Kmeans Clustering of Plutonium (PU) Isotope Composition

Carol Chiu

February 20, 2020

General Assembly

Data Science Immersive

Summer 2019

Who is the Client?

- A power generator company named RadioacX receives an assignment from NASA and SpaceX in which their recent rover project to Mars needs a new energy supply.

- RadioacX currently produces the most powerful energy generators using a chemical compound called plutonium (PU).

What is the problem?

-The problem is the device requires a large amount of PU-239 and RadioacX is quickly running out of PU-239.

-However, they have a large supply of Plutonium including other isotopes not currently being used.

-They need to determine if there is any other isotope of PU is similar to PU-239 that could be used as a fuel instead.

https://www.radioactivity.eu.com/site/pages/Plutonium_Use.htm

What is Kmeans clustering?

- **K-means clustering** is a type of unsupervised learning

- It is used for unlabeled data or data without **defined** categories or groups.

- The goal of this algorithm is to find groups in the data

- The number of groups is represented by the variable **K**

- https://blogs.oracle.com/datascience/introduction-to-k-means-clustering

Convert data to a pandas dataframe

	Unnamed: 0	Pu238	Pu239	Pu240	Pu241
0	1	0.126	75.804	21.204	2.180
1	2	0.133	75.515	21.408	2.240
2	3	0.127	75.175	21.668	2.305
3	4	0.156	78.872	18.428	1.906
4	5	0.503	73.317	20.223	4.128

