

Data Quality Assessment Report

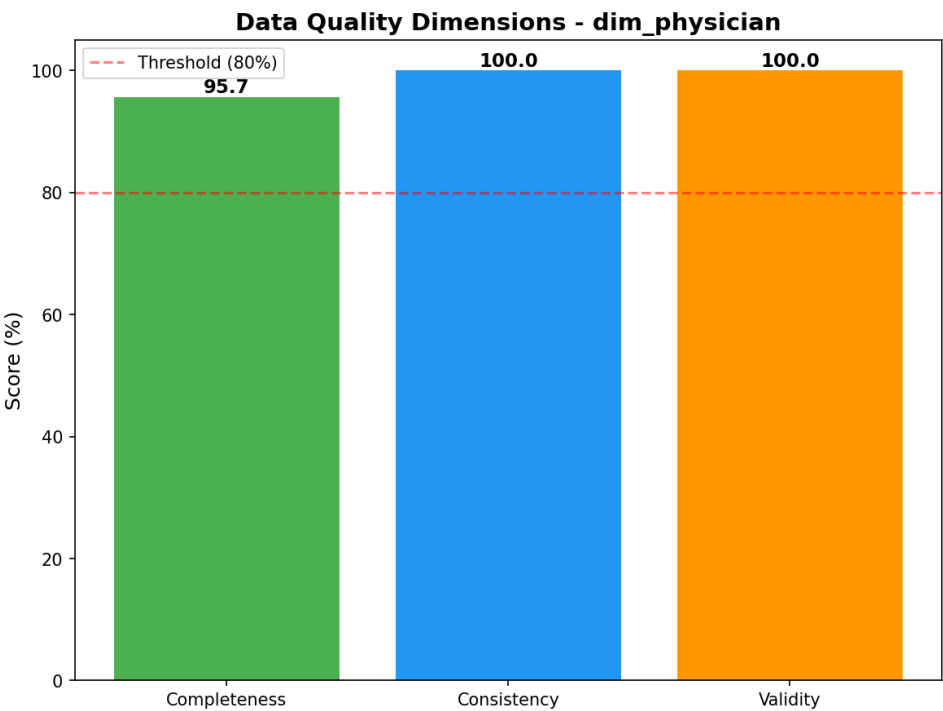
Dataset: dim_physician

Generated: 2025-10-17 09:46:30

Executive Summary

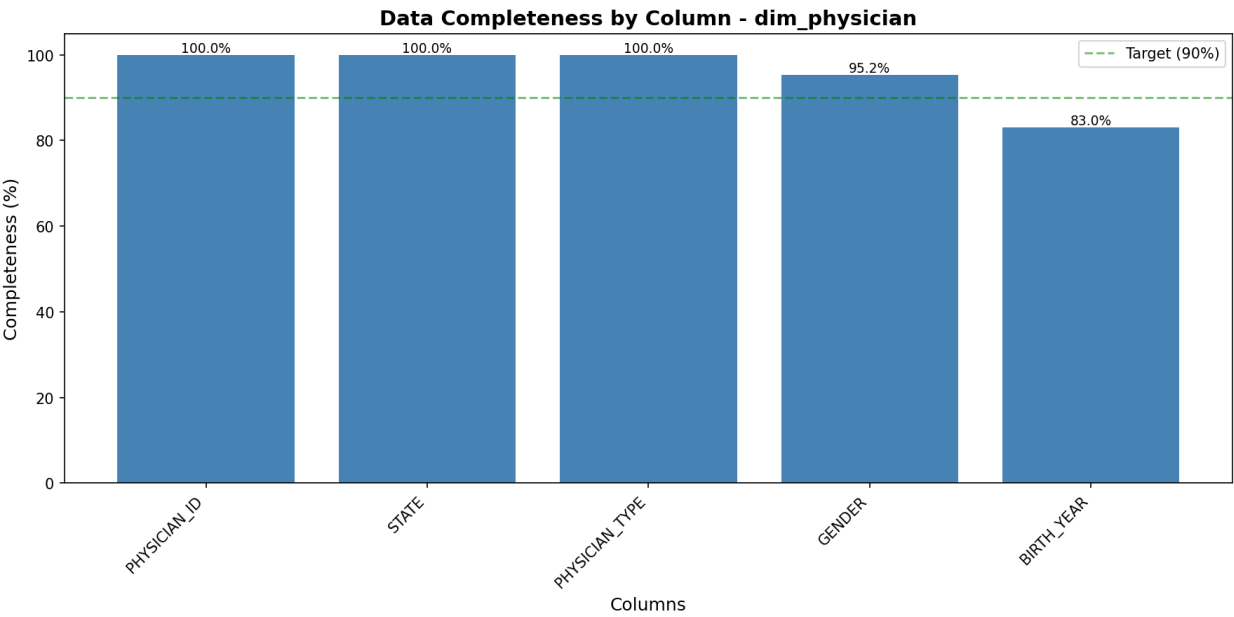
Metric	Value
Overall Quality Score	98.26%
Quality Grade	A (Excellent)
Total Records	25,343
Total Columns	5
Completeness Score	95.65%
Consistency Score	100.0%
Validity Score	100%

Data Quality Visualization



1. Data Completeness Analysis

Column	Total Records	Non-Null	Null Count	Null %	Completeness
PHYSICIAN_ID	25,343	25,343	0	0.0%	100.0%
STATE	25,343	25,341	2	0.01%	99.99%
PHYSICIAN_TYPE	25,343	25,343	0	0.0%	100.0%
GENDER	25,343	24,137	1,206	4.76%	95.24%
BIRTH_YEAR	25,343	21,039	4,304	16.98%	83.02%



2. Data Consistency Analysis

Duplicate Records: 0 (0.0%)

No format issues detected.

3. Data Validity Analysis

Column	Data Type	Unique Values	Unique %
PHYSICIAN_ID	int64	25,343	100.0%
STATE	object	57	0.22%
PHYSICIAN_TYPE	object	199	0.79%
GENDER	object	2	0.01%
BIRTH_YEAR	float64	68	0.27%

Statistical Summary for Numeric Columns

Column	Min	Max	Mean	Median	Std Dev	Outliers
PHYSICIAN_ID	1	25343	12672.0	12672.0	7316.04	0 (0.0%)
BIRTH_YEAR	1900.0	2000.0	1969.88	1971.0	11.42	41 (0.16%)