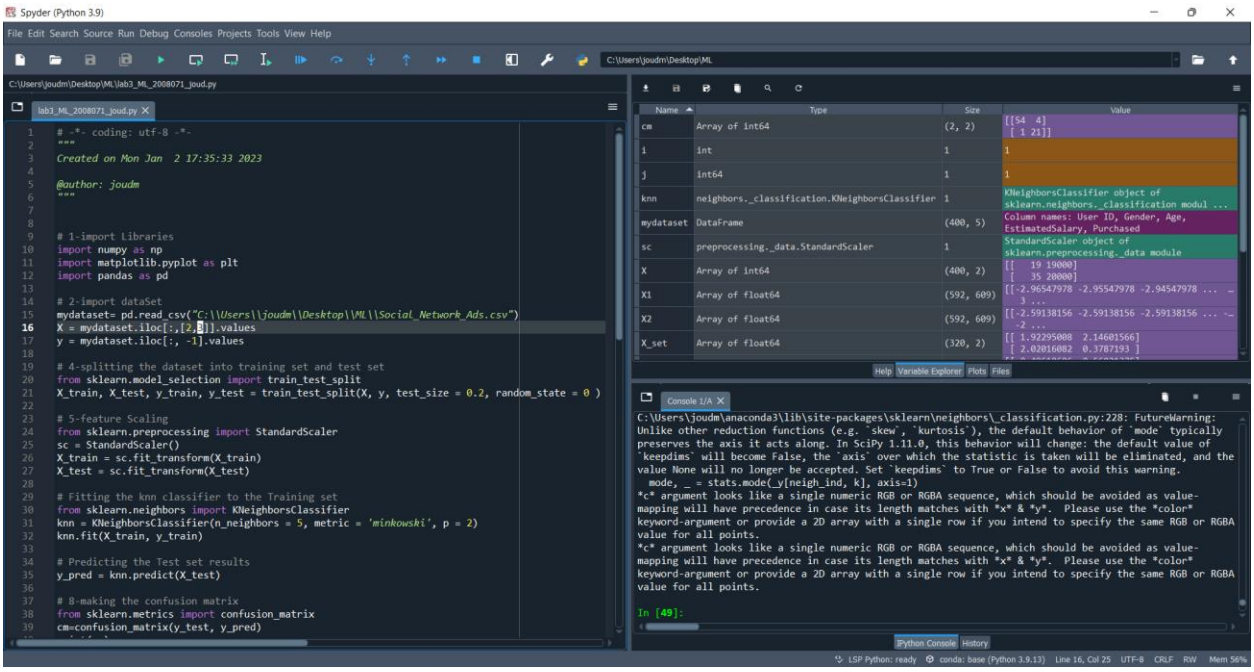
LAB3

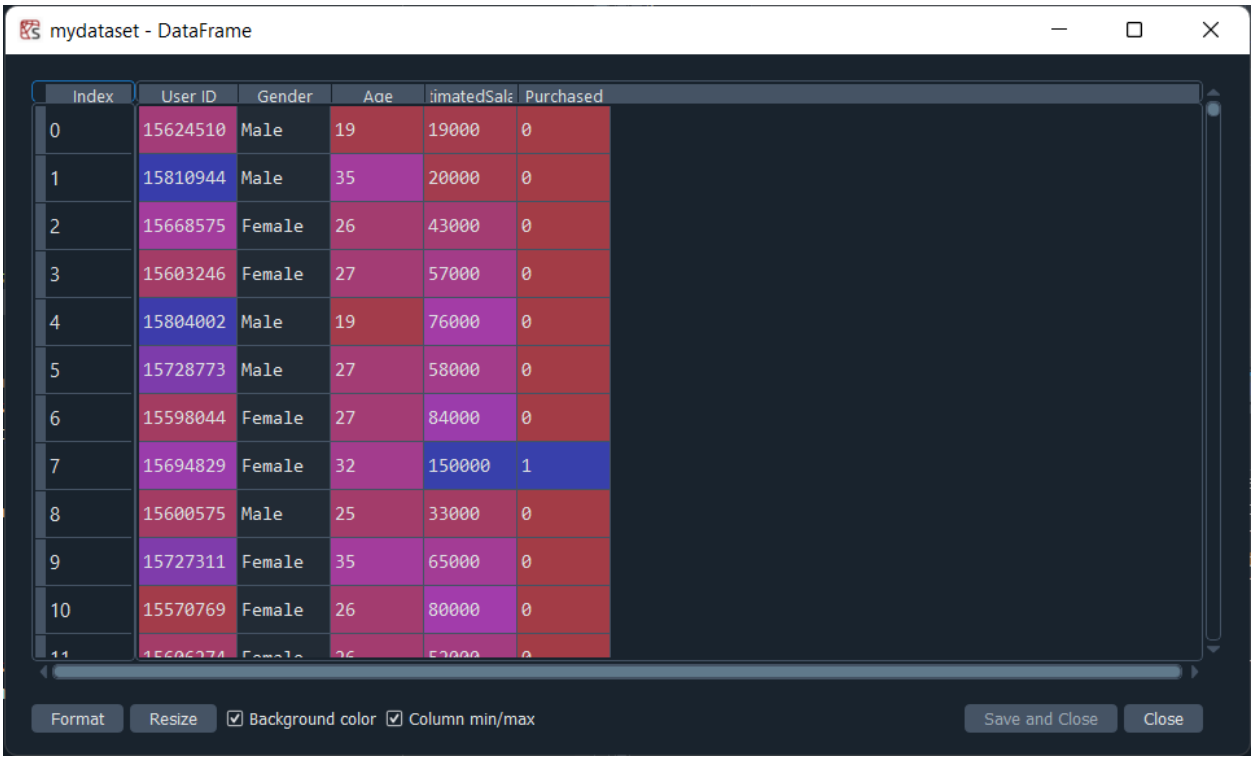
AML

Student Name: Joud Alahmari

Student ID: 2008071

IMPORT THE DATA SET





```
16  X = mydataset.iloc[:,[2,3]].values
17  y = mydataset.iloc[:, -1].values
