تعليم علم البيانات بأقسام ومدارس المكتبات والمعلومات بالجامعات الأمريكية : دراسة تحليلية

اعداد
د. مشيرة أحمد صالح
استاذ علم المكتبات والمعلومات المساعد
كلية الآداب – جامعة عين شمس
moshira444m@yahoo.com

ملخص البحث:

تهدف هذه الدراسة إلى التعرف على مفهوم علم البيانات والبيانات الضخمة ، كما تهدف إلى التعرف على خصائص المقررات الدراسية لعلم البيانات في عينة من أقسام ومدارس المكتبات والمعلومات بالجامعات الأمريكية خلال العام الجامعي ٢٠١٩/٢٠٢٠م ، باستخدام المنهج الوصفي التحليلي . وتوصلت الدراسة إلى عدة نتائج منها : يتوافر ٢١٣ مقرر خاص بعلم البيانات بنسبة ١٣,٦% من إجمالي المقررات المتاحة في مدارس المكتبات والمعلومات عينة الدراسة . وقد ختمت الدراسة بعدة توصيات ، أبرزها إعادة النظر في أهداف ومحتوي مقررات البرامج الحالية لأقسام المكتبات والمعلومات ، والعمل على توافقها مع متطلبات سوق العمل وخاصة في مجال علم البيانات .

الكلمات المفتاحية :

البيانات الضخمة ، علم البيانات ، عالم البيانات ، تدريس علوم المكتبات والمعلومات

تمهيد :

يشير لورايدس وايبرت Louridas and Ebert إلى أنه يتم إنتاج ما يقرب من ١٢٠٠ إكسابايت exabytes من البيانات سنويًا ؛ وبالتالي يواجه كل من الأفراد والمؤسسات على حد سواء ظاهرة طوفان البيانات[١] ، فعلى سبيل المثال في العصر الحالي ، يتم إنتاج وجمع البيانات حول كل نشاط يقوم به الفرد ، سواء كان ذلك أنشطة يومية مثل تصفح الإنترنت أو استخدام الهاتف أو التنقل في سيارة أو أنشطة مؤقتة مثل السجلات الصحية في المستشفي أو التسوق في السوبر ماركت أو طلب قرض من البنك . ويمثل هذا الكم الهائل من البيانات غير المهيكلة تحديات هائلة لقطاع الأعمال وتكنولوجيا المعلومات ؛ لأنها بمثابة مصدر غني لاستخراج المعلومات واستخلاص النتائج سواء لاحتياجات التسويق أو البحث والتطوير .

كما يشير بورتش Burtch إلى أنه في السنوات الأخيرة ، ينعكس استمرار زيادة حجم البيانات المتاحة على ارتفاع الطلب في سوق العمل على علماء البيانات Data Scientists الذين لديهم المهارات اللازمة لإدارة وتحليل البيانات الضخمة ومعالجتها لاستخلاص مفاهيم ذات أهمية علمية أو اقتصادية أو اجتماعية لدعم واتخاذ القرارات في مختلف المجالات[٢].

واستجابة لارتفاع الطلب في سوق العمل على علماء البيانات ، أطلقت العديد من مؤسسات التعليم العالي في الولايات المتحدة الأمريكية برامج مختلطة مثل تحليلات البيانات Data Analytics أو ذكاء الأعمال Business Intelligence أو علم البيانات Data Science . وفي عام ٢٠١١م ، أطلقت جامعة ولاية لويزيانا Louisiana State University برنامج ماجستير العلوم في التحليلات Master of Science program in Analytics . وبعد مرور عام ، أنشأت جامعة ولاية ميشيجان Michigan State University برنامج درجة الماجستير في تحليلات الأعمال Master's degree program in Business Analytics . كما افتتحت جامعة نيويورك New York University مركزًا لعلوم البيانات Center for Data Science لاستضافة برنامج ماجستير العلوم في علوم البيانات Master of

تاريخ استلام البحث : ٢٤ مارس ٢٠٢٠ ، تاريخ القبول : ١٢ مايو ٢٠٢٠

Science program in Data Science ، وتم إطلاق فصلها الافتتاحي في خريف عام ٢٠١٣م . ولقد أوضحت استجابات هذه الجامعات أن المدارس التي تستضيف برامج علوم البيانات متنوعة : وتشمل الأعمال Business ، وعلوم الكمبيوتر Computer Science ، ومدارس المعلومات Information Schools ، ومدارس الفنون والعلوم Arts and Sciences Schools [٣].

أولا : الاطار المنهجي :

١. أهمية الدراسة :

يُقدم علم البيانات Data Science - باعتباره تخصصاً مستحدثاً - اتجاهات تعليمية جديدة وحيوية في مدارس علم المكتبات والمعلومات في الوقت الراهن ، فعلي سبيل المثال : يقدم قسم علم المعلومات والمكتبات بجامعة إنديانا في بلوميجنتون Indiana University in Bloomington ، ومدرسة علوم المعلومات بجامعة إلينويز University of Illiniois في أوربانا شامبين Urbana Champaign تخصص علم البيانات على مستوى درجة الماجستير في علم المكتبات ، ودرجة الماجستير في علم المعلومات . وتتيح مدرسة المعلومات في جامعة واشنطن University of Washington مقررات مثل : مقدمة في علم البيانات Introduction to Data Science ، والبرمجة لعلم البيانات Programming for Data Science ، وتمثيل البيانات Data Visualization ، ومعالجة البيانات Data Processing ضمن أدلة المقررات الدراسية الخاصة بها . وتقدم مدرسة دراسات المعلومات في جامعة سيراكيوز Syracuse University أيضا برنامجاً لعلم البيانات على مستوى الدراسات العليا . ويشير كل ذلك إلى أن مدارس علم المكتبات والمعلومات تضطلع بدور رائد في الارتقاء بتعليم علم البيانات عبر العديد من التخصصات . وبالتالي يتعين على مدارس علم المكتبات والمعلومات Library and Information Science (LIS) ، ومدارس المعلومات iSchools السعي إلى إحداث نوع من التكامل بين كلا من علم البيانات وعلم المعلومات ، لتطوير مدارس علم المكتبات والمعلومات ، والارتقاء بها لمواكبة عصر البيانات الضخمة Big Data [٤].

وتبرز أهمية هذه الدراسة في إلقائها الضوء على خصائص البيانات الضخمة ومفهوم علم البيانات ، إلى جانب المهارات المطلوبة في علماء البيانات . وقد سلطت الدراسة الضوء أيضا على خصائص مقررات علم البيانات في مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة ، كذلك تستمد الدراسة أهميتها من أن تلفت أنظار أقسام المكتبات والمعلومات العربية بصفة عامة ، وأقسام المكتبات والمعلومات المصرية بصفة خاصة ، إلى الاهتمام بتصميم مقررات مستقلة تتعلق بعلم البيانات تطرح في إطار برامج التأهيل الأكاديمي بتلك الأقسام بما يعزز من دور تلك الأقسام في إعداد أخصائي مكتبات ومعلومات قادر على التعامل مع البيانات الضخمة تمهيداً لإعداده لسوق العمل في مجال علم البيانات .

٢. أهداف الدراسة :

تتناول هذه الدراسة واقع مقررات علم البيانات في مدارس المكتبات والمعلومات بالجامعات الأمريكية ، ومن أبرز أهداف هذه الدراسة :

١/٢. التعرف على خصائص البيانات الضخمة ومفهوم علم البيانات .

٢/٢. التعرف على واقع مقررات علم البيانات بمدارس المكتبات والمعلومات بالجامعات الأمريكية .

٣/٢. تصميم مقرر دراسي متكامل يتناول علم البيانات يمكن طرحه في أقسام المكتبات والمعلومات بالجامعات العربية بصفة عامة والمصرية بصفة خاصة .

٣. مشكلة الدراسة وتساؤلاتها :

أدركت مدارس علم المكتبات والمعلومات بصفة عامة ، ومدارس المعلومات في الولايات المتحدة الأمريكية بصفة خاصة النمو المطرد لعلم البيانات ، حيث بدأت بعض مدارس علم المكتبات والمعلومات في إطلاق برامج ماجستير لعلم البيانات ، أو إضافة المزيد من المقررات الدراسية المستقلة تتعلق بعلم البيانات في سبيل سعيها لتطوير الخطط ، والمناهج الدراسية الحالية ، وتتجه هذه المدارس ، من خلال هذا الإجراء ، لتحديد أبعاد مشكلة نقص الأعداد في علماء البيانات ، فلقد كشف تقرير أعده مانيكا وآخرون Manyika et al. في معهد ماكينزي العالمي McKinsey Global Institute (MGI) أنه من المتوقع أن تواجه الولايات المتحدة الأمريكية خلال عام ٢٠١٨م نقصًا خطيرًا في أعداد علماء البيانات بما يتراوح بين ١٤٠,٠٠٠- ١٩٠,٠٠٠ فرد (٥). ويمكن صياغة مشكلة الدراسة في السؤال الرئيسي التالي :

"ما واقع مقررات علم البيانات في مدارس المكتبات والمعلومات بالجامعات الأمريكية؟"

ويتفرع من هذا السؤال ، الأسئلة الفرعية التالية :

١/٣. ما خصائص المقررات الدراسية لعلم البيانات بمدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة من خلال تحليل نوع المقررات ، مستوى المقررات ، أسماء المقررات ، الفئات الموضوعية للمقررات وأهدافها ، طرق التدريس ، التكليفات والمهام الدراسية ، أساليب تقييم الطلاب ؟

٢/٣. ما التصور المقترح لمقرر دراسي يتناول علم البيانات ويمكن طرحه في أقسام المكتبات والمعلومات بالجامعات العربية بصفة عامة والمصرية بصفة خاصة ؟

٤. حدود الدراسة :

تستعرض الدراسة خصائص البيانات الضخمة ومفهوم علم البيانات ، إلى جانب المهارات المطلوبة في علماء البيانات ، كما تتناول الدراسة خصائص المقررات الدراسية لعلم البيانات في عينة من أقسام ومدارس المكتبات والمعلومات بالجامعات الأمريكية على مستوي البكالوريوس والماجستير والدكتوراه ، وذلك من خلال تحليل توصيفات المقررات الدراسية (سواء التفصيلية أو المختصرة) والمتاحة على مواقع مدارس المكتبات والمعلومات عينة الدراسة على شبكة الإنترنت باللغة الإنجليزية للعام الجامعي ٢٠١٩/٢٠٢٠م (وقد جرى تحليل توصيفات المقررات الدراسية خلال الفترة الزمنية من شهر أكتوبر ٢٠١٩م وحتى شهر يناير ٢٠٢٠م) .

٥. منهج الدراسة وأدواتها:

استخدمت الباحثة المنهج الوصفي التحليلي الذي يعد ملائما لمثل هذا النوع من الدراسات ، حيث تم حصر جميع مقررات علم البيانات وتحليل توصيفات تلك المقررات ، وقد اعتمدت الباحثة في هذه الدراسة على أسلوب تحليل المضمون ؛ نظراً لأنه أكثر الأساليب المنهجية تناسباً مع طبيعة هذه الدراسة . واعتمدت الباحثة في جمع بيانات هذه الدراسة على قائمة مراجعة للمقررات الدراسية لعلم البيانات وتشتمل على العناصر التالية : اسم المقرر (المصطلحات الشائعة) ، الجامعة ، المدرسة/القسم ، ونوع المقرر (إجباري/اختياري) ، ومستوي المقرر (بكالوريوس أو ماجستير أو دكتوراه) ، والفئة الموضوعية للمقرر ، وأهداف المقرر ، وطرق التدريس ، والتكليفات ، وأساليب تقييم الطلاب .

٦. مجتمع وعينة الدراسة :

يتمثل مجتمع الدراسة في قائمة أفضل مدارس المكتبات والمعلومات في الجامعات الأمريكية وذلك وفقاً لتصنيف دورية U.S. News and World Report لعام ٢٠١٧م [٦](في أحدث إصدارة لها حتى وقت الدراسة) ، والتي تتضمن و ٥١ مدرسة من مدارس المكتبات والمعلومات في الجامعات الأمريكية .

وبناء على هذا عمدت الباحثة إلى اختيار عينة مقصودة من بين مفردات هذا المجتمع بشرط أن تحتل مدارس وأقسام المكتبات والمعلومات محل الاختيار المراتب العشر الأولى في قائمة أفضل مدارس المكتبات والمعلومات في الولايات المتحدة الأمريكية وفق تصنيف دورية U.S. News and World Report لعام ٢٠١٧م (الجدول رقم ١) ، ولقد بلغ عدد مدارس وأقسام المكتبات والمعلومات عينة الدراسة ١٠ مدارس ؛ حيث يفترض أن هذه المدارس تمثل نموذجاً يمكن الاسترشاد به فيما يتعلق بتصميم المقررات الدراسية .

