

الإحصائيات في برامج (Rapid Miner) بالتطبيق على سرطان الثدي (دراسة حاله مستشفى الذرة – الخرطوم 2010 - 2021)

محاضر كلية علوم الحاسوب وتقانة المعلومات
جامعة القران الكريم والعلوم الإسلامية

أ. مرشد إبراهيم طالب مصطفى

المستخلص:

يهدف البحث الي إستخدام الطرق الإحصائية في برامج (Rapid Miner) في إستخراج النتائج الإحصائية وتحديد نسبة الإصابة بالمرض وفئات الفئات العمرية الأكثر عرضة لهذا المرض من خلال عملية التنقيب عن بيانات المرضى ومعرفة اكثر الاعمار اصابة بالمرض لإجراء فحوصات وقائية مبكرة من المرض وهتمة الدراسة علي العمل علي تحليل دقيق لكمية كبيرة من البيانات المتوفرة لعدد من السنوات وإستخدام الطرق الإحصائية التي تساعد علي اتخاذ قرارات تساعد في المعرفة بمعدلات انتشار المرض في المستقبل وتوفير البيانات الازمة التي تساعد علي ارشادات ونصائح في افضل الطرق لتجنب انتشار سرطان الثدي. و يتبع البحث المنهج الوصفي التحليلي والتجريبي ، حيث يتم جمع البيانات والمعلومات الخاصه بسجلات المراقبة وإعدادها وتصنيفها وتبويبها ومن ثم عرضها وتحليلها ، وتم التوصل للنتائج من خلال إستخدام الطرق الإحصائية في برنامج (Rapid Miner) تم التوصل الي ان ولاية الخرطوم الاكثر إصابة بعدد 1944 اصابة بين ان ولاية النيل الازرق الاقل إصابة بعدد 45 اصابة و ان القبائل الاكثر انتشارا للمرض قبية الجعلين بعدد 1025 اصابة والاقبل قبيلة البرقاوية بعدد 1 اصابة وان الاصابة حسب النوع حيث الإناث الاكثر اصابة بالمرض بعدد 6942 اصابة و الرجال الاقل بعدد 557 اصابة وان الاعوام اكثر إنتشارا للمرض حيث ان العام 2021 الاكثر انتشارا بعدد 3736 اصابة و العام الاقل 2012 بعدد 138 اصابة .

Statistics in programs) Rapid Miner (applied to breast cancer)
(م2010-2021م) case study corn hospital

Mustafa Morshed Ibrahim Talib

Abstract:

The research aims to use statistical methods in (Rapid Miner) programs to extract statistical results and determine the incidence of the disease and the age groups most vulnerable to this disease through the process of excavating patient data and knowing the most affected ages for early preventive examinations of the disease. Careful analysis of a large amount of data available for a number of years and the use of statistical methods that help in making decisions that help in knowing the rates of disease prevalence in the future and providing the necessary data that helps guides and advice on the best ways to avoid

the spread of breast cancer. The research follows the analytical and experimental descriptive approach, where data and information related to monitoring records are collected, prepared, classified and tabulated, and then displayed and analyzed, and the results were reached through the use of statistical methods in the (Rapid Miner) program. The Blue Nile State has the least infection with 45 infections, and the tribes with the most prevalence of the disease have 1025 infections, and the least is the Barqawi tribe with 1 infection, and the infection is according to gender, where females are the most infected with the disease with 6942 infections, and men are the least with 557 infections, and the years are more prevalent for the disease, as the The year 2021 is the most common, with 3,736 infections, and the least common year is 2012, with 138 infections.

أولاً : الإطار المنهجي :

1- المقدمة:

مع وجود كميات كبيرة من البيانات المخزنة في قواعد البيانات ومخازن البيانات ، زادت الحاجة إلى تطوير أدوات قوية لتحليل البيانات واستخراج المعلومات والمعرفة منها. من هنا ، ظهر ما يسمى بالتنقيب في البيانات كتقنية تهدف إلى استخراج المعرفة من كميات هائلة من البيانات وإيجاد علاقة منطقية تلخص البيانات. بطريقة جديدة مفهومة ومفيدة لصاحب البيانات ، هي تقنية حديثة فرضت نفسها بقوة في عصر المعلومات ، واستخدامها يوفر للدولة والشركات والمؤسسات والمستشفيات في جميع المجالات القدرة على استكشاف و التركيز على أهم المعلومات في قواعد البيانات ، وتركز تقنيات الاستكشاف على بناء التنبؤات. بعد أن طور العلماء أجهزة الكمبيوتر، أدرك المجتمع والعالم كله أن هذه الأجهزة الجديدة ستوفر العديد والعديد من الخدمات للبشرية جمعاء ، خاصة في مجال المعلومات والتخزين والمعالجة والاسترجاع ، وبعد هذا التاريخ قبل عقد من الزمن ، الأطباء والمتخصصون وبدوره بدأ بمحاولة الاستفادة من هذه التقنيات بشكل حقيقي من خلال تطوير فكرة إدارة المعلومات ودور الحاسب الآلي في الطب والرعاية الصحية التي تعد من أهم المجالات العلمية وأكثرها انتشارا ومؤثر ، ولا يزال الطب يبحث عن مزيد من التطور باستخدام جميع وسائل العلم المتاحة ، وأهمها التكنولوجيا وأنظمة المعلومات وأدوات التنقيب عن البيانات لتحليل الأمراض ومدى انتشارها وطرق الحماية. تهتم العديد من الدول المتقدمة ببيانات المرضى أكثر من البيانات الأخرى لأنها تدرك تمامًا أهمية الصحة ، ونجد أن الهيكل الأساسي للدولة هو القوة العاملة ، ووجودها بصحة جيدة يعني مدى انتشار الصحة. الوعي ، وهذا بدوره يؤدي إلى تطوير جيل جديد سليم في المجتمع.

