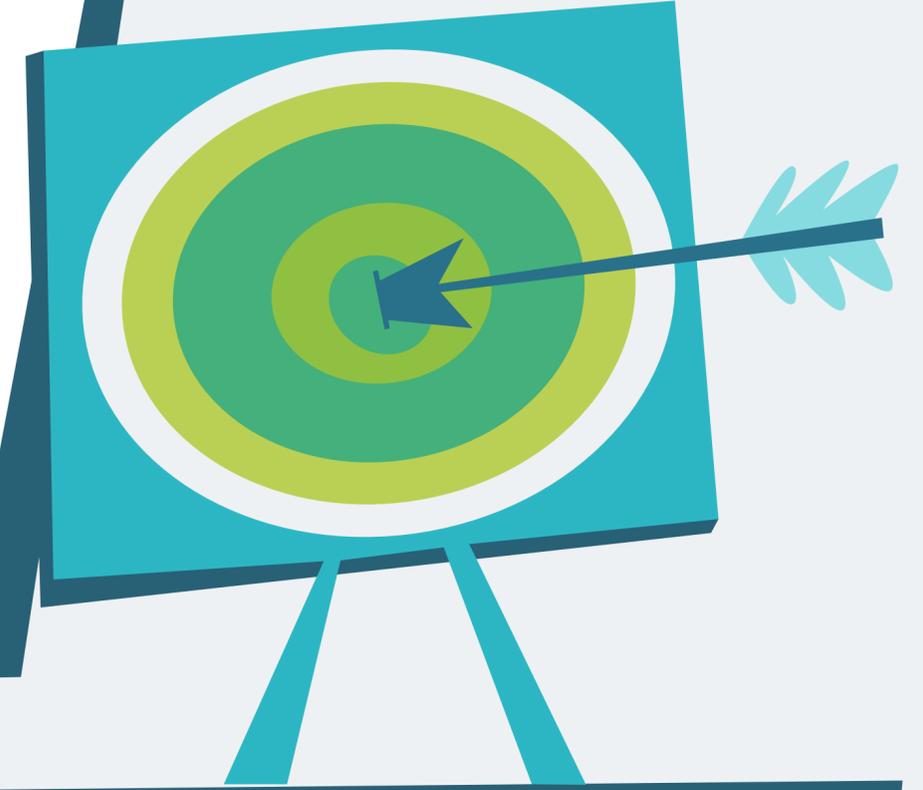


دورة حياة علم البيانات هي عملية منظمة
توجه الممارسين خلال مراحل مختلفة
للحصول على رؤى فعالة للبيانات. تشمل
المراحل الرئيسية ما يلي:



فهم الأعمال

تحديد أهداف المشروع مع أصحاب المصلحة لمواءمة الحلول مع الأهداف



التنقيب عن البيانات (تعددين البيانات)

جمع البيانات ذات الصلة، مع الأخذ بعين الاعتبار الجودة والملاءمة وسهولة الوصول إليها



تنظيف البيانات

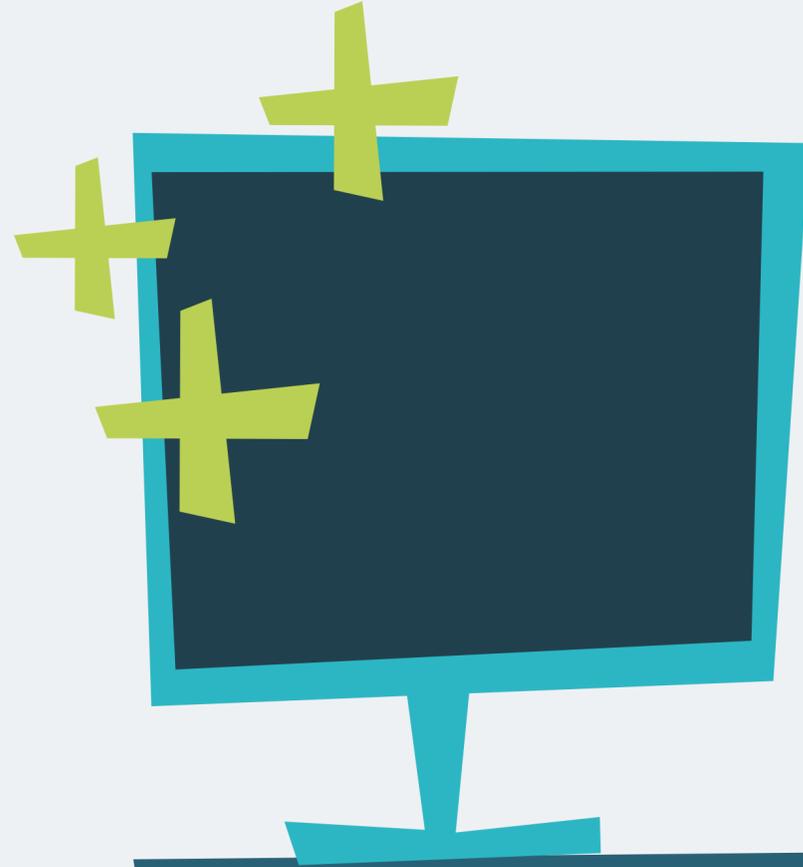
التعامل مع التناقضات والأخطاء والقيم المفقودة للتحليل الدقيق



استكشاف البيانات

تحليل البيانات المُنظَّفة للاتجاهات والرؤى باستخدام الإحصائيات والمرئيات



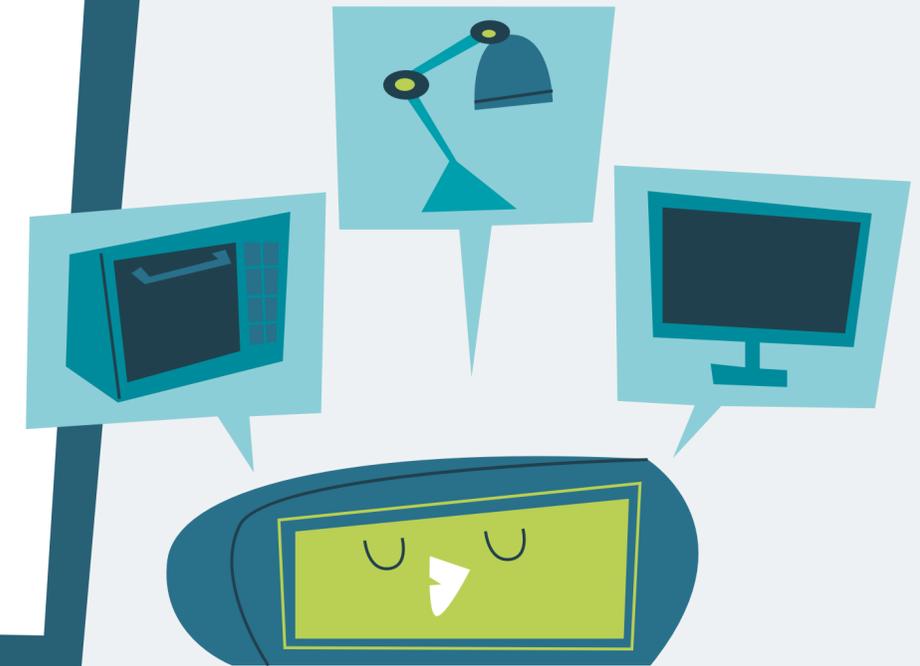


هندسة الخصائص

تحسين أداء النموذج عن طريق تحديد الميزات أو تحويلها

النمذجة التنبؤية

تطوير نماذج التعلم الآلي وضبطها للحصول على تنبؤات دقيقة



التصور البصري للبيانات

استخدم العناصر المرئية لتوصيل النتائج المعقدة إلى أصحاب المصلحة، مما يساعد في اتخاذ القرار



ماذا عن علماء البيانات؟ هل تساءلت يوماً عن أدوارهم؟

يُمكن وصف دور عالم البيانات بأنه "متعدد المهارات"، وذلك بسبب مجموعة المهارات المتنوعة المطلوبة للتميُّز في هذا المجال.

