Exercise 14: Intermediate clusterings - how many clusters?

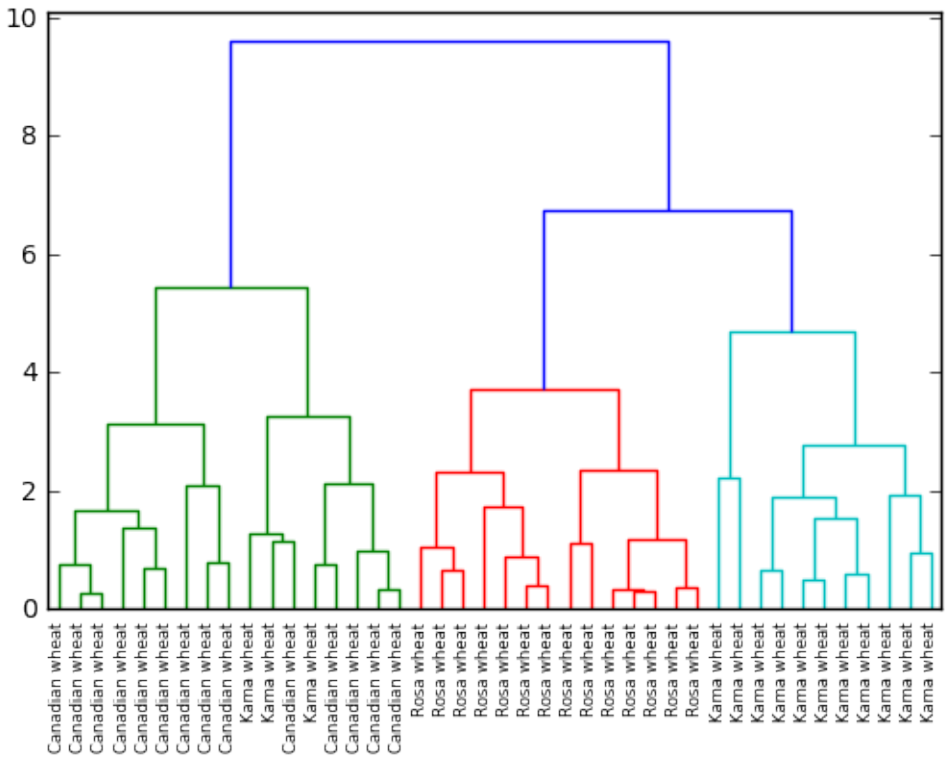
Consider the dendrogram below - it is the result of your hierarchical clustering of some of the grain samples.

Question: If the hierarchical clustering were stopped at height 6 on the dendrogram, how many clusters would there be?

Hint: Imagine a horizontal line at this height.

From the course *Transition to Data Science*. Buy the entire course for just $10 for many more exercises and helpful video lectures.

In [2]:
```
import pandas as pd

seeds_df = pd.read_csv('../datasets/seeds-less-rows.csv')

# remove the grain species from the DataFrame, save for later
varieties = list(seeds_df.pop('grain_variety'))

# extract the measurements as a NumPy array
samples = seeds_df.values

from scipy.cluster.hierarchy import linkage, dendrogram
import matplotlib.pyplot as plt

mergings = linkage(samples, method='complete')

dendrogram(mergings,
           labels=varieties,
           leaf_rotation=90,
           leaf_font_size=6,
)
plt.show()
```

Answer: 3

In [ ]:

