

المملكة العربية السعودية

رؤية  
2030  
المملكة العربية السعودية  
KINGDOM OF SAUDI ARABIA



وزارة التعليم  
Ministry of Education



# علم البيانات

قناة مسار علوم الحاسب والهندسة

أ- يحيى الحكمي

# 1. مقدمة في علم البيانات



سيتعرف الطالب في هذه الوحدة على المفاهيم الأساسية لعلم البيانات، ويشمل ذلك تعريف معنى البيانات والمعلومات والمعرفة والفرق بينهم، وسيتم تناول موضوع دورة حياة علم البيانات وطريقة التعامل مع البيانات الضخمة، وأيضاً سيتم مناقشة بعض الموضوعات الخاصة بالسياسات المتعلقة بالبيانات وحوكمة البيانات.

وفي الختام سيتعرف الطالب على أساسيات علم البيانات مع التركيز على الفرص الوظيفية التي يوفرها علم البيانات.

## أهداف التعلم

بنهاية هذه الوحدة سيكون الطالب قادراً على أن:

- < يعرف مصطلح علم البيانات.
- < يفرق بين البيانات والمعلومات والمعرفة.
- < يفرق بين علم البيانات وذكاء الأعمال.
- < يوضح أوجه التقارب بين علم البيانات والذكاء الاصطناعي.
- < يحدد مراحل دورة حياة علم البيانات.
- < يعرف مصطلح البيانات الضخمة.
- < يحدد خصائص البيانات الضخمة.
- < يصنف تقنيات البيانات الضخمة.
- < يعرف مفهوم إدارة البيانات.
- < يحدد مبادئ حوكمة البيانات.
- < يناقش المهارات والأدوات التي يتطلبها علم البيانات.
- < يحدد المهن المتعلقة بعلم البيانات.
- < يوضح أهمية المجتمعات الرقمية لعلم البيانات.

متطلب البرمجة بلغة البايثون



يتطلب منهجي علم البيانات والهندسة في نظام المسارات معرفة أساسيات البرمجة بلغة البايثون. يرجى مسح رمز الاستجابة السريع أدناه للوصول لمحتوى تعريفي بالبايثون. ولمعرفة الموضوعات المتوفرة والوصول السريع لكل وحدة، يمكنك الاطلاع على الصفحات 208-209.

# ماذا تعلمت

- < المقصود بعلم البيانات.
- < الفرق بين البيانات والمعلومات والمعرفة.
- < كيفية التمييز بين علم البيانات وذكاء الأعمال والذكاء الاصطناعي.
- < تعريف دورة حياة علوم البيانات.
- < المقصود بالبيانات الضخمة.
- < كيفية مساهمة الباحثون أو الأدوات الأخرى في علم البيانات.

## المصطلحات الرئيسية

Artificial Intelligence	ذكاء اصطناعي	Data Visualization	تصوير البيانات
Big Data	البيانات الضخمة	Data Warehouse	مستودع البيانات
Business Intelligence	ذكاء الأعمال	Exploratory Data Analysis	التحليل الاستكشافي للبيانات
Cloud Storage	التخزين السحابي	Information	معلومات
Data	البيانات	In-Memory Computing	الحوسبة في الذاكرة
Data Analysis	تحليل البيانات	Jupyter Notebook	مفكرة جوبيتر
Data Analyst	محلل البيانات	Knowledge	المعرفة
Data Cleaning	تنظيف البيانات	Python	البايثون
Data Collection	جمع البيانات	Raw Data	بيانات أولية
Data Lake	بحيرة بيانات	Value	القيمة
Data Mining	التنقيب في البيانات	Variety	التنوع
Data Preparation	تجهيز البيانات	Velocity	السرعة
Data Science	علم البيانات	Veracity	الموثوقية
Data Science Life Cycle	دورة حياة علم البيانات	Volume	الحجم
Data Scientist	عالم البيانات		

## 2. جمع البيانات والتحقق من صحتها



سيتعرف الطالب في هذه الوحدة على مفاهيم جمع البيانات وطرق التحقق من صحتها. ويشمل ذلك معنى جمع البيانات وأنماطها المختلفة ومصادرها. وسيتم التطرق إلى موضوع ترميز البيانات والتركيز على ميزات وعيوب هذه الطرق، وفي الختام سيتعرف الطالب على خطوات التحقق من صحة البيانات وأنماطها.

### أهداف التعلم

- بنهاية هذه الوحدة سيكون الطالب قادرًا على أن:
- < يعرف مصطلح جمع البيانات.
- < يصنف مصادر البيانات.
- < يصنف معايير جودة المعلومات.
- < يعرف مفهوم منصات البيانات المفتوحة.
- < يتعرف على أهمية الحصول على التصاريح القانونية المناسبة قبل البدء بجمع البيانات.
- < يحدد الأنماط المختلفة للبيانات.
- < يعرف المقصود بترميز البيانات.
- < يطبق خطوات التحقق من صحة البيانات.
- < يصنف الطرق المتنوعة للتحقق من صحة إدخال البيانات.

