Random Forest Classification:

Random Forest Algorithm to classify the cities based on the Geo-code coordinates, Latitudes, Longitudes, Object ID, Feature ID, Zip Code

```
from sklearn.ensemble import RandomForestClassifier,BaggingClassifier
from sklearn.metrics import precision_score,recall_score,f1_score
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split
classifier=RandomForestClassifier(n_estimators=50, random_state=0)
x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size=0.2,random_state = 0)
```

```
scaler = StandardScaler()
x_train = scaler.fit_transform(x_train)
```

```
classifier.fit(x_train,y_train)
```

```
RandomForestClassifier(n_estimators=50, random_state=0)
```

```
y_pred = classifier.predict(x_test)
```

```
accuracy_score(y_pred,y_test)
```

```
0.8303219106957425
```

```
print("precision_score :",precision_score(y_pred,y_test,average='macro'))
print("Recall_score :",recall_score(y_pred,y_test,average = 'macro'))
print("F1_Score :",f1_score(y_pred,y_test,average='macro'))
```

```
precision_score : 0.5938195406377141
Recall_score : 0.5724698589569142
F1_Score : 0.5726849051342382
```

Naïve Bayes Classification:

```
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import cross_val_score
```

```
classifier = GaussianNB()
classifier.fit(x_train, y_train)
pred = classifier.predict(x_test)
scores = cross_val_score(classifier, x_train,y_train, cv=10)
```

```
C:\Users\admin\anaconda3\lib\site-packages\sklearn\model_selection\_split.py:676: UserWarning:
as only 1 members, which is less than n_splits=10.
  warnings.warn(
```

```
np.mean(scores)
```

```
0.378875241062157
```

```
print("precision_score :",precision_score(pred,y_test,average='macro'))
print("Recall_score :",recall_score(pred,y_test,average = 'macro'))
print("F1_Score :",f1_score(pred,y_test,average='macro'))
```

```
precision_score : 0.2269817419555284
Recall_score : 0.15457986404322877
F1_Score : 0.1699799020683929
```

By Comparing Random Forest algorithm and Naïve bayes algorithm, it is observed that Random Forest algorithm is the best suitable classifier based on the accuracy score, precision score, recall score and f1 score.