مركز الإحصاء والتنافسية
Statistics and Competitiveness Center
حكومة عجمان
Government of Ajman

مفاهيم عامة عن البيانات الضخمة 2021

إصدار فبراير 2021

www.scc.ajman.ae

جميع الحقوق محفوظة © مركز الإحصاء والتنافسية

حكومة عجمان - الإمارات العربية المتحدة @ 2021

يمنع نسخ أو استعمال أي جزء من هذا الكتاب من قبل أي شخص أو شركة أو جهة بأية وسيلة تصويرية أو إلكترونية أو ميكانيكية بما في ذلك التسجيل الفوتغرافي والتسجيل على أقراص مقروءة أو بأية وسيلة نشر أخرى بما فيها حفظ المعلومات و استرجاعها دون الحصول على موافقة مسبقة صادرة من مركز عجمان للإحصاء والتنافسية،

حكومة عجمان، دولة الإمارات العربية المتحدة.

في حالة الاقتباس يرجى الإشارة إلى المطبوعة كالتالي:

مركز عجمان للإحصاء والتنافسية - حكومة عجمان

مفاهيم عامة عن البيانات الضخمة 2021

رقم الإصدار 1 - فبراير 2021

للتواصل وطلب البيانات الإحصائية يرجى التواصل:

مركز عجمان للإحصاء والتنافسية

البريد الإلكتروني: info.scc@ajman.ae

رقم الهاتف: +971 6 701 6770

الموقع الإلكتروني: scc.ajman.ae

ص.ب: 6556، عجمان - دولة الإمارات العربية المتحدة

@sccajman

التعريف بمركز عجمان للإحصاء والتنافسية

تم إنشـــاء "مركز عجمان للاحصاء و التنافســـية " استناداً للمرسوم الأميري رقم (28) لسنة 2017 .

ويعتبر المركز هو الجهة المختصة محلياً في إمارة عجمان والمصدر الرئيس والمرجع الوحيد فيها في الشؤون الإحصائية و التنافسية المنصوص عليها في هذا المرسوم.

يهدف المركز إلى تحقيق الغايات التالية:

1. تنظيم وتطوير العمل الإحصائي بما يحقق مصالح الدولة والإمارة.
2. بناء نظام إحصائي محلي متكامل.
3. رفع القدرة التنافسية للإمارة في مختلف القطاعات.
4. المســـاهمة فـــي تعزيـــز مكانـــة الإمـــارة فـــي تقاريـــر التنافســـية المحلية والعالمية.
5. دعـــم منظومـــة إتخـــاذ القـــرار فـــي الحكومـــة ببيانـــات ومعلومـــات دقيقة وحديثة.

الرؤية

بالمعرفة نعزز مستقبل عجمان.

الرسالة

الإرتقـــاء بالعمـــل الإحصائي والتنافســـي من خـــلال تطبيق أفضـــل الممارســـات بإتباع المنهجيات العلمية الإحصائية والمعايير الموصى بها دولياً لتلبي إحتياجات مستخدمي البيانات ومتخذى القرار في الإمارة.

القيم

الجودة / الحيادية / الإحترافية / الموثوقية / الإبداع والابتكار / السرية / الشفافية

مفاهيم عامة عن البيانات الضخمة 2021

المحتويات

مقدمة 6
1- مفهوم البيانات الضخمة 7
2- أهمية البيانات الضخمة 7
3 - أنواع البيانات الضخمة 8
1.3 المنظمة 8
2.3 غير المنظمة 9
3.3 شبه منظمة 9
4-السمات التي تحدد البيانات الضخمة 9
1.4ضخامة الحجم 9
2.4 سرعة إنشائها 9
3.4 التنوع 9
5 - التحديات التي تواجه البيانات الضخمة 10
1.5 نمو سريع للبيانات 10
2.5 التخزين 10
3.5 بيانات غير موثوقة 10
4.5 أمن البيانات 10
5.5 فهم غير كافٍ للبيانات الضخمة 10
6.5 التنوع في تقنيات البيانات الضخمة 10
7.5 تعقيد إدارة جودة البيانات 10
6 – طرق التعامل مع بعض تحديات البيانات الضخمة 11
7- خطوات التعامل مع البيانات الضخمة 11
1.7 وضع استراتيجية البيانات الضخمة 12
2.7 معرفة مصادر البيانات الضخمة 12
3.7 الوصول إلى البيانات الضخمة ومعالجتها وتخزينها 12

4.7 تحليل البيانات الضخمة 12
5.7 اتخاذ قرارات ذكية مستندة إلى البيانات الضخمة 12
8- تطبيقات البيانات الضخمة 13
1.8 تطبيق البيانات الضخمة في القطاع الحكومي 13
2.8 البيانات الضخمة لضمان الأمن القومي 13
3.8 تطبيق البيانات الضخمة في المجال الإقتصادي 13
4.8 البيانات الضخمة في التعليم 13
5.8 تطبيقات البيانات الضخمة في مجال الرعاية الصحية 14
9 – الإستخدامات الإحصائية للبيانات الضخمة 14
10- البيانات الضخمة في مركز عجمان للإحصاء 15
1.10 الأهداف 15
2.10 الأهمية 15
3.10 منهجية التطبيق 16
4.10 التحديات 16
المراجع 19

مقدمة

يسر مركز عجمان للإحصاء والتنافسية أن يقدم دليل البيانات الضخمة،الذي يستعرض معلومات عن مفهوم البيانات الضخمة وسماتها،أنواعها وأهميتها، والتحديات التي من الممكن مواجهتها عند التعامل معها ، طرق التعامل مع بعض التحديات ،وأخيراً خطوات التعامل مع البيانات الضخمة

تعتبر البيانات الضخمة نقطة تحول في عصرنا الحالي، إذ تمكننا من تحويل الحياة الإجتماعية والتي هي ثمرة الإستخدام المكثف للتكنولوجيا إلى بيانات قابلة للقياس، بحيث يمكننا مراقبة الأنشطة البشرية والتفاعلات مع البيئة مما ينتج عنه بصمة رقمية هائلة.