## ماذا تعلمت

- < ما المقصود بجمع البيانات.
- < التمييز بين أنواع البيانات المختلفة.
- < كيفية ترميز البيانات.
- < كيفية التحقق من صحة البيانات.
- < كيفية التحقق من جودة المعلومات.
- < كيفية التمييز بين مصادر البيانات الرئيسة ومصادر البيانات الثانوية.

### المصطلحات الرئيسة

Accuracy	الدقة	Information Quality	جودة المعلومات
Appropriateness	الملاءمة	Isbn	رقم الكتاب المعياري الدولي
Audio Data	البيانات الصوتية	Legal Permissions	الأذونات القانونية
Barcode	الرموز الشريطية	Length Check	التحقق من الطول
Check Digit	التحقق من الرقم	Level Of Detail	مستوى التفاصيل
Completeness	الكفاية	Lookup Check	التحقق من البحث
Data	البيانات	Open Data Platforms	منصات البيانات المفتوحة
Data Coding	ترميز البيانات	Presence Check	التحقق من التواجد
Data Types	أنواع البيانات	QR Code	رمز الاستجابة السريعة
Data Validation	التحقق من صحة البيانات	Range Check	التحقق من النطاق
Fixed Data	بيانات ثابتة	Type Check	التحقق من النوع
Format Check	التحقق من الصيغة	Variable Data	بيانات متغيرة
Graphical Data	البيانات المرئية	Video Data	بيانات الفيديو

# 3. التحليل الاستكشافي للبيانات



في الوحدات السابقة، تعلم الطالب مفهوم البيانات وأنماطها المختلفة وكيفية جمعها بشكل صحيح. وفي هذه الوحدة، سيتعلم الطالب طريقة فحص وتحليل البيانات لفهمها بشكل أفضل.

## أهداف التعلم

- بنهاية هذه الوحدة سيكون الطالب قادرًا على أن:
  - < يصنّف أنماط تحليل البيانات.
  - < يعرف مفهوم التحليل الاستكشافي للبيانات.
  - < يصنّف أنماط التحليل الاستكشافي للبيانات.
  - < يشرح مراحل عملية التحليل الاستكشافي للبيانات.
  - < يعرف مفهوم المكتبات البرمجية.
  - < يطور برنامج لتحليل البيانات باستخدام مكتبات البرمجة.
  - < يستخدم تقنيات إعداد البيانات وتنظيفها في مجموعة بيانات.
  - < يشرح أهمية تصوير البيانات.
  - < يميز الأنماط المختلفة من الرسوم البيانية التي يمكن إنشاؤها باستخدام مكتبات البايثون.

## ماذا تعلمت

- < خطوات عملية تحليل البيانات.
- < استخدام مفكرة جوبيتر كأداة لتحليل البيانات.
- < استخدام مكتبة بانداس لإنشاء الإحصاءات.
- < أهمية تصوير البيانات.
- < استخدام مكتبة مات بلوت لربط لتمثيل البيانات بشكل رسومي.
- < إنشاء مخطط بياني شريطي ومخطط دائري في مفكرة جوبيتر.

### المصطلحات الرئيسية

Attribute	سمة	Grouping	تجميع
Data Cleaning	تنظيف البيانات	Indexing	فهرسة
Data Frame	إطار البيانات	Method	وظيفة
Data Visualization	تصوير البيانات	Multivariate	متعدد المتغيرات
Descriptive Analysis	التحليل الوصفي	Non-Graphical Analysis	تحليل غير رسومي
Diagnostic Analysis	التحليل التشخيصي	Predictive Analysis	التحليل التنبؤي
Exploratory Data Analysis	تحليل البيانات الاستكشافي	Prescriptive Analysis	التحليل التوجيهي
Filtering	تصفية	Programming Library	مكتبة البرمجة
Function	دالة	Series Object	كائن المتسلسلة
Graphical Analysis	تحليل رسومي	Univariate	أحادي المتغير

# 4. نمذجة البيانات التنبؤية والتوقع

سيتعرف الطالب في هذه الوحدة على مفهوم نمذجة البيانات التنبؤية والتوقع، ويشمل ذلك مفهوم النمذجة التنبؤية، وأنماط النماذج التنبؤية وتطبيقاتها. وأيضاً، سيتعلم الطالب مفهوم التوقع، وشرح وتوضيح خطوات الوصول إلى نتائج التوقع، وسيتم التركيز على مفهوم تحسين الحلول، وذلك من خلال صياغة المشكلة وإيجاد الحل الأمثل لها من بين الحلول الممكنة باستخدام أداة إكسل سولفر (Excel Solver).

وفي الختام، سيتعلم الطالب طريقة تقييم النتائج التي يحصل عليها وذلك بهدف الوصول إلى أفضل النتائج والتوصيات المتعلقة بالإجراءات أو العمليات التي قد يتم تنفيذها في المستقبل.

