1. Calories\_consumed-> predict weight gained using calories consumed

Model Table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model #** | **Model** | **R Squared** | **Correlation** | **RMSE** |
| Model 1 | Y~X | 0.8968 | 0.9461 | 103.3 |
| Model 2 | y ~ log(X) | 0.807 | 0.898 | 141 |

**Y~X is a better model with 89.6% R Squared value**

1. Delivery\_time -> Predict delivery time using sorting time

Model Table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model #** | **Model** | **R Squared** | **Correlation** | **RMSE** |
| Model 1 | Y~X | 0.68 | 0.82599 | 2.79165 |
| Model 2 | y ~ log(X) | 0.6954 | 0.8339 | 2.7331 |
| Model 3 | log(Y) ~ X | 0.71 | 0.8431 | 2.94 |
| Model 4 | Y ~ (X + X^2) | 0.6934 | NA | 2.74 |
| Model 5 | log(Y) ~ log(X) | 0.77 | 0.8787 | 2.94 |

**Log Transforming Both variables gives a better R squared value of 0.77**

1. Emp\_data -> Build a prediction model for Churn\_out\_rate

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model #** | **Model** | **R Squared** | **Correlation** | **RMSE** |
| Model 1 | Y~X | 0.83 | -0.911 | 3.997 |
| Model 2 | y ~ log(X) | 0.8486 | -0.92 | 3.786 |
| Model 3 | log(Y) ~ X | 0.8735 | -0.934 | 3.54 |
| Model 4 | Y ~ (X + X^2) | 0.9737 |  | 1.577 |

Model 4 – Quadratic Regression gives a very good R squared value of 0.9737 and the lease RMSE value.

1. Salary\_hike -> Build a prediction model for Salary\_hike

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model #** | **Model** | **R Squared** | **Correlation** | **RMSE** |
| Model 1 | Y~X | 0.957 | 0.978 | 5592 |
| Model 2 | y ~ log(X) | 0.8539 | 0.924 | 10302 |
| Model 3 | log(Y) ~ X | 0.93 | 0.96 | 7213 |
| Model 4 | Y ~ (X + X^2) | 0.957 |  | 5590 |

Model 4 – Quadratic Regression gives a very good R squared value of 0.957 and least RMSE value.