IC 272: DATA SCIENCE - III
LAB ASSIGNMENT – III
Data visualization and statistics from data

Student's Name: Vivek jaiswal **Mobile No: 9887552039**

Roll Number: B20172 **Branch:DSE**

1

Table 1 Mean, median, mode, minimum, maximum and standard deviation for all the attributes

S. No.	Attributes	Mean	Median	Mode	Min.	Max.	S.D.
1	pregs	3.8450	3.00	1	0	17	3.367
2	plas	120.8945	117.0	99	0	199	31.951
3	pres (in mm Hg)	69.1054	72.0	70	0	122	19.343
4	skin (in mm)	20.5364	23.0	0	0	99	15.941
5	test (in mu U/mL)	79.7994	30.5	0	0	846	115.168
6	BMI (in kg/m^2)	31.9925	32.0	32.0	0	67.1	7.879
7	pedi	0.4718	0.3725	0.254	0.078	2.42	0.331
8	Age (in years)	33.2408	29.0	22	21	81	11.752

Inferences:

1. Infer if there is any relation between the magnitude of standard deviation and mean, mode and median values.(Hint : If standard deviation is close to zero; are mean, median and mode close to each other?)
2. Inference 2(You may add or delete the number of inferences)

2 a.

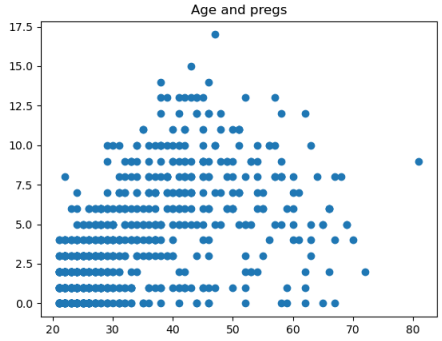

Figure 1 Scatter plot: Age (in years) vs. pregs

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with Age (in years) and y-axis legend with pregs.

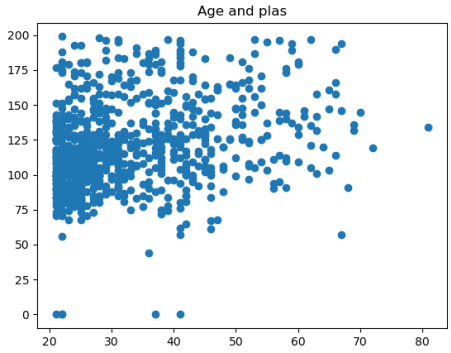

Figure 2 Scatter plot: Age (in years) vs. plas

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with Age (in years) and y-axis legend with plas.

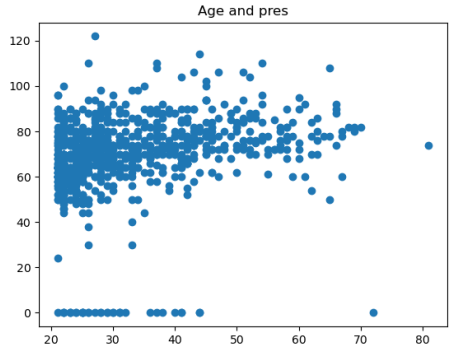

Figure 3 Scatter plot: Age (in years) vs. pres (in mm Hg)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with Age (in years) and y-axis legend with pres (in mm Hg).

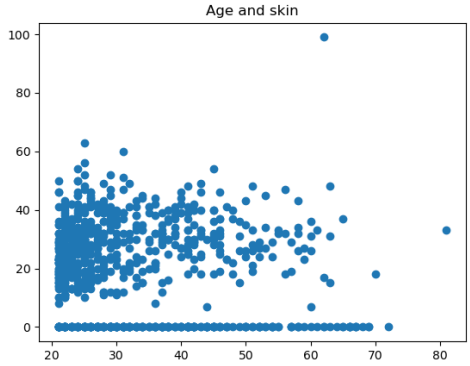

Figure 4 Scatter plot: Age (in years) vs. skin (in mm)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)

Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with Age (in years) and y-axis legend with skin (in mm).