2- مشكلة البحث:

لصعوبة التعامل مع البيانات الكبيرة وصعوبة الوصول الي نتائج مع البيانات الكبيرة ويمكن توضيح مشكلة البحث في الاتي :

1. عدم استخدام أدوات التصنيف لوصف الفئات العمرية الأكثر عرضة للإصابة بمرض انتشار المرض.
2. صعوبة تحليل مجموعة كبيرة من البيانات بالوسائل الاحصائية التقليدية.

3 -أهداف البحث:-

- 1.إستخدام الطرق الإحصائية في برامج (Rapid Miner) في إستخراج النتائج الإحصائية.
- 2.تحديد نسبة الإصابة بالمرض وفئات الفئات العمرية الأكثر عرضة لهذا المرض من خلال عملية التنقيب عن بيانات المرضى.
- 3.معرفة اكثر الاعمار اصابة بالمرض لإجراء فحوصات وقائية مبكرة من المرض.
4. تصنيف المرضى علي اساس الولاية الاكثر انتشاراً .

5.أسئلة البحث:

- 1.ماهي السنوات الاكثر انتشارا للمرض؟
2. ما هي الوظائف الاكثر عرضة للإصابة بالمرض؟
- 3.هل المرض اكثر انتشارا بين المتزوجين ام غير المتزوجين؟

4-اهمية البحث:

العمل علي تحليل دقيق لكمية كبيرة من البيانات المتوفرة لعدد من السنوات وإستخدام الطرق الإحصائية التي تساعد علي اتخاذ قرارات تساعد في المعرفة بمعدلات انتشار المرض في المستقبل وتوفير البيانات الازمة التي تساعد علي ارشادات ونصائح في افضل الطرق لتجنب انتشار سرطان الثدي.

4 - حدود البحث:-

- الحدود المكانية:مستشفى الذرة الخرطوم.
- الحدود الزمانية: 2010الى 2021 م.

6- نطاق البحث:-

مجموعة بيانات لسرطان الثدي في الفترة من 2010-2021.

7- طريقة جمع البيانات:

تم جمع البيانات بناء على المقابلة ، حيث تم جمعها من نظام قاعدة بيانات مستشفى "الذرة".

8- عينة الدراسة:

تضمنت عينة الدراسة 7500 حالة مسجلة لمرضى سرطان الثدي بمستشفى الذرة.

9- منهجية البحث:-

يتبع البحث المنهج الوصفي التحليلي والتجريبي ، حيث يتم جمع البيانات والمعلومات الخاصه بسجلات المراقبة وإعدادها وتصنيفها وتبويبها ومن ثم عرضها وتحليلها، ومن ثم تعتمد المنهج البنائي لبنا نموذج قادر على الاكتشاف بصورة فاعلة

ثانيا: الإطار النظري

1- تنقيب البيانات والعلوم الأخرى:

يُعتبر تنقيب البيانات ملتقى الجهود المبذولة من الباحثين في عدة مجالات من المعرفة، والذي من خلاله يتم تطوير وبناء تقنيات تتعامل مع البيانات وأشكالها المتعددة وأنواعها المختلفة بهدف مواجهة المشكلات في مجالات مختلفة كالهندسة، الأعمال، الصناعة، الطب والعلوم. يجمع تنقيب البيانات بين عدة علوم كالإحصاء وتعليم الآلة وقواعد البيانات وتقنيات الإظهار المرئي، ويتجلى هذا الجمع في مراحل وخطوات تنقيب البيانات بدءاً من تجهيز وتجميع البيانات وحتى النتيجة النهائية والتي تختلف حسب أهداف وأغراض التنقيب.⁽⁵⁾



