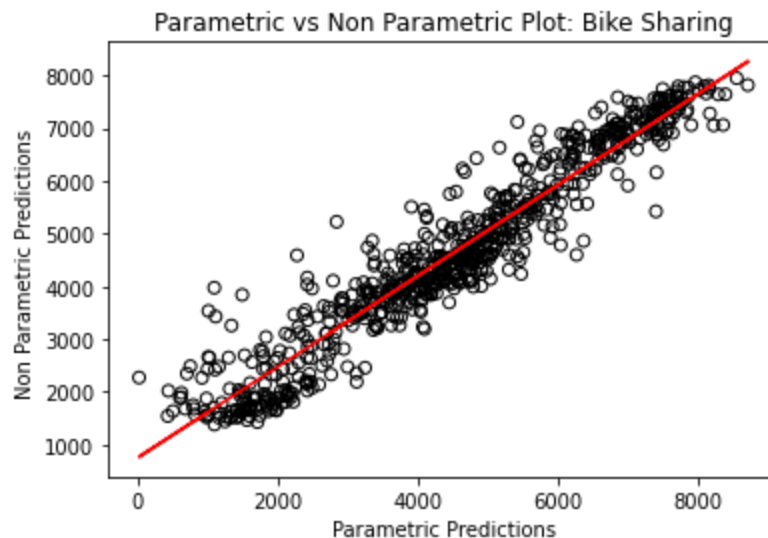


Below a plot that shows two different sets of predictions for the number of bike riders on a given day. The X values were generated by taking the original prediction variables and outcome variable and fitting them to a linear regression model. These X values were then plugged into the model we generated. These Y values were obtained by taking the same prediction and outcome variables and generating predictions using a KNN model. In order to determine how similar or different the outcomes of these two models are, they have been plotted in a Parametric vs Nonparametric Plot.



From the above graph, we can see that the data is fit well, as the plotted X,Y values, seen in black, fall close to, and with the same pattern as the line of best fit, seen in red. The data does not seem to be over or under fit. We can also see that there are no subpopulations in our data since there are no streaks or tails.

Since the slope of the line of best fit is 0.8635, we can infer that the predictions we generated with the parametric model are very similar to those we generated with the nonparametric model. Since the results given both models are highly similar, it would most likely be best to use the linear regression model that we generated when making future predictions about this dataset. The linear regression model is much faster to generate and run than the KNN model is, so since we do not observe a significant difference in the results, we can avoid using the KNN model.