ومع زيادة الإستخدام للتكنولوجيا هناك زيادة موازية في تدفقات المعلومات وجمع البيانات التي تنشأ يومياً ،والسبب الأهم لزيادة حجم البيانات، لأنها تستمر بالتولد بشكل أكبر بكثير من السابق من خلال عدة أجهزة ومصادر، و الأهم أن معظم تلك البيانات ليست مهيكلة، كتغريدات تويتر و الفيديوهات على يوتيوب و تحديثات الحالة على فيس بوك وغيرها، ما يعني أنه لا يمكن إستخدام أدوات إدارة قواعد البيانات وتحليلها التقليدية مع هذه البيانات لأنها ببساطة ليست وفق الهيكل المعتاد الذي يتم التعامل معه كالجداول بالإكسل فرضاً.

وللتعامل مع هذه الزيادات الأسية في البيانات ، لا بد من إتباع نهج وآلية لتحديد معدل النمو في حجم البيانات أو ما يعرف عموماً بانفجار المعلومات[1]، إن البيانات الكبيرة هي إحدى الطرق للتعامل مع هذه الظاهرة.

[1] هو مصطلح استخدم لأول مرة في عام 1941 ، وفقًا لقاموس أكسفورد الإنجليزي

1- مفهوم البيانات الضخمة

البيانات الضخمة تمثل مجموعة البيانات التي يفوق حجمها قدرة أي نظم قواعد بيانات على إجراء العمليات المعتادة عليها من تسجيل ونقل وتحليل وتخزين، بحيث يمكن تحويلها إلى معلومات يمكن الإستفادة منها.

وتشير البيانات الضخمة إلى عملية تُستخدم عندما لا يمكن لتقنيات استخراج البيانات التقليدية الكشف عن معنى البيانات الأساسية. لا يمكن معالجة البيانات غير المهيكلة أو الحساسة للوقت بواسطة محركات قواعد البيانات العادية،ويتطلب هذا النوع من البيانات نهجاً مختلفاً في المعالجة يسمى البيانات الضخمة، يستخدم توازياً هائلاً على الأجهزة المتاحة بسهولة.

وقد تكون بيانات لسلاسل زمنية ومتنوعة تشمل جوانب الحياة كافة حيث تتدفق البيانات بشكل مستمر دون توقف، ،ومن العوامل التي تجعل البيانات الضخمة ضخمة حقاً هي أنها تأتي من مجموعة متنوعة من المصادر أكثر من أي وقت مضى،وتكون غنية بكميات كبيرة من البيانات، والتي يؤدي تحليلها بشكل فعال إلى استخراج معلومات ذات أهمية على قدرٍ عالي من التفصيل .

2- أهمية البيانات الضخمة

تكمن أهمية البيانات الضخمة بعد معالجتها بأنها تغطي كافة المواضيع وجوانب الحياة .

1.2 يمكن الإستفادة من المعلومات الضخمة باتخاذ قرارات ورسم سياسات تؤدي إلى استمرار عجلة التطور والنهضة وذلك بعد تنظيم وربط المعلومات ببعضها .

2.2 تساعد المنظمات الحكومية والخاصة على رسم استراتيجيات لأعمالهم بناءً على البيانات الإجتماعية من محركات البحث ومواقع مثل الفيسبوك، تويتر وذلك بعد أن يتم معالجتها وتحويلها لمخرجات إحصائية.

3.2 يساعد استخدام البيانات الكبيرة في تحسين وتطوير مستوى كافة الخدمات وبكافة المجالات فمثلاً بمجال الصحة تشير التقديرات إلى أنه سيكون هناك واحد في المئة من مكاسب الكفاءة في مجال الرعاية الصحية العالمية من خلال الإستفادة من تحليل البيانات الضخمة.

يستخدم الباحثون الطبيون البيانات الضخمة لتحديد عوامل خطر الإصابة بالمرض للمساعدة في تشخيص الأمراض والظروف لدى المرضى ، بالإضافة إلى ذلك، فإن البيانات المستمدة من السجلات الصحية الإلكترونية ووسائل التواصل الإجتماعي وشبكة الإنترنت وغيرها من المصادر تزود منظمات الرعاية الصحية والوكالات الحكومية بمعلومات مستوفية عن التهديدات أو الفاشيات للأمراض المعدية.

