

Categorical Data Analysis in Python

Verified Python analysis transcript

Test: Categorical Data Analysis in Python

Design: A reproducible Python workflow for a 2 x 4 school-by-final-grade-band table.

Variables: school, G3 grade band

Null hypothesis: School and final-grade band are independent.

Primary formula: $X^2 = \sum (O-E)^2/E$; $V = \sqrt{X^2/(N*\min(r-1,c-1))}$

Valid source rows: 649

Verified results:

statistic: 59.534045213043086

p_value: 7.392451635150176e-13

degrees_of_freedom: 3

cramers_v: 0.30287285267769787

minimum_expected_count: 16.018489984591678

cells_expected_below_5: 0

fisher_freeman_halton_monte_carlo_p: 1.999960000799984e-05

mutual_information_nats: 0.04397165580391293

Interpretation:

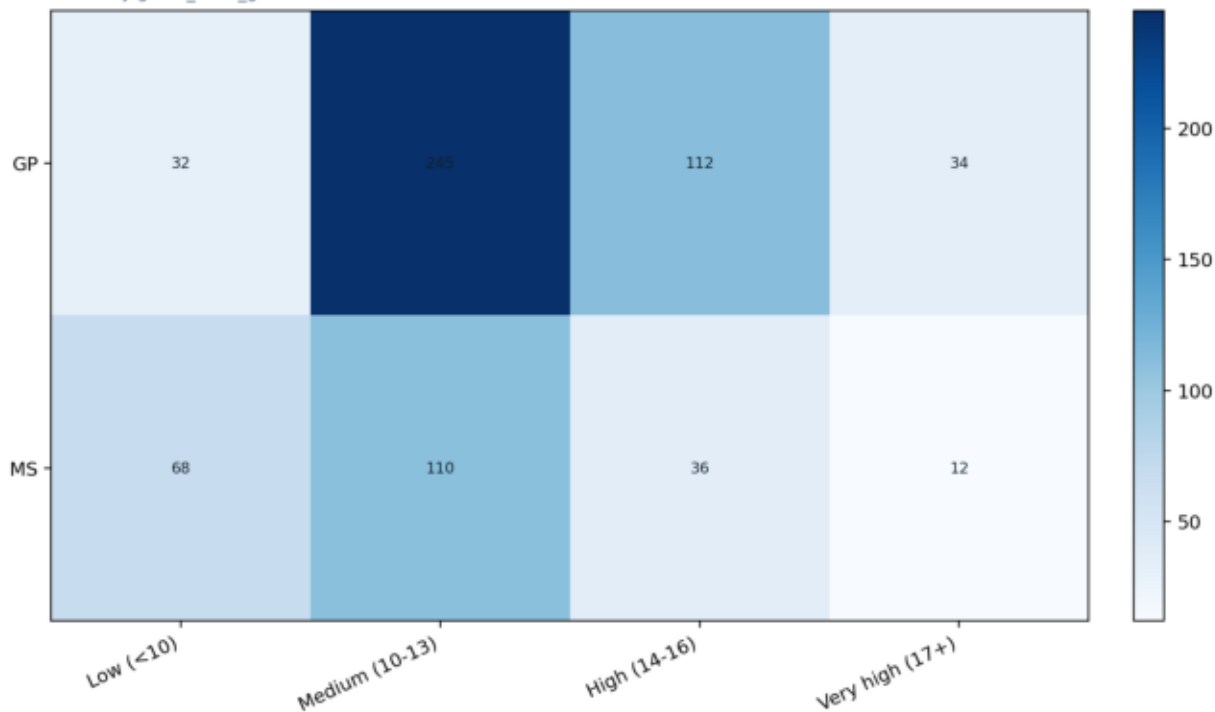
The primary statistic is 59.534045 with $p = 7.39245e-13$.

Cramer's V is 0.302873 and the minimum expected count is 16.018.

Interpret the residual and contribution tables before describing which cells account for the global result.

