**REGRESSION DATASET**

Below dataset contains articles taken from [**www.mashable.com**](http://www.mashable.com)

The dataset is pretty large with **39644** rows and **56** columns.

Perform following tasks on the dataset

1. Load the appropriate libraries in python
2. Perform exploratory data analysis such as finding mean, median, describe functions.
3. Find whether dataset contains any null values or not.
4. Handle null values wisely and justify the reason.
5. Plot several visualizations such as boxplot, kdeplot, scatter plots, pair grid and histograms to give insights about dataset and write observations such as which columns are categorical and which are continuous and do they follow Gaussian distribution.
6. The target variable is **shares.** Choose independent variable(s) wisely and justify the reason for selecting them.
7. Fit a Linear Regression model using **shares** as target variable.
8. Choose whether some columns can be dropped or combined together to get better results with just few / larger columns.
9. Using operations in step 8 again fit a linear regression model.
10. Compare linear regression model fit in step 7 and 9 using mean

squared error.