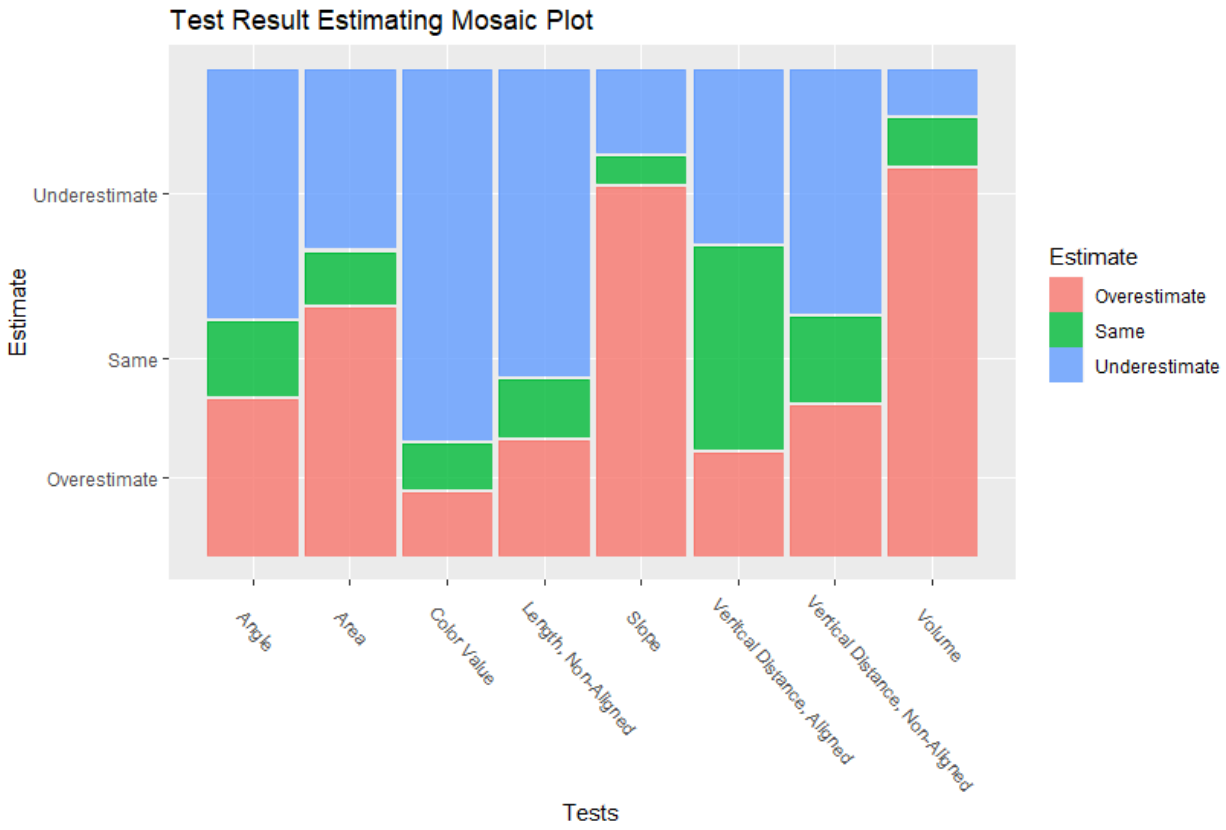
Name: Ashay Kargaonkar

DSC 465 Assignment 3

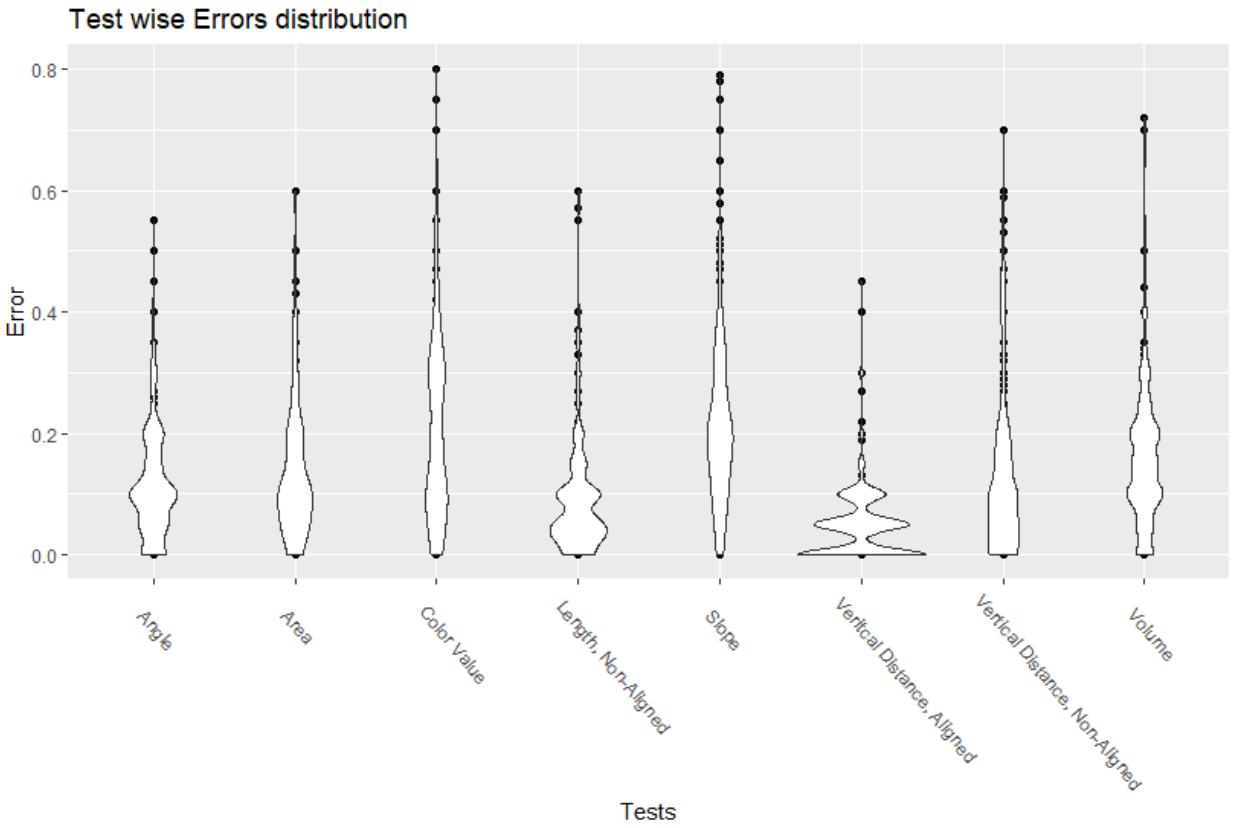
Q1.

a.



