فهرس الدرس

أنواع البيانات وتنسيقاتها المختلفة

1 تعريف البيانات

2 أنواع البيانات

3 البيانات والمعلومات

4 مصادر إضافية

5 تطبيقات

بما أن البيانات هي المادة الخام الأساسية التي سنعتمد عليها في إعداد تقاريرنا واتخاذ قراراتنا الحيوية، فمن المهم أن نتعرف على أنواعها المختلفة. يظن البعض أن البيانات تنقسم إلى نوعين فقط: بيانات كمية، وبيانات وصفية، لكن في الحقيقة يوجد العديد من أنواع البيانات المختلفة التي تأخذ تنسيقات وأنماطًا متعددة، وينبغي أن نتعرف عليها حتى نتمكن من الاستفادة منها وتحديد طريقة التعامل معها.

تعريف البيانات

ينبغي وضع تعريف للبيانات، بما أنها المادة الخام التي سنتعامل معها. ومن الشائع تعريفها بأنها مجموعة من الحروف أو الكلمات أو الرموز أو المؤشرات أو الإحصاءات الأولية المعنية بموضوع محدد، ولا يتحتم أن تكون دائمًا في صيغة رقمية؛ بل من الدارج أن تأخذ صيغًا أخرى غير رقمية.

- بعبارةٍ أخرى: هي كل ما يمكن رصده وتوثيقه حول ظاهرة معينة، ويسهل وضعه في هيكل مُنظَّم، سواء كان رقمًا أو نصًّا أو تاريخًا.

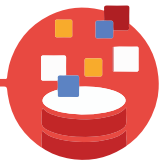
يوجد الآن العديد من المصطلحات المتداولة حول العالم، التي تحمل شِقًّا مُتعلِّقًا بالبيانات، مثل «البيانات مفتوحة المصدر» التي تُنتجها المؤسسات الحكومية في إطار مباشرتها لوظائفها. لا يجب أن تكون هذه البيانات مُدرجة ضمن البيانات المحمية بمقتضى التشريع الجاري العمل به، وبخاصة القانون المُتعلق بحماية الشؤون الشخصية؛ وحقوق الملكية الأدبية والفنية؛ ذلك لأنها بيانات يحق للمواطن الاطلاع عليها.

تأتي «البيانات الضخمة» من بين المصطلحات المتداولة الأخرى، وقد وضع لها معهد ماكينزي العالمي تعريفًا عام 2011 بأنها مجموعة البيانات التي يفوق حجمها قدرة أدوات قواعد البيانات التقليدية على التخزين والتحليل؛ ويتطلب العمل عليها أن يضعها مُطورو البرمجيات في صيغة تُسهِّل عملية تحليلها وعرضها على الجمهور.

أنواع البيانات

يوجد العديد من أنواع البيانات المختلفة التي تأخذ تنسيقات وأنماطًا متعددة، وينبغي أن نتعرف عليها حتى نتمكن من الاستفادة منها وتحديد طريقة التعامل معها ومعالجتها وتحليلها.

من حيث المصدر



أولية - ثانوية

الاسم	التعريف	مثال
أولية	يُمكن جمع البيانات من مصادرها الأساسية عن طريق الرصد والمشاهدة أو التجربة، أو استطلاعات الرأي ومسوح العينات، وكذلك القياس والتحليل. يتسم هذا النوع من البيانات بالدقة، لكنه مُكلِّف وقد يتطلب وقتًا وجهدًا لجمعه مباشرةً.	جمع وزارة الصحة لأعداد الحالات التي تلقَّت اللقاح الخاص بفيروس كورونا، عن طريق رصد الحالات التي تتوجه يوميًّا إلى مراكز التطعيم المنتشرة حول الجمهورية.
ثانوية	هي البيانات المنشورة عن مصدر آخر؛ قام بنقلها هو نفسه عن مصدرها الأول. يكون المصدر هنا ناقلًا للبيانات، وليس جامعًا لها. ويتميز هذا النوع من البيانات بأنه غير مُكلِّف ومتوفر، لكنه لا يتمتع بنفس مستوى الثقة والدقة التي تتمتع بها البيانات الأولية.	البيانات المنشورة عن منظمة الصحة العالمية حول أعداد المُلقَّحين ضد فيروس كورونا بالدول العربية، ذلك أنَّ المصدر الأول للبيانات هو وزارة الصحة بكل دولة باعتبارها الجهة التي ترصد وتحصر أعداد المُلقَّحين وتنشرها في تقاريرها الدورية.

داخلية - خارجية

الاسم	التعريف	مثال
البيانات الداخلية	هي البيانات التي يتم إنتاجها داخل المؤسسة؛ غالبًا ما تكون متعلقة بأداء المؤسسة والعاملين بها.	ساعات العمل والحضور والانصراف، عدد الموظفين الذين تقدموا بطلب من أجل الحصول على إجازة خلال فترة معينة، إنتاجية العاملين، وحجم الإنتاج اليومي.

الاسم	التعريف	مثال
البيانات الخارجية	هي البيانات التي يتم إنتاجها خارج أنظمة المؤسسة وبعيدًا عن أنظم التشغيل المعمول بها داخليًّا؛ غالبًا ما تستعين بها المؤسسة لتحليل حجم السوق ورصد أداء المنافسين.	نسبة إستحواذ المنافسين على السوق، حجم الطلب على سلعة محددة، متوسط الأجور في الشركات الآخرى، معدل التسعير الذي يتّبعه المنافسون لتسعير منتجاتهم.

- يرتبط الأصل في التفرقة بين النوعين بحماية الخصوصية، وحق الحصول على البيانات، ذلك أن البيانات التي يتم إنتاجها داخليًّا ليس من المُعتاد أن تكون متاحة أمام مُحللين من خارج المؤسسة. وعند مشاركتها على سبيل التعاون في مشروع محدد، من الضروري عدم مشاركتها مع أي طرف آخر.

مجموعة آليًّا - مجموعة عن طريق الإنسان

الاسم	التعريف	مثال
البيانات المجموعة آليًّا	بفضل إنترنت الأشياء (IoT)، وأجهزة الاستشعار والتقنيات الأخرى، يمكن جمع البيانات تلقائيًّا عن طريق تكنولوجيا المعلومات والبنية التحتية للاتصالات والسيارات الذكية والأجهزة المحمولة وغيرها. يُعرف هذا النوع من البيانات باسم بيانات الآلة، إذ يتم إنتاجها كليًّا عن طريق الآلات؛ دون تعليمات بشرية.	المخالفات المرورية التي ترصدها كاميرات مراقبة السرعة على الطريق، أو عدد الخطوات التي تمشيها يوميًّا وتحسبها حسّاسات الساعة الذكية أو تطبيقات الهواتف المحمولة.

الاسم	التعريف	مثال
البيانات المجموعة عن طريق الإنسان	يُنتجها الإنسان عن طريق التوثيق اليدوي أو الإدخال المباشر على الحاسب الآلي.	السعرات الحرارية المُتناولة يوميًا التي يوصي طبيب التغذية بتدوينها، أو الوقت المستقطع للانتقال من نقطة جغرافية إلى أخرى داخل المدينة باستخدام المواصلات العامة أو السيارة الخاصة.

من حيث النوع

كمية - نوعية

الاسم	التعريف	مثال
البيانات الكمية	هي البيانات التي تُعبِّر عن كمية، أو رقم، أو شيء يمكن قياسه، أو شيء يمكن إجراء عمليات حسابية عليه.	الطول، والعمر، والوزن، وعدد سنوات الخبرة.
البيانات النوعية (الوصفية)	وتكون بيانات تفسيرية أو وصفية؛ يمكن ملاحظتها لكن لا يمكن قياسها، لأنها تكون واصفةً لأمرٍ ما.	اللون، والطعم، والملمس، والجنس، والعِرق، والدين، والجنس، وغيرها من الصفات التي تُحدد ملامح الأشياء والأشخاص.

ليست كل البيانات العددية كمية

مثال: أرقام الهواتف - أرقام بطاقات الائتمان - رموز الأمان - أكواد الطُلَّاب.

تأخذ كل هذه البيانات صيغة رقمية، لكنها لا تعكس قيمة معينة ولا يمكن الاستفادة من معالجتها حسابيًّا.

هل يمكن اعتبار مجموع القيم داخل عمود يضم أرقام هواتف عملاء أحد المطاعم بيانات كمية؟ أو قيمة ذات دلالة ومعنى؟

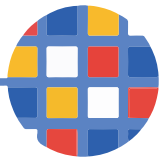
تراتبية - غير تراتبية

الاسم	التعريف	مثال
البيانات النوعية التراتبية	هى بيانات غير عددية، ولا يمكن إجراء عمليات حسابية عليها. الفرق بينها وبين البيانات الاسمية هو عملية المفاضلة والترتيب. يمكن ترتيب البيانات تصاعديًّا أو تنازليًّا، وليس بشكل أبجدي فقط.	البيانات المُتعلقة بتقديرات الطُلَّاب (إمتياز، جيد جدًّا، جيد، مقبول). مستوى التعليم (ابتدائي، إعدادي، ثانوي، جامعي). معدل الدخل الشهري (مرتفع، منخفض، أقل من المنخفض).
البيانات النوعية غير التراتبية	نوع من البيانات النوعية التي لا تُصنَّف بناءً على ترتيب محدد. تكون في صورة غير عددية، ولا يمكن المُفاضلة بينها.	الحالة الاجتماعية (متزوج، أعزب، مطلق، أرمل). النوع (ذكر، أنثى).

مستمرة - منفصلة

الاسم	التعريف	مثال
البيانات المستمرة	هي بيانات عددية يمكن أن تأخذ جميع القيم الممكنة ضمن نطاق يسمى البيانات المستمرة. بعبارة أخرى، هي بيانات يمكن أن تحتوي على كسور في شكل أنصاف وأرباع، ويتم تحديدها من خلال القياس.	البيانات الخاصة بالطول، أو الوزن، بشرط ألَّا يتم تقريبها إلى أقرب عدد صحيح.
البيانات المنفصلة	نوع من البيانات الكمية التي تأخذ قيمة عددية محدودة، وتكون فيها المتغيرات أرقامًا صحيحة يمكن حصرها، ولا يمكن تقسيمها إلى أجزاء أصغر. بعبارة أخرى، لا يكون فيها أنصاف أو أرباع، وتُحدَّدْ عن طريق العد. يتكون هذا النوع من أرقام صحيحة (أرقام موجبة وسالبة، على سبيل المثال: 10، 100، 1000... إلخ).	عدد أفراد الأسرة، عدد المقررات الدراسية، عدد أجهزة الهواتف الذكية المُباعة العام الماضي في أحد المتاجر، عدد لاعبي فريق كرة القدم في المنتخب الوطني.

من حيث الشكل



مُنظَّمة - شبه مُنظَّمة - غير مُنظَّمة

الاسم	التعريف	مثال
البيانات المُنظَّمة	هي بيانات تلتزم بنموذج محدد مُسبَّقًا، وبالتالي يَسهُل تحليلها. عادةً ما تُنتَج وتُنظَّم على هيئة جداول (Tables)؛ تحتوي كل خلية أو حقل على قيمة منفصلة. تكون هذه البيانات مُنظَّمة تنظيمًا إحترافيًّا، بحيث تكون لها جميعًا فهارس مُحددة تساعدنا على استرجاع البيانات بطريقة سريعة ومُنسَّقة.	جداول نتائج الطُلَّاب بنهاية أحد الفصول الدراسية، جداول الناخبين في الدوائر الانتخابية، قواعد بيانات الكتب المتاحة في مكتبة الجامعة أو مكتبة أحد دور النشر.
البيانات شبه المُنظَّمة	هي شكل من أشكال البيانات المُنظَّمة، التي لا تخضع للبنية النموذجية أو الجدولية للبيانات المرتبطة بقواعد البيانات، إذ إنها تفتقر إلى مخطط أو نموذج ثابت للجداول. يحتاج هذا النوع من البيانات إلى معالجة أولًا قبل التحليل.	صفحات الويب، أرشيف المواقع الإلكترونية، الرسائل البريدية، سيناريوهات الأفلام والمسلسلات، قائمة المنتجات الأكثر مبيعًا بأحد المتاجر الرقمية.
البيانات غير المُنظَّمة	هى بيانات لا تخضع لنموذج بيانات مُحدد مسبَّقًا بهيكل مُنظَّم وثابت. وعلى الرغم من صعوبة التعامل مع هذا النوع من البيانات ضمن بيئة قواعد البيانات التقليدية، يمكن أن تضم هذه البيانات معلومات قيِّمة جدًّا ربما تكون أكثر ثراءً من البيانات المنظمة؛ مثل التواريخ والأرقام والحقائق.	ملفات الصوت، ملفات الفيديو، ملفات النصوص، الأخبار في الصحف والمواقع الإلكترونية.

مثال:

تدخل مقاطع فيديوهات الأفلام بصيغة MP4 في نطاق البيانات غير المُنظَّمة، إذا كانت في صورتها الأساسية. لكن مع تفريغها وكتابة اسم المتحدث ونص ما قاله في الفيديو، داخل ملف Word، تُصبح بيانات شبه مُنظَّمة. وتتحول إلى بيانات مُنظَّمة إذا تم تفريغها في شكل جدول يحتوي على عدد من الأعمدة؛ عمود باسم المتحدث، وآخر لوقت التحدث، وثالث للطرف الثاني الموجه له الكلام، وهكذا.

طولية - عرضية

الاسم	التعريف	مثال
البيانات الطولية	هي بيانات تحتوي على أعمدة أقل وصفوف أكثر، ويُمثِّل كل عمود مُتغيرًا بحالات تكراره.	جداول نتائج الطُلَّاب بنهاية أحد الفصول الدراسية، جداول الناخبين في الدوائر الانتخابية، قواعد بيانات الكتب المتاحة في مكتبة الجامعة أو مكتبة أحد دور النشر.
البيانات العرضية	هي بيانات تحتوي على صفوف أقل وأعمدة أكثر، إذ يُعرض كل متغير على عدة أعمدة.	جدول نتائج الانتخابات البرلمانية في مصر. في حالة البيانات العرضية، يُقسَّم الجدول إلى عدد من الأعمدة؛ يكون لكل حزب عمود مستقل، ثم عمود يحتوي على عدد المقاعد، ثم أعمدة متتالية للفترات الزمنية.

مثال: البيانات العرضية

البلد	2019	2020
أ	100	1
ب	80	2

مثال: البيانات الطولية

البلد	العام	الهاتف المحمول
أ	2020	1
ب	2019	80
أ	2019	100
ب	2020	2

تمتاز البيانات الطولية بسهولة تحليلها، إذ تكون عملية التحليل على مستوى العمود الذي يمكن تنقيحه إلى عدة مستويات مختلفة.

تفصيلية - مُجَمَّعة

الاسم	التعريف	مثال
البيانات التفصيلية	هي عبارة عن جدول يحتوي على بيانات مفردة؛ يضم كل سطر أقل درجة من التفاصيل. بعبارة أخرى، هي بيانات مُقسَّمة وفقًا لفئات فرعية تفصيلية، وتكون كبيرة تُعطي تفاصيل لكل شيء.	جدول بيانات لمجموعة من الطُلَّاب؛ يحتوي العمود الأول منها على اسم الطالب، أو رقم جلوسه، أو الكود التعريفي الخاص به في الجامعة.
البيانات المُجَمَّعة	هي عبارة عن بيانات تفصيلية لكنها جمعت بإحدى خصائصها. بعبارةٍ أخرى، هي تلخيص للبيانات التفصيلية؛ أي أنها البيانات القادرة على تجميع سجلات البيانات الكبيرة وتلخيصها.	جدول بيانات لمجموعة من الطُلَّاب؛ يحتوي كل صف على مؤشرات مُجمَّعة مثل الجنسية، التخصص الدراسي، النوع، فئات المجموع التراكمي.

مثال للبيانات المُجَمَّعة:

يحتوي الجدول التالي على أعداد المصابين والمتوفين بفيروس كورونا لكل من المصريين والجنسيات الأخرى:

	المصريون	الأجانب	الإجمالي
عدد الوفيات	##	##	####
عدد الإصابات	##	##	####

مثال للبيانات التفصيلية

من شأن التاريخ أن يُمكِّننا من رصد تطور الحالات يوميًّا، ومعرفة في أي أيام الأسبوع تزداد الحالات وتقل، فضلًا عن تحديد الشهر الأعلى من حيث عدد الحالات والشهر الأقل. أمَّا عمود كود المريض؛ يوفر خصوصية للمرضى، بل ويمكن أن يكون كودًا مرجعيًّا فقط لتنظيم البيانات دون أن يرتبط بأية بيانات شخصية عنهم. وفي عمود الحالة، يمكن تقسيم الحالات إلى مصابين ووفيَّات. بينما نُخصِّص عمود النوع لتوضيح أي الحالات ذكر وأيها أنثى. كذلك يمكن معرفة الشرائح العمرية الأكثر عرضة للوفاة من خانة العمر، فيما يساعد عمود الجنسية في معرفة معلومات أكثر تفصيلًا عن جنسيات المصابين والمتوفين.

التاريخ	كود المريض	الحالة	النوع	العُمر	الجنسية

تذكَّر:

قد تكون البيانات المُجمَّعة مُضَلِّلة في أغلب الحالات. عند الإعلان عن ارتفاع معدلات التدخين في الوطن العربي بنسبة 10% بين عام وآخر، لا يعني هذا بالضرورة ارتفاع المُعدَّل في جميع الدول العربية. ربما تكون بعض النسب قد تراجعت في بعض الدول، وارتفعت في البعض الآخر. وعليه، ينبغي أن نسعى للاطلاع على البيانات التفصيلية التي تعطينا صورة أدق عن كل بند في سجل البيانات.

مقروءة آليًّا - غير مقروءة آليًّا

الاسم	التعريف	مثال
البيانات المقروءة آليًّا	هي البيانات التي يَسهُل معالجتها باستخدام الحاسب الآلي، و إخضاعها لعمليات فرز وتنقيح وبحث؛ ناهيك عن إجراء عمليات حسابية.	جداول البيانات في صيغة Xlx.xlxs
البيانات غير المقروءة آليًّا	هي البيانات التي يمكن للبشر قراءتها بصورة طبيعية، ويَصعُب على الحاسب الآلي فهمها أو إجراء عمليات عليها.	الوثائق المحمية، الوثائق المُصوَّرة هاتفيًّا، أو الملفات المحفوظة عن طريق جهاز الماسح الضوئي في صيغة PDF.

أنواع البيانات من حيث صيغ الملفات

أنواع الصيغ المختلفة لملفات البيانات

تأخذ البيانات صيغًا مختلفة؛ يمكن قراءة بعضها آليًّا ومعالجتها بواسطة الحاسب الآلي. من بين هذه الصيغ: XLS التي تُحفظ بها البيانات في شكل جدول يمكن لبرنامج مايكروسوفت إكسل قراءته، وCSV التي تُحفظ بها البيانات على شكل ملف نصي واضح مع فصل كل إدخال للبيانات بفاصلة يمكن أيضًا لبرنامج مايكروسوفت إكسل قراءته، وTSV التي تُحفظ بها البيانات على شكل قيم مفصولة بعلامة (-).

تُعدُّ هذه الصيغ الأكثر شيوعًا مع ملفات البيانات المُنظَّمة، والأنسب للتحليل المباشر باستخدام برامج الحاسب الآلي، فضلًا عن سهولة التعامل معها داخل مختلف برامج جداول البيانات مثل Google Spreadsheets ومايكروسوفت إكسل. وإذا كان بوسعك الحصول على ملفات البيانات التي تبحث عنها بهذه التنسيقات من البداية، فلن تكون مُضطرًا إلى قضاء وقت طويل في عملية التنقية والتنسيق.

عندما تعثر على البيانات التي كنت تبحث عنها في صيغ وتنسيقات أخرى، مثل الوثائق المكتوبة بخط اليد والملفات المحفوظة في صيغة PDF، ينبغي تحويل الجداول المحفوظة في صيغة PDF إلى تنسيق يمكن فتحه باستخدام برنامج جداول البيانات؛

إذ على الرغم من أنه يحتوي على جداول بيانات يمكن قراءتها بالعين البشرية، يجد الحاسب الآلي صعوبة في الوصول إلى البيانات المجدولة بداخله وقراءته.

يعجز الحاسب الآلي أيضًا عن قراءة البيانات داخل الصور الممسوحة ضوئيًّا، ويتعامل معها باعتبارها جدارًا مُتماسكًا وضخمًا؛ ذلك أن البيانات بداخلها تكون مُشفَّرة في صيغة صور تُقرأ وكأنها كتلة واحدة ضخمة وليست أجزاءً مُنفصِلة. يتطلب هذا النوع من الملفات برامج التعرف الضوئي على الأحرف، حتى يتسنى التعرف على النص داخل الصورة.

	شكل البيانات	الصيغة
	مُجَدوَلة	XLS
	صور	JPEG, PNG
	فيديو	MP4, AVI
	نصوص	DOC, TXT, TSV, CSV, PDF
	خرائط	SHP
	صفحات ويب	XML, HTML

- في لغات البرمجة، مثل بايثون وR، تُقسَّم أنواع البيانات حسب طبيعتها إلى:

النوع	مثال
رقم صحيح/Integer	7, 12, 999
رقم مصحوب بكسور (علامة عشرية مثلًا)/Float	3.15, 9.06, 00.13
قيمة منطقية/Boolean	TRUE, FALSE
نصية/String	«Hello World»

البيانات والمعلومات

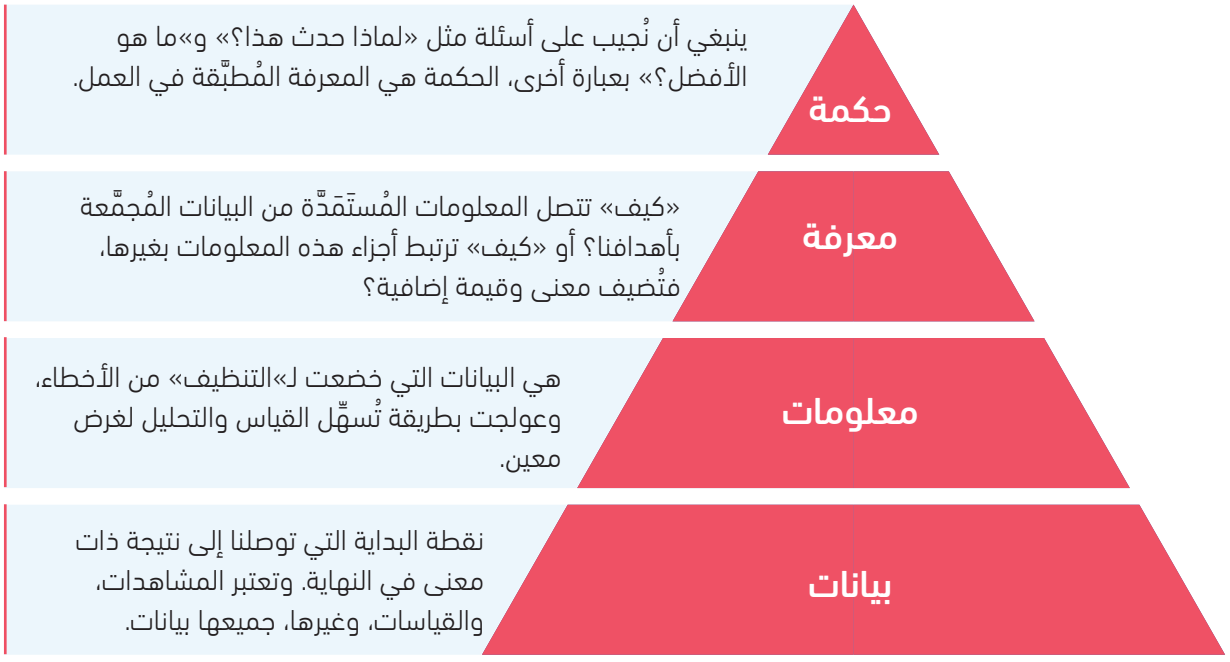
كثيرًا ما يُستخدم مصطلحا البيانات والمعلومات باعتبارهما مترادفين، رغم اختلاف معنى كل منهما. تُعدُّ البيانات المادة الخام لأي عمل، وعادةً ما تأخذ شكل أرقام أو معطيات أولية يُمكن جمعها حول موضوع محدد، ونادرًا ما تكون مفيدة بمفردها.

الفرق بين البيانات والمعلومات

- **البيانات** هي حقائق خام؛ أولية، وغير مُنظَّمة، وينبغي معالجتها حتى تكون ذات مغزى. دائما ما يفسر الناس أو الآلة البيانات لاستنتاج المعاني.
- **المعلومات** هي بيانات خضعت لمعالجة بطريقة ذات مغزى للإجابة على أسئلة مُحددة. معالجة البيانات أو تنظيمها أو تقديمها في سياق معين لجعلها ذات مغزى ومفيدة

هرم الحكمة

- **البيانات** عبارة عن فوضى هائلة لا معنى لها؛ يجب معالجتها حتى تكون ذات مغزى.
- يصف التسلسل الهرمي DIKW pyramid كيف تتحول البيانات إلى معلومات، ثم إلى مجموعة من المعارف، ومن ثم إلى حكمة؛ في سلسلة متتالية هي العناصر المُكَوِّنة لهرم الحكمة.