أما بمجال تحسين خدمة العملاء ،يتم استبدال أنظمة ملاحظات العملاء التقليدية بأنظمة جديدة مصممة بتقنيات البيانات الضخمة. وفي هذه النظم الجديدة، تستخدم تكنولوجيا معالجة البيانات الضخمة واللغات الطبيعية لقراءة وتقييم استجابات المستهلكين.
كذلك تستخدم معظم الشركات تحليلات البيانات لإدارة الطاقة، بما في ذلك إدارة الشبكة الذكية، وتحسين الطاقة، وتوزيع الطاقة.

4.2 تنظيم ودمج البيانات الضخمة مع تحليلات البيانات المنظمة المتوفرة سابقاً ،يساعد على الحصول على بيانات متكاملة على مستوى من التفصيل، وقد ينتج عنها مؤشرات تُمكن من تحقيق المهام المتعلقة بالأعمال مثل تحديد الأسباب الأساسية للمشاكل أو الأخطاء في الوقت الحقيقي.

5.2 تغطي البيانات الضخمة كافة التفاصيل لسلسة زمنية ليس فقط الوقت الحالي، إذ يمكن استخدام البيانات التاريخية والحقيقية لتقييم التفضيلات المتطورة للمستهلكين ، وبالتالي تمكين الشركات من تحديث وتحسين استراتيجياتها التسويقية وتصبح أكثر استجابة لرغبات العملاء واحتياجاتهم.

6.2 تغطي البيانات الضخمة كافة المعلومات حول المناطق الجغرافية، مما يُمَّكن التنبؤ بالأماكن الأكثر عرضة لخطر الكوارث الطبيعية، وبالتالي توجه السلطات لإتخاذ الإجراءات الوقائية بخطوة استباقية للحد من الخسائر.

7.2 يمكن للبيانات الضخمة أن تزود الشركات برؤى قيمة عن عملائها، يمكن استخدامها لتحسين حملات وتقنيات التسويق من أجل زيادة مشاركة العملاء ومعدلات التحويل.

3 - أنواع البيانات الضخمة

للبيانات الضخمة عدة أنواع هي :

1.3 المنظمة

أي أنها بيانات يمكن تخزينها والوصول إليها ومعالجتها في شكل ثابت، حقق التقدم في العلوم التكنولوجية نجاحاً أكبر في تطوير تقنيات للعمل مع هذا النوع من البيانات (حيث يكون الشكل معروفاً مقدماً) ، وقابلاً لإستخلاص قيمة منه.

2.3 غير المنظمة

يتم تصنيف أي بيانات ذات شكل أو بنية غير معروفة على أنها بيانات غير منظمة بالإضافة إلى كونها ذات حجم ضخم ، والبيانات غير المنظمة تطرح تحديات متعددة من حيث معالجتها والخروج بقيمة ذات فائدة منها، مثال على ذلك مصدر بيانات غير متجانس يحوي على نصوص وصور ومقاطع فيديو حيث يبرز التحدي هنا بربط مختلف أنواع هذه البيانات ببعضها بصورة متجانسة تؤدي إلى استخلاص قيمة منها.

3.3 شبه منظمة

يمكن أن تحتوي البيانات شبه المنظمة على كلا النموذجين من البيانات،حيث يمكننا أن نرى البيانات شبه المنظمة على شكل منظم أو عشوائي.

4-السمات التي تحدد البيانات الضخمة

للبيانات الضخمة خصائص محددة تتميز بها عن غيرها من البيانات وهي :

1.4ضخامة الحجم

الصفة المحددة لهذه البيانات هو ضخامة حجمها،إذ يتم قياس هذه البيانات بواسطة أكبر وحدة تخزين،ولأخذ الفائدة من البيانات الضخمة، سيتطلب الأمر معالجة كميات كبيرة من البيانات ذات الكثافة المنخفضة وغير المُهيكلة و يمكن أن تكون هذه البيانات غير معروفة القيمة، مثل عمليات النقر على صفحة ويب أو تطبيق جوال وغيرها.

2.4 سرعة إنشائها

من أهم صفات البيانات الضخمة أيضاً هو سرعة تدفقها بشكل استثنائي، والتي يجب التعامل معه بشكل مناسب،وذلك بتوظيف تقنيات حديثة تستوعب كم البيانات الهائل المتدفق بوقت زمني قصير.

3.4 التنوع

تشير هذه الصفة إلى تعدد أنواع البيانات المتدفقة من مختلف النظم وهي تتراوح بين البيانات المنظمة وغير المنظمة، والرقمية لقواعد البيانات التقليدية إلى الوثائق غير الرقمية أو النصية، ورسائل البريد الإلكتروني، والصوتيات ومقاطع الفيديو، وبيانات مؤشر الأسهم، وبيانات تسجيل الدخول، والبيانات المشفرة ، أو حتى المعاملات المالية، وقد تتميز هذه البيانات بالتباين و يشير ذلك إلى عدم الإتساق الذي يمكن أن تظهره البيانات في بعض الأحيان، مما يعوق عملية القدرة على التعامل مع البيانات وإدارتها بفعالية.

5 - التحديات التي تواجه البيانات الضخمة

هناك الكثير من التحديات التي قد تواجهنا أثناء التعامل مع البيانات الضخمة نذكر منها:

1.5 نمو سريع للبيانات

سرعة النمو في مثل هذا المعدل العالي للبيانات يخلق مشكلة للبحث عن رؤى استخدامه. إذ لا توجد طريقة فعالة بنسبة 100٪ لتصفية البيانات ذات الصلة.