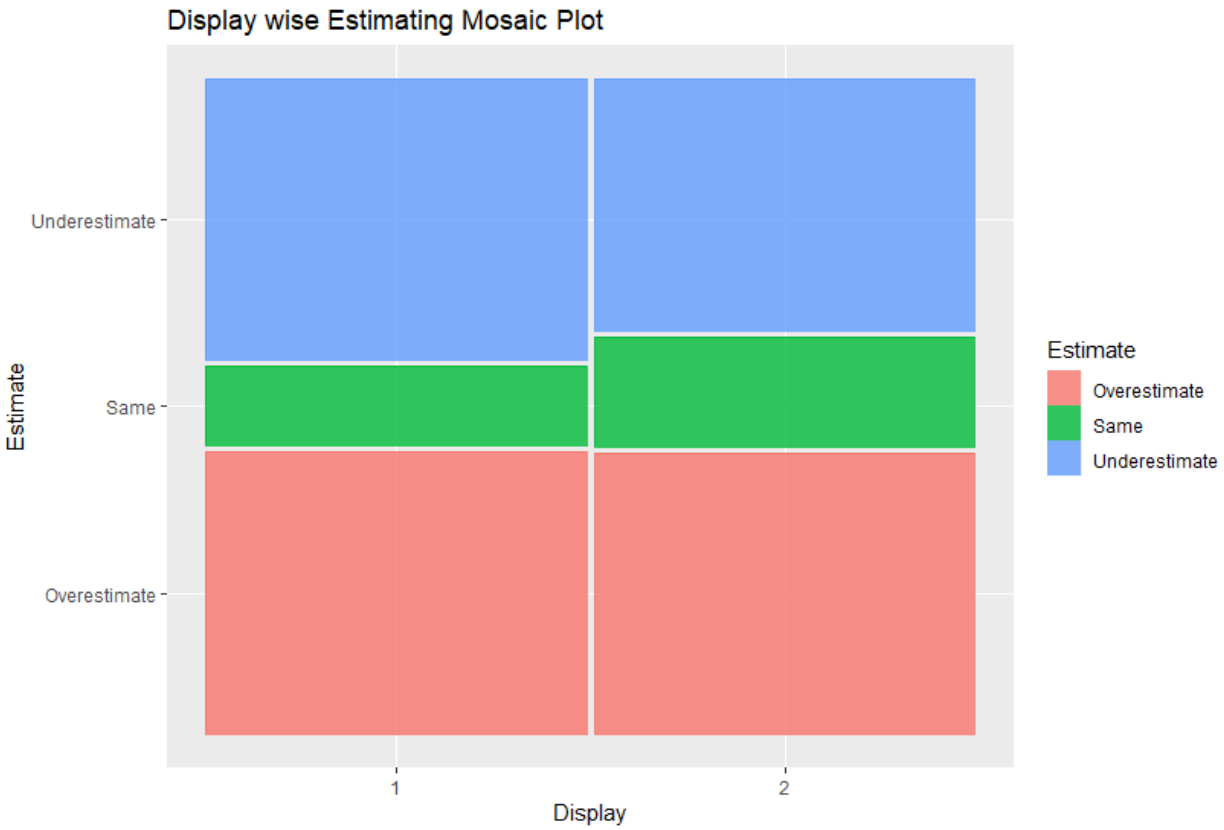
I used error column and distributed the values into 3 groups i.e., Same, Underestimate and Overestimate. If the value is less than 0 then underestimate, if greater than 0 then overestimate else its equal. Then I created a mosaic plot where x-axis contains different tests and y-axis shows about the newly created column (''Estimate").

b.



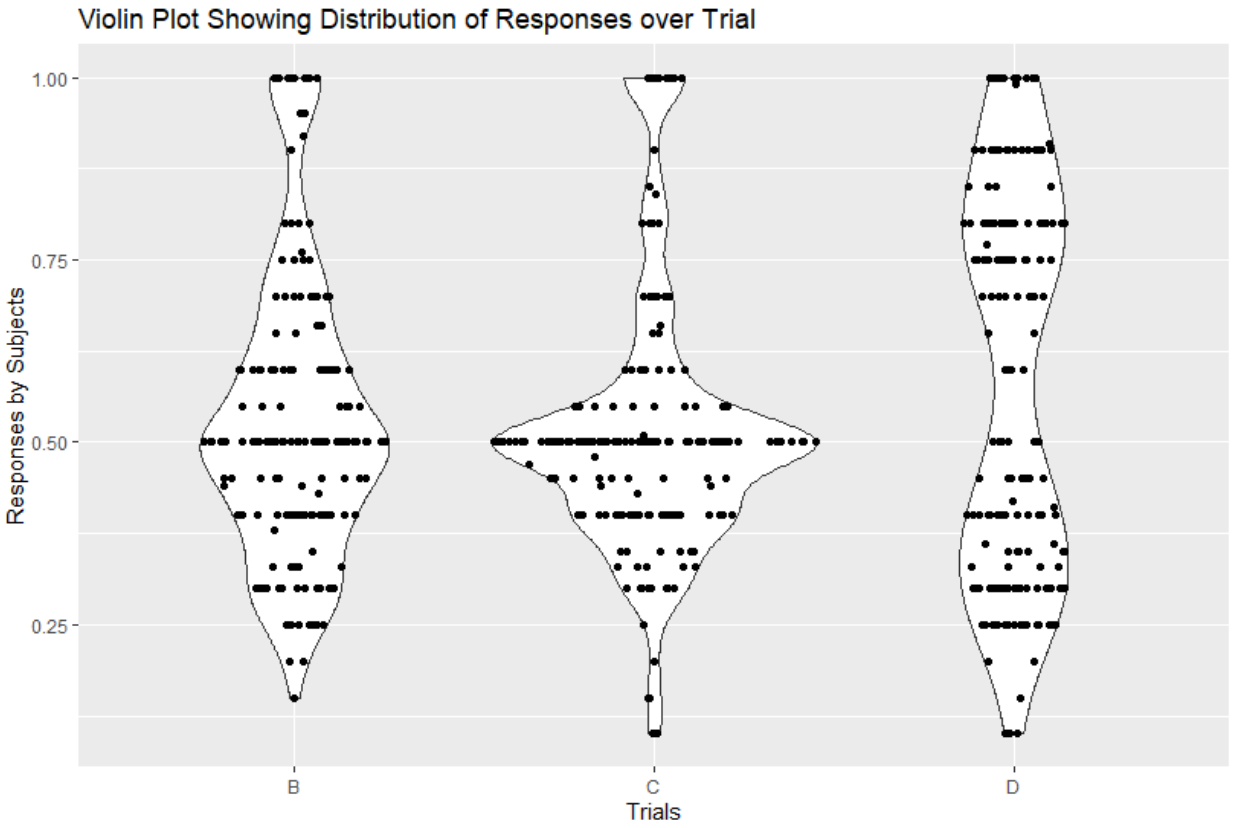
From the above graph we can see that most of the values of error are from 0.0 to 0.2 for all kinds of test. There are very less values which have error more than 0.2.

c.



From the above graph we can see that the results of participants even after doing test once does not improve their score much for the second test. There is a slight decrease in overestimate by underestimate data is same.

