

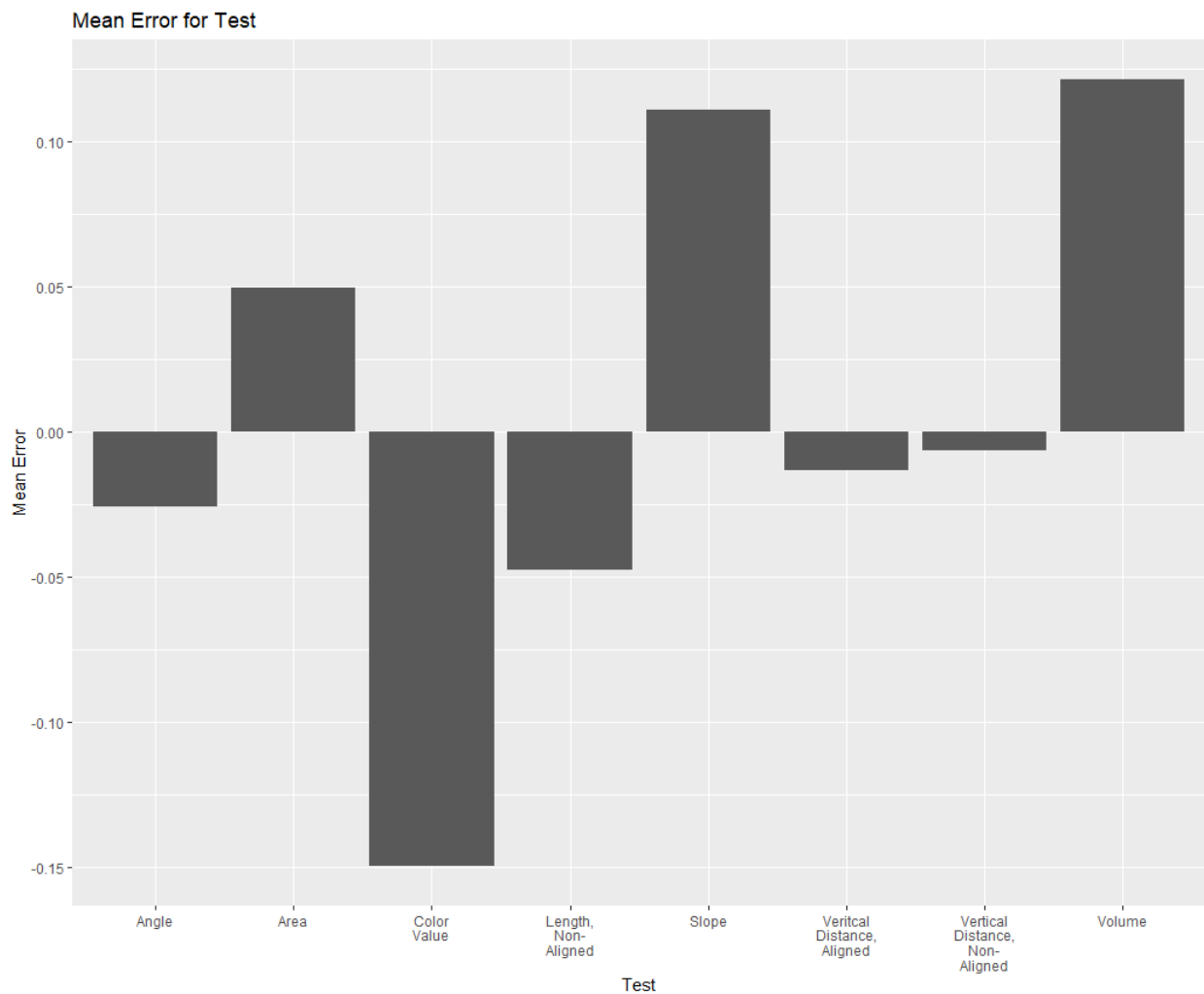
Azat Dovgeldiyev

DSC 465

Assignment 3

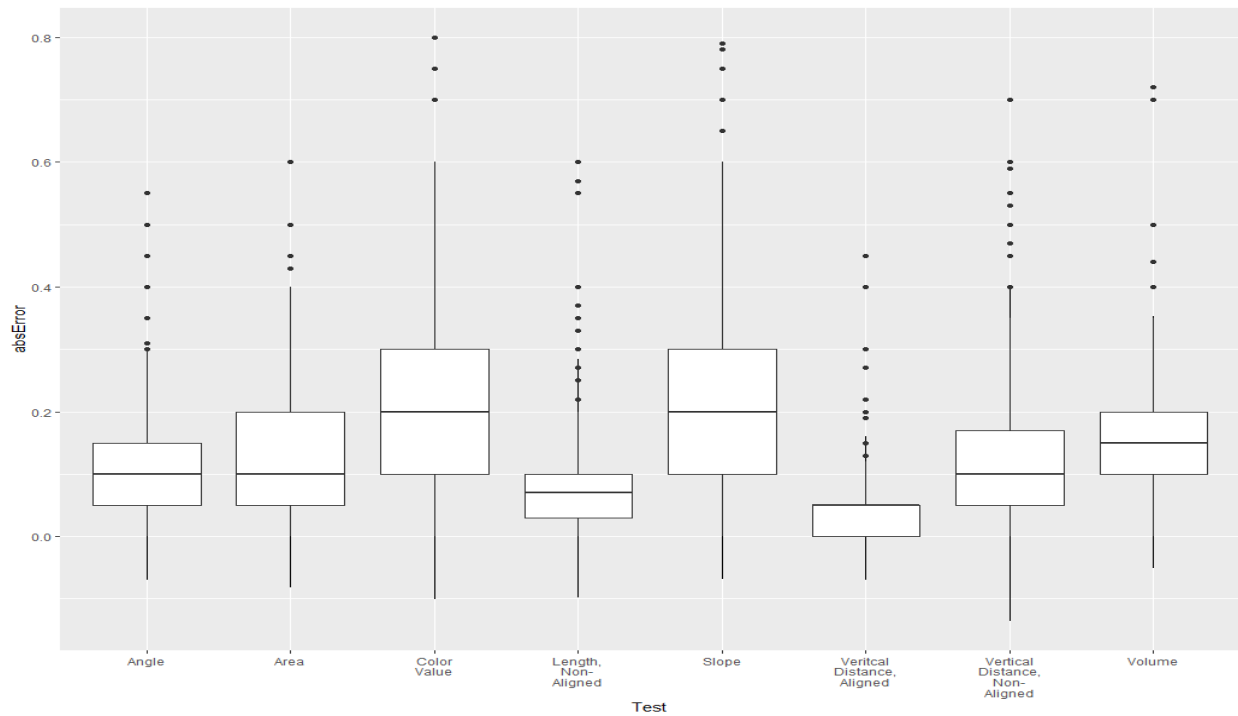
### Problem 1.

a)



