مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة

Essential Data Science Skills for Undergraduate Students Across Disciplines

إعداد

أ / محمد عادل محمد محمد صقر

محاضر الرياضيات والإحصاء بقسم العلوم الأساسية

بالجامعة الخليجية بالبحرين

Mammsakr@gmail.com

أ.د/ وفاء مصطفى كفافى

أستاذ المناهج وطرق تعليم الرياضيات

كلية الدراسات العليا للتربية

جامعة القاهرة

wmkefafe@cu.edu.com

د. صلاح أحمد فؤاد

مدرس المناهج وطرق تدريس الرياضيات

كلية الدراسات العليا للتربية

جامعة القاهرة

salahmath@hotmail.com

مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة

مستخلص البحث:

هدفت هذه الدراسة إلى التعرف على مهارات علم البيانات (وجوانبها المعرفية والأدائية) الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة، ولتحقيق هدف الدراسة، تم استخدام المنهج الوصفي المسحي، وتمثلت أداة الدراسة في استبانة لاستطلاع الرأي لتحديد مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة.

وأظهرت النتائج أن مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة، هي: وضع خطة العمل، وجمع البيانات، ومعالجة البيانات، وتحليل البيانات، وتنقيب البيانات، وعرض النتائج.

وقد أوصت الدراسة بإعداد مقررًا لتنمية مهارات علم البيانات لطلاب المرحلة الجامعية، وإثراء مقررات الإحصاء القائمة في المرحلة الجامعية بمهارات علم البيانات، وتوفير المواد والأدوات والبرامج اللازمة لتنمية مهارات علم البيانات لدى الطلاب، وتنظيم مؤتمرات وورش عمل لأعضاء هيئة التدريس حول كيفية تطعيم مقررات الإحصاء القائمة بأساسيات علم البيانات ومهاراته.

الكلمات المفتاحية: مهارات علم البيانات، طلاب المرحلة الجامعية.

Essential Data Science Skills for Undergraduate Students Across Disciplines

Mr. Mohamed Adel Sakr
Lecturer of Mathematics and Statistics at the Basic Sciences Department at the Gulf University in Bahrain Email
Mammsakr@gmail.com

Prof. Wafaa Mostafa Kafafy
Emeritus professor of Mathematics Curricula and Teaching Methods – Faculty of Graduate Studies for Education Cairo University
wmkefafe@cu.edu.com

Dr. salah ahmed fouad
Lecturer of Curriculum and Instruction Faculty of Graduate Studies for Education- Cairo University
salahmath@hotmail.com

Abstract:

The present study identified the data science skills (and their cognitive and performance aspects) required for undergraduate students. A descriptive survey methodology was employed to achieve this objective. Data collection utilized an opinion poll questionnaire designed to identify the perceived essential data science competencies for undergraduate students.

The results indicated that the following data science skills were deemed essential for undergraduate students: developing an action plan, data collection, data processing, data analysis, data mining, and result presentation.

The study recommends the development of a dedicated data science course, enrichment of existing university-level statistics courses to incorporate data science skills, and the provision of necessary materials, tools, and programs to facilitate students' data science skill development. Additionally, organizing conferences and workshops for faculty members on integrating data science concepts into their courses would be beneficial.

Keywords: data science skills, undergraduate students

مقدمة:

شهد العالم في مطلع القرن الحادي والعشرين ثورة علمية ورقمية وتكنولوجية هائلة أدت إلى انفجار هائل في حجم البيانات المتاحة إلكترونيًا، وتنوعها بشكل مذهل لتشمل: النصوص، والصور، والصوت، والفيديوهات... ولقد أدى هذا التراكم الهائل لـ "البيانات الضخمة" إلى تسارع وتيرة استثمار الدول في هذه الثروة المعرفية الجديدة، التي باتت تُعرف باسم "وقود الاقتصاد الجديد" أو "اقتصاد المعرفة". حيث أدركت الدول إمكانياتها الهائلة في اتخاذ قرارات فعّالة ومدروسة في مختلف المجالات، شملت المجالات الطبية، والتعليمية، والصناعية، والتجارية، والاقتصادية، والعسكرية...

وبات امتلاك مهارات وضع خطط العمل في مشاريع البيانات الضخمة، وجمع تلك البيانات، ومعالجتها، وتحليلها، واستخراج المعرفة منها ضرورة ملحة في عصرنا الحالي

ويُعد علم البيانات مجالًا تعليميًا متعدد التخصصات يجمع بين المعارف في مجالات الرياضيات، والإحصاء، وعلوم الكمبيوتر، إلى جانب المعرفة في مجالات الأعمال المختلفة – التجارية، والصناعية، والفنية. كما يضم هذا المجال المهارات وطرق التفكير والأساليب المنهجية العلمية لكل من هذه المجالات.

ويهدف علم البيانات إلى فهم الظواهر، ووصفها، وتفسيرها، والتنبؤ بالنتائج في مختلف المجالات. ويستخدم البيانات الضخمة لاستخلاص المعلومات والمعرفة، واكتشاف الرؤى، وبناء النماذج الآلية، وابتكار خدمات ومنتجات تناسب متطلبات العصر الحالي.

ويُعدّ عام (1974) نقطة الانطلاق لعلم البيانات، حيثُ ابتكره "بيتر نوير" تحت اسم Datalogy (Press, 2013). بينما ظهر مصطلح "علم البيانات" رسميًا للمرة الأولى عام (1996) في وقائع المؤتمر الخامس للاتحاد الدولي لجمعيات التصنيف في اليابان (Hayashi, et al., 1996).

ويعرَّف علم البيانات بأنه "العلم الذي يستخدم أساليب وإستراتيجيات كَمّية ونوعيّة للتعامل مع البيانات الضخمة بهدف حل المشكلات ذات الصلة، والتنبؤ بالنتائج" (Waller & Fawcett, 2013, p. 78).

ويتطلب التعامل مع البيانات الضخمة، وفهم خصائصها، حزمة من المهارات المُتكاملة التي تُشكل منهجية علم البيانات وعملياته، ووفقا لـ (Wang & Li, 2019, P.131) تتضمن مهارات علم البيانات الأساسية تخزين البيانات ومعالجتها وتحليلها، أما (Baumer, 2015, P.340) فوسع نطاق مهارات علم البيانات ليشمل: الحصول على مصادر البيانات، وتخزين البيانات واسترجاعها، ومعالجتها وإدارتها، ونمذجتها والتنقيب عن البيانات باستخدام الأدوات المُناسبة، وعرض النتاج وتصورها.

ولعلم البيانات إمكانيات هائلة لحلّ المشكلات الواقعية المعاصرة، وابتكار منتجات وخدمات جديدة في مجالات شتّى، منها: " الإعلان، والنشر، ووسائل الإعلام، والفنون، والتصميم الإبداعي، والعلوم وعلم الفلك، والفضاء، والطب، والرعاية الصحية، والمعلوماتية الحيوية Bioinformatics، والبيئة، والعلوم الإنسانية، والاجتماعية، والتعليم، والتجارة الإلكترونية، والتمويل، والتسويق، والمبيعات، والإدارة، والعمليات والتخطيط، والخِدمات الاستشارية، والصناعة، والزراعة، والهندسة، والأمن والسلامة، وصناعة الألعاب والرياضة، والترفيه..." (Cao, 2018, p. 263).

وتُعد البيانات الضخمة ظاهرة تُشكل تحديًا وثروة هائلة في آن واحد، ويكمن التحدي في صعوبة التعامل معها نتيجة لحجمها الهائل ولتنوع تنسيقاتها، بينما تعد البيانات الضخمة ثروة تتمثل في إمكانية استخلاص رؤى قيّمة تُساهم في تحسين مختلف جوانب الحياة. لذلك، برزت الحاجة الماسة إلى أدوات وتقنيات متخصصة تُسهل عملية معالجة البيانات الضخمة وتحويلها إلى معلومات ومعرفة يمكن تطبيقها في المجالات المختلفة، وتتضمن تلك الأدوات: أدوات لجمع البيانات الضخمة وإعدادها، وتخزينها، واسترجاعها مثل قواعد بيانات NoSQL ولغات برمجة SQL وJava وScala والمنصات السحابية Amazon web services و Azure و Hadoop و Google Cloud platform، وتوفر تلك المنصات السحابية خِدمات تحليل البيانات الضخمة، أو يمكن استخدام برامج الجداول الحسابية Spreadsheets في ذلك، وتوجد أيضا أدوات لتصور البيانات، وعرض النتائج على Dashboard مثل: Tableau و Power BI و Looker ، كما يمكن استخدام بعض البرامج مثل: RapidMiner Studio أو لغات البرمجة مثل: Python و لغة R في إجراءات العمليات السابقة، علاوة على استخدامهما في إعداد نماذج الذكاء الاصطناعي، وتعلم الألة والتنقيب في البيانات.

