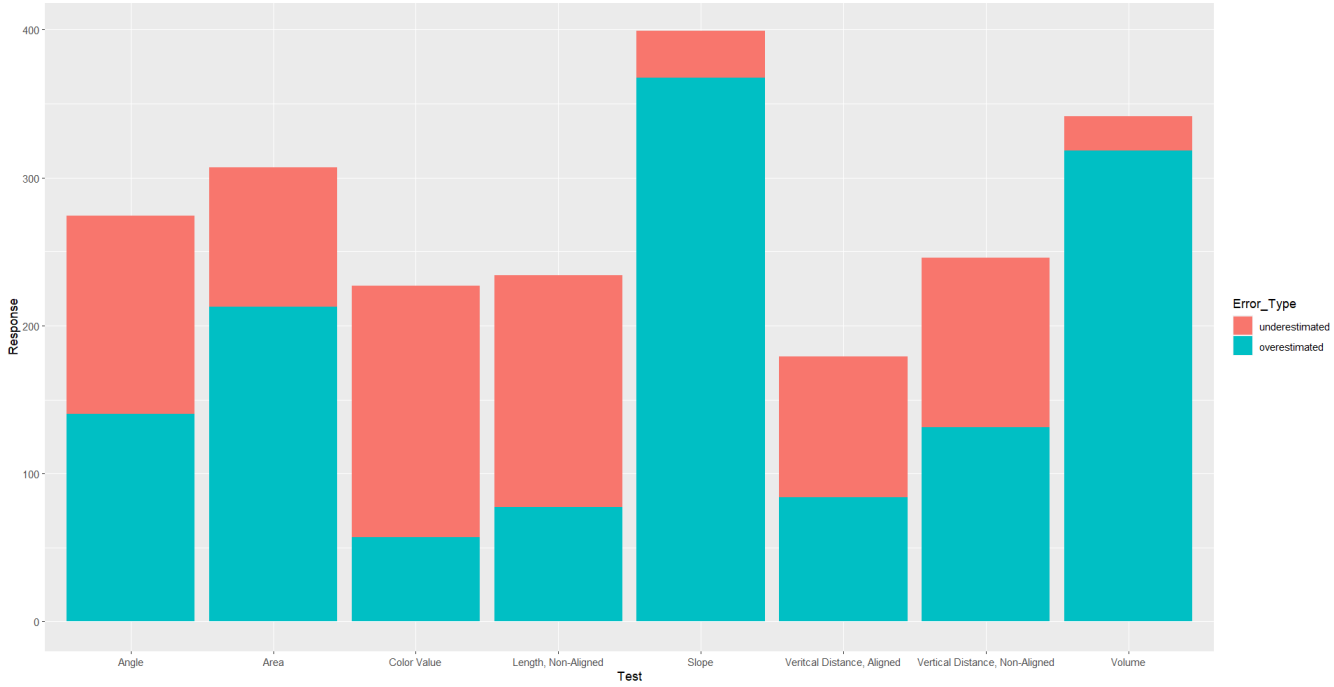
DV_Assignmnet_3

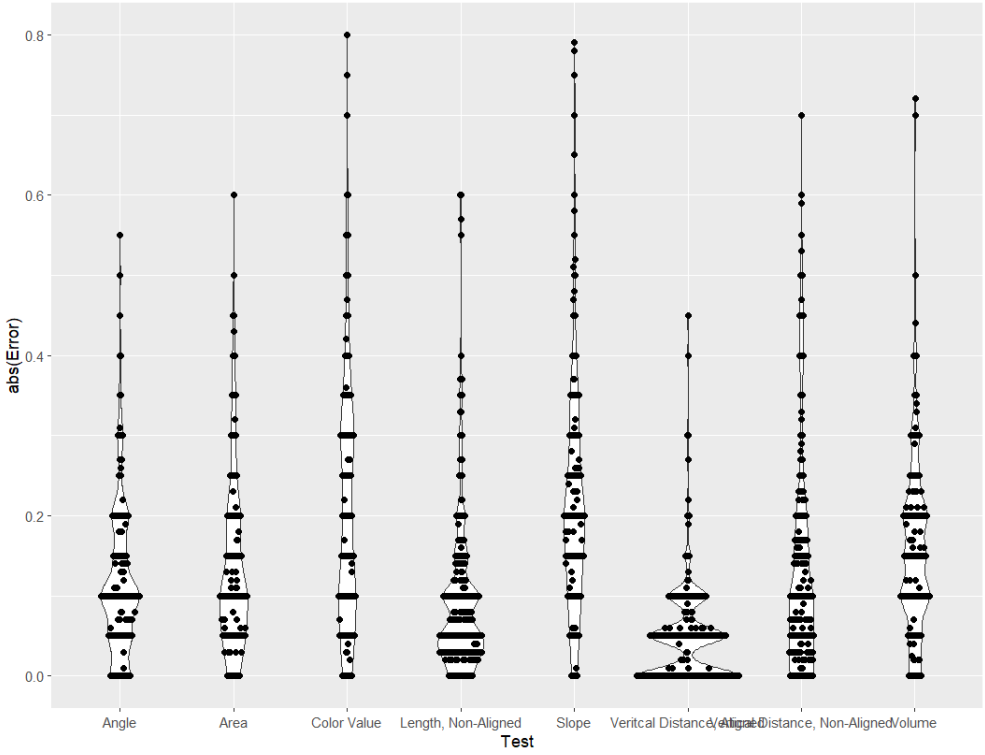
ADITHYA HARSHA

1a. Were there any tests where people generally underestimated or overestimated the data? Explain what field you can graph to test this, what graphical method reveals this clearly. Analyze the results and explain in a short paragraph.