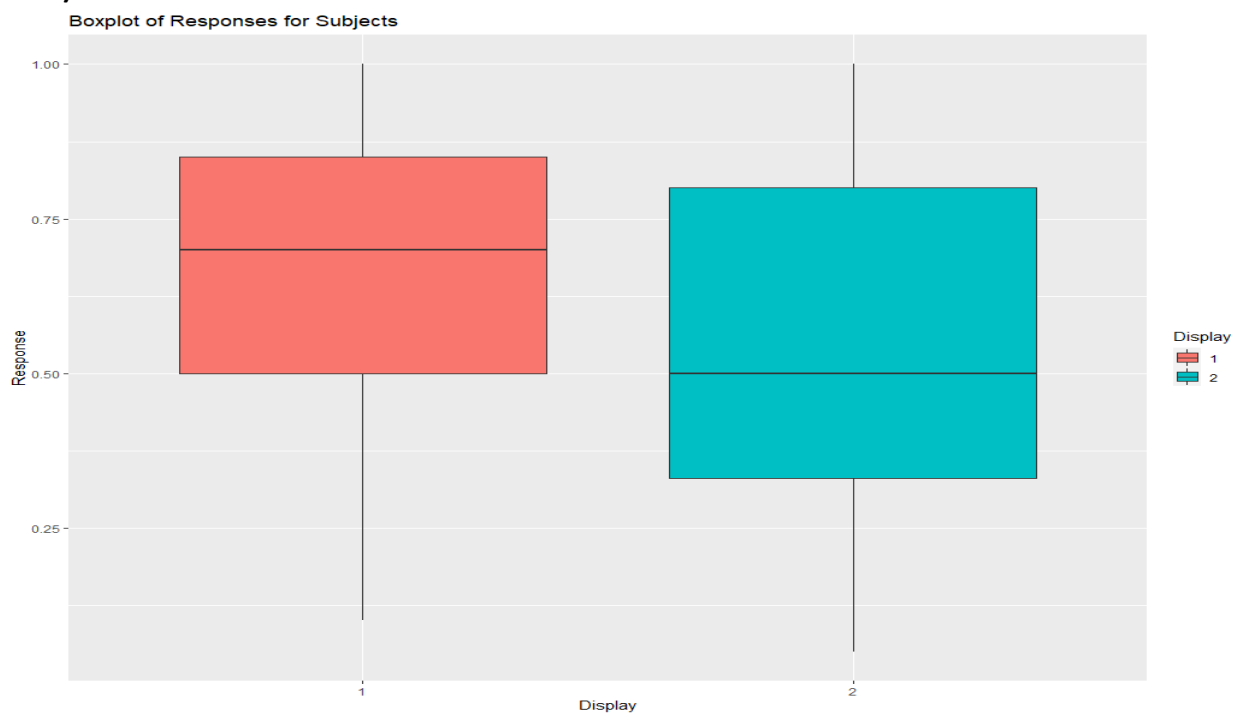
The above table shows the average error score (Response – True Value) for each test. Error scores state that any values below 0 is considered as underestimation, and obviously color value is the most underestimated value, whereas volume is the most overestimated. Because Position along aligned and nonaligned are the most reliable encoding for a human to be able to decode specific value their mean error values are close to 0.

b)



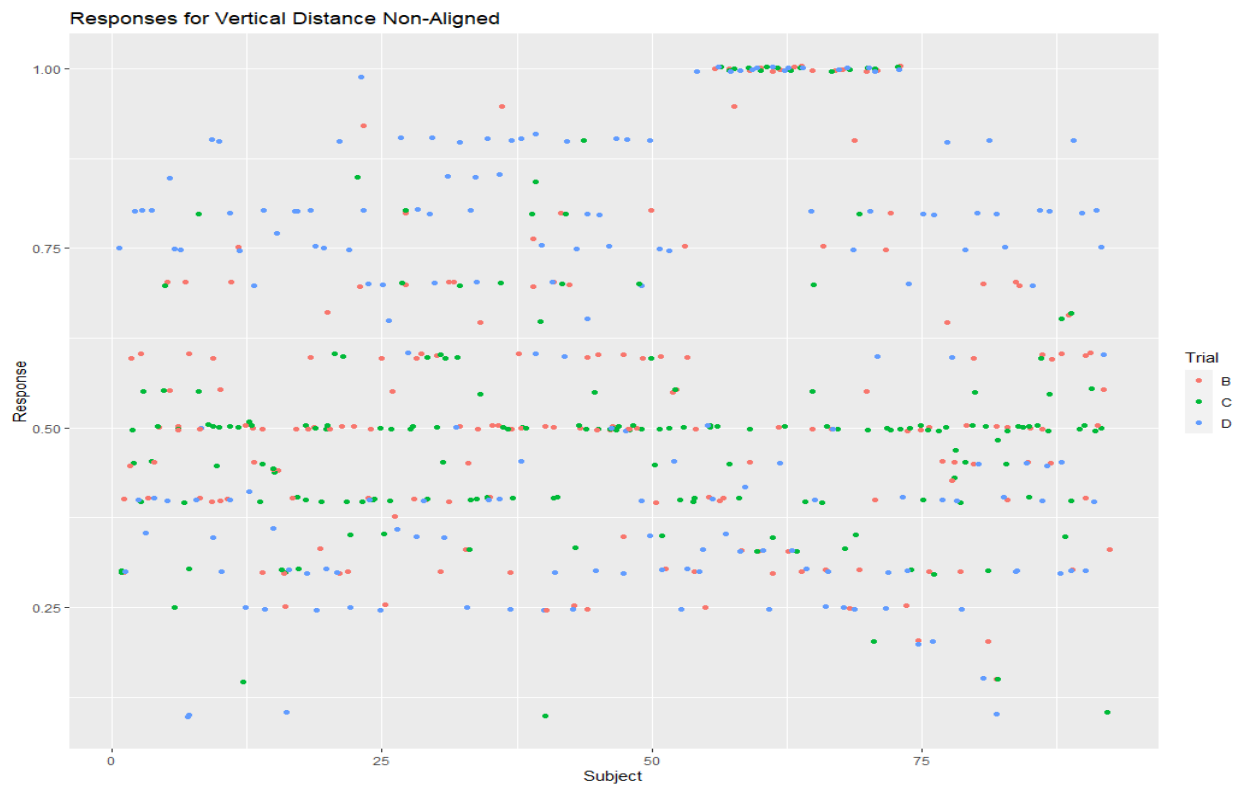
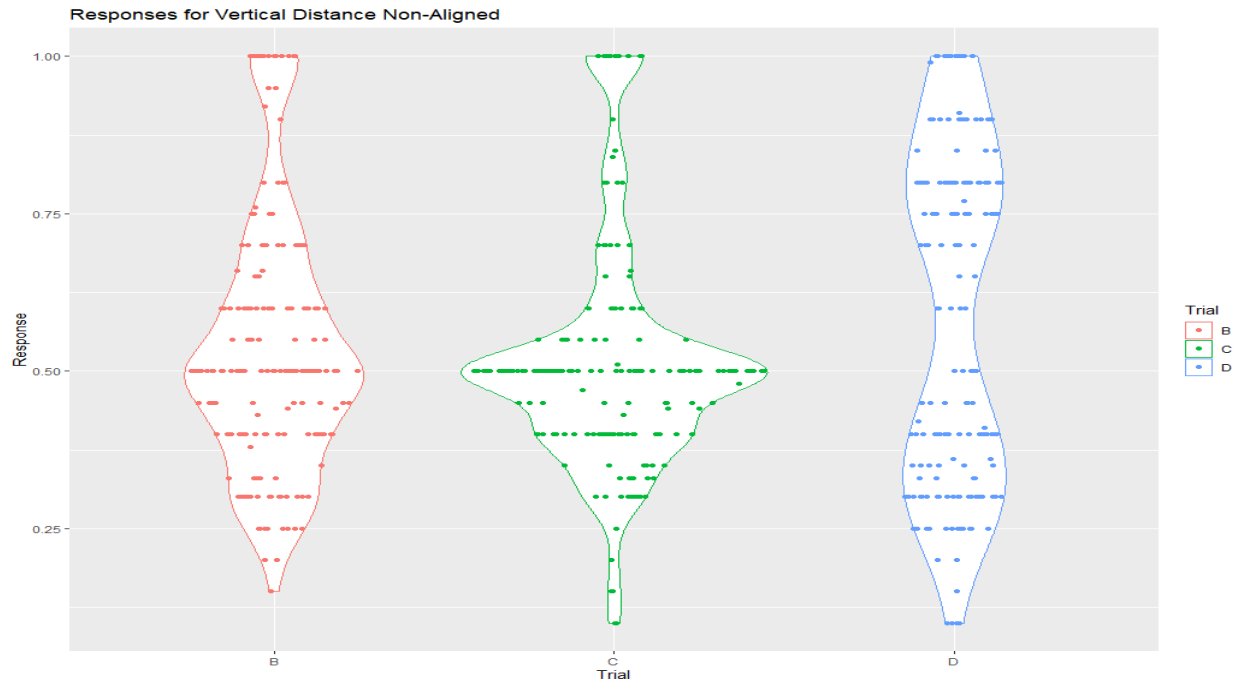
From the graph above we can see that vertical distance aligned lies between 0.1 and 0.2, which means it has the smallest distribution, also length non-aligned with vertical distance non-aligned are in the lower range. Color value, slope shares most points in the upper range.

