

Weighted Kappa

Verified Python analysis transcript

Test: Weighted Kappa

Design: Ordinal agreement between G1 and G3 four-level grade classifications.

Variables: G1 grade band, G3 grade band

Null hypothesis: Observed weighted disagreement equals chance weighted disagreement.

Primary formula: $\text{kappa}_w = 1 - [\sum_{ij} w_{ij} p_{ij}] / [\sum_{ij} w_{ij} p_{i+} + p_{+j}]$

Valid source rows: 649

Verified results:

linear_weighted_kappa: 0.6473984072889338

quadratic_weighted_kappa: 0.7651598550305657

quadratic_observed_disagreement: 0.032357473035439135

quadratic_expected_disagreement: 0.13778510075289996

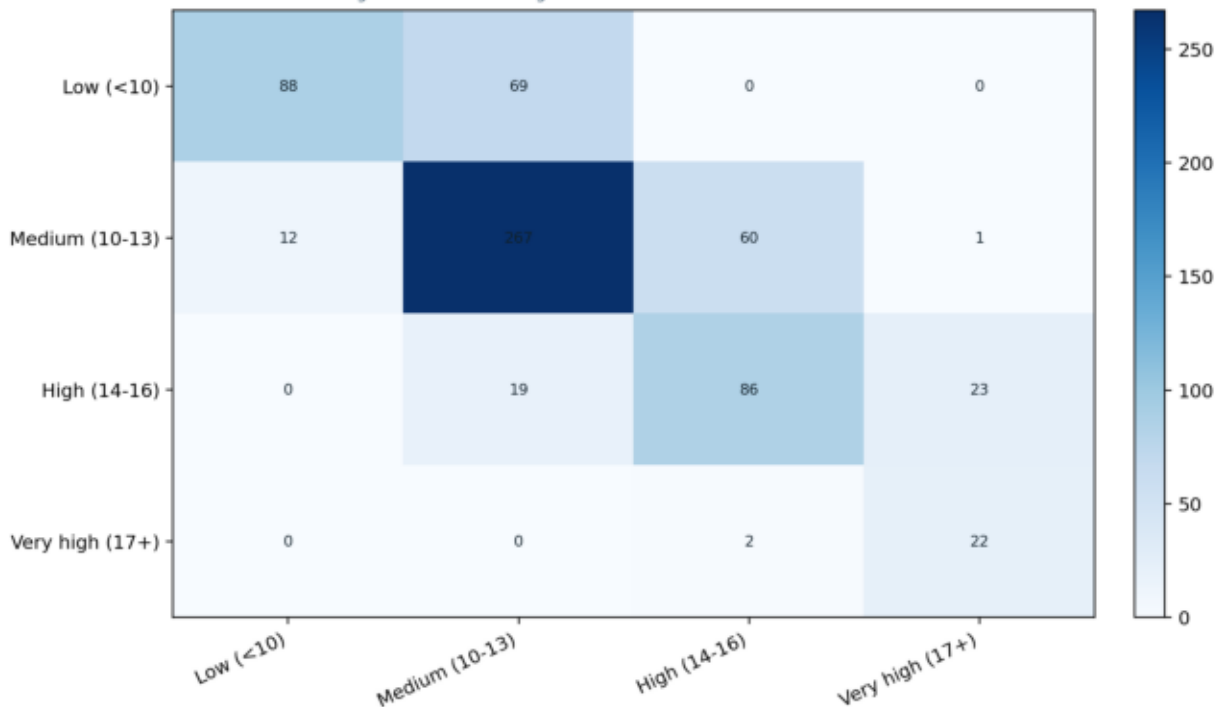
Interpretation:

Linear weighted kappa = 0.647398; quadratic weighted kappa = 0.765160.

Quadratic weighting is primary because distant grade-band errors receive a larger penalty than adjacent errors.