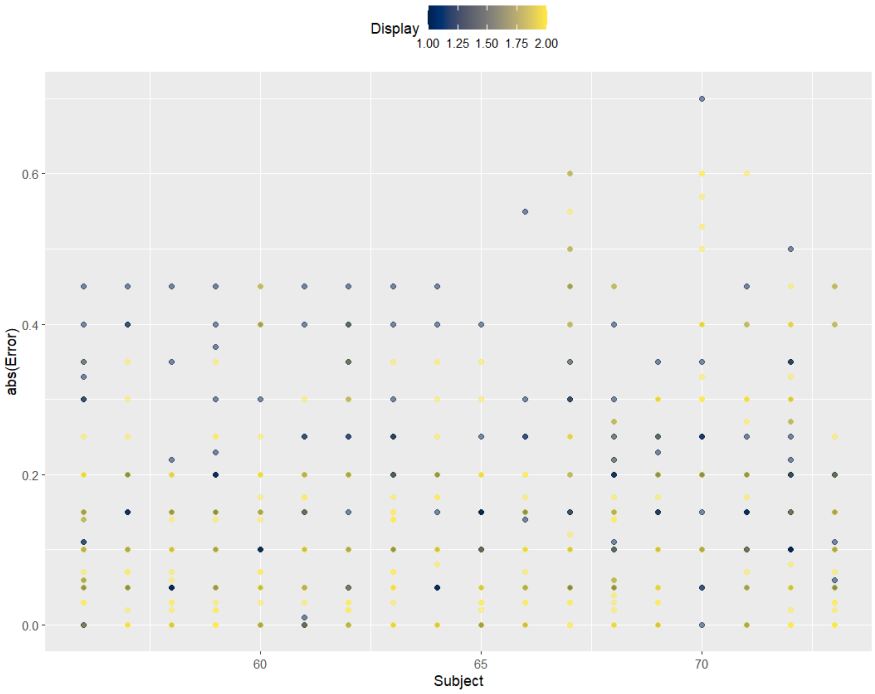
A column chart is a great way to visualize the range of overestimation and underestimation in each test. By analyzing the chart, we can easily identify which tests exhibit the highest levels of overestimation and underestimation. From the chart, we can observe that the "Slope" and "Volume" tests have the highest levels of overestimation.

1 b)



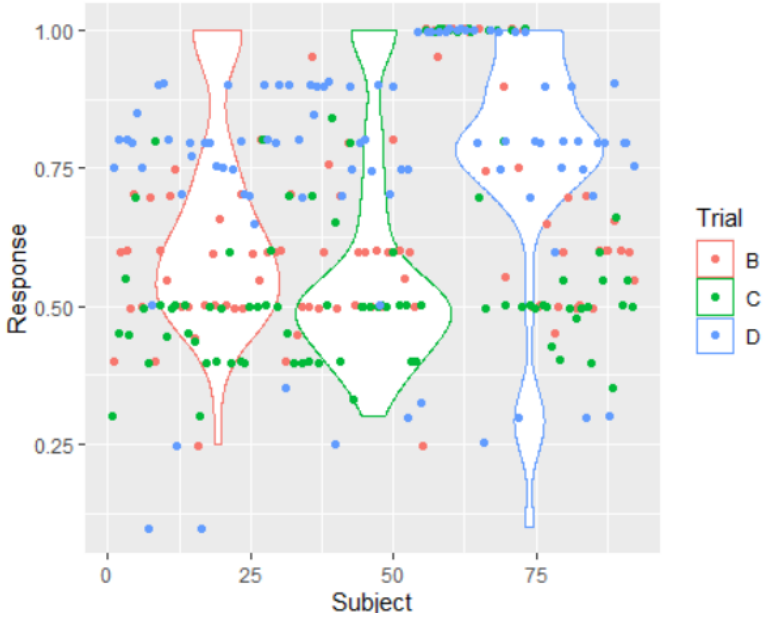
All tests have outliers and similar ranges of errors, except for "Vertical Distance, Aligned" which has a lower range of errors. This information can be important for understanding the variability and performance of the tests.

1 c)



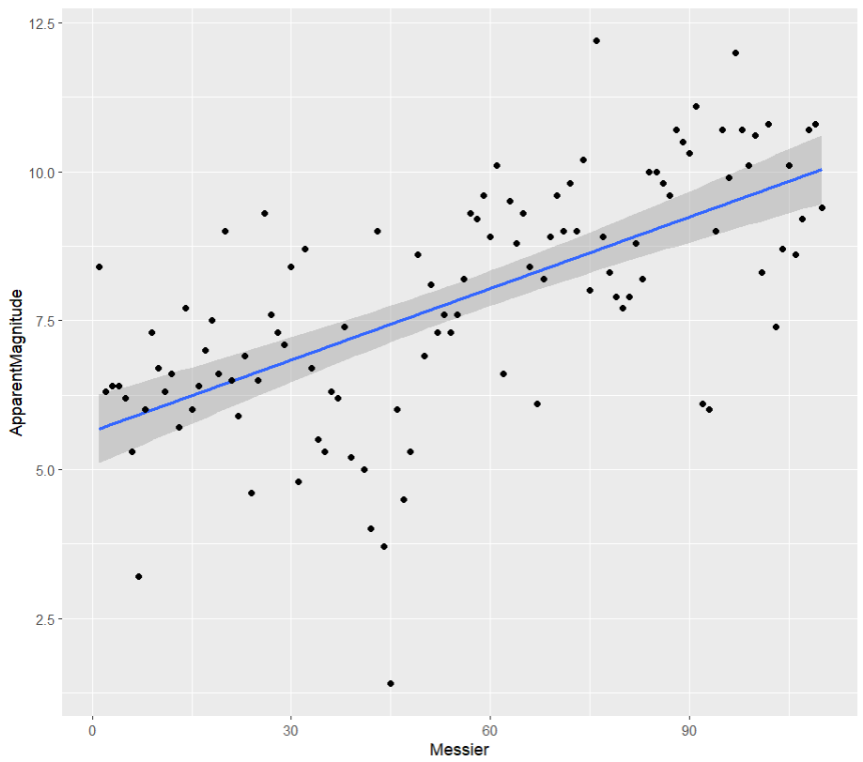
A univariate plot can be used to display the range of errors for each subject across various displays. The subjects seem to struggle with display 2 as well and this is more prominent with subjects greater than 50