يرتدي عالم البيانات قبعات متعددة، حيث يجمع بين الخبرة من مختلف المجالات لاستخراج رؤى ذات معنى من البيانات.





فيما يلي استكشاف للمهارات المختلفة التي تجعل من عالم
البيانات عالمًا موسوعيًا حقيقيًا:

عالِم رياضيات

الدور: يستفيد علماء البيانات من المفاهيم والتقنيات الرياضية لبناء نماذج وخوارزميات قوية. من الجبر الخطي لعمليات المصفوفات إلى حساب التفاضل والتكامل للتحسين، يعد الأساس الرياضي المتين أمراً بالغ الأهمية لفهم نماذج التعلم الآلي المعقدة وتطويرها.



عالم إحصاء



الدور: تعتبر المعرفة الإحصائية أمراً أساسياً لعلماء البيانات فيما يتعلق باختبار الفرضيات ونظرية الاحتمالات والإحصاءات الاستدلالية. يستخدم علماء البيانات الأساليب الإحصائية لاستخلاص استنتاجات موثوقة من البيانات، وإجراء التنبؤات، وتقييم أهمية النتائج التي توصلوا إليها.

علم

الدور: يتبع علماء البيانات منهجًا علميًا في عملهم. إنهم يشكلون الفرضيات، ويصممون التجارب (أو التحليلات)، ويقيمون النتائج بشكل منهجي. تضمن هذه العقلية العلمية موثوقية النتائج وإمكانية تكرارها، مما يساهم في الدقة الشاملة لعملية علم البيانات.



مبرمج

الدور: تعد الكفاءة في البرمجة سمة أساسية لعالم البيانات. فهم يستخدمون لغات البرمجة مثل *R* أو *Python* لمعالجة البيانات وبناء النماذج وإنشاء حلول قابلة للتطوير. تعد مهارات البرمجة ضرورية لتنفيذ الخوارزميات وأتمتة العمليات ودمج الحلول التحليلية في بيئات الإنتاج.



01001
11010

في مجالات العلوم والتكنولوجيا والهندسة
تعتبر خبرة عالم البيانات، (STEM) والرياضيات
الماهر لا تقدر بثمن لتسخير الرؤى المستندة إلى
البيانات لتعزيز البحث وتحسين العمليات واتخاذ
قرارات مستنيرة، والمساهمة في نهاية المطاف
في التقدم العلمي والابتكار.

ما هي فوائد استخدام الأدوات؟

01010100
01011011
01110110

1. سير عمل أفضل:

تعمل *Jupyter Notebooks* على تعزيز سير العمل المتكامل، والجمع بين التعليمات البرمجية والمرئيات والنص لعملية تحليل متماسكة ومتكررة.



2. توفير الوقت:

تعمل أدوات مثل *Apache Spark* على تمكين المعالجة المتوازية لمجموعات البيانات الكبيرة، مما يقلل من أوقات الحساب. تتميز الأتمتة والتحسين بميزة إعفاء علماء البيانات من المهام اليدوية، مما يسمح لهم بالتركيز على العمل عالي المستوى.



3. التناسق:

يعد الاتساق في التحليل أمراً حيوياً للموثوقية، ويعمل على تمكين علماء البيانات من تصور البيانات، *Power BI* والاتصال بمصادر البيانات المتنوعة للتقارير ولوحات المعلومات وواجهته *Microsoft* التفاعلية. يساعد تكامله مع منتجات سهلة الاستخدام في سرد البيانات واتخاذ القرارات بشكل تعاوني.

