

HMIS data quality report

Reporting period 2025-10-01 to 2026-09-30. Generated by the toolkit's `run_checks.py`.

Demonstration. Every person, project and record in this report is synthetic, generated by `generate_synthetic.py` with known problems planted at fixed rates. No client data was used. The files follow a subset of the HUD HMIS CSV export columns (Project, Client, Enrollment, Exit), so the same checks run against a real export.

Summary

Active enrollments	Active clients	Enrollments with no finding	Load time	Check run time	Database integrity	Backup restore test
913	592	553 (60.6%)	20 ms	30 ms	ok	pass

Findings by check

Check	Area	What it finds	Flagged	Records checked	Rate
DQ01	Identifiers	Name missing, unknown or refused	17	913	1.9%
DQ02	Identifiers	SSN missing or not a valid SSN	68	913	7.4%
DQ03	Identifiers	Date of birth missing or impossible	14	913	1.5%
DQ04	Universal data elements	Veteran status missing for adults	13	709	1.8%
DQ05	Universal data elements	Disabling condition missing	21	913	2.3%
DQ06	Universal data elements	Household without exactly one head of household	13	726	1.8%
DQ07	Universal data elements	Prior living situation missing	25	730	3.4%
DQ08	Exits	Exit destination missing	65	651	10.0%
DQ09	Exits	Exit date before entry date	2	651	0.3%
DQ10	Timeliness	Entry recorded late	94	789	11.9%
DQ11	Timeliness	Exit recorded late	57	651	8.8%
DQ12	Consistency	Overlapping residential stays	10	382	2.6%
DQ13	Consistency	Possible duplicate client records	12	712	1.7%

DQ14	Timeliness	Shelter or outreach enrollment open over 90 days	31	86	36.0%
------	------------	--	----	----	-------

Error rate by project

Each cell is the share of that project's checked records flagged, with the count beneath. Agencies get their own column as a monthly scorecard.

Check	Demo Emergency Shelter (entry/exit) ES entry/exit	Demo Night-by-Night Shelter ES night-by-night	Demo Street Outreach Street outreach	Demo Transitional Housing TH	Demo Permanent Supportive Housing PH-PSH	Demo Rapid Re-Housing PH-RRH
DQ01 Name missing, unknown or refused	2.3% 7 of 298	1.0% 1 of 98	2.0% 3 of 150	2.4% 2 of 84	0.0% 0 of 119	2.4% 4 of 164
DQ02 SSN missing or not a valid SSN	6.4% 19 of 298	7.1% 7 of 98	8.7% 13 of 150	7.1% 6 of 84	10.9% 13 of 119	6.1% 10 of 164
DQ03 Date of birth missing or impossible	1.7% 5 of 298	1.0% 1 of 98	2.0% 3 of 150	1.2% 1 of 84	1.7% 2 of 119	1.2% 2 of 164
DQ04 Veteran status missing for adults	1.9% 4 of 206	3.1% 3 of 96	2.0% 3 of 147	0.0% 0 of 61	1.2% 1 of 82	1.7% 2 of 117
DQ05 Disabling condition missing	2.3% 7 of 298	3.1% 3 of 98	2.7% 4 of 150	1.2% 1 of 84	2.5% 3 of 119	1.8% 3 of 164
DQ06 Household without exactly one head of household	1.9% 4 of 211	3.1% 3 of 98	1.3% 2 of 150	1.6% 1 of 62	1.2% 1 of 85	1.7% 2 of 120
DQ07 Prior living situation missing	3.8% 8 of 213	4.1% 4 of 98	4.0% 6 of 150	1.6% 1 of 62	4.7% 4 of 86	1.7% 2 of 121
DQ08 Exit destination missing	3.7% 9 of 244	16.3% 16 of 98	29.7% 35 of 118	0.0% 0 of 50	0.0% 0 of 40	5.0% 5 of 101
DQ09 Exit date before entry date	0.0% 0 of 244	1.0% 1 of 98	0.8% 1 of 118	0.0% 0 of 50	0.0% 0 of 40	0.0% 0 of 101
DQ10 Entry recorded late	12.7% 34 of 268	7.2% 7 of 97	15.5% 20 of 129	8.1% 6 of 74	8.2% 7 of 85	14.7% 20 of 136
DQ11 Exit recorded late	9.0% 22 of 244	7.1% 7 of 98	9.3% 11 of 118	12.0% 6 of 50	12.5% 5 of 40	5.9% 6 of 101

DQ12 Overlapping residential stays	2.7% 8 of 298	n/a 0 of 0	n/a 0 of 0	2.4% 2 of 84	n/a 0 of 0	n/a 0 of 0
DQ14 Shelter or outreach enrollment open over 90 days	25.9% 14 of 54	n/a 0 of 0	53.1% 17 of 32	n/a 0 of 0	n/a 0 of 0	n/a 0 of 0

