F1 Score vs Size of training dataset