Please process the attached “Breast Cancer” dataset (See descriptions on the next page) using python and answer the following questions:

* How many instances are there?

116

* How many are healthy controls and how many are patients

healthy controls: 52, patients: 64

* How many people are in the follow age range

Young 0<age<30

Mid-aged 30<=age<40

Senior 40<= age < 150

Young: 5

Mid-aged: 10

Senior: 101

* Define five ranges of BMI and calculate the number of people in each range

Underweight: 1

Normal: 39

Overweight: 37

Obese: 39

* Using different appropriate visualization method to visualize the dataset. Show one visualization for each of the following attribute: Age, BMI, Glucose, Insulin, HOMA, Leptin, Adiponectin, Resistin and MCP.1. (copy-and-paste the screenshot of each visualization picture)
* (Open-end question) What can you observe from the dataset. Any interpretation? (list three answers)

a. A majority of the patients where seniors

b. Roughly all of the patients where not underweight

c.Based off insulin levels, 35 pateints are in sever risk to have type 2 diabetes

==========

Submission requirements:

* Please write the answer to the questions on this document.
* Please submit your python source code.

**====================================================**

**Data Set Information:**

There are 10 predictors, all quantitative, and a binary dependent variable, indicating the presence or absence of breast cancer.  
The predictors are anthropometric data and parameters which can be gathered in routine blood analysis.  
Prediction models based on these predictors, if accurate, can potentially be used as a biomarker of breast cancer.

**Attribute Information:**

Quantitative Attributes:  
Age (years)  
BMI (kg/m2)  
Glucose (mg/dL)  
Insulin (µU/mL)  
HOMA  
Leptin (ng/mL)  
Adiponectin (µg/mL)  
Resistin (ng/mL)  
MCP-1(pg/dL)  
  
Labels:  
1=Healthy controls  
2=Patients