```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.cluster import AgglomerativeClustering
from scipy.cluster import hierarchy

data = pd.read_csv('penguins.csv')
data.shape
```

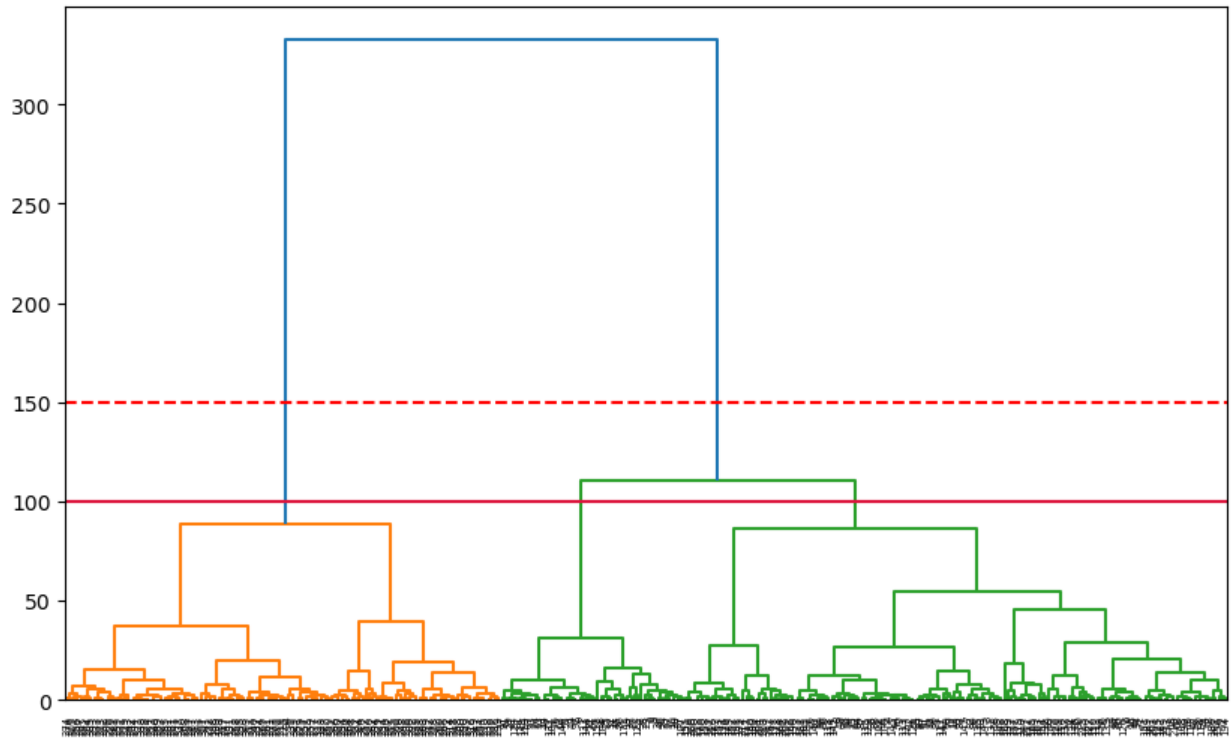
```
(344, 7)
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 344 entries, 0 to 343
Data columns (total 7 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   species            344 non-null    object 
 1   island             344 non-null    object 
 2   bill_length_mm     342 non-null    float64
 3   bill_depth_mm      342 non-null    float64
 4   flipper_length_mm  342 non-null    float64
 5   body_mass_g        342 non-null    float64
 6   sex                334 non-null    object 
dtypes: float64(4), object(3)
memory usage: 18.9+ KB
```

```
df = data[['bill_length_mm', 'flipper_length_mm']]
df = df.dropna(axis=0)
clusters = hierarchy.linkage(df, method='ward')
plt.figure(figsize=(10, 6))
dendrograms = hierarchy.dendrogram(clusters)
plt.axhline(150, color='red', linestyle='--')
plt.axhline(100, color='crimson')
```

```
<matplotlib.lines.Line2D at 0x1733d4f19a0>
```




```
clustering_model = AgglomerativeClustering(n_clusters=3,
linkage='ward')
clustering_model.fit(df)
clustering_model.labels_
fig, axes = plt.subplots(nrows=1, ncols=2, figsize=(15, 5))
sns.scatterplot(ax=axes[0], data=df, x='bill_length_mm',
y='flipper_length_mm').set_title("Without Clustering")
sns.scatterplot(ax=axes[1], data=df, x='bill_length_mm',
y='flipper_length_mm', hue=clustering_model.labels_).set_title("With
Clustering")
```

```
Text(0.5, 1.0, 'With Clustering')
```