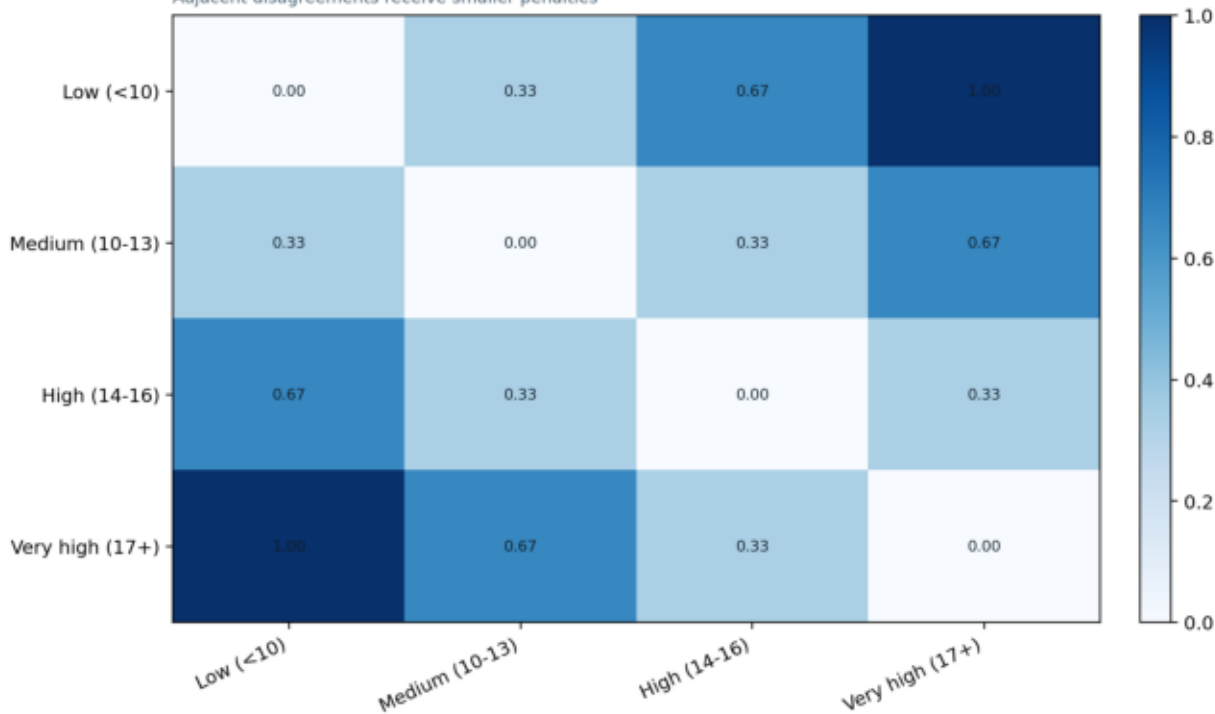
G1 and G3 ordinal agreement matrix

Distance from the diagonal carries meaning



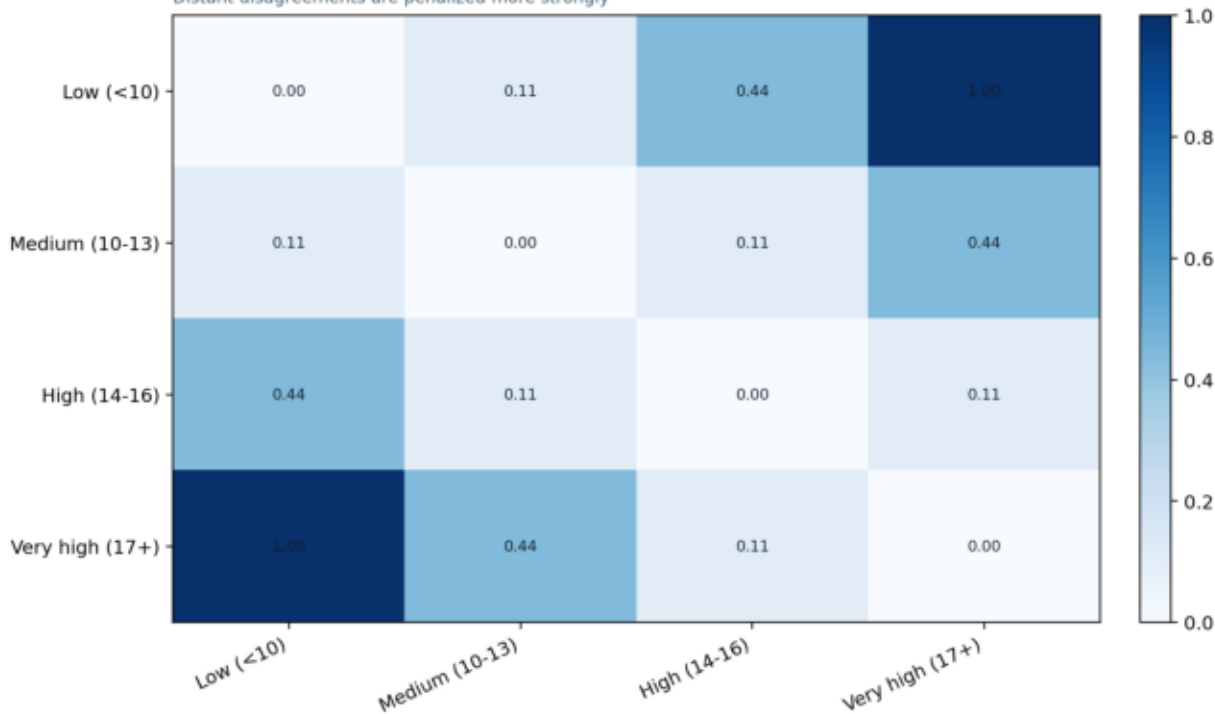
