Hi Suzannah, I have taken the K means ML algorithm for my home task.

* **Is it Supervised/Unsupervised/Reinforcement learning?**

K-means clustering is an [unsupervised learning](https://deepai.org/machine-learning-glossary-and-terms/unsupervised-learning).

* **What does the algorithm do?**

K-means clustering is an [unsupervised learning](https://deepai.org/machine-learning-glossary-and-terms/unsupervised-learning) technique to [classify](https://deepai.org/machine-learning-glossary-and-terms/classifier) unlabelled data by grouping them by features, rather than pre-defined categories. The variable K represents the number of groups or categories created. The goal is to split the data into K different clusters and report the location of the centre of mass for each cluster. Then, a new data point can be assigned a cluster (class) based on the closed centre of mass.

**How does K-Means Clustering Work?**

Each centroid of a cluster is a collection of feature values which define the resulting groups. Examining the centroid feature weights can be used to qualitatively interpret what kind of group each cluster represents.

**Data assignment**: Each cluster is created and defined by its centroid (central collection of features). Each data point is then assigned to its nearest centroid, based on some choice of distance function

**Centroid update**: After all data points are assigned, the centroids are recalculated by taking the mean of all data points assigned to that cluster.

**Repeat**: This assignment and update process repeats until some stopping criteria is met, such as, no change to clusters, the sum of the distances is minimized, or some maximum iteration threshold is reached.

* **In which situations will it be most useful?**

It will helpful in a situation where we have n no of patterns and we need to find out which pattern has got more connections inside and also to find out which patterns are well connected together to do some insight.

* **(Optional) Can you find any examples of where this algorithm has been used**

**Real time approach:**

I have taken super market products as an example for my real time approach. I am assuming each section in the super market as my clusters, where I create a most demanding product as my centroid with the help of the previous data. I am connecting my centroid to all the products in that particular section. Now I can able to find what is the best combination to go with the centroid. So, we increase the product sale if we give more variety or by giving more offers with the most demanding product.

Next, we can measure the distance of each cluster, by doing this we can find which two clusters has more connection. By doing this we can identify, which are the combined products are in demand in that particular shop. By analysing this, we can increase the sales by keeping more demanded products nearer to each product