# <span style="color:red">_Capstone Project Submission_</span>

**Instructions:**
i) Please fill in all the required information.
ii) Avoid grammatical errors.

---

**Name, Email and Contribution:**

Baliram Kumar– baliramkumarmgs@gmail.com

i.     Analysis of Rossmann Store Data.
ii.    Analysis of Store Data.
iii.   Analysis of Datasets and performed Different EDA Operation
iv.    Applied Different Machine learning Model to Test the Data
v.     Performed Analysis on Model.
vi.    Appropriate Model Selection and Deriving conclusions based on all operations that we have performed.
vii.   Project PPT and Team Colab Building
viii.  Check the null values
ix.    Remove the outliers
x.     Performed Analysis on Model.
xi.    Appropriate Model Selection
xii.   Check total number of entries and column types

---

**Please paste the GitHub Repo link.**

Github Link:- https://github.com/Baliram72/Retail-Sales-Prediction

---

**Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**

Predicting sales performance is one of the key challenges every business faces. It is important for firms to predict customer demands to offer the right product at the right time and at the right place. The importance of this issue is underlined by the fact that figuratively a bazillion consulting firms are on the market trying to offer sales forecasting services to businesses of all sizes. Some of these firms rely on advanced data analytics techniques, the kind of which we will be covering in this Project.

Rossmann is the largest drugstore in Germany. Moreover, it operates over 3,000 drugstores in 7 European countries. In 2015, Rossmann store managers are tasked with predicting their daily sales for up to six weeks in advance. Store sales are influenced by many factors, including promotions, competition, school and state holidays, seasonality, and locality. With thousands of individual managers predicting sales based on their unique circumstances, the accuracy of results can be quite varied.

We explored and analyzed the data to discover key factors responsible for engagement and success Sales. In this we had to remove all the null values as well as garbage values so that we can get appropriate data.

So as the second step we divided our task in Five sections:

1. **Analysis of Rossmann Store Data :-**

   In this part we have seen that store, day of week, date, sales, customers, open, promo, state holiday, school holiday these Columns are present in Rossmann data set and how different Columns Affect Sales and promo.

2. **Analysis of Store Data :-**

   In this part we have seen different types of columns present in the dataset such as store, storetype, assortment, competition distance, competition open since month, promo2, promo2 since week, promo2 since year, promo interval.

3. **Analysis of Datasets and performed Different EDA Operation :-**

   In this part we have performed different EDA operations to retrieve more information from the datasets.

4. **Applied Different Machine learning Model :-**

   In this part we have performed different Machine Learning operations to retrieve more information from the datasets.

5. **Appropriate Model Selection :-**

   In this part we have Analysed different models and selected appropriate models for final model selection which have maximum model accuracy and minimum time required for execution.