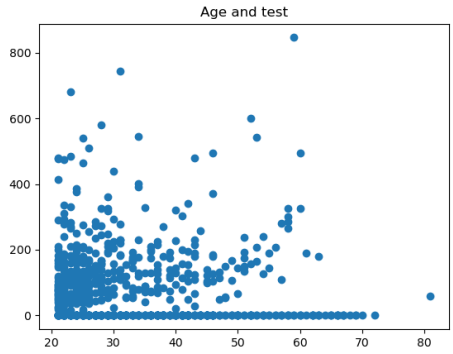

Figure 5 Scatter plot: Age (in years) vs. test (in mm U/mL)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with Age (in years) and y-axis legend with test (in mm U/mL).

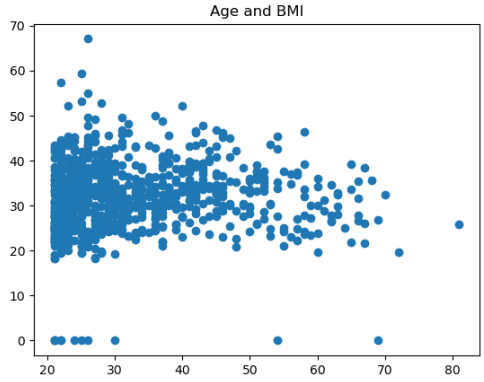

Figure 6 Scatter plot: Age (in years) vs. BMI (in kg/m²)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with Age (in years) and y-axis legend with BMI (in kg/m^2).

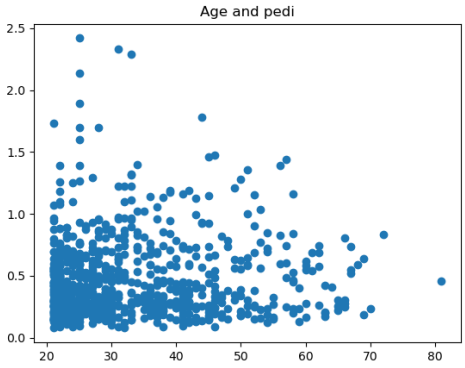

Figure 7 Scatter plot: Age (in years) vs. pedi

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with Age (in years) and y-axis legend with pedi.

b.

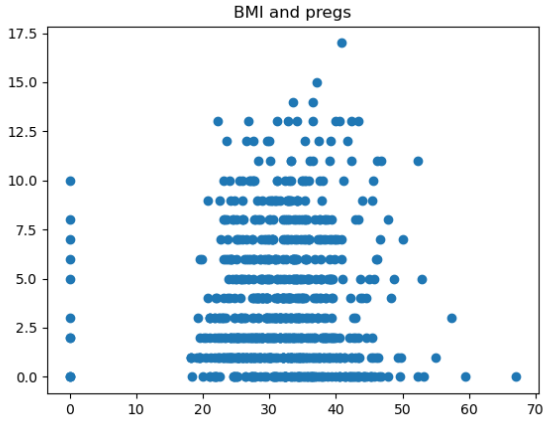

Figure 8 Scatter plot: BMI (in kg/m²) vs. pregs

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with BMI (in kg/m^2) and y-axis legend with pregs.

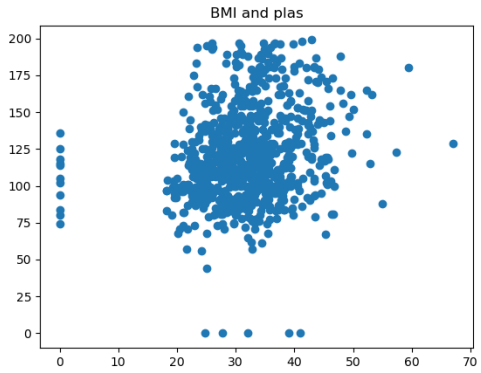

Figure 9 Scatter plot: BMI (in kg/m²) vs. plas

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with BMI (in kg/m^2) and y-axis legend with plas.

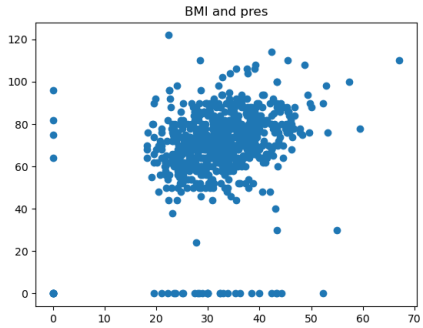

Figure 10 Scatter plot: BMI (in kg/m²) vs. pres (in mm Hg)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)

Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with BMI (in kg/m^2) and y-axis legend with pres (in mm Hg).