c)



Subjects have lower values on the second display where median value falls in 0.5, while 1<sup>st</sup> display has greater values. And it tells that the subjects in given range had better responses on the second display.

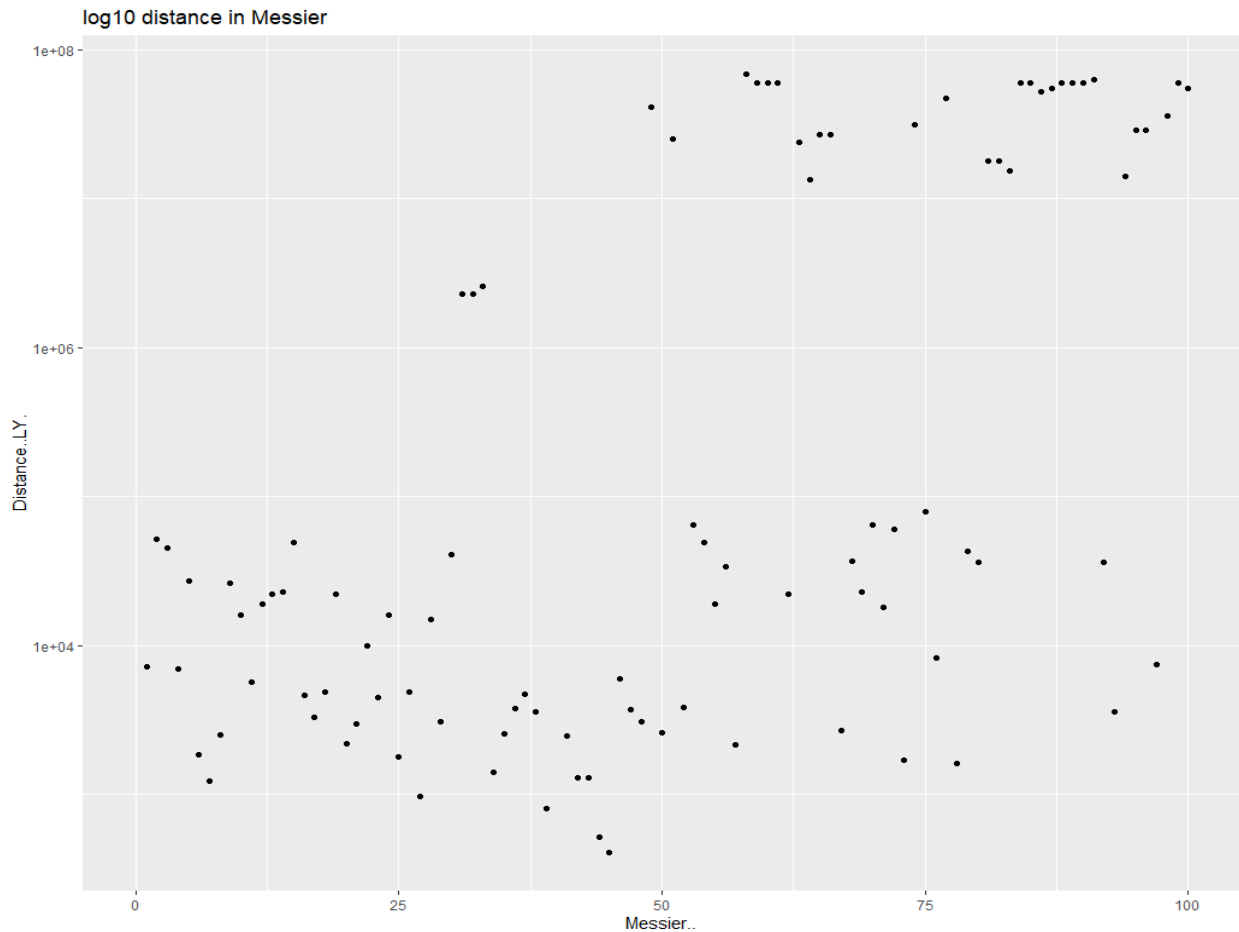
d)



Filtered graphs above are based on Vertical Distance Non-Aligned, 1<sup>st</sup> violin plot shows Responses for each Trial, and each trial differs. Scatterplot in the second graph highlights the anomaly, Responses for Subject, and points are on top of each other and they might be mapped according to Trial D.

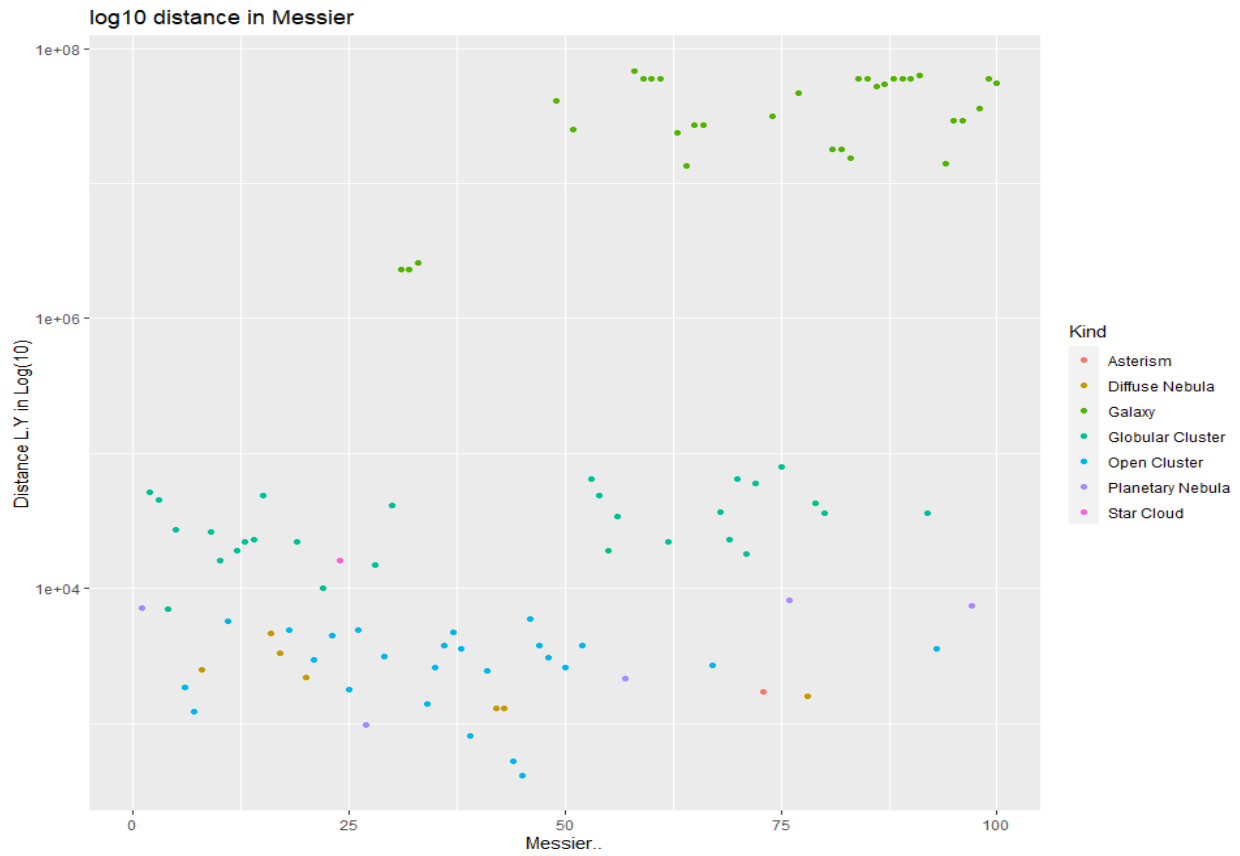
## Problem 2.