جدول (١) مدارس وأقسام المكتبات والمعلومات في الجامعات الأمريكية عينة الدراسة

م	مدارس وأقسام المكتبات والمعلومات	الرتبة
١	University of Illinois - School of Information Sciences	1
٢	University of Washington - The Information School	2
٣	University of North Carolina - School of Information and Library Science	3
٤	Syracuse University - School of Information Studies	4
٥	University of Michigan - School of Information	5
٦	University of Texas - School of Information	5#
٧	Rutgers University - Department of Library and Information Science	7
٨	University of Maryland - College of Information Studies	8
٩	Indiana University–Department of Library and Information Science	9
١٠	University of Pittsburgh - School of Information Sciences	10

كما قامت الباحثة بحصر البرامج والمقررات الدراسية المعتمدة في مدارس وأقسام المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة على مستوي البكالوريوس والماجستير والدكتوراه ؛ للتعرف على مدى توافر مقرر أو أكثر خاص بعلم البيانات سواء تلك المقررات التي صممت خصيصاً لتناول علم البيانات أو تلك التي تتضمن عناصر ترتبط بها بشكل غير مباشر . وتجدر الإشارة إلى أنه تم تجميع البيانات المتعلقة بالمقررات الدراسية في مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة من خلال مواقعها على شبكة الإنترنت ، حيث اعتمدت الباحثة في تحليل مضمون مقررات علم البيانات على التوصيفات الكاملة أو المختصرة لمقررات علم البيانات على مستوي البكالوريوس والماجستير والدكتوراه في مدارس المكتبات والمعلومات بالجامعات الأمريكية ؛ بشرط توافر العناصر الخاصة بأهداف ومحتوى المقررات عينة الدراسة .

جدول (٢) عدد مقررات علم البيانات المتاحة في مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة

الرتبة	مدارس وأقسام المكتبات والمعلومات	إجمالي المقررات		مقررات علم البيانات	
		العدد	النسبة المئوية	العدد	النسبة المئوية
١	University of Illinois - School of Information Sciences	١١٢	٧	٣٥	١٦,٤
٢	University of Washington - The Information School	٢٠٢	١٣	٢٠	٩,٤
٣	University of North Carolina- School of Information and Library Science	١٥٠	١٠	٢٠	٩,٤
٤	Syracuse University - School of Information Studies	٢١٧	١٤	٢٩	١٣,٦
٥	University of Michigan--Ann Arbor - School of Information	١٢٦	٨	١٦	٧,٥
#٥	University of Texas—Austin - School of Information	١٧٩	١١,٤	٩	٤,٢
٧	Rutgers University - Department of Library and Information Science	١٤٢	٩	١١	٥
٨	University of Maryland - College of Information Studies	١٨٩	١٢	٣١	١٤,٥
٩	Indiana University– Department of Library and Information Science	١٣٥	٨,٦	٢٠	٩,٤
١٠	University of Pittsburgh - School of Information Sciences	١١١	٧	٢٢	١٠,٣
الإجمالي		١٥٦٣	١٠٠	٢١٣	١٠٠

ويتبين من الجدول رقم (٢) أن إجمالي عدد المقررات الدراسية المعتمدة في مدارس وأقسام المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة للعام الجامعي ٢٠١٩/٢٠٢٠م قد بلغ ١٥٦٣ مقرراً . كما يتبين من الجدول رقم (٢) أن ١٠ مدارس أي بنسبة ١٠٠% من إجمالي مدارس وأقسام المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة يتوافر بها مقررات خاصة بعلم البيانات ، وأن إجمالي عدد مقررات علم البيانات في مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة بلغ ٢١٣ مقرراً بنسبة ١٣,٦% من إجمالي عدد المقررات في تلك المدارس (١٥٦٣ مقرراً) ؛ وهذا يشير إلى اهتمام مدارس المكتبات والمعلومات التي تحتل المراتب العشر الأولى في قائمة أفضل مدارس المكتبات والمعلومات في الجامعات الأمريكية عينة الدراسة على تبني طرح مقرر أو أكثر يتناول علم البيانات بشكل مباشر أو غير مباشر . ولقد احتلت مدرسة علوم المعلومات School of Information Sciences بجامعة إلينويز University of Illinois المرتبة الأولى حيث بلغ عدد ما تطرحه من مقررات علم البيانات ٣٥ مقرر بنسبة ١٦,٤% من إجمالي مقررات علم البيانات عينة الدراسة (٢١٣ مقرراً) مما يشير إلى وجود علاقة ارتباطية بين حصول مدرسة علوم المعلومات بجامعة إلينويز على الترتيب الأول في تصنيف دورية U.S. News and World Report 2017 وما يوفره من مقررات في علم البيانات . وتأتي في المرتبة الثانية مدرسة دراسات المعلومات College of Information Studies بجامعة

ميريلاند University of Maryland بعدد ٣١ مقرراً بنسبة ١٤,٥% من إجمالي ٢١٣ مقرر من مقررات علم البيانات عينة الدراسة . وتحتل ثلاث مدارس المرتبة الخامسة ، وهي مدرسة المعلومات The Information School بجامعة واشنطن University of Washington ، ومدرسة علوم المكتبات والمعلومات School of Information and Library Science بجامعة ولاية كارولينا الشمالية University of North Carolina ، وقسم علوم المكتبات والمعلومات Department of Library and Information Science بجامعة إنديانا Indiana University حيث بلغ عدد المقررات في كل منهم ٢٠ مقرر بنسبة ٩,٤% ، وجاءت في المرتبة الأخيرة على التوالي كل من قسم علوم المكتبات والمعلومات Department of Library and Information بجامعة ولاية نيوجيرسي The State University of New Jersey بنسبة ٥% (١١ مقرر) ، ومدرسة المعلومات School of Information بجامعة تكساس University of Texas—Austin - بنسبة ٤,٢% (٩ مقررات) من إجمالي ٢١٣ مقرر من مقررات علم البيانات عينة الدراسة .

وتركز الدراسة الحالية على مقررات علم البيانات المتاحة في مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة (٢١٣ مقرراً بنسبة ١٣,٦%) سواء تلك المقررات التي صممت خصيصاً لتناول علم البيانات بشكل مباشر أو تلك المقررات التي تتضمن عناصر ترتبط بعلم البيانات بشكل غير مباشر ؛ للتعرف على الخصائص التي يجب توافرها في مقررات علم البيانات .

٧. مصطلحات الدراسة :

تناولت الباحثة المصطلحات الأساسية للدراسة من خلال وضع تعريف إجرائي بسيط لكل مصطلح ، وهو التعريف الذي التزمت به في هذه الدراسة .

١/٧. علم البيانات :

لقد حاولت العديد من الدراسات تعريف علم البيانات ، حيث عرف جورج وآخرون[٧] George. et al علم البيانات بأنه مجال ناشئ متعدد التخصصات لاشتماله على : علم الإحصاء Statistics Science ، والتنقيب في البيانات Data mining ، وتعلم الآلة Machine Learning ، وتحليل البيانات Data Analysis . وأشار سونج وزهو[٨] Song & Zhu إلى ضرورة الإلمام بالرياضيات والإحصاء ، وعلم الحاسب ، وتعلم الآلة فهي العناصر الأساسية لعلم البيانات ، كما أكد الباحثان على أن علم البيانات يهدف إلى تحويل البيانات إلى معلومات ، ومعرفة ، وحكمة ، وأن الركائز الأساسية لعلم البيانات ، هي : البيانات ، والتقنية ، والأفراد . في حين أشار تشاو ولو[٩] Chao & Lu إلى أن العناصر الأساسية لعمل البيانات ، هي : علم الإحصاء ، تعلم الآلة ، تمثيل البيانات ، تجهيز البيانات ، والحوسبة ، والإدارة ، وتحليل البيانات ، إلى جانب تطبيقات علم البيانات في مجالات المعرفة المختلفة .

وعلى الرغم من اختلاف الآراء فيما يتعلق بمفهوم علم البيانات ، ألا أن هناك نوع من الإجماع على أن علم البيانات هو أحد المجالات متعددة التخصصات لاشتماله على : علم الإحصاء ، والتنقيب في البيانات ، وتعلم الآلة ، وتحليل البيانات . ويهدف علم البيانات إلى استخدام الأساليب العلمية ، والمعالجات والخوارزميات والنظم لرصد مشكلات البيانات الضخمة واستخراج المعلومات والمعرفة من تلك البيانات سواء المهيكلة أو غير المهيكلة من خلال تقنيات تعلم الآلة والذكاء الصناعي وبرامج معالجة البيانات الضخمة .

٢/٧. البيانات الضخمة :

البيانات الضخمة هي مفهوم حديث نسبياً ، ولقد استخدم لأول مرة من قبل لاني Laney عام ٢٠٠١م ؛ حيث عرف البيانات الضخمة بأنها : مجموعة من البيانات الضخمة والمعقدة ، والتي تتميز بمجموعة من الخصائص الفريدة (5Vs) ، وهي : الحجم Volume ، وسرعة النمو Velocity ، والتنوع Variety ، والصدق Veracity ، والقيمة Value ، ولا تستطيع أساليب إدارة البيانات التقليدية التعامل معها ، حيث تتطلب تلك البيانات تقنيات متطورة لتجميعها ومعالجتها وتخزينها وتحليلها والبحث فيها ومشاركتها ونقلها وتمثيلها لدعم واتخاذ القرارات في المجالات المختلفة [١٠].

٨. الدراسات السابقة وأدب الموضوع :

على حد علم الباحثة فإنه لا يوجد دراسات عربية تناولت موضوع تعليم علم البيانات بأقسام ومدارس المكتبات والمعلومات بالجامعات الأمريكية ، وإذا تحدثنا عن الدراسات العربية السابقة من منظور الدراسات التي تناولت موضوع البيانات الضخمة من خلال رصد علاقة البيانات الضخمة بقطاع المكتبات والمعلومات وأثر هذا المصطلح الجديد على مهنة المكتبات والمعلومات من حيث مؤسساتها أو العاملين فيها ، فهي بضعة دراسات قليلة ، مثل : دراسة أحمد أحمد المزين (٢٠١٩)[١١] ، ودراسة خالد عتيق سعيد عبد الله ، عبد الله بن سالم الهنائي (٢٠١٨)[١٢] ، ودراسة عبدالرحمن فراج (٢٠١٨)[١٣] ، ودراسة زينب بن الطيب ، سليمان بن إبراهيم الرباعي (٢٠١٨)[١٤] ، ودراسة يونس أحمد إسماعيل الشوابكة (٢٠١٨)[١٥] ، ودراسة عمر حسن فتوح (٢٠١٧)[١٦] وغير ذلك من دراسات لا يحسن التنويه إليها في هذا المقام ، إذ تبدو بعيدة شيئاً ما عما تهدف إليه الدراسة الحالية .

أما بالنسبة للدراسات الأجنبية فقد اتضح بمراجعة الإنتاج الفكري المرتبط بموضوع الدراسة وجود عدد من الدراسات التي تتناول موضوع تعليم علم البيانات وتأهيل علماء البيانات في شتي المجالات ، مما يشير إلى أهمية الموضوع وبصفة خاصة الاهتمام بالجوانب التطبيقية لعلم البيانات في المجالات المختلفة . كما أسفر البحث عن وجود عدد قليل من الدراسات التي تهتم بالتأهيل الأكاديمي لطلاب مدارس المكتبات والمعلومات للتعامل مع علم البيانات . وتم التوصل إلى الدراسات التالية والتي تم ترتيبها ترتيباً زمنيا من الاحدث إلى الأقدم ثم ترتيبا هجائيا بالمؤلف في حالة تكرار نفس الفترة الزمنية :

دراسة كروس وآخرون Kross, Sean et al. (2019) [١٧]، بعنوان The democratization of data science education . تناولت الدراسة برنامج علم البيانات في جامعة جونز هوبكنز Johns Hopkins University والمتاح أيضا على منصة كورسيرا Coursera ، ويتكون البرنامج من ٩ دورات تدريبية للطلاب في تخصصات الرياضيات وعلوم الكمبيوتر . وتهدف الدراسة إلى وصف البرنامج ومقارنته بمناهج علم البيانات القياسية في الفترة ما بين عامي ٢٠١٤م و ٢٠١٥م . كما تناقش الدراسة تأثير تخصص علم البيانات على تعليم علم البيانات في الولايات المتحدة الأمريكية . وانتهت الدراسة بعدة نتائج ، منها : أن تخصص علم بيانات في جامعة جونز هوبكنز أصبح لديه الآن أكثر من ٤ ملايين تسجيل على مدار السنوات الخمس الماضية . وأن برنامج علم البيانات يركز على تدريب الطلاب في مجالات الإحصاء والتعلم الآلي وتحليل البيانات . كما تشير الدراسة إلى أن تخصص علم البيانات يعتبر من أكثر التخصصات طلباً الأن في سوق العمل . وأن الغالبية العظمى من برامج علم البيانات تستهدف طلاب الدراسات العليا ممن لديهم خلفية تعليمية في تخصصات الرياضيات وعلوم الكمبيوتر ، ولكن هناك فرصة كبيرة لتطوير وتعليم علم البيانات خارج هذه المجموعات والتواصل مع مختلف التخصصات .

