In [1]:
```
import pandas as pd
data = pd.read_csv('/Users/dhruva/Desktop/Anaconda/CPRG 109_Fish.csv')
print(data)
```

```
    Species  Weight  Length1  Length2  Length3   Height   Width
0     Bream   242.0     23.2     25.4     30.0  11.5200  4.0200
1     Bream   290.0     24.0     26.3     31.2  12.4800  4.3056
2     Bream   340.0     23.9     26.5     31.1  12.3778  4.6961
3     Bream   363.0     26.3     29.0     33.5  12.7300  4.4555
4     Bream   430.0     26.5     29.0     34.0  12.4440  5.1340
..      ...     ...      ...      ...      ...      ...     ...
154   Smelt    12.2     11.5     12.2     13.4   2.0904  1.3936
155   Smelt    13.4     11.7     12.4     13.5   2.4300  1.2690
156   Smelt    12.2     12.1     13.0     13.8   2.2770  1.2558
157   Smelt    19.7     13.2     14.3     15.2   2.8728  2.0672
158   Smelt    19.9     13.8     15.0     16.2   2.9322  1.8792

[159 rows x 7 columns]
```

In [ ]:
```
```

In [32]:
```
data = data[['Weight','Length1','Length2','Length3','Height','Width']]
print(data.mean())
```

```
Weight     398.326415
Length1     26.247170
Length2     28.415723
Length3     31.227044
Height       8.970994
Width        4.417486
dtype: float64
```

In [31]:
```
print(data.std())
```

```
Weight     357.978317
Length1      9.996441
Length2     10.716328
Length3     11.610246
Height       4.286208
Width        1.685804
dtype: float64
/var/folders/kl/sznc2w850f98v6f1gdmc0br00000gn/T/ipykernel_17829/691007357.py:1: FutureWarning: The default value of numeric_only in DataFrame.std is deprecated. In a future version, it will default to False. In addition, specifying 'numeric_only=None' is deprecated. Select only valid columns or specify the value of numeric_only to silence this warning.
  print(data.std())
```

In [18]:
```
data.var()
```

```
/var/folders/kl/sznc2w850f98v6f1gdmc0br00000gn/T/ipykernel_17829/445316826.py:1: FutureWarning: The default value of numeric_only in DataFrame.var is deprecated. In a future version, it will default to False. In addition, specifying 'numeric_only=None' is deprecated. Select only valid columns or specify the value of numeric_only to silence this warning.
  data.var()
```

Out[18]:

```
Weight     128148.475121
Length1        99.928837
Length2       114.839688
Length3       134.797808
Height         18.371576
Width           2.841935
dtype: float64
```

In [19]:

```
data.cov()
```

```
/var/folders/kl/sznc2w850f98v6f1gdmc0br00000gn/T/ipykernel_17829/4020818503.py:1: FutureWarning: The default value of numeric_only in DataFrame.cov is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_only to silence this warning.
  data.cov()
```

Out[19]:

	Weight	Length1	Length2	Length3	Height	Width
Weight	128148.475121	3276.882797	3524.013253	3836.368648	1111.413300	534.990098
Length1	3276.882797	99.928837	107.073431	115.136248	26.795457	14.611556
Length2	3524.013253	107.073431	114.839688	123.685458	29.416988	15.781169
Length3	3836.368648	115.136248	123.685458	134.797808	35.004389	17.194921
Height	1111.413300	26.795457	29.416988	35.004389	18.371576	5.729125
Width	534.990098	14.611556	15.781169	17.194921	5.729125	2.841935

