المملكة العربية السعودية

وزارة التعليم
Ministry of Education

قررت وزارة التعليم تدريس
هذا الكتاب وطبعه على نفقتها

علم البيانات

التعليم الثانوي - نظام المسارات

السنة الثانية

الحلول اون لاين
hulul.online

يوزع مجاناً ولا يباع

البيانات والمعلومات والمعرفة

الحلول اون لاين
hulul.online

تمرينات

1

حدد الجملة الصحيحة والجملة الخاطئة فيما يلي:	صحيحة	خاطئة
1. علم البيانات هو حقل متعدد التخصصات يركّز على استخراج معلومات ذات فائدة من البيانات.	✓	
2. يُطلق اسم المعرفة على البيانات عند تحليلها وتنظيمها وهيكلتها لتصبح ذات معنى. **يطلق اسم المعلومات على البيانات عند تحليلها وتنظيمها وهيكلتها لتصبح ذات معنى.**		✓
3. يتم الحصول على المعلومات من خلال عمليات تحليل البيانات. **تحليل البيانات لا يكفي ليحولها إلى معلومات، بل لابد أن تكون مفيدة وذات معنى.**		✓
4. يطلق اسم المعرفة على عملية جمع البيانات بطريقة صحيحة تجعلها ذات فائدة. **اسم المعرفة يطلق على مجموعة المعلومات التي يتم استخدامها لتقديم فائدة ذات معنى.**		✓
5. تُعدُّ الرسوم البيانية والمخططات من وسائل عرض المعلومات.	✓	
6. تعتبر معلومات حالة الطقس المقدمة من الأرصاد الجوية بمثابة معرفة.	✓	
7. إن علم البيانات، والذكاء الاصطناعي، وذكاء الأعمال هي ثلاث مجالات مختلفة ومستقلة عن بعضها البعض. **علم البيانات، والذكاء الاصطناعي، وذكاء الأعمال مجالات متداخلة ومكملة لبعضها البعض.**		✓
8. يساعد استخدام التمثيل المرئي لتحليل البيانات على استنباط رؤى أفضل مما يعني اكتساب معرفة أفضل بمعنى تلك البيانات.	✓	
9. تُعدُّ نُظم التوصية الذكية وتحليل الانحدار من أفضل طرق تخزين البيانات. **تعد نظم التوصية الذكية وتحليل الانحدار من أشهر أنواع تحليلات البيانات.**		✓
10. تُعدُّ قواعد بيانات السلاسل الزمنية وقواعد البيانات غير العلائقية (NoSQL) جزءًا من وسائل جمع البيانات. **قواعد بيانات السلاسل الزمنية وقواعد البيانات غير العلائقية (NoSQL) تنسيقات تخزين البيانات.**		✓

الحلول اون لاين
hulul.online

يمكن تقديم إجابات إضافية من قبل الطلبة

2 أنشئ قائمة من البيانات ثم حول تلك البيانات إلى معلومات مفيدة، ثم وضّح كيف يحوّل الحاسب البيانات إلى معلومات؟

يمكن أن تكون قائمة البيانات عبارة عن قياسات الطقس، مثل: درجة الحرارة، وقوة الرياح، واتجاه الرياح، وما إلى ذلك. ويمكن تحويلها إلى معلومات مفيدة من خلال تحليل البيانات الاستكشافية.

3 وضّح الفروق الثلاثة الرئيسة بين علم البيانات والذكاء الاصطناعي، وادعم إجاباتك ببعض الأمثلة.

يمكن استخدام كل تقنية من هاتين التقنيتين بصورة منفصلة وكذلك يمكن لكل منهما إكمال بعضهما البعض.

- يختص علم البيانات بمعالجة البيانات التاريخية بينما يحاكي الذكاء الاصطناعي الطريقة التي يفكر بها البشر والتي يقومون بناء عليها باتخاذ القرارات وحل المشكلات ومعالجة اللغة الطبيعية والإدراك.
- يركّز علم البيانات على استخدام أدوات حسابية للقيام بالتحليل الوصفي والتنبؤي والتوجيهي للبيانات بينما يركز الذكاء الاصطناعي على استخدام عناصر المعرفة والذكاء والحوسبة المعرفية.
- لا تتناسب تقنيات علم البيانات التقليدية مع العمل ببيانات غير كاملة أو غير دقيقة لذا يستخدم الجيل القادم من أدوات علم البيانات تقنيات الذكاء الاصطناعي وتعلّم الآلة للقيام بتنبؤات أكثر دقة.

فعلى سبيل المثال: يمكن استخدام علم البيانات لفهم كيفية تدفق حركة المرور الحالية عبر المدينة، واستخدام الذكاء الاصطناعي لبناء نموذج للتنبؤ بكيفية تدفق حركة المرور المستقبلية.

حلول أون لاين
hulul.online

4 وضّح وقارن بين علم البيانات وذكاء الأعمال، وإذا كان لديك شركة استثمارية، أي الحقلين سيكون خيارك المفضل للاستثمار؟

الرجوع للصفحة 12 من الكتاب

يمكن تقديم إجابات إضافية من قبل الطلبة

5 ما مدى فعالية تقارب علم البيانات والذكاء الاصطناعي؟ ابحث في الإنترنت واذكر مثالين ناجحين على ذلك.

