**Dataset Directory**

Suppose you are working as an analyst in an e-commerce company, and you have been given the two datasets containing the details of women's apparels sold during the last year.

Attribute Dataset: This dataset contains the different features of women's apparels.

Dress Sales: This data set contains the number of sales for a particular dress ID on a particular date.

In this data set, there are a total of 13 dress-related features.

**Dress\_ID:** This represents the ID number of a particular dress. This is a unique identification number for different dresses.

**Style:** This represents the style of a particular dress according to the occasion, like a party, a vintage event, etc.

**Price:** Each dress ID can fall in a particular price bucket, which can be Low, Medium and High.

**Rating:** This is the average rating given by the customers for a particular dress ID.

**Size:** The size column represents the majority of the size bought by the customers for that particular dress ID in the previous sale.

**Season:** This depicts the season in which a particular dress is suitable, for example, summer, winter, etc.

**Neckline:** This contains the type of neck in the dress, like V-neck, round-neck, etc.

**SleeveLength:** This represents the type of sleeve of the dresses. Half sleeves, full sleeves, cap sleeves, etc.

**Material:** This contains information regarding which material the dress has been made of, like cotton, nylon, polyester, silk, etc.

**FabricType:** This contains information regarding the type of fabric of the dress, for example, chiffon, broadcloth, jersey, etc.

**Decoration:** This represents the kind of decoration around the dress, like ruffles, bow, embroidery, etc.

**PatternType:** This represents the type of pattern a particular dress has. Pattern may be solid colours, geometric designs, printed or patchwork.

**Recommendation:** This is the target variable. 'Recommendation' is based on the features and sales of the dress in the previous year. This is either 1 (means yes) or 0 (means no). This represents whether a particular dress is suitable for sale to the customers or not.

Now, moving to the second data set.

This particular data set represents the number of sales of a particular dress ID on a certain date, where columns represent the dates on which a particular dress ID has been sold.

Now, based on the above two datasets, you are expected to perform the EDA and draw useful insights from that. Based on the EDA analysis, answer the graded questions for this module.

You have gone through the data cleaning part with an example of bank marketing data set in the previous segments. Now, let’s answer the following questions based on all that you learnt in this session.