In [1]:
```
import pandas as pd
import seaborn as sns
```

In [5]:
```
df = pd.read_csv("diabetes.csv")
df
```

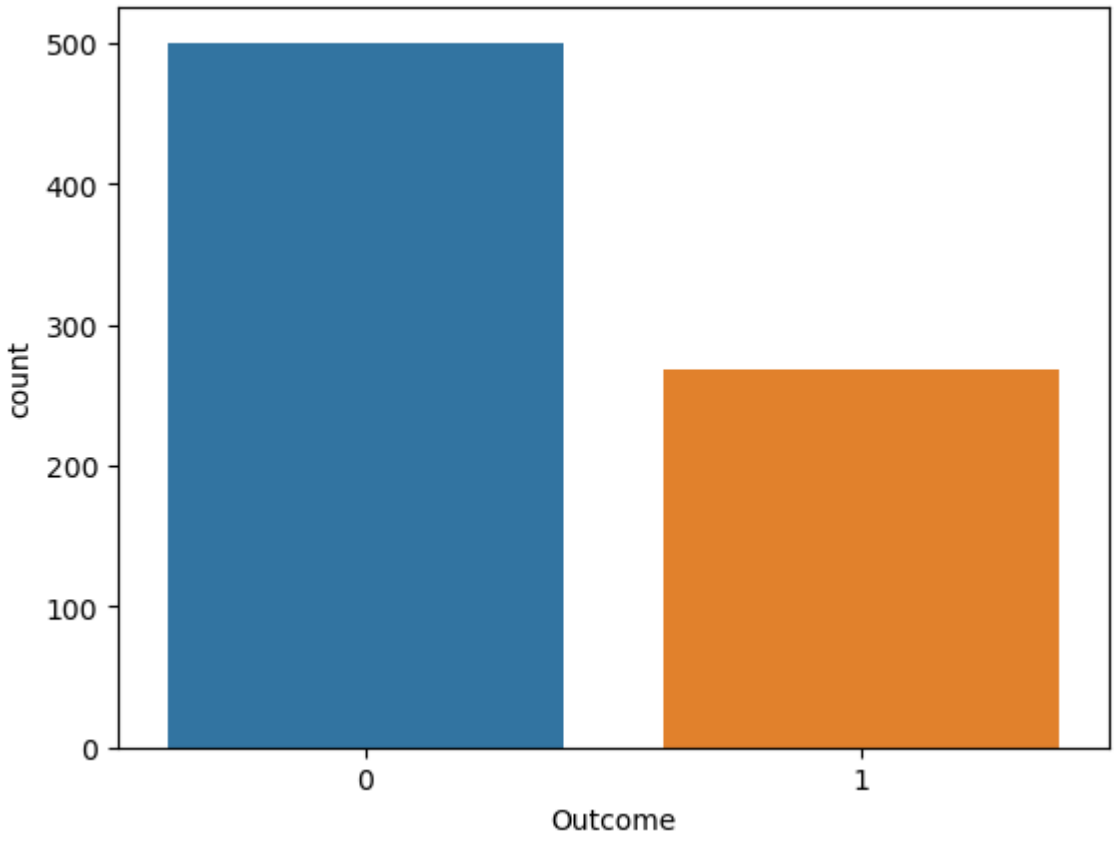
Out[5]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	Pedigree	Age	Out
0	6	148	72	35	0	33.6	0.627	50	
1	1	85	66	29	0	26.6	0.351	31	
2	8	183	64	0	0	23.3	0.672	32	
3	1	89	66	23	94	28.1	0.167	21	
4	0	137	40	35	168	43.1	2.288	33	
...	...	...	...	...	...	...	...	...	
763	10	101	76	48	180	32.9	0.171	63	
764	2	122	70	27	0	36.8	0.340	27	
765	5	121	72	23	112	26.2	0.245	30	
766	1	126	60	0	0	30.1	0.349	47	
767	1	93	70	31	0	30.4	0.315	23	

768 rows × 9 columns

In [7]:
```
x = df.drop('Outcome', axis = 1)
y = df['Outcome']
```

In [8]:
```
1 sns.countplot(x=y);
```



In [9]:
```
1 y.value_counts()
```

Out[9]:
```
0    500
1    268
Name: Outcome, dtype: int64
```

In [11]:
```
1 from sklearn.preprocessing import MinMaxScaler
2 scaler = MinMaxScaler()
3 x_scaled = scaler.fit_transform(x)
```

In [14]:
```
1 from sklearn.model_selection import train_test_split
2 x_train,x_test,y_train,y_test = train_test_split(x_scaled,y,random_stat
```

In [15]:
```
1 x.shape
```

Out[15]: (768, 8)

In [16]:
```
1 x_train.shape
```

Out[16]: (576, 8)