فيما يلي بعض أدوات علم البيانات الرئيسية الأخرى:

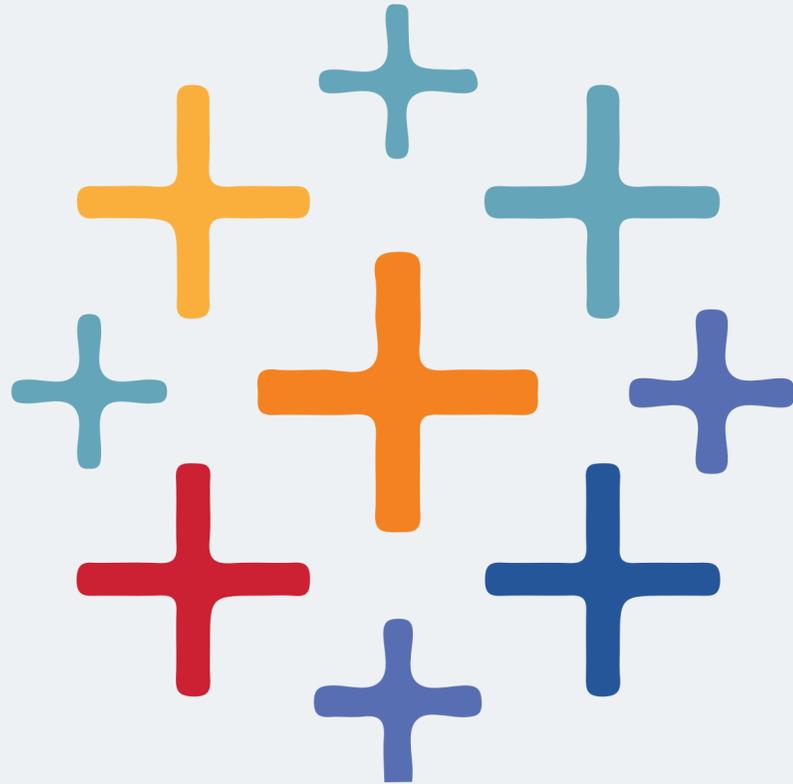
EXCEL

يعتبر برنامج *Excel* ، وهو برنامج جداول بيانات يستخدم على نطاق واسع، ضرورياً لمعالجة البيانات الأساسية وتحليلها. تعمل واجهته البديهية على تسهيل استكشاف مجموعة البيانات، وإجراء العمليات الحسابية، والتصورات البسيطة، مما يجعلها أحد الأصول القيمة في مجموعة أدوات عالم البيانات.



TABLEAU

Tableau هي أداة بديهية لتصور البيانات تمكن علماء البيانات من إنشاء لوحات معلومات تفاعلية ومفيدة مع واجهة السحب والإفلات، مما يجعل البيانات المعقدة سهلة الفهم لجميع المستويات التقنية.





عبر مختلف المجالات، بما في ذلك الأعمال التجارية والرعاية الصحية والتمويل، يسهل علم البيانات اتخاذ قرارات أكثر استنارة واستراتيجية من خلال عملية منهجية من خمس خطوات:

1. حدد أهدافك:

حدد بوضوح أهداف عملية اتخاذ القرار المستندة إلى البيانات، وحدد التحديات أو الفرص بما يتماشى مع الأهداف التنظيمية.

2. حدد مصادر البيانات:

قم بتقييم وتحديد مصادر البيانات، مع الأخذ في الاعتبار مجموعات البيانات الداخلية أو المصادر الخارجية أو المزيج. قيم جودة البيانات وإمكانية الوصول إليها وأهميتها.



3. بناء / شراء حل لمعالجة البيانات:

تطوير أو تحديد حل لمعالجة البيانات، مما يضمن الكفاءة وقابلية التوسع وأمن البيانات والتكامل مع المصادر المتنوعة.



4. استخراج القيمة من البيانات:

قم بإجراء التحليل باستخدام الأساليب الإحصائية والتعلم الآلي لاستخلاص رؤى قابلة للتنفيذ، بما يتماشى مع الأهداف المحددة.



5. تعزيز الثقافة القائمة على البيانات

في مؤسستك:

زراعة ثقافة تنظيمية تقدر عملية صنع القرار المبنية على البيانات. رفع مستوى الوعي وتشجيع التعاون وتعزيز المعرفة بالبيانات.