مثال: درجات الحرارة

البيانات	سجل بدرجات الحرارة المتوقعة خلال الأسبوع القادم.
المعلومات	أي الأيام أعلى في درجات الحرارة، وأيها أقل.
المعرفة	التباين الشديد في حالة الجو، وعدم استقرارها.
الحكمة	إتخاذ القرار المناسب المُتَمثِّل في ارتداء ملابس ملائمة لكل يوم، حسب درجة الحرارة.

مثال: تذاكر الطيران

البيانات	أسعار تذاكر الطيران، من القاهرة إلى جدة، وفقًا للوقت والوزن المُصرَّح به.
المعلومات	نطاق أسعار تذاكر السفر من القاهرة إلى جدة.
المعرفة	أفضل سعر تذكرة، وأفضل يوم للسفر.
الحكمة	إختيار أفضل رحلة، من حيث التوقيت ووالسعر والوزن المُصرَّح به.

الفرق بين المعلومات والمعرفة

أوجه الاختلاف	المعلومات	المعرفة
المعنى	بيانات تُعالج لتصبح ذات سياق مفهوم.	إستنتاجات مُستقاة من المعلومات، وتساعد في اتخاذ القرار.
القدرة على التنبؤ	بمفردها، لا تكفي للتوصل إلى نتائج.	تُساهم في القدرة على التنبؤ، واتخاذ القرارات.
إيصال النتائج إلى الآخرين	يَسهُل توصيلها شفهيًّا أو ورقيًّا أو إلكترونيًّا.	تتطلب إطِّلاعًا على الموضوع المحدد.
النتائج	يَنتُج عنها الفهم.	يَنتُج عنها الإدراك.
الهدف	تُجيب على أسئلة مثل: من؟ ومتى؟ وماذا؟ وأين؟	تجيب على أسئلة مثل: كيف؟ ولماذا؟

مصادر إضافية:

- دليل مبرمج، ما هي البيانات وما هي أنواعها.

https://www.dalilmobarmg.com/tutorials/algorithms/fundamentals/datatypes/

- المجلة العربية للعلوم ونشر الأبحاث: ما هي البيانات في التحليلات الإحصائية.

https://blog.ajsrp.com/%D8%A7%D986%%D988%%D8%A7%D8%B9-%D8%A7%D984%%D8%A8%D98%A%D8%A7%D986%%D8%A7%D8%AA-%D981%%D98%A-%D8%A7%D984%%D8%AA%D8%AD%D984%%D98%A%D984-%%D8%A7%D984%%D8%A7%D8%AD%D8%B5%D8%A7%D8%A6%D98%A/

تمارين

1 - حَدِّد نوع البيانات المناسب وضعه أمام ما يناسبه في الجانب الآخر، إذ يحتوي الجانب الآخر على أسماء رؤوس أعمدة لقاعدة بيانات الأطباء المتوفين جراء الإصابة بفيروس كورونا:

الاختيارات: (بيانات كمية - بيانات وصفية)

البند	الإجابة الصحيحة
العمر	
الجنس	
سنوات الخبرة	
الحالة الاجتماعية	
الوزن	

2 - حَدِّد ما إذا كانت البيانات الآتية منفصلة Discrete أم مستمرة Continuous:

عدد رسائل البريد الإلكتروني المُرسَلة خلال شهر	
أطوال لاعبي فريق كرة السلة، دون تقريب الرقم	
قراءات ضغط الدم	
عدد الطُلَّاب المُتَغَيِّبين عن أداء الاختبار النهائي	
درجة الحرارة	
سرعة السيارة	
عدد السيارات في المرآب	
عدد الطائرات عند مدرج الهبوط	

3 - ضع علامة صح أو خطأ أمام العبارات الآتية:

- ☐ تُعتبر درجات الطُلَّاب وتقديراتهم بيانات كمية Quantitative Data.
- ☐ تُقاس البيانات المستمرة Continuous Data، ولها عدد محدود من القيم.
- ☐ لا يمكن أن تكون البيانات داخل المستندات الممسوحة ضوئيًّا مُنظَّمة Structured Data.
- ☐ تُعتبر المُخالفات المرورية التي يجمعها رجل الشرطة بيانات مجموعة آليًّا Machine Gathered Data.
- ☐ مُعدل التوظيف في منطقة الشرق الأوسط وشمال إفريقيا هو رقم مُجمَّع Aggregated Data.
- ☐ فيما يتعلق بالبيانات، يُعتبَر النص الموجود داخل ملف Word بيانات مُنظَّمة.
- ☐ تؤدي البيانات إلى المعرفة، وتؤدي المعرفة إلى المعلومات.
- ☐ يُطلق اسم المعرفة على البيانات، عند تحليلها وتنظيمها وهيكلتها لتصبح ذات معنى؟
- ☐ يتم الحصول على المعلومات من خلال عملية تحليل البيانات؟

4 - من هذه الصورة، حَدِّد 4 مُتغيرات لبيانات كَمِّية و4 مُتغيرات لبيانات وصفية.

.....

.....

.....

.....

5 - يحتوي الجدول التالي على بيانات طلبة جامعيين. وتشتمل هذه البيانات على الاسم والسن والنوع وسنة الميلاد. اِطَّلِع على الجدول، واستخدم المعلومات الموجودة بداخله، للإجابة على الأسئلة التالية:

أ	ب	ج	د
الاسم	السن	النوع	سنة الميلاد
عُمر	22	ذكر	1998
يوسف	0.19	ذكر	3120
سارة	20		2000
أميرة	0		1999

البيانات في هذا الجدول:

- ☐ مُنظَّمة.
- ☐ شبه مُنظَّمة.
- ☐ غير مُنظَّمة.

البيانات في العمود ب:

- ☐ وصفية.
- ☐ كمية.

البيانات في العمود ج:

- ☐ وصفية.
- ☐ كمية.

إذا كان هذا الجدول في صيغة CSV، كيف سيبدو؟

- ☐ عُمر/22/ذكر/1998
- ☐ عُمر,22,ذكر,1998
- ☐ عُمر::22::ذكر::1998
- ☐ عُمر /22/ ذكر /1998

إذا أردت إضافة عمود آخر (z) لعرض المستوى الأكاديمي لكل طالب (بيانات وصفية)، فعلى ماذا ستحتوي مُدخلات البيانات؟

- ☐ التقدير التراكمي بالأرقام (4.0, 3.8 ...)
- ☐ التقدير التراكمي بالمستوى (امتياز، جيد جدًّا...)
- ☐ النسبة المئوية (95%، 90%...)

للحفاظ على خصوصية الأشخاص في الجدول السابق، يمكن استبدال عمود الأسماء بـ:

- ☐ رقم الهوية.
- ☐ رقم الجلوس.
- ☐ رقم الهاتف.
- ☐ كود الطالب.

تكليف:

البيانات في مكان ما...

البيانات في المدرسة/المستشفى/سكن الطلبة الجامعيين/عيادة طبيب النساء والتوليد/مركز إطفاء الحريق.

الخطوات:

- إختر مكانًا من الأماكن السالف ذكرها (المدرسة/المستشفى....)، وقُم بإعداد جدول بيانات تفصيلية من شأنها أن تكون ضمن قاعدة بيانات ذلك المكان، ثم إقتراح 4 بنود بيانات وصفية؛ بشرط أن تحتوي على بيانات تراتبية وغير تراتبية، و4 بنود أخرى كمية؛ بشرط أن تحتوي على بيانات مستمرة وبيانات منفصلة.
- أنشِئ جدول جوجل Google Sheet، وأَدخِل البنود به.
- انظر إلى جدول البيانات النهائي الذي أنشأته؛ هل هناك أي شيء مفقود؟
- قُم بإعداد نُسخة مُجمَّعة من البيانات التي أعدتها في الخطوة السابقة.
- ما الأسئلة التي قد نتمكن من الإجابة عليها إذا استطعنا ملء الجدول بكل البيانات التي نبحث عنها؟

الجواب:

- الجدول الأول (جدول بيانات تفصيلية):

بيانات كمية				بيانات وصفية			

- رابط الجدول على جوجل:

.....

- الجدول الثاني (النسخة المجمعة):

- الأسئلة التي قد نتمكن من الإجابة عليها إذا استطعنا ملء الجدولين:

.....

.....

.....

تمارين