# Step 1: Load the dataset and split it

In [1]:
```
from sklearn.datasets import load_breast_cancer
from sklearn.model_selection import train_test_split

dataset = load_breast_cancer()
x_train, x_test, y_train, y_test = train_test_split(dataset.data, dataset.targ
print(x_train.shape, y_train.shape)
print(x_test.shape, y_test.shape)
```

```
(426, 30) (426,)
(143, 30) (143,)
```

# Step 2: Find training accuracy and testing accuracy scores for a range of K values

In [2]:
```
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score

train_accuracy = []
test_accuracy = []
krange = range(1, 26)

for k in krange:
    model = KNeighborsClassifier(n_neighbors = k)
    model.fit(x_train, y_train)

    train_accuracy.append(accuracy_score(y_train, model.predict(x_train)))
    test_accuracy.append(accuracy_score(y_test, model.predict(x_test)))

print(train_accuracy)
print(test_accuracy)
```

```
[1.0, 0.960093896713615, 0.9553990610328639, 0.9483568075117371, 0.9413145539
906104, 0.9366197183098591, 0.9366197183098591, 0.9366197183098591, 0.9389671
361502347, 0.9413145539906104, 0.9413145539906104, 0.9413145539906104, 0.9342
723004694836, 0.9389671361502347, 0.9272300469483568, 0.9272300469483568, 0.9
272300469483568, 0.9295774647887324, 0.9295774647887324, 0.9295774647887324,
0.9272300469483568, 0.9272300469483568, 0.9272300469483568, 0.927230046948356
8, 0.9272300469483568]
[0.9090909090909091, 0.9230769230769231, 0.9230769230769231, 0.93706293706293
71, 0.951048951048951, 0.951048951048951, 0.9440559440559441, 0.9440559440559
441, 0.9440559440559441, 0.9440559440559441, 0.9370629370629371, 0.9370629370
629371, 0.9370629370629371, 0.9370629370629371, 0.9440559440559441, 0.9440559
440559441, 0.9300699300699301, 0.9300699300699301, 0.9230769230769231, 0.9230
769230769231, 0.9230769230769231, 0.9230769230769231, 0.9230769230769231, 0.9
230769230769231, 0.9230769230769231]
```

# Step 3: Plot training and testing accuracies against the range of K values and find the optimal value from the plot

In [3]:
```
import matplotlib.pyplot as plt

plt.plot(krange, train_accuracy, color = "green")
plt.plot(krange, test_accuracy, color = "blue")
plt.xlabel("K")
plt.ylabel("Accuracy")
plt.legend(["Training Accuracy", "Testing Accuracy"])
plt.show()
```