# الكلية التطبيقية

## تحليل البيانات ونظم دعم القرار

سنة رابعة

مدرس المقرر

د.م. عبدالرزاق المحمد المحمود

2026 - 2025

# (1)

## مستودعات البيانات وتقانات معالجة التحليلات المباشرة OLAP

## Data warehouse and OLAP technology

### مقدمة

يعتبر تحليل البيانات أحد أهم المجالات في عصر المعلومات، التي تُمكّن الشركات، المؤسسات، والباحثين من اتخاذ قرارات مستنيرة بناء على الأنماط والمعلومات المستخلصة من البيانات.

مع تزايد حجم البيانات المتولدة يوميا، أصبح تحليل البيانات والتنقيب في البيانات من الأدوات الحيوية في مجالات مختلفة مثل الطب، التسويق، التكنولوجيا، والعلوم الاجتماعية.

في هذه المحاضرة، سنستعرض المفاهيم الأساسية لتحليل البيانات، أنواعه، أهدافه، ومراحله، بالإضافة إلى التنقيب في البيانات وتطبيقاته العملية ومقدمة عن مستودعات البيانات.

### مفهوم البيانات

البيانات Data هي مجموعة من القيم أو الحقائق التي تمثل معلومات حول كائنات أو أحداث أو عمليات. يمكن أن تكون البيانات أرقاماً، نصوصاً، صوراً، أو حتى تسجيلات صوتية، تنقسم البيانات إلى أنواع رئيسية:

- **البيانات المنظمة Structured Data** يقصد بها البيانات المخزنة في جداول أو قواعد بيانات، مثل بيانات العملاء في جدول Excel .
- **البيانات غير المنظمة Unstructured Data** يقصد بها البيانات المحفوظة بصيغ مختلفة (كالنصوص، الصور، الفيديو)، من الأمثلة المعروفة (منشورات وسائل التواصل الاجتماعي).
- **البيانات شبه المنظمة Semi-Structured Data** هي بيانات تحتوي على بعض التنظيم، مثل ملفات JSON أو XML .

## مفهوم تحليل البيانات وأهدافه

تحليل البيانات Data Analysis هو عملية فحص البيانات، تنظيمها، وتحويلها إلى معلومات مفيدة تساهم في اتخاذ القرارات الصحيحة. وهذا الأمر يتطلب مهارات خاصة في الإحصاء، البرمجة، بالإضافة لفهم المجال المعرفي الذي يطبق عليه التحليل (الاقتصاد، الطب، التجارة،،،).

يهدف تحليل البيانات إلى تحقيق عدة أغراض منها:

1. **دعم اتخاذ القرارات:** توفير رؤى دقيقة لاتخاذ قرارات استراتيجية، (اختيار موقع جديد لمتجر).
2. **تحسين الأداء:** تحديد نقاط الضعف في العمليات (تأخير في سلسلة التوريد ◄ اقتراح حلول).
3. **اكتشاف الأنماط:** الكشف عن اتجاهات أو علاقات غير واضحة، (ارتباط المبيعات بموسم معين).
4. **تقليل المخاطر:** التنبؤ بالمشكلات المحتملة، (مخاطر الاحتيال المالي في البنوك).

## مراحل تحليل البيانات

يتكون تحليل البيانات من سلسلة مراحل منظمة لضمان الدقة والفعالية

1. **تحديد الأهداف:** فهم المشكلة التي تحتاج إلى إجابة (ما الذي يؤثر على انخفاض المبيعات؟).
2. **جمع البيانات:** الحصول على البيانات من مصادر موثوقة (قواعد البيانات، استبيانات).
3. **تنظيف البيانات:** إزالة القيم المفقودة، التكرارات، أو الأخطاء لضمان جودة البيانات.
4. **تحليل البيانات:** باستخدام تقنيات إحصائية أو نماذج تعلم الآلة (Excel, Python).
5. **تفسير النتائج:** ترجمة النتائج إلى رؤى مفهومة، (اكتشاف أن المبيعات تنخفض بسبب نقص المخزون)
6. **تصور النتائج:** عرض النتائج باستخدام الرسوم البيانية، التقارير، أو لوحات المعلومات، لتسهيل التواصل مع أصحاب القرار.