In [20]:

```
data.corr(method='kendall')
```

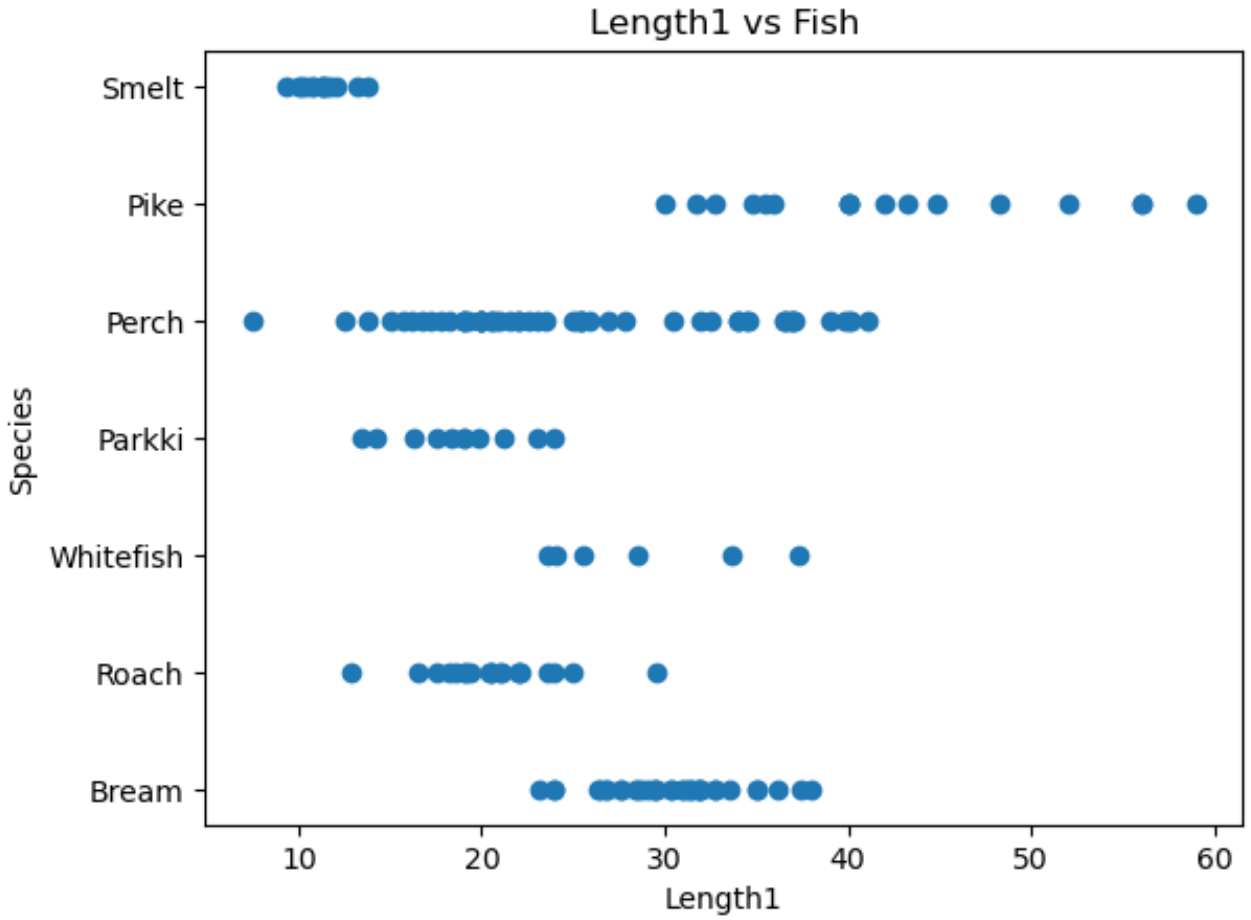
```
/var/folders/kl/sznc2w850f98v6f1gdmc0br00000gn/T/ipykernel_17829/1809062968.py:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_only to silence this warning.
  data.corr(method='kendall')
```

Out[20]:

