Chapter 5 - Outlier Analysis

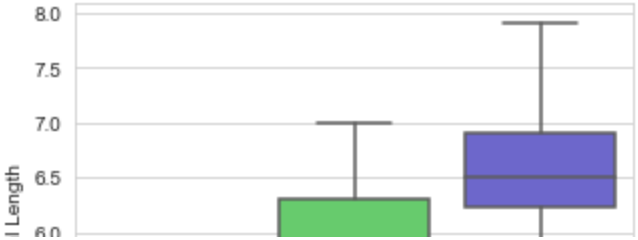
Segment 9 - Multivariate analysis for outlier detection

```
import pandas as pd

import matplotlib.pyplot as plt
from pylab import rcParams
import seaborn as sb


%matplotlib inline
rcParams['figure.figsize'] = 5, 4
sb.set_style('whitegrid')
```

Visually inspecting boxplots

```
df = pd.read_csv(filepath_or_buffer='C:/Users/Lillian/Desktop/ExerciseFiles/Data/iris.data.csv', header=None, sep=',')

df.columns=['Sepal Length','Sepal Width','Petal Length','Petal Width', 'Species']


data = df.iloc[:,0:4].values
target = df.iloc[:,4].values

df[:5]

sb.boxplot(x='Species', y='Sepal Length', data=df, palette='hls')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b1ffc4ab00>
```



Looking at the scatterplot matrix

```
sb.pairplot(df, hue='Species', palette='hls')
```

```
<seaborn.axisgrid.PairGrid at 0x1b1ffd0ec88>
```