ووفقًا لتقرير "مستقبل الوظائف" الصادر عن المنتدى الاقتصادي العالمي لعام (2020)، احتلت "الوظائف والأعمال المرتبطة بعلم البيانات" المرتبة الثالثة ضمن أكثر الوظائف الواعدة مستقبلًا (W.E.F., 2020, p. 62).

ويفرض ذلك الواقع على المؤسسات التعليمية، وخاصةً الجامعات بناءً على دورها المحوري في إعداد خريجين مؤهلين بالمعارف والمهارات اللازمة لسوق العمل العصري، ضرورة التجاوب مع معطيات هذا العصر ومتطلباته.

وعليه شهدت مناهج الإحصاء بالمرحلة الجامعية في بعض دول العالم تطورات وتغيرات ملموسة، حيث قامت هذه الدول بإعادة النظر في محتوى مناهج الإحصاء لديها؛ سعيًا لملاحقة التطورات المتسارعة في العلم، وتطبيقاته، وفي السير قدمًا نحو الرقي والتقدم، والتحول نحو الاقتصاد المبنيّ على البيانات الضخمة والمعرفة، "وأيضًا في حل مشكلات عصرية في عالم البيانات، والمعلومات، والتكنولوجيا" (نظلة خضر، 2020، ص 13)، "وتتسم مشكلات هذا العصر بالتعقيد، والتداخل بين فروع المعرفة المختلفة، الأمر الذي يتطلب ضرورة التخلي عن النظرة الخطية في إعداد المقررات بالمرحلة الجامعية، وبناء برامج تدعم اتجاه التكامل بين مختلف فروع المعرفة" (فايز مينا، 2013، ص 18) كعلم البيانات.

وعليه أقدمت بعض مؤسسات التعليم الجامعي حول العالم على إضافة تخصص منفرد لدراسة علم البيانات، أو وضع مقرر له ضمن الخطة الدراسية في التخصصات المختلفة، أو برنامج حر تقدمه؛ بهدف التطوير العلمي والمهني لطلابها (Press G. , 2013; Ceccucci, Tamarkin, & Jones, 2015).

وفي ضوء ما سبق، تتضح أهمية اكتساب مهارات علم البيانات لطلاب المرحلة الجامعية، وهو ما يبرز أهمية الدارسة الحالية في التعرف مهارات علم البيانات(بجوانبها المعرفية والأدائية) الواجب توافراها لدى طلاب المرحلة الجامعية.

مشكلة الدراسة:

نبع الإحساس بمشكلة الدراسة من خلال:

ما أوصت به الدراسات السابقة التي نادت بأهمية علم البيانات، واكتساب مهاراته في المراحل التعليمية المختلفة، وبمستويات مناسبة، بالإضافة إلى أن إدخال موضوعات جديدة يؤكد على أن علم الإحصاء علم حي متجدد متطور يتماشى ومتطلبات العصر، ومن هذه الدراسات:

(Baumer, 2015; Ceccucci, Tamarkin, & Jones, 2015; Saltz & Heckman, 2016; Brunner & Kim, 2016; McCoy & Shih, 2016; West, 2018; Alturkistani, et al., 2018; Guzman, Pennell, & Nikelski, 2019; Kahn, 2019; Yan & Davis, 2019; Salas-Rueda, Ricardo-Adán, Érika-Patricia, & Rodrigo-David, 2020 & Sami, Sinclair, Stein, & Medsker, 2020).

- توصيات المنتديات العلمية عن أهمية اكتساب مهارات علم البيانات مثل: المنتدى الاقتصادي العالمي لعام (2020)، والذي صنف الوظائف والأعمال المرتبطة بعلم البيانات في المرتبة الثالثة ضمن أكثر الوظائف الواعدة مستقبلًا (W.E.F., 2020, p. 62)، وضرورة اكتساب الأفراد لمهارات علم البيانات.

- خبرة الباحث في الميدان من خلال العمل كمحاضر لمقررات الإحصاء بالمرحلة الجامعية بمملكة البحرين، حيث لاحظ ضعف مستوى الطلاب في مهارات علم البيانات الذي ظهر من خلال تكليفهم في محاضرات الإحصاء ببعض المهام المرتبطة بتلك المهارات، ومناقشاتهم وملاحظة أدائهم فيها وتتضمن تلك المهام: جمع البيانات، وتخزينها، واسترجاعها، ومعالجة البيانات مثل: حذف التكرار في الحالات، والتعامل مع البيانات المفقودة والمتطرفة، وتحليل البيانات، وإجراء التنبؤ بالنتائج، وعرضها.
- اطلاع الباحث على الدراسات العربية والأجنبية السابقة لم يجد أي دراسة تحاول التعرف على مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة.

وعليه، تتمثل مشكلة هذه الدراسة عدم وجود قائمة تشتمل على مهارات علم البيانات (وجوانبها المعرفية والأدائية) الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة.

أسئلة الدراسة:

تسعى الدراسة للإجابة عن السؤال التالي:

ما مهارات علم البيانات (جوانبها المعرفية والأدائية) الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة؟

أهداف الدراسة:

تسعى الدراسة لتحقيق الهدف التالي:

تعرف مهارات علم البيانات (وجوانبها المعرفية والأدائية) الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة.

أهمية الدراسة:

تتضح أهمية الدراسة في البعدين التاليين:

1. الأهمية العلمية النظرية:

تتضح الأهمية العلمية للدراسة في النقاط الآتية:

تُعدّ هذه الدراسة من الدراسات الحديثة التي اهتمت بتحديد مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة، واستجابة للاتجاهات الحديثة التي تنادي بأهمية علم البيانات، وهو ما يجعلها إضافة علمية للمكتبة العربية التربوية.

- تأتي هذه الدراسة استجابة لتوصيات الدراسات السابقة المتخصصة.
- أهمية التعرف على مهارات علم البيانات واكتسابها لدى الطلاب في المرحلة الجامعية، لإعداد الطالب الجامعي المواكب لمتطلبات سوق العمل والقادر على التعامل مع معطيات التطور العلمي والرقمي والتكنولوجي ومتطلباته.

2. الأهمية العلمية التطبيقية:

تتضح الأهمية العملية للدراسة من خلال العوائد التي تنعكس على المستفيدين من نتائج الدراسة، ومن ذلك:

- توفير قائمة بمهارات علم البيانات الواجب توافراها لدى طلاب المرحلة الجامعية في التخصصات المختلفة.
- توجيه نظر أساتذة الرياضيات والإحصاء بالمرحلة الجامعية لعلم البيانات كأحد العلوم الحديثة والعصرية، ومهاراته التي يمكن تطعيم مقرراتهم بأجزاء منه، أو بتخصيص مقرر له في برامجهم.
- توجيه نظر مخططي ومطوري مناهج الرياضيات والإحصاء بالمراحل التعليمية المختلفة إلى ضرورة الاهتمام بمهارات علم البيانات وتضمينها ضمن المناهج التعليمية.
- تقديم مجموعة من المقترحات لبحوث في مجالات علم البيانات ومهاراته.

حدود الدراسة:

1. الحدود الموضوعية:

تحديد المهارات الأساسية والفرعية لعلم البيانات الواجب توافرها لدى الطلاب الدارسين لمقرر مبادئ الإحصاء بالمرحلة الجامعية.

2. الحدود الزمنية:

إجريت هذه الدراسة في الفصل الربيعي للعام الدراسي(2024/ 2023).

مصطلحات الدراسة:

1- علم البيانات Data Science

يعرفه الباحث إجرائيًا: بأنه العلم الذي يطبق فيه الطالب الأساليب المناسبة للتعامل مع البيانات الضخمة؛ لفهم الظواهر المرتبطة بتلك البيانات، ووصفها، وتفسيرها، واكتشاف الرؤي منها، وتوليد المعرفة ونشرها، واكتساب الحكمة في اتخاذ القرارات المناسبة لحل المشكلات ذات الصلة بها.

2- مهارات علم البيانات Data Science Skills

يعرفها الباحث إجرائيًا: بأنها تلك المهارات التي يمارسها الطالب، وتمكنه من التعامل مع البيانات الضخمة، وتشمل هذه المهارات الجوانب المعرفية والأدائية لمهارات: وضع خطة العمل، وجمع البيانات، ومعالجتها، وتحليلها، وتنقيبها، وعرض النتائج بهدف التواصل مع الآخرين بشأن تلك النتائج وتصورها. وتقاس بالدرجة التي يحصل عليها الطالب في اختبار مهارات علم البيانات الذي يتكون من اختبار الجانب المعرفي واختبار الجانب الأدائي لمهارات علم البيانات.