1 d)

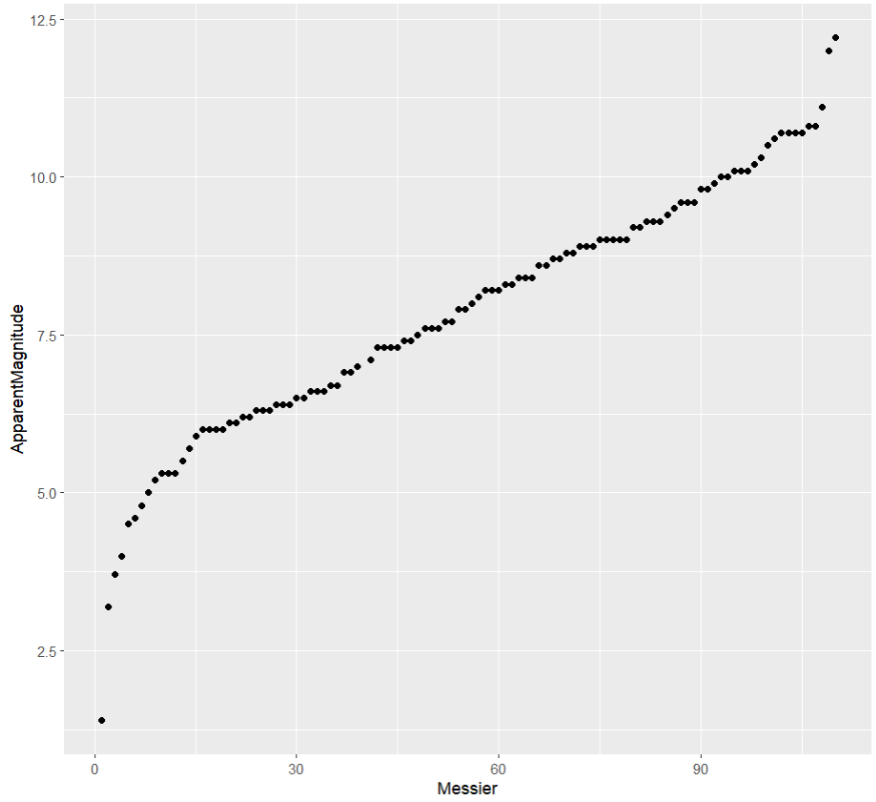


We can use a violin plot to examine the distribution of responses and identify patterns across different subjects. From the plot, it is evident that the anomalous sequence of "1" responses is more prevalent among subjects over the age of 55.

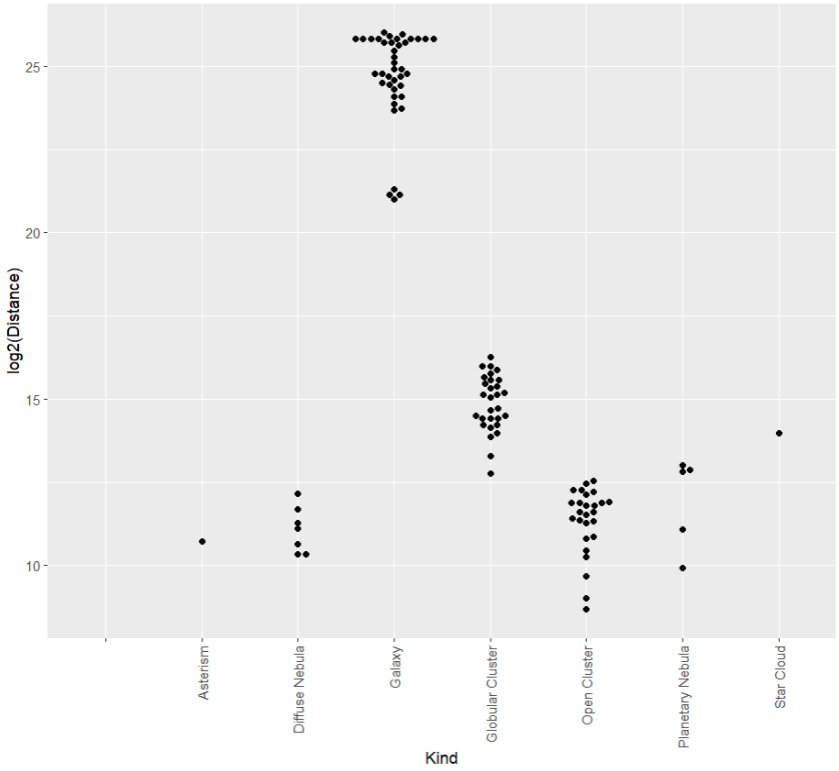
2 a)



Based on the QQ plot below, there appears to be a moderately positive correlation between Messier number and Apparent Magnitude. Below I have performed qq plot to compare the distribution of messier and apparent magnitude

