The dataset [Education - Post 12th Standard.csv](https://olympus.mygreatlearning.com/courses/78180/files/5510379/download?verifier=V7FXa5qDDiX6kKrYf6QnDsNFLacD0bddD8fB2orc&wrap=1) contains information on various colleges. You are expected to do a Principal Component Analysis for this case study according to the instructions given. The data dictionary of the 'Education - Post 12th Standard.csv' can be found in the following file: [Data Dictionary.xlsx](https://olympus.mygreatlearning.com/courses/78180/files/5510378/download?verifier=VV7P65zCUC8DHgEGTjlwSDaCH3PcDbJLHkBOqa4S&wrap=1).

* Perform Exploratory Data Analysis. What insight do you draw from the EDA?
* Is scaling necessary for PCA in this case? Give justification and perform scaling.
* Comment on the comparison between the covariance and the correlation matrices from this data [on scaled data].
* Check the dataset for outliers before and after scaling. What insight do you derive here?
* Extract the eigenvalues and eigenvectors.
* Perform PCA and export the data of the Principal Component (eigenvectors) into a data frame with the original features
* Write down the explicit form of the first PC (in terms of the eigenvectors
* Consider the cumulative values of the eigenvalues. How does it help you to decide on the optimum number of principal components? What do the eigenvectors indicate?
* Explain the business implication of using the Principal Component Analysis for this case study. How may PCs help in the further analysis?