2.5 التخزين

إن توليد مثل هذا الكم الهائل من البيانات يحتاج إلى مساحة للتخزين، وتواجه المنظمات تحديات للتعامل مع هذه البيانات الضخمة دون أدوات وتكنولوجيا مناسبة.

3.5 بيانات غير موثوقة

لا يمكن ضمان أن تكون البيانات الضخمة التي تم جمعها وتحليلها هي دقيقة تماماً (100%) ، البيانات الزائدة عن الحاجة، أو البيانات المتضاربة، أو البيانات غير المكتملة هي تحديات لا تزال موجودة.

4.5 أمن البيانات

يمكن أن تكون الهيئات التي تخزن مثل هذه البيانات الضخمة (للمستخدمين) هدفًا لمجرمي الإنترنت ، وهناك خطر سرقة البيانات. وبالتالي، فإن تشفير مثل هذه البيانات الضخمة يشكل أيضا تحدياً جديداً.

5.5 فهم غير كافٍ للبيانات الضخمة

في كثير من الأحيان ، تفشل المؤسسات في معرفة حتى الأساسيات حول البيانات الضخمة مثلاً ما هي البيانات الضخمة في الواقع ، وما هي فوائدها ، وما هي البنية التحتية المطلوبة ، وما إلى ذلك .

و بدون فهم واضح لها ، فإن مشروع اعتماد البيانات الضخمة قد يكون محكوماً عليه بالفشل، قد تهدر المؤسسات الكثير من الوقت والموارد في بيانات لا تعرف حتى كيفية استخدامها.

6.5 التنوع في تقنيات البيانات الضخمة

يوجد مجموعة متنوعة من تقنيات التعامل مع البيانات الضخمة (Spark ، Hadoop MapReduce وبالنسبة لتخزين البيانات Cassandra ، HBase) ومن المحتمل اختيار برنامج بشكل سيئ ، يؤخر عملية معالجة البيانات.

7.5 تعقيد إدارة جودة البيانات

هناك مشكلة في تكامل البيانات ، لأن البيانات التي تحتاج إلى تحليل تأتي من مصادر متنوعة في مجموعة متنوعة من التنسيقات المختلفة. وبالتالي فإن معالجة البيانات ستختلف، وقد تكون مطابقتها مشكلة.

6 – طرق التعامل مع بعض تحديات البيانات الضخمة

يتم تأهيل البيانات الضخمة في الأغلب من قبل خبراء،يقومون بمعالجة كل التحديات التي تنشأ عن التعامل مع البيانات الضخمة بشكل فردي.

فمثلاً يتم وضع خطة لمقدار البيانات التي سيتم استثمارها، وكيف وأين سيتم تخزينها لمعالجة ضخامة هذه البيانات، كذلك تُحدد جميع المصادر المختلفة للبيانات بواسطة نظام تكنولوجي وتُحدد الأدوات المناسبة للتعامل معها.

أما بالنسبة لسرعة تدفق البيانات وهو أمر بالغ الأهمية ،في وقتنا هذا الذي يعتمد فيه أغلب الناس على آلية أخذ المعلومات بأبسط الوسائل وذلك عن طريقة مواقع ومحركات البحث على الإنترنت مما يسبب زيادة هائلة في تدفق البيانات الضخمة ،لذلك لا بد من توظيف التكنولوجيا المناسبة لضمان تطوير صورة البيانات الضخمة الواردة بأسرع وقت.

ليست كل المعلومات الواردة و المجمعة متساوية الأهمية، لذا لا بد من بناء بيانات كبيرة قابلة للمعالجة بطرق سهلة، والتأكد من أن البيانات المعالجة دقيقة ونظيفة.

ولضمان فهم البيانات الضخمة وقبولها على جميع المستويات ، تحتاج أقسام تكنولوجيا المعلومات إلى تنظيم العديد من الدورات التدريبية وورش العمل، لمعرفة المزيد من قبول البيانات الضخمة ، ويجب مراقبة تنفيذ واستخدام حل البيانات الضخمة الجديد والتحكم فيه.

ولتحديد التقنية المناسبة للتعامل مع البيانات الضخمة في حالة أنها كانت المرة الأولى للتعامل معها، فستكون محاولة البحث عن مُسَاعدة احترافية هي الطريقة الصحيحة ، حيث يمكن تعيين خبير في التعامل مع البيانات الضخمة.

مشكلة تعقيد إدارة جودة البيانات تُحل بتصميم خوارزميات البيانات الضخمة مع مراعاة الترقية المستقبلية. ولكن بالإضافة إلى ذلك لا بدَّ من التخطيط لصيانة النظام ودعمه بحيث يتم التعامل مع أي تغييرات متعلقة بنمو البيانات بشكل صحيح. علاوة على ذلك ، فإن إجراء عمليات تدقيق منتظمة للأداء يمكن أن يساعد في تحديد نقاط الضعف وبالتالي معالجتها في الوقت المناسب.

7- خطوات التعامل مع البيانات الضخمة

قبل أن يتم استثمار البيانات الضخمة بالعمل لصالح أي مؤسسة ، يجب أن يتم النظر في كيفية تدفقها بين العديد من المواقع والمصادر والأنظمة والمستخدمين .هناك خمس خطوات رئيسية لتولي مسؤولية هذا النسيج من البيانات الكبير الذي يتضمن البيانات التقليدية المنظمة إلى جانب البيانات غير المنظمة وشبه المنظمة .

