

**IE 7280: Statistical Methods in Engineering**

**Project Proposal**

**Topic 1: Regression Analysis**

**Dataset:** [**https://www.kaggle.com/mehdidag/black-friday**](https://www.kaggle.com/mehdidag/black-friday)

**Description**: This dataset contains more than 500000 rows of transactions in a retail store. It has many independent variables like Age, Occupation, Marital Status, Gender etc. of the customer and a dependent variable – Purchase amount.

**Objective:** To find relationship between dependent variable (Purchase Amount) and independent variables (Age, Occupation, Marital Status, etc.)

**Topic 2: ANOVA**

**Dataset:** [**https://www.kaggle.com/uciml/iris**](https://www.kaggle.com/uciml/iris)

**Description: This dataset contains data of three “iris” species containing information such as their petal and sepal length and width.**

**Objective:** to compare means of sepal width and petal length based on setosa versicolor and virginica (group means) and see if they are significant or is there any difference among their means. Similarly compare the average with and length for sepal and petal in one-way, 2-way anova to check for difference within or among means

**STUDENT NAME:** Hemanth Lakshman Raju

Umeash Vaithyalingam Gangatharan

**PROFESSOR:** Nasser Fard