الإطار النظري والدراسات السابقة:

أولًا: الإطار النظري

1. نشأة علم البيانات ومفهومه

لم تكن نشأة علم البيانات مرتبطة بالتطور العلمي والرقمي والتكنولوجي المتسارع في الفترة الأخيرة فحسب، وإنما أيضًا كمحصلة لتراكم المعرفة والتطورات التي توصل لها العلماء في العقود الستة الماضية في مجالات الرياضيات والإحصاء وعلوم الكمبيوتر.

فقد شهدت تلك الفترة، منذ ستينيات القرن العشرين، عملية تحول مرحلية في استبدال مصطلح "تحليل البيانات Data Analysis" – تلك العملية التي تبدأ بجمع البيانات ثم التعامل معها بالأساليب المناسبة بهدف اتخاذ قرارات فعالة – بمصطلح "علم البيانات Data Science"؛ وكان ذلك نتيجة للتطورات المتتالية الحادثة في تلك المجالات الثلاثة، وظهور تطبيقات إنترنت الأشياء وتعلم الآلة والتعلم العميق، والاستفادة منها في التعامل مع البيانات الضخمة– التي يتم جمعها آليا– لحل المشكلات العصرية والتوصل لرؤى واتخاذ قرارات فعالة في المجالات المختلفة. ويعد علم البيانات مشتقا في أصله من علم الإحصاء وامتداد طبيعي له نتيجة التغير الحادث في طبيعة البيانات وأساليب معالجتها، وأيضا كنتيجة لتطبيق الجديد في علوم الكمبيوتر والبرمجة والطرق الرياضياتية المتقدمة في التعامل مع البيانات (Cao, 2018, P.10).

ويعد عالم الفلك والكمبيوتر Peter Naur أول من استخدم مصطلح "علم البيانات Data Science" في عام (1960) كمصطلح يشير إلى المجال الذي يهتم باستخراج المعرفة من البيانات وتعميمها (Wainer, 2015, P.2).

وتلي ذلك، ما نشره عالم الرياضيات والإحصاء John Tukey (1962, P. 64) في مقالة تناولت مجموعة من الأساليب المبتكرة لتحليل للبيانات التي تشبه الأساليب المطبقة في علم البيانات بصورته الحالية، وكان الهدف من تلك الأساليب مواجهة التحديات وحل المشكلات المتزايدة المتعلقة بطبيعة البيانات وتنوع مصادرها وتصنيفاتها وخصائصها، فضلًا عن ضمان واقعية النتائج في المجالات المختلفة.

ومع استمرار دراسات Peter Naur لابتكار أساليب جديدة في التعامل مع البيانات، استخدم مصطلح "Datalogy" بديلًا لمصطلح "Data Science" في بحثه المنشور عام (1968) بالمؤتمر العلمي للاتحاد الدولي لمعالجة البيانات[1] (Naur, 1968)، إلى أن أستقر على استخدام مصطلح "Data Science" مرة أخرى عام (1974) في كتابه "Concise Survey of Computer Methods" للدلالة على العلم الذي يعالج البيانات، ويربط بين علوم الكمبيوتر والعلوم الأخرى (Naur, 1974).

وتبع ذلك، استخدام عالم الإحصاء Chien-Fu Wu (1986) مصطلح "علم البيانات Data Science" للإشارة إلى علم الإحصاء في محاضرة ألقاها في الأكاديمية الصينية للعلوم في العاصمة بكين.

عَقب ذلك، ما قام به كل من Escoufier (1995) و Hayashi في تطوير مفهوم علم البيانات ووصفه بأنه نظام بيني التخصصات يتضمن المعارف والأساليب من علوم الرياضيات والإحصاء والكمبيوتر، وقد تم اعتماد "علم البيانات Data Science" رسميًا كعلم مستقل لأول مرة في وقائع المؤتمر الخامس للاتحاد الدولي لجمعيات التصنيف في اليابان.

وشهد علم البيانات منذ ذلك الحين اهتمامًا متزايدًا، تجلى ذلك في إقامة العديد من المؤتمرات وإصدار الدوريات العلمية المتخصصة في مجالاته وتطبيقاته المتنوعة، ومن أبرز الأمثلة على ذلك: المؤتمر العلمي السنوي لعلم البيانات بنيويورك[2] بالولايات المتحدة الأمريكية، والمؤتمر العالمي لعلوم البيانات والإحصاء بأمستردام في هولندا[3]، ودورية علم البيانات التي صدر عددها الأول في أبريل (2002) عن اللجنة المعنية بالبيانات في العلوم والتكنولوجيا للمجلس الدولي للعلوم[4]، ودورية علم البيانات التي صدر عددها الأول في يناير (2003) عن مركز الإحصاء التطبيقية في كلية الإحصاء في جامعة رينمين في جمهورية الصين الشعبية[5]. ولم يقتصر الاهتمام بعلم البيانات على ذلك، بل شمل أيضًا اهتمام العديد من الجامعات والكليات والمعاهد في مختلف دول

[1] International Federation for Information Processing (IFIP)
[2] Data Science Salon NYC
[3] World Conference on Data Science & Statistics, Amsterdam, Netherlands
[4] Committee on Data (CODATA) of the International Science Council (ISC)
[5] The Center for Applied Statistics (CAS) of the School of Statistics at Renmin University of China

العالم بمواكبة التطورات الحادثة في مجال البيانات باستحداث أقسام خاصة لعلم البيانات كتخصص جامعي منفرد، أو كتخصص فرعي في الأقسام المتاحة فيها، أو تطوير مناهجها من خلال دمج موضوعات علم البيانات بالتخصصات المتعلقة بالإحصاء والبيانات، وكان ذلك على مستوى المرحلة الجامعية الأولى أو أيضًا في مراحل الدراسات العليا(Press G. , 2013; Ceccucci, Tamarkin, & Jones, 2015).

ومن ناحية أخرى، شهد تعريف علم البيانات تطورات مستمرة، مواكبًة للتطورات المتسارعة في طبيعة البيانات ومجالات معالجتها. وقد قدمت العديد من المحاولات لتعريفه، حتى ذهب البعض إلى شمولية هذا العلم وقدرته على استيعاب الكثير من المصطلحات الأخرى ذات الصلة، على سبيل المثال مصطلحات: البيانات الضخمة، وتحليل البيانات، والتنقيب في البيانات، وتعلم الآلة... وفيما يلي أبرز التعريفات المعبرة عن دلالة هذا المصطلح.

فيُشير كل من Saltz (2016, P.85) و Heckman إلى علم البيانات كمجال ناشئ يجمع بين مجموعة من المعارف والأساليب في مجالات مختلفة، بما في ذلك علوم الكمبيوتر، وإدارة البيانات، والتواصل الفعال، وعلم الإحصاء، ويشمل ذلك أيضًا "البيانات الضخمة Big Data" التي تُعتبر غالبًا فرعًا من علم البيانات، ويمكن تطبيق أساليب علم البيانات وتقنياته وإجراءاته على أي عينة من البيانات، سواء كانت محدودة أو ضخمة، وتتميز البيانات الضخمة عن البيانات المحدودة بالزيادة الهائلة في الحجم(Volume) المتاح إلكترونيا نتيجة ضخ المجالات المعرفية المختلفة لها وحاجتها لسعات تخزينية كبيرة، وأيضا سرعة (Velocity) تراكمها وتحديثها وتغيرها، كما تتميز البيانات الضخمة بتنوعها (Variety) ما بين أعداد ونصوص وصور وفيديوهات ومواقع جغرافية ... وأيضا مصداقيتها(Veracity) العالية ودرجة الثقة في نتائجها، علاوة على قيمتها (Value)الكبيرة في الإضافة للمعرفة والتوصل إلى رؤى واقعية واتخاذ قرارات صائبة.

وقد وصف Nawaz (2020, P. 9) علم البيانات بأنه "مجال بيني التخصصات يتعلق بالمعارف في مجالات البيانات الضخمة، والرياضيات، والإحصاء، والبرمجة، والتنقيب عن البيانات، والتعلم الآلي، ويطبق فيه أساليب التعلم الآلي والخوارزميات والطرق الإحصائية؛ لعمل تنبؤات من البيانات".

وعرف كل من(2013, P.78) Waller وFawcett علم البيانات بأنه "العلم الذي يستخدم أساليب وإستراتيجيات كمية ونوعية للتعامل مع البيانات الضخمة؛ بهدف حل المشكلات ذات الصلة، والتنبؤ بالنتائج".