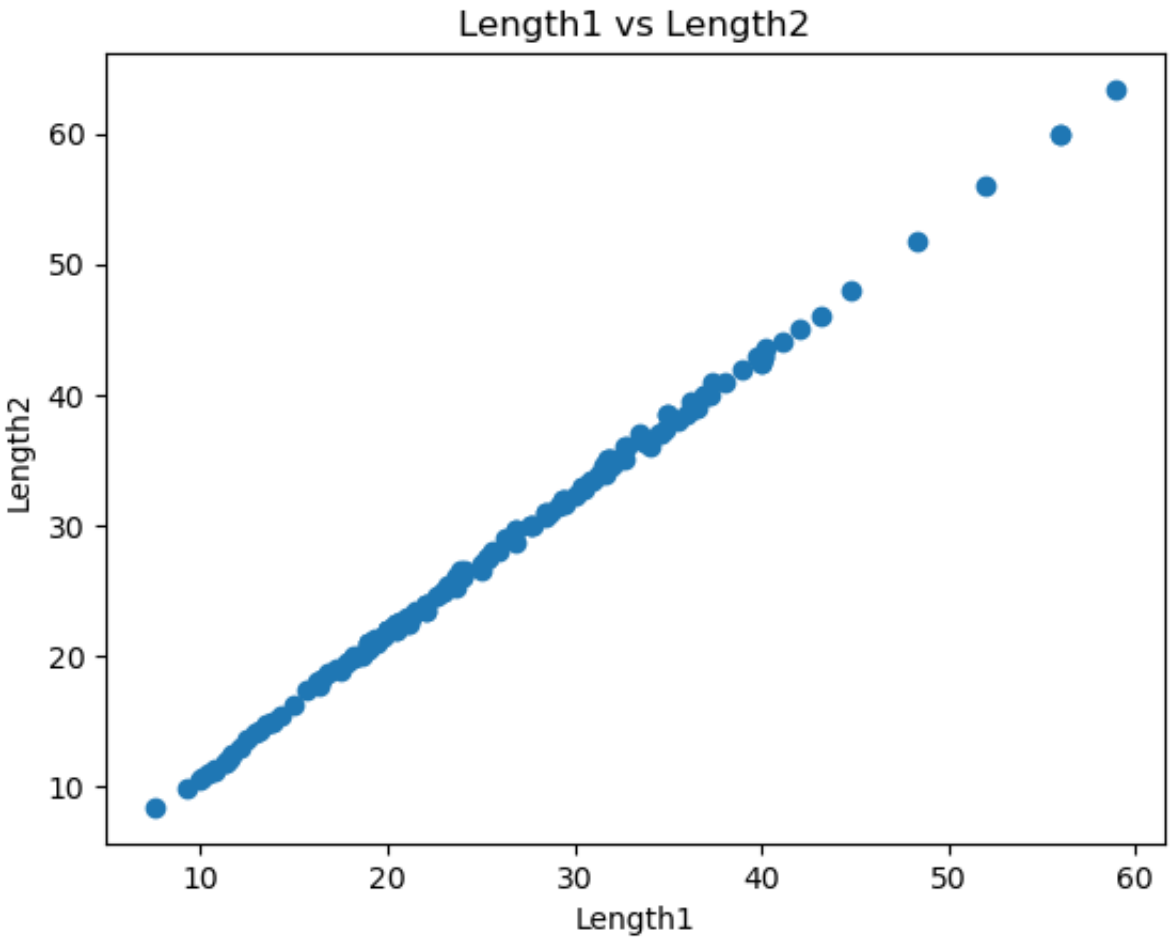
	Weight	Length1	Length2	Length3	Height	Width
Weight	1.000000	0.850766	0.857476	0.867304	0.689884	0.856201
Length1	0.850766	1.000000	0.988813	0.944258	0.573255	0.801216
Length2	0.857476	0.988813	1.000000	0.946828	0.578396	0.805882
Length3	0.867304	0.944258	0.946828	1.000000	0.623375	0.785177
Height	0.689884	0.573255	0.578396	0.623375	1.000000	0.658330
Width	0.856201	0.801216	0.805882	0.785177	0.658330	1.000000

In [43]:
```
from matplotlib import pyplot as plot
data = pd.read_csv('/Users/dhruva/Desktop/Anaconda/CPRG 109_Fish.csv')
plot.title('Length1 vs Fish')
plot.xlabel('Length1')
plot.ylabel('Species')
#Create the scatter plot
plot.scatter(data['Length1'], data['Species'])
#Show scatter plot x-axis = Length1 , y-axis = Length2 plot.show()
```