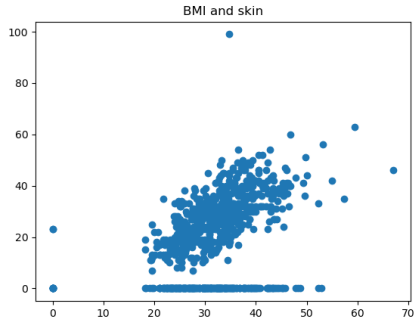

Figure 11 Scatter plot: BMI (in kg/m^2) vs. skin (in mm)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)

   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with BMI (in kg/m^2) and y-axis legend with skin (in mm).

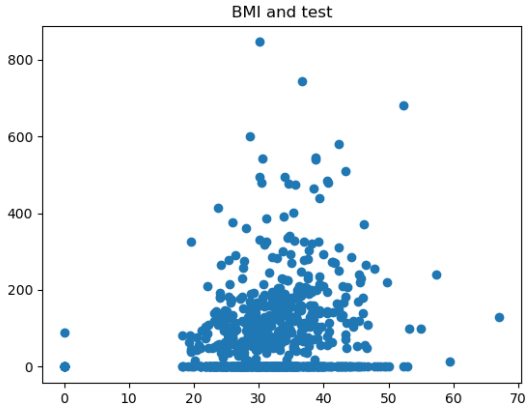

Figure 12 Scatter plot: BMI (in kg/m^2) vs. test (in mm U/mL)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with $^{\text{BMI (in kg/m2)}}$ and y-axis legend with x2.

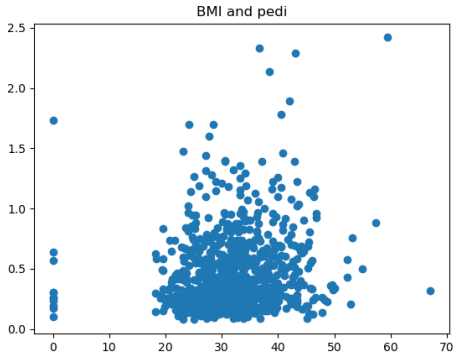

Figure 13 Scatter plot: BMI (in kg/m^2) vs. pedi

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with BMI (in kg/m^2) and y-axis legend with pedi.

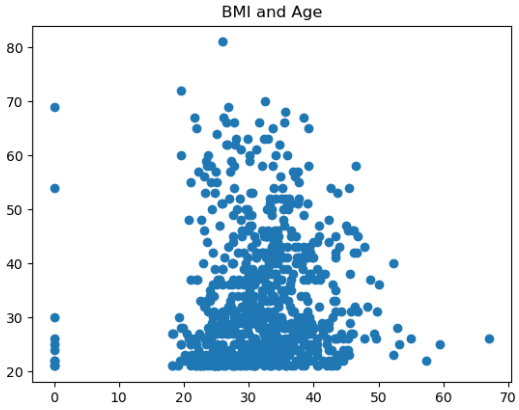

Figure 14 Scatter plot: BMI (in kg/m²) vs. Age (in years)

Inferences:

1. Infer how the attribute 1 is correlated to attribute 2 based upon spread of the data points
2. Inference based on density of points
3. Inference 3(You may add or delete the number of inferences)
   Note: The scatter plot above is for illustration purpose. Replace it with the scatter plot obtained by you. Rename x-axis legend with BMI (in kg/m^2) and y-axis legend with Age (in years).

3 a.

Table 3 Correlation coefficient value computed between age and all other attributes

S. No.	Attributes	Correlation Coefficient Value
1	pregs	0.544
2	plas	0.263
3	pres (in mm Hg)	0.239
4	skin (in mm)	-0.11
5	test (in mu U/mL)	-0.04
6	BMI (in kg/m^2)	0.036
7	pedi	0.033
8	Age (in years)	1.0

Inferences:

1. From the magnitude of correlation coefficient value, comment on the degree of correlation between age and each of the attribute.
2. From the sign of correlation coefficient value, comment whether with increase or decrease in age each of the attributes will increase or decrease.
3. Relate and comment on the value of correlation coefficient with corresponding scatter plot.
4. Inference 4(You may add or delete the number of inferences)

b.