كما عرفه Baumer بأنه ذلك "العلم الذي يجمع بين علوم الرياضيات والإحصاء والكمبيوتر وأحد العلوم التطبيقية؛ بغرض التوصل لمعلومات ذات مغزى من البيانات الضخمة التي يمكن جمعها من مصادر متنوعة "(Baumer, 2015, P.1).

على الرغم من تفاوت التعريفات التي قدمت حول علم البيانات، والتي تختلف وفقًا لوجهات نظر أصحابها وتخصصاتهم، سواء كانوا من تخصصات الإحصاء أو الرياضيات أو علوم الكمبيوتر، أو من ناحية المجال أو السياق المعرفي المطبق فيه علم البيانات، إلا أن هناك شبه إجماع على أن علم البيانات هو مجال جديد وعصري بيني التخصصات، ويقوم علم البيانات على استخراج الأنماط من البيانات التي تشكل المستوى القاعدي من الهرم المعرفي (DIKW)، الذي يبدأ بجمع البيانات (Data) ومعالجتها بتطبيق الأساليب المبتكرة في التعامل مع البيانات، ثم التوصل إلى المعلومات (Information) منها، ثم اكتساب المعرفة (Knowledge)، وأخيرا اكتساب الحكمة (Wisdom) من خلال تلك المعرفة والتي تعد أعلى قمة هذا الهرم.

ومما تقدم يمكن ووضع تعريف إجرائي لعلم البيانات: بأنه ذلك العلم الذي يطبق فيه الطالب الأساليب المناسبة للتعامل مع البيانات الضخمة؛ بهدف فهم الظواهر المرتبطة بتلك البيانات، ووصفها، وتفسيرها، واكتشاف الرؤى منها، وتوليد المعرفة ونشرها، واكتساب الحكمة في اتخاذ القرارات المناسبة لحل المشكلات ذات الصلة بها.

2. طبيعة علم البيانات

شهدت العقود الستة الأخيرة ثورة كبيرة في العلم حيث ظهر ما يسمى بالعلوم العصرية التي نمت بتقدم علوم الكمبيوتر، وأساليبه، وتطبيقاته في الرسوم والنمذجة، ويعد علم البيانات مثالا للعلوم العصرية الذي تعكس طبيعة الفكر الذي أسهم في نموها، وهي طبيعة نصف عملية إنسانية تطبيقية.

وترى نظله خضر أن "من يتبنى هذا الفكر يريد أن يقول لنا ما المعنى لعمل العلم؟ ويرى أن العلم متغير أي يتغير مع الزمن، وإنساني أي أنه ما نقوم بعمله نحن بنو البشر، وسياسي أي أنه دالة للمكان وللحضارة والثقافة المحلية، واجتماعي أي يتفاعل الأفراد مع بعضهم البعض في عمله، وما دام العلم من صنع الإنسان، فالإنسان يخطئ، ولكنه يصحح من أخطائه" (نظله خضر، 2020، ص. 21).

ويستمد علم البيانات هذه الطبيعة من كونه علمًا تطبيقيًا عصريًا مبنيًا على النموذج الرابع للعلم المتعلق بالاكتشافات العلمية التي تم التوصل إليها من خلال عملية استكشاف البيانات الضخمة التي تقوم بشكل أساسي على الوصول إلى سيل كبير من البيانات المتدفقة التي يتم جمعها بالأدوات التكنولوجية أو توليدها باستخدام

المستشعرات، ثم معالجتها، وإدارتها، وتخزينها، وتحليلها آليا. وقد عزز هذا النموذج الجديد النماذج العلمية الثلاثة السابقة، ويظهر في الشكل التالي وصف للنماذج العلمية والتي تتضمن: النموذج الأول للعلم (التجريبي) والذي ظهر منذ ألف عام ويعتمد على وصف الظواهر الطبيعية من خلال التجارب، والنموذج الثاني للعلم (النظري) الذي ظهر منذ مئات السنين ويعتمد على استخدام النماذج والتعميمات، والنموذج الثالث للعلم (الحاسوبي) والذي ظهر في العقود القليلة الماضية ويعتمد على نمذجة الظواهر المعقدة، وأخيرًا النموذج الرابع للعلم (الحالي) ويعتمد على جمع البيانات باستخدام الأدوات التكنولوجية وتوليدها بالمستشعرات، وتخزين البيانات والمعلومات والمعرفة في أجهزة الكمبيوتر سواء على ملفات أو قواعد البيانات، وتحليلها، استكشاف البيانات بالأساليب المناسبة، وقد جمع النموذج الرابع للعلم بين التجريب والتنظير والمحاكاة (Hey, Tansley & Tolle, 2009).

وعلى الجانب الآخر، تنبع طبيعة علم البيانات من طبيعته كعلم بيني التخصصات يضم المعارف والمهارات في علوم الرياضيات والإحصاء والاحتمالات والكمبيوتر والبرمجة والذكاء الاصطناعي، وأيضا في مجالات أو سياقات الأعمال المختلفة– التجارية، والصناعية، والفنية...– والتي ستتم معالجة بياناتها (Escoufier & Hayashi, 1995; Baumer, 2015, P.1; Nawaz, 2020, P. 9).

وتعقيبًا على ذلك، تلعب الرياضيات دورًا جوهريًا في علم البيانات فهي تُتيح لعلماء البيانات فهمَ البيانات والتعامل معها بفعالية من خلال توظيف الأدوات والأساليب الرياضياتية المُناسبة من مختلف فروع الرياضيات، مثل: الجبر الخطي والرياضيات المتقطعة وحساب التفاضل والتكامل. أما من ناحية دور كل من الإحصاء والاحتمالات في علم البيانات، يبدأ علماء البيانات مهامهم بجمع كميات هائلة من البيانات، ويُوظّف علماء البيانات مهاراتهم الإحصائية لاختيار عينات مُمثلة للمجتمع، وتصميم التجارب الإحصائية، وتقييم احتمالات الأحداث، وتحديد مقدار الخطأ المتعلق بقرار معين ونوعه، واستكشاف خصائص البيانات من خلال حساب مقاييس النزعة المركزية والتشتت والانحراف والتفرطح، والكشف عن الأنماط والاتجاهات، وقياس العلاقات بين المتغيرات، وعرض النتائج بالأشكال والمخططات البيانية المناسبة، وعلى صعيد الذكاء الاصطناعي، يُستخدم تعلم الآلة لبناء نماذج ذكية قادرة على التنبؤ والتصنيف، بينما يُوظّف التعلم العميق لفهم البيانات المُعقدة واستخلاص رؤى جديدة. وتُستخدم تقنيات الذكاء الاصطناعي أيضًا لقياس دقة نماذج التعلم الآلة وتحسينها باستمرار. وأخيرًا، تُستخدم الأساليب الإحصائية لاختبار الفرضيات وتحديد العلاقات ذات الدلالة الإحصائية، مما يُؤكد على صحة النتائج ويُعزّز موثوقيتها. ومن جانب علوم الكمبيوتر والبرمجة، يقوم علماء البيانات باستخدام الأدوات والتقنيات والبرامج ولغات البرمجة المناسبة، واستخدام قواعد البيانات والأكواد البرمجية للتعامل

بشكل سلس مع البيانات الضخمة التي يصعب معالجتها بالطرق العادية. وأخيرا من جانب السياق الذي يطبق فيه علم البيانات، كما يحتاج علماء البيانات إلى فهم عميق للمجال أو السياق الذي سيعملون فيه من أجل التطبيق الفعال لمبادئ وتقنيات علم البيانات. كما تؤدي خبرة السياق أيضا إلى فهم علماء البيانات لطبيعة البيانات، ومعناها، وكيفية التعامل معها، واستخلاص رؤى ذات مغزى منها، وعادة ما يتم اكتساب خبرة في تطبيق عمليات علم البيانات في مجال أو سياق ما بعد عدد سنوات كافي من البحث والدراسة فيه.