a)



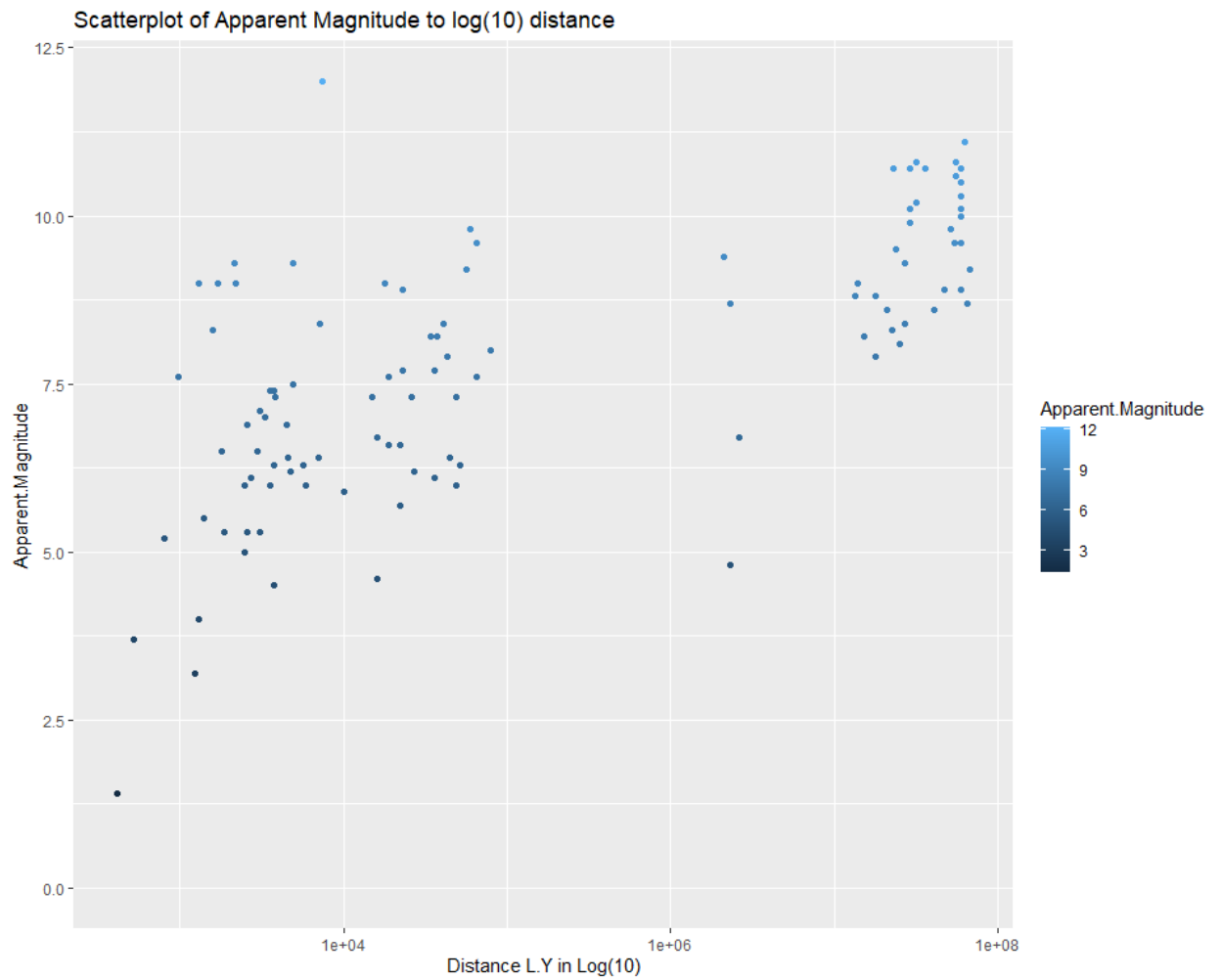
Because the Distance in Light Years gives very large values, it is reasonable to use Log base 10 (astronomy). It appears that points divided into two clusters lower and upper values, far distance points are consistent starting from Messier number range 50.

b)



