الدرس 1 (السداسي الثاني)

تحليل البيانات

BOURAS Meriem

جامعة أم البواقي-الجزائر

العام الدراسي 2024-2025

الوحدة 1

مقدمة في تحليل البيانات

مهارة ضرورية في العصر الرقمي

I. تعريف تحليل البيانات

تحليلات البيانات تُحوّل البيانات الخام إلى رؤى قابلة للتنفيذ.

تحليل البيانات هو عملية استخراج المعلومات المفيدة من البيانات من خلال تطبيق تقنيات إحصائية ورياضيتيه. يُستخدم في مجالات متعددة مثل الاقتصاد، المالية، البحث العلمي، والذكاء الاصطناعي لتحليل الاتجاهات واتخاذ قرارات مبنية على الأدلة.

تحليل مجموعة من الملاحظات بهدف إيجاد علاقات غير متوقعة وتلخيص البيانات بطريقة جديدة بحيث تكون أكثر قابلية للفهم ومفيدة لحامليها.

• أهمية تحليل البيانات

01

تحسين اتخاذ القرار من خلال توفير رؤى مبنية على البيانات.

02

الكشف عن الأنماط والعلاقات بين المتغيرات.

03

تحسين كفاءة العمليات وتقليل التكاليف من خلال تحليل البيانات التشغيلية.

04

التنبؤ بالمستقبل عبر النماذج الإحصائية والتعلم الآلي.

• مراحل تحليل البيانات

جمع البيانات

- تحديد مصادر البيانات (استطلاعات، قواعد بيانات، بيانات مالية...).
- التأكد من دقة وجودة البيانات قبل معالجتها.

تنظيف البيانات

- إزالة القيم الشاذة والتعامل مع البيانات المفقودة.
- تحويل البيانات إلى تنسيقات مناسبة للتحليل.

استكشاف البيانات

- تحليل التوزيع الإحصائي للمتغيرات.
- البحث عن العلاقات والارتباطات بين المتغيرات.

بناء النماذج وتحليلها

- اختيار النموذج الإحصائي المناسب.
- تدريب النموذج على البيانات.
- تقييم دقة النموذج

تفسير النتائج واتخاذ القرارات

- تحليل النتائج الإحصائية وتقديم التوصيات.
- إعداد التقارير لمساعدة صناع القرار.

• أنواع تحليل البيانات

01

التحليل الاستكشافي للبيانات

يهدف إلى فهم توزيع البيانات والعلاقات بين المتغيرات باستخدام:

- الإحصائيات الوصفية مثل المتوسط، الوسيط، الانحراف المعياري.
- تقنيات التصور البياني مثل المخططات الصندوقية، المخططات الشريطية، ورسوم التشتت.

02

التحليل العاملي (Factor Analysis)

تقنية تُستخدم لتقليل أبعاد البيانات الضخمة واستخراج العوامل الرئيسية

- **تحليل المكونات الرئيسية (PCA):** يستخدم لاختزال البيانات المعقدة إلى متغيرات أقل عددًا مع الحفاظ على أكبر قدر من المعلومات
- **تحليل المراسلات العاملية (AFC):** يُستخدم مع البيانات النوعية لفهم العلاقات بين الفئات المختلفة.

03

تحليل العنقودي (Clustering Analysis)

يهدف إلى تصنيف البيانات إلى مجموعات بناءً على التشابه بينها.

- التجميع الهرمي التصاعدي (CAH): يُستخدم لإنشاء هرم تصنيفي لمجموعات البيانات.
- خوارزمية K-Means: تُستخدم لتقسيم البيانات إلى عدد محدد من المجموعات بناءً على القرب الإحصائي.

04

تحليل الانحدار(Regression Analysis)

يُستخدم لفهم العلاقة بين متغير تابع ومتغيرات مستقلة، مثل:

- الانحدار الخطي: يُستخدم للتنبؤ بالقيم الرقمية مثل المبيعات أو الأرباح.
- الانحدار اللوجستي: يُستخدم لتصنيف البيانات إلى فئات مثل "ناجح/فاشل".

05

تحليل السلاسل الزمنية (Time Series Analysis)

يُستخدم لتحليل البيانات المتغيرة عبر الزمن والتنبؤ بالقيم المستقبلية، مثل.

- نماذج ARIMA: تُستخدم لتوقع القيم المستقبلية بناءً على البيانات التاريخية.