1.7 وضع استراتيجية البيانات الضخمة

تصميم استراتيجية البيانات الضخمة وهي خطة مصممة للمساعدة على الإشراف وتحسين الطريقة التي يتم بها الحصول على البيانات وتخزينها وإدارتها ومشاركتها واستخدامها داخل وخارج المؤسسة.

عند وضع استراتيجية من المهم أن تُدرس أهداف ومبادرات الأعمال والتكنولوجيا الحالية والمستقبلية وهذا يستدعي التعامل مع البيانات الضخمة كأنها أصول في المؤسسة.

2.7 معرفة مصادر البيانات الضخمة

لا بدّ من معرفة وتحديد مصادر البيانات الضخمة المتدفقة والمستخدمة ضمن منظومة المؤسسة.

فالبيانات قد تأتي من بيانات البث من الإنترنت والأجهزة المتصلة الأخرى والتي تتدفق إلى أنظمة تكنولوجيا المعلومات من السيارات الذكية والأجهزة الطبية والمعدات الصناعية وغيرها.

يمكن تحليل هذه البيانات الكبيرة عند وصولها وتحديد البيانات التي يجب الاحتفاظ بها، والبيانات التي تحتاج إلى مزيد من التحليل.

ويشمل التدفق الهائل للبيانات الناتجة عن استخدام مواقع التواصل الاجتماعي على صور ومقاطع فيديو وصوت ونص وصوت ، وغالباً ما تكون هذه البيانات في أشكال غير منظمة أو شبه منظمة، لذلك فإنه يشكل تحدياً فريداً للتحليل واستخراج النتائج.

3.7 الوصول إلى البيانات الضخمة ومعالجتها وتخزينها

توفر أنظمة الحوسبة الحديثة السرعة والقوة والمرونة اللازمة للوصول بسرعة إلى كميات هائلة وأنواع مختلفة من البيانات الضخمة، وإلى جانب الوصول الموثوق، يجب توفر طرق لدمج البيانات وضمان جودة البيانات، وتوفير إدارة البيانات وتخزينها، وإعداد البيانات من تنظيف لها وتعديل وتجهيزها بشكل قابل للتحليل. قد يتم تخزين بعض البيانات في أماكن العمل في مستودع البيانات التقليدي ،ولكن هناك أيضاً خيارات مرنة ومنخفضة التكلفة لتخزين البيانات الضخمة والتعامل معها عبر حلول السحابة وبحيرات البيانات وهادوب.

4.7 تحليل البيانات الضخمة

مع التكنولوجيا عالية الأداء مثل الحوسبة الشبكية ، يمكن للمؤسسات استخدام كل البيانات الضخمة الخاصة بهم وتحليلها. وثمة نهج آخر هو تحديد البيانات ذات الصلة مقدماً قبل تحليلها. وفي كلتا الحالتين، فإن تحليلات البيانات الضخمة هي كيفية اكتساب المؤسسات للقيمة والرؤى من البيانات.

وعلى نحو متزايد تغذي البيانات الضخمة مساعي التحليلات المتقدمة اليوم مثل الذكاء الاصطناعي.

5.7 اتخاذ قرارات ذكية مستندة إلى البيانات الضخمة

تؤدي البيانات الموثوق بها التي تتم إدارتها بشكل جيد إلى تحليلات وقرارات موثوق بها، ولكي تظل المؤسسات قادرة على المنافسة يتعين عليها أن تستخدم البيانات الضخمة بشكل تام لإتخاذ القرارات و بناء الإستراتيجيات.

8- تطبيقات البيانات الضخمة

تُستخدم البيانات الضخمة لزيادة الكفاءة والإنتاجية على مستوى جميع القطاعات في المجتمع، ومن أهم القطاعات التي طَبقت استخدام البيانات الضخمة ما يلي :

1.8 تطبيق البيانات الضخمة في القطاع الحكومي

تحتاج الحكومة إلى التعامل مع مختلف القضايا المعقدة المحلية والوطنية والعالمية يومياً، إن تطبيق البيانات الضخمة يمكن أن يترك أثراً هائلاً على هذا القطاع من خلال جمع كل المعلومات عن ملايين الأشخاص والتي تساعد على اتخاذ أي قرار بالنظر إلى السكان المحليين. فهو يسمح بتحليل تأثير ورأي أي قرار وتقرير ما إذا كان هناك حاجة إلى أي تغيير أم لا.

2.8 البيانات الضخمة لضمان الأمن القومي

في العديد من مراكز حفظ الأمن في دول العالم، يتم استخدام البيانات الضخمة لتحسين سير العمل والعمليات، وقد وظفت البلدان المتقدمة البيانات الضخمة من أنشطتها الإجتماعية والأمنية منذ فترة طويلة، حيث تقوم الحكومات بجمع المعلومات الخاصة بالمواطنين، ويتم تخزينها في قاعدة بيانات ، ويمكن تقييم كثافة السكان في موقع محدد وتحديد المخاطر المحتملة التي تشكل تهديد حتى قبل حدوثه ، ويمكن لموظفي الأمن استخدام مجموعة البيانات هذه للعثور على أي مجرم والكشف عن الأنشطة الإحتيالية في أي منطقة من مناطق البلد.

