**Data Set** – diabetic

**Data Information:-**

The datasets consist of several medical predictor (independent) variables and one target (dependent) variable, **Outcome**. Independent variables include the number of pregnancies the patient has had, their BMI, insulin level, age, and so on.

**Attribute information :-**

Pregnancies - Number of times pregnant

Glucose - Plasma glucose concentration a 2 hours in an oral glucose tolerance test

BloodPressure - Diastolic blood pressure (mm Hg)

SkinThickness - Triceps skin fold thickness (mm)

Insulin - 2-Hour serum insulin (mu U/ml)

BMI - Body mass index (weight in kg/(height in m)^2)

DiabetesPedigreeFunction - Diabetes pedigree function

Age - Age (years)

Outcome - Class variable (0 or 1) 268 of 768 are 1, the others are 0

**In class Assignment Expectations/Steps -**

* Apply Data Cleaning to the Datasets and then apply PCA Find pattern and the choose the number of desired Principal components.
* Provide the inferences for the above analysis.