Linear disagreement weights

Adjacent disagreements receive smaller penalties



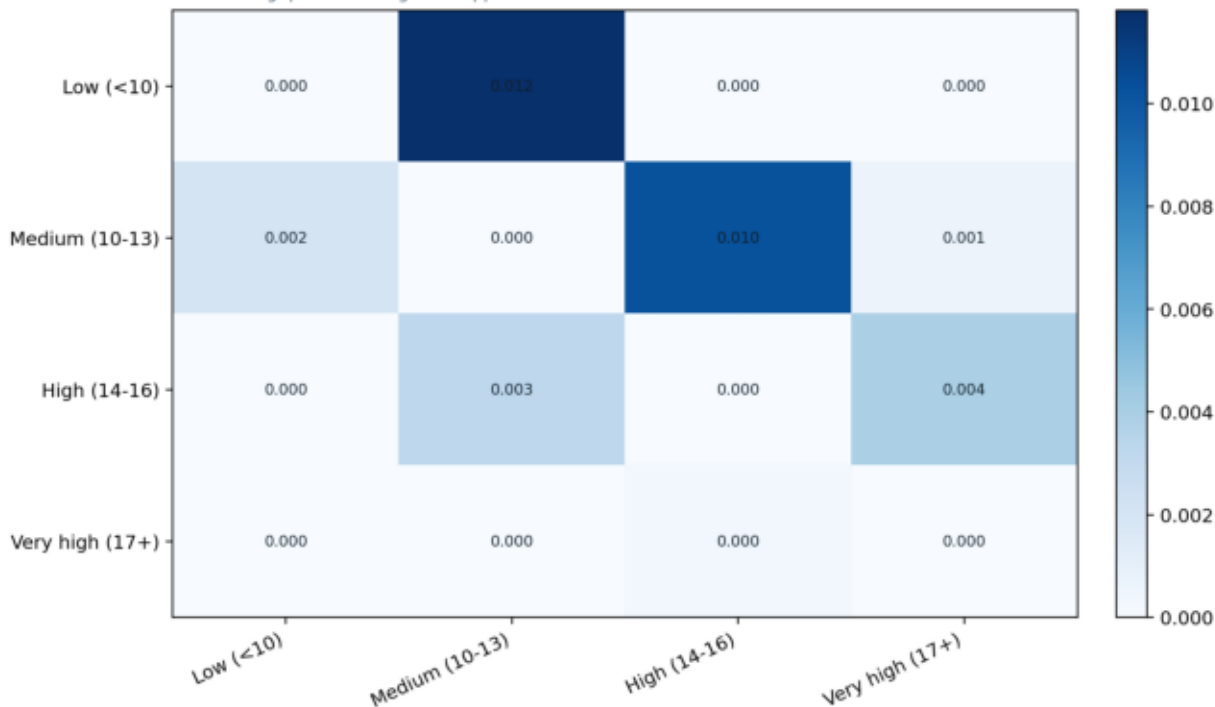
Quadratic disagreement weights