Table 4 Correlation coefficient value computed between BMI and all other attributes

S. No.	Attributes	Correlation Coefficient Value
1	pregs	0.017
2	plas	0.22
3	pres (in mm Hg)	0.28
4	skin (in mm)	0.39
5	test (in mu U/mL)	0.19
6	BMI (in kg/m^2)	1.0
7	pedi	0.14
8	Age (in years)	0.036

Inferences:

1. From the magnitude of correlation coefficient value, comment on the degree of correlation between age and each of the attribute.
2. From the sign of correlation coefficient value, comment whether with increase or decrease in age each of the attributes will increase or decrease.
3. Relate and comment on the value of correlation coefficient with corresponding scatter plot.
4. Inference 4(You may add or delete the number of inferences)

4 a.

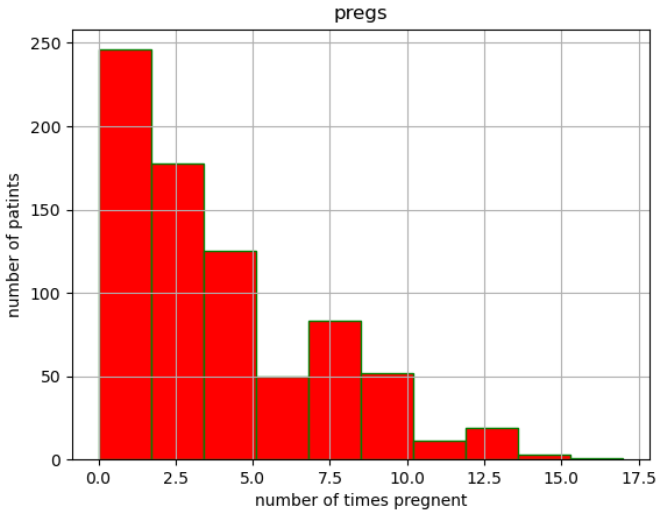

Figure 15 Histogram depiction of attribute pregs

Inferences:

1. Infer the frequency of each bin referring to its height.
2. From the histogram, infer in which of the bins mode of the attribute skin lies.
3. Inference 3(You may add or delete the number of inferences)
   Note: The histogram plot above is for illustration purpose. Replace it with the histogram plot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

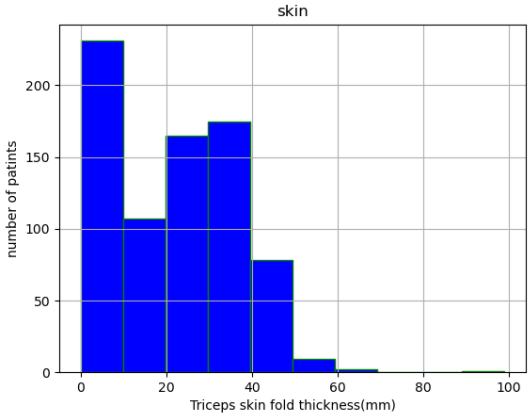

Figure 16 Histogram depiction of attribute skin

Inferences:

1. Infer the frequency of each bin referring to its height.
2. From the histogram, infer in which of the bins mode of the attribute skin lies.
3. Inference 3(You may add or delete the number of inferences)

Note: The histogram plot above is for illustration purpose. Replace it with the histogram plot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

5

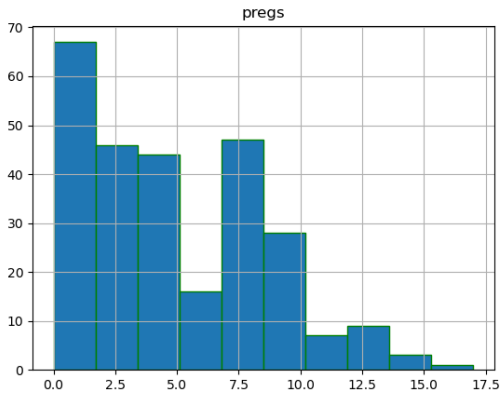

Figure 17 Histogram depiction of attribute pregs for class 0

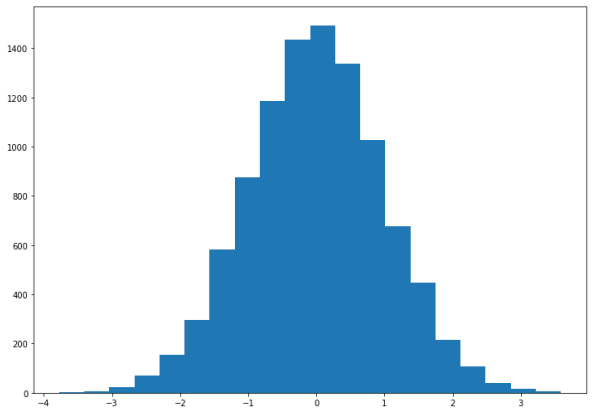

Figure 18 Histogram depiction of attribute pregs for class 1

Inferences:

1. From the histogram, infer in which of the bins mode of the attribute pregs lies for class 0 and 1.
2. Compare and contrast the frequency referring to the height of each bin for class 0 and 1
3. Inference 3(You may add or delete the number of inferences)
   Note: The histogram plot above is for illustration purpose. Replace it with the histogram plot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

6

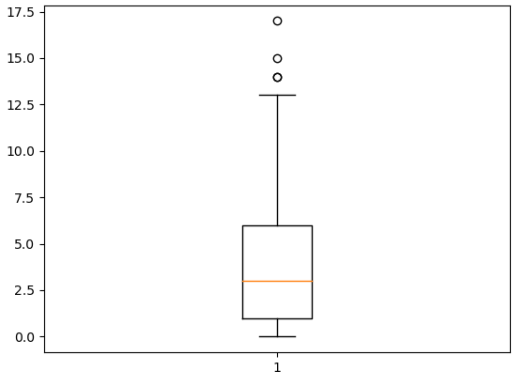

Figure 19 Boxplot for attribute pregs

Inferences:

1. Inference on outliers and their values.

2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Relate with the values from Q1. for this attribute.
6. Inference 6(You may add or delete the number of inferences)

   Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

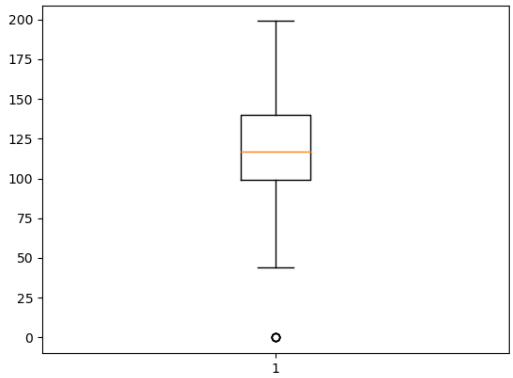

Figure 20 Boxplot for attribute plas

Inferences:

1. Inference on outliers and their values.
2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Relate with the values from Q1. for this attribute.
6. Inference 6(You may add or delete the number of inferences)

   Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

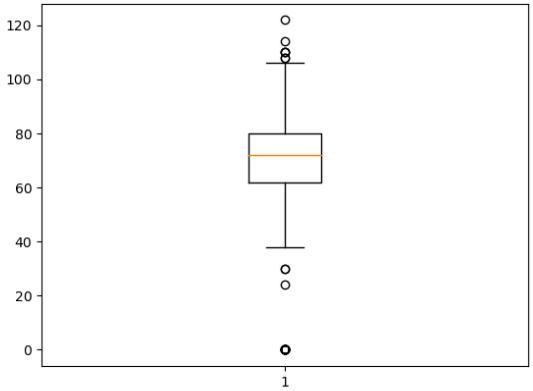

Figure 21 Boxplot for attribute pres(in mm Hg)

Inferences:

1. Inference on outliers and their values.
2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Relate with the values from Q1. for this attribute.
6. Inference 6(You may add or delete the number of inferences)

   Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

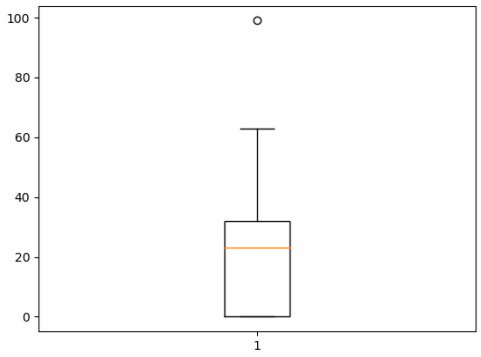

Figure 22 Boxplot for attribute skin(in mm)

Inferences:

1. Inference on outliers and their values.
2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Relate with the values from Q1. for this attribute.
6. Inference 6(You may add or delete the number of inferences)

   Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

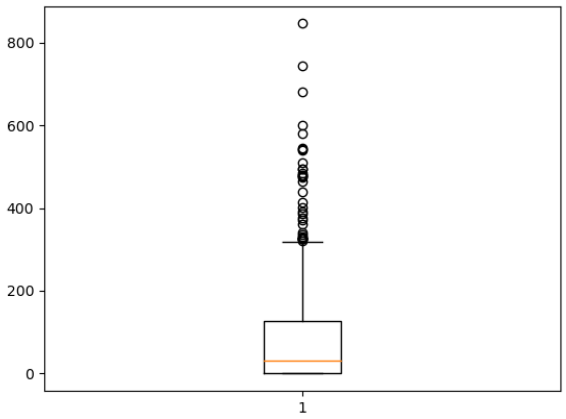

Figure 23 Boxplot for attribute test (mu U/mL)

Inferences:

1. Inference on outliers and their values.
2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Relate with the values from Q1. for this attribute.
6. Inference 6(You may add or delete the number of inferences)

   Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

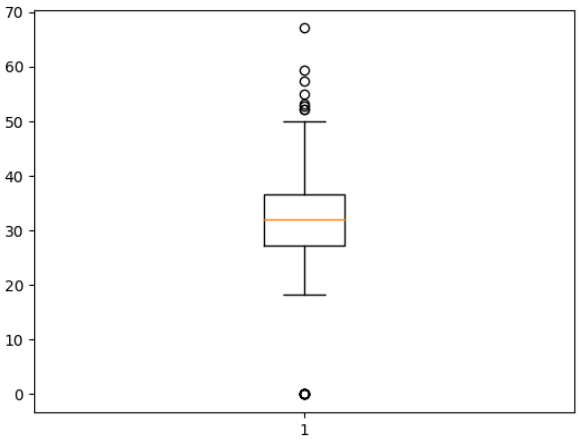

Figure 24 Boxplot for attribute BMI (in kg/m²)

Inferences:

1. Inference on outliers and their values.
2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Relate with the values from Q1. for this attribute.
6. Inference 6(You may add or delete the number of inferences)

   Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

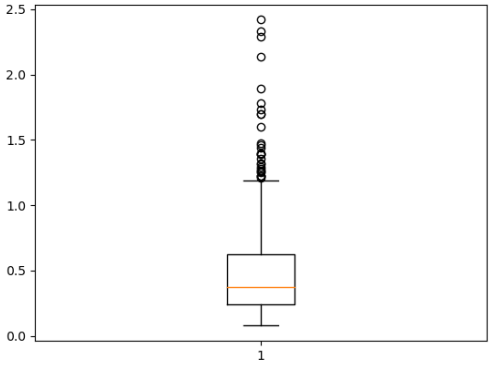

Figure 25 Boxplot for attribute pedi

Inferences:

1. Inference on outliers and their values.
2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Relate with the values from Q1. for this attribute.
6. Inference 6(You may add or delete the number of inferences)

Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units.

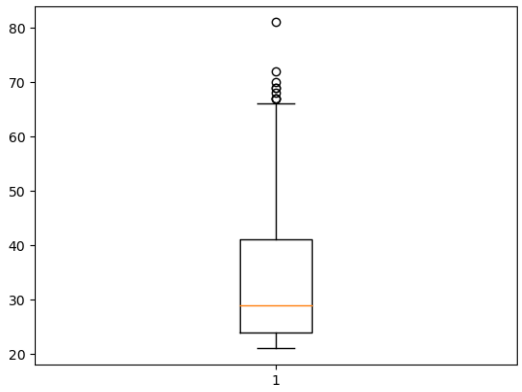

Figure 26 Boxplot for attribute Age (in years)

Inferences:

1. Inference on outliers and their values.
2. Infer the Inter quartile range.
3. Infer the variability of attribute.
4. Infer the skewness of the data.
5. Inference 5(You may add or delete the number of inferences)

Note: The boxplot above is for illustration purpose. Replace it with the boxplot obtained by you. Rename x-axis legend and y-axis legends with appropriate attribute names with units

Guidelines for Report (Delete this while you submit the report):

- **The plot/graph/figure/table should be centre justified with sequence number and caption.**
- **Inferences should be written as a numbered list.**

- **Use specific and technical terms to write inferences.**
- **Values observed/calculated should be rounded off to three decimal places.**
- **The quantities which have units should be written with units.**