## أنواع تحليل البيانات

تتنوع أشكال تحليل البيانات بناء على الهدف والأدوات المستخدمة تشمل الأشكال الرئيسية :

1. **التحليل الوصفي Descriptive Analysis** يركز على وصف البيانات الحالية باستخدام مقاييس مثل المتوسط، الوسيط، والانحراف المعياري ، مثال: تحليل متوسط مبيعات متجر خلال عام.
2. **التحليل التشخيصي Diagnostic Analysis** يهدف إلى فهم أسباب الظواهر مثال تحديد سبب انخفاض المبيعات.
3. **التحليل التنبؤي Predictive Analysis** يستخدم نماذج إحصائية أو تعلم الآلة للتنبؤ بالأحداث المستقبلية ، مثال: التنبؤ بطلب العملاء في الموسم القادم.
4. **التحليل التوجيهي Prescriptive Analysis** يقدم توصيات بناء على التحليل مثال اقتراح استراتيجيات تسويقية لزيادة المبيعات بناء على التنبؤات.

## التنقيب في البيانات

التنقيب في البيانات Data Mining هو عملية استخدام لخوارزميات وتقنيات متقدمة لاكتشاف الأنماط، العلاقات، أو الاتجاهات المخفية في عينات البيانات الكبيرة، يعتمد التنقيب على تقنيات مثل التعلم الآلي، الإحصاء، والذكاء الاصطناعي.

على سبيل المثال، يمكن لشركة تجارة إلكترونية استخدام التنقيب لاكتشاف أن العملاء الذين يشترون منتجاً معيناً، غالباً ما يشترون منتجا آخر.

يختلف التنقيب في البيانات عن تحليل البيانات التقليدي بتركيزه على اكتشاف الأنماط غير الواضحة في البيانات الضخمة Big Data .

## أهداف التنقيب في البيانات

تشمل أهداف التنقيب في البيانات

1. **التصنيف Classification** تصنيف البيانات إلى فئات محددة، مثل تصنيف العملاء إلى "محتمل الشراء" أو "غير محتمل".
2. **التجميع أو العنقدة Clustering** تجميع البيانات المتشابهة دون تصنيف مسبق، مثل تجميع العملاء بناء على سلوك الشراء.
3. **اكتشاف العلاقات Association Rule Mining** تحديد العلاقات بين الأحداث، مثل "الذين يشترون الحليب غالبا يشترون الخبز".
4. **التنبؤ Prediction** التنبؤ بالقيم المستقبلية، مثل التنبؤ بأسعار الأسهم.
5. **كشف الانحرافات Anomaly Detection** تحديد البيانات غير الطبيعية، مثل اكتشاف معاملات احتيالية في البنوك.

## المراحل الرئيسية لعملية التنقيب في البيانات

تتبع عملية التنقيب في البيانات نموذجا قياسيا مثل CRISP-DM
(Cross-Industry Standard Process for Data Mining) ويشمل المراحل:

1. **فهم المجال Business Understanding** تحديد أهداف المشروع ومتطلباته، مثل زيادة المبيعات .
2. **فهم البيانات Data Understanding** جمع البيانات الأولية واستكشاف خصائصها مثل الحجم، النوع، والجودة .
3. **إعداد البيانات Data Preparation** تنظيف البيانات، تحويلها، واختيار الجزء المناسب للتحليل
4. **النمذجة Modeling** اختيار وتطبيق خوارزميات التنقيب مثل شجرة القرار، K-Means، أو الشبكات العصبية.
5. **التقييم Evaluation** تقييم النماذج للتأكد من دقتها وملاءمتها للأهداف .
6. **النشر Deployment** تطبيق النتائج في بيئة العمل، مثل دمج نموذج تنبؤي في نظام التسويق .

## كيفية دراسة البيانات

تتطلب دراسة البيانات منهجية منظمة، تشمل الأدوات والتقنيات التالية :

1. **جمع البيانات** استخدام مصادر مثل قواعد البيانات SQL ، واجهات برمجة التطبيقات APIs ، أو أدوات الزحف Web Scraping
2. **استكشاف البيانات** **Exploratory Data Analysis – EDA** استخدام الرسوم البيانية مثل الهيستوغرام، المخططات المنتشرة والإحصائيات مثل المتوسط، الانحراف المعياري لفهم البيانات
3. **تنظيف البيانات** معالجة القيم المفقودة، إزالة التكرارات، وتصحيح الأخطاء باستخدام أدوات مثل Pandas في Python
4. **تحويل البيانات** إلى صيغ مناسبة، مثل تحويل النصوص إلى أرقام Encoding أو تقليل الأبعاد PCA
5. **اختيار الأدوات** استخدام برمجيات مثل
   - **Python** مكتبات مثل Pandas ، NumPy، Scikit-learn
   - **R** للتحليل الإحصائي
   - **Tableau/Power BI** لتصور البيانات
6. **التقييم والتحقق** استخدام مقاييس الدقة Accuracy ، Precision ، والاستدعاء Recall لتقييم النماذج