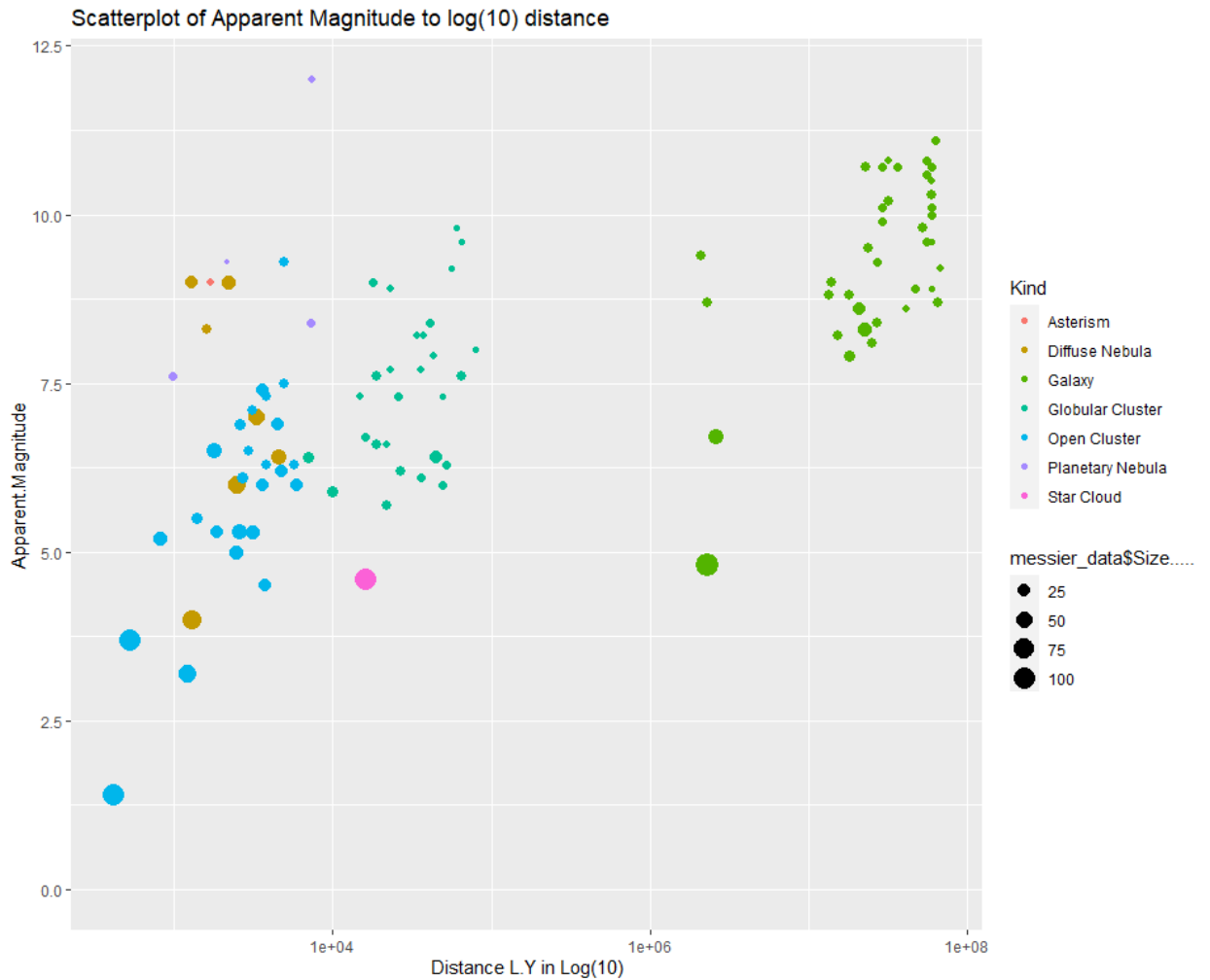
The Galaxy has the furthest distance for its center, following Globular Cluster with the second furthest distance has the widest spread, and open cluster has the shortest distance, so there are mainly 3 clusters.

c)



Scatter plot with Apparent magnitude to log base 10 distance shows that the higher the number the fainter the object is in the sky. The maximum value of Apparent Magnitude is 12.5.

d)

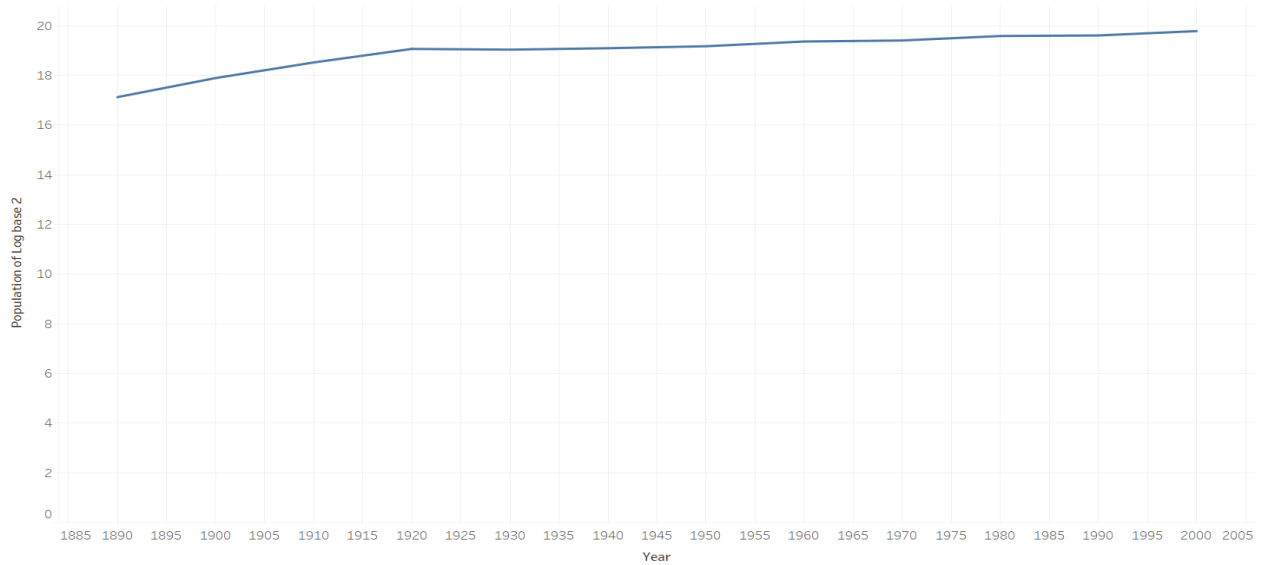


The graph makes it very easy to analyze all aspects of the graph. We can visualize that the furthest objects are fainter and smaller in size and belong to the Galaxy, the objects that closer in distance are brighter and relatively bigger in size, mostly belong to the open cluster. Overall scatterplot shows clear visualization in terms of encoding.

### Problem 3.

a)

Population Growth of Montana

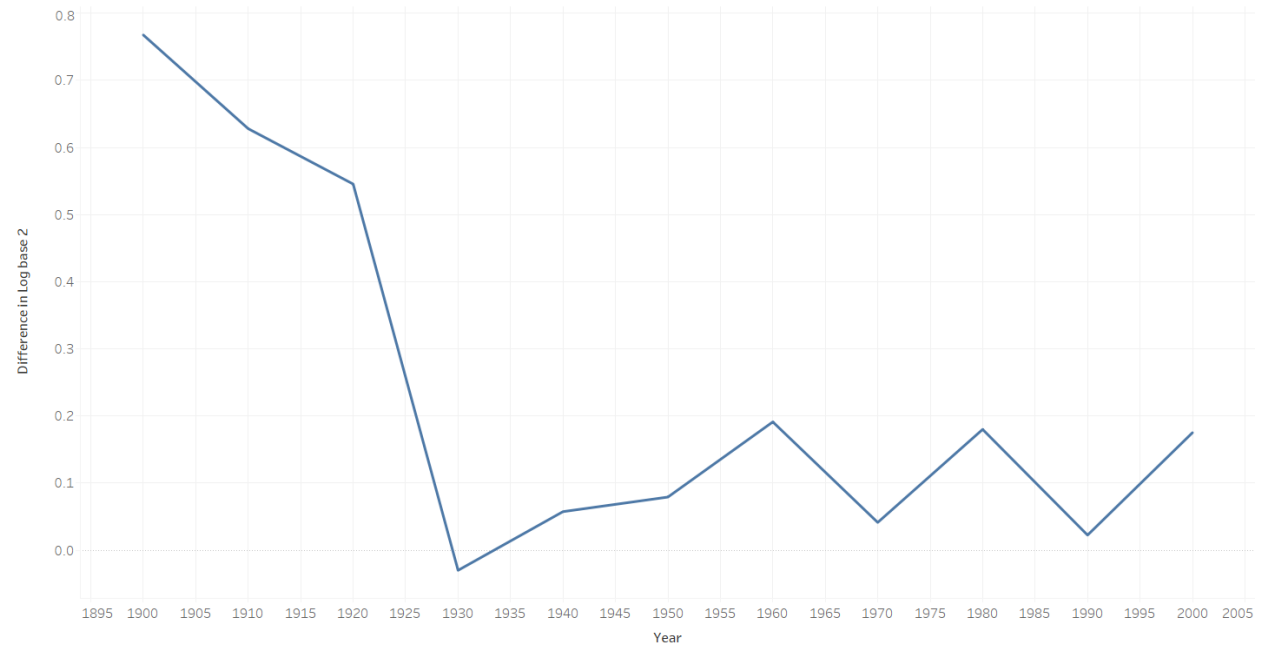


The trend of sum of Log base 2 for Year.

