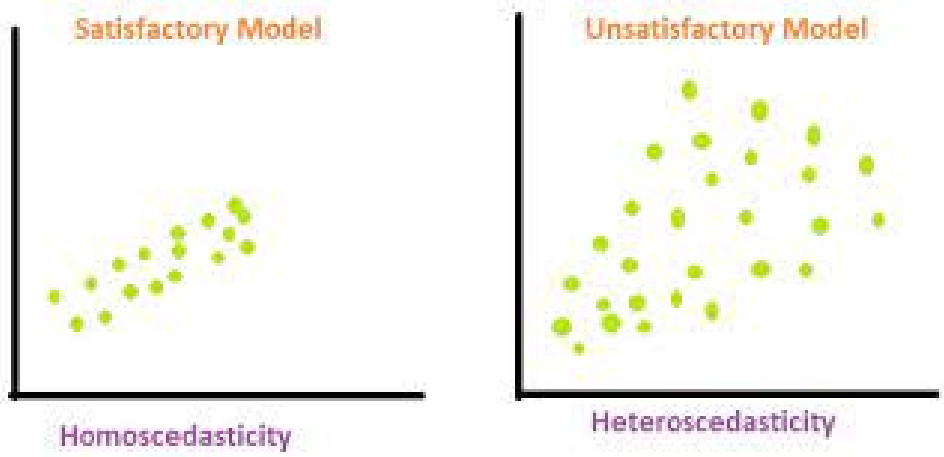
HOPE AI

Homoscedasticity & Heteroscedasticity Document

General rule of thumb for homoscedasticity: If the ratio of the largest sample variance to the smallest sample variance does not exceed 1.5

1) **Homoscedasticity / homogeneity of variances:**
   - In a data, having the same scatter, the points are in the same distance.
2) **Heteroscedasticity:**
   - In a data, having the different scatters, the points are in the widely varying distance from the regression line.



3) **Is Homoscedasticity or Heteroscedasticity good?**
   - Homoscedasticity, or getting close to it, is the gold standard for who wants to get a working model
   - To test the null hypothesis that the standard deviations of the measurement variable are the same for the different groups, homoscedasticity is good.
4) **To test homo & heteroscedasticity:**
   - Make a scatterplot with the residuals against the dependent (output) variable / column.
5) **Formulae:**
   - Homoscedasticity is expressed as $\text{var}(\varepsilon) = I\sigma^2$
   - Heteroscedasticity as $\text{var}(\varepsilon) = \text{diag}[\sigma_1^2, \sigma_2^2, \ldots, \sigma_I^2]$

     where we again assumed that the errors are uncorrelated

   (so the off-diagonal terms of the variance–covariance matrix are zero).