3. مهارات علم البيانات

يُعدّ علم البيانات مجالًا واسعًا يتطلب مزيجًا فريدًا من المهارات التقنية التخصصية والمهارات الشخصية الإنسانية، فمن الناحية التقنية التخصصية، لا بد من امتلاك علماء البيانات مهارات وضع خطة العمل لتحديد أهداف تحليل البيانات ووضع استراتيجية فعّالة لتحقيقها، ومهارات إدارة البيانات من خلال جمع البيانات من مختلف المصادر ، وتنظيفها، وتحويلها، وتخزينها، واسترجاعها بكفاءة، ومهارات تحليل البيانات من خلال تطبيق تقنيات إحصائية متقدمة لفهم أنماط البيانات واستخراج المعرفة منها، ومهارات نمذجة البيانات وتنقيبها من خلال بناء نماذج تعلم آلي وتعلم عميق للتنبؤ بالنتائج واتخاذ القرارات وتقييم دقة تلك النماذج وفعاليتها وتحسينها باستمرار، ومهارات تصور البيانات من خلال تحويل البيانات إلى تنسيقات وأشكال ومخططات بيانية مرئية سهلة الفهم للتواصل الفعّال مع الآخرين. وأيضًا لا بد من امتلاك علماء البيانات لمجموعة من المهارات الشخصية الإنسانية، تشمل مهارات التركيز والانتباه من خلال القدرة على التركيز على التفاصيل الدقيقة وتحليل البيانات بدقة، ومهارات التفكير النقدي من خلال تقييم المعلومات وتحليلها بشكل موضوعي واستخلاص النتائج المنطقية، ومهارات الابتكار من خلال إيجاد حلول جديدة ومبتكرة لمشكلات تحليل البيانات المعقدة، ومهارات حل المشكلات من خلال القدرة على تحديد المشكلات وتحليلها وتطوير حلول فعّالة لها، ومهارات التعلم الذاتي من خلال مواكبة التطورات المتسارعة في مجال علم البيانات من خلال التعلم المستمر، ومهارات التواصل والتعاون من خلال العمل بفعالية مع فرق العمل وتقديم نتائج تحليلات البيانات بشكل واضح وموجز. إنّ مزيجًا من هذه المهارات يُمكّن علماء البيانات من تحويل البيانات إلى معلومات قيّمة تُساهم في اتخاذ قرارات استراتيجية فعّالة تُعزّز نجاح المؤسسات في مختلف المجالات.

وبالاطلاع على الأدبيات ذات الصلة بمهارات علم البيانات، بما في ذلك المصادر الأكاديمية والدراسات السابقة وما توصلت إليه من نتائج في هذا المجال، نجد أن المتخصصون في المجال اقترحوا نظريًا العديد من المهارات التي ترتبط بمنهجية علم البيانات وعملياته، وتمحورت هذه المهارات حول إجراءات التعامل مع البيانات، وفهم خصائصها، وتعقيداتها.

حيث أشار Cao (2018, P. 303) إلى أن مهارات علم البيانات تتضمن مجموعة من المهارات المترابطة، وهي: مهارات التفكير في البيانات كمهارات التفكير الابتكاري والنقدي والتخيلي والاستنتاجي، والمهارات التقنية لجمع البيانات، وتخزينها في مستودعات البيانات، واسترجاعها، وإدارتها، ونمذجتها والتطبيق الآلي لتلك العمليات، وأيضا مهارات التواصل وعرض البيانات، والمهارات القيادية كالتخطيط الإستراتيجي واتخاذ القرارات.

بينما ذكر Cielen et al. (2016) أن الجوانب المعرفية والأدائية لمهارات علم البيانات تشمل مهارات تحديد هدف مشروع علم البيانات، ومهارات استرجاع البيانات من المصادر المتنوعة ومهارات عرضها ونمذجتها، وتكرار تلك العمليات بشكل آلي ومستمر.

ويرى كل من Deshpande (2019) و Kotu أن مهارات علم البيانات تتضمن الجوانب المعرفية والأدائية لمهارات فهم مجال العمل، ومهارات تحديد عينة البيانات، ومهارات فهم البيانات وجمعها وإعدادها ونمذجتها، وتقويم ما تم القيام به من عمليات، ثم التطبيق الآلي لها.

وأشار Nawaz (2020) إلى أن تلك المهارات تشمل مهارات الحصول على البيانات، ومهارات إعداد البيانات وتحليلها ونمذجتها وتقويمها، وتفسير النتائج، وأخيرا إعداد التقارير.

أما على الجانب الإجرائي، فقد اختلف الباحثون في تحديد مهارات علم البيانات، ويعزى ذلك لتباين الأهداف التي يسعون لتحقيقها من خلال الدراسة والبحث.

فقد حددها كل من Heckman (2016) و Saltz في ثلاث مهارات أساسية، هي: إعداد البيانات وتحليلها، وعرض النتائج.

وتوصل Baumer (2015) إلى مجموعة من المهارات تتضمن الجوانب المعرفية والأدائية لمهارات: مهارة طرح سؤال حول مجموعة من البيانات، ومهارات تحديد مصادر للبيانات وجمعها وإدارتها ومعالجتها والاستعلام عنها وتحليلها وتصورها، والتنقيب في البيانات Data Mining باستخدام الأدوات المناسبة، وكذلك التعبير عن النتائج بطرق متنوعة.

بينما حددها Wang (2018, P. 131) في مهارات تخزين البيانات، ومعالجتها، وتحليلها.

وتوصل كل من Davis (2019) و Yan إلى مجموعة من المهارات شملت مهارات: جمع البيانات، صياغة الأسئلة المرتبطة بالمتغيرات والبيانات، التحليل الاستكشافي للبيانات، تصور البيانات والرسوم البيانية، تحويل البيانات وهندستها، نمذجة البيانات، اختبار الفرضيات.

ومن خلال ما تقدم تم وضع تعريف إجرائي لمهارات علم البيانات بأنها: تلك المهارات التي يمارسها الطالب، وتتضمن الجوانب المعرفية والأدائية لمهارات: وضع خطة العمل، وجمع البيانات، ومعالجتها، وتحليلها، وتنقيبها، وعرض النتائج بهدف التواصل مع الآخرين بشأن تلك النتائج وتصورها.

ويتضمن جدول (1) قائمة مهارات علم البيانات وجوانبها المعرفية والأدائية كالتالي:

جدول(1) قائمة مهارات علم البيانات وجوانبها المعرفية والأدائية

المهارة	الجانب المعرفي	الجانب الأدائي
وضع خطة العمل	• التعرف على طبيعة البيانات • التعرف على طبيعة علم البيانات • التعرف على الأدوات الرقمية لعلم البيانات • التعرف على خطة العمل في مشاريع علم البيانات	
جمع البيانات	• التعرف على مهارة جمع البيانات • التعرف على مصادر الحصول على البيانات	• الحصول على البيانات
معالجة البيانات	• التعرف على مهارة معالجة البيانات • تطبيع البيانات بالطرق الرياضياتية	• تنظيف البيانات • تحويل البيانات
تحليل البيانات	• التعرف على مهارة تحليل البيانات • التعرف على التحليل الاستكشافي للبيانات • حساب مقاييس النزعة المركزية والتشتت للبيانات والعلاقات بين المتغيرات بالطرق الرياضياتية • التعرف على التحليل المرئي للبيانات	• التحليل الاستكشافي للبيانات • التحليل المرئي للبيانات
تنقيب البيانات	• التعرف على مهارة تنقيب البيانات • التعرف على تصنيف البيانات باستخدام نماذج تعلم الآلة • التعرف على تجميع البيانات باستخدام نماذج تعلم الآلة	• تصنيف البيانات باستخدام نماذج التعلم الآلي • تجميع البيانات باستخدام نماذج التعلم الآلي
عرض النتائج	• التعرف على مهارة عرض النتائج • كتابة تقرير عرض النتائج	• إنشاء لوحة التحكم Dashboard لتقرير عرض النتائج

وفيما يلي عرض لتلك المهارات وجوانبها المعرفية والأدائية:

(Baumer, 2015; Cielen et al., 2016; Saltz & Heckman, 2016; Wang, 2018, P. 131; Cao, 2018, P. 303; Kotu & Deshpande, 2019; Yan & Davis, 2019; Nawaz, 2020).

- **وضع خطة العمل** Developing action plan

وهي قدرة الطالب على تحديد الخطوات والأهداف والموارد اللازمة لإنجاز مشروع علم البيانات بنجاح، وتتضمن مهارة وضع خطة العمل الجوانب المعرفية التالية:

أ. التعرف على طبيعة البيانات

ب. التعرف على طبيعة علم البيانات

ج.التعرف على الأدوات الرقمية لعلم البيانات

د. التعرف على خطة العمل في مشاريع علم البيانات

- **جمع البيانات** Data Collection

وهي قدرة الطالب في الحصول على البيانات الخام المرتبطة بالمشكلة المطروحة من المصادر المختلفة المتاحة، وتتضمن مهارة جمع البيانات الجانبين المعرفين التاليين:

أ- التعرف على مهارة جمع البيانات

ب- التعرف على مصادر الحصول على البيانات

كما تتضمن مهارة جمع البيانات الجانب الأدائي التالي:

الحصول على البيانات

- **معالجة البيانات** Presenting Results

وهي قدرة الطالب على التعامل مع البيانات الخام؛ لتجهيزها للتحليل والنمذجة؛ وللحصول منها على معلومات ومعرفة غير مضللة. وتتضمن مهارة معالجة البيانات الجانبين المعرفين التاليين:

أ. التعرف على مهارة معالجة البيانات

ب. تطبيع البيانات بالطرق الرياضياتية

كما تتضمن مهارة معالجة البيانات الجانبين الأداءين التاليين:

أ. تنظيف البيانات

ب. تحويل البيانات

- تحليل البيانات Data Analysis

وهي قدرة الطالب على استخدام طرق التحليل الاستكشافي Exploratory Data Analysis بحساب مقاييس الإحصاء الوصفي للبيانات، والتي تتضمن: مقاييس النزعة المركزية والتشتت، وحساب معامل الارتباط بين المتغيرات، وأيضًا باستخدام طرق التحليل المرئي للبيانات Visual Data Analysis بتصور تمثيلات البيانات Data Visualization بالمخططات والأشكال البيانية؛ لوصف توزيع البيانات وتشتتها، وتحديد العلاقة بين المتغيرات.