"

> الإكمال التلقائي في محركات البحث: استنادًا إلى عمليات البحث التاريخية للمستخدم، تقوم خوارزميات خاصة بعمل تنبؤات حول ما قد يرغب المستخدم في البحث عنه وتقديم قائمة من الاقتراحات لهذا المستخدم المحدد.

> روبوت دردشة (Chatbots): يقوم روبوت دردشة بتحليل البيانات من العديد من المحادثات السابقة لتكون قادرة على محاكاة المحادثة التي قد يجريها المستخدم مع إنسان فعلي. بهذه الطريقة، فإنها توفر ما يحتاجه المستخدم دون تدخل إنسان آخر.

6 وضّح المقصود بعلم البيانات، واذكر ثلاثة تطبيقات حياتية في المجال الصحي، ومجال الأعمال التجارية، والترفيه، ثمّ بيّن لماذا يُعدُّ علم البيانات ضروريًا لمثل هذه المجالات؟

علم البيانات هو مجال الدراسة الذي يتعامل مع كميات هائلة من البيانات باستخدام الأدوات والتقنيات الحديثة لإيجاد أنماط غير بديهية داخل تلك البيانات، وللوصول إلى معلومات مهمّة يمكن أن تساهم في اتخاذ القرارات المتعلقة بكافة الأعمال.

> الصحة: تحليل تطور المرض داخل جسم المريض.

> الأعمال: حملات إعلانية مستهدفة في تطبيقات وسائل التواصل الاجتماعي.

> الترفيه: توصيات المحتوى في خدمات البث.

يُعدُّ علم البيانات ضروريًا لهذه المجالات والعديد من المجالات الأخرى لأنه يساعد على تطورها وتوسعها من خلال تحليل كميات كبيرة من البيانات لاستخراج رؤى وتوصيات قيمة.

الحلول أون لاين
hulul.online

7 وضّح وقارن بين مجموعات البيانات المعالجة وغير المعالجة التي تَصِفُ الدرجات الفصلية للطالب وأدائه خلال العام الدراسي.

مثال على البيانات غير المعالجة: أحمد، 78، 98، 56، 89، علم البيانات، الهندسة، إنترنت الأشياء 1، إنترنت الأشياء 2، 1 ، 1، 2، 3

مثال على البيانات المعالجة:

اسم الطالب	المادة الدراسية	الفصل الدراسي	الدرجات
أحمد	علم البيانات	1	78
أحمد	الهندسة	1	98
أحمد	إنترنت الأشياء 1	2	56
أحمد	إنترنت الأشياء 2	3	89

من بيانات مثل هذه للعديد من الطلاب، يمكننا استخلاص رؤى حول أدائهم بناء على المواد التي درسوها أو فصلهم الدراسي أو المعلمين المحددين.

هذه البيانات وحدها قد لا تكفي للتنبؤ بالأداء الأكاديمي، حيث لم يتم قياس العديد من العوامل الخارجية التي قد تؤثر على الأداء الأكاديمي.

يمكن تقديم إجابات إضافية من قبل الطلبة

8 ابحث عن مزيد من المعلومات حول "مصنع أرامكو السعودية الرقمي" وحدد ثلاثة أمثلة لاستخدام الذكاء الاصطناعي في استخراج البيانات، ومن ثم اعطينا رأيك في مدى تأثير هذه التقنيات على العمليات التشغيلية للشركة؟

9 ابحث في الإنترنت بصورة مفصّلة عن نماذج دورة حياة علم البيانات التي تصف المراحل الرئيسة المذكورة في هذا الدرس، ومن ثم اختر إحداها وحدد المراحل الإضافية واشرحها بإيجاز.

تشجيع الطلاب على البحث عن المصطلحات على الإنترنت "نماذج دورة حياة علوم البيانات". ومن الأمثلة على ذلك ما يلي:

- فهم الأعمال: فهم متطلبات العمل وراء مشكلة تجارية معينة.
- الحصول على البيانات وفهمها: تخزين البيانات، وإنشاء خطوط أنابيب معالجة البيانات، وتنظيف البيانات.
- النمذجة: تحويل البيانات لنموذج التحليل المعين الذي سيتم استخدامه، وتصميم النموذج، والتدريب، وتقييم النموذج.
- النشر: نشر النموذج كخدمة لاستخدامها من قبل التطبيقات الأخرى ومراقبته.

الحلول أون لاين
hulul.online

التعامل مع البيانات

حلول
الحلول اون لاين
hulul.online

تمرينات

1

حدد الجملة الصحيحة والجملة الخاطئة فيما يلي:	صحيحة	خاطئة
1. تشير البيانات الضخمة إلى البيانات الكبيرة جدًا أو المعقدة التي لا يمكن معالجتها باستخدام الطرق التقليدية.	✓	
2. من التقنيات الخمس التي تمكن من إدارة البيانات الضخمة السرعة والمصداقية ومستودع البيانات. السرعة والمصداقية من خصائص البيانات الضخمة وليست من التقنيات المستخدمة في إدارة البيانات الضخمة.		✓
3. يعتبر اكتشاف المعرفة عملية بسيطة لا تتطلب أي خطوات محددة. اكتشاف المعرفة عملية معقدة وتمر بسبع خطوات للوصول للمعرفة المراد اكتشافها.		✓
4. التخزين السحابي هو طريقة التخزين الوحيدة المستخدمة لكم البيانات الكبير مثل البيانات الضخمة. التخزين السحابي ليس الطريقة الوحيدة المستخدمة لتخزين البيانات الضخمة حيث يمكن استخدام التخزين الداخلي أيضا.		✓
5. تُعدُّ قابلية التوسع والتكلفة المنخفضة لتحليل البيانات من المزايا العديدة لتخزين البيانات الضخمة سحابيًا.	✓	
6. مستودع البيانات هو مستودع يتوافر سحابيًا عادةً لتخزين كميات هائلة من البيانات الأولية وغير المعالجة.	✓	
7. الحوسبة في الذاكرة هي طريقة لتسهيل تحليل البيانات الضخمة لاعتمادها بصورة أساسية على ذاكرة الحاسب الرئيسة لتخزين البيانات.	✓	
8. تشير بحيرة البيانات إلى قاعدة البيانات التي تخزن البيانات الحالية والتاريخية الناتجة عن أنظمة المعاملات الأساسية في المؤسسة. بحيرة البيانات هي مستودع بيانات عادة ما يكون سحابيًا يستخدم لتخزين كميات هائلة من البيانات الأولية وغير المعالجة.		✓
9. اختيار البيانات هو تحديد جزء من مجموعة البيانات الذي نريد استخدامه لعملية اكتشاف المعرفة.	✓	
10. تمثيل المعرفة هو عملية استخراج البيانات من التحليلات من خلال الأنماط. تمثيل المعرفة هو تمثيل النتائج التي تم الحصول عليها من خلال التقارير والرسوم البيانية والمخططات الواضحة والمختصرة.		✓

الحلول اون لاين
hulul.online

يمكن تقديم إجابات إضافية من قبل الطلبة

2 اذكر ثلاثة أمثلة على استخدام البيانات الضخمة للمساعدة في الأعمال التجارية.

تلميح:

وجّه الطلبة للاستعانة بكتاب الطالب صفحة 22 و شجعهم على البحث على الإنترنت عن مزيد من الأمثلة لاستخدام البيانات الضخمة في الأعمال التجارية مثل:

- التسويق الموجّه للمنتجات بحيث يتم تسويق أصناف المنتجات المختلفة لأكثر الفئات المهتمة بكل صنف.
- اقتراح منتجات إضافية قد يرغب العميل بشراءها مثل "قسم منتجات أخرى قد ترغب بها" أو "المستخدمون الآخرون قاموا أيضاً بشراء.."
- تحسين إدارة سلاسل الإمداد.

3 ابحث في الإنترنت عن مزودي خدمات الحوسبة السحابية الأكثر شهرة اليوم في السوق العالمية والمُستخدمة في تخزين ومعالجة البيانات الضخمة.

تلميح:

ساعد الطلبة على استخدام الكلمات المفتاحية الصحيحة للبحث عن مزودي خدمات الحوسبة السحابية، وناقشهم في نتائج بحثهم.

يمكن استخدام كلمات مفتاحية للبحث، مثل:"حوسبة سحابية"، "تخزين سحابي".

الحلول اون لاين
hulul.online

4 اشرح بإيجاز مدى مساعدة التخزين السحابي لنا في التعامل مع مشكلة تخزين الكم الهائل من البيانات الناتج عن البيانات الضخمة.

يتطلب التعامل مع البيانات الضخمة الكثير من مساحة التخزين وقوة المعالجة. من حيث القدرة على تخزين الأشياء، تقوم السحابة بذلك. يمكن للمؤسسات شراء خدمات التخزين التي تجعل التوسع صعودا وهبوطا أسهل. علاوة على ذلك، يمكن لهذه الخدمات السحابية أيضا تلبية احتياجات البيانات الضخمة في الحوسبة.

يمكن تقديم إجابات إضافية من قبل الطلبة

5 البيانات الضخمة هي تطور حديث في تاريخ الحوسبة، فهل يمكنك تحديد عاملين ساهما في هذا النمو المفاجئ في جمع البيانات؟

> يتيح الاعتماد الواسع على الإنترنت للمُستخدِمين والتطبيقات نقل البيانات بسرعة.

> تتيح الخدمات السحابية للعديد من المستخدمين والمنظمات خدمات حوسبة عالية الكفاءة.

6 قارن بين التقنيات الثلاث لتخزين البيانات الضخمة، وإذا قمت بتطوير تطبيق يتطلب وصولًا سريعًا جدًا إلى البيانات، فما هي التقنية التي ستختارها؟

> مستودعات البيانات: هي قاعدة البيانات التي تخزّن البيانات الحالية والتاريخية التي نتجت عن العديد من المعاملات التشغيلية

> الحوسبة في الذاكرة: تعتمد على ذاكرة الحاسب الرئيسة وبالتالي تسمح بتجاوز معوقات استرداد البيانات.

> بحيرة البيانات: هي مستودع بيانات سحابي لتخزين البيانات الأولية وغير المعالجة.

في حال تطوير تطبيق يتطلب وصولًا سريعًا جدًا إلى البيانات سأقوم باختيار تقنية الحوسبة في الذاكرة لأنها تقلل وقت الاستعلام.

الحلول اون لاين
hulul.online

7 لماذا يُعدُّ تقييم أنماط البيانات مهمًا في عملية التنقيب عن البيانات؟

تلميح:

وجّه الطلبة لحل التمرين، والاستعانة بالجدول رقم 1.6 صفحة 23 بكتاب الطالب عند الحاجة.

يمكن تقديم إجابات إضافية من قبل الطلبة

8 اشرح كيفية عمل قابلية التوسع في تخزين البيانات السحابية، ثم ابحث عن خدمتين لتخزين البيانات سحابيًا على شبكة الإنترنت.

ميزة قابلية التوسع تعني أن مزوّد الخدمة يوفر مساحة تخزين إضافية للمستفيد من الخدمة حسب الطلب أو حسب الحاجة فلا يضطر المستفيد لدفع مقابل للخدمة إلا عندما يحتاجها فعلًا.

ومن الخدمات الشائعة للتخزين السحابي:

- Microsoft Azure Cosmos DB.
- Amazon Web Services S3.

9 ما الغرض من حوكمة البيانات؟ وهل تعتبر حوكمة البيانات مرادفة لإدارة البيانات؟

حوكمة البيانات ليست مرادفة لإدارة البيانات بل هي أحد مكوناتها، فحكومة البيانات تحدد جميع الضوابط والسياسات والعمليات والتي تنفذ بواسطة إدارة البيانات والتي مهمتها جمع البيانات واستخدامها.

حلول
الحلول اون لاين
hulul.online

10 ابحث في الإنترنت عن معلومات حول ضوابط أو قوانين إدارة البيانات الصحية في المملكة العربية السعودية، ثم اذكر العواقب المترتبة على تسرب بيانات أحد مرافق الرعاية الصحية؟

نصت وثيقة سياسات حوكمة البيانات الوطنية الصادرة من مكتب إدارة البيانات الوطنية على أن البيانات الصحية تعتبر من البيانات الحساسة وبالتالي فإن تسرب بيانات أحد مرافق الرعاية الصحية يعتبر خرق لخصوصية بيانات المرضى وقد يؤدي إلى التنبؤ بشكل مباشر أو غير مباشر بالظروف الصحية لهم.

يمكن تقديم إجابات إضافية من قبل الطلبة

11 أنشئ تقريرًا عن تغير المناخ من خلال مقارنة بيانات الطقس التاريخية لدولتين من اختيارك، ثم حدّد أين ستبحث عن المعلومات على الإنترنت، واشرح العوامل المؤثرة وراء قرارك هذا.

تلميح:

ذكّر الطلبة أنه ينبغي عليهم جمع البيانات من مصادر موثوقة، مثل: البيانات الحكومية المفتوحة وما إلى ذلك. ولإنشاء تقرير أكثر موثوقية، يجب عليهم اختيار دولتين لا تتشابهان في المناخ؛ للعثور على اختلافات بينهما.

12 ما مخاوف الخصوصية التي يمكنك التفكير بها عند تعامل مؤسسة كبيرة مع البيانات الضخمة؟

من التحديات الحديثة المرتبطة بالبيانات الضخمة هي مخاوف الخصوصية الناتجة عن التخزين السحابي للبيانات الضخمة وما يرتبط بها من مخاوف تسرب البيانات أو تعرضها للهجمات الإلكترونية.

13 هل لديك فكرة عن المعلومات التي تمتلكها شبكة التواصل الاجتماعي التي انضممت إليها حول عائلتك وأصدقائك؟ إذا كانت الإجابة نعم، اكتب قائمة قصيرة بهذه المعلومات.

تلميح:

لإنشاء هذه القائمة، ساعد الطلبة على التفكير بالمعلومات التي تقوم شبكات التواصل الاجتماعي بجمعها من المستخدمين ثم اضرب لهم أحد الأمثلة.

مثال: أحد تطبيقات الوسائط الاجتماعية الشائعة جدا هو فيس بوك. فيما يلي قائمة قصيرة بالمعلومات التي يمكن أن يجمعها فيسبوك من المُستخدِم:

- من هم عائلتك وأصدقائك.
- حالة علاقتهما.
- هواياتهم واهتماماتهم.

حلول الحلول أون لين
hulul.online

أساسيات علم البيانات

الحلول اون لاين
hulul.online

تمرينات

1

حدد الجملة الصحيحة والجملة الخاطئة فيما يلي:	صحيحة	خاطئة
1. يتم تحويل جميع الأرقام والمعلومات إلى مصفوفات في نماذج وخوارزميات تعلّم الآلة.	✓	
2. لكي يتمكن عالم البيانات من فهم وتفسير البيانات عند إنشاءها، فإنه يحتاج إلى معرفة عملية بالإحصاء والاحتمالات.	✓	
3. تتخصص الرياضيات المتقطعة في طرق المنطق والاستنتاج، وهي جوانب أساسية في تصميم الخوارزمية والتي تُعدُّ الأساس لتعلّم الآلة.	✓	
4. بعض المجتمعات على الإنترنت مدعومة من قبل الحكومات، والبعض الآخر يديره متطوعون.	✓	
5. مهندس الأنظمة هو الشخص الذي يصمم أنظمة المعلومات للمؤسسات والشركات. **معماري التطبيقات هو من يصمم أنظمة المعلومات للمؤسسات والشركات.**		✓
6. عالم البيانات هو محترف يأخذ الرؤى من مجموعات البيانات التي تمت معالجتها ويقوم بإنشاء تقارير وتصورات وتحليلات أخرى متنوعة تتماشى مع الأهداف الرئيسة التي تحتاج إلى حل مبني على علم البيانات لتحقيقها.	✓	
7. محلل البيانات هو محترف مسؤول عن تخزين وتدفق المعلومات في شركة أو مؤسسة، ويعمل مع علماء ومهندسي البيانات لبناء طرق نقل البيانات المناسبة لإدخال مجموعة البيانات وتحليلها وإخراج النتائج. **من يقوم بهذه المهمة هو معماري البيانات.**		✓

الحلول اون لاين
hulul online

يمكن تقديم إجابات إضافية من قبل الطلبة

2 وضّح كيف تُمكّن لغة البايثون من مساعدة المتخصص في علم البيانات.

تساعد لغة البايثون المتخصص في علم البيانات من خلال العديد من المكتبات القياسية التي توفر دوال قوية سهلة الاستخدام تغطي مجموعة متنوعة من الاحتياجات مثل استخراج وجمع ومعالجة البيانات والنمذجة التنبؤية وتمثيل البيانات.

3 وضّح كيف يمكن لجوبيتر نوت بوك مساعدة المتخصص في علم البيانات.

يعد جوبيتر نوت بوك أحد تطبيقات الويب مفتوحة المصدر المستخدمة لتطوير مشاريع علم البيانات باستخدام البايثون حيث يتيح للمتخصصين في علم البيانات بيئة تفاعلية تدمج أوامر البايثون وتمثيلات البيانات في مستند واحد.

حلول
الحلول اون لاين
hulul.online

4 عدّد أهم أدوات علم البيانات، وكيف يمكن لكل أداة المساهمة في كل خطوة من خطوات علم البيانات؟

تلميح:

وجّه الطلبة لحل التمرين، والاستعانة بكتاب الطالب صفحة 36 عند الحاجة.

يمكن تقديم إجابات إضافية من قبل الطلبة

5 لماذا يعتبر فهم الإحصاء مهارة أساسية لعالم البيانات؟ وهل يمكنك التفكير في مثال يتضمن تحليل البيانات؟

تلميح:

يمكن للطلبة الاطلاع على فروع علم الرياضيات التي يحتاجها عالم البيانات في الصفحة 34، ومنها الإحصاء. ساعدهم بالعثور على مثال تكون فيه الحاجة للإحصاءات، مثل: استكشاف مدى تشابه ميزات مجموعة البيانات، أو كيف يمكن أن يرتفع سعر السهم أو ينخفض... إلخ.

6 البايثون هي لغة برمجة متعددة الاستخدامات، فهل تُعدُّ كافية لمشروعات علوم البيانات؟

تتطلب العديد من حلول علوم البيانات الكثير من المهام المعقدة. قد تكون بايثون لغة قوية بسبب استخدامها للمكتبات، ولكن يتم إنجاز بعض جوانب مشروع علوم البيانات بتقنيات أخرى لتخزين البيانات وتطوير خوارزميات ونماذج التعلم الآلي.

حلول
الحلول اون لاين
hulul.online

7 ابحث عن ثلاث مكتبات البايثون تحظى بشعبية كبيرة بين علماء البيانات على الإنترنت، ثم اشرح باختصار سبب ذلك.

- باندا (Pandas): مكتبة قياسية تعمل مع البيانات المجدولة وتستخدم لتحليل البيانات وتمثيلها.
- نمباي (NumPy): مكتبة قياسية للعمل مع البيانات العددية في البايثون، ويمكن استخدامها لإجراء مجموعة من العمليات الرياضية على المصفوفات.
- مات بلوت ليب (Matplotlib): مكتبة قياسية تقدم أدوات متنوعة لتمثيل البيانات وتصويرها.

يمكن تقديم إجابات إضافية من قبل الطلبة

8 قارن بين بيئة التطوير المتكاملة و بيئة جوبيتر نوت بوك، ثم اذكر الاختلافات الرئيسة بينهما؟

يتم استخدام بيئة التطوير المتكاملة على الجهاز الشخصي بينما يمكن استخدام بيئة جوبيتر نوت بوك عن طريق المتصفح ويدعم إنشاء ملفات بايثون تفاعلية يتم فيها الدمج بين أوامر البايثون و تمثيلات البيانات.

9 على فرض أنك عالم بيانات جديد تتقن لغة البايثون، فما الأدوات الأخرى التي تحتاجها من أدوات علم البيانات؟

تلميح:

وجّه الطلبة إلى الاستعانة بكتاب الطالب صفحة 36 للتعرف على الأدوات المستخدمة من قبل عالم البيانات.

10 يوجد في هذا الدرس قائمة بالمهن المتعلقة بعلم البيانات، فأي منها تفضّل؟ ولماذا؟ وما التحديات التي تعتقد بأنك ستواجهها في هذه المهنة؟

تلميح:

اطلب من الطلبة التفكير في وظائف المستقبل المتعلقة بعلم البيانات مستعينًا بكتاب الطالب صفحة 37، وتحديد أسباب اختيارهم، وساعدهم على استكشاف التحديات والمشاكل التي تعالجها الوظيفة المختارة.

11 قم بزيارة أحد مجتمعات علم البيانات عبر الإنترنت وابحث عن دورة تدريبية بسيطة للدراسة الذاتية لتعزيز معرفتك بعلم البيانات، ثم قم بتقييم مدى ملاءمة الدورة التدريبية لمستوى معرفتك.

تلميح:

وجّه الطلبة لحل التمرين، والاستعانة بكتاب الطالب صفحة 38 عند الحاجة، حيث يوجد بعض الأمثلة لمجتمعات علم البيانات.

يمكن استخدام كلمات مفتاحية للبحث، مثل: "مجتمعات علم البيانات"، "دورات علم البيانات".

حلول الطالب اون لاين
hulul.online

الوحدة الثانية

جمع البيانات والتحقق من صحتها

حلول
الحلول اون لاين
hulul.online

تمرينات

1

حدد الجملة الصحيحة والجملة الخاطئة فيما يلي:	صحيحة	خاطئة
1. جمع البيانات هو عملية جمع وقياس القراءات أو الحقائق.	✓	
2. هناك نوعان من التصنيفات الأساسية لمصادر جمع البيانات: الرئيسة والثانوية.	✓	
3. يعتبر تاريخ نشر المعلومات من أهم معايير جودة المعلومات.	✓	
4. الملاءمة تعني أنه كلما كانت المعلومات ليست ذات صلة بما يتم البحث عنه، كانت جودتها أسوأ.	✓	
5. تعتبر مستويات التفاصيل والدقة معايير جودة للمعلومات.	✓	
6. تساعدك معايير الجودة الخمس في التحقق من موثوقية المعلومات.	✓	
7. ليس للحكومة سلطة على منصات البيانات المفتوحة. تستضيف الجهات الحكومية منصات البيانات المفتوحة وتسن التشريعات الخاصة باستخدامها.		✓
8. تختلف التصاريح القانونية لجمع واستخدام البيانات بناءًا على عدة متغيرات.	✓	
9. يتم استخدام البحث الموجه عندما نريد التركيز على قضايا محددة ظهرت من بحثنا الأساسي.	✓	
10. يمكن إجراء مقارنة البيانات عندما يكون لديك أكثر من مجموعة بيانات واحدة مع بيانات مسجلة من نفس المنطقة ومن فترات زمنية مماثلة.	✓	

الحلول اون لاين
hulul.online

يمكن تقديم إجابات إضافية من قبل الطلبة

2 قارن بصورة مختصرة بين مصادر البيانات الرئيسة والثانوية.

تلميح:

شجّع الطلبة على حل التمرين باستخدام بعض الأمثلة من المصادر الأولية والثانوية، وليس فقط من خلال تقديم تعريف المصطلحين.

3 صِفْ باختصار كل سمة جودة يمكن استخدامها للتحقق من جودة المعلومات.

تلميح:

وجّه الطلبة لحل التمرين، والاستعانة بكتاب الطالب صفحة 50 عند الحاجة.

الحلول اون لاين
hulul.online

4 أعطِ مثالًا على البحث الموجه ومقارنة البيانات.

من أمثلة البحث الموجه استخدام قيم درجة الحرارة والرياح للتنبؤ بالطقس في مدينة ما، ثم ملاحظة أن مناطق معينة من هذه المدينة سجلت قيمًا قصوى لدرجات الحرارة، وفي هذه الحالة يجب إجراء بحث حول تلك المناطق لتقييم تأثير العوامل الأخرى على درجة الحرارة فيها.

وفي حالة وجود مجموعة بيانات لقيم درجات الحرارة المسجلة لمدينة جدة في مارس 2021 ومجموعة أخرى مسجلة في مارس 2022، فيمكن بسهولة إجراء مقارنة البيانات من أجل اكتشاف التغيرات في درجات الحرارة أو التغيرات خلال سنوات.

يمكن تقديم إجابات إضافية من قبل الطلبة

5 قارن بين مصادر البيانات الرئيسة والثانوية في حالة الطقس مع طرح بعض الأمثلة.

يحتوي مصدر البيانات الرئيسة على بيانات لم تجمع من قبل ويمكن جمعها من خلال أجهزة الإستشعار وأجهزة تسجيل البيانات وكذلك من خلال الاستبانات، ومن الأمثلة التي تدل على مصادر البيانات الرئيسة مقياس سرعة الرياح الذي يقيس سرعة الرياح، بينما يتم إنشاء البيانات الثانوية من خلال استخدام بيانات أخرى من مصدر البيانات الرئيس، فعلى سبيل المثال، يمكن استخدام بيانات درجة حرارة الهواء وسرعة الرياح من جهازي استشعار مختلفين للحصول على بيانات لمُعامِل آخر يسمى درجة حرارة الرياح الباردة (WIND-CHILL TEMPERATURE).

6 زُر منصة البيانات المفتوحة data.gov.sa وابحث عن المعلومات المتعلقة بالحصول على تصاريح باستخدام تلك البيانات. هل هناك أي استثناء؟

يجب على مستخدمي البيانات التي يتم استضافتها على منصة البيانات المفتوحة في المملكة العربية السعودية استخدام هذه البيانات وفق شروط سياسة البيانات المفتوحة. وإذا نقلت مجموعة البيانات هذه علنًا، أو أي مجموعة بيانات مشتقة أو مجموعة بيانات كجزء من مجموعة بيانات جماعية، فيجب:

- القيام بذلك فقط بموجب شروط ترخيص البيانات المفتوحة.
- عدم استخدام البيانات لأغراض سياسية، أو لدعم نشاط غير قانوني أو إجرامي، أو لاستخدامها في تعبيرات عنصرية أو تمييزية، أو لإحداث تأثير سلبي في الثقافة أو المساواة، أو في إثارة أي سلوك غير قانوني، أو مخالف لعادات وتقاليد المملكة.

الحلول اون لاين
hulul.online

7 ابحث على الإنترنت عن منصات البيانات المفتوحة في البلدان الأخرى، وهل يمكنك العثور على معلومات شخصية عبر هذه المنصات؟

تستضيف الجهات الحكومية منصات البيانات المفتوحة التي تتيح وصول الجمهور إلى البيانات، ولكن لا يمكن العثور على معلومات شخصية عبر هذه المنصات لأنها تتطلب أذونات من كل فرد تم إدراجه وجمع بياناته.

8 حدد موقعين على الإنترنت أحدهما حكومي والآخر خاص، ثم قارن بين جودة المعلومات بينهما بناءً على المعايير الخمس التي تعرفت عليها.

تلميح:

ذكّر الطلبة بمعايير الجودة الخمس للمعلومات عند الحاجة.

حلول
الحلول أون لاين
hulul.online

أنواع البيانات

الحلول اون لاين
hulul.online

تمرينات

1

حدد الجملة الصحيحة والجملة الخاطئة فيما يلي:	صحيحة	خاطئة
1. البيانات الفئوية هي نوع كمي من البيانات. البيانات الفئوية هي بيانات غير كمية، وإنما توصف بأنها اسمية أو ترتيبية.		✓
2. يتم تعريف البيانات الترتيبية على أنها البيانات المستخدمة لتسمية المتغيرات أو توسيمها. البيانات الاسمية هي المستخدمة لتسمية المتغيرات أو توسيمها.		✓
3. تمثل البيانات المنفصلة عناصر قابلة للعد ويمكن أن تأخذ قيمًا معينة فقط.	✓	
4. رموز المطارات ورموز العملات هي أمثلة على ترميز البيانات.	✓	
5. البيانات الثابتة هي البيانات التي قد تتغير بعد تسجيلها ويجب تحديثها باستمرار. البيانات الثابتة هي التي لاتتغير بعد تسجيلها.		✓
6. البيانات المتغيرة (الديناميكية) هي البيانات التي لا تتغير بعد تسجيلها. البيانات المتغيرة هي التي قد تتغير بعد تسجيلها ويجب تحديثها باستمرار.		✓
7. غالبًا ما يتم إجراء تشفير البيانات على البيانات في شكلها الأولي، والتي يتم الحصول عليها من التجارب أو من خلال الاستطلاعات.	✓	
8. يتكون رقم ISBN من 10 مجموعات متتالية من الأرقام. يتكون رقم ISBN من خمس مجموعات متتالية من الأرقام.		✓
9. الرمز الشريطي عبارة عن ملصق به خطوط سوداء رفيعة، مع مجموعة متنوعة من الأرقام.	✓	
10. يتكون رمز الاستجابة السريعة من مربعات سوداء تحتوي على معلومات.	✓	

الحلول أونلاين
hulul.online

يمكن تقديم إجابات إضافية من قبل الطلبة

2 صِفْ باختصار البيانات الثابتة والبيانات المتغيرة.

تلميح:

وجّه الطلبة لحل التمرين، والاستعانة بصفحة 58 من كتاب الطالب عند الحاجة.

3 اذكر بعض الأمثلة على منتجات مميزة برموز شريطية أو رموز الاستجابة السريعة.

تلميح:

ساعد الطلبة على ذكر أمثلة من حياتهم اليومية.

الحلول اون لاين
hulul.online

4 وضّح باختصار المقصود بترميز البيانات.

ترميز البيانات هو عملية تتيح تنظيم البيانات وترتيبها بطريقة محددة وذلك باستخدام رموز مختلفة مثل الأرقام أو الحروف أو الكلمات القصيرة، والتي تصف سياقًا معينًا أو تعبّر عن عبارة أو فقرة بأكملها.

5 ابحث عن موقع الكتروني يمكن من خلاله إنشاء رموز QR مجانية، وأنشئ رمز QR للصفحة الرئيسة لذلك الموقع، ولصفحة ويب أخرى لموقع الكتروني من اختيارك. هل يمكنك ملاحظة الاختلافات في المربعات السوداء لكل رمز QR؟

هناك اختلافات ملحوظة في المربعات السوداء بين رمزي QR للصفحة الرئيسية وصفحة الويب الأخرى.

الصفحة الرئيسية – رؤية السعودية 2030

تعرف على المملكة – رؤية السعودية 2030

6 تحتفظ المنظمة الدولية للتوحيد القياسي بالقائمة الرسمية لرموز الدول من خلال معيار (ISO 3166). ابحث عن رموز دول مجلس التعاون الخليجي (GCC) المكونة من حرفين. هل يمكنك إعطاء أمثلة على استخدام هذه الرموز؟

المثالان هما: SA هو رمز دولة المملكة العربية السعودية ، و AE هو رمز دولة الإمارات العربية المتحدة، ويمكن استخدام الرمزين في جدول بيانات أو في رسم بياني لتمثيل أسماء البلدان باختصار.

الحلول اون لاين
hulul.online

7 ابحث عن رمز ISBN لهذا الكتاب، وهل يمكنك تحديد أرقام الدولة والناشر؟

رمز ISBN لهذا الكتاب هو: 978-603-511-237-6

رقم الدولة (رقم التسجيل) هو: 603

رقم الناشر (رقم المُسجل) هو: 511

64

التحقق من صحة إدخال البيانات

الحلول اون لاين
hulul.online

تمرينات

1

حدد الجملة الصحيحة والجملة الخاطئة فيما يلي:	صحيحة	خاطئة
1. يشير التحقق من صحة البيانات إلى الإجراء الذي يحذف تلقائيًا أي بيانات أولية لا تفي بمعايير معينة. **يشير التحقق من صحة البيانات إلى أي نشاط يتحقق من أن البيانات المدخلة تنبثق من مجموعة من القيم المعتمدة، وتتوافق مع القواعد المقبولة للبيانات.**		✓
2. لا يوجد سوى خمسة أنواع للتحقق من صحة البيانات. **للتحقق من صحة البيانات أنواع عديدة يشتهر منها ستة أنواع تم ذكرها في الكتاب.**		✓
3. يساعد التحقق من التواجد على تقليل الأخطاء باستخدام قائمة محدودة من القيم المحددة مسبقًا. **التحقق من التواجد يجعل عملية الإدخال في الخلية إلزامية مما يضمن عدم تركها فارغة.**		✓
4. يهدف التحقق من البحث إلى التأكد من أن الرموز والحروف تدخل بنطاق طول محدد. **يهدف التحقق من البحث إلى تقليل الأخطاء باستخدام قائمة محدودة النطاق تحتوي على قيم مدخلة مسبقًا.**		✓
5. يُستخدم فحص النطاق للتأكد من أن الأرقام التي تُدخل تقع ضمن نطاق معين.	✓	
6. يُستخدم التحقق من الصيغة للتأكد من أن البيانات تأتي بصيغة محددة مسبقًا.	✓	
7. يساعد التحقق من النوع في تقليل أخطاء اللغة. **التحقق من النوع يضمن إدخال المستخدمين لنوع القيمة الصحيح في حقل محدد.**		✓
8. يتم استخدام رقم التحقق إذا أردت التأكد من إدخال مجموعة من الأرقام بشكل صحيح.	✓	
9. يعتبر مايكروسوفت إكسل الأداة الوحيدة التي يمكن استخدامها للتحقق من صحة البيانات. **توجد العديد من الأدوات المستخدمة للتحقق من صحة البيانات، مثل جداول بيانات قوقل (GOOGLE SHEETS)، ونحوها.**		✓
10. يمكن إجراء التحقق من صحة البيانات بعد إدخال القيم في برنامج التحقق من صحة البيانات.	✓	

حلول hulul.online

يمكن تقديم إجابات إضافية من قبل الطلبة

2 اشرح بايجاز المقصود بعملية التحقق من صحة البيانات.

تلميح:

وجّه الطلبة لحل التمرين، والاستعانة بكتاب الطالب صفحة 65 عند الحاجة.

3 صِف باختصار الخطوات التي يجب اتباعها في إكسل لإكمال عملية التحقق من صحة البيانات.

تلميح:

نبّه الطلبة على أن أول ما عليهم فعله هو إنشاء جدول بيانات في إكسل. ويمكنهم الرجوع إلى صفحة 68 من كتاب الطالب؛ حيث يوجد ملخص لإجراءات التحقق من صحة البيانات للمثال.

الحلول أون لاين
hulul.online

يمكن تقديم إجابات إضافية من قبل الطلبة

4 أنشئ دليل عناوين لمعلومات أصدقائك يتضمن الحقول التالية: الاسم، والهاتف، وعنوان المنزل، وعنوان البريد الإلكتروني، وتاريخ الميلاد، والهواية. اكتب نوع التحقق من صحة البيانات الذي يجب إجراؤه على كل حقل من الحقول السابقة.

تلميح:

باستخدام برنامج الإكسل، وجّه الطلبة لإنشاء جدول دفتر العناوين، والذي سيتكون من ستة أعمدة، وساعدهم على استخدام كل نوع من أنواع التحقق من صحة البيانات، إذا وجدوا أي صعوبات.

5 قارن ما يلي: (أ) فحص الطول مقابل فحص النطاق، (ب) فحص الصيغة مقابل فحص النوع. أعطِ أمثلة على استخدام كل نوع من أنواع التحقق السابقة.

(أ) فحص الطول يهدف إلى التأكد من أن الرموز والحروف تُدخل بنطاق طول محدد، بينما فحص النطاق يُستخدم للتأكد من أن الأرقام المدخلة تقع ضمن نطاق معين ويشمل حدين هما: **الحد الأقصى** (MAXIMUM LIMIT) **والحد الأدنى** (MINIMUM LIMIT).

مثال على فحص الطول: يُستخدم عند وجود أرقام ISBN أو رموز العملات أو رموز البلدان ISO التي لها طول ثابت مكون من 13 و 3 و 2 رقمًا أو حرفًا، على التوالي.

مثال على فحص النطاق: يُستخدم إذا تم إدخال عمر الشخص، فيجب على النظام قبول الأرقام الموجبة بحد أعلى 140 فقط، وأي شيء آخر يتجاوز هذا النطاق سَيُعدُّ بيانات غير صالحة.

(ب) فحص الصيغة يُستخدم للتأكد من أن البيانات تأتي بصيغة محددة مسبقًا ولن يسمح بأي صيغة أخرى يتم إدخالها في الخلية، بينما فحص النوع يضمن إدخال المستخدمين لنوع القيمة الصحيح في حقل محدد.

مثال على فحص الصيغة: يكون مفيد عند استخدام حقول البيانات للرموز البريدية أو أرقام الهواتف، وفي كلتا الحالتين، يتوقع النظام تنسيق بيانات محدد للغاية.

مثال على فحص النوع: إذا تم تمييز حقل البيانات على أنه حقل رقمي، فلن تتمكن من تخزين قيمة نصية.

hulul.online