Distant disagreements are penalized more strongly



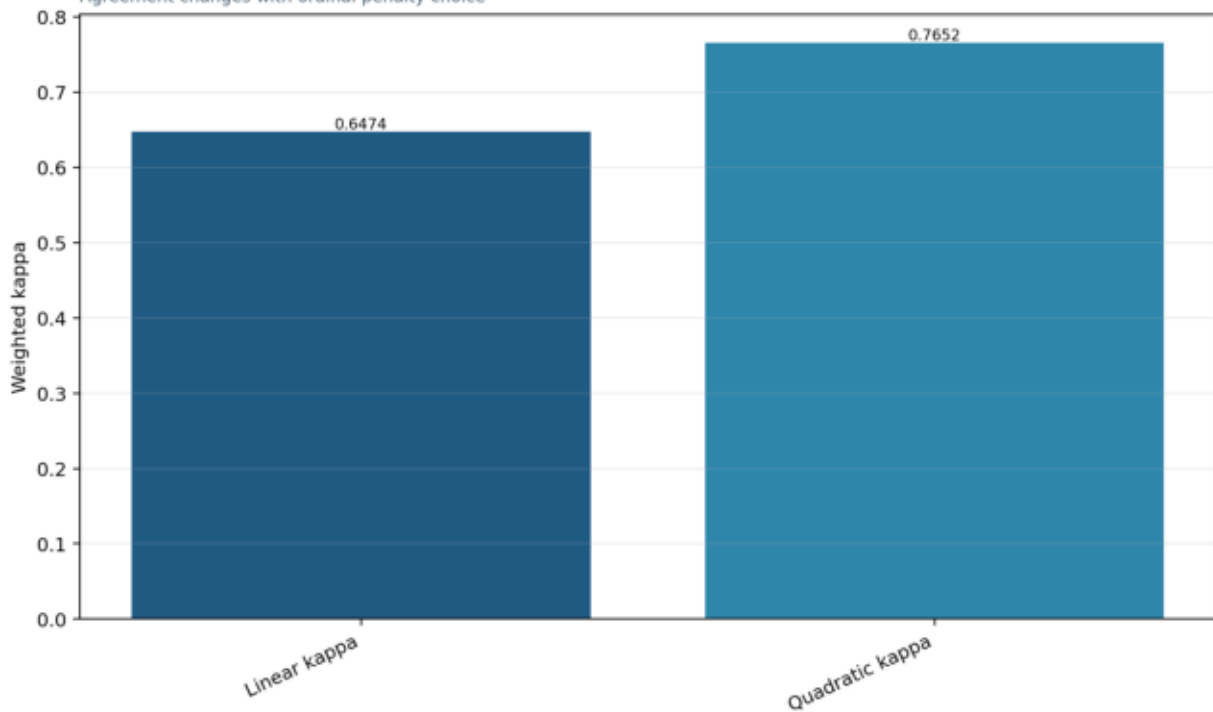
Observed weighted disagreement contributions

