

Cohen's Kappa

Verified Python analysis transcript

Test: Cohen's Kappa

Design: Nominal agreement between G1 and G2 four-category grade classifications.

Variables: G1 grade band, G2 grade band

Null hypothesis: Agreement between the two classifications equals chance agreement.

Primary formula: $\text{kappa} = (P_o - P_e) / (1 - P_e)$

Valid source rows: 649

Verified results:

cohens_kappa: 0.610017494485434

z_value: 23.177634403770856

p_value: 7.65551073960406e-119

observed_agreement: 0.7565485362095532

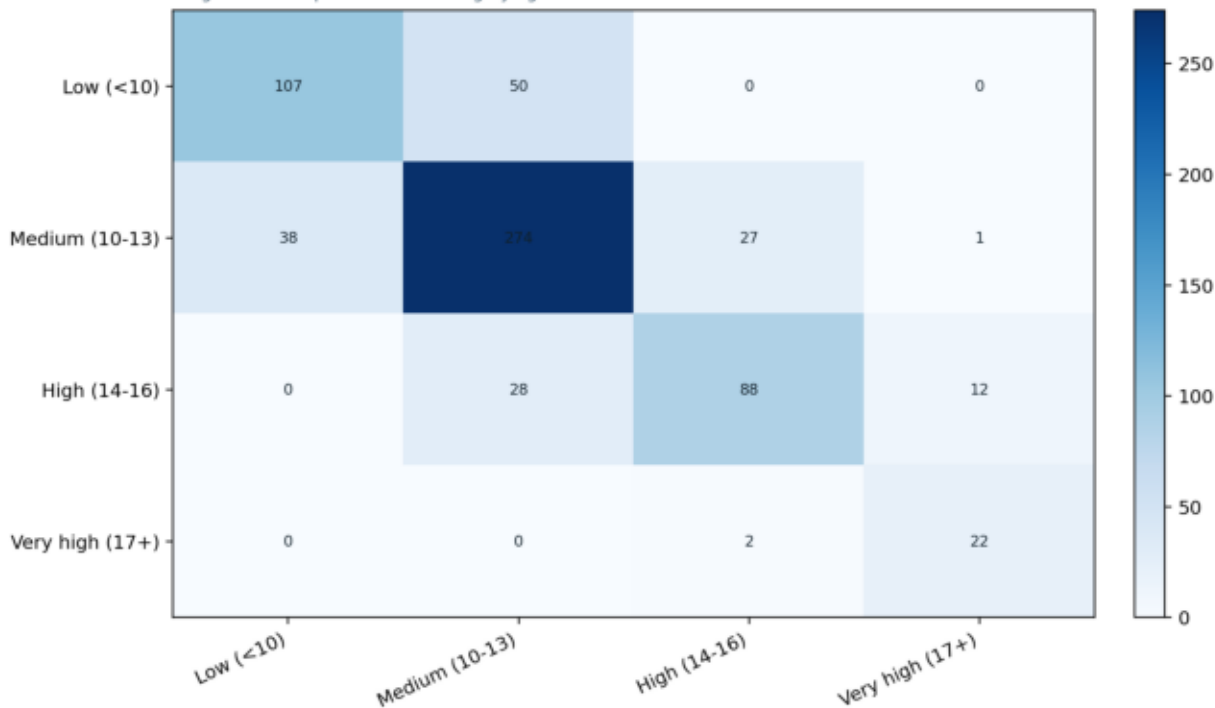
Interpretation:

Cohen's kappa = 0.610017, z = 23.177634, p = 7.65551e-119.

Because this transcript treats categories as nominal, disagreement distance is not weighted.