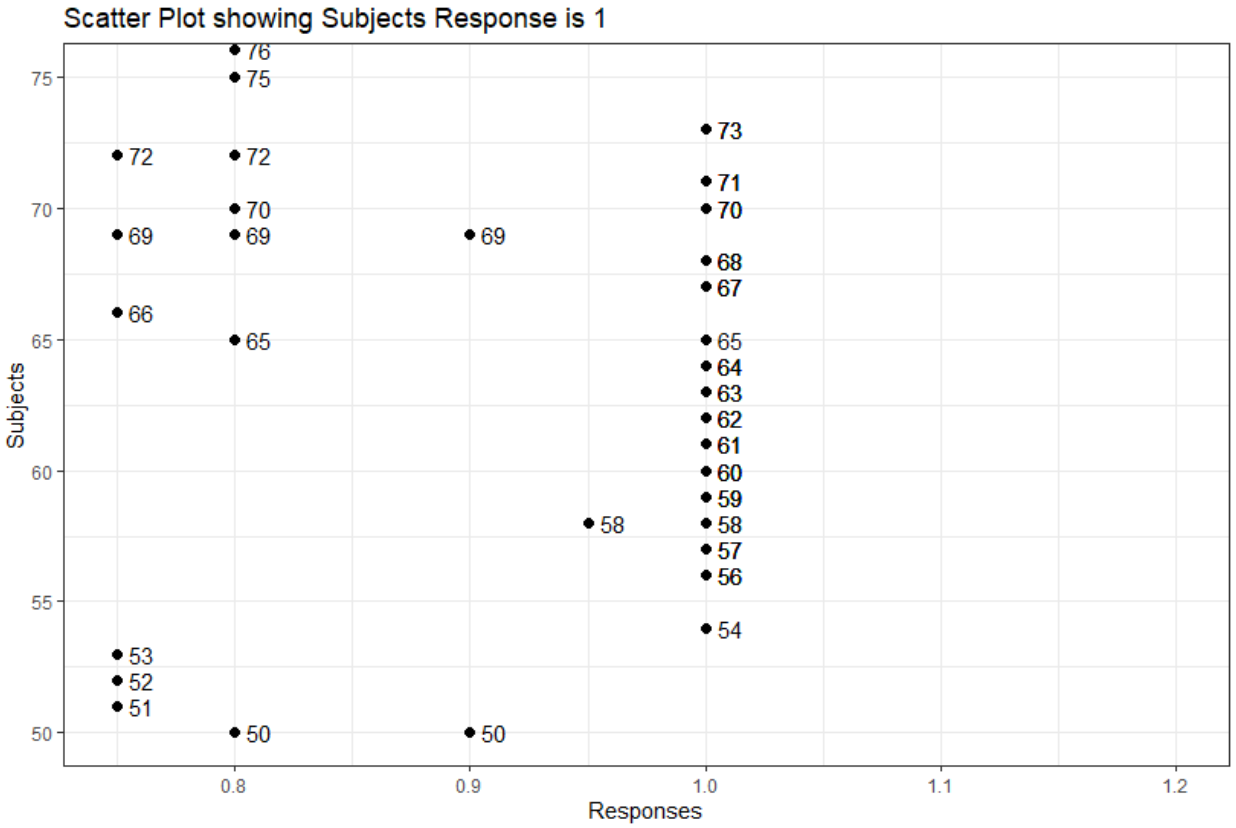
d.

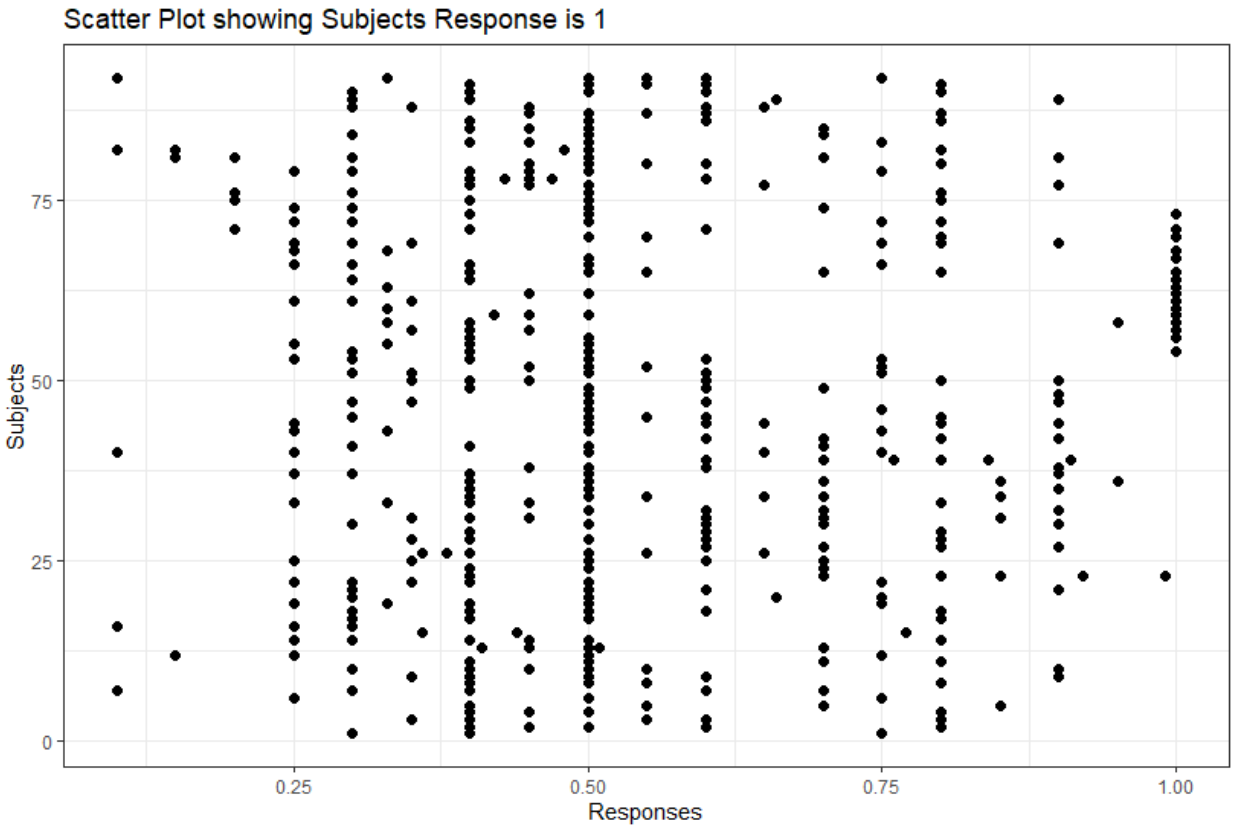


I have used violin plots with sina to display the anomalies in subject's result by showing the errors present in each trial. The graph is represented based on 3 trails B, C and D but only from Display 1.

Also, to support my further argument I have attached 2 more graphs below which captures the outliers with the labels of the subject numbers.