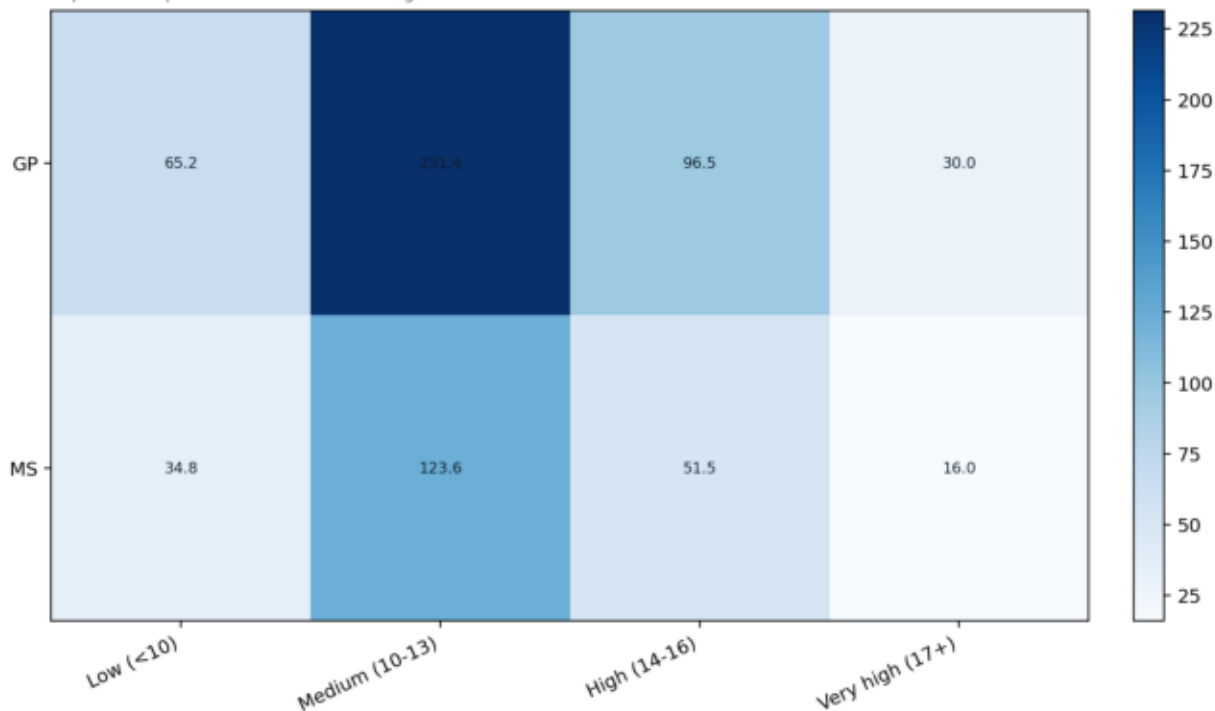
Observed categorical counts

school by grade_band_g3



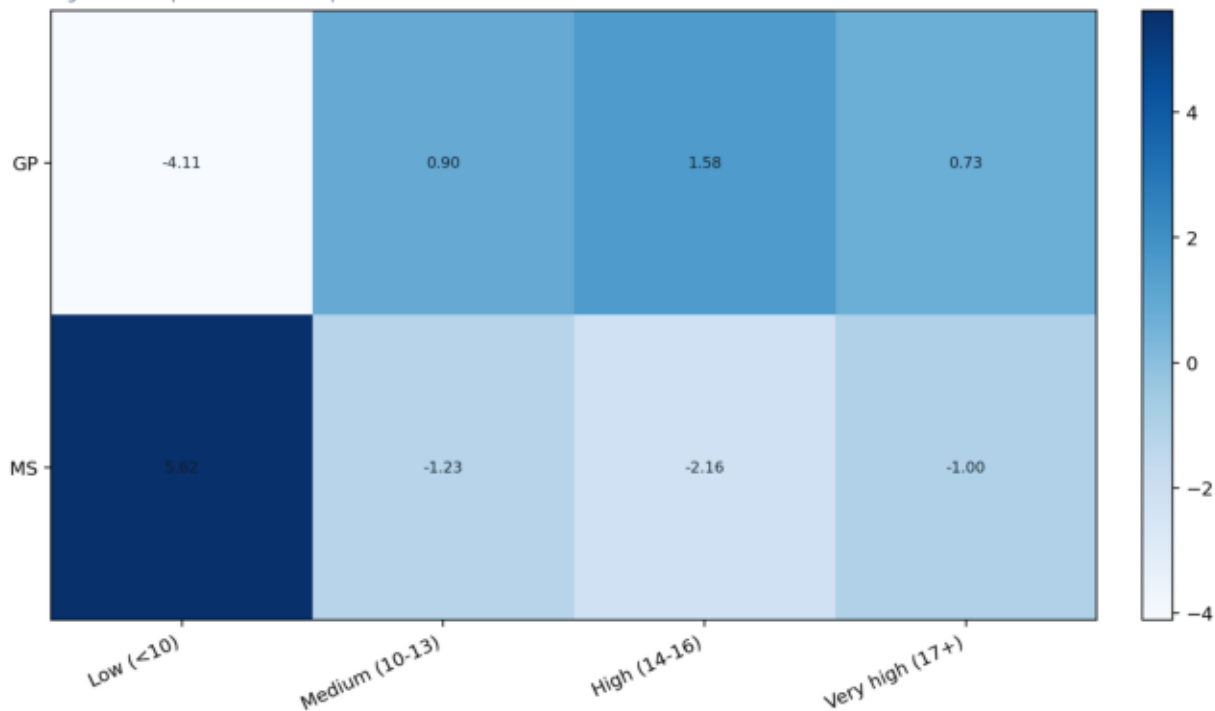
Expected counts under the null

Expected frequencies from the fitted margins



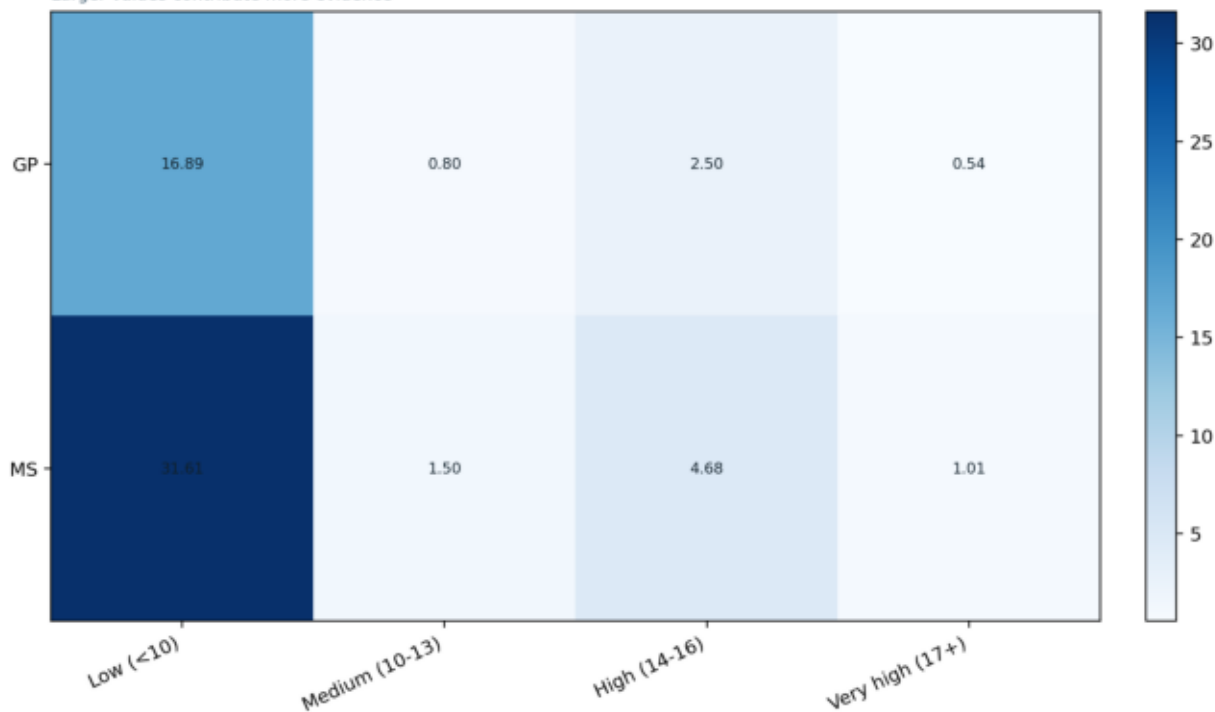
Pearson residual pattern

Signed cell departures from independence



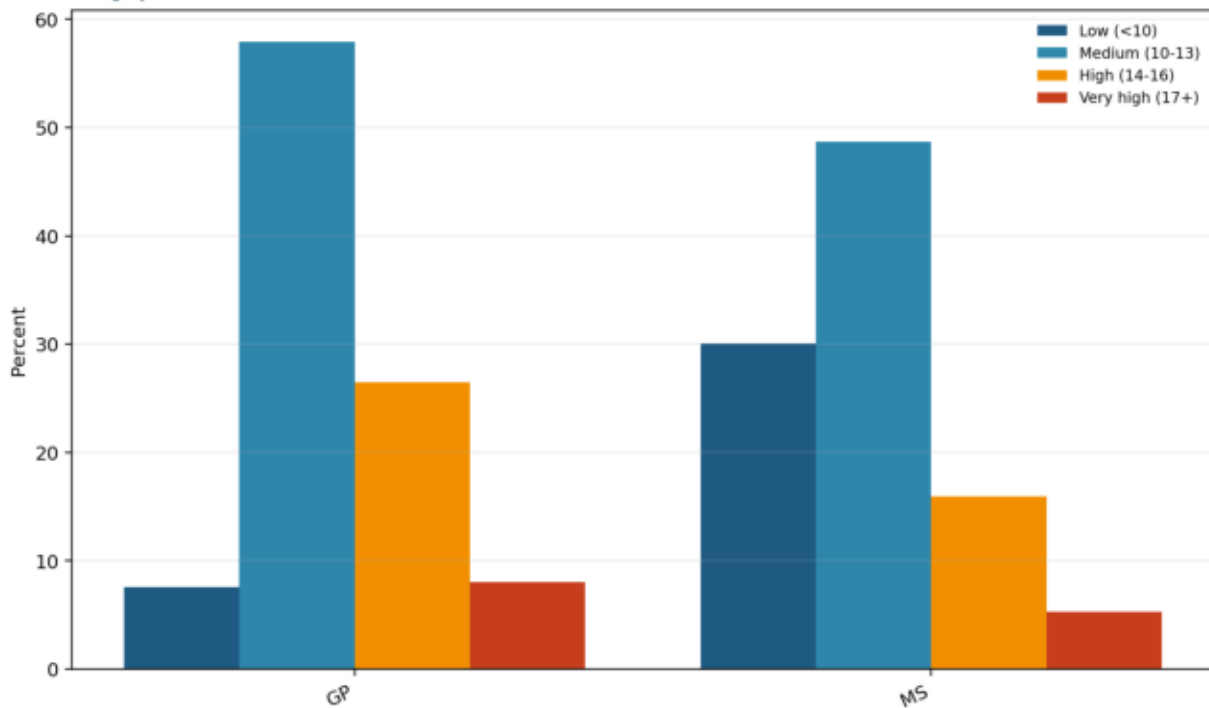
Cell contributions to the statistic