The population doubled (2.5 times) since 1890 using log base 2.

b)

Difference in growth by years



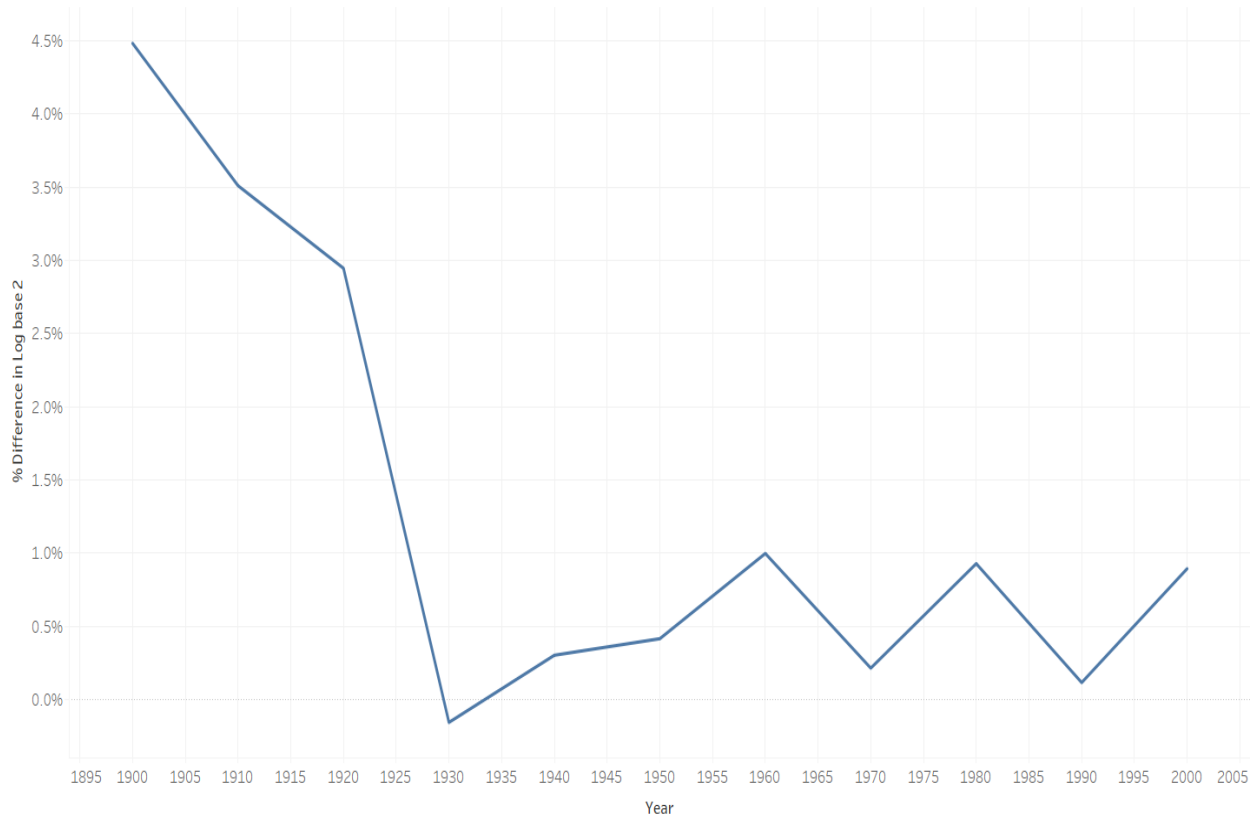
The trend of Difference in Log base 2 for Year. The view is filtered on Year, which ranges from 1890 to 2000.

The percentage rate of change increased and decreased over the years, however rate of change significantly decreased from 1900 to 1930. The biggest percentage wise in years 1900, 1910, and 1920.



c)

Difference in growth by years

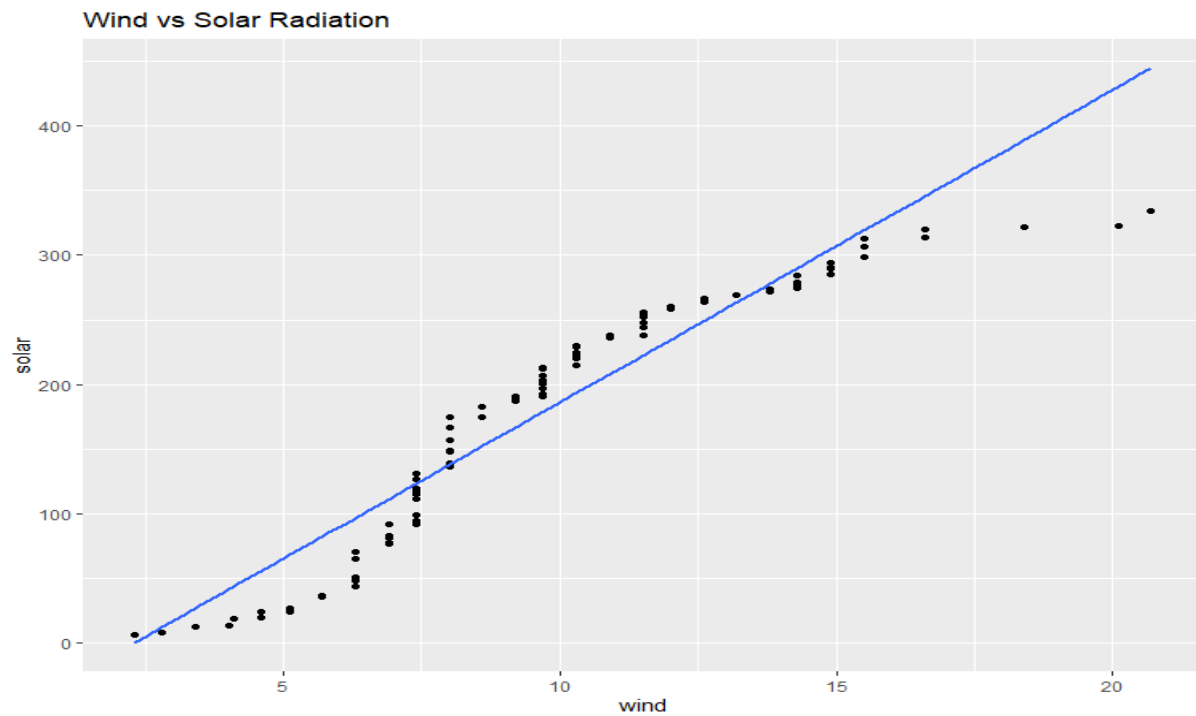


The trend of % Difference in Log base 2 for Year. The view is filtered on Year, which ranges from 1890 to 2000.

**Population percentage increase greater than 15% was in 1900, 1910 and 1920. Percentage increases were just below 15% in 1960, 1980 and 2000.**

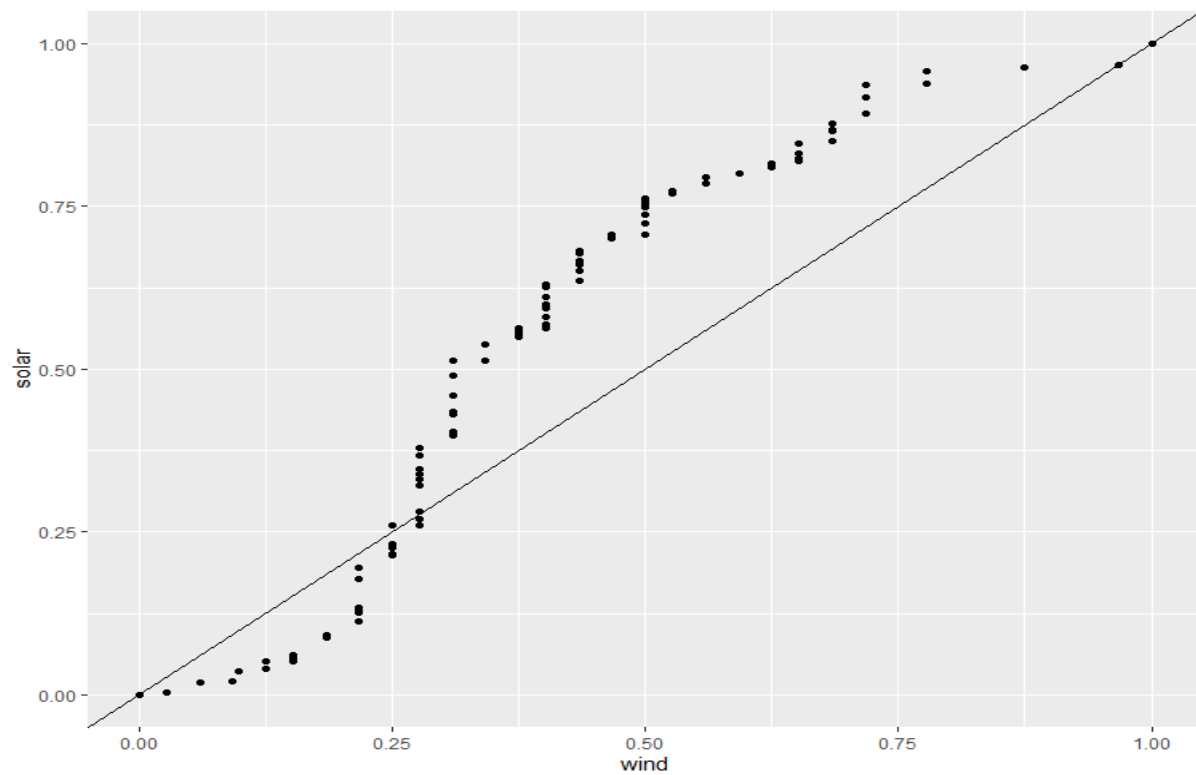
#### Problem 4.

a)



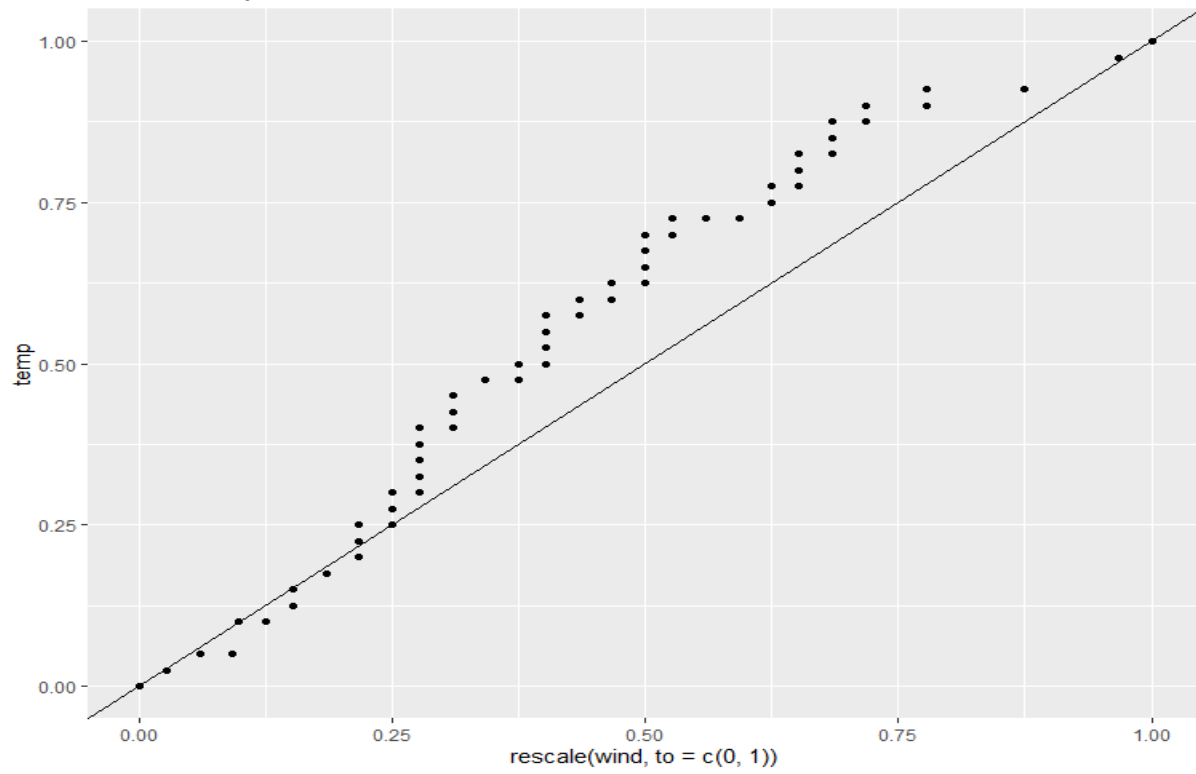
As the solar radiation increases wind increases.

b) Fine detailed distribution show axis scale ranges between 0 and 1, with slope of a plot.

