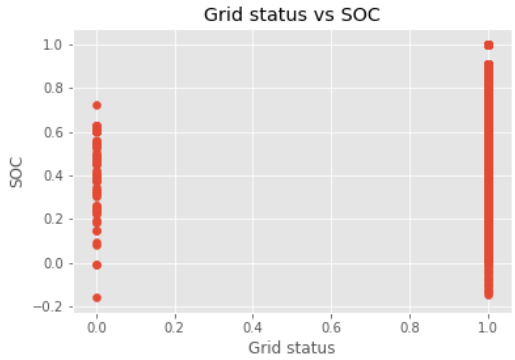
Correlation coefficient and explanation of the relationship

1. Correlation between Grid status and SOC

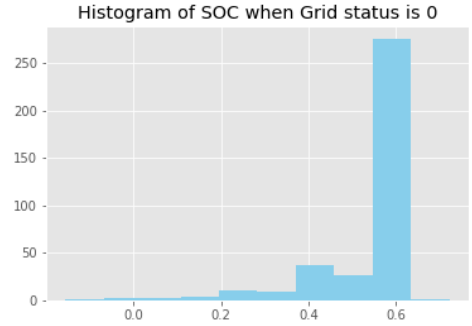
The variable ‘Grid status’ is a categorical variable and ‘SOC’ is a continuous variable. So point biserial correlation would be a measure to find correlation between these two variables. As ‘Grid status’ takes the values 0 and 1 (dichotomous variable) the point-biserial correlation is mathematically equivalent to the Pearson's correlation coefficient.

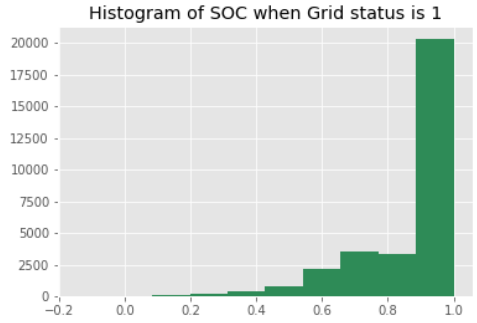
The correlation coefficient between ‘Grid status’ and ‘SOC’ is **0.2279**. There is a **positive correlation** between ‘Grid status’ and ‘SOC’ but it is **not too high**.

If we plot these two variables it would look like a bar diagram containing two columns as one variable is of binary type.

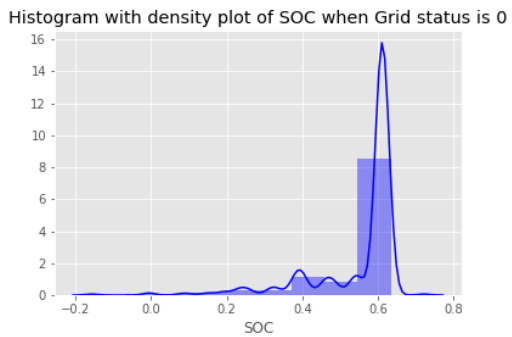


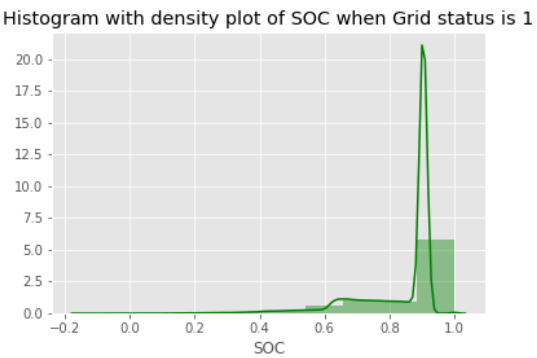
We have plotted the variable ‘SOC’ when ‘Grid status’ is 0 and 1.





We have also plotted the Histogram with density plot



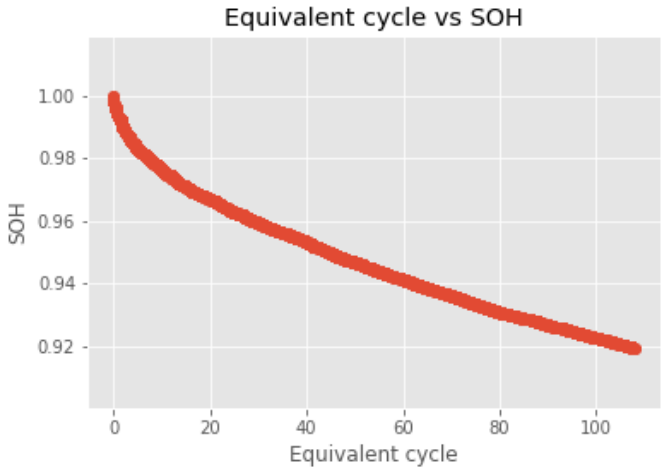


From the plots we see the distribution is same for Grid status = 0 and Grid status = 1. But there is a shift in origin in the curve when Grid status = 1 from the curve when Grid status = 0. It implies there is a positive correlation between Grid status and SOC. The value of the correlation coefficient is also claiming the same.

2. Correlation between 'Equivalent cycle' and 'SOH'

The correlation coefficient between 'Equivalent cycle' and 'SOH' is **-0.9843**. We find **highly negative linear correlation** between Equivalent cycle and SOH. We can say one variable is explaining the other. Increase in 'Equivalent cycle' indicates the decrease in 'SOH'.

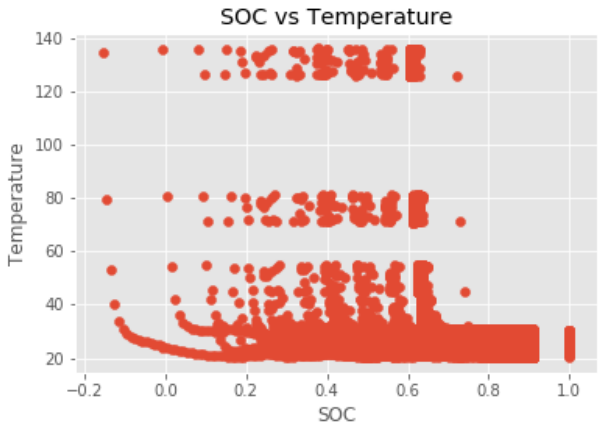
The scatterplot between 'Equivalent cycle' and 'SOH' also shows a linear negative relationship between them.



3. Correlation between 'SOC' and 'Temperature'

Correlation coefficient between these two variables is **-0.3690**. We find a **negative correlation** between 'SOC' and 'Temperature', but **not too high**. As 'SOC' increases 'Temperature' decreases.

The scatterplot looks like



From the plot we can see nonlinear relationship between 'SOC' and 'Temperature'. The curve is like negative exponential curve. So a logarithmic transformation may give some linear relationship between them.