دراسة مبويلو و كيمارو و خوستو Mbwilo, Betty & Kimaro, Honest & Justo, Godfrey (٢٠١٩) [١٨]، بعنوان Data Science Postgraduate Education at University of Dar es Salaam in Tanzania: Current Demands and Opportunities . هدفت الدراسة إلى تحليل متطلبات أصحاب سوق العمل من خلال المعارف والمهارات والاتجاهات التكنولوجية المطلوبة في مجال علم البيانات ؛ لإعداد برنامج دراسات عليا لعلم البيانات في جامعة دار السلام University of Dar es Salaam في تنزانيا ، وتم توزيع استبيان على ٨٥ مؤسسة تم تحديدها لالتماس معلومات عن احتياجات سوق العمل لعلماء البيانات والفجوات الموجودة في المهارات . وتم تلقي ما مجموعه ٦١ استجابة من الاستبيانات (بنسبة ٧٠%) . وانتهت الدراسة بعدة نتائج ، منها : أنه لا يوجد عدد كافٍ من الأفراد في سوق العمل يتمتعون بمهارات عالية في مجال علم البيانات ، وحتى الخريجين في تخصص تكنولوجيا المعلومات والاتصالات لا يمتلكون المهارات التي يتطلبها أصحاب العمل ؛ مما يشير إلى ارتفاع الطلب في سوق العمل على مهارات عالم البيانات في تنقيب وتحليل البيانات لمساعدة أصحاب العمل على اتخاذ القرارات الاستراتيجية بشكل أفضل . كما تشير الدراسة إلى أنه تمت الموافقة على برنامج جديد في ماجستير العلوم في علم البيانات وتأسيسه منذ أبريل عام ٢٠١٨م في جامعة دار السلام في تنزانيا ، وعلى الرغم من موافقته المتأخرة ودون الإعلان الرسمي على وسائل الإعلام العامة ، فقد اجتذب البرنامج عددًا كبيرًا من المتقدمين للعام الدراسي ٢٠١٨/٢٠١٩م مقارنةً ببرامج الدراسات العليا الأخرى في تكنولوجيا المعلومات والاتصالات المقدمة في جامعة دار السلام .

دراسة ميكروياني‌ديس وآخرون Mikroyannidis, Alexander et al. (٢٠١٨) [١٩]، بعنوان Designing and Delivering a Curriculum for Data Science Education Across Europe . تناولت الدراسة تجربة الأكاديمية الأوروبية لعلوم البيانات the European Data Science Academy (EDSA) ؛ وهي مشروع الاتحاد الأوروبي الذي استمر ما بين فبراير عام ٢٠١٥ وحتى يناير عام ٢٠١٨م ، وكان الهدف من المشروع هو تصميم وتقديم منهج لتعليم علم البيانات في جميع أنحاء أوروبا لسد فجوة المهارات في علم البيانات في الاتحاد الأوروبي وتدريب جيل جديد من علماء البيانات الرائدين في العالم ، بالإضافة إلى تقديم دورة تدريبية عن المهارات المطلوبة لمحللي البيانات عبر القطاعات الصناعية الرئيسية في أوروبا .

دراسة وانج ، لين Wang, Lin (2018) [٢٠]، بعنوان Twinning data science with information science in schools of library and information science . هدفت الدراسة إلى التعرف على علم البيانات باعتباره تخصصاً ناشئاً ، حيث يمثل علم البيانات اتجاه تعليمي جديد وحيوي في مدارس علم المكتبات والمعلومات في الوقت الراهن . وكذلك التعرف أيضا على كيفية ارتباط علم البيانات بعلم المعلومات في مدارس علم المكتبات والمعلومات . وتناولت الدراسة تحليل البيان الخاص برسالة وطبيعة كل من : علم البيانات وعلم المعلومات من خلال مراجعة العمل القائم في كلا التخصصين ، ووضع تسلسل هرمي للبيانات ، والمعلومات ، والمعرفة ، والحكمة Data – Information – Knowledge – Wisdom (DIKW) ، كما تتناول الدراسة الطرق التي تستحضر من خلالها الرؤى ، والأفكار الجديدة ، ونظريات علم المعلومات ، وإلقاء الضوء على أساسيات علم البيانات . وانتهت الدراسة بعدة نتائج منها : أن علم البيانات وعلم المعلومات هما تخصصات بمثابة التوأم من حيث الطبيعة والرسالة ، فكلاهما يتداخل مع الآخر ويشاركه نفس الاهتمامات إلى حد بعيد ، إلى جانب أن هناك نوع من التكامل ما بين علم البيانات وعلم المعلومات في مدارس علم المكتبات والمعلومات LIS ، ومدارس المعلومات iSchools .

دراسة سيميلر وبينتو وروزادوس Semeler, Alexandre Ribas & Pinto, Adilson Luiz & Rozados, Helen Beatriz Frota (٢٠١٧) [٢١]، بعنوان Data science in data librarianship: Core competencies of a data librarian . هدفت الدراسة إلى تحديد المهارات التقنية المطلوبة لأمناء مكتبات البيانات من خلال التعرف على مفاهيم البيانات وعلم البيانات وبيانات البحث ومكتبة البيانات ، كما ناقشت الدراسة الكفاءات والمهارات التي يتطلبها علماء البيانات وأمناء مكتبات البيانات . وانتهت الدراسة بعدة نتائج منها : أن علم البيانات يوفر أساليب وممارسات جديدة لأمناء مكتبات البيانات ، ولا ينبغي أن يصبح أمين مكتبة البيانات مبرمجًا أو إحصائيًا أو مدير قاعدة بيانات ، ولكن يجب أن يهتم بالتعرف على لغات البرمجة وقواعد البيانات وأدوات استرجاع المعلومات . وأن أمناء المكتبات يجب أن يتعلموا المزيد حول إدارة البيانات وتحليلها من خلال اكتساب بعض مهارات عالم البيانات والتي تشمل : المهارات التقنية ، والمهارات التحليلية ، ومهارات عرض النتائج .

دراسة تشانج وآخرون Zhang, Jilong et.al (٢٠١٧) [٢٢]، بعنوان The development of data science education in China from the LIS perspective . هدفت الدراسة إلى إدخال تطوير علم البيانات في التعليم العالي في الصين ، من خلال تحليل تعليم علم البيانات في خمس مدارس لعلوم المكتبات والمعلومات في الصين ، وهي : مدرسة إدارة المعلومات School of Information Management بجامعة وهان Wuhan University ، ومدرسة إدارة المعلومات School of Information Management بجامعة نانجينغ Nanjing University ، ومدرسة إدارة مصادر المعلومات School of Information Resources Management بجامعة رنمين الصينية Renmin University of China ، وقسم إدارة المعلومات Department of Information Management بجامعة بكين Peking University ، ومكتبة العلوم الوطنية National Science Library بجامعة الأكاديمية الصينية للعلوم Chinese Academy of Sciences . باستثناء المكتبة الوطنية للعلوم التابعة للأكاديمية الصينية للعلوم ، والتي تقدم برامج الدراسات العليا فقط ، تقدم المدارس الأربع الأخرى برامج البكالوريوس والماجستير والدكتوراه . وتناولت الدراسة تحليل توصيفات البرنامج والمقررات الدراسية المتاحة على المواقع الإلكترونية لمدارس المكتبات والمعلومات عينة الدراسة . وانتهت الدراسة بعدة نتائج ، منها : أن مدارس المكتبات والمعلومات في الصين تبدو أكثر ثقة في استخدام مصطلح المعلومات Information بدلا من مصطلح البيانات لوصف برامجها ؛ ونتيجة لذلك نادراً ما تظهر كلمة "data" في اسم البرامج التي تتيحها مدارس المكتبات والمعلومات في الصين على الرغم من أنه في بعض الأحيان يمكن العثور على كلمة "البيانات" في مسميات المناهج الدراسية ومحتوى الدورة التدريبية ، فعلى سبيل المثال يتيح برنامج المكتبات والمعلومات بجامعة بكين مقررات مثل : الإحصاء Statistics وتحليل البيانات Data Analysis وتنقيب البيانات Data Mining وهياكل البيانات والخوارزميات Data Structures and Algorithms . كما اشارت الدراسة إلى أن هناك حاجة إلى إضافة المزيد من الدورات التدريبية نحو تطوير المهارات الإحصائية والحوسبة لدى الطلاب في المستقبل ، بحيث يمكن للخريجين أن يكونوا أكثر قدرة على المنافسة مقارنة بالخريجين من برامج علم البيانات في تخصصات أخرى . كما تشير الدراسة إلى أن تعليم علم البيانات في برامج المكتبات والمعلومات الصينية لا يزال في مرحلة التطوير الأولى .

دراسة مارشال وإلوف Marshall, Linda & Eloff, Jan H.P. (٢٠١٦) [٢٣]، بعنوان Towards an Interdisciplinary Master's Degree Programme in Big Data and Data Science: A South African Perspective . تناولت الدراسة إطارًا نظرياً لتصميم برنامج أكاديمي متعدد التخصصات على مستوي الماجستير في البيانات الضخمة وعلم البيانات . وانتهت الدراسة بعدة نتائج ، منها : أن مجالات المعرفة التالية ذات صلة بمنهج علوم البيانات والبيانات الضخمة ، وهي : التعلم

الآلي وتمثيل البيانات ونمذجة البيانات Data modeling والتعلم الإحصائي وتطبيقات علم البيانات في مجالات المعرفة المختلفة . كما اشارت الدراسة إلى أن هناك علاقة ترابط ما بين المعرفة المكتسبة وبين مهارات الاتصال والبرمجة وتحليل البيانات في مجال علم البيانات .

دراسة ميلز وتشودوبا وأولسن Mills, Robert J. & Chudoba, Katherine M. & Olsen, David H. (2016) [٢٤]، بعنوان IS Programs Responding to Industry Demands for Data Scientists: A Comparison between 2011 – 2016 . تناولت الدراسة تحليل محتوي برامج نظم المعلومات Information Systems (IS) programs التي تقدمها جمعية تطوير كليات إدارة الأعمال Association to Advance Collegiate Schools of Business (AACSB) في ١٢٩ جامعة بالولايات المتحدة الأمريكية ، خلال الفترة من عام ٢٠١١م وحتى عام ٢٠١٦م ، وتم اختيار عينة عشوائية تكونت من ١١٨ برنامج من إجمالي ما يزيد عن ٤٨٥ برنامجًا من برامج AACSB . واعتمدت الدراسة على توصيفات البرامج المتاحة على المواقع الإلكترونية للجامعات ، إلى جانب المقابلات الهاتفية مع المرشدين الأكاديميين في حالة عدم توافر بعض البيانات على المواقع الإلكترونية للجامعات . وانتهت الدراسة بعدة نتائج ، منها : أن غالبية برامج نظم المعلومات بنسبة ٦٠% تقدم دورات متعلقة بعلوم البيانات ما بين عامي ٢٠١١ و ٢٠١٦ وخاصة فيما يتعلق بموضوعات تحليل البيانات وذكاء الأعمال Business intelligence والبيانات الضخمة . كما تشير الدراسة إلى أن العديد من برامج نظم المعلومات استجابت لاحتياجات الصناعة وسوق العمل من خلال إضافة دورات تدريبية في مجال تحليل البيانات الضخمة .

دراسة سونج وتشو Song, Il-Yeol & Zhu, Yongjun(٢٠١٦) [٢٥]، بعنوان Big data and data science: what should we teach? . هدفت الدراسة إلى مناقشة خصائص البيانات الضخمة ، ووجهات النظر المختلفة حول علم البيانات ، وأدوار علماء البيانات . كما تناولت استعراض للوضع الحالي لتعليم علم البيانات في الولايات المتحدة الأمريكية من خلال تحليل ٤٢ برنامج أكاديمي في كل من تخصص علم البيانات ، وتخصص تحليل البيانات خلال عام ٢٠١٤م . وانتهت الدراسة بعدة نتائج ، منها : أن تعليم علم البيانات على مستوي برامج البكالوريوس في الولايات المتحدة الأمريكية غير شائع جداً حتى الآن ، وهناك عدد قليل فقط من الجامعات التي تقدم هذه الأنواع من البرامج ، مثل برنامج البكالوريوس بجامعة ولاية أوهايو Ohio State University حول تحليلات البيانات . كما اشارت الدراسة إلى أنه يتم تقديم معظم برامج الماجستير في علم البيانات بالتعاون مع أقسام علوم الكمبيوتر . كما تعد مقررات تنقيب البيانات والتعلم الآلي وتمثيل البيانات أكثر المقررات الأساسية شيوعًا ، في حين تعد الإحصاء وقواعد البيانات مقررين أساسيين في كل من برامج البكالوريوس والماجستير.