Out[43]:
```
<matplotlib.collections.PathCollection at 0x14f298520>
```



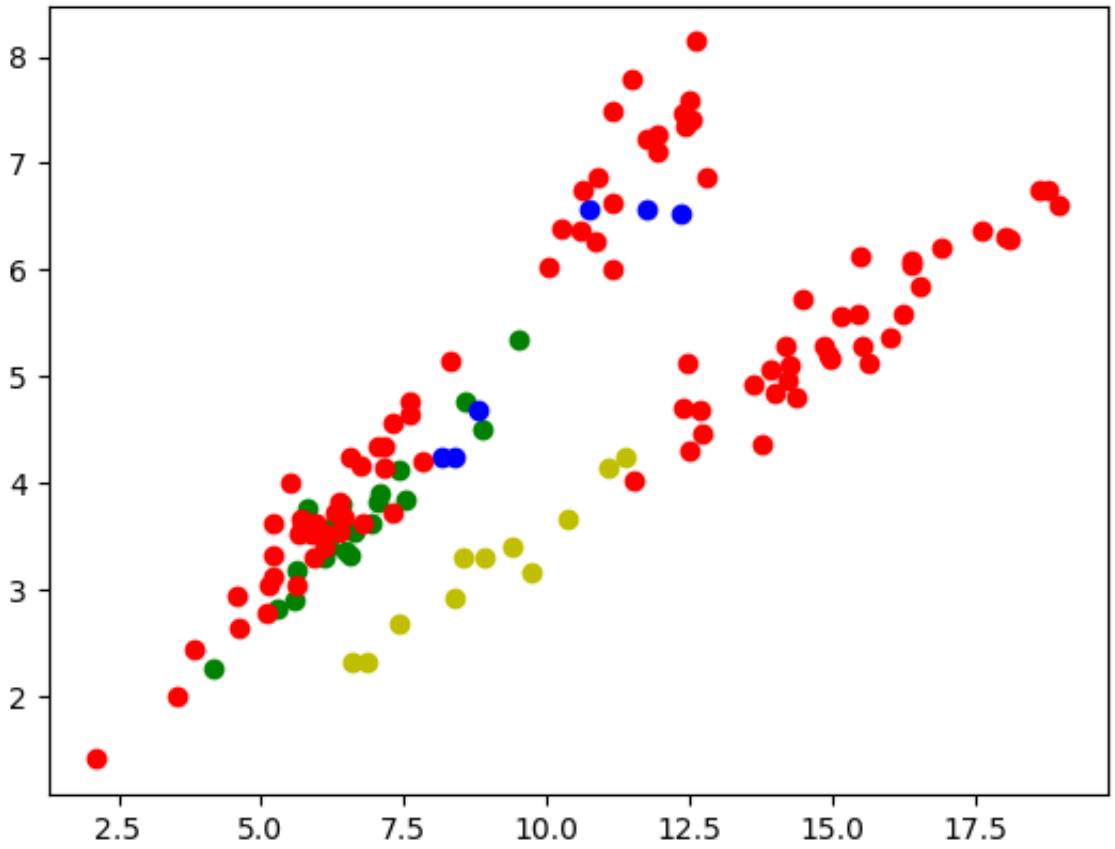
In [36]:
```
from matplotlib import pyplot as plot
plot.title('Length1 vs Length2')
plot.xlabel('Length1')
plot.ylabel('Length2')
#Create the scatter plot
plot.scatter(data['Length1'], data['Length2'])
#Show scatter plot x-axis = Length1 , y-axis = Length2
plot.show()
```



In [27]:
```
from matplotlib import pyplot as plot
data = pd.read_csv('/Users/dhruva/Desktop/Anaconda/CPRG 109_Fish.csv')
colors={'Bream': 'r', 'Roach': 'g','Whitefish':'b', 'Parkki': 'y', 'Perch':'
fig, ax=plot.subplots()
for i in range(len(data['Height'])):
   ax.scatter(data['Height'][i], data['Width'][i], color=colors[data['Specie
#set a title and labels
ax.set_title('Fish Dataset')
ax.set_xlabel('Height')
ax.set_ylabel('Width')
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
Cell In[27], line 6
      4 fig, ax=plot.subplots()
      5 for i in range(len(data['Height'])):
----> 6     ax.scatter(data['Height'][i], data['Width'][i], color=colors[data
['Species'][i]])
      7 #set a title and labels
      8 ax.set_title('Fish Dataset')

KeyError: 'Pike'
```



In [ ]:

In [ ]: