

❖ الوحدة الأولى: علم البيانات

تعريف علم البيانات:

هو علم يجمع بين عدة مجالات (على سبيل المثال: علوم الحاسب والإحصاء والرياضيات) ويعمل على تحليل البيانات لاستخراج معلومات ذات مغزى تؤدي إلى معرفة محددة.

تعريف البيانات:

هو مجموعة من الحقائق أو الكلمات أو الأرقام أو حتى وصف لأشياء لم يتم تحليلها أو معالجتها بأي شكل من الأشكال، وتسمى أيضا البيانات الأولية، حيث تعني كلمة أولية أنها غير معالجة.

تعريف المعلومات:

عندما تُعالج البيانات الأولية فإنها تتحول إلى معلومات، ولذلك يشير مصطلح المعلومات إلى البيانات المعالجة التي لها معنى في سياق محدد ومفيد، بينما يسمى إجراء هذه المعالجة: معالجة البيانات.

أوجه الاختلاف بين البيانات والمعلومات:

البيانات	المعلومات
ليس لها معنى بصورتها الحقيقية	يجب أن تحمل معنى منطقيا
كلمات وأرقام غير معالجة	بيانات تمت معالجتها
مادة أولية	منتج نهائي
أكثر عمومية	أكثر تحديداً
تستخدم كمدخلات للحاسب	تستخدم كمخرجات للحاسب

تعريف المعرفة:

تنتج المعرفة من معالجة المعلومات وفهمها ويؤدي ذلك إلى استنتاجات وقرارات مختلفة.

أوجه الاختلاف بين المعلومات و المعرفة :

المعلومات	المعرفة
بيانات تمت معالجتها لتصبح ذات سياق مفهوم	استنتاج من المعلومات يساعد في اتخاذ القرارات.
وحدها لا تكفي للتوصل إلى الاستنتاجات أو القرارات حول مسألة معينة	توفر القدرة على اتخاذ تنبؤات واتخاذ قرارات
يتم الحصول على نفس المعلومات عند تحليل نفس البيانات	تختلف باختلاف العالم أو الباحث الذي يدرس المعلومات

مثال على كل من البيانات والمعلومات والمعرفة



أنواع البيانات:

- 1/ **بيانات رقمية:** مثل عدد الفعاليات التي تقام في مدينة
- 2/ **بيانات ابداعية:** مثل اسم دولة "المملكة العربية السعودية"
- 3/ **بيانات ابداعية رقمية:** مثل: #، و\$، او تاريخ 3-مارس-2020
- 4/ **بيانات صوتية:** مثل التسجيلات الصوتية الإرشادية
- 5/ **بيانات رسومية:** مثل الرسم البياني الخاص بأعداد الزوار
- 6/ **بيانات مقاطع فيديو:** مثل الإعلان التلفزيوني الخاص بحملة سياحية

اشكال عرض البيانات:

- 1/ **بيانات ثابتة:** هي البيانات التي لا تتغير بعد تسجيلها
- 2/ **بيانات متغيرة:** هي البيانات التي قد تتغير بعد تسجيلها ويجب تحديثها باستمرار

مفهوم ترميز البيانات:

تتيح عملية الترميز تنظيم البيانات وترتيبها بطريقة محددة وذلك باستخدام رموز مختلفة مثل الأرقام أو الحروف أو الكلمات القصيرة
مثل: رموز المطارات ورموز العملات

مزايا ترميز البيانات:

1. إدخال أسرع للبيانات
2. تأخذ مساحة أقل
3. تسريع عمليات البحث عن البيانات

عيوب ترميز البيانات:

1. معنى غامض للبيانات
2. صعوبة فهم الترميز
3. الرموز المستخدمة قد تستنفد

الرموز الشريطية BarcodeS

الرمز الشريطي هو ملصق به خطوط سوداء رفيعة إلى جانب التنوع في الأرقام المختلفة. تُستخدم للمساعدة في تنظيم المعلومات وفهرستها أو وضع علامة على أسعار المنتجات.

رموز الاستجابة السريعة QR

هو الجيل الثاني من الرمز الشريطي والذي يتكون من خطوط سوداء متجاورة ومختلفة السمك ويحتوي على مزيد من المعلومات. قد يشير رمز الاستجابة السريعة إلى محتوى إلكتروني مثل: المواقع الإلكترونية، أو مقاطع الفيديو، أو الملفات الرقمية، ويمكن قراءة هذا الرمز باستخدام كاميرات الهواتف الذكية.



رقم الكتاب المعياري الدولي ISBN

وهو رقم فريد يستخدمه الناشر والمكتبات ومحلات بيع الكتب لتحديد عناوين الكتب وإصداراتها.