الشكل رقم(1) يوضح علاقة تنقيب البيانات بالعلوم الاخرى

تنقيب البيانات يستخدم بعض تقنيات تعليم الآلة مثل الشبكات العصبية وشجرة القرار، ويختلف هدف تنقيب البيانات عن هدف تعليم الآلة، فتعليم الآلة هدفه إعطاء الحواسيب القدرة على تنفيذ مهام يقوم بها البشر عبر تعليمها، بمعنى آخر: استبدال الدور البشري، ولكن تنقيب البيانات هدفه مساعدة الدور البشري ودعمه وليس استبداله. يتمثل الدور الرئيسي لقواعد البيانات في حفظ البيانات والحصول عليها عند الحاجة، بينما دور تنقيب البيانات يتمثل في القدرة على قراءة هذه البيانات وتحليلها للمساعدة في اتخاذ القرار المناسب. تقنيات الإظهار المرئي يتم استخدامها كأداة في مرحلة تحضير البيانات أو مرحلة ما بعد تنقيب البيانات لإظهار النتائج.⁽⁵⁾

2- مفهوم البيانات الضخمة (Big Data):

البيانات الضخمة هي مصطلح يصف الحجم الكبير للبيانات سواء المهيكلة أو غير المهيكلة التي تغمر العمل على أساس يومي. ولكن ليس مقدار البيانات هو المهم، يمكن تحليل البيانات الضخمة للحصول على رؤى تؤدي إلى قرارات أفضل وتحركات تجارية استراتيجية.⁽⁶⁾ أصبح استخدام البيانات الضخمة أمراً شائعاً في هذه الأيام من قبل الشركات للتفوق على أقرانهم، في معظم الصناعات، سيستخدم المنافسون الحاليون والوافدون الجدد على حد سواء الإستراتيجيات الناتجة عن البيانات التي تم تحليلها للمنافسة والابتكار والحصول على القيمة.⁽⁶⁾ تساعد البيانات الضخمة المؤسسات على خلق فرص نمو جديدة وفئات جديدة تماماً من الشركات التي يمكنها الجمع بين بيانات الصناعة وتحليلها، هذه الشركات لديها معلومات وافرة عن المنتجات والخدمات والمشتريين والموردين وتفضيلات المستهلكين التي يمكن التقاطها وتحليلها.⁽⁶⁾

مصطلح البيانات الضخمة:

في حين أن مصطلح «البيانات الضخمة» جديد نسبياً، فإن عملية جمع وتخزين كميات كبيرة من المعلومات لتحليلها في النهاية قديمة. اكتسب هذا المفهوم زخماً في أوائل العقد الأول من القرن الحادي والعشرين عندما أوضح محلل الصناعة دوغ لاني التعريف السائد الآن للبيانات الضخمة على أنها ثلاث مراحل:

1. الصوت: تجمع المؤسسات البيانات من مجموعة متنوعة من المصادر، بما في ذلك المعاملات التجارية ووسائل التواصل الاجتماعي والمعلومات من بيانات أجهزة الاستشعار أو البيانات من جهاز إلى جهاز، في الماضي كان تخزينها يمثل مشكلة - ولكن التقنيات الجديدة خففت العبء.⁽⁶⁾ يرتبط اسم «البيانات الكبيرة» بحد ذاته بحجم ضخم، يلعب حجم البيانات دوراً بالغ الأهمية في تحديد القيمة خارج البيانات، ما إذا كان يمكن اعتبار بيانات معينة في الواقع بيانات كبيرة أم لا، يعتمد على حجم البيانات، ومن ثم، فإن «الحجم» هو أحد الخصائص التي يجب مراعاتها أثناء التعامل مع «البيانات الكبيرة».

2. السرعة: تدفقات البيانات بسرعة غير مسبوقه ويجب التعامل معها في الوقت المناسب، تدفع علامات البيانات والمستشعرات والقياس الذكي الحاجة للتعامل مع السيول من البيانات في الوقت الفعلي تقريبًا.

يشير مصطلح «السرعة» إلى سرعة توليد البيانات، يحدد مدى سرعة إنشاء البيانات ومعالجتها لتلبية الطلبات، الإمكانيات الحقيقية في البيانات، تتعامل سرعة البيانات الضخمة مع السرعة التي تتدفق بها البيانات من مصادر مثل العمليات التجارية، وسجلات التطبيقات والشبكات، ومواقع الوسائط الاجتماعية، وأجهزة الاستشعار، والأجهزة المحمولة، وما إلى ذلك، إن تدفق البيانات ضخم ومستمر.⁽⁶⁾

3 - الشكل: تأتي البيانات بجميع أنواع التنسيقات - من مجموعات البيانات المنظمة ومنها البيانات الرقمية في قواعد البيانات التقليدية إلى المستندات النصية غير المهيكلة والبريد الإلكتروني والفيديو والصوت وبيانات مؤشر الأسهم والمعاملات المالية. يشير التنوع إلى المصادر غير المتجانسة وطبيعة البيانات، سواء المنظمة وغير المنظمة، خلال الأيام السابقة، كانت جداول البيانات وقواعد البيانات هي المصادر الوحيدة للبيانات التي نظرت فيها معظم التطبيقات، و الآن يتم النظر أيضًا في البيانات في شكل رسائل بريد إلكتروني وصور ومقاطع فيديو وأجهزة مراقبة وملفات PDF وصوت وما إلى ذلك في تطبيقات التحليل، يطرح هذا التنوع في البيانات غير المهيكلة بعض المشكلات المتعلقة بالتخزين والتعدين وتحليل البيانات.⁽⁶⁾