دراسة تانج وساي ليم Tang, Rong & Sae-Lim, Watinee (2016) [٢٦]، بعنوان Data science programs in U.S. higher education: An exploratory content analysis of program description, curriculum structure, and course focus . تناولت الدراسة تحليل محتوى مضمون البرامج الدراسية المرتبطة بعلم البيانات والمعتمدة في الجامعات الأمريكية في ثمانية تخصصات مختلفة ، وهي : الفنون والعلوم ، والتجارة ، وعلوم الكمبيوتر ، والهندسة ، وعلوم البيانات ، وعلوم المعلومات ، الرياضيات والاحصاء ، والدراسات المهنية Professional Studies . وتم اختيار عينة عشوائية طبقية تكونت من ٣٠ برنامج أكاديمي في علوم البيانات على مستوى الدراسات العليا . وانتهت الدراسة بعدة نتائج ، منها : أن برامج علم البيانات تتطلب أعدادًا متفاوتة من الساعات المعتمدة ، بما في ذلك التدريب العملي حيث أن متوسط عدد الساعات المعتمدة المطلوبة في ٣٠ برنامجًا لعلم البيانات أكثر من ٤٠ ساعة ، مع ما يقرب من ٣٠ ساعة معتمدة في المقررات الأساسية وأكثر من ١٠ ساعات في

المقررات الاختيارية . كما أن هناك اختلافات من حيث طول توصيف البرنامج وعدد الكلمات والعبارات الفريدة والمصطلحات الأكثر شيوعًا المستخدمة . ولم تقدم جميع برامج علم البيانات تعريفا لعلوم البيانات . كما كانت الدورة التدريبية المقدمة في مجال قوانين الإحصاء في الرياضيات ضعيفة جدًا في مدارس المعلومات iSchools . في حين وفرت المقررات الأساسية في مدارس المعلومات مهارات الاتصال وتمثيل البيانات ، ولم تتناول المقررات الاختيارية هذه المهارات . هذه النتائج لها آثار على تحسين تعليم علوم البيانات في مدارس المعلومات وعبر التخصصات الأخرى .

دراسة يي ، ميونغهو Yi, Myongho (٢٠١٦) [٢٧]، بعنوان A Study on the Curriculums of Data Science . هدفت الدراسة إلى مقارنة سبعة برامج لعلم البيانات في كوريا وعشرة برامج لعلم البيانات في الولايات المتحدة الأمريكية . وانتهت الدراسة بعدة نتائج ، منها : أن ١٤ برنامجًا لعلم البيانات موجودة على مستوي الدراسات العليا . وأن ١٠% من برامج علم البيانات في كوريا و ٢٦% في الولايات المتحدة تندرج تحت مجال الرياضيات والإحصاء . كما لا يُظهر تحليل محتوي المقررات في برامج علم البيانات في كل من كوريا والولايات المتحدة الأمريكية الكثير من الاختلافات من حيث المحتوي الموضوعي للمقررات وتوزيع الدرجات للمقررات .

دراسة عشاييم وآخرون Aasheim, Cheryl L. et al. (٢٠١٥) [٢٨]، بعنوان Data Analytics vs. Data Science: A Study of Similarities and Differences in Undergraduate Programs Based on Course Descriptions . هدفت الدراسة إلى التعرف على أوجه التشابه والاختلاف في البرامج الأكاديمية لكل من تخصص تحليل البيانات وتخصص علم البيانات data analytics and data science programs . وتكونت عينة الدراسة من ١٣ جامعة من إجمالي ٨٦ جامعة في الولايات المتحدة الأمريكية حيث اقتصرت الدراسة على الجامعات التي تتيح على مواقعها الإلكترونية توصيفات كاملة وتفصيلية للبرامج الأكاديمية لكل من تخصص تحليل البيانات وتخصص علوم البيانات . وانتهت الدراسة بعدة نتائج ، منها : أن برامج علوم البيانات تفتقر إلى مقررات تغطي مهارات صنع القرار وإدارة البيانات . كما اشارت الدراسة إلى أنه لا يوجد تغطية للاعتبارات الأخلاقية والقانونية في أي من المقررات المتعلقة بتحليل البيانات أو علم البيانات في البرامج الثلاثة عشر التي تم فحصها ، على الرغم من أن هذه الموضوعات قد تم ذكرها مرارًا وتكرارًا في الأدب كمجال ذي أهمية .

ثانيا : الإطار النظري للدراسة :

١. خصائص البيانات الضخمة :

السرعة الرهيبة في تضاعف حجم البيانات في سنوات قليلة مرده إلى تنوع وتعدد مصادر هذه البيانات والجهات المساهمة في إنتاجها بهذه الأحجام الخيالية . فظهور الإنترنت وظهور الجيل الثاني للويب WEB2.0 ساهم بشكل كبير في مضاعفة حجم البيانات علاوة على التطور التكنولوجي السريع الحاصل إذ أن أجهزة الكمبيوتر والأجهزة الذكية والهواتف النقالة وأجهزة الاستشعار عن بعد تساهم بشكل كبير في إنتاج وتراكم كميات ضخمة من البيانات ؛ وبالتالي لزم استحداث طرق جديدة أكثر كفاءة لتنظيمها ومعالجتها حاسوبيا ليأتي بعد ذلك دور عالم البيانات ويستخرج منها معلومات مفيدة . فالبيانات الضخمة تتعذر معالجتها في قواعد البيانات التقليدية إما لكبر حجمها ، أو بسبب حركتها السريعة ، أو لأنها لا تلائم بنية قواعد البيانات التقليدية ، وتنطوي البيانات الضخمة على خمس خصائص ويتم الإشارة إليها باللغة الإنجليزية بـــ 5Vs ، وهي [٢٩]:

١/١. الحجم Volume :

يقصد به كمية البيانات المتوفرة لدى المؤسسة أي حجم البيانات المستخرجة من مصدر ما ، وهو ما يحدد قيمة وإمكانيات البيانات لكي تصنف ضمن البيانات الضخمة ؛ فحين نشير إلى حجم البيانات العادية فإننا نستخدم وحدات كالميجابايت Megabyte والجيجابايت Gigabyte ، أما البيانات الضخمة فقد يصل حجمها إلى التيرابايت Terabyte والإكسابايت Exabyte والزيتابايت Zettabyte .

٢/١. سرعة النمو Velocity :

يقصد بها سرعة إنتاج واستخراج البيانات لتغطية الطلب عليها حيث تعتبر السرعة عنصراً أساسياً في اتخاذ القرار بناءً على هذه البيانات ، وهو الوقت الذي تستغرقه من لحظة وصول هذه البيانات إلى لحظة الخروج بالقرار بناءً عليها .

٣/١. التنوع Variety :

يقصد به تنوع البيانات المستخرجة ؛ وتتضمن بيانات مهيكلة (منظمة) وبيانات غير مهيكلة (غير منظمة) مثل : الصور والفيديو ورسائل البريد الإلكتروني وسجلات المكالمات وبيانات الخرائط GPS وغيرها من البيانات التي تتطلب وقتاً وجهداً لمعالجتها وتحليلها لاستخراج نتائج منها .

٤/١. الموثوقية Veracity :

يقصد بها مدى جودة وصحة البيانات التي يتم الحصول عليها ؛ أي تشير إلى مدى الاعتمادية على تلك البيانات للمساعدة في اتخاذ القرارات .

٥/١. القيمة Value :

إن قيمة البيانات تكمن في الاستفادة منها واستخدامها بعد معالجتها ، وللبيانات الضخمة أهمية بالغة فهي تساعد على تحقيق الميزة التنافسية للمؤسسات من خلال فهم أعمق لاحتياجات العملاء وبالتالي اتخاذ القرارات المناسبة في الوقت المناسب .

٢. متطلبات عالم البيانات :

يشير سونج وزهو[٣٠] Song & Zhu إلى أن الركائز الأساسية لعلم البيانات ، هي : البيانات (وتشمل البيانات المهيكلة وشبه المهيكلة) ، والتقنيات (مثل ، أنظمة هادوب Hadoop ، ونظام قواعد البيانات غير العلائقية Non-Relational Database Management System ، تقنيات تنقيب البيانات ، ونظم الحوسبة السحابية ، وتقنيات التعلم الآلي ، ...إلخ) ، والأفراد (وتشمل : علماء الكمبيوتر والإحصاء وخبراء المجال وعلماء البيانات ومحللي الأعمال ، ... إلخ) .

وظهر مصطلح عالم البيانات Data Scientist لأول مرة في عام ٢٠٠٥م ، ولكنه أصبح شائعًا في عام ٢٠١٢م ، في مقال بعنوان "علم البيانات: المهنة الأكثر جاذبية في القرن الحادي والعشرين" "Data science: The sexiest profession of the 21st century" ، التي نشرت في مجلة هارفارد بيزنس ريفيو the Harvard Business Review ، ويشير فيها المؤلف إلى أن وظيفة عالم البيانات ستكون الوظيفة الأكثر طلباً وجاذبية في القرن الواحد والعشرين [٣١].

أحيانا يتم استخدام مصطلح "عالم البيانات" كمرادف لمحللي البيانات data analysts والإحصائيين statisticians ومترجمي البيانات data interpreters ومحللي ذكاء الأعمال business intelligence analysts وباحثي البيانات data researchers ومطوري البيانات data

developers ومهندسي البيانات data engineers ، إلا أن هناك اختلافات بينهما . عالم البيانات هو شخص يجيد الإحصاء أفضل من أي مهندس برمجيات وفي نفس الوقت يجيد هندسة البرمجيات أفضل من أي خبير في الاحصاء ، فهو يتسم بمجموعة من المهارات تعتمد بشكل أساسي على الرياضيات والخوارزميات ، إلى جانب المعرفة العميقة للقطاع أو المجال الذي يعمل فيه لصياغة النتائج بشكل مناسب . وهناك عدد من المهارات والمتطلبات المهنية التي يجب على علماء البيانات إتقانها ، من أهمها [٣٢، ٣٣]:

- ١/٢. الرياضيات والاحصاء : الإلمام بالجبر والجبر والخطي وحساب التفاضل والتكامل ، إلى جانب الإلمام بالإحصاء الوصفي والاستدلالي ونظرية الاحتمالات .
- ٢/٢. المهارات الأساسية في استخدام أدوات وتقنيات التعلم الآلي Machine learning ، وبصفة خاصة خوارزميات التعلم الآلي مثل : الخوارزميات الخاضعة للإشراف Supervised Learning ، والخوارزميات غير الخاضعة للإشراف Unsupervised Learning .
- ٣/٢. هندسة البرمجيات : الإلمام بالحوسبة الموزعة distributed computing ، الخوارزميات وهياكل البيانات algorithms and data structures .
- ٤/٢. مهارات استخراج والتنقيب في البيانات Data mining لاستنتاج واستخراج المعرفة من مجموعة من البيانات الضخمة .
- ٥/٢. مهارات تنظيف البيانات ودمجها Data cleaning and merging ويقصد بها استكمال البيانات المفقودة أو إزالة البيانات المتنافرة أو غير المتناسقة ؛ فهي تهدف إلى تجهيز البيانات قبل البدء في مرحلة التحليل الإحصائي للبيانات وتنقيب البيانات .
- ٦/٢. مهارات تمثيل البيانات Data visualization : أي القدرة على تقديم نتائج البيانات بأشكال مرئية واضحة ومنسقة الألوان ؛ لجذب اهتمام الأفراد المتخصصين وغير المتخصصين في علم البيانات .
- ٧/٢. مهارات استخدام قواعد البيانات العلائقية وغير العلائقية ولغات الاستعلام وبصفة خاصة SQL .
- ٨/٢. مهارات استخدام لغات البرمجة : وتعد لغة بايثون Python ولغة آر R من أكثر اللغات استخداماً ، فعلي سبيل المثال تستخدم لغة بايثون لقدرتها على التحليل الإحصائي ، واحتوائها على حزم متعددة للتعلم الآلي ، وتمثيل البيانات ، وتحليل البيانات .
- ٩/٢. القدرة على العمل على البيانات غير المهيكلة من مصادر متنوعة مثل مواقع التواصل الاجتماعي .
- ١٠/٢. وأن يتمتع بمهارات الاتصال الكتابية والشفوية أي القدرة على ترجمة نتائج البيانات بتنسيق يستطيع الآخرون من غير المتخصصين في علم البيانات من فهمه لاتخاذ القرار .
- ١١/٢. بالإضافة إلى المهارات الإحصائية والتقنية ، سيحتاج علماء البيانات إلى التعامل مع القضايا الأخلاقية والقانونية
- ١٢/٢. أخيرًا ، "السمة السائدة بين علماء البيانات هي الفضول الشديد والرغبة في طرح التساؤلات لتحديد المشكلة والوصول لحل المشكلات .

