Comparison of Mean F1 Weighted Scores Across Models