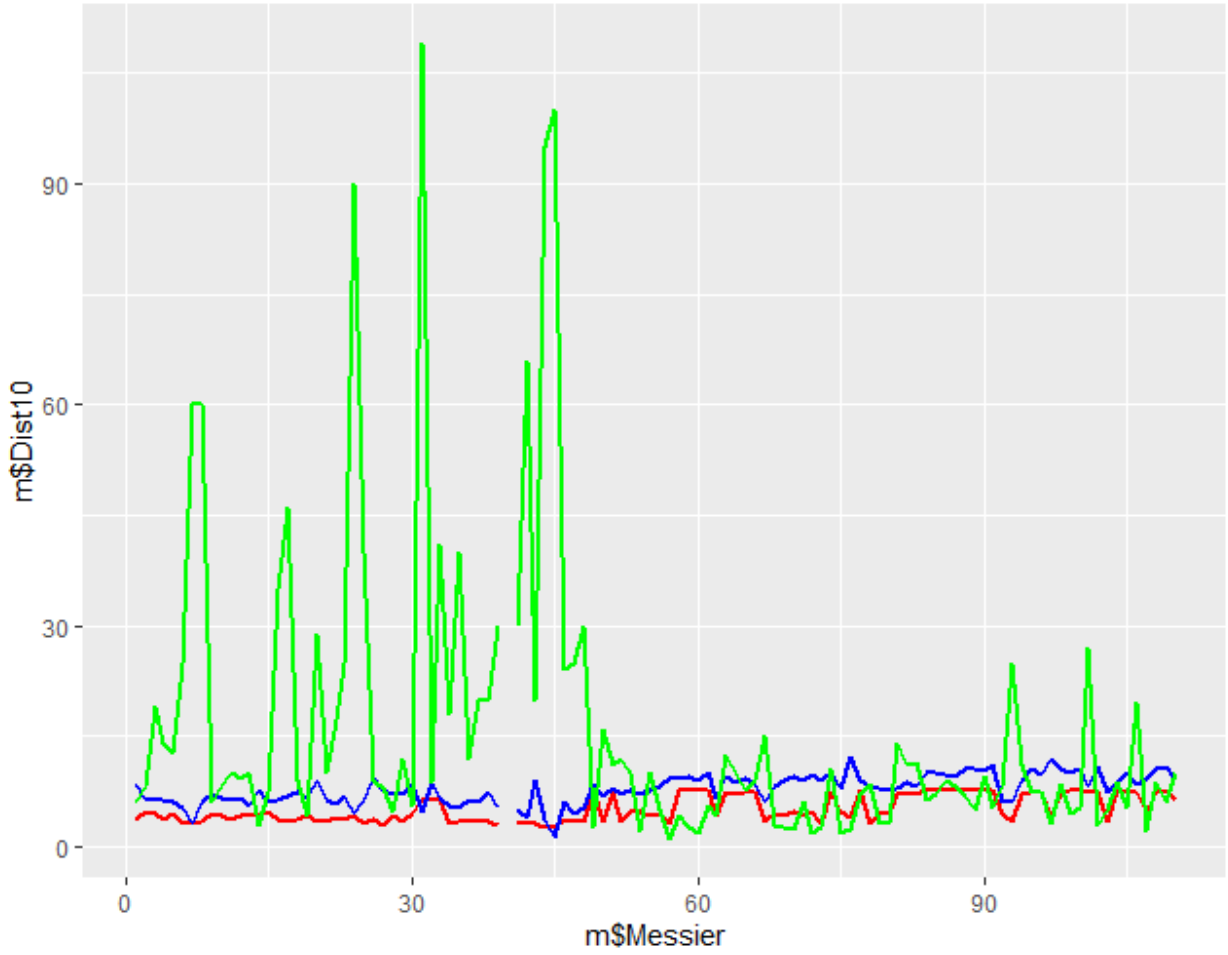
You can see that the datapoints shown on responses 1 are same and is showing the same patter


Scatter Plot showing Subjects Response is 1
Subjects
Responses
75
70
65
60
55
50
0.8
0.9
1.0
1.1
1.2
76
75
73
72
72
71
70
70
69
69
69
68
67
66
65
65
64
63
62
61
60
59
58
58
57
56
54
53
52
51
50
50


Scatter Plot showing Subjects Response is 1
Subjects
Responses
75
50
25
0
0.25
0.50
0.75
1.00

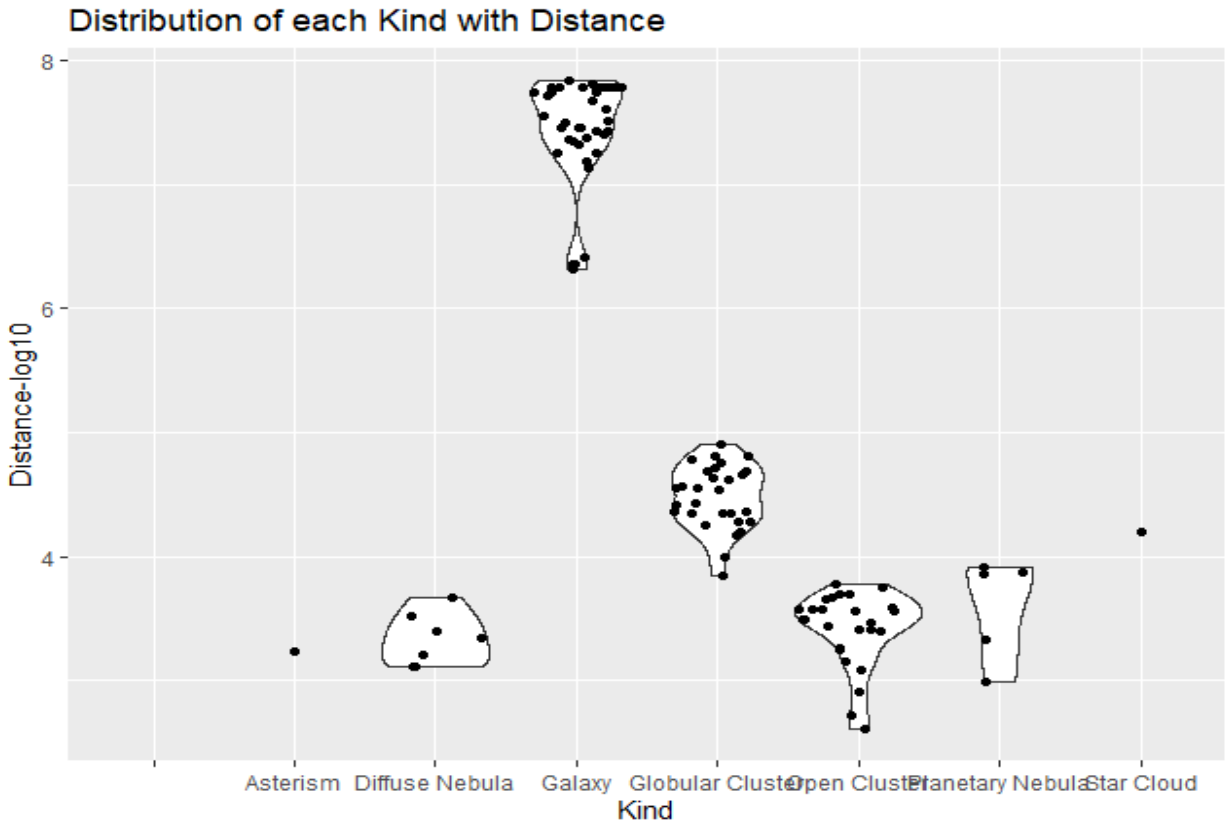
Q2.

a.