وتجدر الإشارة إلى أن كل هذه المهارات والتقنيات لا يمكن أن تحصل على تغطية كاملة ضمن برنامج أكاديمي واحد للحصول على درجة البكالوريوس أو الماجستير أو الدكتوراه ، ويتمثل التحدي الذي يواجه الأكاديميين في اختيار أي من هذه المهارات يجب تغطيتها وإلى أي مدى .

ثالثا : الدراسة الميدانية :

تناولت الدراسة تحليل توصيفات مقررات علم البيانات بالبرامج الأكاديمية في مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة كما وردت في مواقعها على الإنترنت سواء كانت مختصرة أو تفصيلية من حيث : نوع المقررات ، المصطلحات الشائعة في أسماء المقررات ، مستوى المقررات ، الفئات الموضوعية للمقررات وأهدافها ، طرق التدريس ، التكليفات والمهام الدراسية ، أساليب تقييم الطلاب . ويتم في الأجزاء التالية تناول نتائج الدراسة بشأن تلك المقررات .

١. مقررات علم البيانات وفقاً لنوع المقرر :

وقد وضعت الباحثة تعريفا اجرائيا لكل من المقررات الاجبارية والمقررات الاختيارية تمثلت فيما يلى :

- **المقررات الإجبارية** : هي تلك المقررات التي تمثل المعرفة الأساسية التي يجب على الطالب اجتيازها في البرنامج الأكاديمي لمدارس علوم المكتبات والمعلومات أو مدارس المعلومات .
- **المقررات الاختيارية** : هي مقررات متخصصة ولكنها تعطي فرصة التنوع لتلبية رغبة كل طالب في اختيار المقررات المناسبة للمجال الذي يرغب في العمل به بعد التخرج .

جدول (٣) عدد مقررات علم البيانات في مدارس المكتبات والمعلومات عينة الدراسة طبقاً لنوع المقرر

الرتبة	مدارس المكتبات والمعلومات	المقررات الإجبارية		المقررات الاختيارية	
		العدد	النسبة %	العدد	النسبة %
١	University of Illinois - School of Information Sciences	٤	٢	٣١	١٤,٥
٢	University of Washington - The Information School	٢	٠,١	١٨	٨
٣	University of North Carolina - School of Information and Library Science	٣	١,٤	١٧	٨
٤	Syracuse University - School of Information Studies	١٣	٦	١٦	٧,٥
٥	University of Michigan - School of Information	٨	٤	٨	٤
#٥	University of Texas - School of Information	-	-	٩	٤
٧	Rutgers University - Department of Library and Information Science	٥	٢,٣	٦	٣
٨	University of Maryland - College of Information Studies	٦	٣	٢٥	١٢
٩	Indiana University - Department of Library and Information Science	٩	٤,٢	١١	٥
١٠	University of Pittsburgh - School of Information Sciences	١٤	٧	٨	٤
	الإجمالي	٦٤	٣٠	١٤٩	٧٠

يتبين من الجدول رقم (٣) ، أن الغالبية العظمي من مقررات علم البيانات كانت مقررات إختيارية بنسبة ٧٠% (١٤٩ مقرر) ، ولقد لاحظت الباحثة توافر مقررات علم البيانات كمقررات إجبارية في ٩ مدارس بنسبة ٩٠% من إجمالي مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة حيث لم يتوافر أي مقررات إجبارية لعلم البيانات في مدرسة المعلومات بجامعة تكساس University of Texas ، بينما تتوافر مقررات علم البيانات كمقررات اختيارية في جميع مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة أي بنسبة ١٠٠% .

٢. المصطلحات الشائعة في أسماء مقررات علم البيانات :

جدول (٤) توزيع المصطلحات الشائعة في أسماء مقررات علم البيانات عينة الدراسة

م	المصطلح	ك	النسبة %
١	Data	٥٧	٢٧
٢	Data Science	٣٨	١٨
٣	Information	٣٢	١٥
٤	Other	٩٠	٤٢

يتبين من الجدول رقم (٤) ، أن هناك اختلاف في استخدام المصطلحات الشائعة في أسماء مقررات علم البيانات عينة الدراسة (٢١٣ مقرر) ؛ ولقد لاحظت الباحثة ظهور مصطلح "Data" في ٩٥ مقرر من إجمالي مقررات علم البيانات عينة الدراسة بنسبة ٤٥% حيث جاء في المرتبة الأولي مصطلح "Data" فقط بنسبة ٢٧% فهو يظهر في أسماء ٥٧ مقرر من إجمالي مقررات علم البيانات عينة الدراسة ، منها على سبيل المثال لا الحصر : مقرر "Data Mining: Methods & Applications" بمدرسة علوم المكتبات والمعلومات بجامعة كارولينا الشمالية University of North Carolina ، ومقرر "Foundations of Data Curation" بمدرسة علوم المعلومات بجامعة إلينويز University of Illinois ، ومقرر "Data in Society" في مدرسة دراسات المعلومات بجامعة سيراكيوز Syracuse University ، ومقرر "Data Structures" بجامعة إنديانا Indiana University . وجاء في المرتبة الثانية مصطلح "Data Science" بنسبة ١٨% فهو يظهر في أسماء ٣٨ مقرر من إجمالي مقررات علم البيانات عينة الدراسة ، منها على سبيل المثال لا الحصر : مقرر "Core Methods in Data Science" في مدرسة المعلومات بجامعة واشنطن University of Washington ، ومقرر "Introduction to Applied Data Science" في مدرسة دراسات المعلومات بجامعة سيراكيوز . وتم استخدام مصطلح "Information" في أسماء ٣٢ مقرر بنسبة ١٥% . كما تبين من الجدول رقم (٤) أن هناك ٩٠ مقرر بنسبة ٤٢% من إجمالي مقررات علم البيانات عينة الدراسة تستخدم مصطلحات أخرى للتعبير ضمنياً عن محتواها ، منها على سبيل المثال لا الحصر : مقرر "Algorithm Design" بمدرسة علوم المعلومات بجامعة بيتسبرج University of Pittsburgh ، ومقرر "Advanced Topics in Machine Learning & Social Computing" بمدرسة علوم المعلومات بجامعة إلينويز .

٣. مقررات علم البيانات وفقا للمرحلة الجامعية :

جدول (٥) عدد مقررات علم البيانات في مدارس المكتبات والمعلومات عينة الدراسة طبقاً للمرحلة الجامعية

الرتبة	مدارس المكتبات والمعلومات	البكالوريوس		الماجستير		الدكتوراه	
		العدد	النسبة %	العدد	النسبة %	العدد	النسبة %
١	University of Illinois - School of Information Sciences	٦	٣	١٤	٦,٥	١٥	٧
٢	University of Washington - The Information School	٧	٣,٢	١٣	٦	-	-
٣	University of North Carolina - School of Information and Library Science	٧	٣,٢	١٣	٦	-	-
٤	Syracuse University - School of Information Studies	٧	٣,٢	٢٢	١٠,٣	-	-
٥	University of Michigan - School of Information	٤	٢	١١	٥	١	٠,٥
#٥	University of Texas - School of Information	-	-	٩	٤,٢	-	-
٧	Rutgers University - Department of Library and Information Science	٥	٢,٣	٦	٣	-	-
٨	University of Maryland - College of Information Studies	١٤	٦,٥	١٦	٧,٥	١	٠,٥
٩	Indiana University– Luddy School of Informatics, Computing, and Engineering - Department of Library and Information Science	٥	٢,٣	١٥	٧	-	-
١٠	University of Pittsburgh - School of Information Sciences	٦	٣	١٢	٥,٦	٤	٢
	الإجمالي	٦١	٢٨,٦	١٣١	٦١,٥	٢١	١٠

يتبين من الجدول رقم (٥) ، أن الغالبية العظمي من مقررات علم البيانات تتوافر بمرحلة الماجستير بنسبة ٦١,٥% من إجمالي مقررات علم البيانات المتاحة في جميع مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة ، كما تتوافر مقررات علم البيانات بمرحلة البكالوريوس بنسبة ٢٨,٦% من إجمالي مقررات علم البيانات المتاحة في جميع مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة ما عدا مدرسة المعلومات بجامعة تكساس University of Texas لا تتيح أي مقرر لعلم البيانات على مستوي البكالوريوس ، ويتبين من الجدول رقم (٥) أيضا ، أن مقررات علم البيانات بمرحلة الدكتوراه تتوافر بنسبة ١٠% من إجمالي مقررات علم البيانات عينة الدراسة في أربع جامعات فقط من إجمالي مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة ، وهي مدرسة علوم المعلومات بجامعة إلينويز University of Illinois ، ومدرسة المعلومات بجامعة ميشيجان University of Michigan ، ومدرسة دراسات المعلومات بجامعة ميريلاند University of Maryland ، ومدرسة علوم المعلومات بجامعة بيتسبرج University of Pittsburgh .

كما لاحظت الباحثة أن مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة يتوافر بها أكثر من مقرر دراسي خاص بعلم البيانات في المستوي الدراسي الواحد فبعضها يتاح للبكالوريوس والبعض الآخر يتاح للماجستير والدكتوراه ، فمنها على سبيل المثال لا الحصر مدرسة المعلومات بجامعة ميشيجان يتاح بها ١٦ مقرر حيث تقدم ٤ مقررات في مستوي البكالوريوس ، وهي : مقرر "Data-Oriented Programming" ، ومقرر "Data Exploration" ، ومقرر "Algorithms and Society" ، ومقرر "Data Manipulation" ، كما تقدم مدرسة المعلومات بجامعة ميشيجان ١١ مقرر في مستوي الماجستير ، منها على سبيل المثال لا الحصر : مقرر ""Information Visualization" ، ومقرر "Applied Machine Learning" ، بينما تقدم مقرر واحد للدكتوراه وهو : "Data Mining Methods and Applications" .

كما لاحظت الباحثة أن هناك أكثر من مقرر خاص بعلم البيانات في المستوي الدراسي الواحد ولكنها ذات مستويات مختلفة منها على سبيل المثال في مدرسة دراسات المعلومات بجامعة ميريلاند حيث يتوافر ١٦ مقرر لعلم البيانات على مستوي الماجستير ولكنها ذات مستويات مختلفة ، منها على سبيل المثال لا الحصر : مقرر "Introduction to Data Science" ، ومقرر "Advanced Data Science" .

٤. مقررات علم البيانات وفقا للفئات الموضوعية :

اعتمدت الباحثة في تحديد الفئات الموضوعية لمقررات علم البيانات على كل من التوصيفات الكاملة والتوصيفات المختصرة للمقررات بالبرامج الأكاديمية في مدارس المكتبات والمعلومات بالجامعات الأمريكية عينة الدراسة كما وردت في مواقعها على الإنترنت . وقد تبين أن المقررات التي تتناول موضوع علم البيانات يمكن أن تصنف تحت فئتين أساسيتين ، وهما :

الفئة الأولي : المقررات الأساسية : وهي المقررات التي تتخذ من علم البيانات موضوعاً رئيسياً

الفئة الثانية : المقررات ذات العلاقة : وهي مقررات لا تتناول علم البيانات مباشرة ولكنها تعالجه باعتباره جزءاً من منظومة أكبر ، ويمكن تصنيف هذه المقررات إلى ٦ فئات موضوعية ، هي :

١) مقررات جمع ومعالجة البيانات. ٢) مقررات تحليل ونمذجة البيانات.

٣) مقررات تمثيل البيانات. ٤) مقررات الحوسبة السحابية والنظم الموزعة.

٥) مقررات تطبيقات علم البيانات في مجالات المعرفة المختلفة.

٦) مقررات أخلاقيات البيانات .