وإلى جانب ذلك، يمكن التنبؤ باحتمال انتشار أي فيروس أو أمراض واتخاذ الإجراءات اللازمة للوقاية.

3.8 تطبيق البيانات الضخمة في المجال الإقتصادي

أصبح بإمكان الشركات والمؤسسات الإستفادة من تحليل حركة العملاء، حيث أن العملاء هم أبرز الأصول لأي عمل تجاري. فإذا تم إرضاء الزبائن وفهم احتياجاتهم، سيكون بمقدور المؤسسة أن تنافس المؤسسات الموجودة في السوق .

يمكن من تحديد متطلبات العملاء، وما يريدونه ويركز على تقديم أفضل الخدمات لتحقيق مطالبهم. و يمكن للمؤسسات أن تحلل سلوك العميل، واتباع اتجاهاتها لإنتاج منتجات موجهة نحو العملاء.

4.8 البيانات الضخمة في التعليم

يمكن للبيانات الضخمة أن تولد نتائج غير مسبوقة وأن تستحدث نُهجاً مبتكرة تعتمد على البيانات لتعليم الطلاب. وفي كثير من البلدان، تُعدّ استخدامات البيانات الضخمة في المدارس والكليات شائعة.

إن استخدام البيانات الضخمة المخزنة التي تتضمن سجلات الطلاب وغيرها من البيانات حول الطلاب ، توفر بيانات مؤثرة على الأنشطة التي تقوم بها الفصول الدراسية ويساعد في اتخاذ القرارات للمنظمات.

كما باستخدام كاميرات عالية الدقة ولقطات فيديو ومعالجة الصور ، يمكن للبيانات الضخمة تقييم تعبير الوجه للطالب ويمكنه تتبع تحركاته، وتحفيز الطلاب من خلال تحديد المشاكل وتقديم أفضل تعليم ممكن للأطفال.

5.8 تطبيقات البيانات الضخمة في مجال الرعاية الصحية

علم البيانات يهيمن على تحسين الرعاية الصحية في الوقت الحاضر،حيث أن البيانات الضخمة لم يتم إدخالها فقط لتحديد العلاج ولكن أيضا لتحسين عملية تقديم الرعاية الصحية، وللبيانات الضخمة تأثير كبير على الحد من إهدار المال والوقت .وإلى جانب ذلك، تستخدم الحكومات البيانات الضخمة لتطوير بنى تحتية جديدة وخدمات طبية في حالات الطوارئ.

وتمكن مديري التحولات من التنبؤ بالأطباء المطلوبين في أوقات محددة ، كما أنه باستخدام الأجهزة الرقمية القابلة للإرتداء، يمكن للبيانات الضخمة مراقبة المرضى وإرسال التقارير إلى الأطباء المرتبطين بها، كذلك تُمكن البيانات الضخمة من تقييم الأعراض وتحديد العديد من الأمراض في المراحل المبكرة.

9 – الإستخدامات الإحصائية للبيانات الضخمة

يمكن للبيانات الضخمة أن تلعب دوراً كبيراً في المؤسسات الإحصائية التي تواجه مطالب ببيانات أكثر وأفضل وبموارد ووقت أقل ،حيث تساعد على استكمال الإحصائيات عندما تكون المصادر التقليدية ليست كافية،وتقلل من التكلفة والوقت اللازم لجمع البيانات الإحصائية المطلوبة على الرغم من وجود بعض التحديات التي قد تظهر عند استخدام البيانات الضخمة كتوفر الآليات التكنولوجية المناسبة للتعامل مع البيانات الضخمة وتنظيمها وتنقيحها بحيث تصبح قابلة للاستخدام الإحصائي ، وإكساب الموظفين المهارات والتدريب الكافي للتعامل مع هذا النوع من البيانات،وتكون هذه العملية ككل ضمن أطر الجودة المحددة لدى المؤسسة الإحصائية،و تكون هذه الآليات ضمن الميزانية المالية للمؤسسات الإحصائية ، بالإضافة إلى تحديد موارد البيانات الضخمة.

وقد أدرك المجتمع الإحصائي إمكانات البيانات الضخمة في تحسين العمل الإحصائي وخفض تكاليف الإحصاءات الرسمية.

ففي عام 2014 ، أنشأت الأمم المتحدة مجموعة عمل عالمية (GWG) لتقوم بتقديم استراتيجية واتجاهاً وتنسيقاً لبرنامج عالمي حول استخدام مصادر البيانات الجديدة والتقنيات الجديدة ، والتي تعتبر ضرورية للأنظمة الإحصائية الوطنية لتبقى ذات صلة في مشهد البيانات سريع الحركة

من أبرز الأمثلة على تطبيق بعض الدول لنظام البيانات الضخمة في إنتاج المؤشرات الإحصائية:

- استخدام بيانات الإنترنت لإنتاج مؤشر أسعار المستهلك .
- استخدام عمليات البحث في" Google للتنبؤ الآني" لحالة الاقتصاد.