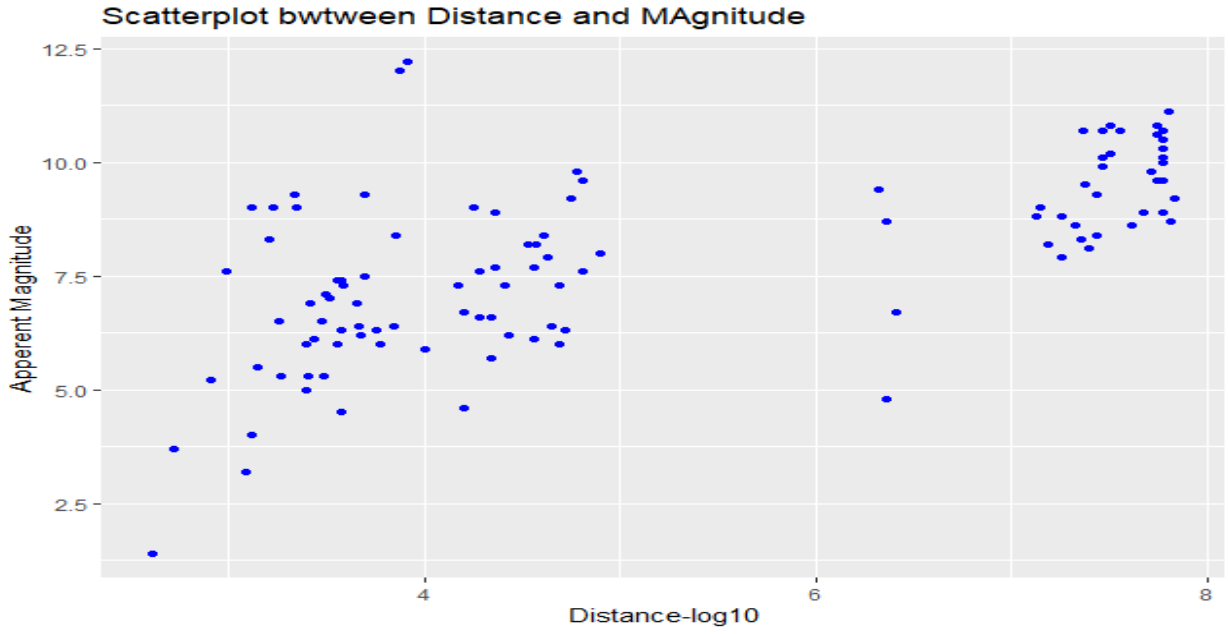


From the graph we can see that at initial stages of messier number size of objects were high. But after around half messier number distances become quite low but the distance started to increase in a small amount.

b.

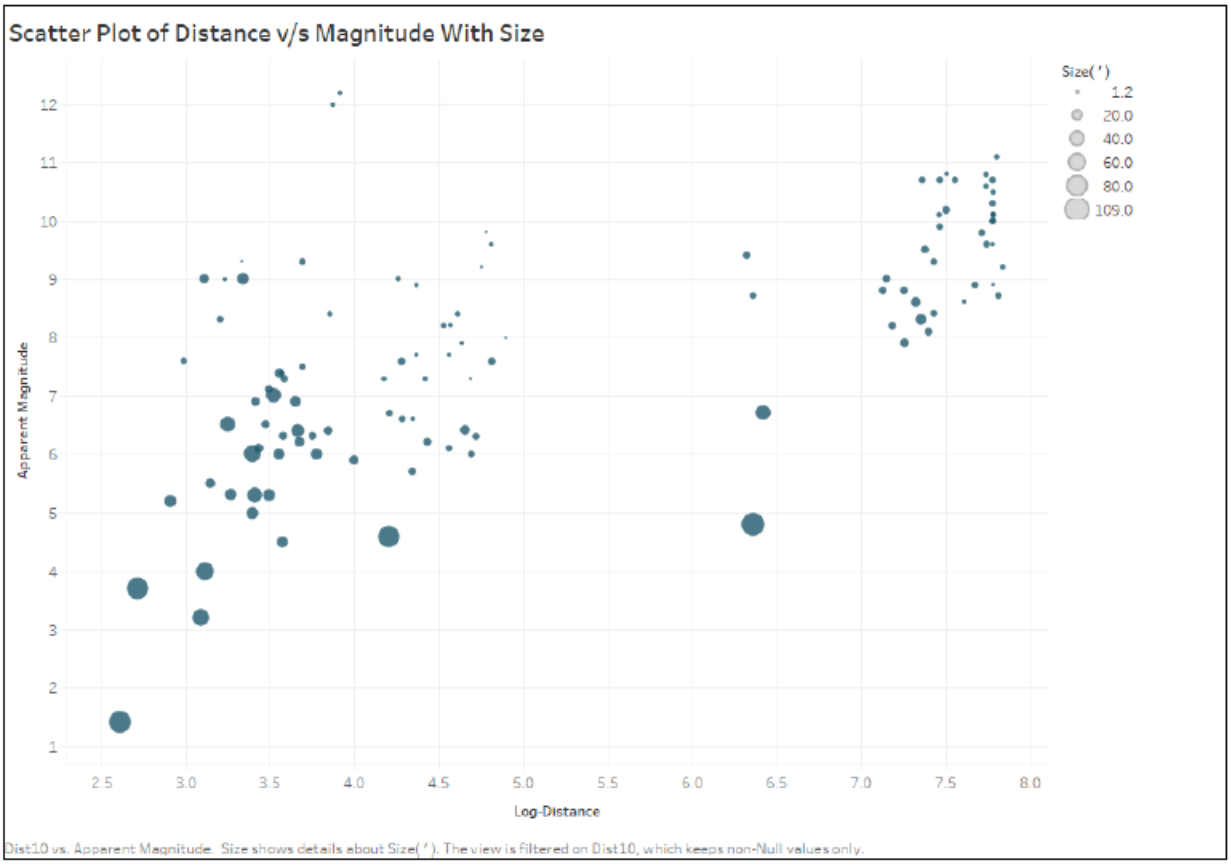

Distribution of each Kind with Distance
Distance-log10
8
6
4
Asterism
Diffuse Nebula
Galaxy
Globular Cluster
Open Cluster
Planetary Nebula
Star Cloud
Kind

c.


Scatterplot bwtween Distance and MAgnitude
Apperent Magnitude
12.5
10.0
7.5
5.0
2.5
4
6
8
Distance-log10

From the above graph we can see that there are like 2 clusters form. One less than log of distance less than 5 and other greater than. By the graph we can say that lesser the distance less the magnitude or light is darker and vice versa.

d.

