**1. Introduction and Problem statement**

Assignment was to build a Multi linear regression model from dataset of 50 startups.

* Building model for predicting Profit of a startups.
* Three different influencing factors, R&D spending, administerial spending, and marketing spending are given.

**2. Methodology**

* Dataset were cleaned by checking null and Nan values in dataset. Duplicates were dropped.
* Correlation factors of profit with other three features were checked to gauge relation with profit
* Also, different plots were used to explore various features of dataset.
* A Multi linear regression model of profit with all three features was build and validated using Sklean python module.
* To understand the significance of individual features, linear regression models for those were also evaluated.

**3. Results and Analysis**

The Strength of the Model (r2 score) is: 0.9413262690696323

**4. Challenges Faced**

It was first multi regression model. Other than technical difficulty of how to code certain thing, major challenge faced was how to understand pre-process the data. What kind of plot to use and how to interpret information from that.

**5. Complexity**

* It was fairly simple and good assignment.
* All three features had different kind of influence and very distinct relation with profit.
* Getting good r2 score took some time.

**6. Time invested**

On and off it took **6 hours**