Que: The following data are numbers of passengers on flights of Delta Air Lines between San Francisco and Seattle over 33 days in April and early May. 128, 121, 134, 136, 136, 118, 123, 109, 120, 116, 125, 128, 121, 129, 130, 131, 127, 119, 114, 134, 110, 136, 134, 125, 128, 123, 128, 133, 132, 136, 134, 129, 132 Find the Mean, Median, and Mode of this data set. Plot the distribution with mean, median, mode overlayed.

In [20]:
```
NumberofPassengersofDeltaAirlines = [128, 121, 134, 136, 136, 118, 123, 109
132]

import numpy as np

#NumberofPassengersofDeltaAirlines = np.array(NumberofPassengersofDeltaAirl
```

Importing Matplotlib and Numpy and statistics

In [4]:
```
import numpy as np

np.__version__
```

Out [4]: '2.0.2'

In [6]:
```
import matplotlib.pyplot as plt
matplotlib.__version__
```

```
Matplotlib is building the font cache; this may take a moment.
```

Out [6]: '3.9.2'

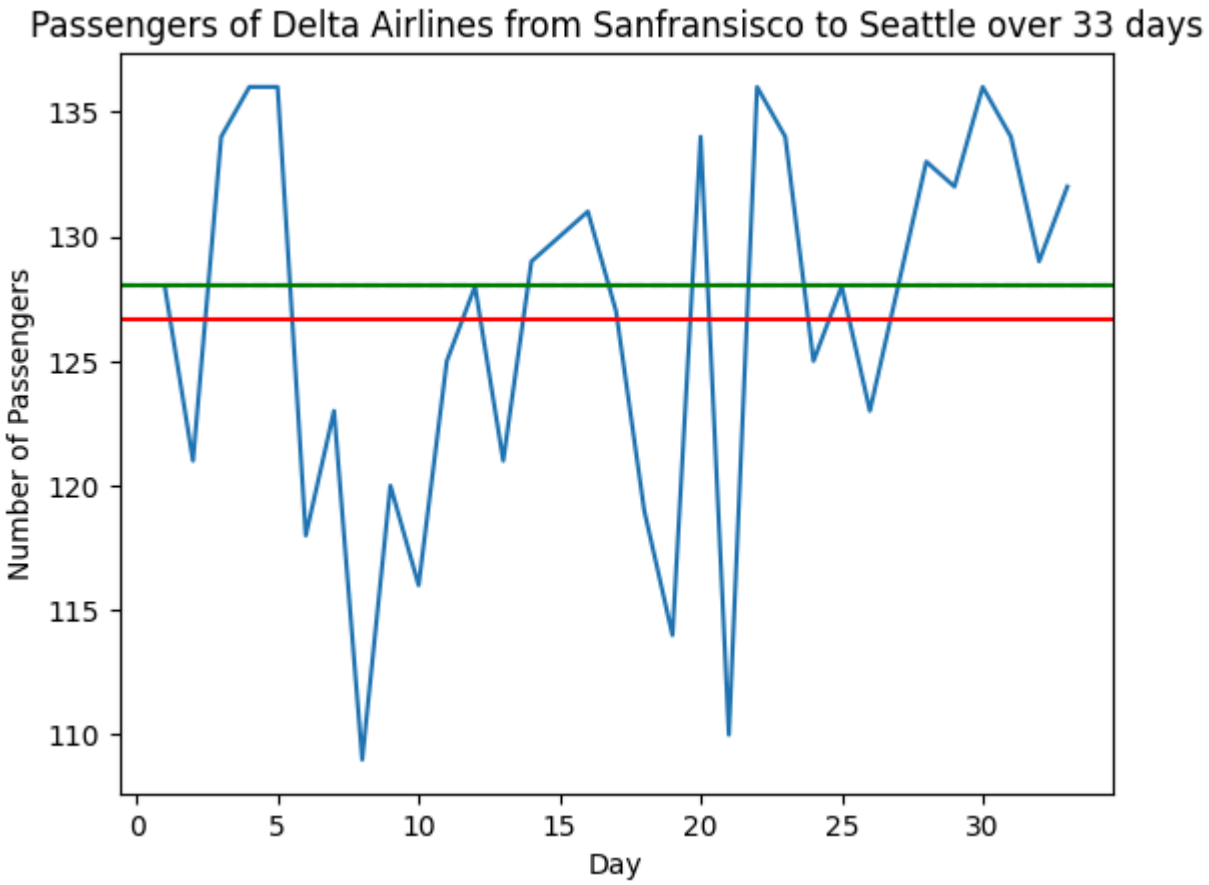
In [31]:
```
from   statistics import median,mean
from statistics import mode
mean_value = mean(NumberofPassengersofDeltaAirlines)
median_value = median(NumberofPassengersofDeltaAirlines)
mode_value = mode(NumberofPassengersofDeltaAirlines)
```