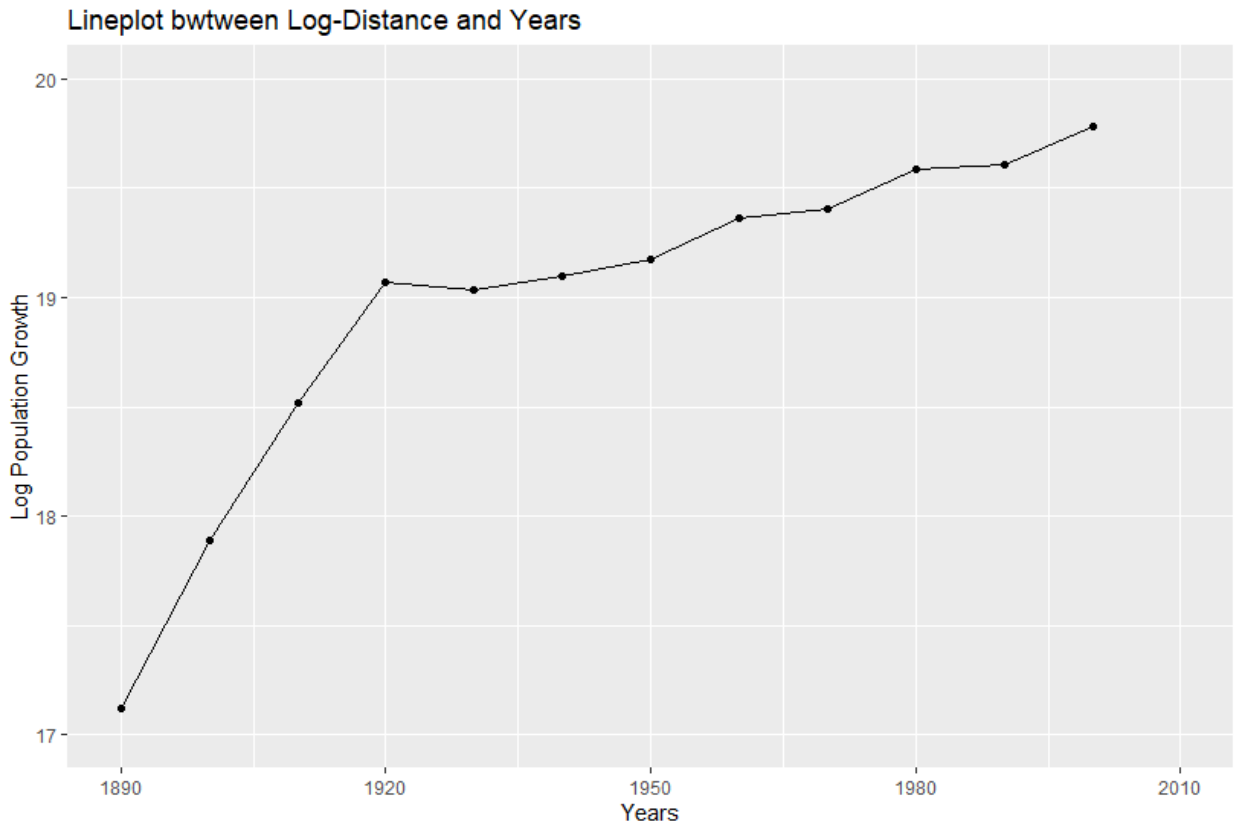


The above graph is made in Tableau. The size of data points denotes the size of object. It means bigger the data point's size bigger the size of object is or vice versa.

Also by looking at the graph we can see that at lower distances bigger size of objects are present.

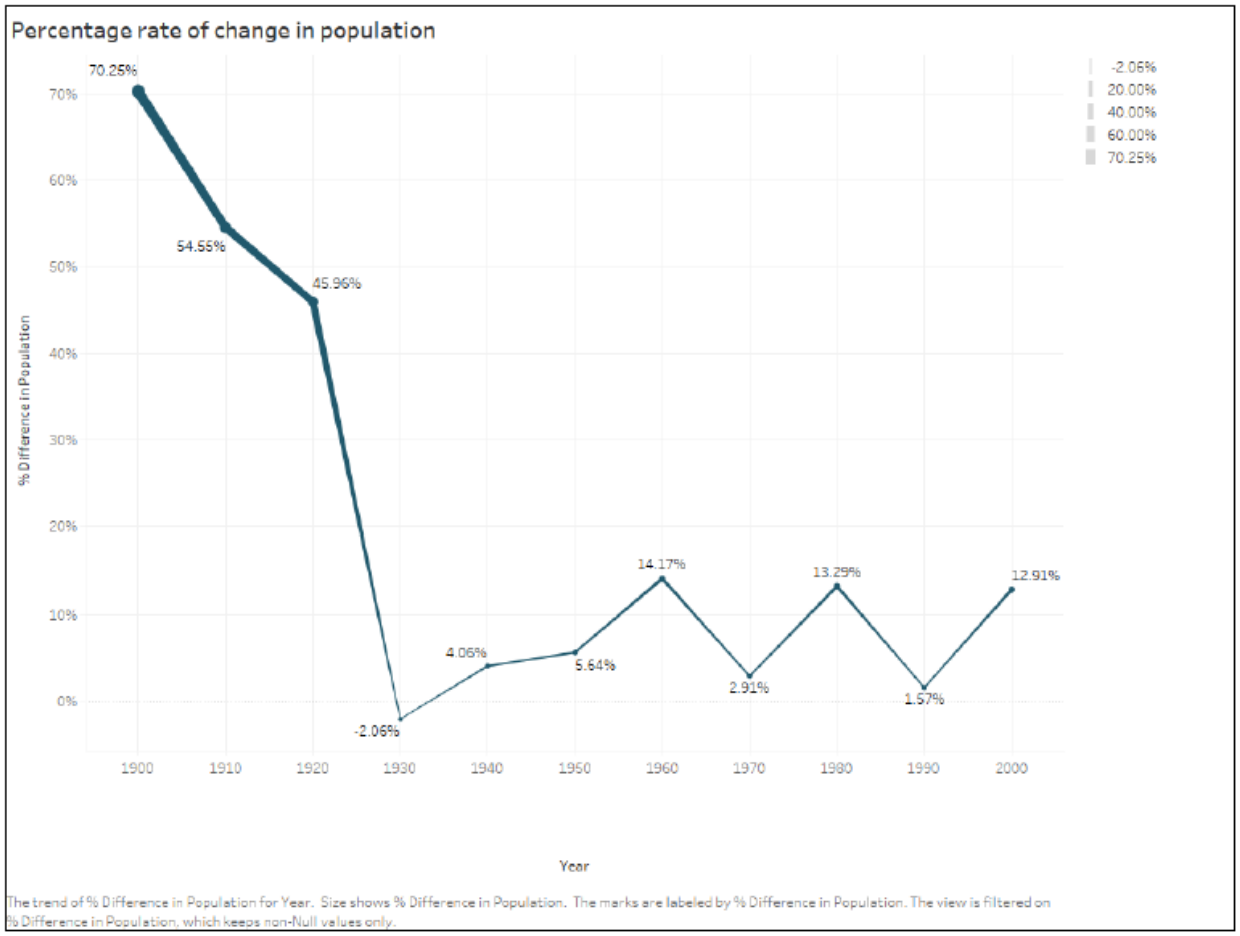
Q3.

a.



