بيانات مهيكلة مُعالجة، Data Warehouse يُشبه كلاهما مستودعات بيانات ضخمة، لكنهما يختلفان في الهيكل والغرض. يخزن وهو مُصمم لمحللي الأعمال لإنشاء تقارير، بتكلفة .ETL وعملية Schema-on-Write مثل بيانات مبيعات مُنسقة، باستخدام جميع أنواع البيانات (مهيكلة، شبه مهيكلة، وغير مهيكلة) Data Lake أعلى وأداء أسرع، ولكنه محدود المرونة. بينما يخزن وهو مُوجه لعلماء البيانات لاستكشاف البيانات وتطوير نماذج، .ELT و عملية Schema-on-Read بشكلها الخام، باستخدام لتحليل أعمال منظّم، بينما يُستخدم Data Warehouse بتكلفة أقل ومرونة عالية، لكن أدائه أقل للاستعلامات العشوائية. يُستخدم حلاً هجيناً يجمع مزايا كليهما. اختيار Data Lakehouse في مشاريع تعلم الآلة وتحليل البيانات الضخمة. يُشكل Data Lake للبيانات المتنوعة والمشاريع Data Lake للتقارير السريعة، و Data Warehouse: الأنسب يعتمد على احتياجات المؤسسة المعقدة.

Summarized by © lakhasly.com