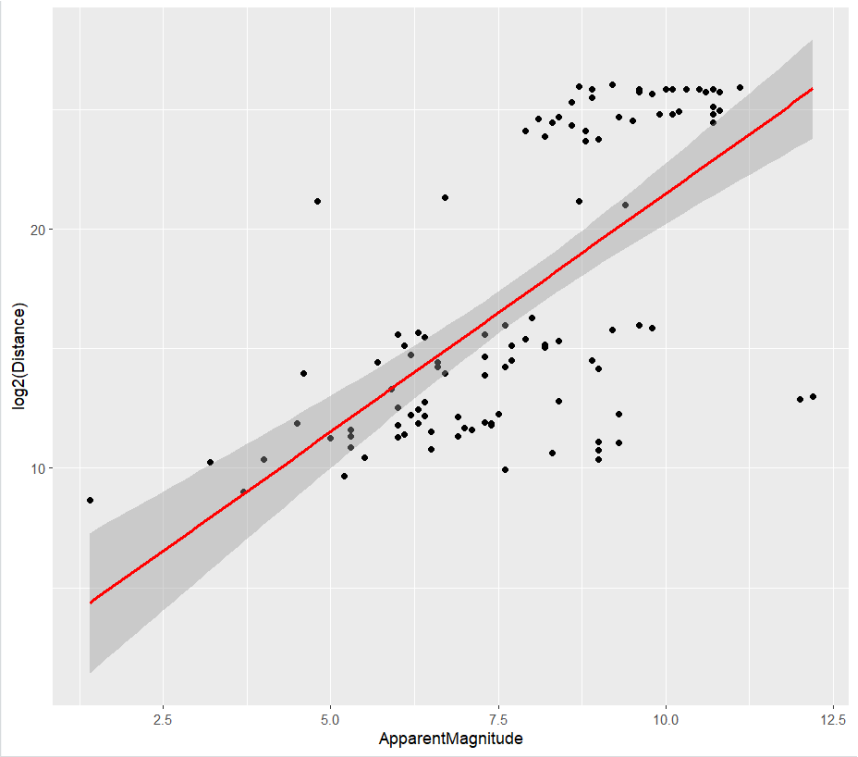


2 b)

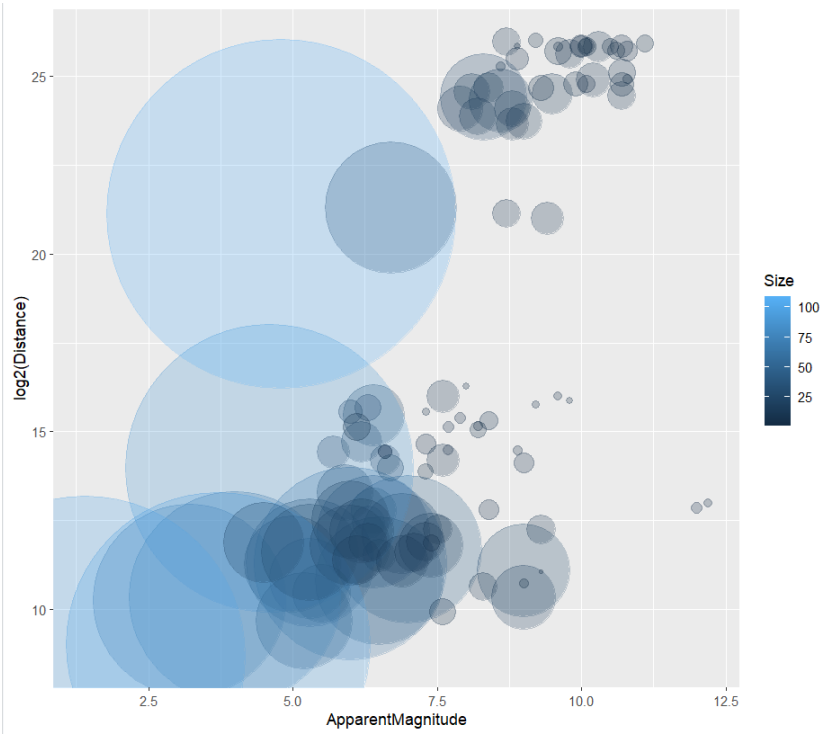


Using beeswarm plot we can visualize the distribution of kind w.r.t to log2(Distance) Galaxy has higher distance

2 c)

