**M3 – D5 – Task Notes**

* 3 Galaxies, separate them.
  + ~~3 Major clutters from a scatter graph~~
* Uppermost galaxy
  + ~~Single out the topmost Major cluster =~~ Galaxy 0
  + MUST:
    - ~~Find the topmost major cluster mathematically, not through the legend and visuals.~~
* Baby Yoda is in right most part of selected cluster.
  + ~~Select the right most planet / coordinates from the selected cluster.~~
  + MUST:
    - ~~Store Yoda’s coordinates into its own variable,~~ **~~do not~~** ~~use the printed number,~~ **~~do not~~** ~~call the index of galaxy\_0.~~
* Input coordinates into the force finder
  + ~~Jons software –~~ Force Token = Access to 2nd CSV
    - Code must be clean, presentation ready. **REMOVE things that do not work.** **KEEP everything that gets us the answers.**
* From planet data you will need to extract the force concentration and its two principal components
  + ~~Map the 2 PCA’s of the data for the planet selected.~~
* Find baby yoda in the closest point to the gravitational centre of the two PCAs. Highlight on map the gravitational centre and its closest point.
  + ~~Find and plot gravitational centre of the two PCA lines. Highlight their gravitational centre.~~
* Deliver these final coordinates to teacher.
  + As a Jupiter notebook file, not GitHub – x = 0.14618599 y = -0.0314799

**KMeans function breakdown:**

* Choose number of clusters.
* Pick random points based on cluster amount.
  + Points need to be within parameters of largest x and largest y.
* Pick data points at random and check them against the cluster points. Cluster point closest to data point then moves on top of that data point **OR** cluster points move next to that data point.
  + Save each data point into that clusters list.
* Decide when KMeans is finished sorting?
  + Mean of all the data against the total?