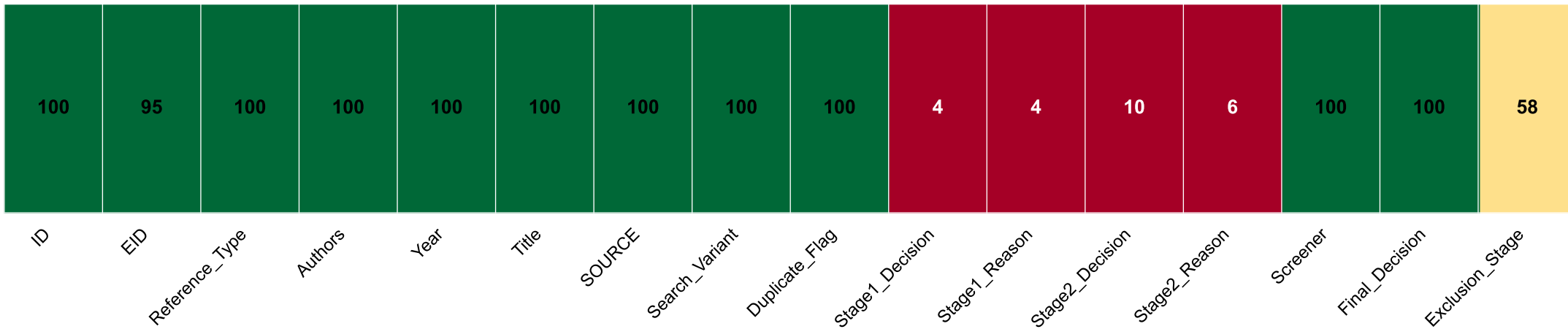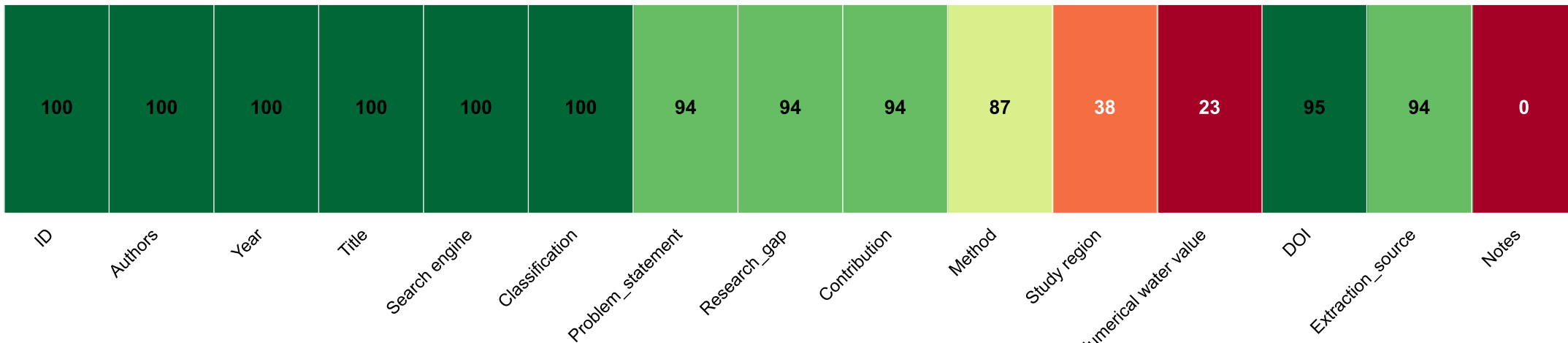# Data Completeness Across Dataset Files

(Percentage of non-null values per column — original data columns only)
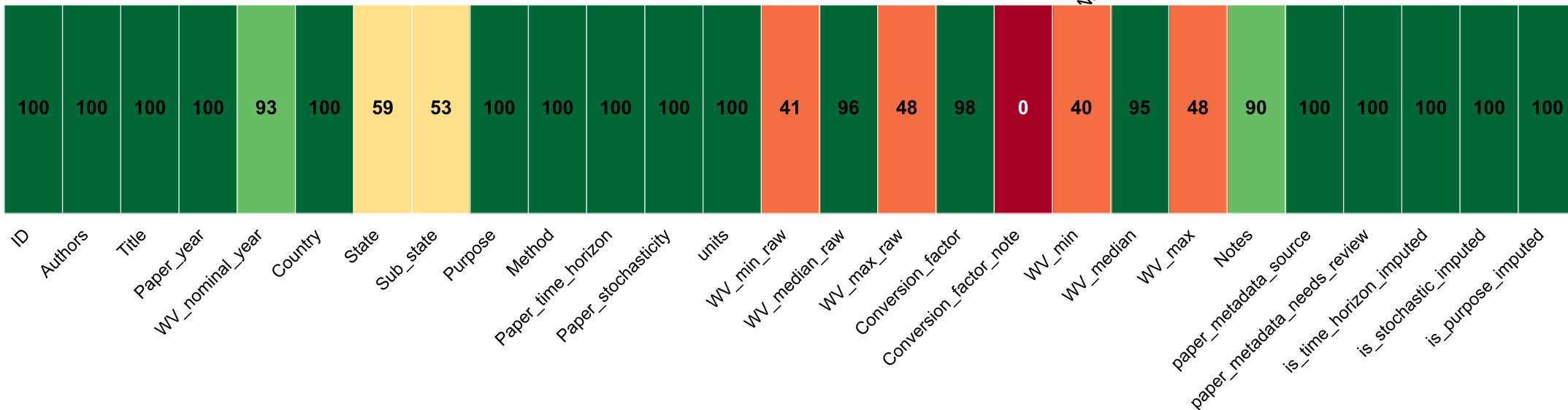
01_Screening

ID	EID	Reference_Type	Authors	Year	Title	SOURCE	Search_Variant	Duplicate_Flag	Stage1_Decision	Stage1_Reason	Stage2_Decision	Stage2_Reason	Screener	Final_Decision	Exclusion_Stage
100	95	100	100	100	100	100	100	100	4	4	10	6	100	100	58

02_Classification

ID	Authors	Year	Title	Search engine	Classification	Problem_statement	Research_gap	Contribution	Method	Study region	Numerical water value	DOI	Extraction_source	Notes
100	100	100	100	100	100	94	94	94	87	38	23	95	94	0

03_WaterValue

ID	Authors	Title	Paper_year	WV_nominal_year	Country	State	Sub_state	Purpose	Method	Paper_time_horizon	Paper_stochasticity	units	WV_min_raw	WV_median_raw	WV_max_raw	Conversion_factor	Conversion_factor_note	WV_min	WV_median	WV_max	Notes	paper_metadata_source	paper_metadata_needs_review	is_time_horizon_imputed	is_stochastic_imputed	is_purpose_imputed
100	100	100	100	93	100	59	53	100	100	100	100	100	41	96	48	98	0	40	95	48	90	100	100	100	100	100

Completeness (%): 0 – 25 – 50 – 75 – 100