Below are the instructions for your assignment. Please make sure your code is written correctly, efficiently and clearly. Please submit the code in R or Python (R is preferred, but use whichever you are most confident in) and submit it in a Jupyter notebook, python file or R file via the Greenhouse link below. Also, please make sure that we can see the code as well - not just the output.

The dataframe attached has records of 1,000 objects with 3 columns:

1. shape - square, triangle (equilateral) or circle

2. color - blue, red, green or yellow

3. area - area size in square inches

Questions:

1. Draw a boxplot showing the area size distribution for each shape.

2. Calculate the mean, max, and standard deviation of the area size of each color.

3. What is the average area size of a yellow square?

4. Which shape is most likely to be green?

5. Given the fact that the object is red, with an area size larger than 3,000 - what are the chances the object is a square? a triangle? a circle?

6. Write a function that calculates the side or radius of an object, depending on the shape and area of the object [for an equilateral triangle - area = (side ^ 2) * sqrt(3) / 4].

7. Add a column to the dataset called "side" that shows the size matching the area in each row, round that number to the closest integer (shape side or radius).

8. Draw a boxplot showing the side size distribution for each shape - what can you infer from this plot?

9. Make a scatter plot with "side" on the x axis, "area" on the y axis with a different color for each shape.

10. Create a dataframe, table or list that show for each shape:

    a. The proportion of red objects within the shape

    b. The proportion of blue area out of the shape's total area (sum of square inch blue area of the shape over sum of all shape size).

11. Create a function that calculates 10. b. for a given shape and color.