وتتضمن مهارة تحليل البيانات الجوانب المعرفية التالية:

أ. التعرف على مهارة تحليل البيانات

ب. التعرف على التحليل الاستكشافي للبيانات

ج. حساب مقاييس النزعة المركزية والتشتت للبيانات والعلاقات بين المتغيرات بالطرق الرياضياتية

د. التعرف على التحليل المرئي للبيانات

كما تتضمن مهارة تحليل البيانات الجانبين الأداءين التاليين:

أ. التحليل الاستكشافي للبيانات

ب. التحليل المرئي للبيانات

- **تنقيب البيانات** Data Mining

وهي قدرة الطالب على تحديد الأنماط واستخراج المعرفة من البيانات من خلال تطبيق النموذج المناسب للمشكلة المطروحة من نماذج تعلم الآلة سواء بتصنيف البيانات أو تجميعها، وتقويم فاعلية النموذج المطبق.

وتضمن مهارة تنقيب البيانات الجوانب المعرفية التالية:

أ. التعرف على مهارة تنقيب البيانات

ب. التعرف على تصنيف البيانات باستخدام نماذج تعلم الآلة

ج. التعرف على تجميع البيانات باستخدام نماذج تعلم الآلة

كما تتضمن مهارة تحليل البيانات الجانبين الأداءين التاليين:

أ. تصنيف البيانات باستخدام نماذج التعلم الآلي

ب. تجميع البيانات باستخدام نماذج التعلم الآلي

- **عرض النتائج** Results Presentation

وتعرف بأنها قدرة الطالب على عرض النتائج من خلال إنشاء تقرير يتضمن: كتابة مقدمة تقرير عرض النتائج، وإنشاء لوحة تحكم Dashboard تحوي مجموعة من المخططات والأشكال البيانية المناسبة لطبيعة

البيانات والنتائج، وأخيرًا كتابة خاتمة تقرير عرض النتائج؛ وذلك بهدف التوصل مع الآخرين بشأن تلك النتائج وتصورها. وتتضمن مهارة عرض النتائج الجوانب المعرفية التالية:

أ. التعرف على مهارة عرض النتائج

ب. كتابة تقرير عرض النتائج

كما تتضمن مهارة عرض النتائج الجانب الأدائي التالي:

إنشاء لوحة التحكم Dashboard لتقرير عرض النتائج.

4. علم البيانات وتطوير مناهج الإحصاء

سريعا ما حلت الوظائف والأعمال المرتبطة بعلم البيانات في المرتبة الثالثة على مستوى الوظائف الواعدة في تقرير مستقبل الوظائف لعام (2020) والذي ينشره المنتدى الاقتصادي العالمي (W.E.F., 2020, p. 62).

الأمر الذي يلقي على عاتق المؤسسات التعليمية ضرورة التجاوب مع معطيات ذلك التحول بتطوير مناهجها. وتعد الجامعة الكيان الذي يقع على عاتقه المسؤولية الأكبر في إعداد الخريجين المؤهلين بالمعارف والمهارات لشغل الوظائف والأعمال العصرية، والمساهمة في تحقيق اقتصاد تنافسي يستخدم البيانات الضخمة لتحقيق التنمية في جميع الجوانب الاقتصادية، والاجتماعية...

لذا شهدت مناهج الإحصاء بالمرحلة الجامعية في العديد من دول العالم تطورات ملموسة، وتغيرات سريعة في الآونة الأخيرة، حيث قامت هذه الدول بإعادة النظر في محتوى مناهج الإحصاء لديها لتأتي منسجمة مع حاجات مجتمعاتها وتطلعات أفرادها، ولملاحقة التطورات المتسارعة في العلم وتطبيقاته، وفي السير قدما نحو الرقي والتقدم، والتحول نحو الاقتصاد المبني على البيانات الضخمة والمعرفة، "وأيضا في حل مشكلات عصرية في عالم البيانات، والمعلومات، والتكنولوجيا"(نظله خضر، 2020، ص. 13)، "وتتسم مشكلات هذا العصر بالتعقيد، والتداخل بين فروع المعرفة المختلفة، الأمر الذي يتطلب ضرورة التخلي عن النظرة الخطية في إعداد المقررات بالمرحلة الجامعية، وبناء برامج تدعم اتجاه التكامل بين مختلف فروع المعرفة المختلفة"(فايز مينا، 2013 ، ص. 18) كعلم البيانات، وبتنمية الفكر المعاصر الذي ابتكارها.

وعليه، أقدمت بعض مؤسسات التعليم الجامعي حول العالم على استحداث تخصص منفرد لدراسة علم البيانات، أو وضع مقرر له ضمن الخطة الدراسية في التخصصات المختلفة، أو إتاحة برنامج حر تقدمه؛ بهدف التطوير العلمي والمهني(Press, 2013; Ceccucci et al., 2015).

الأمر الذي يستدعي وضع برنامج يهدف إلى تدريس علم البيانات لطلاب المرحلة الجامعية الأولى بمملكة البحرين وتنمية مهاراته، بحيث يناسب مستوى طلاب التخصصات المختلفة، واختيار أفضل المداخل لتدريسه.

ثانيًا: الدراسات السابقة المرتبطة بمهارات علم البيانات

لأهداف هذه الدراسة سيتم عرض أهم الدراسات التي حاولت تحديد مهارات علم البيانات الواجب توافرها في المراحل التعليمية المختلفة، ومن هذه الدراسات:

1. دراسة (Baumer, 2015)

هدفت إلى بناء مقرر لتنمية مهارات علم البيانات لدى طلاب المرحلة الجامعية بكلية سميث بولاية ماساتشوستس الأمريكية.

وحدد الباحث مجموعة من المهارات لعلم البيانات تضمنت مهارات: طرح سؤال حول مجموعة من البيانات، وجمعها، وإدارتها، ومعالجتها، والاستعلام عنها، وتحليلها، وتصورها، وكذلك التعبير عن النتائج بطرق مختلفة.

وأوصى الباحث بضرورة بناء مقررات في علم البيانات تناسب الطلاب في التخصصات المختلفة؛ لتنمية مهاراتهم في التعامل مع البيانات...

2. دراسة (Yan & Davis, 2019)

هدفت إلى تصميم مقرر أساسيات دورة عمل علم البيانات لحل مشكلات حياتية واقعية، وفاعلية تدريسه باستخدام التعلم النشط لطلاب البكالوريوس في علم البيانات، تكونت عينة الدراسة من جميع طلاب الفرقة الأولى بتخصص علم البيانات في جامعة ماساتشوستس الأمريكية، وتوصلت الدراسة إلى فاعلية التعلم النشط في تدريس مقدمة علم البيانات لطلاب الفرقة الأولى.

3. دراسة (Schwab-McCoy et al., 2020)

هدفت إلى مسح دورات وبرامج علم البيانات في مرحلة التعليم العالي للتعرف على محتوى مناهجها، والمصادر والمواد والأدوات الرقمية المستخدمة فيها، والتحديات التي واجهات أعضاء هيئة التدريس والطلاب داخل القاعات الدراسية في تعليم وتعلم علم البيانات.

وشملت عينة الدارسة (69) مؤسسة من تخصصات أكاديمية مختلفة من مؤسسات التعليم العالي بالولايات المتحدة الأمريكية، وقد اتبع في الدراسة المنهج الكيفي من خلال التصميم الوصفي التحليلي لتوصيف الدورات والبرامج التي شملتها الدراسة، علاوة على تطبيق استبانة أعضاء هيئة التدريس.