جدول (٦) توزيع مقررات الفئة الموضوعية الأولي : المقررات الأساسية لعلم البيانات في مدارس المكتبات والمعلومات عينة الدراسة

الرتبة	مدارس المكتبات والمعلومات	عدد المقررات	النسبة %
١	University of Illinois - School of Information Sciences	٦	٣
٢	University of Washington - The Information School	٥	2
٣	University of North Carolina - School of Information and Library Science	-	-
٤	Syracuse University - School of Information Studies	٢	١
٥	University of Michigan - School of Information	-	-

الرتبة	مدارس المكتبات والمعلومات	عدد المقررات	النسبة %
#٥	University of Texas - School of Information	١	٠,٥
٧	Rutgers University - Department of Library and Information Science	١	٠,٥
٨	University of Maryland - College of Information Studies	٥	٢
٩	Indiana University - Department of Library and Information Science	-	-
١٠	University of Pittsburgh - School of Information Sciences	-	-
إجمالي مقررات علم البيانات (٢١٣ مقرر)		٢٠	٩

يتبين من الجدول رقم (٦) ، أن عدد مقررات الفئة الموضوعية الأولى : المقررات الأساسية لعلم البيانات بلغ ٢٠ مقرراً بنسبة ٩% من إجمالي مقررات علم البيانات (٢١٣ مقرراً) بمدارس المكتبات والمعلومات عينة الدراسة . ولقد لاحظت الباحثة ظهور مصطلح "Data Science" في مسميات ٢٠ مقرر أي بنسبة ١٠٠% من إجمالي مقررات الفئة الأولي . كما تبين أن ٦ مدارس بنسبة ٦٠% من إجمالي مدارس المكتبات والمعلومات عينة الدراسة تحرص على تقديم مقررات الفئة الموضوعية الأولى : المقررات الأساسية لعلم البيانات ، كما احتلت مدرسة المعلومات بجامعة إلينويز المرتبة الأولي فيما يتعلق بعدد مقررات الفئة الأولى لعلم البيانات حيث بلغ عدد المقررات ٦ مقررات بنسبة ٣٠% من إجمالي مقررات هذه الفئة (٢٠ مقرر) وبنسبة ٣% من إجمالي مقررات علم البيانات عينة الدراسة والبالغ ٢١٣ مقرراً ، وجاءت في المرتبة الثانية كل من مدرسة المعلومات بجامعة واشنطن ، ومدرسة دراسات المعلومات بجامعة ميريلاند بعدد ٥ مقررات لكل منهما بنسبة ٢٥% من إجمالي مقررات الفئة الأولى (٢٠ مقرر) وبنسبة ٢% من إجمالي مقررات علم البيانات عينة الدراسة والبالغ ٢١٣ مقرراً .

جدول (٧) توزيع مقررات الفئة الموضوعية الثانية : المقررات ذات العلاقة بعلم البيانات في مدارس المكتبات والمعلومات عينة الدراسة

الرتبة	مدارس المكتبات والمعلومات	الفئة الموضوعية الثانية : المقررات ذات العلاقة												الإجمالي	
		جمع ومعالجة البيانات		تحليل ونمذجة البيانات		الحوسبة السحابية والنظم الموزعة		تطبيقات العالم الحقيقي		أخلاقيات البيانات		تمثيل وتصوير البيانات			
		العدد	%	العدد	%	العدد	%	العدد	%	العدد	%	العدد	%	العدد	%
١	University of Illinois	١٢	٥,٦	٩	٤	١	٠,٥	٤	٢	٢	١	١	٠,٥	٢٩	١٤
٢	University of Wash-ington	٩	٤	٢	١	-	-	٢	١	-	-	٢	١	١٥	٧
٣	University of North Carolina	٩	٤	٨	٣,٧	١	٠,٥	-	-	-	-	2	١	٢٠	٩
٤	Syracuse University	٧	٣	٩	٤	٢	١	٣	٦	١	٠,٥	2	١	٢٧	١٢,٥

الرتبة	مدارس المكتبات والمعلومات	الفئة الموضوعية الثانية : المقررات ذات العلاقة										الإجمالي	
		جمع ومعالجة البيانات		تحليل ونمذجة البيانات		الحوسبة السحابية والنظم الموزعة		تطبيقات العالم الحقيقي		أخلاقيات البيانات		تمثيل وتصوير البيانات	
		العدد	%	العدد	%	العدد	%	العدد	%	العدد	%	العدد	%	العدد	%
٥	University of Michigan	٧	٣	٨	٣,٧	-	-	-	-	-	-	١	٠,٥	١٦	٧,٥
٦	University of Texas	٢	١	٥	٢,٣	-	-	-	-	-	-	١	٠,٥	٨	٤
٧	Rutgers University	٤	٢	٣	١,٤	-	-	١	٠,٥	-	-	٢	١	١٠	٥
٨	University of Maryland	١٠	٤,٧	٩	٤	-	-	-	-	٣	١,٤	٤	٢	٢٦	١٢
٩	Indiana University	٩	٤	٨	٣,٧	٣	١,٤	-	-	-	-	١	٠,٥	٢١	١٠
١٠	University of Pitts-burgh	٨	٣,٧	٨	٣,٧	٣	١,٤	-	-	-	-	٢	١	٢١	١٠
الإجمالي		٧٧	٣٦,١	٦٩	٣٢	١٠	٥	١٣	٦,٥	٦	٣	١٨	٨	١٩٣	٩١

كما يتبين من الجدول رقم (٧) أن مقررات الفئة الموضوعية الثانية (المقررات ذات العلاقة بعلم البيانات) تمثل الغالبية العظمي من إجمالي المقررات عينة الدراسة ، حيث تبلغ ١٩٣ مقرراً بنسبة ٩١% . كما احتلت مدرسة المعلومات بجامعة إلينويز المرتبة الأولي فيما يتعلق بعدد مقررات هذه الفئة (٢٩ مقرر) بنسبة ١٤% من إجمالي مقررات علم البيانات عينة الدراسة والبالغ ٢١٣ مقرراً وبنسبة ١٥% من إجمالي مقررات الفئة الموضوعية الثانية (١٩٣ مقرر) ، وجاءت في المرتبة الثانية مدرسة المعلومات بجامعة سيراكيوز حيث بلغ عدد المقررات ٢٧ مقرر بنسبة ١٢,٥% من إجمالي مقررات علم البيانات عينة الدراسة وبنسبة ١٤% من إجمالي مقررات هذه الفئة (١٩٣ مقرر) ، ثم جاءت في المرتبة الثالثة مدرسة دراسات المعلومات بجامعة ميريلاند حيث بلغ عدد المقررات ٢٦ مقرر بنسبة ١٢% من إجمالي مقررات علم البيانات عينة الدراسة وبنسبة ١٣,٤% من إجمالي مقررات هذه الفئة (١٩٣ مقرر) .

كما لاحظت الباحثة أن مقررات "جمع ومعالجة البيانات" جاءت في المرتبة الأولى حيث تبلغ ٧٧ مقرراً بنسبة ٤٠% من إجمالي مقررات الفئة الموضوعية الثانية ذات العلاقة (١٩٣ مقرراً) وبنسبة ٣٦,١% من إجمالي المقررات عينة الدراسة (٢١٣ مقرر) . وجاءت في المرتبة الثانية مقررات "تحليل ونمذجة البيانات" حيث تبلغ ٦٩ مقرراً بنسبة ٣٥,٧% من إجمالي مقررات الفئة الموضوعية الثانية ذات العلاقة (١٩٣ مقرراً) وبنسبة ٣٢% من إجمالي المقررات عينة الدراسة (٢١٣ مقرر) ؛ ويشير ارتفاع نسبة تلك المقررات إلى اهتمام مدارس المكتبات والمعلومات عينة الدراسة بتغطية الموضوعات المتعلقة بجمع ومعالجة البيانات ، إلى جانب موضوعات تحليل ونمذجة البيانات .

كما تبين أن مقررات "أخلاقيات البيانات" جاءت في المرتبة السادسة والأخيرة حيث تبلغ ٦ مقررات بنسبة 3% من إجمالي مقررات الفئة الموضوعية الثانية ذات العلاقة (١٩٣ مقرراً) وبنفس النسبة ٣% من

إجمالي المقررات عينة الدراسة (٢١٣ مقرر) ؛ ويشير انخفاض نسبة تلك المقررات إلى عدم اهتمام مدارس المكتبات والمعلومات عينة الدراسة بتغطية القضايا القانونية والأخلاقية المتعلقة بإدارة البيانات .

كما تبين من خلال حصر وتحليل المقررات للبرامج الدراسية في مدارس المكتبات والمعلومات عينة الدراسة تقديم ٥ من مدارس المكتبات والمعلومات برامج مستقلة لمنح درجة الماجستير في تخصص علم البيانات ، وهي : مدرسة المعلومات بجامعة إلينويز ، ومدرسة المعلومات بجامعة سيراكيوز ، ومدرسة المعلومات بجامعة ميتشجان ، وقسم علوم المكتبات والمعلومات بجامعة روتجرز Rutgers University ، وقسم علم المعلومات والمكتبات بجامعة إنديانا .

٥. أهداف مقررات علم البيانات :

تم حصر أهداف مقررات علم البيانات المتاحة بالتوصيفات سواء الكاملة أو المختصرة ، للتعرف على الأهداف الرئيسية التي تسعي تلك المقررات للتوصل إليها ، ولقد اعتمدت الباحثة في تجميع هذه الأهداف على كل من العناصر التالية : توصيف المقرر الدراسي ، وأهدافه ، وقد تم توحيد بعض صيغ الأهداف ذات المغزى الواحد والمتاحة بصيغ مختلفة ، وتتمثل أهداف مقررات علم البيانات في الجدول التالي :

جدول (٨) أهداف مقررات علم البيانات مدارس المكتبات والمعلومات عينة الدراسة

م	الأهداف	التكرارات	النسبة %
١	استخدام لغات برمجية مفتوحة المصدر ، مثل : بايثون Python وآر R لمعالجة وتحليل مجموعات ضخمة من البيانات المهيكلة وغير المهيكلة في العالم الحقيقي	٧١	٣٣,٣
٢	نمذجة البيانات وتصميم مخططات قواعد البيانات باستخدام لغة استعلام SQL	٤٦	٢١,٥
٣	استخدام الخوارزميات وهياكل البيانات المختلفة والشائعة في علوم الحاسب للتحليل والتنقيب في البيانات واستخلاص المعرفة من البيانات المهيكلة وغير المهيكلة ، مثل : خوارزميات الترتيب والبحث ، المصفوفات ، تراكيب البيانات ، الدوال ، أشجار القرار	٣٩	١٨,٣
٤	تطبيق أساليب تنقيب البيانات ، وتشمل : تصنيف وتحويل البيانات ، تنظيف البيانات ، والدمج ، والاختزال	٣٦	١٧
٥	استخدام أدوات وتقنيات تحليل البيانات ، مثل : البرمجة النصية ، تحليل النص ، لغة الاستعلام المنظمة ، التعبيرات العادية ، الرسوم البيانية ، وطرق التجميع لاستكشاف البيانات	٣٣	١٥,٥
٦	تطبيق تقنيات علم البيانات (مثل : الخوارزميات ، لغات البرمجة ، تقنيات التعلم الآلي) على مجموعة من البيانات غير المهيكلة تمثل مشكلة في العالم الحقيقي في القطاعات المختلفة	٢٩	١٤
٧	اختيار وتطبيق الأساليب الإحصائية والرياضية المناسبة لنمذجة وتحليل البيانات وإدارتها ، وتشمل : الاحصاء الوصفي والاستدلالي ، القياس ، تحليل العينات ، التحليل الخطي ، تحليل الانحدار والسلاسل الزمنية ، النماذج الإحصائية الخطية ، نماذج الاحتمالات وتطبيق الفروض	٢٨	١٣
٨	استخدام أدوات تمثيل البيانات في المخططات والتمثيلات المرئية لدعم واتخاذ القرار	٢٥	١٢

م	الأهداف	التكرارات	النسبة %
٩	التعرف على الآثار الأخلاقية والقانونية والاجتماعية التي تنشأ من جمع البيانات ومعالجتها وتحليلها ، مثل : الخصوصية ، تبادل البيانات العلمية ونشرها وحفظها ، ومراقبة جودة ونشر البيانات	٢٣	١١
١٠	جمع وتخزين واسترجاع (جمع وإدارة وتنظيم البيانات) كميات هائلة من البيانات الناشئة من مصادر متعددة ، بما في ذلك وسائل التواصل الاجتماعي ، وشبكة الإنترنت	٢١	١٠
١١	تنظيم وإدارة البيانات في القطاعات المختلفة	٢١	١٠
١٢	التعرف على المفاهيم الأساسية في علوم البيانات ، وتشمل : تحليل البيانات ، تنقيب البيانات ، نمذجة البيانات ، قواعد البيانات ، إدارة البيانات الضخمة ، تقنيات التعلم الآلي ، تصوير البيانات ، تحليل البيانات	١٤	٦,٥
١٣	تحديد وتقييم أنواع مختلفة من البيانات وتفسيرها واستخلاص النتائج منها لدعم اتخاذ القرار	١٠	٥
١٤	اساسيات الحوسبة السحابية والنظم الموزعة	١٠	٥
١٥	التعرف على أنواع البيانات وكيفية تجهيزها وعرضها في أشكال موحدة وهياكل منظمة	٦	٣
١٦	اختيار هياكل البيانات المناسبة لتخزين وتنظيم البيانات المتنوعة	٥	٢,٣
١٧	استخدام أساليب تقييم جودة البيانات وربطها بتقنيات تمثيل البيانات	٤	٢
١٨	التعامل مع نظم وطرق تخزين البيانات المختلفة ، وإيجاد العلاقات ما بين تلك البيانات	٤	٢
١٩	التعرف على المهارات الأساسية والجوانب المختلفة لمهنة عالم البيانات	٣	١,٤

