Learning Outcomes:
You will learn to create multiple ML models on a dataset using

1. PCA,
2. Regularization,
3. Linear regression
4. and/or combinations of these

Dataset -
You are given the data OnlineNewsPopularity.csv which has attributes as given OnlineNewsPopularity.names The requirement is to predict, given the details about a news article how many times would it be shared (i.e., 'shares' is the target column). The task is to apply your knowledge of regression, PCA and Regularization and report your evaluations on each experiment.

1. **Set seed and Split the data into train and test**

2. **In each experiment you should**

0. Ensure that preprocessing on test data must only use whatever parameters were used for train preprocessing

1. Create a model on train data
2. Get predictions for train data and test data from the model
3. Evaluate `train predictions with train target` and `test predictions with test target`
4. Store the evaluated results

3. **Run the following experiments**

1. Baseline - using linear regression (lm)
2. Remove columns based on VIF and apply lm
3. AIC - forward, backward and both
4. Elastic Net - Find the best alpha and lambda using cv.glm
5. PCA preprocessing -

Decompose the train data into principle components

Rotate both train and test data using `predict(pcaObject, data)` and obtain data in principal component space

Apply lm on the rotated train data and evaluate the model on the rotated train and test data

4. **Report all the errors in a single data frame with column names c('Expt Name', 'Train RMSE', 'Test RMSE')**