In [17]:
```
1 x_test.shape
```

Out[17]: (192, 8)

In [18]:
```
1 from sklearn.neighbors import KNeighborsClassifier
```

In [19]:
```
1 knn =  KNeighborsClassifier(n_neighbors = 5)
```

In [20]:
```
1 knn.fit(x_train, y_train)
```

Out[20]:
```
KNeighborsClassifier()
```

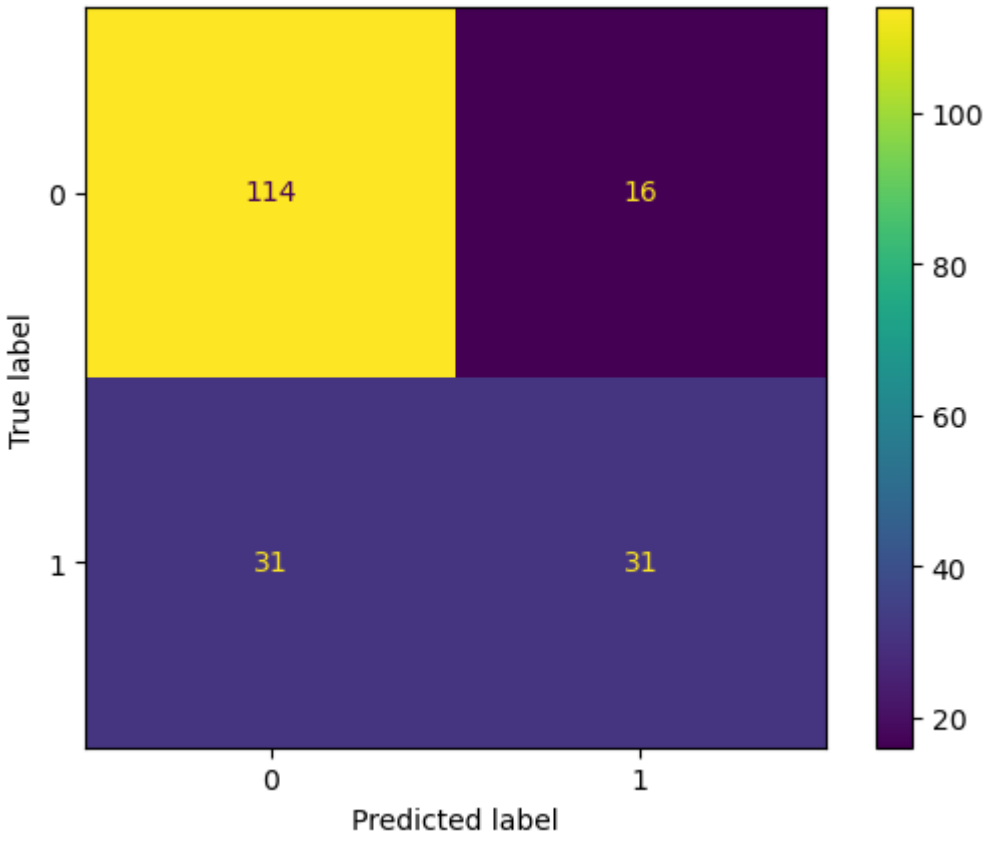
In [23]:
```
1 from sklearn.metrics import accuracy_score , ConfusionMatrixDisplay
2 from sklearn.metrics import classification_report
```

In [24]:
```
1 y_pred = knn.predict(x_test)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\neighbors\_classificati
on.py:228: FutureWarning: Unlike other reduction functions (e.g. `skew`, `
kurtosis`), the default behavior of `mode` typically preserves the axis it
acts along. In SciPy 1.11.0, this behavior will change: the default value
of `keepdims` will become False, the `axis` over which the statistic is ta
ken will be eliminated, and the value None will no longer be accepted. Set
`keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
```

In [26]:
```
1  ConfusionMatrixDisplay.from_predictions(y_test,y_pred)
```

Out[26]:
```
<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x24c112
54f10>
```



In [27]:
```
print(classification_report(y_test,y_pred))
```

```
              precision    recall  f1-score   support

           0       0.79      0.88      0.83       130
           1       0.66      0.50      0.57        62

    accuracy                           0.76       192
   macro avg       0.72      0.69      0.70       192
weighted avg       0.75      0.76      0.75       192
```