وقد أشارت النتائج إلى أن أكثر موضوعات علم البيانات تكرارا كانت على الترتيب تنازليا: تصور البيانات، معالجة البيانات، أخلاقيات علم البيانات، إدارة البيانات، الطرق الرياضياتية والإحصائية لعلم البيانات، عرض النتائج، تحليل البيانات، هندسة البيانات، تعلم الآلة. وكانت لغتي البرمجة R و Python على قائمة الأدوات الرقمية المستخدمة.

4. دراسة (Çetinkaya-Rundel & Ellison, 2021)

هدفت إلى بناء مقرر اختياري لمبادئ علم البيانات لطلاب المرحلة الجامعية الأولى في التخصصات العلمية والإنسانية والاجتماعية بجامعة ديوك الأمريكية.

واعتمد الباحثان في بناء المقرر على مراجعة: مقرر Foundations of Data Science من جامعة كاليفورنيا، ومقرر Foundations of Data Science من جامعة كامبريدج، ومقرر Introduction to Data Science من كلية سميث، ومقرر Data Science 101 من جامعة ستانفورد.

وتوصلت النتائج إلى فاعلية طرق التدريس المستخدمة في المقرر، وأظهرت نتائج التقويمات تحقيق الطلاب لأهداف المقرر بنسب مرتفعة؛ مما أدي لتصميم مقرر اختياري متقدم مكمل لهذا المقرر، وأشار الباحثان إلى أن من أبرز التحديات التي واجهتهما كان اختيار الموضوعات الواجب توافرها في وحدة استخلاص النتائج لتنوع الطرق التي يمكن تضمينها؛ مما أدى للتركيز على طريقة اختبار الفرضيات باستخدام لغة R دون تقديم الكثير من الطرق الإحصائية المرتبطة.

التعليق على الدراسات المتعلقة بعلم البيانات وتنمية مهاراته

من خلال عرض الدراسات السابقة التي تناولها هذا المحور يتضح ما يلي:

1. تناولت الدراسات السابقة بعض الموضوعات المرتبطة بمفاهيم علم البيانات ومهاراته مثل:
   أ. تضمنت دراسة (Baumer, 2015) مهارات طرح سؤال حول البيانات، وجمعها، وإدارتها، ومعالجتها، والاستعلام عنها، وتحليلها، وتصورها.
   ب. تضمنت دراسة (Yan & Davis, 2019) موضوعات دورة عمل علم البيانات: صياغة الأسئلة المرتبطة بالمتغيرات والبيانات، وجمع البيانات، والتحليل الاستكشافي لها، ومقاييس الإحصاء الوصفي، وتصور البيانات، وتحويلها وهندستها، وإنشاء نموذجي التجميع والانحدار البسيط، وضع الفرضيات واختبارها.
2. أوصت معظم الدراسات السابقة بتضمين موضوعات علم البيانات في مراحل التعليم الجامعي في التخصصات المختلفة بمستويات ملائمة للدرجة العلمية والتخصص الأكاديمي، وإجراء برامج تدريبية بعد التخرج لتطبيق مفاهيم علم البيانات ومهاراته لتطوير العمل داخل المؤسسات المختلفة، واستخدام

المداخل ولإستراتيجيات والطرق المناسبة لتقديمها لما له من أهمية تطبيقية ودلالة في الحياة المعاصرة ومتطلبات المستقبل مما يثري العملية التعليمية.

وقد استفاد الباحث من هذه الدراسات في:

1. تحديد الأساسيات الخاصة بعلم البيانات المناسبة لمستوى طلاب المرحلة الجامعية الأولى في التخصصات المختلفة.
2. عرض أساس نظري لعلم البيانات وأهميته لتطوير مناهج الإحصاء بالمرحلة الجامعية.
3. التعرف على بعض المهارات المرتبطة بعلم البيانات التي يمكن الاستفادة منها في تحديد مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية الأولى في التخصصات المختلفة والخاصة بهذه الدراسة.

أما عن علاقة هذه الدراسة بالدراسات التي تم عرضها نجد أنه يتفق معها من حيث:

محاولة التعرف على مهارات علم البيانات الواجب توافرها لدى الطلاب في المرحلة الجامعية.

وتختلف هذ الدراسة مع الدراسات التي تم عرضها في:

1. التعرف على التعرف على مهارات علم البيانات الواجب توافرها لدى الطلاب في المرحلة الجامعية في البيئة العربية.
2. لم تتناول أي من الدراسات السابقة مهارات علم البيانات الخاصة بهذه الدراسة، وهي: وضع خطة العمل، وجمع البيانات ومعالجتها، وتحليلها، وتنقيبها، وعرض النتائج.

الإجراءات المنهجية للدراسة والنتائج:

أولًا: منهج الدراسة:

اتبعت الدراسة المنهج الوصفي المسحي، لمناسبته من حيث الهدف في تحديد مهارات علم البيانات (وجوانبها المعرفية والأدائية) الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة.

ثانيًا: أداة الدراسة:

لما كان الهدف من هذه الدراسة هو تحديد مهارات علم البيانات (وجوانبها المعرفية والأدائية) الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة؛ لذلك قام الباحث بإعداد قائمة مهارات علم البيانات الواجب توافرها لدى طلاب المرحلة الجامعية في التخصصات المختلفة:

1. مصادر تحديد مهارات علم البيانات

أطلع الباحث على العديد من الأدبيات وما تضمنتها من مصادر علمية في مجال علم البيانات، بالإضافة إلى الدراسات السابقة المرتبطة بهذا المجال وما توصلت إليه من نتائج بهدف وضع قائمة أولية تتضمن مهارات علم البيانات في صورة إجرائية.

2. التحقق من مناسبة تلك المهارات لمستوى طلاب المرحلة الجامعية بالتخصصات المختلفة

بعد وضع القائمة الأولية لمهارات علم البيانات، قام الباحث بإعداد استبانة لاستطلاع رأي الخبراء والمتخصصين في مجال إعداد برامج علم البيانات وتدريسه عن مناسبة المهارات التي تم تحديدها لطلاب المرحلة الجامعية بالتخصصات المختلفة. وتم عرض الاستبانة على مجموعة من المحكمين، وقد أبدت الرأي من حيث:

- مدى ارتباط المهارات التي تم تحديدها(بجوانبها المعرفية والأدائية) بعلم البيانات.
- مدى مناسبتها لمستوى الطلاب بالمرحلة الجامعية بالتخصصات المختلفة.
- إضافة أي اقتراحات أخرى.

وفي ضوء مقترحات وآراء المحكمين، قام الباحث بحذف بعض المهارات، ودمج البعض الآخر، وبإعادة صياغة عدد منها؛ حتى توصل إلى الصورة النهائية لقائمة مهارات علم البيانات المناسبة لمستوى طلاب المرحلة الجامعية في التخصصات المختلفة.

ويوضح جدول (2) مهارات علم البيانات (وجوانبها المعرفية والأدائية) والنسب المئوية للاتفاق بين المحكمين على تلك المهارات (وجوانبها) بعد إعادة عرضها عليهم كما يلي:

جدول(2) نسب الاتفاق بين المحكمين في تحديد مهارات علم البيانات
(وجوانبها المعرفية والأدائية)

المهارة (وجوانبها المعرفية والأدائية)	النسبة المئوية للاتفاق
1) وضع خطة العمل	100%
الجوانب المعرفية:	
أ. التعرف على طبيعة البيانات	100%
ب. التعرف على طبيعة علم البيانات	85.71%
ج. التعرف على الأدوات الرقمية لعلم البيانات	100%

المهارة (وجوانبها المعرفية والأدائية)	النسبة المئوية للاتفاق
د. التعرف على خطة العمل في مشاريع علم البيانات	100%
2) جمع البيانات	100%
الجوانب المعرفية: أ. التعرف على مهارة جمع البيانات ب. التعرف على مصادر الحصول على البيانات الجوانب الأدائية: أ. الحصول على البيانات	 73.43% 100% 100%
3) معالجة البيانات	100%
الجوانب المعرفية: أ. التعرف على مهارة معالجة البيانات ب. تطبيع البيانات بالطرق الرياضياتية الجوانب الأدائية: أ. تنظيف البيانات ب. تحويل البيانات	 71.43% 100% 100% 100%
4) تحليل البيانات	100%
الجوانب المعرفية: أ. التعرف على مهارة تحليل البيانات ب. التعرف على مفاهيم التحليل الاستكشافي للبيانات ج. حساب مقاييس النزعة المركزية والتشتت للبيانات والعلاقات بين المتغيرات بالطرق الرياضياتية الجوانب الأدائية: أ. التحليل الاستكشافي للبيانات ب. التحليل المرئي للبيانات	 71.43% 71.43% 100% 100% 100%