Rules

Check	Rule
DQ01	Name data quality is 8, 9 or 99, or first or last name is blank.
DQ02	SSN data quality is 8, 9 or 99, or a "full SSN" fails format rules (000/666/9xx area, 00 group, 0000 serial, repeated or sequential digits).
DQ03	DOB quality is 8, 9 or 99, DOB is blank, DOB falls after the entry date, or age at entry is over 105.
DQ04	Client is 18 or older at entry and veteran status is 8, 9, 99 or blank.
DQ05	Disabling condition is 8, 9, 99 or blank.
DQ06	A household's active enrollments contain zero or more than one RelationshipToHoH = 1.
DQ07	For heads of household and adults, prior living situation is 8, 9, 99 or blank.
DQ08	An exit in the period has destination 8, 9, 99, 30 or blank.
DQ09	The exit date is earlier than the enrollment's entry date.
DQ10	The enrollment was created more than 4 business days after the entry date (weekends and listed holidays excluded).
DQ11	The exit was created more than 4 business days after the exit date.
DQ12	The same person has two emergency shelter (entry/exit), transitional housing or safe haven stays whose dates overlap.
DQ13	Two PersonIDs share a valid full SSN, or share DOB, last name and the first three letters of first name.
DQ14	An emergency shelter (entry/exit) or street outreach enrollment has no exit and started more than 90 days before period end; review for a missed exit.

Missing-value codes: 8 = client doesn't know, 9 = client prefers not to answer, 99 = data not collected; 30 = no exit interview completed. Business days exclude weekends and the dates in `holidays.csv`.

Example flagged records

Two per check. The full list is in `flagged_records.csv`, one row per record, ready to split by agency.

Check	Project	Record	Detail
-------	---------	--------	--------

DQ01	101	E000062	PersonalID C00622, name quality 99
DQ01	101	E000305	PersonalID C00092, name quality 9
DQ02	101	E000031	PersonalID C00198, SSN quality 8
DQ02	101	E000041	PersonalID C00513, SSN quality 8
DQ03	101	E000527	PersonalID C00089, DOB missing
DQ03	101	E000653	PersonalID C00295, DOB missing
DQ04	101	E000318	PersonalID C00658, veteran status 8
DQ04	101	E000776	PersonalID C00371, veteran status 9
DQ05	101	E000077	disabling condition 8
DQ05	101	E000114	disabling condition 9
DQ06	101	H00082	0 heads of household among 1 members
DQ06	101	H00259	0 heads of household among 1 members
DQ07	101	E000041	prior living situation 9
DQ07	101	E000081	prior living situation 9
DQ08	101	E000044	exit 2025-11-03, destination 99
DQ08	101	E000479	exit 2026-04-26, destination 99
DQ09	102	E000184	entry 2026-09-22, exit 2026-09-16
DQ09	201	E000613	entry 2026-04-17, exit 2026-04-06
DQ10	101	E000018	entry 2026-06-16, created 2026-07-15, 19 business days
DQ10	101	E000062	entry 2026-07-21, created 2026-08-25, 25 business days
DQ11	101	E000044	exit 2025-11-03, created 2025-12-08, 23 business days
DQ11	101	E000077	exit 2026-09-27, created 2026-10-08, 9 business days
DQ12	101	E000144/E001034	PersonalID C00559: 2025-12-07 to 2025-12-18 overlaps 2025-12-08 to open
DQ12	101	E000234/E001038	PersonalID C00057: 2026-09-02 to open overlaps 2026-09-07 to open
DQ13		C00009/C00708	same DOB, last name and first-name start
DQ13		C00236/C00712	same SSN

DQ14	101	E000027	entry 2026-03-23, open 191 days
DQ14	101	E000060	entry 2026-05-31, open 122 days

Database administration checks

Read-only export snapshot. An attempted edit to a loaded client record: blocked as designed: export snapshot is read-only; correct the record in the HMIS.

Follow-up log. Changing a finding's status from open to "sent to agency" to "corrected in HMIS" wrote 2 history rows through the status trigger.

Backup and restore. The database was copied with SQLite's online backup API to `hmis_reporting_2026-09-30.sqlite`; the copy passed `PRAGMA integrity_check` (ok) and every table's row count and content hash matched the live database:

Table	Rows	Restore comparison
project	6	match
client	712	match
enrollment	1038	match
exit	776	match
dq_finding	442	match
holiday	12	match

Index use. Query plans for the two self-join checks:

DQ12: SCAN e; SEARCH p USING INDEX sqlite_autoindex_project_1 (ProjectID=?); SEARCH c USING COVERING INDEX sqlite_autoindex_client_1 (PersonalID=?); SEARCH x USING INDEX sqlite_autoindex_exit_2 (EnrollmentID=?) LEFT-JOIN; SEARCH c USING COVERING INDEX sqlite_autoindex_client_1 (PersonalID=?); SEARCH e USING INDEX ix_enrollment_person (PersonalID=? AND EntryDate<?); SEARCH p USING INDEX sqlite_autoindex_project_1 (ProjectID=?); SEARCH x USING INDEX sqlite_autoindex_exit_2 (EnrollmentID=?) LEFT-JOIN

DQ13: SCAN b; MULTI-INDEX OR; INDEX 1; SEARCH a USING INDEX ix_client_ssn (SSN=?); INDEX 2; SEARCH a USING INDEX ix_client_dob_name (DOB=?)

Load checks. Foreign key errors after load: 0. Source file SHA-256 hashes are stored with the load record.