Appendix D

Download the Jupyter Notebook file called "Kmeans_FromCSV.ipynb" here:
https://github.com/charraden/Mineralogy-by-K-Means

Python package prerequisites (user to install): pandas, scikit-learn, matplotlib, tkinter, seaborn

This Jupyter notebook contains Python code to run the k-means clustering method to estimate mineralogy from µXRF elemental data. The code first reads in a table of µXRF X-ray counts and uses these as inputs to the k-means analysis. The code first executes the k-means cluster algorithm and provides the user with a plot showing the number of clusters (k) plotted against the sum of squared distances and delta values and prompts the user to select the number of clusters. The user then uses this plot to determine the appropriate number of clusters from the dataset, enters the number of clusters, and the algorithm then assigns cluster numbers to the data, appends this information to the X-ray count table, and exports a new file containing the clustering information. The tabular data can be viewed either in external software (e.g. QGIS, or ioGAS) or the user can add additional Python code to plot the tabular data as an image within the Python platform.

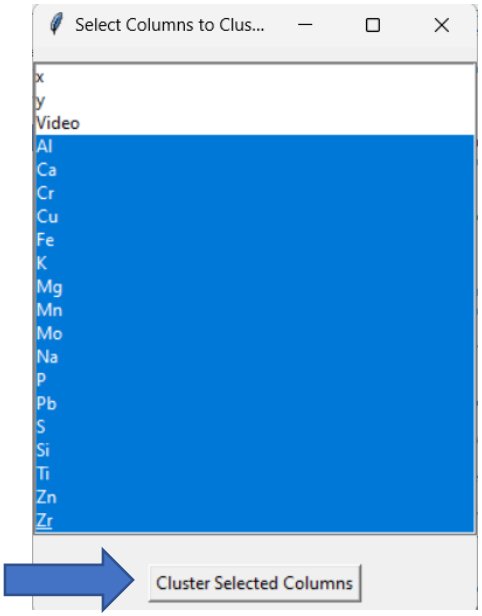
The code requires the elemental data comes in the same format as is exported from the code in Appendix A, with the element column headers containing the element abbreviation only. The instructions below outline how to use the k-means Jupyter Notebook:

1) Save or copy the .csv file containing the X-ray count data into the same folder as the Jupyter Notebook.
2) Add the incoming file path (ensuring that any "\" are changed to "/" – this is a Python requirement) in line 12. This will also be the folder that the output file is saved into.
3) Add the incoming file name in line 13, making sure to put it in quotes and adding the ".csv" file extension. Add the input output file name in line 14, making sure to put it in quotes and adding the ".csv" file extension.

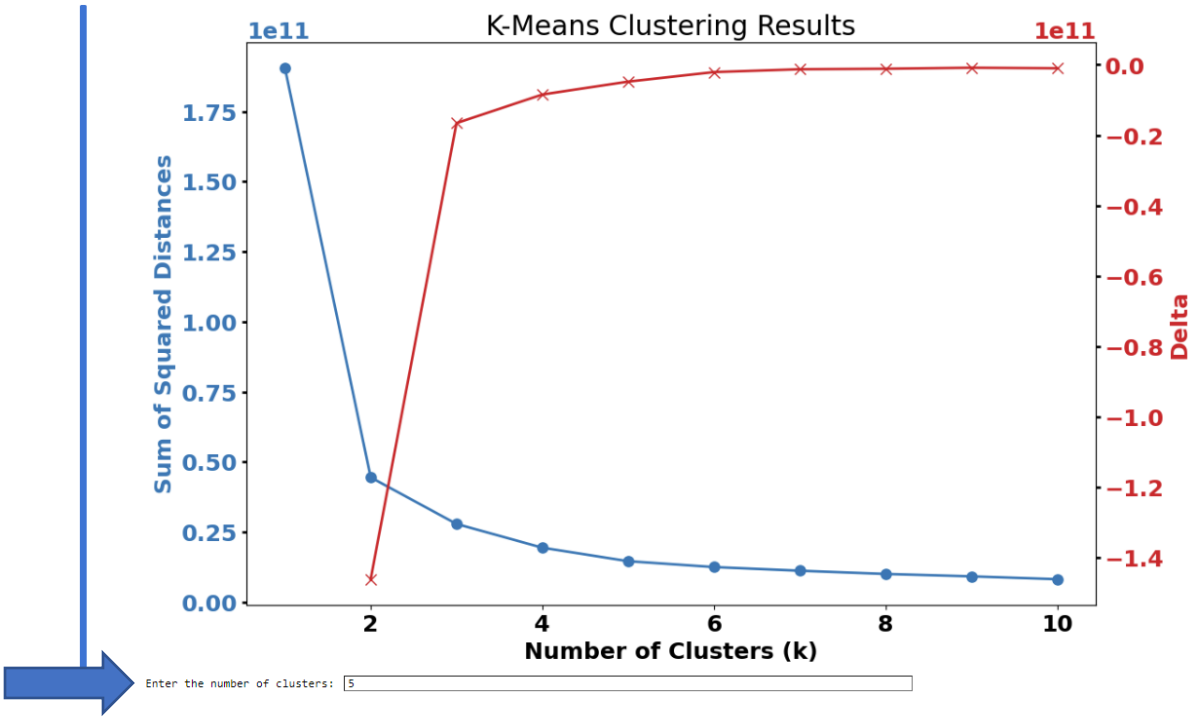
```
10
11 # Define file paths
12 folder_path = 'C:/Users/charraden/OneDrive - UBC/BCPorphyryProject/CalculatedMineralogy/LP_Paper_24April2024/CodeDistribution_AppendixD'
13 importfile_name = '19FB-131_Counts.csv'
14 exportfile_name = '19FB-131_Counts_KClusters.csv'
15
```

4) Click inside the code box once, then hold down the SHIFT key, and hit ENTER on your keyboard to run the code.

5) A picklist of columns contained in the incoming table will appear. Select the elements to include in the clustering, and click "Cluster selected columns".



6) When the clustering is finished, the statistical plot will appear showing the sum of squared distances (SSD) and delta values for various k-clusters. Select the appropriate number of clusters, type that value into the prompt box, and hit ENTER on the keyboard.



7) When the code is finished, the .csv file containing the cluster assignments appended to the X-ray count table will appear in the same folder as Jupyter Notebook.