For creating this graph, I took log to the base 2 of the population. If we see that the y -value changes from around 17 to 20. We can say that the population has doubled almost 3 times but not full 3 times. Its because the value ranges from more than 17 but less than 20, so around 3 times.

b.



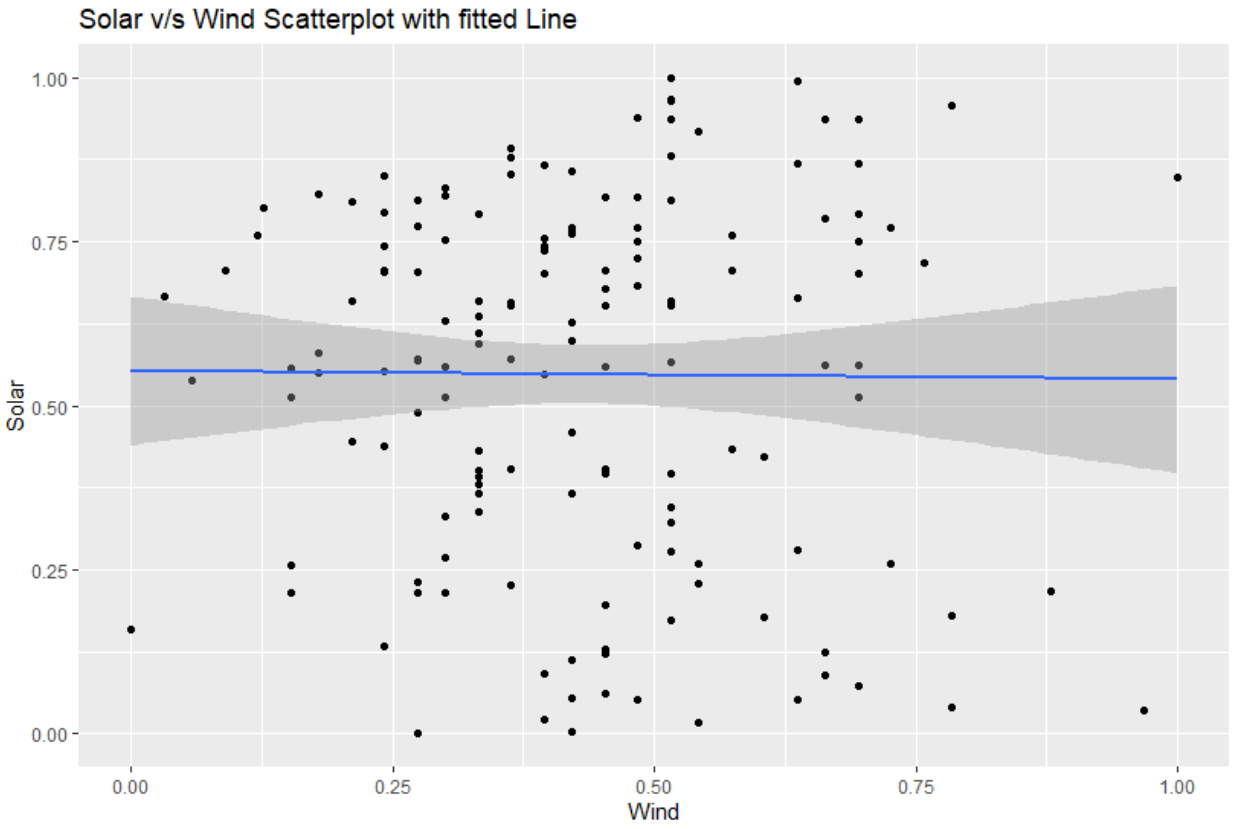
From 1900 to 1930 there was a sharp decrease in rate of population but later from 1930 to 1960 there is the greatest increase in population percentage wise. Later it keeps fluctuating after 10 years.

I have created the above graph in Tableau.

c. There was a population percentage increase greater than 15% from 1930 to 1960.

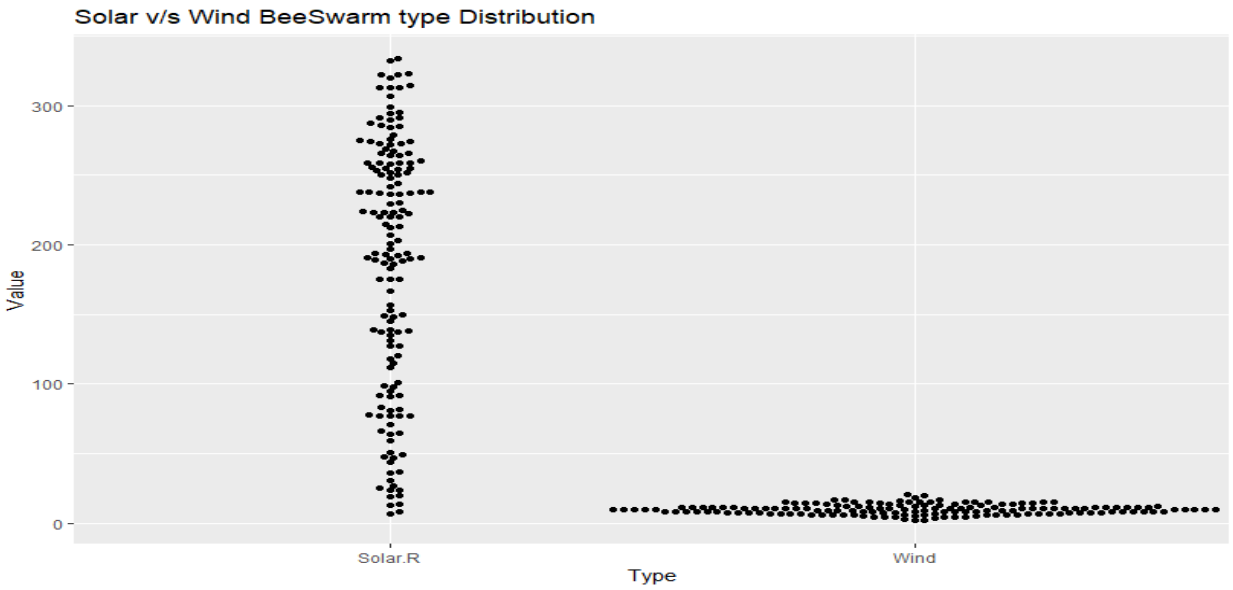
Q4.

a.



