## Assignment Instructions

1. Load and inspect the file 'ASA All NBA Raw Data.csv', which is raw NBA data for the time range 2019-10-22 to 2022-02-27.
2. Conduct a cluster analysis and divide the players into 5 groups.
3. Use inertia to determine what is the optimal number of clusters.
4. Conduct a cluster analysis and divide the players into the optimal number of groups.
5. Characterize each of these groups. Can you characterize each group?
6. Conduct a PCA analysis on the data
    - Reduce your data into 3 components.
    - Use 95% variance and let the PCA choose the number of components.
7. Use non-linear dimensionality reduction to display the data.

## What to produce?

1. Generate a plotter chart of your five clusters.
2. Try to define what characterizes each of the five cluster?
3. What is the optimal number of clusters based on your inertia?
4. Show a figure of your updated clusters.
5. What is the first number that you see in your explained variance? What is its meaning? What is the total variance explained by this model?
6. How many components are required to explain 95% of the variance? Which component accounts for the highest explained variance?