فوائد معالجة البيانات الضخمة:

- 1- تجلب القدرة على معالجة «البيانات الضخمة» فوائد متعددة مثل:
 - 1- يمكن للشركات الاستفادة من الذكاء الخارجي أثناء اتخاذ القرارات
 - 2- خدمة عملاء محسنة
 - 3- التحديد المبكر للمخاطر على المنتج الخدمات ، إن وجدت
 - 4- كفاءة تشغيلية أفضل⁽⁶⁾

جمع المعلومات والبيانات الضخمة:

نظرًا لأن البيانات الضخمة تشير إلى جمع البيانات من مصادر متباينة، فإن هذه الميزة تخلق حالة استخدام حاسمة لصناعة التأمين للانقضاض عليها، مثال: عندما ينوي العميل شراء وثيقة للتأمين على السيارات، يمكن للشركات الحصول على معلومات يمكن من خلالها حساب مستويات الأمان للقيادة في محيط المشتري وسجلات قيادته السابقة، على هذا الأساس يمكنهم حساب تكلفة التأمين على السيارة بشكل فعال من خلال اكتساب رؤية العملاء وذلك بعد تحديد تجربة العملاء وجعل العملاء مركز جذب الشركة له أهمية قصوى للمؤسسات.⁽⁶⁾

ومن الضروري الكشف عن الغش من خلال عمليات الاحتيال التي قد تطرأ على التأمين من الحوادث الشائعة، تعتبر حالة استخدام البيانات الضخمة للحد من الاحتيال فعالة للغاية. وعندما تباع وكالة التأمين تأمينًا، فإنها تريد أن تكون على دراية بكل احتمالات حدوث أشياء غير مؤاتية مع عملائها، مما يجعلها تقدم مطالبة وبالتالي فإن استخدام البيانات الضخمة مفيد جدا لنشاط الأعمال.⁽⁶⁾

3- استخدام تقنيات تنقيب البيانات في المجالات الصحية:

مقدمة:

سرعان ما أصبحت السجلات الصحية الإلكترونية (EHR) أكثر شيوعًا بين مرافق الرعاية الصحية. مع زيادة إمكانية الوصول إلى كمية كبيرة من بيانات المرضى ، يمكن لمقدمي الرعاية الصحية الآن تحسين كفاءة ونوعية مؤسساتهم باستخدام استخراج البيانات. منذ تسعينيات القرن الماضي ، استخدمت الشركات تعدين البيانات لأشياء مثل سجل الائتمان والكشف عن الاحتيال. والآن ، بدأ عدد من مؤسسات الرعاية الصحية بمشاهدة الفوائد المحتملة لتنقيب البيانات والتحليلات التنبؤية.

في مجال الرعاية الصحية ، أثبتت عملية استخراج البيانات فعاليتها في مجالات مثل الطب التنبؤي ، وإدارة علاقات العملاء ، واكتشاف الاحتيال وإساءة الاستخدام ، وإدارة الرعاية الصحية وقياس فعالية بعض العلاجات. الغرض من استخراج البيانات ، سواء كان يتم استخدامه في الرعاية الصحية أو الأعمال التجارية ، هو تحديد أنماط مفيدة ومفهومة من خلال تحليل مجموعات كبيرة من البيانات. وتساعد أنماط البيانات هذه على التنبؤ باتجاهات الصناعة أو المعلومات ، ثم تحديد ما يجب فعله حيالها.

التنقيب عن البيانات في مجال الصحة :

تعد تقنيات التحليل المتقدمة واستخراج البيانات من أهم الأدوات التي يتم استخدامها في المجالات الطبية والصيدلانية وخاصة في مجال استكشاف وتقييم الظروف الصحية السائدة والبحث في أسباب الأمراض واستكشاف السلوك المرضي في المجتمع ، في من أجل المساهمة في تطوير الخطط والسياسات الطبية والصحية المناسبة والحد من انتشار الأمراض ، حيثما وجدت البيانات الطبية والصحية ، يمكن استخدام تقنيات التحليل والاستكشاف لدراساتها وتحليلها واستكشاف كل ما من شأنه أن يساهم في تحسين الوضع الصحي بشكل عام في المجتمع ، وتطوير أداء المؤسسات الصحية ، وتقليل مخاطر التعرض للأمراض ، وتطوير سياسات وإجراءات التوعية الصحية للفرد والأسرة. في استكشاف وتوصيف الأمراض الأكثر شيوعًا في مناطق محددة ، أو أوقات معينة ، أو ظروف وظروف معينة ، بهدف تطوير الحلول المناسبة واتخاذ التدابير الوقائية اللازمة للحد من انتشار الأمراض ، بالإضافة إلى إعداد البحوث المتخصصة في دراسة الأدوية والعلاجات الطبية وسبل تطويرها وتحديثها ورفع كفاءتها وفعاليتها وصلاحياتها وقدرتها على العلاج.⁽⁶⁾