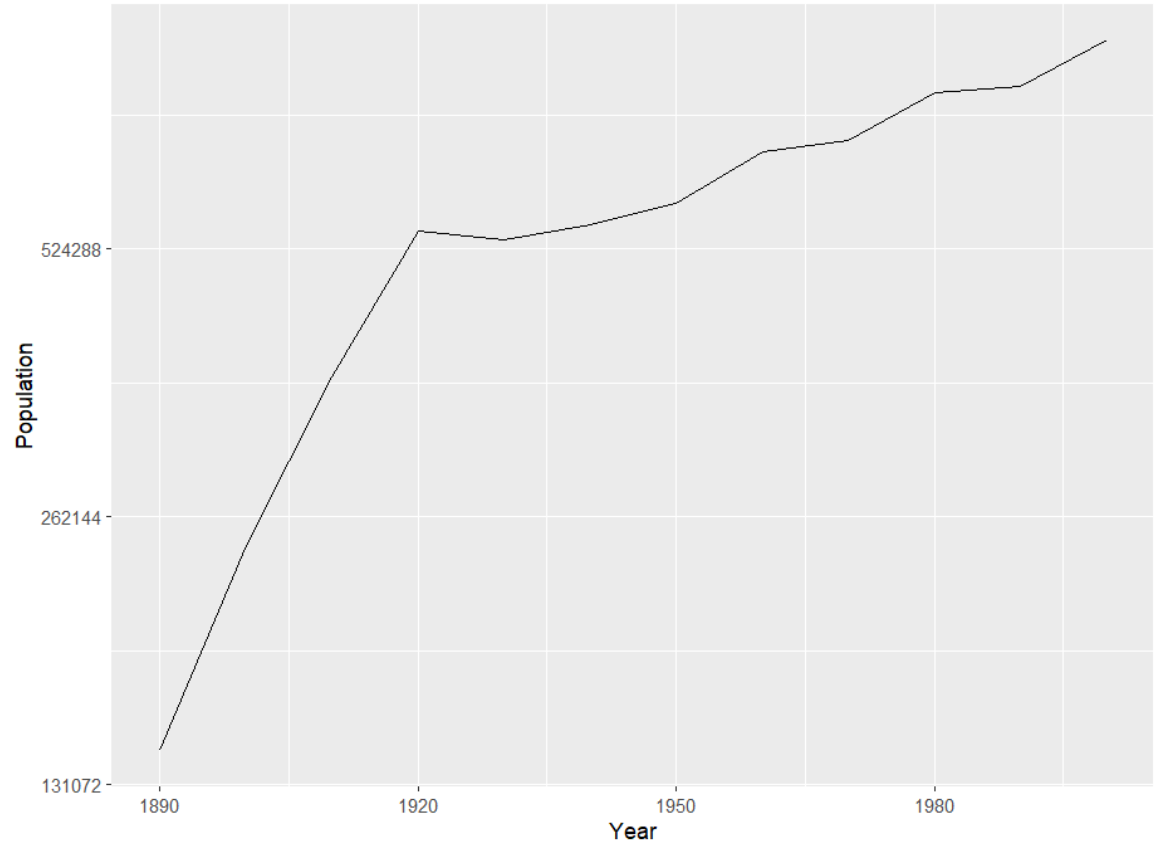


2 d)



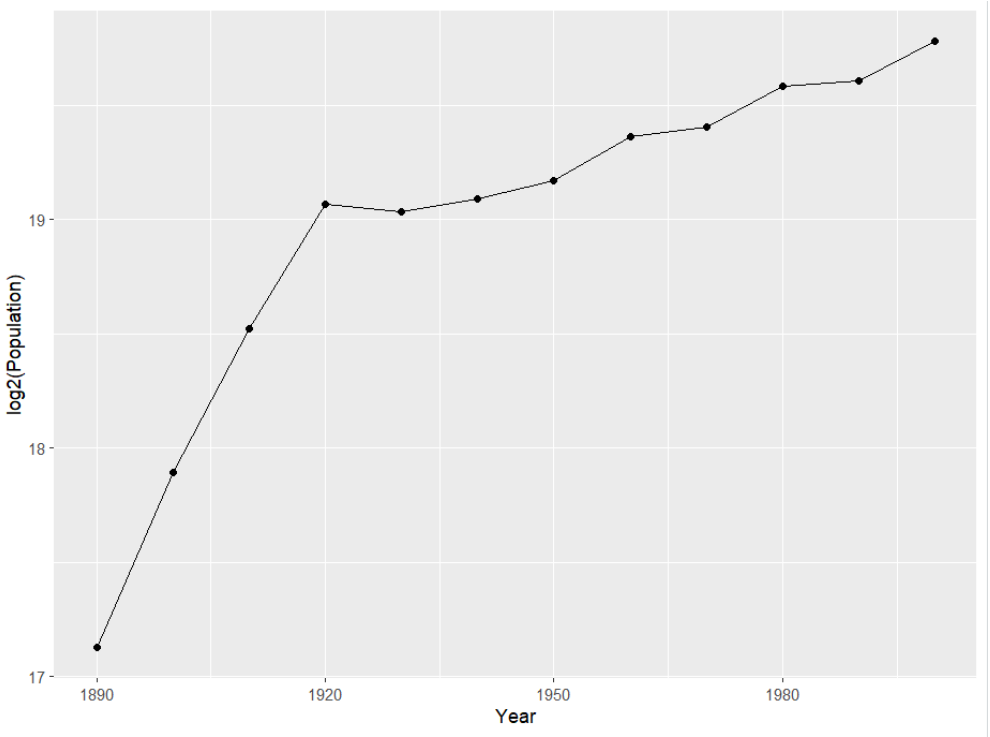
3a. How many times has the population doubled since 1890?