Cells driving quadratic weighted kappa



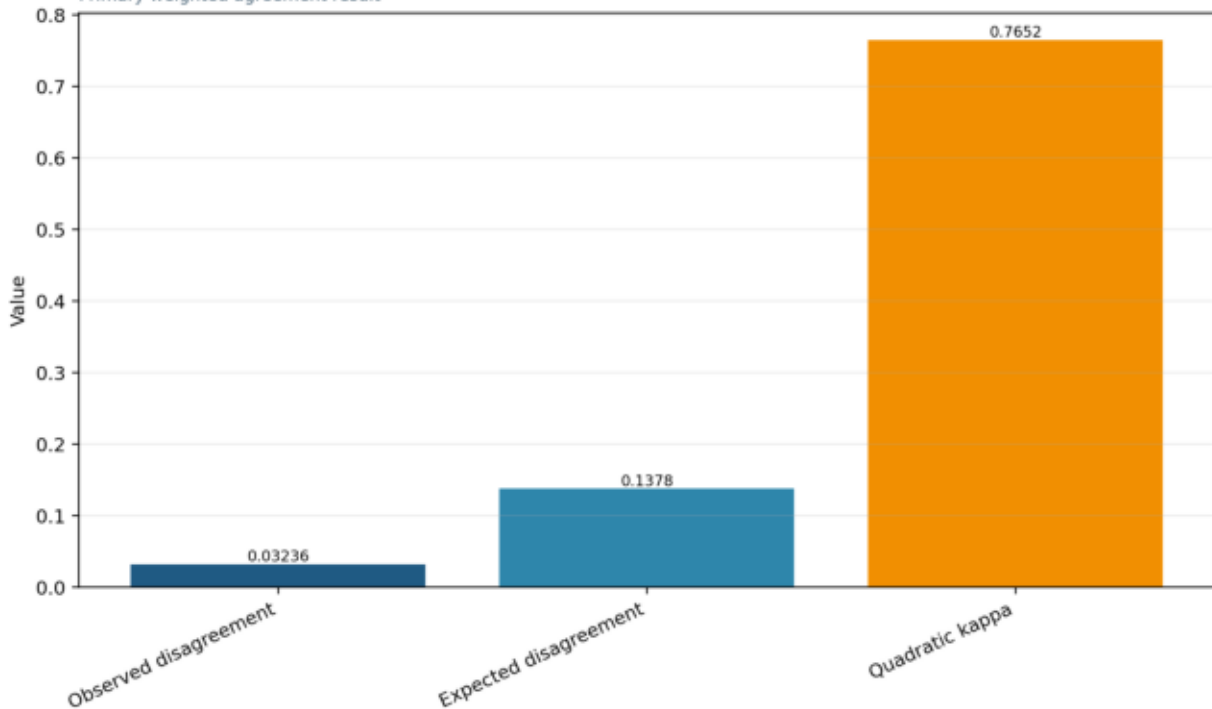
Weighting-scheme sensitivity

Agreement changes with ordinal penalty choice



Quadratic kappa components

Primary weighted agreement result



Ordinal Agreement Table

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
88	69	0	0
12	267	60	1
0	19	86	23
0	0	2	22

Linear Disagreement Weights

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
0	0.333333	0.666667	1
0.333333	0	0.333333	0.666667
0.666667	0.333333	0	0.333333
1	0.666667	0.333333	0

Quadratic Disagreement Weights

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
0	0.111111	0.444444	1
0.111111	0	0.111111	0.444444
0.444444	0.111111	0	0.111111
1	0.444444	0.111111	0

Quadratic Weighted Observed

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
0	0.011813	0	0
0.00205444	0	0.0102722	0.000684814
0	0.00325287	0	0.00393768
0	0	0.000342407	0