- استخدام بيانات الأقمار الصناعية لتقدير إحصاءات المحاصيل الزراعية (نوع المحاصيل والتنبؤ بالإنتاجية).
- استخدام بيانات الهاتف المحمول لقياس إحصاءات السياحة.
- استخدام بيانات الماسح الضوئي لحساب إحصاءات الأسعار.
- وتجميع إحصاءات النقل البحري لمعرفة مدى تأثير جائحة كورونا على مجالات التجارة والسفر والنقل. باستخدام بيانات تتبع سفينة AIS.
- استخدام البيانات من شبكات الهاتف المحمول لتتبع تحركات الناس في جميع أنحاء البلاد للتنبؤ بإحصاءات تفشي الأمراض.

10- البيانات الضخمة في مركز عجمان للإحصاء

بناءً على توجه مركز عجمان للإحصاء والتنافسية لتطبيق مفاهيم البيانات الضخمة وتقنياتها لخدمة الإمارة والتي توفر بيانات تناسب المتطلبات في أقل وقت ممكن وعلى درجة من الحداثة والتكامل مع البيانات من المصادر التقليدية يأتي استخدام البيانات الضخمة في القطاع الحكومي لإبتكار حلول فعالة و سريعة من شأنها تحسين الأداء الحكومي والإنتاجية، كما تسهم في تقديم خدمات وبناء سياسات ذكية.

1.10 الأهداف

يهدف تطبيق البيانات الضخمة في مركز عجمان للإحصاء و التنافسية إلى جمع وتوفير بيانات ومعلومات إحصائية حديثة ذات جودة عالية تساهم في صنع القرارات ورسم السياسات وتقييم الأداء، وعلى هذا النحو يحقق الأهداف الرئيسية التالية:

1. تحسين مخرجات الإحصاءات الرسمية.
2. توفير مؤشرات تنافسية عجمان.
3. ربط الخدمات الحكومية في مختلف الجهات و تكامل و ترابط البيانات المتوفرة منها.
4. معرفة متطلبات المواطن في الإمارة.

2.10 الأهمية

1. يوفر لصانعي القرار المؤشرات الهامة التي تساعد في وضع الخطط ورسم السياسات الاقتصادية.
2. تقديم خدمات حكومية مبنية على أسس و بيانات متكاملة.
3. توفير معلومات وبيانات يمكن الإعتماد عليها في إعداد الدراسات والتنبؤات المستقبلية للتنمية في إمارة عجمان والعوامل المؤثرة فيها.

4. الإستجابة لمتطلبات المواطن في إمارة عجمان و العمل على توفير مستوى عالي من الرفاهية.
5. تقليل حالات عدم الإستجابة في المسوح الميدانية.

3.10 منهجية التطبيق

تم وضع المراحل التالية لتطبيق البيانات الضخمة في مركز عجمان للإحصاء و التنافسية:

1.3.10 مرحلة تقييم الوضع الحالي

يتم في هذه المرحلة دراسة الوضع الحالي لإمارة عجمان والتي على ضوءها يتم حصر مصادر البيانات المتوفرة في الإمارة والتقنيات التي يمكن تطبيقها بالإضافة لتحديد متطلبات العمل و هي كالتالي:

1. الإطلاع ودراسة أفضل الممارسات العالمية في البيانات الضخمة التي يمكن تطبيقها في الإمارة.
2. عقد ورشات عمل تدريبية للفريق للإلمام بآليات العمل من تقنيين وإحصائيين.
3. حصر مصادر البيانات.
4. التنسيق مع الجهات المعنية بمشروع البيانات الضخمة.
5. عقد ورشات عمل للجهات المعنية في الإمارة لشرح المشروع وآلية العمل.
6. توفير الخوادم والتقنيات لتخزين البيانات ومعالجتها.
7. توفير البرامج اللازمة لتحليل البيانات.
8. إعداد ميزانية تناسب المشروع ككل.

2.3.10 مرحلة دراسة أفضل الممارسات العالمية

يتم في هذه المرحلة دراسة أفضل الممارسات العالمية في مجال البيانات الضخمة ومحاولة تطبيق ما يتناسب مع إمارة عجمان و الخروج بالتوصيات.

3.3.10 المرحلة التنفيذية

بناءً على مخرجات المرحلتين السابقتين يتم تنفيذ المشروع حيث يتم الربط مع الجهات المعنية واستيراد البيانات و معالجتها من ثم تحليلها واستخراج النتائج والتوصيات، والتي يتم عرضها على أصحاب القرار للخروج بالسياسات و الاستراتيجيات.

4.10 التحديات

قد تواجه المركز بعض التحديات كحال الكثير من الجهات الإحصائية في التعامل مع البيانات الضخمة مثل:

1. توفر البيانات و تغطيتها.
2. عدم فهم أهمية البيانات الضخمة وفوائدها والبنية التحتية الي تحتاجها

3. تعاون الجهات المعنية بالربط.
4. المحافظة على امن المعلومات
5. توفر الموارد البشرية المتخصصة
6. التحديات التقنية (كتخزين البيانات و تحليلها).
7. التكلفة المالية للبرمجيات.
8. جودة ودقة البيانات.

4.11 التوجهات الحالية للتغلب على التحديات

يُشير مصطلح البيانات الضخمة إلى كميات هائلة من المعلومات المختلفة، التي يصعب جمعها وتقييمها عبر التقنيات التقليدية، بالإضافة إلى أنها تتميز بالحاجة للمعالجة السريعة، بحيث يمكن من خلالها عرض النقاط المشتركة، والتوجهات، والأنماط في سلوك المجموعة المستهدفة.