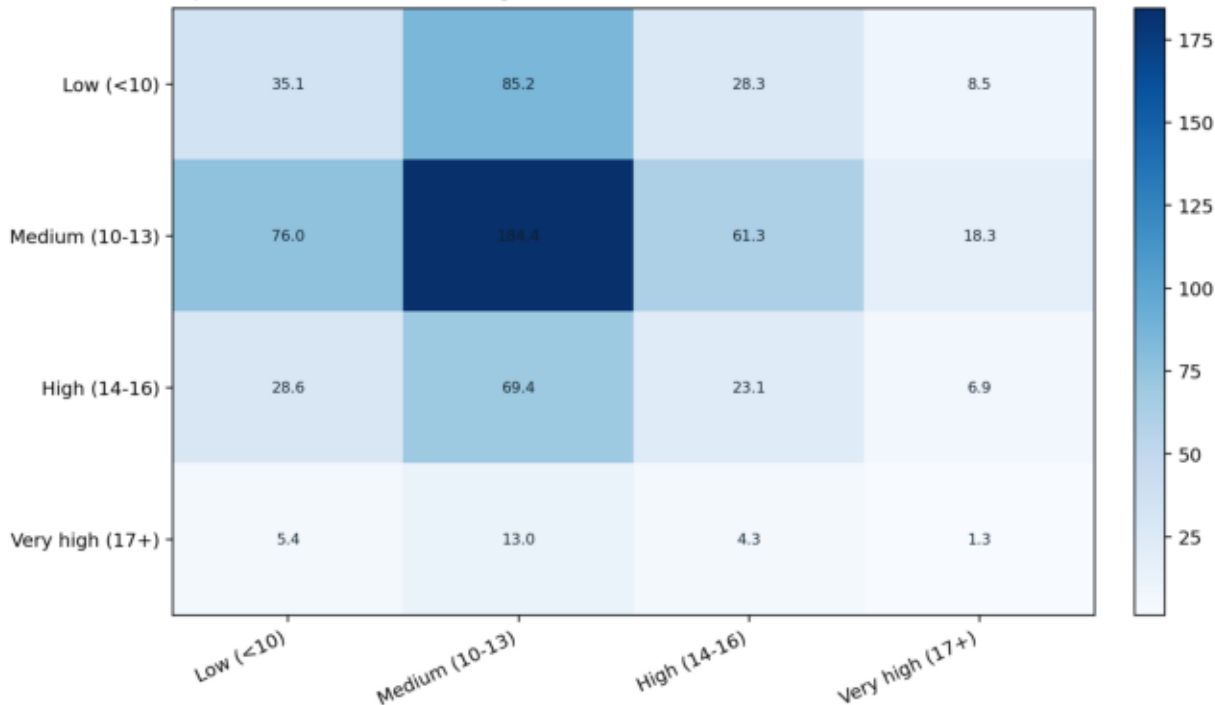
G1 and G2 agreement matrix

Diagonal cells represent exact category agreement



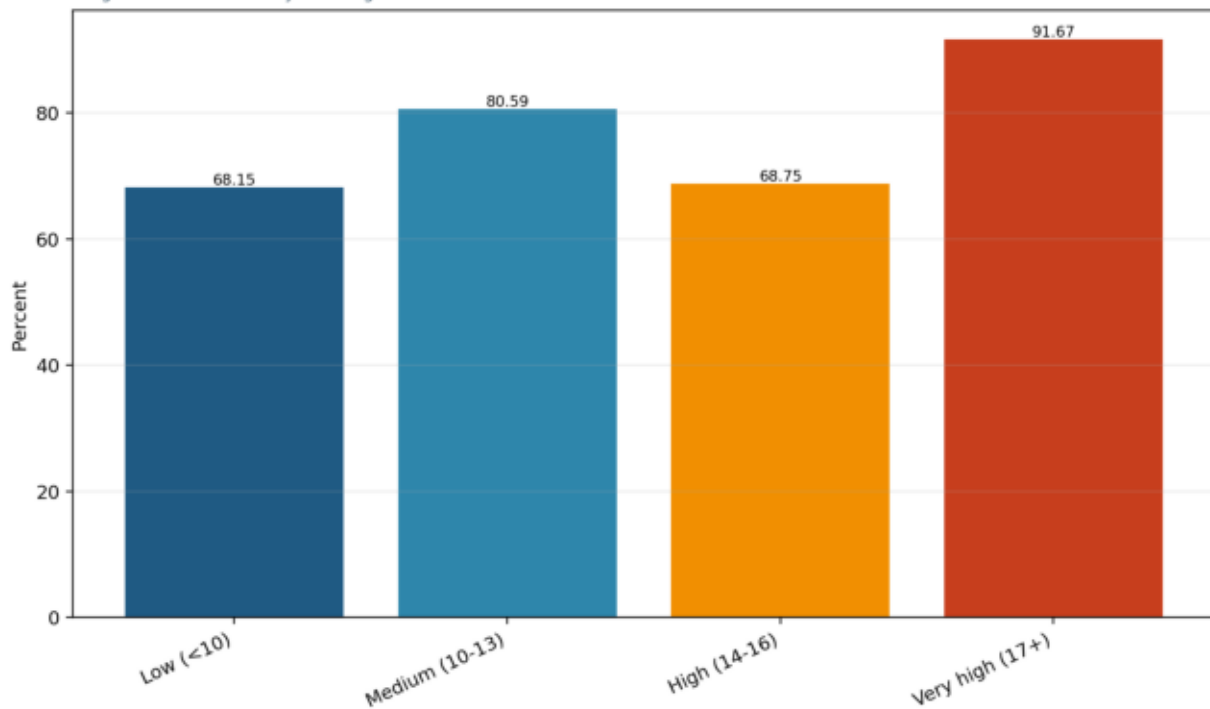
Chance-expected agreement counts

Expected table from classification margins



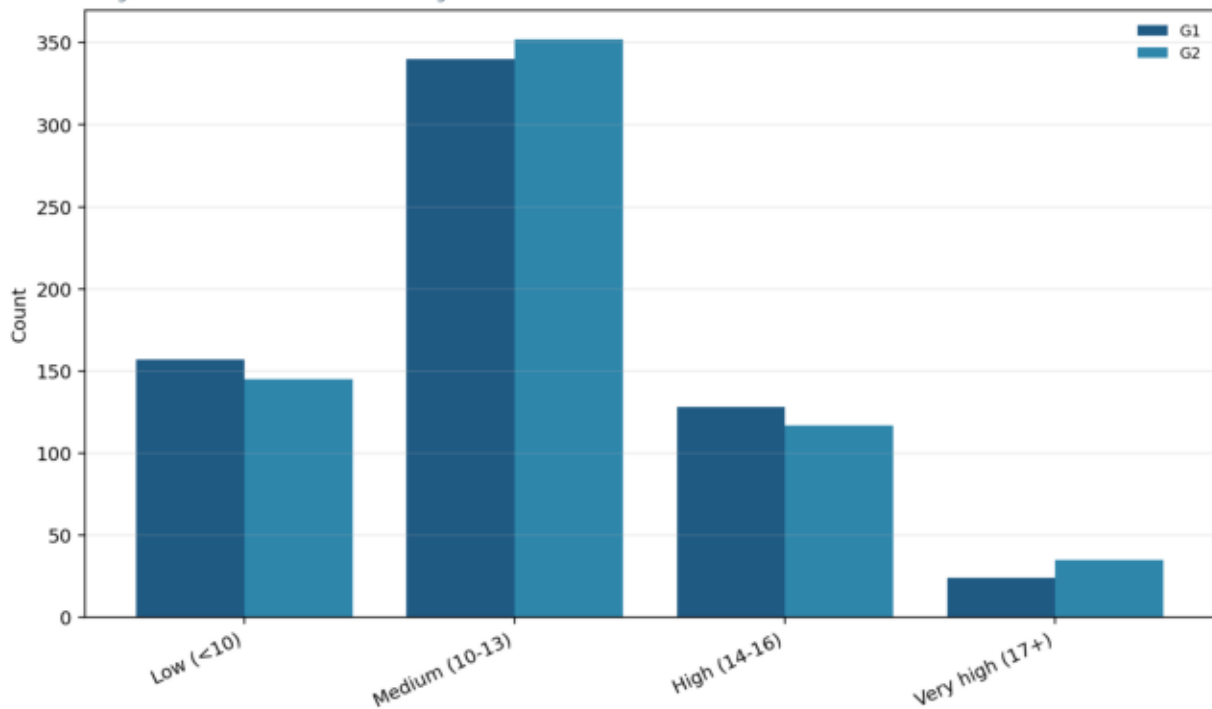
Agreement within each G1 category

Diagonal count divided by G1 marginal count