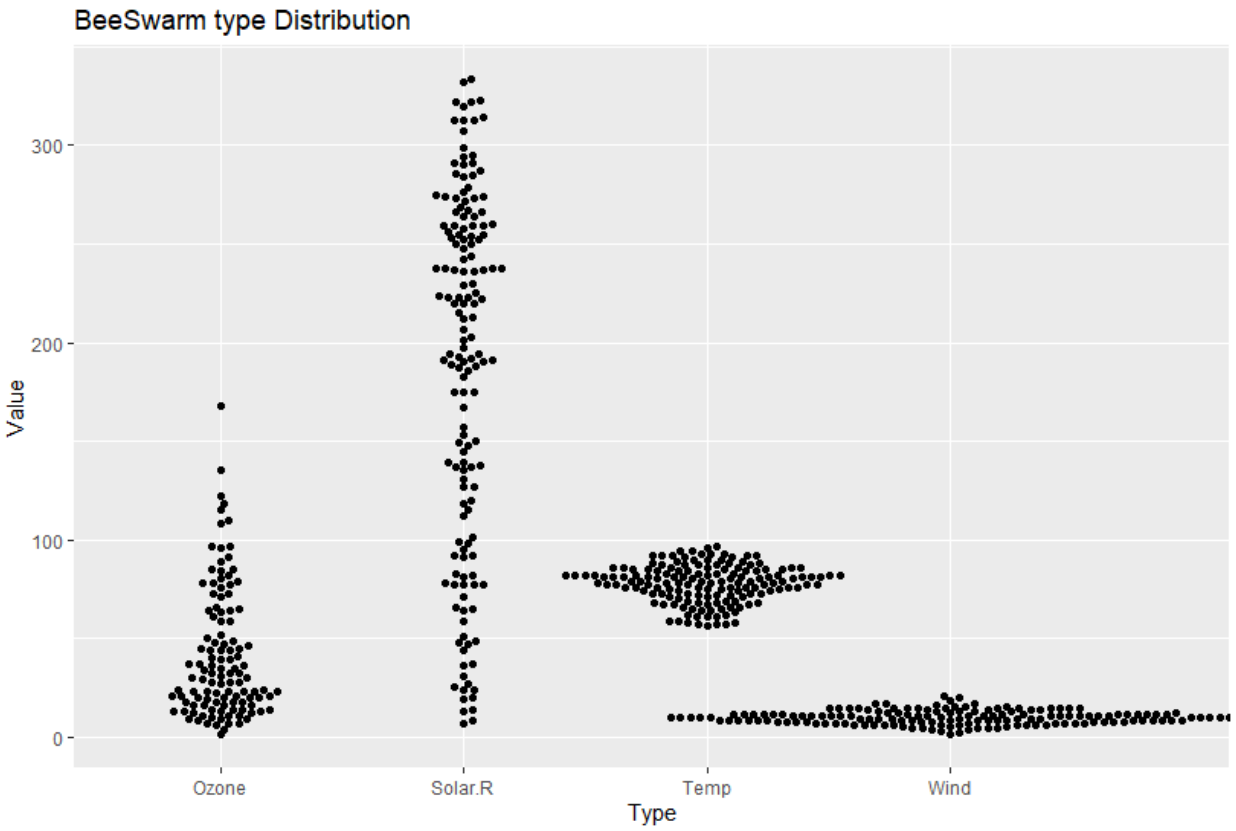
The above graph is the required graph made in R.

b.

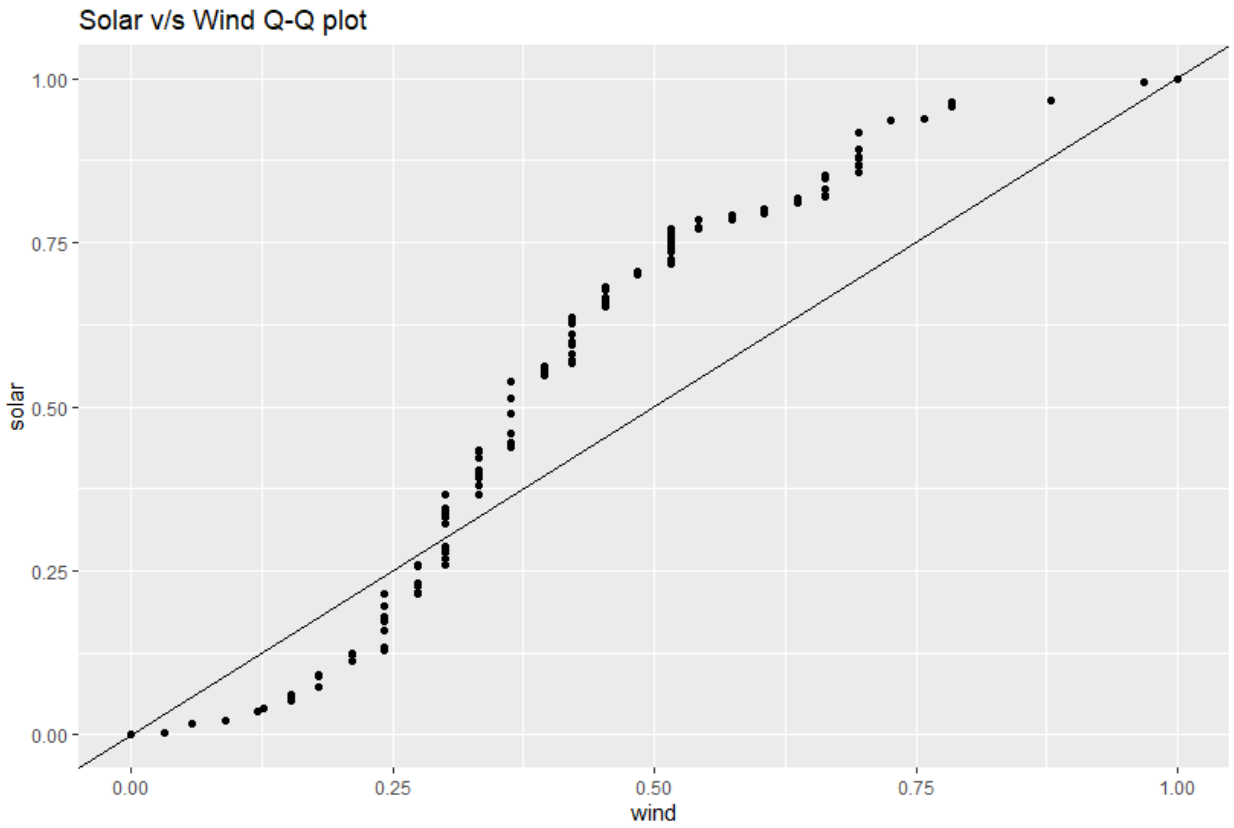


For this I have used bee-swarm plot on a new dataset which was created by pivoting the wind and solar datapoints from columns into rows.
As we can see from the graph that the values of wind are scattered because they have less value range from 0 to 25 unlike the solar whose value range is from 0 to 340. That's why solar shows a vertical line.

c.



d.



The above QQ plot is created between the wind and the solar. For the first quarter the distribution falls under the normal line and after that it goes above the normal line. Overall, we can say that the distribution is not normal for both wind and solar.