ويتبع مركز عجمان للإحصاء والتنافسية منهجية معينة في إدارة البيانات الضخمة. وتتمثل في تحديد البيانات المتوفرة وبحث كيفية استخدامها لإنتاج إحصاءات جديدة ملائمة للاحتياجات المطلوبة أو لتوفير بدائل أرخص، وأنسب من حيث التوقيت مقارنة بالمصادر التقليدية.

بالإضافة الى ذلك سيبحث المركز عن فرص جديدة للتعاون مع احدى المنظمات الإحصائية، وأصحاب البيانات، والخبراء في مجال البيانات الضخمة وغيرهم من الاساتذه الأكاديميين المتخصصين في مجال البيانات ..

كما ان هناك بعض التحديات التي ستواجه المركز عند إدارة البيانات الضخمة ولكن سيتم وضع الحلول المناسبة لها من اجل معالجة المعوقات التي قد نواجهها.

في كثير من الأحيان يتعذر علينا فهم ماهية البيانات الضخمة، وفوائدها، والبنية التحتية اللازمة لاعتمادها..ومن ثم توضيح مدى أهميتها للموظفين و لضمان فهم أهمية الأمر. لذا سيتم تنظيم عددا من الدورات التدريبية وورش العمل من اجل تعريف الموظفين المعنين ورفع مستوى كفاءة اداءهم.

والتحدي الاخر الكلفة المالية لادارة البيانات الضخمة من تكاليف الأجهزة الجديدة، وتعيين موظفين جدد مثل: مديري النظم، والمطورين، ودفع تكاليف تطوير البرامج الجديدة، وإعدادها، وصيانتها. ولمعالجة ذلك يتطلب الاعتماد على الحلول السحابية، وعلى البرمجيات الداخلية، التي يمكن أن تكون أيضًا فعالة من حيث التكلفة. أو يمكن اللجوء إلى إستراتيجيات جديدة، أو الخوارزميات المُحسنة، وإذا تم ذلك بشكل

صحيح يمكن أن يوفر أيضًا الكثير من المال.واستخدام الخوارزميات المحسّنة يمكن أن تقلل من استهلاك الطاقة الحاسوبية بمقدار 5 إلى 100 مرة، أو أكثر.

واما بخصوص المشاكل الأمنية يجب إعطاء التدابير اللازمة لمواجهة تحديات أمان البيانات أولوية كبرى، حيث إن أمن البيانات أمر مهم جدًا خاصة في مرحلة تصميم البنية التحتية ، تجنبًا لحدوث أي شيء غير متوقع قد يؤدي إلى فشل المشروع بالكامل.

وفي الكثير من الأحيان؛ تُؤجل مشاريعُ تبني البيانات الضخمة موضوعَ الأمن الإلكتروني إلى مراحل لاحقة.

المراجع

1- مركز أبو ظبي للإحصاء _ مفاهيم عامة حول البيانات الكبيرة _ أدلة المنهجية والجودة -دليل رقم (13)
13دليل البيانات الكبيرة النسخة النهائية(scad.gov.ae) pdf.

2- الدليل الإرشادي لدمج البيانات الضخمة في الأعمال الإغاثية_سبتمبر 2015

(PDF)الدليل الارشادي لدمج البيانات الضخمة في العمليات الانسانية | Andrej Verity and Kahtan Al-Shabi - Academia.edu

3- موقع البيانات الضخمة العالمي
Home — UN GWG for Big Data

4- معالجة البيانات الضخمة_ اكتشاف المعرفة في البيانات الضخمة _فيليب نيومان
Big Data Processing - an overview | ScienceDirect Topics

5- البيانات الضخمة _ ويكيبديا الموسوعة الحرة
Big data - Wikipedia

6- تحليلات البيانات الضخمة_ يناير 2021-22
ما هي تحليلات البيانات الضخمة/Big Data Analysis/ ؟ – عرب ميديا(arabdia.com)

7-التقنيات المستخدمة مع البيانات الضخمة- بتول عتوم _ نوفمبر -2-2020
التقنيات المستخدمة مع البيانات الضخمةe3arabi –

8- بيانات كبيرة قيد التنفيذ للحكومة - مجموعة البنك الدولي

http://documents1.worldbank.org/curated/en/765114912873809 86/pdf/114011-BRI-3-4-2017-11-49-44-WGSBigDataGovernmentFinal.pdf

9- تحليلات البيانات الضخمة -فيليب روسوم- الربع الأول 2011

https://tdwi.org/research/2011/09/~/media/TDWI/TDWI/Research/BPR/2011/TDWI_BPReport_Q411_Big_Data_Analytics_Web/TDWI_BPReport_Q411_Big%20Data_ExecSummary.ashx

10-تطبيقات البيانات الضخمة في القطاع الحكومي: تحليل مقارن بين الدول الرائدة - جانج هون كيم-مارس 2014

https://www.researchgate.net/publication/260865566_Big_Data_Applications_in_the_Government_Sector_A_Comparative_Analysis_among_Leading_Countries

11-تحليل الشبكات الاجتماعية - ويكيبيديا

https://en.wikipedia.org/wiki/Social_network_analysis#:~:text=Social%20network%20analysis%20(SNA)%20is,or%20interactions)%20that%20connect%20them
.