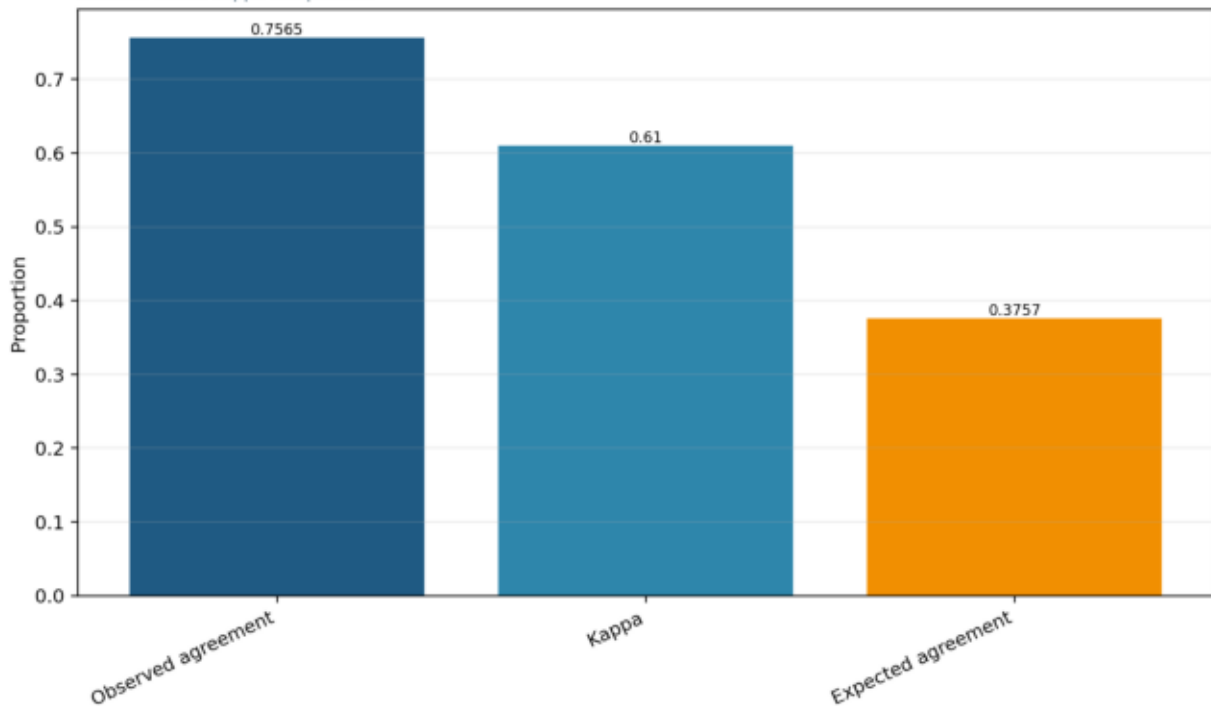
Classifier marginal distributions

Marginal imbalance determines chance agreement



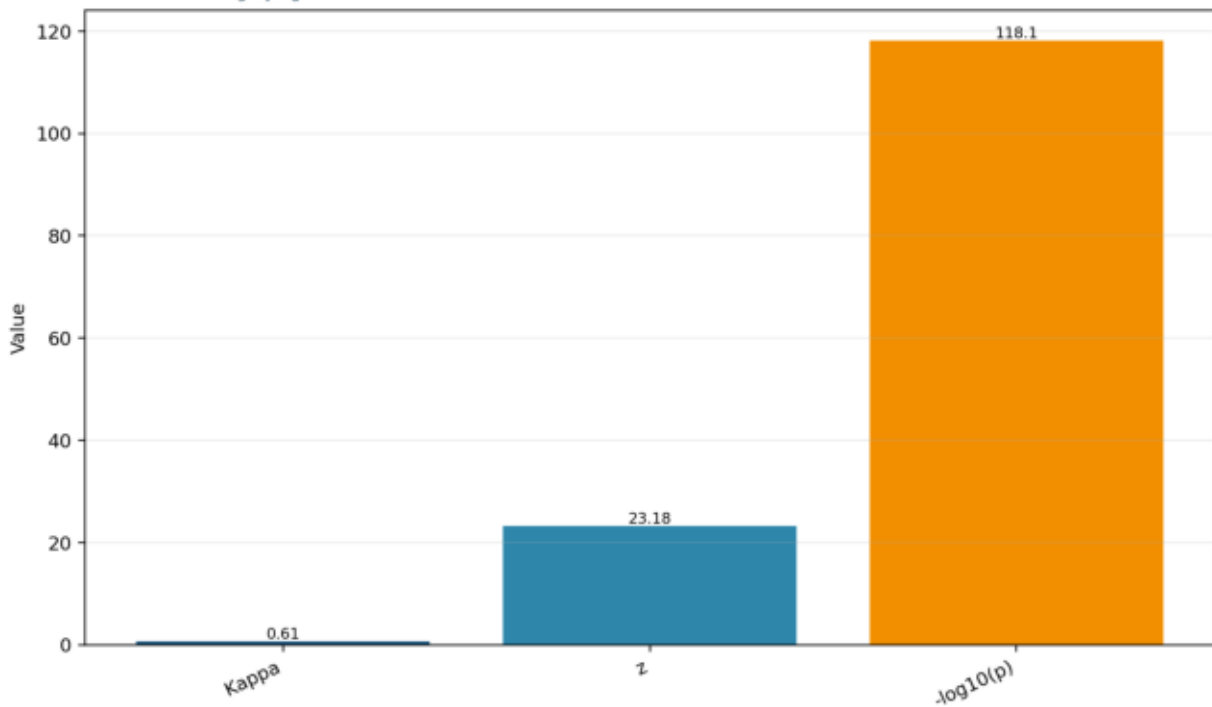
Observed and chance-corrected agreement

Nominal Cohen kappa components



Cohen kappa inference

Nominal four-category agreement



Agreement Table

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
107	50	0	0
38	274	27	1
0	28	88	12
0	0	2	22

Chance Expected Counts

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
35.077	85.1525	28.3035	8.46687
75.963	184.407	61.2943	18.3359
28.5978	69.4237	23.0755	6.90293
5.3621	13.0169	4.32666	1.2943

Category Agreement

within_g1_agreement
0.681529
0.805882
0.6875
0.916667