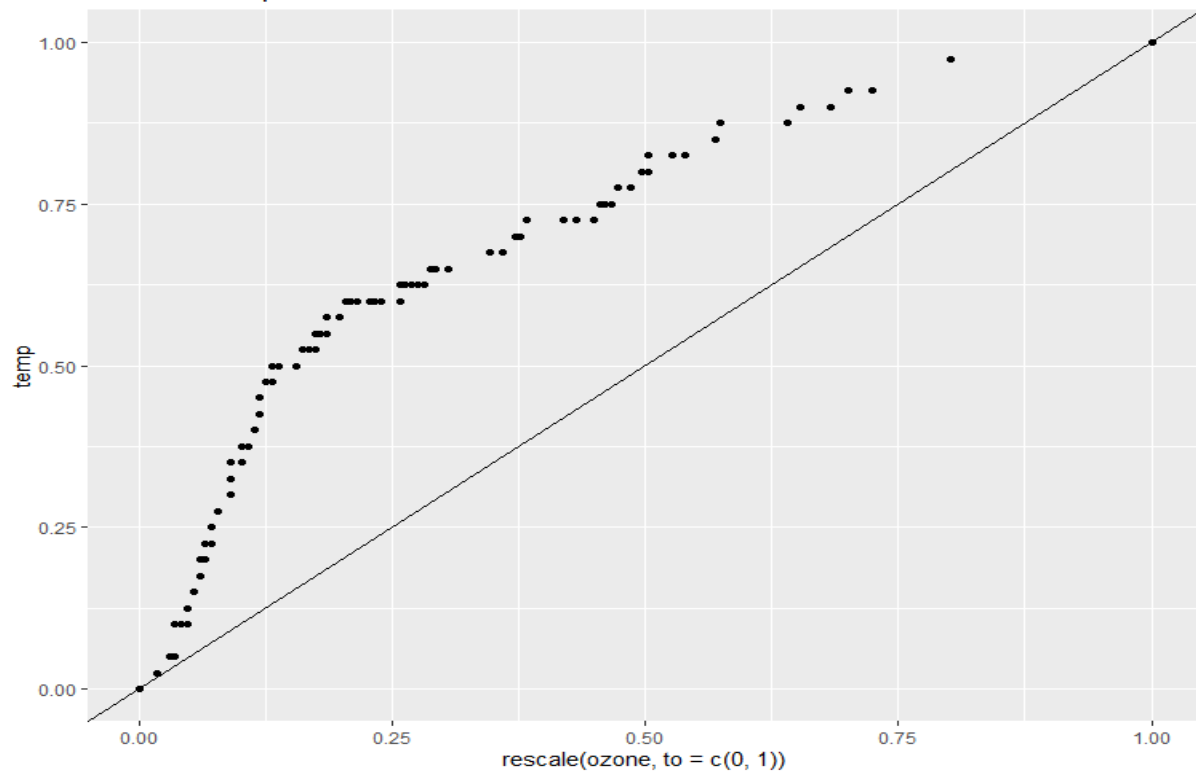


c)

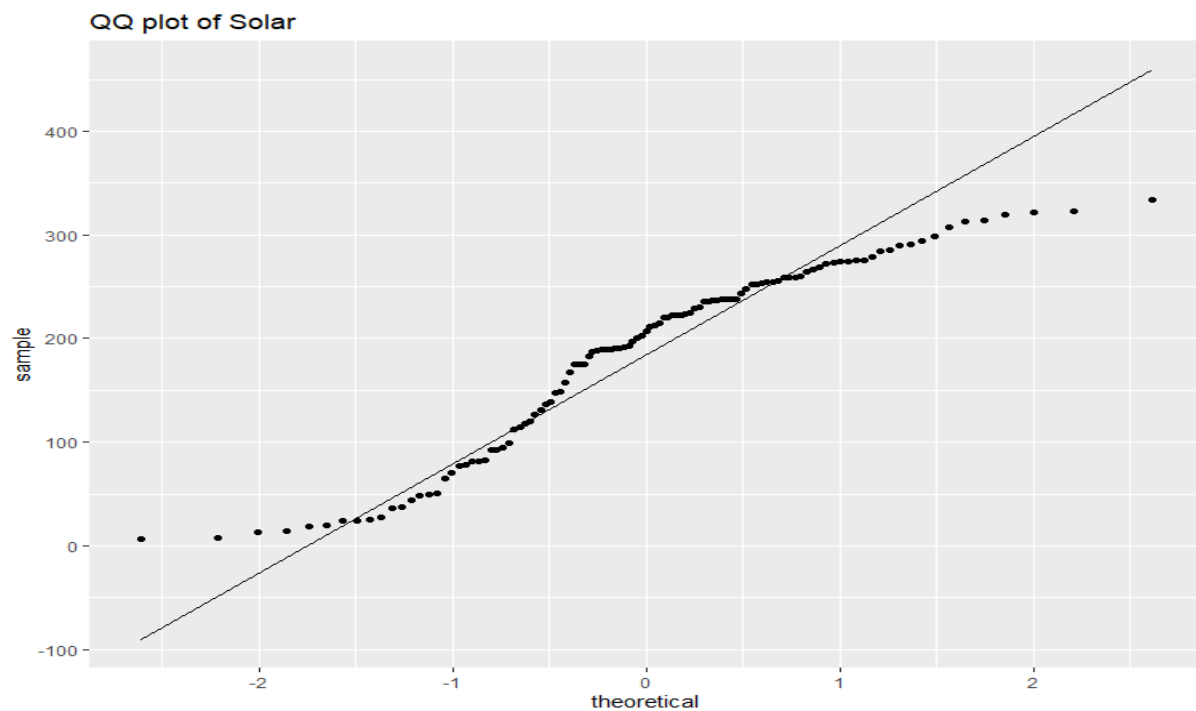
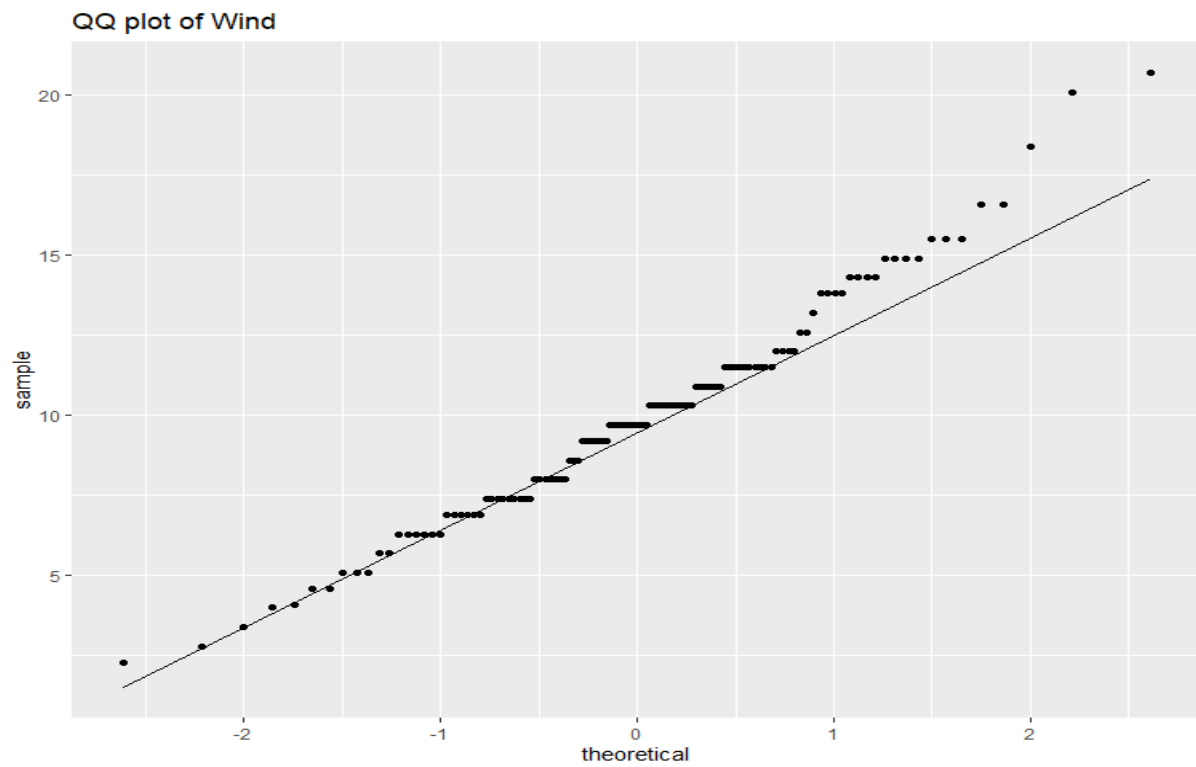
Wind vs temp



Ozone vs temp



d)



The points in QQ plot of Wind variable seem to fall about a straight line, and the linearity of the points suggests that the data are almost normally distributed, whereas qqplot of solar wind is not normally distributed, as the upper and below points do not fall into the straight line.