In [41]:
```
print("Mean " ,mean_value)
print("Mode",mode_value)
print("Median",median_value)
```

```
Mean  126.63636363636364
Mode 128
Median 128
```

In [61]:
```
plt.plot(range(1,34,1),NumberofPassengersofDeltaAirlines)
plt.axhline(mean_value, color = 'r')
plt.axhline(median_value,linestyle= "--",color = 'b')
plt.axhline(mode_value,color = 'g')
plt.title("Passengers of Delta Airlines from Sanfransisco to Seattle over 3
plt.xlabel("Day")
plt.ylabel("Number of Passengers")
```

```
plt.legend()
plt.show()
```



The following data are numbers of passengers on flights of Delta Air Lines between San Francisco and Seattle over 33 days in April and early May. 128, 121, 134, 136, 136, 118, 123, 109, 120, 116, 125, 128, 121, 129, 130, 131, 127, 119, 114, 134, 110, 136, 134, 125, 128, 123, 128, 133, 132, 136, 134, 129, 132

Find the median, the first and third quartile and hence find IQR. What would you like to say about the spread (dispersion) of dataset. Plot a box plot to display 5-summary points graphically. Comment on the symmetry (skweness) of dataset.

In [76]:
```
from numpy import percentile


NumberofPassengersofDeltaAirlines = np.array(NumberofPassengersofDeltaAirli

quartiles = percentile(NumberofPassengersofDeltaAirlines,[25,50,75])
print("1st quartiles",quartiles[0])
print("2nd quartiles",quartiles[1])
print("3rd quartiles",quartiles[2])

IQR = quartiles[0]- quartiles[2]


print("IQR  ", abs(IQR))
```

