Big Data Analytics
Assignment-03 (Spring 2024)

Note: Late Submission will be not be graded.

Submission to be done on portal

Due Date: Sunday, June 9, 2024, before 11:59 PM

- **Take screen shoots of R code and paste on the word file. Export world file to .pdf**
- **Upload .pdf file on portal**
- **[Note: Title page must have Student Name, Registration Number, Section, and Date of Submission.**
- **Don't use register pages and assignment pages**
- **`Submission will be on portal only`**

Regression and Clustering in R

Objective:
This assignment aims to help you understand and implement different types of regression and clustering techniques in R. You will generate simple datasets and apply Linear Regression, Multiple Regression, Polynomial Regression, Logistic Regression, and K-Means Clustering.

Instructions:

Generate datasets:

- For each regression method, generate a simple dataset that demonstrates the concept.
- For K-Means Clustering, create a dataset with clear clusters.

Tasks:

1. Linear Regression

- **Dataset**: Create a simple dataset with a linear relationship between variables.
- **Task**: Fit a linear regression model and visualize the results.

2. Multiple Regression

- **Dataset**: Create a dataset with multiple predictors.
- **Task**: Fit a multiple regression model and summarize the results.

3. Polynomial Regression

- **Dataset**: Create a dataset with a polynomial relationship.
- **Task**: Fit a polynomial regression model and visualize the results.

4. Logistic Regression

- **Dataset**: Create a binary outcome dataset.
- **Task**: Fit a logistic regression model and summarize the results.

5. K-Means Clustering

- **Dataset**: Create a dataset with clear clusters.
- **Task**: Apply K-Means clustering and visualize the clusters.