يتبين من الجدول رقم (٨) أنه يوجد تباين في أهداف مقررات علم البيانات التي تسعي لتحقيقها حيث يعد أكثر الأهداف توافراً في مقررات علم البيانات هو "استخدام لغات برمجية مفتوحة المصدر ، مثل : بايثون Python وآر R لمعالجة وتحليل مجموعات ضخمة من البيانات المهيكلة وغير المهيكلة في العالم الحقيقي" وهو متاح في ٧١ مقرر أي بنسبة ٣٣,٣% من إجمالي المقررات عينة الدراسة (٢١٣ مقرر) ، ثم يليه كل من : هدف " نمذجة البيانات وتصميم مخططات قواعد البيانات باستخدام لغة استعلام SQL" حيث يتوافر في ٤٦ مقرر بنسبة ٢١,٥% وهدف "استخدام الخوارزميات وهياكل البيانات المختلفة والشائعة في علوم الحاسب للتحليل والتنقيب في البيانات واستخلاص المعرفة من البيانات المهيكلة وغير المهيكلة ، مثل : خوارزميات الترتيب والبحث ، المصفوفات ، تراكيب البيانات ، الدوال ، أشجار القرار" حيث يتوافر في ٣٩ مقرر أي بنسبة ١٨,٣% ، وهدف "تطبيق أساليب تنقيب البيانات ، وتشمل : تصنيف وتحويل البيانات ، تنظيف البيانات ، والدمج ، والاختزال" حيث يتوافر في ٣٦ مقرر بنسبة ١٧% ؛ والحقيقة أن هذه الأهداف الأربعة من العناصر الأساسية التي لا غني عنها لإعداد برنامج لعلم البيانات .

وعلى الجانب الأخر يوجد العديد من الأهداف الفرعية لعلم البيانات والتي تتوافر في ثلاث مقررات أو أربعة منها على سبيل المثال لا الحصر هدف "التعرف على المهارات الأساسية والجوانب المختلفة لمهنة عالم البيانات" حيث يتوافر في ٣ مقررات بنسبة ١,٤% .

٦. طرق التدريس بمقررات علم البيانات :

حاولت الباحثة تحديد طرق التدريس في مقررات علم البيانات بمدارس المكتبات والمعلومات عينة الدراسة إلا أن غالبية توصيفات المقررات (١٩٠ مقرر بنسبة ٨٩,٢%) هي توصيفات مختصرة مما حال دون إتمام الدراسة على هذه التوصيفات ، ولقد اقتصرت الدراسة هنا على تحليل ٢٣ توصيف لمقررات علم البيانات وهي التوصيفات الكاملة التي اشارت إلى طرق التدريس بالمقرر (بنسبة ١١% من إجمالي مقررات عينة الدراسة) . وقد تبين من خلال تحليل مضمون تلك المقررات أن جميع مقررات علم البيانات (٢٣ مقرر بنسبة ١١%) تستخدم أكثر من طريقة لتدريس المقرر ؛ وتشمل هذه الطرق : محاضرات في قاعات الدراسة ويتم تعزيزها بالأمثلة والمناقشات والعصف الذهني لتعزيز التفاعل ما بين المحاضر والطالب ، إلى جانب محاضرات مسجلة بالفيديو مع إمكانية التواصل مع المحاضر بعد مشاهدة الفيديو . ويلاحظ أن غالبية مقررات علم البيانات بمدارس المكتبات والمعلومات عينة الدراسة تعتمد على أسلوب حل المشكلات من خلال عرض مشاكل حقيقية في المجالات المختلفة (الصحة ، التعليم ، الاقتصاد ، ...إلخ) والتوصل إلى حلول لها من خلال تطبيقات علم البيانات .

٧. التكليفات والمهام الدراسية بمقررات علم البيانات :

تبين من خلال دراسة تحليل مضمون ٢٣ توصيف لمقررات علم البيانات وهي التوصيفات الكاملة التي اشارت إلى التكليفات والمهام الدراسية بالمقرر (بنسبة ١١% من إجمالي مقررات عينة الدراسة) ، أن التكليفات والمهام الدراسية لجميع مقررات علم البيانات (٢٣ مقرر بنسبة ١١%) تتراوح ما بين تطبيق أو بناء خوارزميات تعليم الآلة ، أو تحليل مجموعة من البيانات الضخمة الحقيقية في المجالات المختلفة ، أو إعداد مشروع لأحد المشكلات في المجالات المختلفة (الصحة ، التعليم ، الاقتصاد ، ...إلخ) والتوصل إلى حلول لها من خلال تطبيقات علم البيانات .

٨. أساليب تقييم الطلاب بمقررات علم البيانات :

تبين من خلال دراسة تحليل مضمون ٢٣ توصيف لمقررات علم البيانات وهي التوصيفات الكاملة التي اشارت إلى التكليفات والمهام الدراسية بالمقرر (بنسبة ١١% من إجمالي مقررات عينة الدراسة) ، أن هناك تنوع في أساليب تقييم الطلاب ؛ وتشمل : التقييم الذاتي من خلال الاعتماد على الأقران المسجلين في المقرر في تصحيح التكليفات ورصد الدرجات لبعضهم البعض مع إشراف المحاضر . وفي نهاية كل مقرر يتم تكليف الطلاب بمشروع عملي للتأكد من استيعابهم لجميع المهارات المستهدفة من المقرر .

٩. المتطلبات التكنولوجية بمقررات علم البيانات :

تعد المتطلبات التكنولوجية من الركائز الأساسية اللازمة لتعليم مقررات علم البيانات ، فهي من العوامل الأساسية لنجاح أهداف المقرر الدراسي ، وقد تبين من خلال دراسة تحليل المضمون لتوصيفات مقررات علم البيانات أن المتطلبات التكنولوجية لتدريس علم البيانات تمثلت فيما يلي :

١/٩. الأجهزة :

ضرورة توافر معمل حاسب آلي أو إمكانية أن يحضر الطلاب أجهزة الحاسب المحمولة الخاصة بهم والتي تعمل بنظام تشغيل الويندوز Windows أو نظام ماك Mac OS X أو نظام لينكس Linux حتي يستطيع الطالب التطبيق العملي لما يتم دراسته . أيضا اوصت العديد من مقررات علم البيانات ، مثل مقرر اكتشاف علوم البيانات "Data Science Discovery" بمدرسة علوم المعلومات بجامعة إلينويز باستخدام أجهزة نظم **الاستجابة الشخصية** iClicker ويهدف استخدام هذه التقنية في القاعات الدراسية إلى تحسين العملية التعليمية وزيادة التفاعل بين عضو هيئة التدريس والطلبة عن طريق طرح الأسئلة أو عمل

امتحانات قصيرة خلال المحاضرة وإظهار النتائج فوريا وبوقت قياسي ؛ وبالتالي تساعد تلك التقنية أعضاء هيئة التدريس على قياس مدى استيعاب طلابهم الفعلي للمادة العلمية أثناء المحاضرة وكذلك الاحتفاظ بسجل عن مستوي الطالب خلال الفصل الدراسي كتغذية مرتدة . بالإضافة إلى ذلك ضرورة توافر بعض الأجهزة الأخرى مثل : أجهزة عرض البيانات Data Show وأجهزة الفيديو Video .

٩/٢. البرامج :

تعد لغات وأدوات البرمجة مثل أوامر UNIX و MySQL و Python و R لجمع البيانات وتنظيفها ومعالجتها وتحليلها من الركائز الأساسية اللازمة لتعليم مقررات علم البيانات . ويعتبر من أهم المتطلبات أيضا في مقررات علم البيانات تحميل وتنصيب مكتبات وأدوات لغة البرمجة بايثون ومن أشهرها : Numpy, Pandas, Scipy, Matplotlib, Scikit-learn وهي عبارة عن تراكيب وأدوات برمجية لتحليل ومعالجة البيانات وتعليم الآلة حيث تقدم إمكانيات هائلة في معالجة البيانات والصور والعمليات الحسابية المعقدة ، وكذلك تقدم مكتبات بايثون مجموعة واسعة من الخوارزميات الخاصة بتعليم الآلة ، كما تبين في العديد من مقررات علم البيانات عينة الدراسة ، مثل : مقرر "Intensive Python Programming for Information Students" في مدرسة المعلومات بجامعة ميتشجان . هذا بالإضافة إلى ضرورة توافر برامج خاصة بالتعامل مع البيانات الضخمة ، مثل برنامج جبتر نوت Jupyter Notebooks [٣٥] وهو عبارة عن برنامج مفتوح المصدر للتعامل مع الأكواد البرمجية والرسومات التوضيحية والإحصائيات وهو يدعم عدة لغات للبرمجة مثل بايثون وآر ، إلى جانب الاعتماد على برنامج IPython Notebooks حيث يستخدم لتعلم لغة بايثون وخوارزميات تعلم الآلة ، وأشارت العديد من مقررات علم البيانات عينة الدراسة أن مثل هذه البرامج متاحة على موقع خدمة GitHub وهي خدمة لإستضافة وإدارة المشاريع البرمجية ، مثل مقرر "Intermediate Programming" في مدرسة المعلومات بجامعة ميتشجان ، ومقرر "Advanced Data Science" مدرسة المعلومات بجامعة إلينويز .

٩/٣. البريد الإلكتروني :

يتطلب في العديد من المقررات الدراسية توافر بريد إلكتروني لكل طالب حتي يتم التواصل معه .

١٠. تصور مقترح لمقرر دراسي لعلم البيانات :

لاحظت الباحثة أن جميع أقسام المكتبات والمعلومات في الجامعات المصرية على مستوي الليسانس والماجستير والدكتوراه تفتقر لمقرر خاص بعلم البيانات يهدف إلى تهيئة خريجي أقسام المكتبات والمعلومات للعمل في مجال علم البيانات . لذا تقترح توافر مقرر لعلم البيانات في مرحلة الليسانس يهدف إلى تأهيل خريجي أقسام المكتبات والمعلومات للعمل في مجال علم البيانات . وتتسم ملامح المقرر المقترح بما يلي :

نموذج مقترح لتوصيف مقرر مقدمة في علم البيانات

جامعة / عين شمس
كلية / الآداب
قسم : المكتبات والمعلومات

بيانات المقرر :		
الرمز الكودي:	اسم المقرر : مقدمة في علم البيانات	المستوى: الثامن
التخصص: المكتبات والمعلومات	عدد الوحدات الدراسية : نظري (٣) عملي : (١)	

٢- هدف المقرر	تعريف الطالب بمفهوم علم البيانات ، وكذلك أهم لغات البرمجة لعلم البيانات مثل : بايثون ، وكيفية بناء وتطبيق خوارزميات تعليم الآلة الشائعة ، بالإضافة إلى استخدام أدوات وبرمجيات التمثيل المرئي للبيانات .
٣- المستهدف من تدريس المقرر:	
أ- المعلومات والمفاهيم :	في نهاية المقرر يكون الطالب قادراً على أن : أ١. يتعرف على مفهوم علم البيانات وأهميته أ٢. يلم بأنواع البيانات وخصائصها أ٣. يتعرف على التقنيات الأساسية في علم البيانات ، وتشمل : تحليل البيانات ، نمذجة البيانات ، قواعد البيانات العلائقية ، معالجة البيانات ، تمثيل البيانات ، جودة البيانات . أ٤. يتعرف على المهارات اللازمة لعالم البيانات
ب – المهارات الذهنية :	في نهاية المقرر يكون الطالب قادراً على أن : ب١. يدرك الفروق بين علم البيانات وتحليل البيانات والتعلم الآلي . ب٢. يجمع البيانات من مجموعة متنوعة من المصادر بما في ذلك شبكة الإنترنت . ب٣. يجد بدائل الحلول المختلفة للمشكلات في المجالات المختلفة . ب٤. يدرك الآثار القانونية والأخلاقية المترتبة على البيانات .
ج – المهارات المهنية الخاصة بالمقرر :	في نهاية المقرر يكون الطالب قادراً على أن : ج١. يستخدم لغة البرمجة بايثون لمعالجة مجموعة من البيانات الضخمة وتمثيلها . ج٢. يطبق خوارزميات لتحليل البيانات واستخراج المعرفة منها ج٣. يقيس جودة أداء الخوارزميات وتحسين نتائجها .
د- المهارات العامة :	في نهاية المقرر يكون الطالب قادراً على أن : د١. يعد مشروع لأحد المشكلات الواقعية في أحد المجالات وتقديم حلول لها باستخدام الخوارزميات .

