



مستودعات البيانات والنمذجة البعدية: مسارات ETL وELT وبحيرات البيانات ومستودع البحيرة



مستودعات البيانات والنمذجة البعدية: مسارات ETL وELT وبحيرات البيانات ومستودع البحيرة

مقدمة:

ما زالت مؤسسات كثيرة تبني تقاريرها من جداول بيانات مجمعة يدويا من الانظمة التشغيلية، فتتعارض الارقام ويضيع التاريخ مع كل تحديث، ويتطلب كل سؤال جديد استخراجا جديدا. تعالج مستودعات البيانات والنمذجة البعدية هذه المشكلة بدمج بيانات المصادر في مخزن تحليلي موحد مصمم لتقارير سريعة ومتسقة. تؤهل هذه الدورة من كور كونسبت المختصين لتصميم مخططات النجمة، وبناء مسارات ETL وELT موثوقة، وتحديد موقع بحيرة البيانات أو مستودع البحيرة في البنية، وتأمين المنصة وضبط أدائها وتكلفتها. ويخرج المشاركون بمخطط بنية مستودع البيانات والنموذج البعدي لمجال أعمال ذي أولوية.

أهداف الدورة:

- تقييم الاحتياجات التحليلية وانظمة المصدر لاختيار بنية المستودع أو بحيرة البيانات أو مستودع البحيرة المناسبة للمؤسسة
- تصميم مخططات النجمة وندفة الثلج بدرجة تفصيل صحيحة ونوع جدول حقائق مناسب وأبعاد موحدة ومفاتيح بدلة
- بناء إجراءات تحميل ETL وELT تعالج منطقة التجويز والتحميل الترايدي والتقاط تغيرات البيانات والابعاد بطيئة التغير
- دمج اختبارات جودة البيانات وفحوص المطابقة وتتبع النسب داخل مسارات المستودع
- تطبيق ضوابط الوصول وأمن مستوى الصف والإخفاء والتشفير مع ضبط التقسيم والتجميع وموارد الحوسبة لتحسين الأداء وخفض التكلفة
- إعداد مخطط بنية مستودع البيانات والنموذج البعدي جاهزا للعرض على لجنة مراجعة البنية

الفئة المستهدفة:

- المدبرون والقادة المسؤولون عن منصات ذكاء الاعمال وتسليم التقارير على مستوى المؤسسة
- قادة هندسة البيانات المسؤولون عن مسارات الاستيعاب والتحويل والتحميل
- مهندسو الحلول والبيانات المسؤولون عن قرارات تصميم المنصات التحليلية
- مديرو التحليلات الذين تعتمد فرقهم على بيانات تاريخية ومشتركة بين الاقسام يوثق بها
- مديرو عمليات المنصات والحوسبة السحابية المسؤولون عن توفر المستودع وأمنه وتكلفة تشغيله

محاوِر الدورة:

اليوم 1: أسس مستودعات البيانات وتشخيص الوضع التحليلي الحالي

- أحمال العمل التشغيلية مقابل التحليلية: خصائص OLTP و OLAP
- مصنع المعلومات المؤسسي Inmon Corporate Information Factory مقابل بنية الناقل البعدية لدى Kimball
- مصفوفة المقارنة بين مستودع البيانات وبحيرة البيانات ومستودع البحيرة
- حصر انظمة المصدر وتقييم نقاط الخلل في التقارير
- ترتيب أولويات عمليات الاعمال بمصفوفة الناقل المؤسسي Enterprise Bus Matrix

اليوم 2: النمذجة البعدية والبنى المرجعية

- خطوات Kimball الاربعة للتصميم البعدي: عملية الاعمال ودرجة التفصيل والابعاد والحقائق
- المفاضلة بين مخطط النجمة ومخطط ندفة الثلج في التصميم
- جداول حقائق المعاملات واللقطات الدورية واللقطات التراكمية
- الابعاد الموحدة والمفاتيح البديلة وأبعاد التاريخ متعددة الادوار
- المحاور والروابط والتوابع في Data Vault 2.0 منهجية نمذجة طبقة التكامل بوصفها طبقة تكامل

اليوم 3: بناء مسارات ETL و ELT

- اختيار نمط ETL أو ELT وتصميم منطقة التجهيز
- منطق تحميل الابعاد بطيئة التغير من النوع 1 والنوع 2 والنوع 3
- التحميل التزايدى بعلاّمة التحميل الاخير والتقاط تغيرات البيانات
- بنية الميدياليات Medallion Architecture: مسؤوليات الطبقات البرونزية والفضية والذهبية
- تنسيق المسارات وجدولة الاعتماديات والتعافي من الاخفاق

اليوم 4: منصات مستودع البحيرة وأمن المستودع والأداء والتكلفة

- مقارنة صيغ الجداول المفتوحة: Apache Hudi و Apache Iceberg و Delta Lake
- اختبارات جودة البيانات داخل المسار ومطابقة الاعداد وتتبع النسب على مستوى العمود
- أمن المستودع: الوصول القائم على الادوار وأمن مستوى الصف والإخفاء الديناميكي والتشفير
- التقسيم والتجميع والعروض المادية لتحسين أداء الاستعلامات
- فصل التخزين عن الحوسبة وعزل أحمال العمل وضوابط تكلفة الحوسبة السحابية

اليوم 5: حالات مختبرية ومخطط بنية المستودع

- مختبر مبيعات التجزئة: بناء مخطط نجمة من بيانات نقاط البيع
- مختبر العمليات الصحية: الحقائق المتأخرة الوصول وتتبع تاريخ الإبعاد
- مختبر الطبقة الدلالية: مقاييس مشتركة وربط أدوات ذكاء الأعمال بالطبقة الذهبية
- صياغة مخطط بنية مستودع البيانات والنموذج البعدي
- عرض المخطط والدفاع عنه أمام لجنة مراجعة البنية

المهارات التي ستكتسبها:

- النمذجة البعدية للبيانات
- اختيار بنية المستودع
- هندسة مسارات ETL و ELT
- التقاط تغيرات البيانات
- تقييم صيغ جداول مستودع البحيرة
- أمن الوصول إلى المستودع
- ضبط أداء الاستعلامات
- إدارة تكلفة منصات البيانات السحابية

لماذا تحضر هذه الدورة:

- العودة بمخطط بنية مستودع البيانات والنموذج البعدي لمجال أعمال تحتاج مؤسستك إلى تقاريره
- إيقاف فقدان التاريخ وإعادة بناء المستخرجات عبر تحميل الإبعاد والحقائق بطريقة متكررة قابلة للتدقيق
- إقناع الرعاة بالخيار المناسب بين المستودع وبحيرة البيانات ومستودع البحيرة وتكلفة تشغيله
- اختبار قرارات التصميم في مختبرات عملية مع مختصي بيانات من قطاعات التجزئة والرعاية الصحية والمالية والقطاع العام

الخاتمة:

يكسب مستودع البيانات ثقة المستخدمين حين يطابق نموذجه طريقة قياس المؤسسة لأدائها، وتحمل مساراته التاريخ بالطريقة نفسها كل ليلة. تنتقل الدورة من مفاهيم المستودع وخيارات البنية، إلى النمذجة البعدية وبناء مسارات ETL و ELT، ثم صيغ جداول مستودع البحيرة وفحوص الجودة داخل المسار والأمن والأداء والتكلفة. ويطبق اليوم الأخير ذلك في مختبرات من التجزئة والرعاية الصحية، ليخرج المشاركون بمخطط بنية مستودع البيانات والنموذج البعدي يعودون به إلى فريق المنصة في مؤسساتهم.