Comparison of Average F1 between Two Models