جودة المعلومات:

تُعدُّ جودة المعلومات عاملاً مهماً وتعبّر عن مدى استخدام المعلومات في اتخاذ القرارات. يساعد ضمان جودة المعلومات على التحديد الدقيق للاحتياجات الفعلية لتنفيذ المشروعات، وكذلك لتوجيه الخدمات بشكل فعال، وزيادة الكفاءة في كل يوم عمل. وبالمقارنة

يمكن أن تتسبب المعلومات غير الدقيقة في حدوث خللٍ في الأعمال، وتقلل من الكفاءة وتؤدي إلى التأخير في إنجاز المشروعات

معايير جودة المعلومات:

- 1/ **الدقة:** الدقة هي التأكد من صحة المعلومات
- 2/ **الملاءمة:** يجب أن تكون المعلومات مرتبطة بموضوعك أو بالسؤال البحثي
- 3/ **التوقيت:** مدى حداثة المعلومات ومناسبتها لموضوع البحث،
- 4/ **مستوى التفاصيل:** يعد وجود الكثير من التفاصيل على المعلومات المطلوبة أمراً صعباً، بينما تؤدي قلة التفاصيل إلى صعوبة فهم المعلومات
- 5/ **الكفاية:** التأكد من أن المعلومات المقدمة تعطي صورة كاملة عن الواقع

جمع البيانات

وهي عملية جمع الحقائق والأرقام والكلمات للمتغيرات المستهدفة وتحسينها ويمكن جمع البيانات باستخدام أجهزة مختلفة مثل المستشعرات ومسجلات البيانات.



مصادر البيانات الرئيسية والثانوية

1. **مصادر البيانات الرئيسية:** بيانات لم تجمع من قبل ويمكن جمعها من المستشعرات ومسجلات البيانات وحتى من الاستبيانات
 2. **مصادر البيانات الثانوية:** يأتي هذا النوع من البيانات عندما تستخدم مصدر البيانات الرئيس لإنتاج بيانات أخرى
- ♦ يمكن تصنيف مصادر البيانات الثانوية إلى مصادر داخلية ومصادر خارجية

التحقق من صحة إدخال البيانات

هو أي نشاط يتحقق من أن البيانات المدخلة تأتي من مجموعة من القيم المعتمدة، وتتوافق مع القواعد المقبولة للبيانات

فإذا كانت البيانات متوافقة مع القواعد ستقبل، وإلا فسترفض.

أنواع التحقق من صحة إدخال البيانات

التحقق من البحث (Lookup check)	يساعد على تقليل الأخطاء باستخدام قائمة محدودة من القيم المحددة مسبقًا.
التحقق من التواجد (Presence check)	يجعل عملية الإدخال إلزامية في الخلية مما يضمن عدم تركها فارغة.
التحقق من الطول (Length check)	يهدف إلى التأكد من أن الرموز والحروف تُدخل بنطاق طول محدد.
التحقق من النطاق (Range check)	يُستخدم للتأكد من أن الأرقام التي تُدخل تقع ضمن نطاق معين ويشمل حدين هما: الحد الأدنى (Minimum limit) والحد الأعلى (Maximum limit).
التحقق من الصيغة (Format check)	يُستخدم للتأكد من أن البيانات تأتي بصيغة محددة مسبقًا ولن يُسمح بأي صيغة أخرى يتم إدخالها في الخلية.
التحقق من النوع (Type check)	يضمن إدخال المستخدمين لنوع القيمة الصحيح في حقل محدد.

التنبؤ Forecasting

هو عملية بناء التوقعات المستقبلية بناءً على البيانات السابقة

*إن مصطلحي التنبؤ والتوقع متشابهان، ولكنهما غير متطابقين، حيث أن التنبؤ مصطلح أكثر عمومية.

(بإمكاننا التنبؤ باستخدام برنامج مايكروسوفت إكسل)

أنواع مخططات التنبؤ

المخطط الخطي	المخطط العمودي
يقدم تحليل سريع للبيانات	تساعد في توضيح المقارنة بين مجموعات البيانات
يمكن ملاحظة التغييرات بسهولة خلال فترة زمنية محددة	تلخص كمية كبيرة من البيانات في شكل مرئي يسهل تفسيره
يناسب مجموعات البيانات التي يصل عددها إلى ٥٠ قيمة	تجعل الاتجاهات الإحصائية أسهل في الملاحظة
يساعد في عمل تنبؤات حول نتائج البيانات التي لم تُسجل بعد	تساعد في دراسة الأنماط على مدى فترة طويلة من الزمن

فاصل الثقة

يُعرف فاصل الثقة في الإحصاء على أنه: نطاق من القيم المقدرة لمعامل غير معروف كل التنبؤات يوجد بها قدر من عدم اليقين فيها، فهي ليست قيمًا "حقيقية" تم قياسها أو تم الحصول عليها من البحث، إنها قيم "تقديرية"، مما يعني أنها قيم غير موجودة بالفعل.

الانحدار الخطي

هو نوع أساسي وأكثر استخدامًا في تحليل التنبؤ؛ لأنه يسمح لك بتلخيص ودراسة العلاقات بين متغيرين نوعيين qualitative أو كمييين quantitative

التشفير Encryption

التشفير هو وسيلة لحماية البيانات عن طريق إخفائها عن الأشخاص غير المرغوب بهم

هناك نوعان رئيسيان من التشفير:

- 1/ التشفير المتماثل: وفيه يستخدم نفس المفتاح لتشفير وفك تشفير
- 2/ التشفير غير المتماثل: وفيه تشفير البيانات بمفتاح عام أولاً ثم فك تشفيرها باستخدام مفتاح خاص

(من المهم تشفير رسائل البريد الإلكتروني قبل إرسالها وتشفير القرص الصلب لحماية وحدة التخزين الداخلية الموجودة في الحاسب بكاملها)