In [3]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn import svm
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score
```

In [4]:
```
data = pd.read_csv(r"C:\Users\Anusha V\Documents\fish.dataset.csv")
```

In [7]:
```
print(data.head(20))
```

```
   Species  Weight  Length  Diagonal   Height   Width
0    Bream   242.0    25.4      30.0  11.5200  4.0200
1    Bream   290.0    26.3      31.2  12.4800  4.3056
2    Bream   340.0    26.5      31.1  12.3778  4.6961
3    Bream   363.0    29.0      33.5  12.7300  4.4555
4    Bream   430.0    29.0      34.0  12.4440  5.1340
5    Bream   450.0    29.7      34.7  13.6024  4.9274
6    Bream   500.0    29.7      34.5  14.1795  5.2785
7    Bream   390.0    30.0      35.0  12.6700  4.6900
8    Bream   450.0    30.0      35.1  14.0049  4.8438
9    Bream   500.0    30.7      36.2  14.2266  4.9594
10   Bream   475.0    31.0      36.2  14.2628  5.1042
11   Bream   500.0    31.0      36.2  14.3714  4.8146
12   Bream   500.0    31.5      36.4  13.7592  4.3680
13   Bream   340.0    32.0      37.3  13.9129  5.0728
14   Bream   600.0    32.0      37.2  14.9544  5.1708
15   Bream   600.0    32.0      37.2  15.4380  5.5800
16   Bream   700.0    33.0      38.3  14.8604  5.2854
17   Bream   700.0    33.0      38.5  14.9380  5.1975
18   Bream   610.0    33.5      38.6  15.6330  5.1338
19   Bream   650.0    33.5      38.7  14.4738  5.7276
```

In [15]:
```
X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.2,
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call las
t)
~\AppData\Local\Temp\ipykernel_5672\781037704.py in <module>
----> 1 X_train, X_test, y_train, y_test = train_test_split(X,y, test_size
=0.2, random_state=43)

NameError: name 'X' is not defined
```

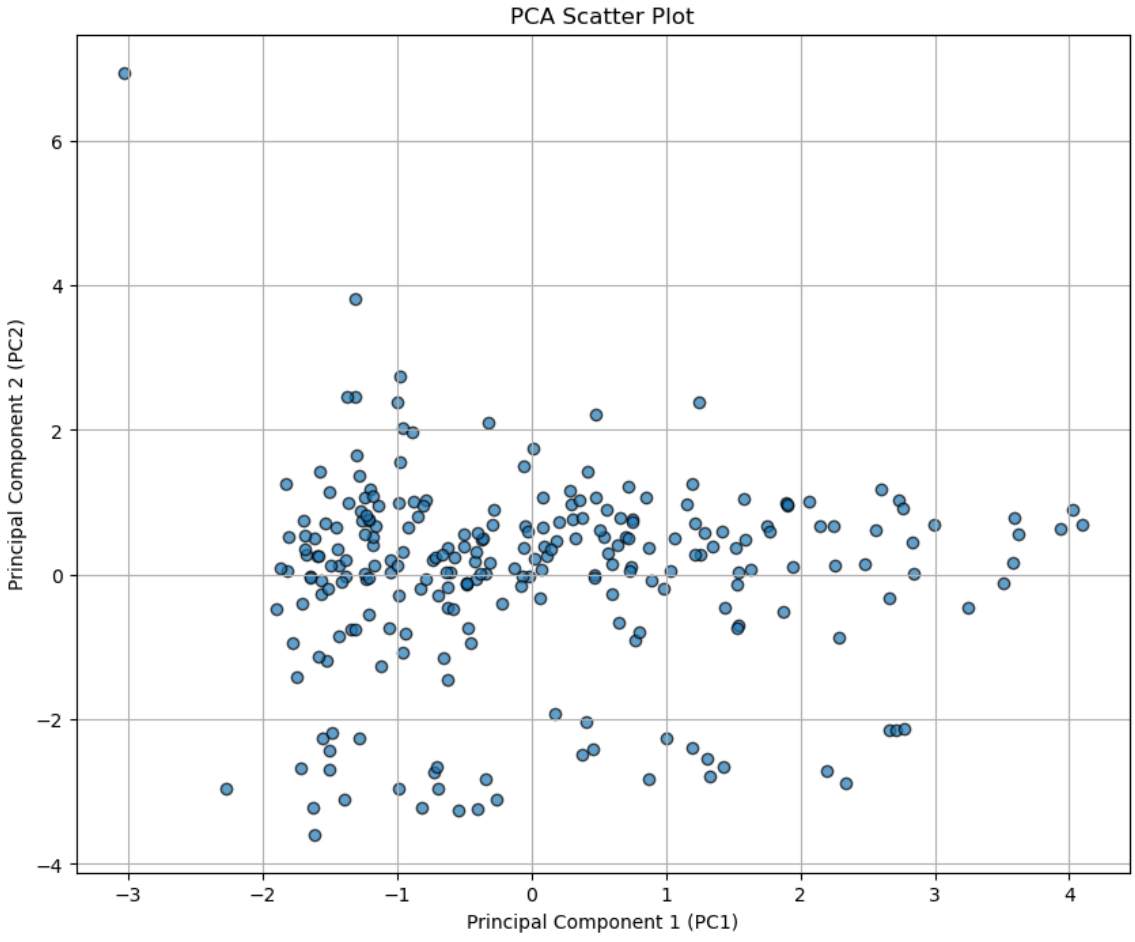
In [5]:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score
data = pd.read_csv(r"C:\Users\Anusha V\Documents\ashaaaa.csv")
print(data.head())
X = data[['nofish', 'livebait', 'camper']]
y = data['persons'] # Target
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)
print(f"Accuracy of the SVM model")
```

```
   nofish  livebait  camper  persons  child        xb        zg  count
0       1         0       0        1      0 -0.896315  3.050405      0
1       0         1       1        1      0 -0.558345  1.746149      0
2       0         1       0        1      0 -0.401731  0.279939      0
3       0         1       1        2      1 -0.956298 -0.601526      0
4       0         1       0        1      0  0.436891  0.527709      1
Accuracy of the SVM model
```

In [6]:

```
import pandas as pd
import numpy as np
from sklearn.decomposition import PCA
import matplotlib.pyplot as plt
data = pd.read_csv(r"C:\Users\Anusha V\Documents\ashaaaa.csv")
X = data.drop('persons', axis=1)
X_standardized = (X - X.mean()) / X.std()
pca = PCA(n_components=2)
X_pca = pca.fit_transform(X_standardized)
pca_df = pd.DataFrame(data=X_pca, columns=['PC1', 'PC2'])
plt.figure(figsize=(10, 8))
plt.scatter(pca_df['PC1'], pca_df['PC2'], edgecolors='k', alpha=0.7)
plt.title('PCA Scatter Plot')
plt.xlabel('Principal Component 1 (PC1)')
plt.ylabel('Principal Component 2 (PC2)')
plt.grid(True)
plt.show()
```



In [ ]:

```
1
```