المهارة (وجوانبها المعرفية والأدائية)	النسبة المئوية للاتفاق
5) تنقيب البيانات	100%
الجوانب المعرفية: أ. التعرف على مهارة تنقيب البيانات ب. التعرف على تصنيف البيانات باستخدام نماذج تعلم الآلة ج. التعرف على تجميع البيانات باستخدام نماذج تعلم الآلة الجوانب الأدائية: أ. تصنيف البيانات باستخدام نماذج التعلم الآلي ب. تجميع البيانات باستخدام نماذج التعلم الآلي	 71.43% 100% 100% 100% 100%
6) عرض النتائج	100%
الجوانب المعرفية: أ. التعرف على مهارة عرض النتائج ب. كتابة تقرير عرض النتائج الجوانب الأدائية: إنشاء لوحة التحكم Dashboard لتقرير عرض النتائج	 71.43% 100% 100%

التوصيات والمقترحات:

أولًا: التوصيات

يمكن تقديم بعض التوصيات في ضوء النتائج التي تم التوصل إليها:

1. إعداد مقررًا في علم البيانات لتنمية مهارات علم البيانات طلاب المرحلة الجامعية.
2. إثراء مقررات الإحصاء القائمة في المرحلة الجامعية بمهارات علم البيانات.
3. توفير المواد والأدوات والبرامج اللازمة لتنمية مهارات علم البيانات لدى الطلاب.
4. تنظيم مؤتمرات وورش عمل لأعضاء هيئة التدريس حول كيفية تطعيم مقررات الإحصاء القائمة بأساسيات علم البيانات ومهاراته.

ثانيًا: المقترحات:

في ضوء نتائج الدراسة وتوصياتها، يمكن اقتراح بعض الدراسات العلمية التي يتوقع أن تثري موضوع الدراسة وتستكمل الجهود البحثية في هذا المجال:

1. تحليل محتوى مناهج المرحلة الجامعية في التخصصات المختلفة الدراسية الحالية لتقييم مدى تضمينها لمهارات علم البيانات.
2. تحديد الفجوات بين المهارات المطلوبة في سوق العمل والمهارات التي يتم تدريسها للطلاب.
3. تطوير مقرر مبادئ الإحصاء في المناهج الدراسية الجامعية في ضوء مهارات علم البيانات.
4. مقارنة فاعلية أساليب التدريس المختلفة لتنمية مهارات علم البيانات للطلاب من مختلف التخصصات.
5. أثر استخدام التعلم النشط والتعلم القائم على المشاريع لتنمية مهارات علم البيانات.
6. دراسة آراء أصحاب العمل حول قيمة مهارات علم البيانات لخريجي الجامعات من مختلف التخصصات.
7. تقييم فعالية استخدام أدوات ومنصات التعلم عبر الإنترنت لتعليم علم البيانات.
8. تطوير نماذج تعليمية هجينة تجمع بين التعلم التقليدي والتعلم عبر الإنترنت لتعليم علم البيانات بشكل فعال.
9. تحديد التحديات التي تواجه الجامعات في تعليم مهارات علم البيانات، مثل نقص الموارد البشرية المؤهلة والبنية التحتية التكنولوجية.

المراجع العربية والأجنبية

أولاً المراجع العربية :

- نظلة حسن خضر .(2020). تجديدات في أصول تدريس الرياضيات وفي أنشطة لصنع الرياضيات المتجددة. عالم الكتب.
- فايز مراد مينا. (2013). قضايا وآراء في البحث التربوي. مكتبة الأنجلو المصرية.

ثانياً المراجع الأجنبية :

- Alturkistani, A., Car, J., Majeed, A., Brindley, D., Wells, G. & Meinert, E. (2018). Determing The Effectiveness of a Massive Open Online Course in Data Science for Health. *International Conference E-Learning*, 27-34.
- Baumer, B. (2015): A Data Science Course for Undergraduates: Thinking with Data, The American Statistician, 69(4), 334-342, https://doi.org/10.1080/00031305.2015.1081105
- Brunner, R. & Kim, E. (2016). Teaching Data Science. The International Conference on Computational Science(ICCS 2016). 80,1947-1956.Elsevier. https://doi.org/10.1016/j.procs.2016.05.513
- Cao, L. (2018). Data Science thinking: The Next Scientific, Technological and Economic Revolution. Springer International Publishing AG. https://doi.org/10.1007/978-3-319-95092-1
- Ceccucci, W., Tamarkin, D., & Jones, K. (2015, July). The Effectiveness of Data Science as a means to achieve Proficiency in Scientific Literacy. Information Systems Education Journal, 13(4), 64-70.
- Çetinkaya-Rundel, M., & Ellison, V. (2021). A fresh look at introductory data science. Journal of Statistics and Data Science Education, 29(sup1), S16-S26.
- Cielen, D., Meysman, A. & Ali, M. (2016). Introducing data science: big data, machine learning, and more, using Python tools. Simon and Schuster.
- Escoufier, Y., & Hayashi, C. (Eds.). (1995). Data science and its applications. Academic Press/Harcourt Brace.
- Guzman, L. M., Pennell, M. W., & Nikelski, E. (2019). Successful Integration of Data Science in Undergraduate Biostatistics Courses Using Cognitive Load Theory. Life Sciences Education, 1-10.

- Hayashi, C., Yajima, K., Bock, H. H., Ohsumi, N., Tanaka, Y., & Baba, Y. (1996). Data Science, Classification, and Related Methods. Proceedings of the Fifth Conference of the International, 40-51.
- Kahn, J. (2019). Learning at the Intersection of Self and Society: The Family Geobiography as a Context for Data Science Education. Journal of the Learning Sciences, 1–24. https://doi.org/10.1080/10508406.2019.1693377
- Kotu, V., & Deshpande, B. (2019). Data science: concepts and practice.2nd edition. Morgan Kaufmann
- McCoy, C., & Shih, P. C. (2016). Teachers as Producers of Data Analytics: A Case Study of a Teacher-Focused Educational Data Science Program. Journal of Learning Analytics, 3(3), 193–214. http://dx.doi.org/10.18608/jla.2016.33.10
- Naur, p. (1974). Concise Survey of Computer Methods. Lund, Sweden: Studentlitteratur.
- Nawaz, M. W. (2020). Data Science Crash Course for Beginners: Fundamentals and Practices with Python, AI Publishing
- Piatetsky G., KDnuggets. (May 30, 2019), Python leads the 11 top Data Science, Machine Learning platforms: Trends and Analysis, https://www.kdnuggets.com/2019/05/poll-top-data-science-machine-learning-platforms.html
- Piatetsky, G.(May 30, 2019). Top Analytics, Data Science, Machine Learning Software 2017-2019. KDnuggets Annual Polls.
- Press, G. (2013, May 28). A very short history of data science. Forbes. https://www.forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data science/?sh=5b06337755cf
- Salas-Rueda, Ricardo-Adán, Érika-Patricia, & Rodrigo-David. (2020). Analysis and Design of the Web Game on Descriptive Statistics through the ADDIE Model, Data Science and Machine Learning. International Journal of Education in Mathematics, Science and Technology, 8(3), 245-260.
- Saltz, J., & Heckman, R. (2016). Big Data science education: A case study of a project-focused introductory course. Themes in Science & Technology Education, 8(2), 85-94.
- Sami, J., Sinclair, K., Stein, Z., & Medsker, L. (2020, January-May). Data Science Outreach Educational Program for High School Students Focused in Agriculture. Journal of STEM Education, 21(1), 18-33.

- Tukey, J. W. (1962), "The Future of Data Analysis," The Annals of Mathematical Statistics, 33, 1–67
- Wainer, H. (2015). Truth or Truthiness: Distinguishing Fact from Fiction by Learning to Think Like a Data Scientist. Cambridge: Cambridge University Press, MA.
- Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. Journal of Business Logistics, 34(2), 77-84.
- Wang, L. (2018). Twinning data science with information science in schools of library and information science. Journal of Documentation, 74(6), 1243 – 1257
- Wang, Y., & Li, J. (2019). The Curriculum System Construction of E-commerce Major in the Era of Intelligence. 4th Annual International Conference on Education Science and Education Management. 376, 128-134. Destech Publications.
- W.E.F., W. E. (2020). The Future of Jobs Report. WEF.
- West, J. (2018). Teaching data science: an objective approach to curriculum validation Computer Science Education. https://doi.org/10.1080/08993408.2018.1486120
- Wu, C. F. Jeff (1986). "Future directions of statistical research in China: a historical perspective" . Application of Statistics and Management. 1: 1–7. https://www2.isye.gatech.edu/~jeffwu/publications/fazhan.pdf
- Yan, D., & Davis, G. E. (2019). A First Course in Data Science. Journal of Statistics Education, 27(2), 99-109. https://doi.org/10.1080/10691898.2019.1623136