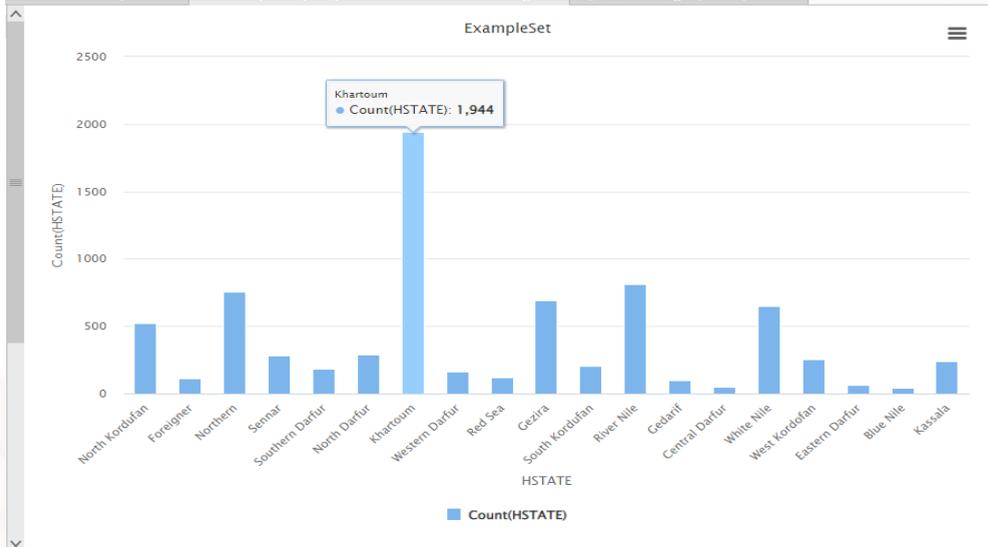
ثالثاً: الأشكال الإحصائية لتوزيع البيانات: 1- الشكل يوضح البيانات الإحصائية لنتائج الدراسة

Label	Polynomial	0	Least	Most	Values
STATUS	Polynomial	0	Child (15)	Married (5362)	Married (5362), Widower (1089)
AGE	Integer	0	Min 13	Max 96	Average 51.310
TRIBE	Polynomial	0	Bargho (1)	Jalia (1025)	Jalia (1025), Shaigi (736), ...[95]
JOB	Polynomial	0	farmer (1)	House wife (5482)	House wife (5482), Teacher (6)
HSTATE	Polynomial	0	Blue Nile (45)	Khartoum (1944)	Khartoum (1944), River Nile (8)
HCITY	Polynomial	0	Qurashi Twenty six (1)	Khartoum (1230)	Khartoum (1230), Omdurman (6)
GENDER	Polynomial	0	male (557)	Female (6942)	Female (6942), male (557)

Showing attributes 1 - 8 Examples: 7,499 Special Attributes: 1 Regular Attributes: 7

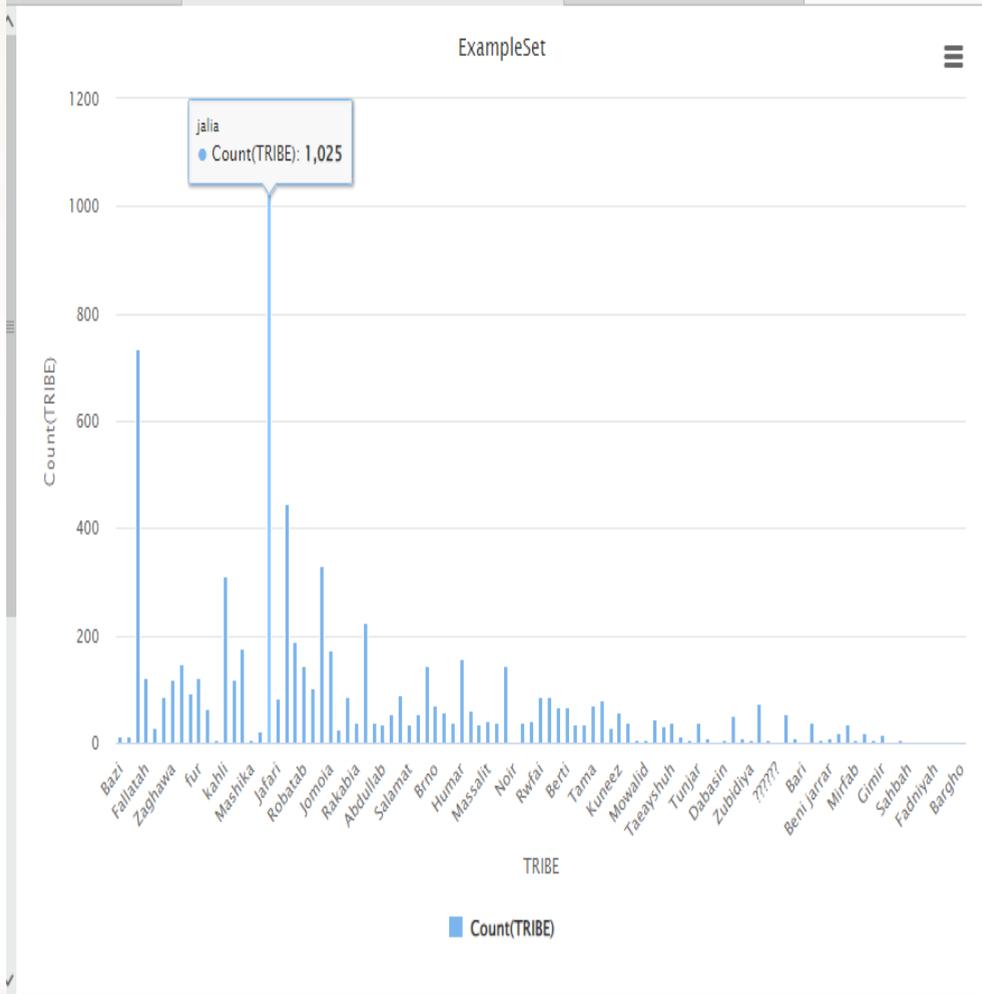
الشكل رقم (2) يوضح البيانات الإحصائية للدراسة التاريخ 2021 / 9 / 15م-المصدر مستشفى الذرة إدارة الإحصاء

