**Description of Dataset and Problem Statement**

This dataset contains data of performance of a broadband gateway network (BNG) device. BNG device is the access point for subscribers, through which they connect to the broadband network. When a connection is established between BNG device and Customer Premise Equipment (CPE), the subscriber can access the broadband services provided by the Network Service Provider (NSP) or Internet Service Provider (ISP). A network management company monitors the performance of each of the BNG devices to provide better services. The company records different performance measures (PMs) for every 15 minutes. A description of different columns i.e. attributes (PMs) in this dataset is detailed below.

**CreationTime**: Date and time of the recording of sample

**AuthenticateCount**: Number of active subscribers authenticated their connection.

**ActiveCount**:Number of active subscribers connected to the device.

**DisconnectCount**: Number of active subscribers disconnected from the device.

**CPUUtil**: contains the % of usage of processor in the device.

**MemoryUsed**: Total memory in Bytes used in the device

**MemoryFree**: Total of memory in Bytes free in the device

**TempMin**: Minimum temperatures among the temperatures recorded from the different slots in the device

**TempMax**: Maximum temperatures among the temperatures recorded from the different slots in the device

**TempAvg**: Average temperatures among the temperatures recorded from the different slots in the device

**InBandwidth**: Total bandwidth utilization in Bytes from the input ports of all the interfaces.

**OutBandwidth**: Total bandwidth utilization in Bytes from the output ports of all the interfaces.

**InTotalPPS**: Total packets per second transmitted from the input ports of all the interfaces.

**OutTotalPPS**: Total packets per second transmitted from the output ports of all the interfaces.

--------------------------------------------------------------------------------------------------------------------

Perform the data cleaning if required. Perform the descriptive analytics to understand the data and infer from the data.

Perform the predictive analytics (regressive analysis) on predicting the **InBandwidth** using different regression techniques and compare the results. Perform the data pre-processing (normalization, standardization, correlation analysis & feature section, dimension reduction using PCA) and compare the results of regression with unprocessed data.

Consider 70% of data for training and remaining 30% of data for testing.

Infer the results obtained from descriptive and predictive analytics.