3a answer: The population has doubled twice since 1890 to 1920

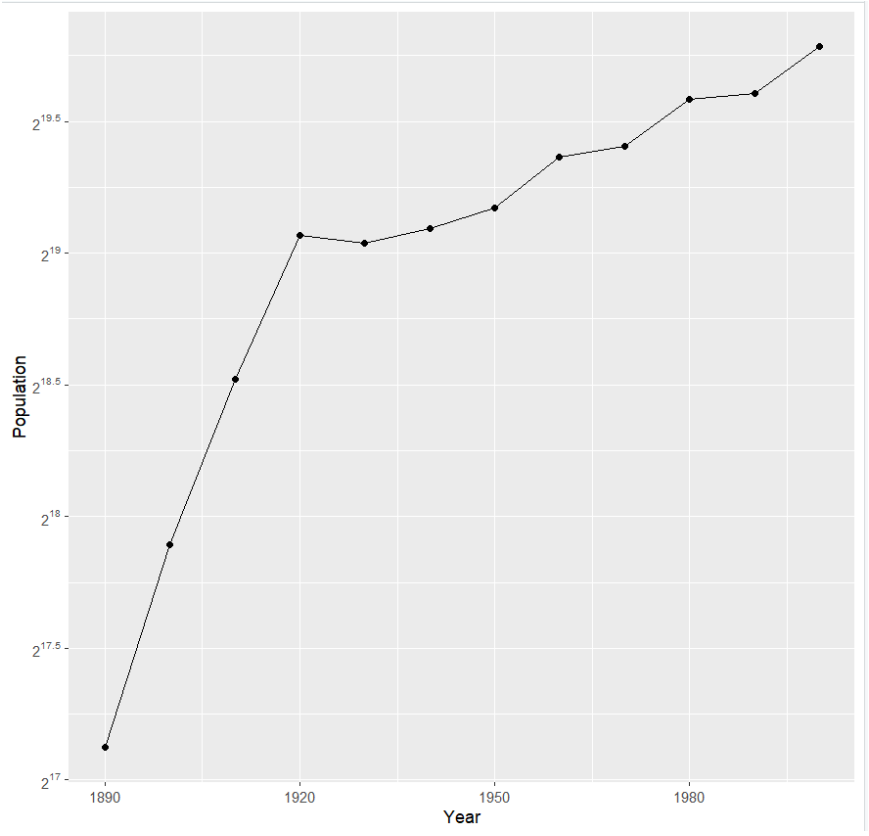


3b. Has the percentage rate of change in the population increased or decreased over the years? What years had the greatest increase in population %-wise?

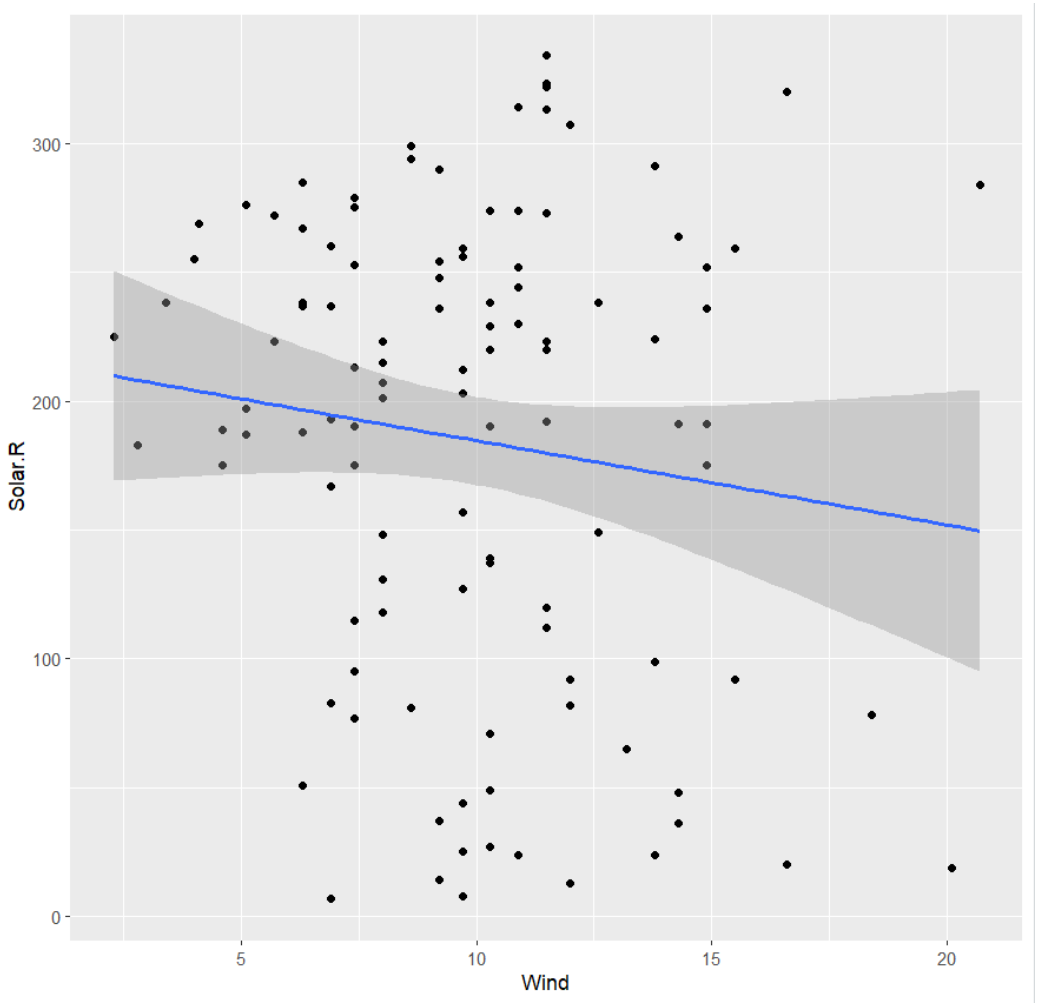
3b answer: The population doubled twice from 1890 to 1920 and then there was a decline in the population rise.



3c answer: 1890 to 1920 and from 1950 to 1980 had a population increase greater than 15%

