**NLP - Topic Modeling**

**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: NLP - Topic Modeling**

**Guidelines:**

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

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

**3. Exploratory Data Analysis (EDA).**

**4. Model Building**

**4.1 Perform Data Cleaning, Stemming, Lemmatization, Topic Modelling and Text Summarization.**

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

**Problem Statement-1**

1. Perform NLP – Topic Modelling and Text Summarization by following all the steps as mentioned below.
2. Data Cleaning using regular expressions, Countvectorizer, POS Tagging, NER, Topic modeling (LDA, LSA) and Text Summarization.

Hint: - the Data.csv file given in the hands-on material.

Text

Description automatically generated

**Problem Statement-2**

Perform Topic Modelling and Text Summarization on the given text data. Use the NLP-TM text file.

**A picture containing letter

Description automatically generated**