2- الشكل يوضح عدد الاصابة في ولايات السودان حيث ان ولاية الخرطوم الاكثر إصابة بعدد 1944 إصابة بين ان ولاية النيل الازرق الاقل إصابة بعدد 45 إصابة .



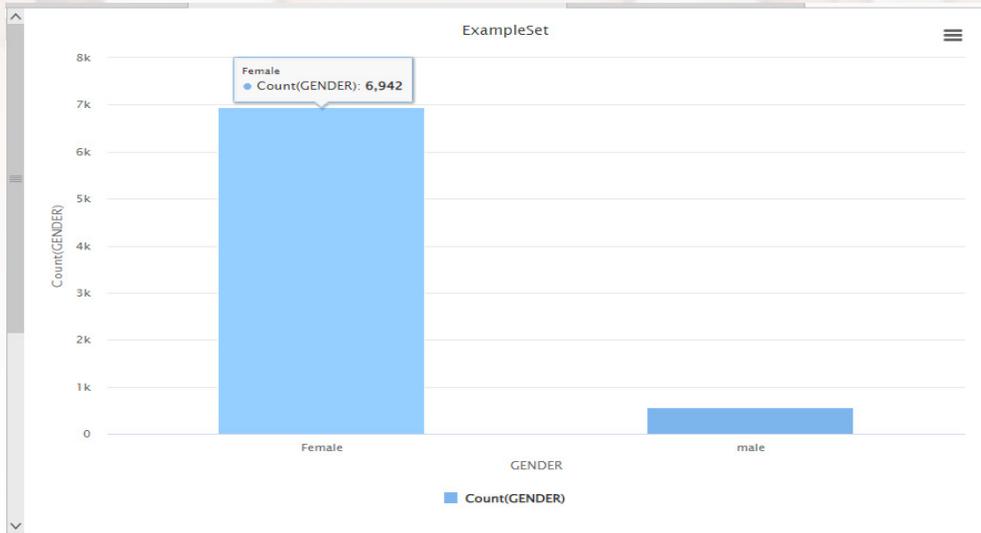
الشكل رقم (3) يوضح الولايات السودان التاريخ 2021 / 9 / 15م-المصدر مستشفى الذرة إدارة الإحصاء

3- الشكل يوضح القبائل الأكثر انتشارا للمرض حيث ان قبيلة الجعليين الأكثر بعدد 1025 إصابة والاقبل قبيلة البرقاوية (البرقو) بعدد 1 إصابة, التاريخ 15/9/2021م-المصدر مستشفى الذرة إدارة الاحصاء