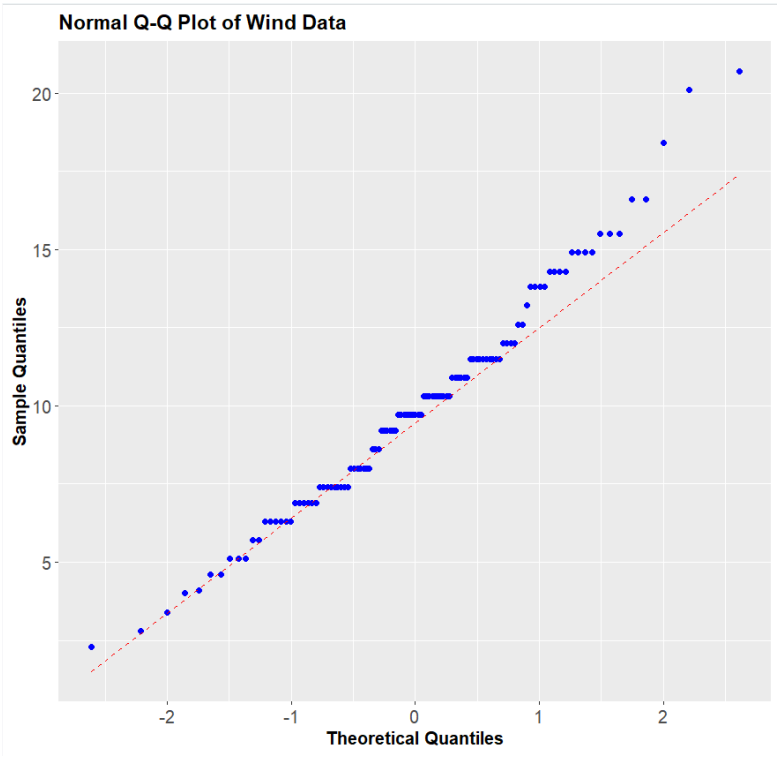


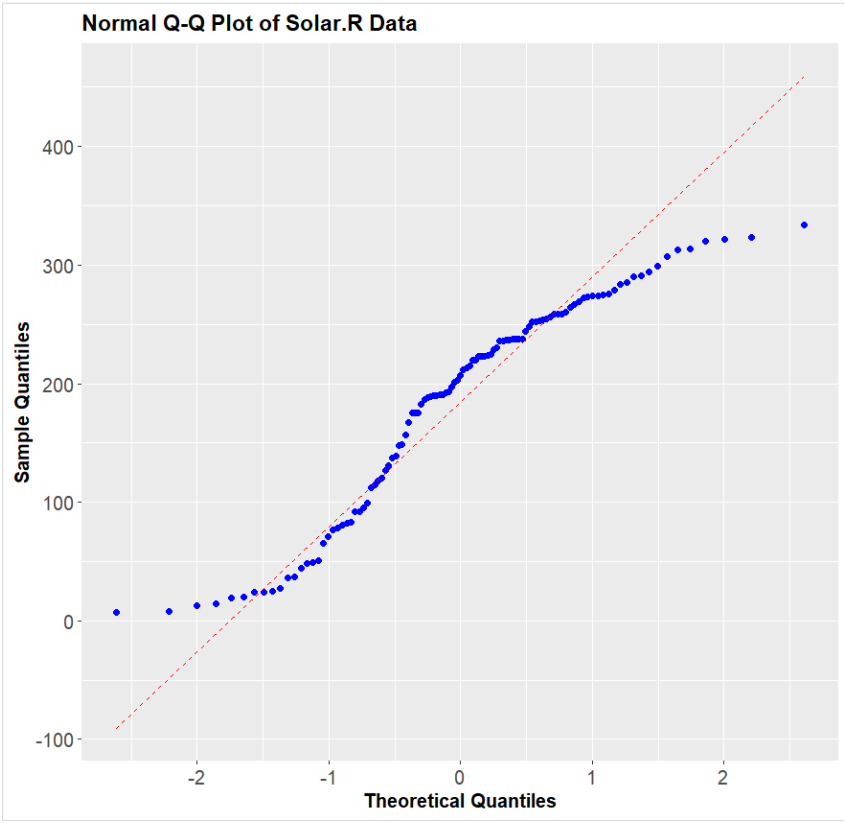
4) a)

