

DATA WAREHOUSE IM VERGLEICH ZU EINEM DATA LAKE

Merkmale	Data Warehouse	Data Lake
Daten	Relationale Daten aus Transaktionssystemen, Betriebsdatenbanken und branchenrelevanten Anwendungen	Nichtrelationale und relationale Daten aus IoT-Geräten, Webseiten, mobilen Apps, sozialen Medien und Unternehmensanwendungen
Schema	Wird vor der Implementierung des Data Warehouse entwickelt (Schema-on-Write)	Wird zum Analysezeitpunkt geschrieben (Schema-on-Read)
Preis/Leistung	Die schnellsten Abfrageergebnisse mit teurerem Speicher	Immer schnellere Abfrageergebnisse mit günstigem Speicher
Datenqualität	Sorgfältig kuratierte Daten, die als zentraler wahrer Datenbestand gelten	Jegliche Daten, ob kuratiert oder nicht (z. B. Rohdaten)
Benutzer	Businessanalysten, Datenwissenschaftler und Datenentwickler	Datenwissenschaftler, Datenentwickler und Businessanalysten (mit kuratierten Daten)
Analysen	Batch-Berichte, BI und grafische Darstellung	Maschinelles Lernen, prädiktive Analyse, Daten-Discovery und Profilierung

Quelle: [Amazon \(2019\)](#)

RELATIONAL DATABASE MANAGEMENT SYSTEM (RDBMS)

- > DW besteht in der Regel aus **relationalen** Datenbanken (Tabellen) und enthält nur **strukturierte** Daten
- > Wird dann als relationales Datenbankmanagementsystem oder **RDBMS** (Relational Database Management System) bezeichnet
- > Zum Abfragen und Manipulieren der Daten wird überwiegend die Datenbanksprache **SQL** (Structured Query Language) eingesetzt

Relational Model