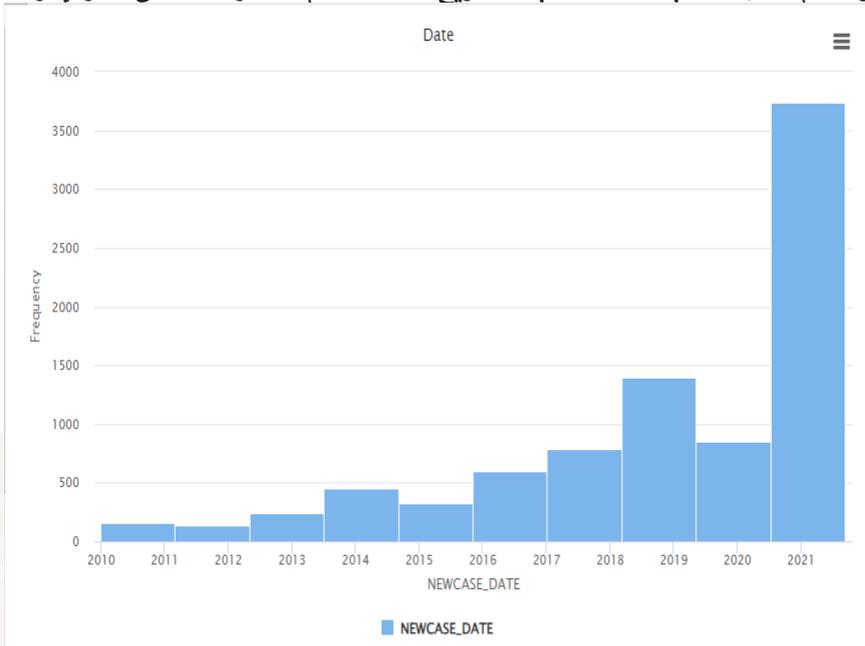


الشكل رقم (4) يوضح لاقبائل الأكثر إصابة بالمرض

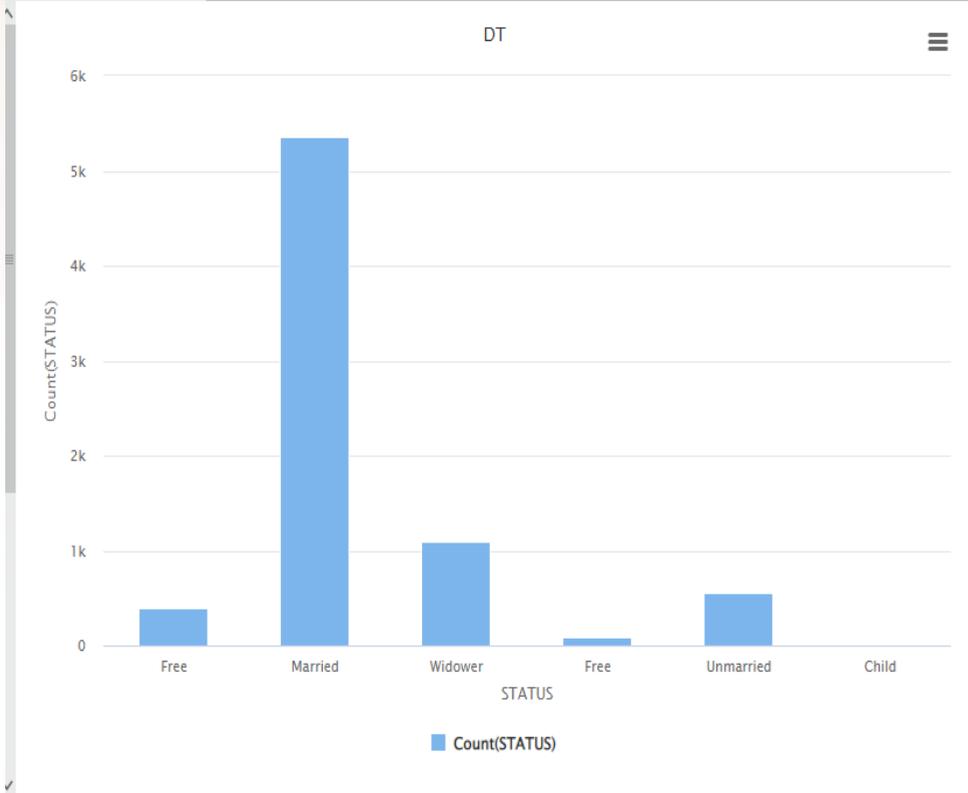
4- الشكل يوضح الإصابة حسب النوع حيث الإناث الأكثر إصابة بالمرض بعدد 6942 إصابة و الرجال الأقل بعدد 557 إصابة, التاريخ 2021 /9/15م-المصدر مستشفى الذرة إدارة الاحصاء



الشكل رقم (5) يوضح عدد الإصابة علي حسب النوع
5- الشكل يوضح أكثر الاعوام انتشارا للمرض حيث ان العام 2021 الأكثر انتشارا بعدد 3736 إصابة و العام الأقل 2012 بعدد 138 إصابة , التاريخ 2021 /9/15م-المصدر مستشفى الذرة إدارة الاحصاء



الشكل رقم (6) يوضح عدد الإصابات حسب الاعوام
 6- الشكل يوضح عدد الاصابة حسب الحالة الاجتماعية وكان العدد الاكبر بين المتزوجين
 بعدد 5357 اصابة والعدد الاقل 15 بين دون سن البلوغ . . التاريخ 2021 / 9/15م-المصدر مستشفى
 الذرة إدارة الاحصاء