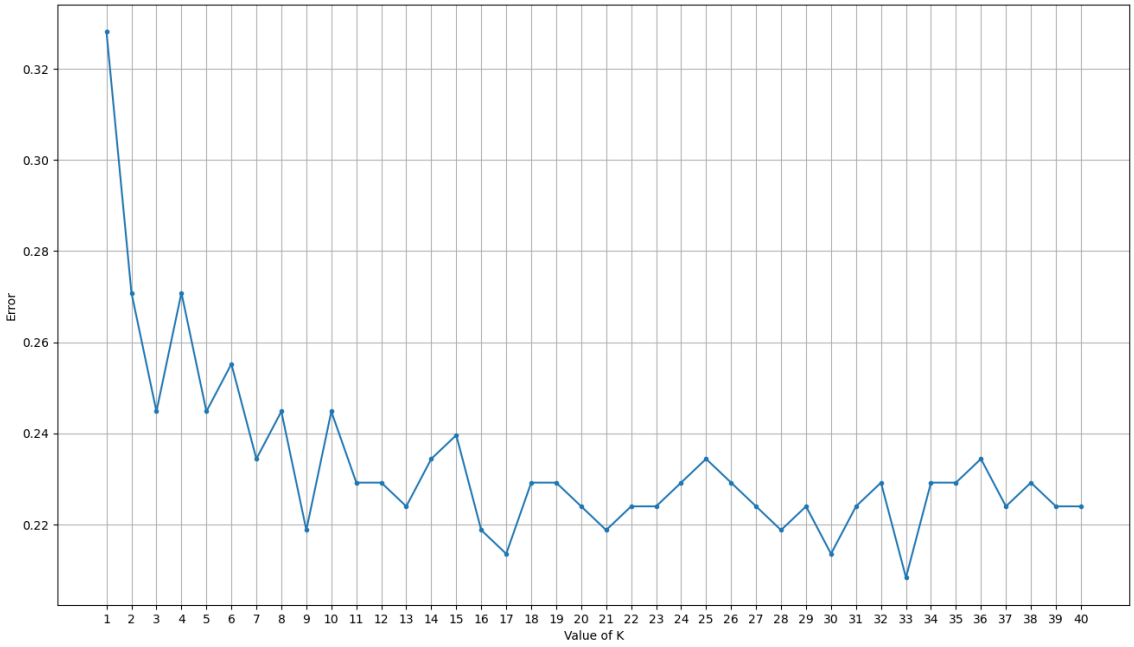
In [33]:
```
import matplotlib.pyplot as plt
import numpy as np
```

In [36]:
```
error = []
for k in range (1,41):
    knn = KNeighborsClassifier(n_neighbors = k)
    knn.fit(x_train, y_train)
    pred=knn.predict(x_test)
    error.append(np.mean(pred!=y_test))
```

```
statistic is taken will be eliminated, and the value None will no longe
r be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\neighbors\_classific
ation.py:228: FutureWarning: Unlike other reduction functions (e.g. `sk
ew`, `kurtosis`), the default behavior of `mode` typically preserves th
e axis it acts along. In SciPy 1.11.0, this behavior will change: the d
efault value of `keepdims` will become False, the `axis` over which the
statistic is taken will be eliminated, and the value None will no longe
r be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\neighbors\_classific
ation.py:228: FutureWarning: Unlike other reduction functions (e.g. `sk
ew`, `kurtosis`), the default behavior of `mode` typically preserves th
e axis it acts along. In SciPy 1.11.0, this behavior will change: the d
efault value of `keepdims` will become False, the `axis` over which the
statistic is taken will be eliminated, and the value None will no longe
r be accepted. Set `keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\neighbors\_classific
```

In [38]:
```
plt.figure(figsize=(16,9))
plt.xlabel('Value of K')
plt.ylabel('Error')
plt.grid()
plt.xticks(range(1,41))
plt.plot(range(1,41),error,marker='.')
```

Out[38]: [<matplotlib.lines.Line2D at 0x24c115ed8e0>]



In [39]:
```
knn = KNeighborsClassifier(n_neighbors = 33)
```

In [40]:
```
knn.fit(x_train, y_train)
```

Out[40]: KNeighborsClassifier(n_neighbors=33)

In [41]:
```
 y_pred=knn.predict(x_test)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\neighbors\_classificati
on.py:228: FutureWarning: Unlike other reduction functions (e.g. `skew`, `
kurtosis`), the default behavior of `mode` typically preserves the axis it
acts along. In SciPy 1.11.0, this behavior will change: the default value
of `keepdims` will become False, the `axis` over which the statistic is ta
ken will be eliminated, and the value None will no longer be accepted. Set
`keepdims` to True or False to avoid this warning.
  mode, _ = stats.mode(_y[neigh_ind, k], axis=1)
```

In [42]:
```
print(classification_report(y_test,y_pred))
```

```
              precision    recall  f1-score   support

           0       0.79      0.94      0.86       130
           1       0.79      0.48      0.60        62

    accuracy                           0.79       192
   macro avg       0.79      0.71      0.73       192
weighted avg       0.79      0.79      0.78       192
```

In [ ]:
```
1
```