Prediction model for salary hike

There are no outliers present in the data.

From the scatter plot it is observed that linear relationship

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **model** | **COR** | **R Squared** | **RMSE** | **Sum of errors** |
| **Model (Y~X)** | 0.97 | 0.95 | 5592 | 0 |
| **Model (Y~log(X))** | 0.92 | 0.85 | 10302 | 0 |
| **Model (log(Y) ~ X)** | 0.96 | 0.93 | 7213 |  |

We don’t go for quadratic model as there is no curvilinear relationship.

Best fit model is simple linear regression

Salary = 25792 + 9450 (years experience)