18
```

X - NumPy object array

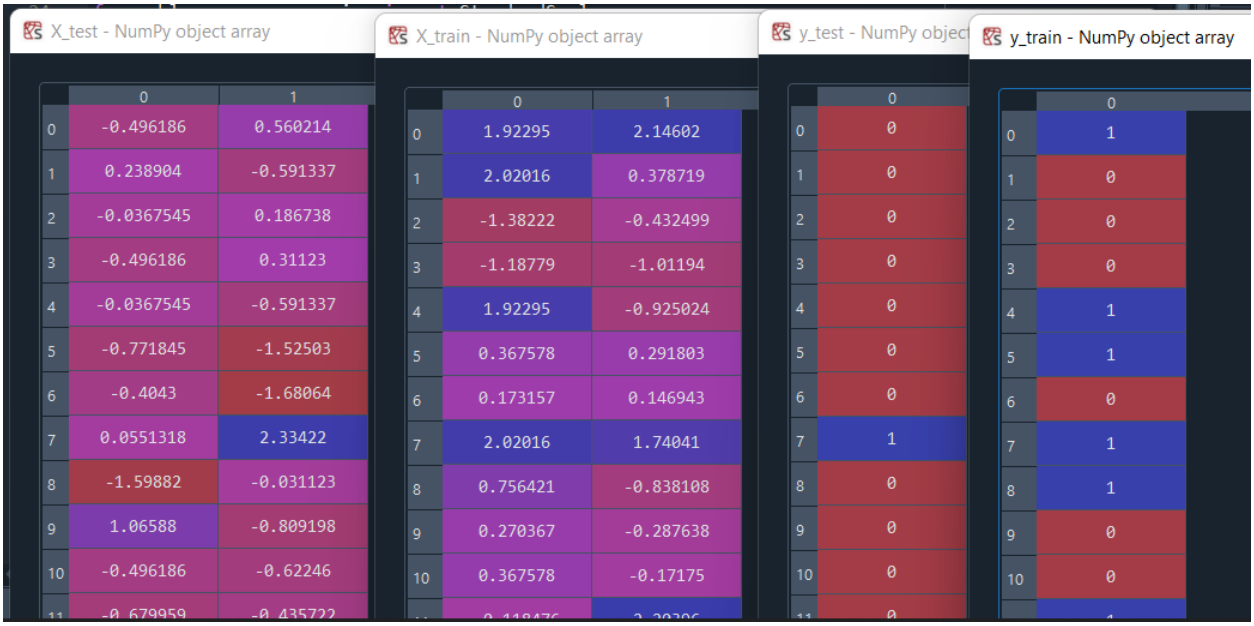
	0	1
0	19	19000
1	35	20000
2	26	43000
3	27	57000
4	19	76000
5	27	58000
6	27	84000
7	32	150000
8	25	33000
9	35	65000
10	26	80000
11	26	52000

