**K Means Cluster Problem**

For this project, we will attempt to use KMeans Clustering to cluster Universities into two groups, Private and Public.

**It is very important to note, we actually have the labels for this data set, but we will NOT use them for the KMeans clustering algorithm, since that is an unsupervised learning algorithm.**

When using the Kmeans algorithm under normal circumstances, it is because you don't have labels. In this case we will use the labels to try to get an idea of how well the algorithm performed, but you won't usually do this for Kmeans, so the classification report and confusion matrix at the end of this project, don't truly make sense in a real-world setting!

**The Data**

We will use a data frame with 777 observations on the following 18 variables.

* Private A factor with levels No and Yes indicating private or public university
* Apps Number of applications received
* Accept Number of applications accepted
* Enroll Number of new students enrolled
* Top10perc Pct. new students from top 10% of H.S. class
* Top25perc Pct. new students from top 25% of H.S. class
* F.Undergrad Number of fulltime undergraduates
* P.Undergrad Number of parttime undergraduates
* Outstate Out-of-state tuition
* Room.Board Room and board costs
* Books Estimated book costs
* Personal Estimated personal spending
* PhD Pct. of faculty with Ph.D.’s
* Terminal Pct. of faculty with terminal degree
* S.F.Ratio Student/faculty ratio
* perc.alumni Pct. alumni who donate
* Expend Instructional expenditure per student
* Grad.Rate Graduation rate

**Here take the dataset:** [**K Means Cluster Problem: Data Set**](https://docs.google.com/document/d/1Ta9BnFJpqOeuy9flapPXCHcsYOPWtj5pRpSOILh9a7o/edit?usp=drivesdk)

**Steps that you can follow:**

1. Read the dataset as a pandas dataset
2. Do some primary analysis on that by using head, info, describe etc.
3. Plot the graphs.
4. Here in this dataset, the result is also there, so on fitting the data, do drop result.
5. With the result, you can see the score of your model