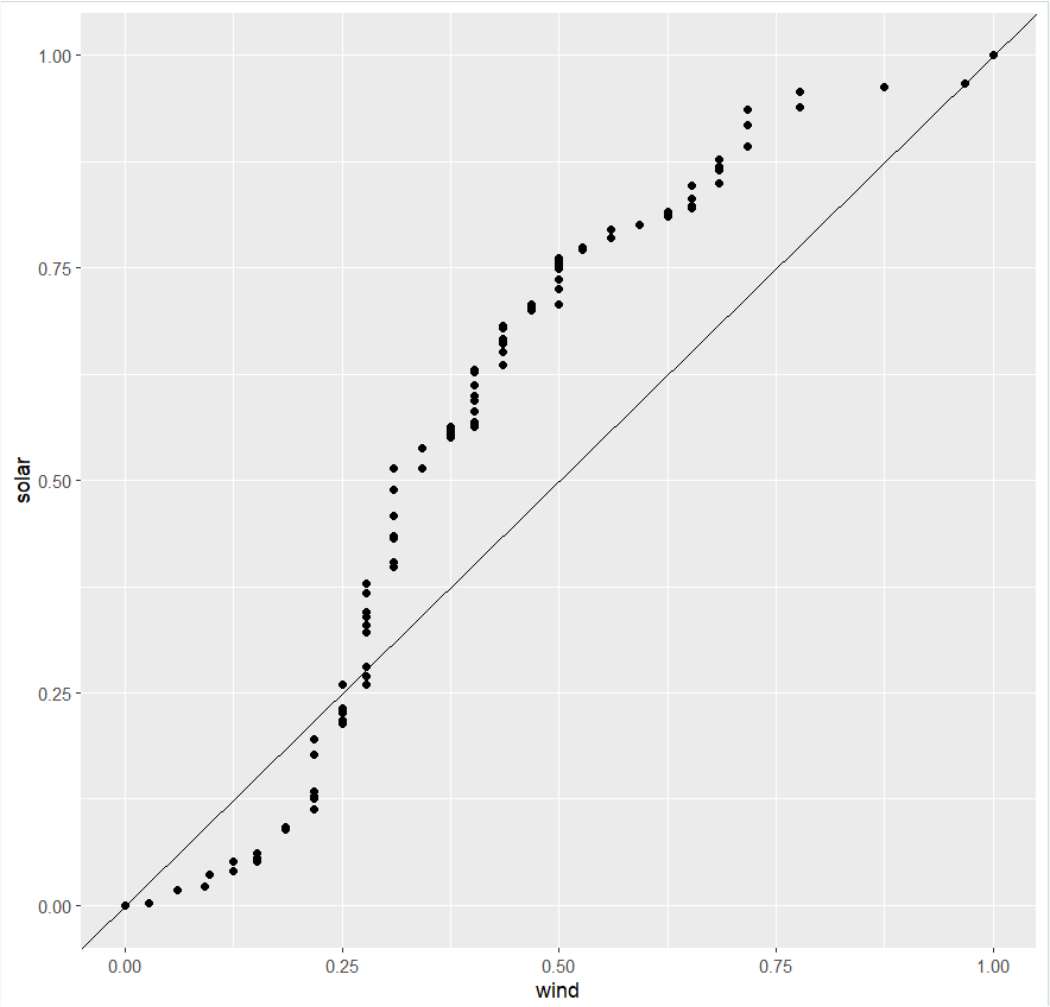


4) a) answer We cannot observe a clear distribution among the two variable

4b. Use a plot that will show the distributions of Wind and Solar.R and allow you to compare with fine detail

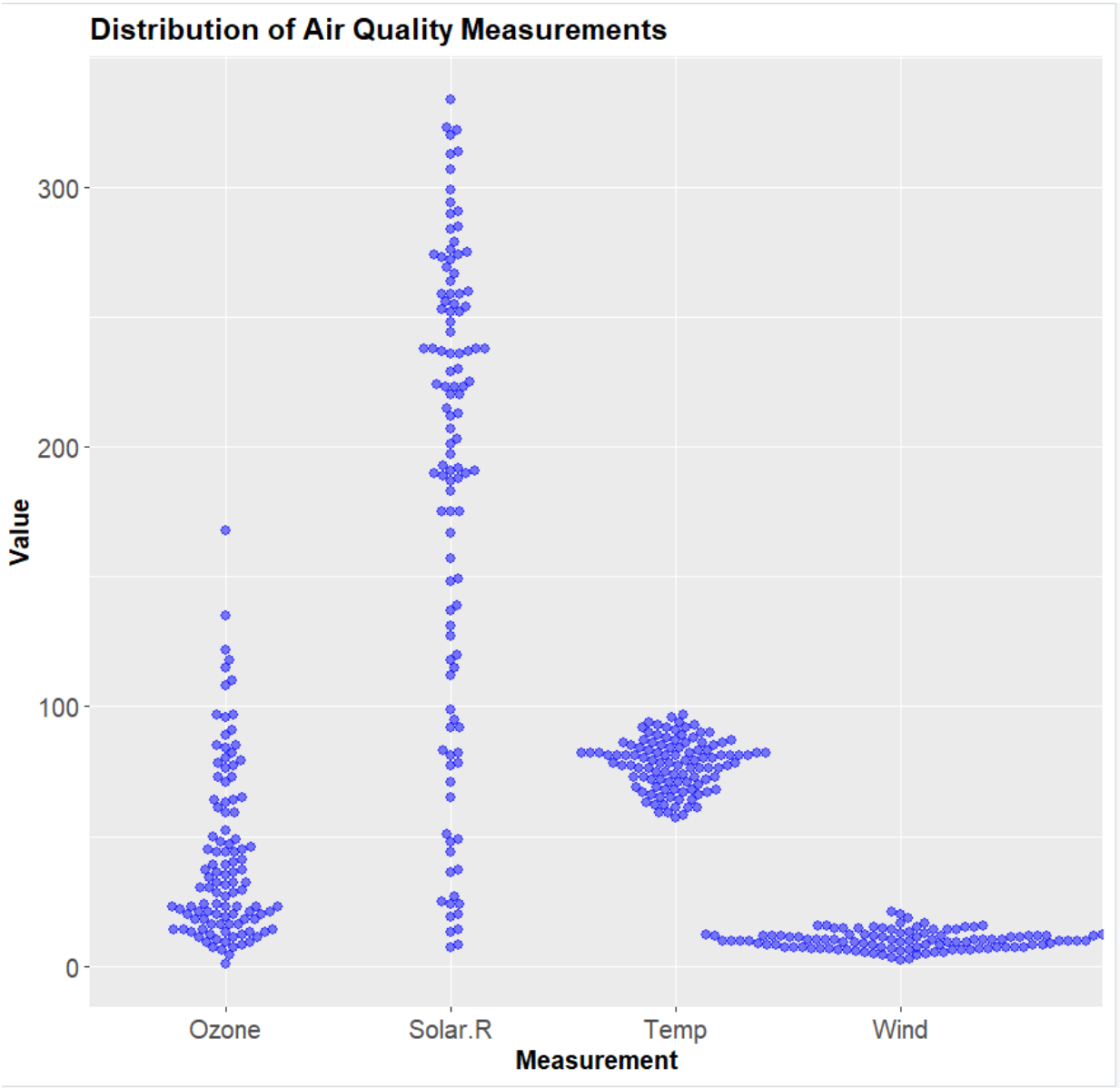






Solar and wind has been scaled to 0 to 1 to compare the distributions clearly. (it is skewed with higher radiation left skewed)

4) C)



Here we can see that Solar radiation has a wider range of distribution and wind has the least.

4d. For extra credit, compare Wind and Solar.R again with a QQ plot. What does this tell you?