y - NumPy object array

	0
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	1
8	0
9	0
10	0
11	0

SPLIT THE DATA SET

```
# 4-splitting the dataset into training set and test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0 )
```



X_test - NumPy object array

	0	1
0	-0.496186	0.560214
1	0.238904	-0.591337
2	-0.0367545	0.186738
3	-0.496186	0.31123
4	-0.0367545	-0.591337
5	-0.771845	-1.52503
6	-0.4043	-1.68064
7	0.0551318	2.33422
8	-1.59882	-0.031123
9	1.06588	-0.809198
10	-0.496186	-0.62246
11	-0.679959	-0.435722

X_train - NumPy object array

	0	1
0	1.92295	2.14602
1	2.02016	0.378719
2	-1.38222	-0.432499
3	-1.18779	-1.01194
4	1.92295	-0.925024
5	0.367578	0.291803
6	0.173157	0.146943
7	2.02016	1.74041
8	0.756421	-0.838108
9	0.270367	-0.287638
10	0.367578	-0.17175

y_test - NumPy object

	0
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	1
8	0
9	0
10	0
11	0

y_train - NumPy object array

	0
0	1
1	0
2	0
3	0
4	1
5	1
6	0
7	1
8	1
9	0
10	0

FEATURE SCALING

```
# 5-feature Scaling
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.fit_transform(X_test)
```

FITTING THE KNN TO THE TRAINING SET

```
# Fitting the knn classifier to the Training set
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p = 2)
knn.fit(X_train, y_train)
```

PREDICT THE TEST SET RESULTS

```
# Predicting the Test set results
y_pred = knn.predict(X_test)
```

y_pred - NumPy object array

	0
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	1
8	0
9	1
10	0
11	0

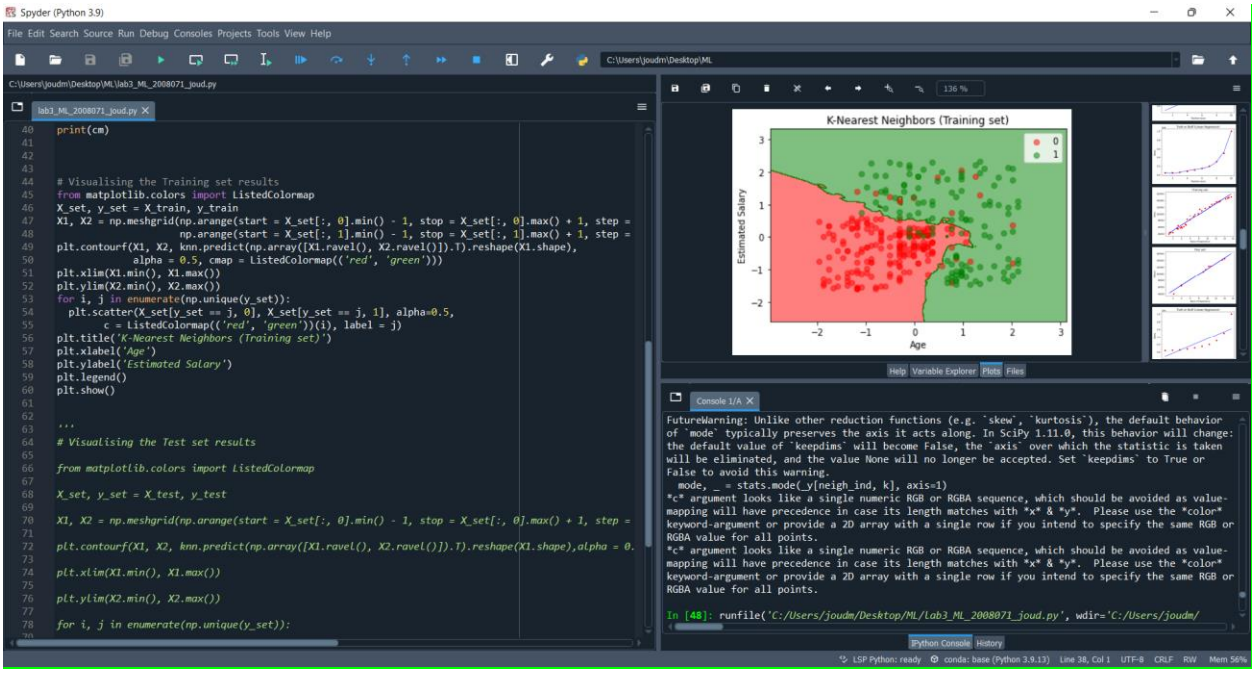
MAKE THE CONFUSION MATRIX

```
# 8-making the confusion matrix
from sklearn.metrics import confusion_matrix
cm=confusion_matrix(y_test, y_pred)
print(cm)
```