## أمثلة على تطبيقات تنقيب البيانات

تُستخدم تقنيات التنقيب في البيانات في مجالات متنوعة مثل:

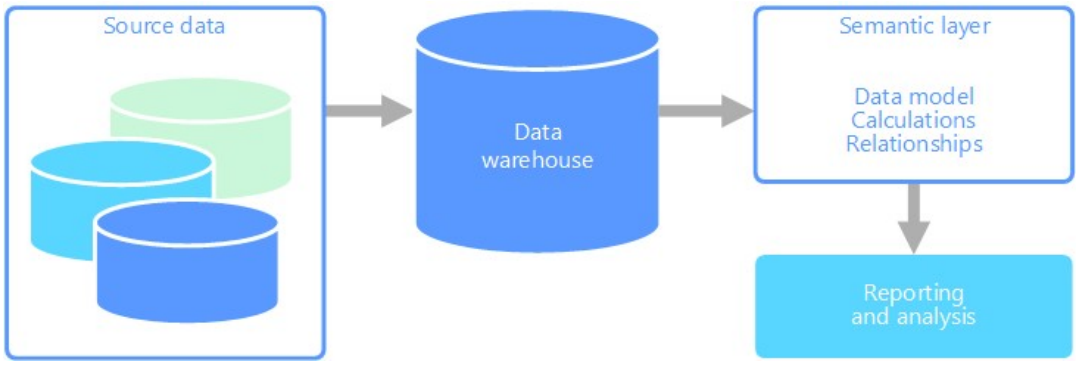
1. **التجارة الإلكترونية** تستخدم أمازون التنقيب لاكتشاف العلاقات بين المنتجات مثل "الذين اشتروا هذا الغرض اشتروا أيضا ذاك الغرض" باستخدام قواعد الارتباط
2. **الرعاية الصحية** تحليل بيانات المرضى للتنبؤ بأمراض مثل السكري باستخدام خوارزميات التصنيف

3. **البنوك** الكشف عن الاحتيال المالي عبر اكتشاف الانحرافات في المعاملات
4. **التسويق** تجميع العملاء إلى مجموعات بناء على سلوك الشراء لتصميم حملات تسويقية مستهدفة
5. **الاتصالات** التنبؤ بخسارة العملاء، لتحديد العملاء المحتمل رحيلهم وتقديم عروض للاحتفاظ بهم
6. **التعليم** تحليل أداء الطلاب لاكتشاف الأنماط التي تساعد في تحسين المناهج الدراسية

ننصح بتجربة أدوات مثل Python لتطبيق تقنيات التحليل و التنقيب عملياً.

## مستودعات البيانات وتقنية معالجة التحليلات المباشرة OLAP

يعتمد تحليل البيانات المتقدمة، على تقنيات تسمح بتخزين كميات هائلة من البيانات وتحليلها بكفاءة، من أهم هذه التقنيات، مستودعات البيانات Data Warehouses وتقنية معالجة التحليلات المباشرة OLAP Online Analytical Processing تُستخدم هذه التقنيات في مجالات مختلفة مثل التجارة، الرعاية الصحية، وإدارة سلسلة التوريد.




## مستودع البيانات

هو قاعدة بيانات مركزية مصممة لتخزين وإدارة كميات ضخمة من البيانات ومن مصادر متنوعة لأغراض التحليل وصنع القرار، يتميز مستودع البيانات عن قواعد البيانات التقليدية OLTP Online Transaction Processing بالخصائص التالية:

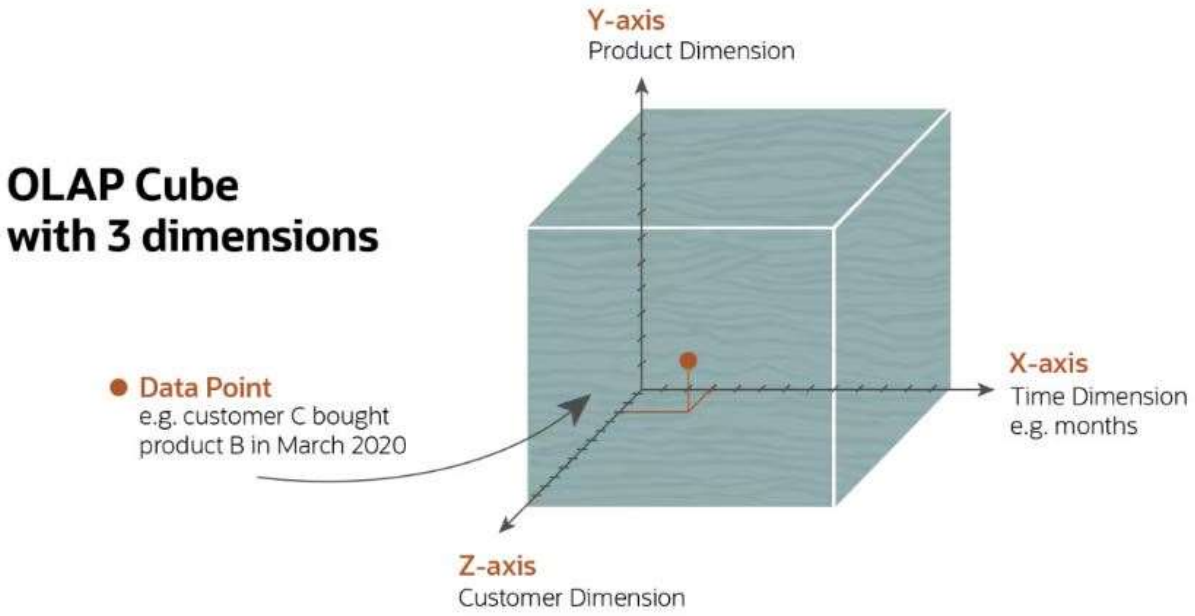
- **موضوعي التوجه Subject-Oriented** فهو يركز على موضوعات محددة مثل المبيعات، العملاء، أو المخزون، بدلاً من العمليات اليومية.

- **متكامل Integrated** يجمع البيانات من مصادر متنوعة مثل قواعد بيانات، ملفات Excel ، أنظمة ERP ويوَّحد تنسيقاتها.
- **غير متغير بمرور الوقت Time-Variant** يحتفظ بالبيانات التاريخية لفترات طويلة لتحليل الاتجاهات.
- **مُستقر Non-Volatile** البيانات في المستودع لا تُحدَّث أو تُحذف كما في قواعد البيانات التشغيلية.

على سبيل المثال، يمكن استخدام مستودع البيانات لتخزين معلومات المبيعات، البضائع، العملاء لتحليل الأداء على مدى سنوات.

## تقنية OLAP

معالجة التحليلات المباشرة Online Analytical Processing (OLAP) هي تقنية تتيح تحليل البيانات متعدد الأبعاد بسرعة ومرونة، تُستخدم OLAP لتنفيذ استعلامات معقدة على مستودعات البيانات، مثل تحليل المبيعات حسب المنطقة، المنتج، والفترة الزمنية، تعتمد OLAP على نموذج البيانات متعدد الأبعاد Multidimensional Data Model ، حيث يتم تنظيم البيانات في **مكعبات بيانات Data Cubes**



**خصائص الـ OLAP**

- **تحليل متعدد الأبعاد:** يسمح بتحليل البيانات من زوايا مختلفة، مثل الوقت، الموقع، والمنتج.
- **أداء عالي:** يعتمد على تقنيات مثل التخزين المؤقت Caching وفهرسة البيانات.
- **تنوع العمليات:**

✓ التجميع Roll-Up : تلخيص البيانات مثل (من المبيعات اليومية إلى الشهرية) .
✓ التفصيل Drill-Down : التعمق في التفاصيل مثل من المبيعات السنوية إلى اليومية.
✓ التقطيع Slice : اختيار بُعد معين مثل مبيعات منطقة معينة .
✓ التكعيب Dice : اختيار نطاق من البيانات مثل مبيعات منتج معين في فترة زمنية.
✓ الدوران Pivot : إعادة ترتيب الأبعاد لعرض البيانات بطريقة مختلفة.

**أمثلة على أدواتOLAP**

✓ Microsoft SQL Server Analysis Services (SSAS)
✓ Oracle OLAP
✓ IBM Cognos
✓ أدوات مفتوحة المصدر مثل Apache Kylin