Larger values contribute more evidence



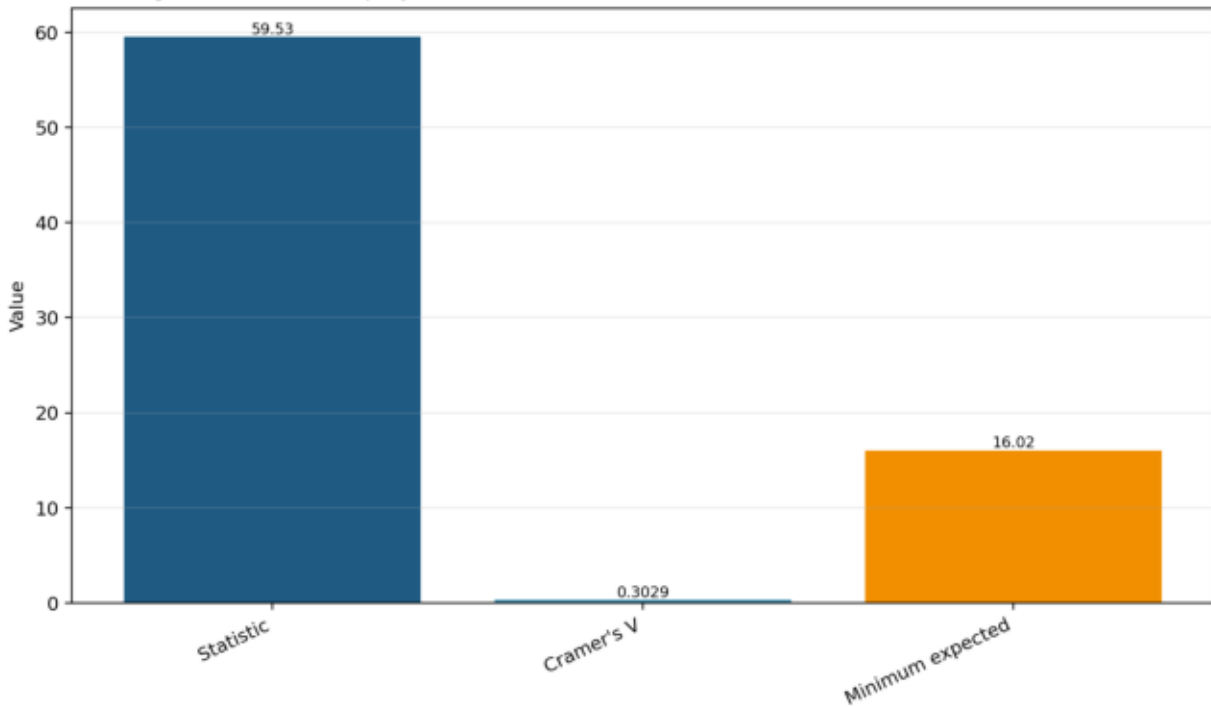
Row percentage comparison

Category distribution within each row



Model diagnostics

test strength, effect size, and adequacy



Asymptotic and exact evidence

Reproducible Python cross-check



Observed Counts

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
32	245	112	34
68	110	36	12

Expected Counts

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
65.1772	231.379	96.4622	29.9815
34.8228	123.621	51.5378	16.0185

Pearson Residuals

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
-4.10953	0.895459	1.58201	0.733899
5.62222	-1.22507	-2.16434	-1.00404

Cell Contributions

Low (<10)	Medium (10-13)	High (14-16)	Very high (17+)
16.8882	0.801846	2.50276	0.538607
31.6094	1.5008	4.68437	1.0081