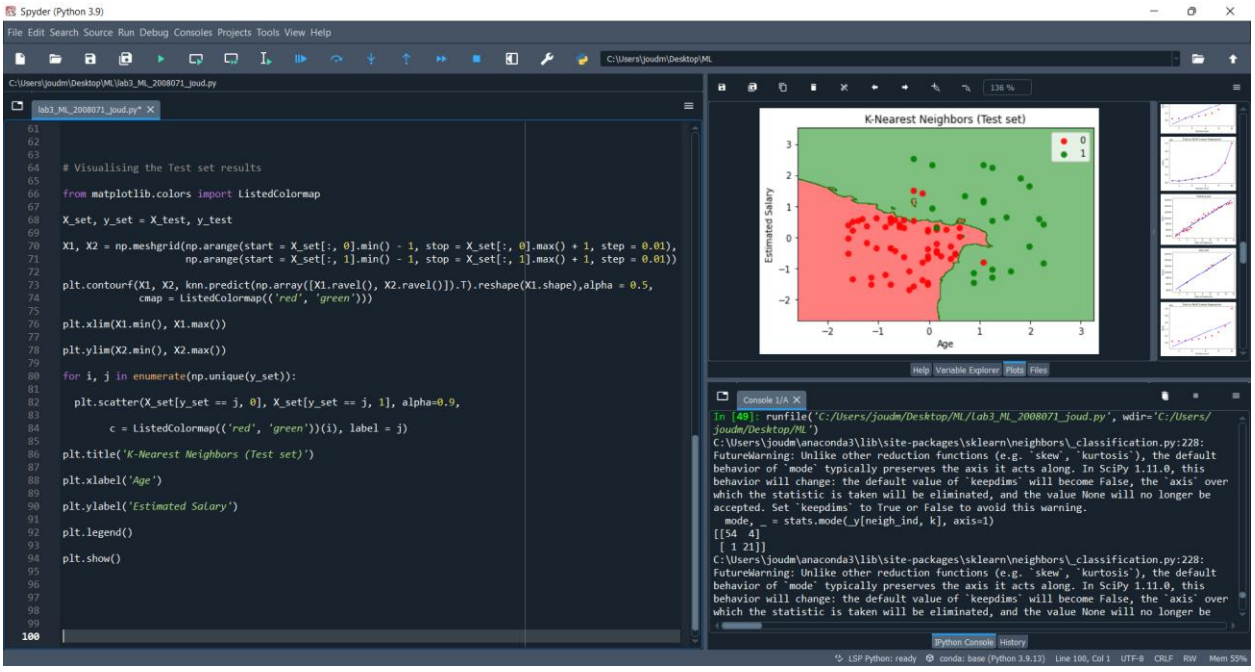
cm - NumPy object array

	0	1
0	54	4
1	1	21

Visualising the Training set results



Visualising the Test set results



9. Write your opinion about the two models' performance.

In the test set model the nodes are less and not crowded as the training set and easy to view and shows the predicted data