Create a Scatterplot
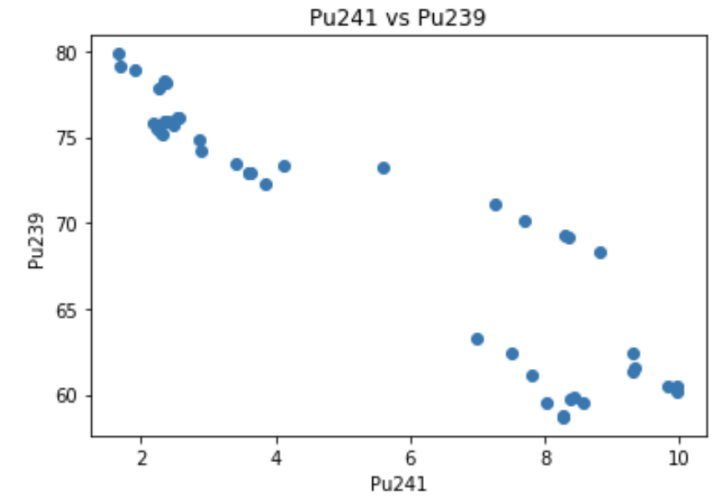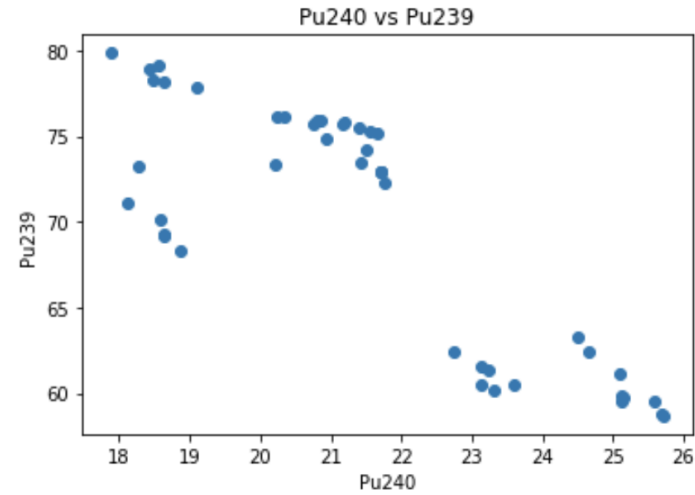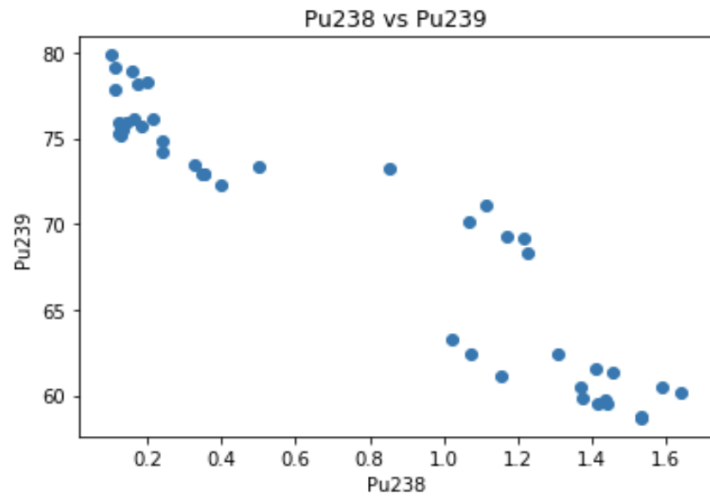
See Clustering – Pu238 VS Pu239
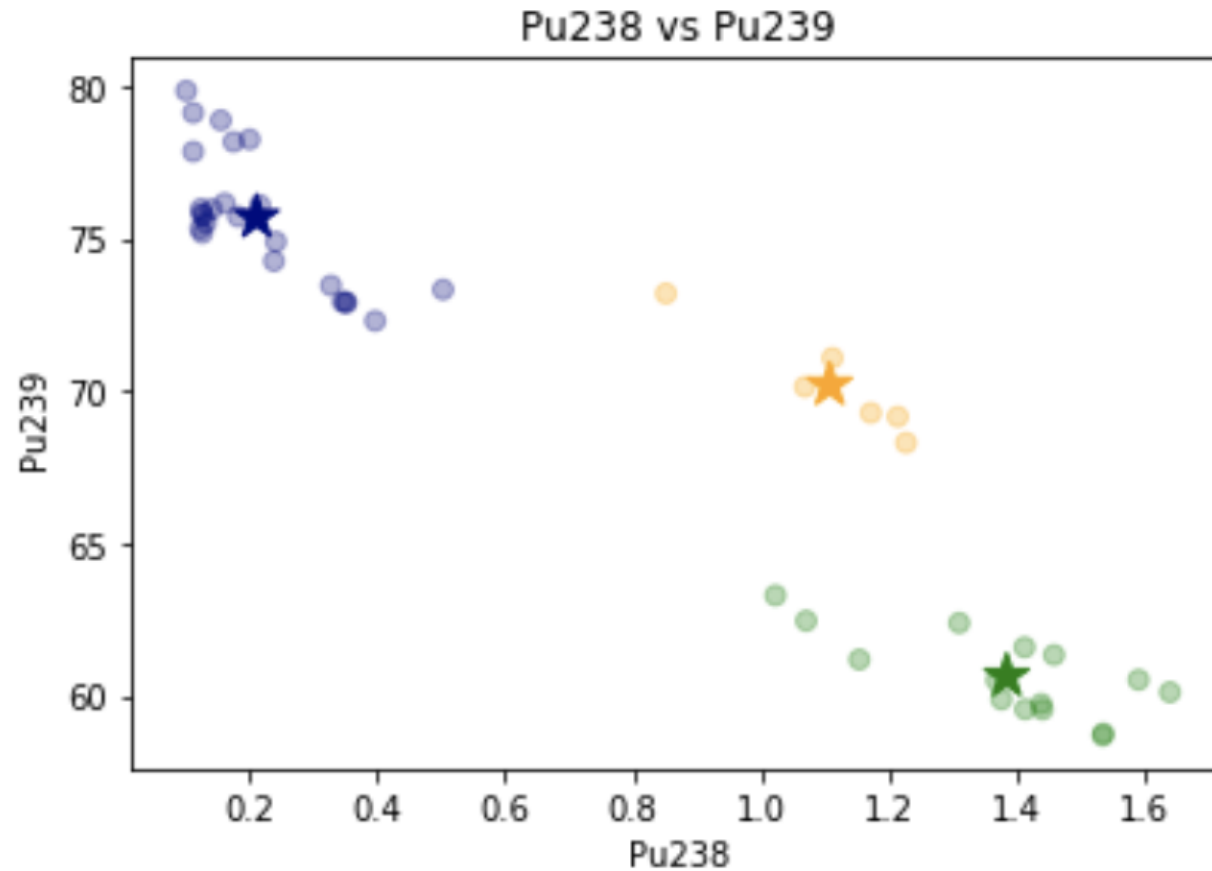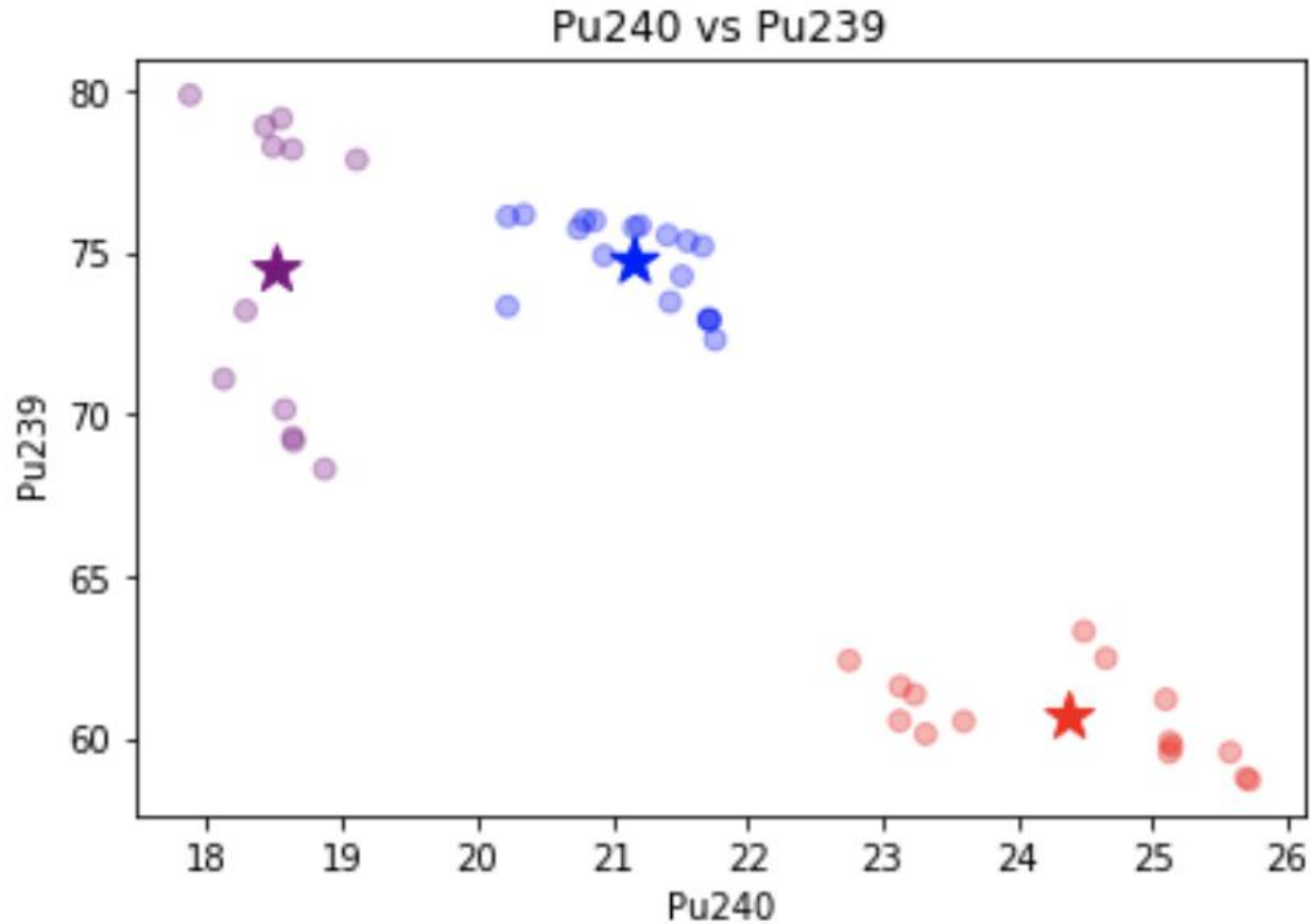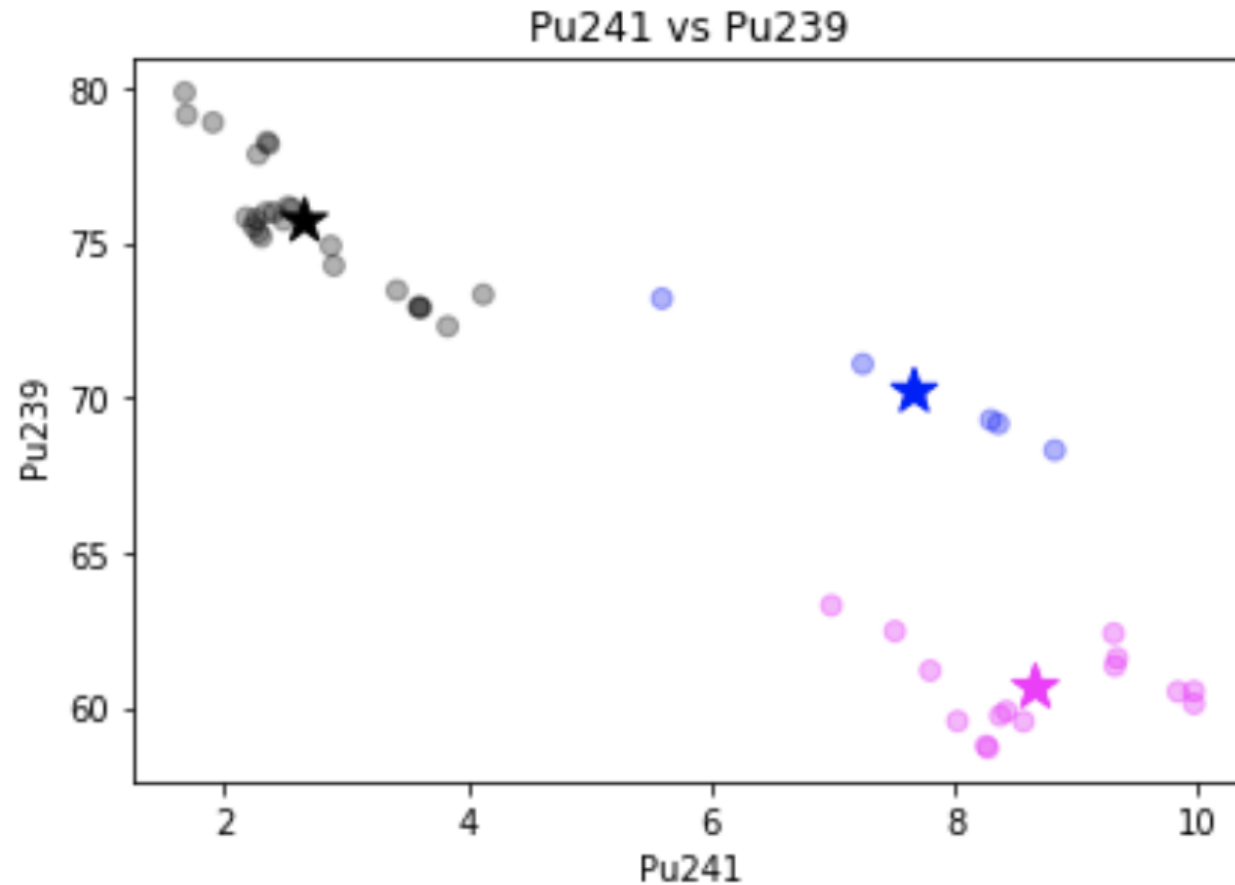

Pu238 vs Pu239

Silhouette score for Pu-238 and Pu-239 is 0.7138299182180218

See Clustering – Pu240 VS Pu239

Silhouette score for Pu-240 and Pu-239 is 0.633217477686232

See Clustering – Pu241 VS Pu239

Silhouette score for Pu-241 and Pu-239 is 0.6985505682786127

Results/Summary:

-Pu-238 and Pu-241 appear to be the closest in space to Pu-239.

-Also, they have the lowest silhouette scores, meaning they are not as easily separable from Pu-239 which we want because we want them to be as similar as possible.

References

1) Trevino, Andrea, "Introduction to Kmeans Clustering,"**Oracle Data Science Blog**, https://blogs.oracle.com/datascience/introduction-to-k-means-clustering, February 20, 2020

2) "Plutonium Use", Radioactivity.ed.com, https://www.radioactivity.eu.com/site/pages/Plutonium_Use.htm, February 20, 2020

3) "Silhouette Clustering," Wikipedia.com, https://en.wikipedia.org/wiki/Silhouette_(clustering), February 20, 2020

4) "8.01 Kmeans Isotopes Lab", General Assembly Data Science Intensive, Atlanta, Georgia, February 20, 2020