بنهاية هذه الوحدة سيكون الطالب قادراً على أن:

- < يُعرّف مصطلح النمذجة التنبؤية.
- < يُقدّم وصفاً واضحاً لفئات النمذجة التنبؤية.
- < يُحدّد خطوات عملية النمذجة التنبؤية.
- < يُعدّد ميزات وعيوب النمذجة التنبؤية.
- < يُعرّف مصطلح التوقع.
- < يُحدّد الخطوات المتبعة في عمليات التنبؤ.
- < يُنفذ عملية التوقع في مايكروسوفت إكسل (Microsoft Excel).
- < يُعرّف مفهوم نطاق الثقة.
- < يُصنّف المخططات أو الرسوم البيانية المتنوعة للتوقع.
- < يُعرّف مفهوم نموذج التحسين.
- < يُطبّق عملية التحسين باستخدام أداة إكسل سولفر.
- < يُقيم نتائج عملية التحسين ويُحدّد التوصيات المستقبلية.

## ماذا تعلمت

- < مصطلح النمذجة التنبؤية.
- < الاختلاف بين النمذجة المعاملية وغير المعاملية.
- < الأنواع المختلفة للنماذج التنبؤية.
- < عملية إنشاء نموذج تنبؤي.
- < ميزات وتحديات النمذجة التنبؤية.
- < تطبيقات النمذجة التنبؤية.
- < مصطلح التوقع.
- < الأنواع المختلفة لمخططات التوقع.
- < خطوات تطبيق توقع على بيانات محددة.
- < مفهوم نطاق الثقة.
- < مصطلح التحسين وإيجاد الحلول المثالية للمشكلة.
- < خطوات إجراء التحسينات باستخدام أداة إكسل سولفر.
- < خطوات تقييم نتائج أداة إكسل سولفر.

## المصطلحات الرئيسية

Classification Model	نموذج التصنيف	Lower Confidence Bound	الحد الأدنى للثقة
Clustered Column Chart	مخطط عمودي متفاوت المسافات	Model Formulation	تكوين النموذج
Clustering Model	نموذج التجميع	Neural Network	شبكة عصبية
Column Chart	مخطط عمودي	Non-Parametric	غير مُعاملي
Confidence Interval	نطاق الثقة	Objective Cell	الخلية الهدف
Constraints	القيود	Optimization Problem	مشكلة التحسين
Data Collection	جمع البيانات	Outlier Detection Model	نموذج اكتشاف القيم الشاذة
Data Purification	تنقية البيانات	Parametric	مُعاملي
Data Transformation	تحويل البيانات	Predictive Data Modeling	نمذجة البيانات التنبؤية
Decision Tree	شجرة القرار	Prophet Model	نموذج بروفيت
Excel Solver	أداة إكسل سولفر	Quality Enhancement	تحسين الجودة
Forecast	توقع	Risk Assessment	تقييم المخاطر
Forecast Model	نموذج التوقع	Stacked Column Chart	مخطط عمودي مكس
General Linear Model	نموذج خطي عام	Time Series Model	نموذج السلاسل الزمنية
Gradient Boosted Model	نموذج التعزيز الاشتقاقي	Upper Confidence Bound	الحد الأعلى للثقة
Line Chart	مخطط خطي	Variable Cells	خلايا المتغير
Linear Regression	الانحدار الخطي		

# متطلب البرمجة بلغة بايثون

تعد البرمجة أحد أهم المهارات التي ينبغي اكتسابها من قبل الطلاب المتحقين بمسار علوم الحاسب والهندسة حيث تعد متطلب لعدد من المناهج في هذا المسار ومنها منهجي الهندسة وعلم البيانات. لتسهيل اكتساب الطالب لأساسيات البرمجة بلغة البايثون، فقد تم تصميم المحتوى التالي الذي يمكن الوصول إليه بمسح رمز الاستجابة السريع الخاص بكل موضوع. وينصح الطالب بوضع خطة زمنية لإتمام الاطلاع على هذه الوحدات ويمكن الاسترشاد بالمدة الزمنية المقترحة لكل وحدة كما يمكن للطالب وضع علامة (V) لتعليم الوحدات التي أتمها.

هل أتممت الوحدة؟	رمز الاستجابة السريع	المدة الزمنية المقترحة	الوحدة
		يوم واحد	1. مقدمة في البايثون Introduction to Python
		يوم واحد	2. المدخلات والمخرجات والعمليات الحسابية Input-Output and Mathematical Operations
		يومان	3. الجمل الشرطية Conditional Statements

هل أتممت الوحدة؟	رمز الاستجابة السريع	المدة الزمنية المقترحة	الوحدة
		يومان	4. التكرارات والدوال Loops and Functions
		أسبوع	5. القوائم وصفوف البيانات والمكتبات البرمجية Lists, Tuples and Python Libraries
		أسبوع	6. القواميس والقوائم المتداخلة وملفات البيانات Dictionary, Nested Lists and Data Files
		أسبوعان	7. هياكل البيانات المتقدمة ودوال الاستدعاء الذاتي Advanced Data Structures and Recursion
		أسبوعان	8. مقدمة في البرمجة الكائنية Introduction to Object Oriented Programming