

Obs	P	_FREQ_	ERROR_MEAN	ERROR_REG	ERROR_POI	ERROR_NB	ERROR_ZIP	ERROR_ZINB	ERROR_HURDLE
1	1	12795	1.55924	1.06725	1.16770	1.16770	0.95785	0.96336	0.96745

Obs	P	_FREQ_	ERROR_MEAN	ERROR_REG	ERROR_POI	ERROR_NB	ERROR_ZIP	ERROR_ZINB	ERROR_HURDLE
1	1.5	12795	1.76929	1.20251	1.29497	1.29497	1.11741	1.12014	1.12284

Obs	P	_FREQ_	ERROR_MEAN	ERROR_REG	ERROR_POI	ERROR_NB	ERROR_ZIP	ERROR_ZINB	ERROR_HURDLE
1	2	12795	1.92629	1.32335	1.40481	1.40481	1.27334	1.27290	1.27344

Above are the error metrics generated for all 6 models that we built. Based on the previous discussion of the AICs as well as the nature of the Zero Inflated counting data at hand, we are not surprised to find that the Zero Inflated Poisson regression model performed the best. We calculated the error values for the different values of P (P=1, P=1.5, and P=2), based on the difference between the number of cases sold as predicted by the model and the actual number of cases sold both using the training data. For all values of P, all 6 of our models performed better than the mean value model. For P=1 and P=1.5, the error values for the ZIP model were the lowest. For P=2, the error value for the ZIP model is actually tied with the Hurdle Model and slightly higher than the error value for the ZINB model. P=1 will gives the average error, so we see that the average error of the ZIP Model is the lowest.

Interesting enough, the linear regression model performed with better accuracy than the Poisson and Negative Binomial models. This isn't what we would expect since our target variable is of the counting data variety, which usually isn't recommended for linear regression. Further exploration into the transformation of our variables could suggest why the linear regression model performed better.