الشكل رقم (7) يوضح الاصابة حسب الحالة الاجتماعية , التاريخ 2021 / 9/15م-المصدر
 مستشفى الذرة إدارة الاحصاء

رابعا : النتائج :

الخاتمة :

1.بعد الدراسة في بيانات سلطان الشدي في مستشفى الذرة الخرطوم فقد تبين ان ولاية
 الخرطوم الاكثر إصابة بعدد 1944 اصابة بين ان ولاية النيل الازرق الاقل إصابة بعدد 45 اصابة و
 ان القبائل الاكثر انتشارا للمرض قبية الجعلين بعدد 1025 اصابة والاقبل قبيلة البرقو بعدد 1 اصابة,
 وقد وصينا انه يجب تخزين بيانات المريض بطريقة أفضل لتكون في متناول اليد ، مما يساعد في
 استكمال عملية البحث في هذا المجال والتوعية والإرشاد للكشف المبكر عن هذا المرض و تحليل
 بيانات الاصابة في الاعوام القادمة و مقارنتها مع الاعوام السابقة.

النتائج :

- بعد استخدام الطرق الإحصائية في برنامج (Miner Rapid) تم التوصل للنتائج الآتية
1. ان ولاية الخرطوم الاكثر إصابة بعدد 1944 اصابة بين ان ولاية النيل الازرق الاقل إصابة بعدد 45 اصابة.
 2. ان القبائل الاكثر انتشارا للمرض قبية الجعلين بعدد 1025 اصابة والاقل قبيلة البرقو بعدد 1 اصابة.
 3. ان الاصابة حسب النوع حيث الإناث الاكثر اصابة بالمرض بعدد 6942 اصابة
 4. و الرجال الاقل بعدد 557 اصابة.
 5. ان الاعوام اكثر إنتشارا للمرض حيث ان العام 2021 الاكثر انتشارا بعدد 3736 اصابة و العام الاقل 2012 بعدد 138 اصابة.
 6. ان عدد الاصابة حسب الحالة الاجتماعية كان العدد الاكبر بين المتزوجين بعدد 5357 اصابة والعدد الاقل 15 بين دون سن البلوغ .

التوصيات:

1. إستخدم برنامج آخر غير (Rapid miner) وقارن بينها وبين برنامج النتائج مع بعضها البعض .
2. يجب تخزين بيانات المريض بطريقة أفضل لتكون في متناول اليد ، مما يساعد في استكمال عملية البحث في هذا المجال .
3. التوعية والإرشاد للكشف المبكر عن هذا المرض
4. تحليل بيانات الاصابة في الاعوام القادمة و مقارنتها مع الاعوام السابقة.

المصادر و المراجع :

1-المراجع العربية:

- (1) 1. د.زياد عبدالكريم القاضي، «مقدمة في تصميم قواعد البيانات»، «دار صفاء للطباعة والنشر والتوزيع»، 2004م.
- (2) 2. ياسر مطيع،محمد الراميني،تامر جلال،محمد نصرالله، «أساسيات قواعد البيانات»، «دار صفاء للطباعة والنشر والتوزيع»، 2005م.
- (3) 3. مراد شلباية،نهلة درويش،وائل أبو مغلي، «مفاهيم أساسية في قواعد البيانات»، «دار المسيرة للنشر والتوزيع و الطباعة»، 2002م.
- (4) 4. حجازي، محمد عثمان ، « محاضرات في برمجة قواعد البيانات» ، جامعة حائل كلية التربية، «2006
- (5) 5. د. أروى عيسى الياسري ، «إستخراج البيانات Data Mining اتجاه جديد في استرجاع المعلومات» ، «مجلة المعلوماتية - العدد السادس عشر» ، 2006.

2-المراجع الالكترونية:

- (1) <http://hdl.handle.net/1234567892076/> Date: 201810-01- Time: 01:33 .
- (2) <http://repository.sustech.edu/handle/1234567892076/> Date: 201818-01- Time: 01:35.
- (3) Liao M-N, Chen S-C, Lin Y-C, Chen M-F, Wang C-H, Jane S-W. Education and psychological support meet the supportive care needs of Taiwanese women three months after surgery for newly diagnosed breast cancer: A non-randomised quasi-experimental study. International journal of nursing studies. 2017; 51(3): 3909-.
- (4) Nikbakhsh N, Moudi S, Abbasian S, Khafri S. Prevalence of depression and anxiety among cancer patients. Caspian J Intern Med. 2017; 5(3): 16770-.
- (5) Haghghi M, Rahmati-Najarkolaei F, Ansarian A. Correlation between Spiritual Wellbeing and Religious Orientation among Staffs of one Military Medical University. Journal of Health Policy and Sustainable Health. 2018; 1(4): 13740-.

(6) <https://mafhome.com/%D985%D8%A7%D987%D98%A%D8%A7%D984%D8%A8%D98%A%D8%A7%D986%D8%A7%D8%AA%D8%A7%D984%D8%B6%D8%AE%D985%D8%A9big-data%D89%F>