# Topic: Support Vector Machines (SVM)

**Instructions:**

Please share your answers filled in-line in the word document. Submit code separately wherever applicable.

Please ensure you update all the details:

**Name: \_\_\_\_\_\_\_\_\_\_\_\_\_ Batch ID:** \_\_\_\_\_\_\_\_\_\_\_

**Topic: SVM**

**Guidelines:**

**1. An assignment submission is considered complete only when the correct and executable code(s) and documentation explaining the method and results are submitted. Failing to submit either of those will be considered an invalid submission and not a correct submission.**

**2. Ensure that you submit your assignments correctly and in full. Resubmission is not allowed.**

**3. Post the submission you can evaluate your work by referring to the keys provided. (will be available only post the submission).**

**Hints:**

1. **Business Problem**
   1. **What is the business objective?**
   2. **Are there any constraints?**
2. **Work on each feature of the dataset to create a data dictionary as displayed in the below image:**

**Make a table as shown above and provide information about the features such as its data type and its relevance to the model building. And if not relevant, provide reasons and a description of the feature.**

**3. Data Pre-processing**

**3.1 Data Cleaning, Feature Engineering, etc.**

**3.2 Outlier Treatment.**

**4. Exploratory Data Analysis (EDA):**

**4.1 Summary.**

**4.2 Univariate analysis.**

**4.3 Bivariate analysis.**

**5. Model Building**

* 1. **Build the model on the scaled data (try multiple options)**
  2. **Use the SVM algorithm.**
  3. **Train and test the model and compare accuracies by building a confusion matrix and using different hyperparameters.**
  4. **Briefly explain the model output in the documentation.**

**6. Write about the benefits/impact of the solution - in what way does the business (client) benefit from the solution provided?**

**Problem Statement: -**

1. A construction firm wants to develop a suburban locality with new infrastructure but they might incur losses if they cannot sell the properties. To overcome this, they consult an analytics firm to get insights into how densely the area is populated and the income levels of residents. Use the Support Vector Machines algorithm on the given dataset and draw out insights and also comment on the viability of investing in that area.

