Lesson 21 (Clustering Seeds) Useful commands are listed below

```
colors = np.array(['blue','red','green','cyan','magenta','yellow','black'])
marker_size = 100
marker_alpha = 0.5

df1['clusters'] = colors[labels]

df1.plot.scatter(x='asymmetry',y='grove',c=df1.clusters.values,s=marker_size,alpha=marker_alpha)
```

(a) Download `seeds_preprocessed.csv` from Moodle. This file contains attributes for several varieties of wheat seeds. Plot a scatter plot of the `asymmetry` vs `grove` attributes.

(b) Use the `asymmetry` and `grove` attributes to cluster the seeds in two clusters using the K-Means algorithm. How good are the resulting clusters?

(c) Try to improve the clustering by increasing the number of K-Means runs (`n_init=...`).

(d) Try to improve the clustering by increasing the number of clusters.

(e) Try to improve the clustering by standardizing the attributes first.