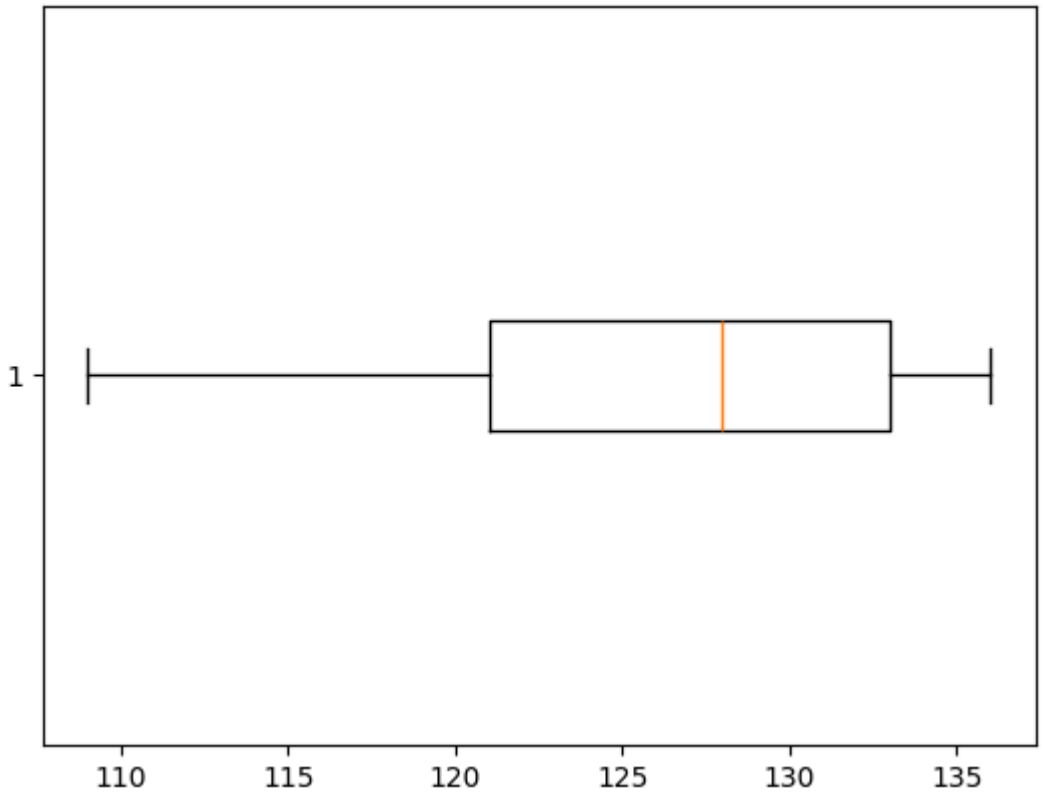
```
1st quartiles 121.0
2nd quartiles 128.0
3rd quartiles 133.0
IQR   12.0
```

In [82]:
```
Variance = NumberofPassengersofDeltaAirlines.var()

print("Variance - ", Variance)
```

```
Variance -  55.988980716253444
```

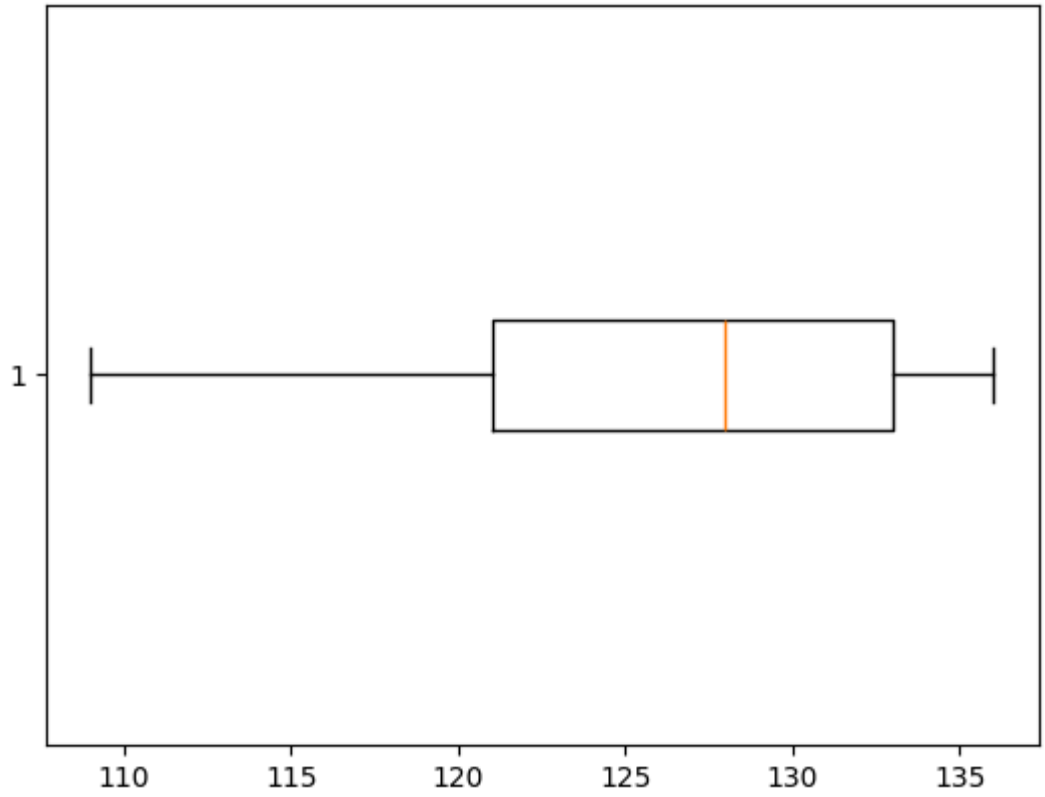
In [104]:
```
plt.boxplot(NumberofPassengersofDeltaAirlines,vert=0)
plt.show()
```

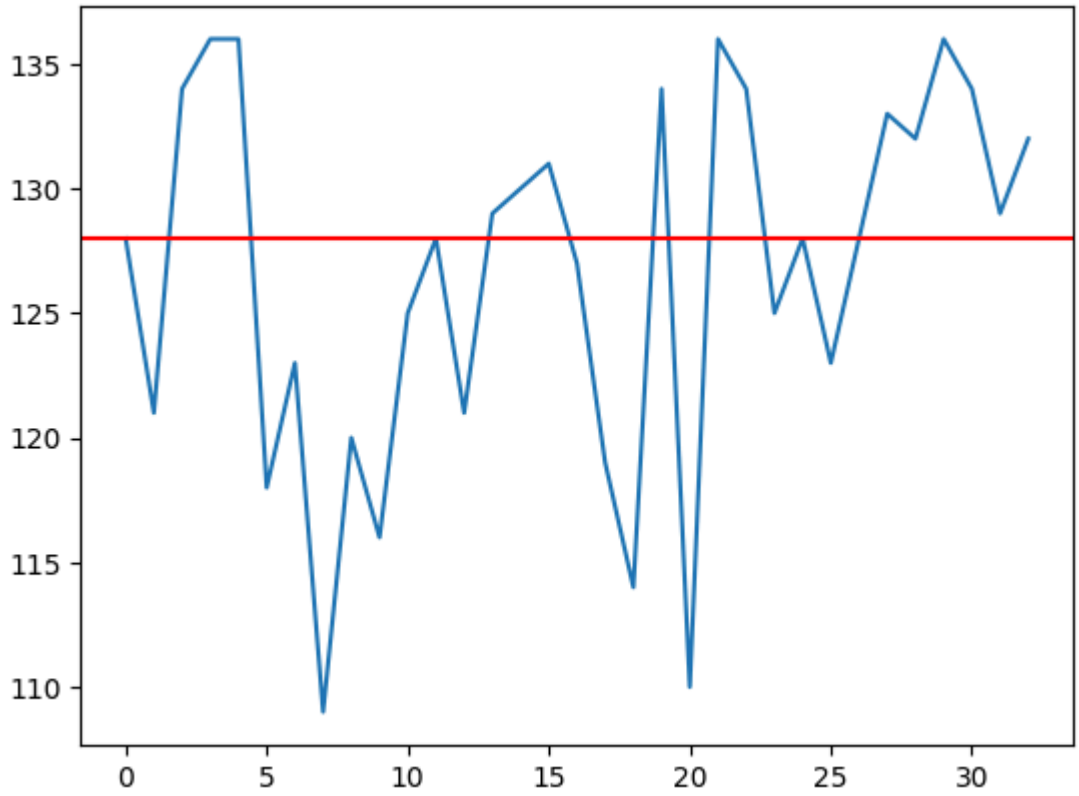


In [103]:
```
plt.subplots(1,1)
plt.boxplot(NumberofPassengersofDeltaAirlines,vert=0)
plt.subplots(1,1)
plt.plot(NumberofPassengersofDeltaAirlines)
plt.axhline(median_value, color = 'r')



plt.show()
```





Looking at the Plots we can see that the data is skewed having most of the passengers in 33 days were lesser than the median value. This says the data is skewed(Left Skewed)