٤- محتوي المقرر :	الوحدة الأولى: التقنيات الأساسية في علم البيانات ، وتشمل : تحليل البيانات ، تعليم الآلة ، نمذجة البيانات ، قواعد البيانات العلائقية ، معالجة البيانات ، تمثيل البيانات ، جودة البيانات ، ... إلخ الوحدة الثانية: وظائف ومهارات عالم البيانات الوحدة الثالثة: اساسيات لغة البحث المهيكل SQL ، ولغة البرمجة بايثون مع التركيز على أشهر حزم بايثون ، مثل : Numpy, Pandas الوحدة الرابعة: خوارزميات التعليم الآلي الشائعة الوحدة الخامسة: مصادر مفتوحة المصدر لتحليل البيانات والتوصل إلى نتائج . الوحدة السادسة: أدوات وبرمجيات التمثيل المرئي للبيانات . الوحدة السابعة : تطبيقات عملية لمشكلات واقعية في المجالات المختلفة . الوحدة الثامنة : قوانين وأخلاقيات التعامل مع البيانات في المجالات المختلفة
٥- أساليب التعليم والتعلم :	المحاضرات النظرية . العصف الذهني . الحوار والنقاش . التكليفات ودراسات تطبيقية .
٦- أساليب التعليم والتعلم للطلاب ذوي القدرات المحدود :	لا يوجد وفي حالة تواجدهم يتم تخصيص ساعات مكتبية لهم
٧- تقويم الطلاب :	
أ- الأساليب المستخدمة :	الحضور والمشاركة تطبيقات وتدريبات امتحان تحريري (الأسبوع السابع) الامتحان النهائي
ب – التوقيت:	امتحان نظري نهاية الفصل الدراسي الثاني . التكليفات على مدار الفصل الدراسي الثاني
ج – توزيع الدرجات :	(٥) درجات للحضور والمشاركة (٥%) (١٠) درجات للتطبيقات والتدريبات (١٠%) (١٠) درجات للامتحان التحريري (الأسبوع السابع) (١٠%) (٧٥) درجة للامتحان النهائي (٧٥%)

١١. نتائج الدراسة :

توصلت الدراسة إلى العديد من النتائج من أهمها :

١/١١. يتوافر ٢١٣ مقرر خاص بعلم البيانات بنسبة ١٣,٦% من إجمالي المقررات المتاحة في مدارس المكتبات والمعلومات عينة الدراسة .

٢/١١. يتوافر ١٤٩ مقرر بنسبة ٧٠% من مقررات علم البيانات عينة الدراسة كمقرر اختياري .

٣/١١. يتوافر الغالبية العظمي من مقررات علم البيانات في مرحلة الماجستير بنسبة ٦١,٥% من إجمالي مقررات علم البيانات عينة الدراسة .

٤/١١. يتوافر مصطلح Data Science في ٣٨ مقرراً بنسبة ١٨% من إجمالي مقررات علم البيانات عينة الدراسة .

٥/١١. تصنف مقررات علم البيانات عينة الدراسة إلى فئتين أساسيتين ، وهما الفئة الأولي : المقررات الأساسية وتبلغ ٢٠ مقرراً بنسبة ٩% ، والفئة الثانية : المقررات ذات العلاقة وتبلغ ١٩٣ مقرراً بنسبة ٩١% .

٦/١١. اتضح أن هناك اختلافاً في أهداف مقررات علم البيانات ، ولكن يعد أكثر الأهداف تكراراً في مقررات عينة الدراسة هو : استخدام لغات برمجية مفتوحة المصدر ، مثل : بايثون Python وآر R لمعالجة وتحليل مجموعات ضخمة من البيانات المهيكلة وغير المهيكلة في العالم الحقيقي .

٧/١١. اتضح أن هناك تنوعاً في طرق تدريس مقررات علم البيانات منها المحاضرات ، والتطبيقات والتكليفات العملية ، ... إلخ .

٨/١١. اتضح أن هناك تنوعاً في أساليب تقييم الطلاب بمقررات علم البيانات ، مثل : التقييم الذاتي ، وامتحانات اسبوعية ، وامتحانات نهائية ، وإعداد مشروع ، ...إلخ .

٩/١١. اتضح أن هناك متطلبات تكنولوجية لتدريس مقررات علم البيانات ، منها : معمل حاسب آلي ، تحميل وتنصيب مكتبات وأدوات لغة البرمجة بايثون ، ...إلخ .

١٢. توصيات الدراسة :

١/١٢. إعادة النظر في أهداف ومحتوي مقررات البرامج الحالية لأقسام المكتبات والمعلومات ، والعمل على توافقها مع متطلبات سوق العمل وخاصة في مجال علم البيانات .

٢/١٢. إضافة بعض المقررات الخاصة بعلم البيانات إلى برامج أقسام المكتبات والمعلومات المصرية ، مثل : مثل : مقدمة في علم البيانات ، والبرمجة لعلم البيانات ، وتمثيل البيانات ، ومعالجة البيانات ، والخوارزميات وهياكل البيانات .

* مصادر الدراسة :

1. Louridas, P. and Ebert, C. (2013) . “Embedded Analytics and Statistics for Big Data” . IEEE Software, V.30 , PP.33-39 .

2. Tang, Rong & Sae-Lim, Watinee . (2016) . “Data science programs in U.S. higher education: Anexploratory content analysis of program descrip-

tion,curriculum structure, and course focus" . Education for Information, V.32 , p.270

3. Ibid. PP.269–290

٤. وانج ، لين (٢٠١٩) . توأمة علم البيانات مع علم المعلومات بمدارس علم المكتبات والمعلومات . ترجمة : محمد إبراهيم حسن الصبحي . المجلة العلمية للمكتبات والوثائق والمعلومات . مج١ ، ع٢ . ص٣١١-٣١٣

5. Tang, Rong & Sae-Lim, Watinee . (2016). op.cit . PP.269–290

6. U.S. News and World Report (2017) . Available at :

https://www.usnews.com/best-graduate-schools/top-library-information-science-programs/library-information-science-rankings . Last accessed : 1/1/2019

7. George,G et.al (2016) . "Big data and data science methods for management research" . Academy Management Journal, V.59, N.5 : PP.1493-1507

8. Song, Il-Yeol & Zhu, Yongjun . (2016) . "Big data and data science: what should we teach?" . Expert Systems, V.33, N.4, P.364-373

9. Chao, Lemen & Lu Xiaobin.(2017) . "Data Science and Its Implication on Information Science" . Journal of the China Society for Scientific and Technical Information, V.36, N.8 : PP.761-771

10. Grossman, Robert L. & Siegel, Kevin P.. (2014). "Organizational Models for Big Data and Analytics" . Journal of Organization Design (JOD), V.3, N.1, pp.20-25

١١. أحمد أحمد المزين . (٢٠١٩) . البيانات الضخمة والتكامل المعرفي في المكتبات الوطنية : مكتبة الكويت نموذجاً . المجلة العلمية للمكتبات والوثائق والمعلومات . مج١ ، ع٢ : ص٢٤٧-٢٨٢

١٢. خالد عتيق سعيد عبد الله ، عبد الله بن سالم الهنائي (٢٠١٨) . البيانات الضخمة في مكتبات جامعة السلطان قابوس واقعها وأثر دور المدراء كمتغير وسيط للاستفادة منها في تحسين الخدمات . المجلة العراقية لتكنولوجيا المعلومات . مج٩ ، ع١ : ص٢٣-٥٢

١٣. عبدالرحمن فراج . (٢٠١٨) . البيانات الضخمة في تخصص المكتبات والمعلومات. أحوال المعرفة. ع٩٣. ص ٧٦-٧٩.

١٤. زينب بن الطيب ، سليمان بن إبراهيم الرباعي . (٢٠١٨) . الأدوار الجديدة لأخصائي المعلومات للتعامل مع البيانات الضخمة . مجلة دراسات المعلومات والتكنولوجيا . ع٢ : ص١-١٥

١٥. يونس أحمد إسماعيل الشوابكة . الوعي بمفهوم البيانات الضخمة Big Data لدى العاملين في المكتبات الأكاديمية : دراسة حالة لمكتبة الجامعة الأردنية . في : المؤتمر السنوي الرابع والعشرين لجمعية المكتبات المتخصصة فرع الخليج العربي بعنوان "البيانات الضخمة وآفاق استثمارها : الطريق نحو التكامل المعرفي" ٦-٨ مارس ٢٠١٨ (مسقط) . ص١-١٣

١٦. عمر حسن فتوح . (٢٠١٧) . تحليلات البيانات الضخمة ودورها في دعم اتخاذ القرار بالمكتبات: دراسة استكشافية . مجلة المكتبات والمعلومات العربية . ع٤(أكتوبر ٢٠١٧ .)

17. Kross, Sean et al. (2019) ."The democratization of data science education" . The American Statistician, DOI:10.1080/00031305.2019.1668849

18. Mbwilo, Betty & Kimaro, Honest & Justo, Godfrey . (2019) . "Data Science Postgraduate Education at University of Dar es Salaam in Tanzania: Current Demands and Opportunities" . IN: International Conference on Social Implications of Computers in Developing Countries . Part of the IFIP Advances in Information and Communication Technology book series (IFIPAICT, volume 552) ,pp.349-360

19. Mikroyannidis, Alexander et al. (2018) . "Designing and Delivering a Curriculum for Data Science Education Across Europe" . Advances in Intelligent Systems and Computing, v.716, pp.540-550

20. Wang, Lin. (2018). "Twinning data science with information science in schools of libraryand information science" . Journal of Documentation, V.74, N.6, PP.1243 – 1257

21. Semeler, Alexandre Ribas & Pinto, Adilson Luiz & Rozados, Helen Beatriz Frota . (2017) . "Data science in data librarianship: Core competencies of a data librarian" . Journal of Librarianship and Information Science, P.1-10

22. Zhang, Jilong et.al . (2017). "The development of data science education in China from the LIS perspective" . International Journal of Librarianship, V.2, N.2, PP.3-17

23. Marshall, Linda & Eloff, Jan H.P.. (2016)."Towards an Interdisciplinary Master's Degree Programme in Big Data and Data Science: A South African Perspective" .SACLA, CCIS 642, pp.131–139. DOI: 10.1007/978-3-319-47680-3 13

24. Mills, Robert J. & Chudoba, Katherine M. & Olsen, David H. (2016) . "IS Programs Responding to Industry Demands for Data Scientists: A Comparison between 2011-2016" . Journal of Information Systems Education, V.27, N.2, PP.131-140

25. Song, Il-Yeol & Yongjun Zhu . (2016) . op.cit., PP.364-373

26. Tang, Rong & Watinee Sae-Lim . (2016) . op.cit., PP.269–290

27. Yi, Myongho . (2016) . "A Study on the Curriculums of Data Science" . Journal of the Korean BIBLIA Society for library and Information Science, V.27, N.1, PP.263-290

28. Aasheim, Cheryl L. et al. (2015) . "Data Analytics vs. Data Science: A Study of Similarities and Differences in Undergraduate Programs Based on Course

Descriptions" . Journal of Information Systems Education, V.26, N.2, PP.103-115

29. Song, Il-Yeol & Yongjun Zhu . (2016) . op.cit., PP.364-373
30. Ibid. P.366
31. Semeler, Alexandre Ribas & Adilson Luiz Pinto & Helen Beatriz Frota Rozados . (2017) . op.cit., P.4
32. MISNEVS, Boriss & YATSKIV, Irina.(2016). "Data Science: Professional Requirements and Competence Evaluation" . Baltic J. Modern Computing, V.4, N.3, PP.444-445
33. Semeler, Alexandre Ribas & Adilson Luiz Pinto & Helen Beatriz Frota Rozados . (2017) . op.cit., P.4
34. Aasheim, Cheryl L. et al. (2015) . op.cit., PP.105-106
35. Bradford, Laurence . (2019) . "What is GitHub, and Why Should I Use It?". Available at: www.thebalancecareers.com . Last accessed : 1/1/2019