استخدام تقنيات تنقيب البيانات في التنبؤ بحجم الصادرات والواردات
دراسة حالة جمارك الصمغ العربي في السودان

هبة إبراهيم محمد الحسن وسليمان مصطفى محمد
دائرة تقنية المعلومات، هيئة الجمارك، الخرطوم، السودان
مركز المعلومات الوطني، وزارة الداخلية، المملكة العربية السعودية

Using data mining techniques to forecast the volume of exports and imports
Case study custom value for Arabic Gum in Sudan

Hiba Ibrahim Mohammed Elhassan and Sulaiman Mustafa Mohammed
Affiliation: Sudan Customs- IT Department
Affiliation- Kingdom of Saudi Arabia
Ministry of Interior- National Data Center

المستخلص

تعتبر الجمارك أداة اقتصادية وجهازاً نظامياً لتنفيذ العديد من السياسات الاقتصادية التي تؤدي إلى التنمية الاقتصادية والاجتماعية، منها تنفيذ السياسات التجارية والصناعية (حماية المنتجات الوطنية) والسياسة النقدية والتمويلية، وتشجيع وجذب الاستثمار الأجنبي المباشر والوطني وتسهيل حركة الصادرات وحماية الاقتصاد الوطني مع توفير إحصائيات التجارة الخارجية والمساهمة في صياغة الاتفاقيات التجارية. يعد السودان أكبر دولة منتجة ومصدرة للصمغ العربي في العالم. ومعظم إنتاج الصمغ العربي من شجرتي الهشاب والطلح، كما أن الصمغ العربي من أهم السلع المصدرة في السودان إلي جانب القطن والثروة الحيوانية والبترول. تقدم هذه الورقة تحليلاً لبيانات من دورة الشهادة الجمركية لصادر ووارد الصمغ العربي في السودان في الفترة (من يناير 2006م إلى سبتمبر 2015م شهرياً)، وهي الفترة التي وجدت لها بيانات متصلة لفترات طويلة في الصادر والوارد، لدعم القرار بوزارة المالية والاقتصاد الوطني. تم استخراج بيانات السلعة من قاعدة البيانات

الرئيسية للجمارك السودانية كقيمة مالية وكميات، وبما أن القيمة المالية تحدد تبعاً للكميات، فقد تم أخذ القيمة المالية متغيراً تابعاً، والكميات متغيراً مستقلاً، فاعتمدت الكميات في التحليل، حيث كانت مساهمة الدراسة في دعم القرار تنبع من الحاجة الماسة لرئاسة مجلس الوزراء في السودان لتقارير هذه السلعة من صادر ووارد بالكميات. أظهرت النتائج أن حجم الصادرات أعلى بنسبة كبيرة من الواردات وأن كميات الصادرات والواردات متذبذبة مع الزمن وليس لها سلوك محدد.

كلمات مفتاحيه: تنقيب البيانات، تنبؤ، دورة الشهادة الجمركية، جمارك الصمغ العربي.

Abstract:

The Sudanese Customs Department is an important economic tool. It is the official organization for the implementation of many social and economic policies that include: commercial and industrial policies designed for the protection of domestic commodities, monetary and fiscal policies to attract local and foreign investments and policies for the promotion and easy movement exports. Other duties for the Sudanese customs department include completion of foreign trade statistics and participation in drafting foreign trade agreements.

Sudan is the largest producer and exporter of Gum Arabic in the world, which is extracted from Acacia trees. Sudan also exports cotton and livestock beside Gum Arabic. This paper analyzed the performance of monthly data obtained from customs Single Administration Document (SAD) cycle of custom value for Gum Arabic (both export and import) in Sudan for the period between January 2006 to September 2015. The data was obtained from the central data base of the Sudanese Customs Department. Quantitative analysis was used in the exercise, the data was required by the Sudanese council of ministers, and many economic policies depended on it. The results showed that the exported quantities were far

more than the imported quantities. The exported and the imported quantities both export and import quantities differ quality from month to month.

Keywords: Data Mining, Forecast, Single Administration Document (SAD) cycle, Gum Arabic Custom's value.

المقدمة

تعد هيئة الجمارك من أكبر البوابات الاقتصادية والأمنية بالسودان، حيث تهتم بمراقبة التبادلات التجارية على اختلاف طرقها ومواقعها من موانئ بحرية أو برية، مطارات، محطات جمركية حدودية كانت أم داخلية، وتحصيل الرسوم الجمركية والضريبية على الواردات والصادرات المقيدة، وتلعب دوراً مهماً في جذب المستثمرين وحماية المنتج المحلي، حيث يخضع نشاط الجمارك للقوانين المحلية وبعض الاتفاقيات الدولية عبر المنظمات العالمية مثل منظمة الجمارك العالمية، واتفاقيات التبادل الحر.

تكمن أهمية هيئة الجمارك الاقتصادية في وضع الخطط والسياسات وتطبيق الأنظمة الجمركية على البضائع المسموح بها عبر المنافذ البرية والبحرية والجوية، ومكافحة تهريب السلع غير المسموح بها للداخل أو الخارج.

كل ما سبق ذكره أوضح الحاجة إلى توافر نظام قوي يغطي جميع أنشطة الجمارك على اختلافها، وموزع على جميع محطاتها المختلفة، مع إمكانيات أمنية هائلة للطبيعة الأمنية للجمارك، وذلك للاعتماد الكبير على الجمارك كعمود فقري اقتصادي للسودان، ولموائمة التغيرات الاقتصادية المحلية والعالمية التابعة للاتفاقيات.

أولاً: نظم التنقيب عن البيانات

يهدف التنقيب في قواعد البيانات إلى استخلاص المعلومات المخبأة فيها، وهي تكنولوجيا حديثة فرضت نفسها بقوة في عصر المعلوماتية، وفي ظل التطور التكنولوجي الكبير وانتشار استخدام قواعد البيانات واستخدامها، الأمر الذي يوفر للدولة والمؤسسات

في جميع المجالات القدرة على استكشاف أهم المعلومات التي تتوفر في قواعد البيانات، حيث إنها تراقب السلوك، والأحداث لاستنتاج العلاقات بينها وبناء الارتباطات. كما تتولى البحث عن علاقات وأنماط موجودة في قواعد البيانات، ولكنها مخفية في البيانات المتراكمة، حيث يستفاد من هذه الأنماط المكتشفة في المقارنة أو المفاضلة بين أكثر من بديل.

يمكن تعريف نظم التنقيب عن البيانات كما يأتي: "الاستكشاف الآلي أو المؤتمت لأنماط شائعة ومخفية في قاعدة بيانات معينة": (قنديلجي، 2006، ص 157). كما تعرف بأنها:"سيرورة تحليل دقيقة وذكية، تفاعلية وتسلسلية، تسمح لمسيري النشاطات عند استخدام هذه السيرورة باتخاذ قرارات، والقيام بأعمال ملائمة في صالح نشاط المسؤولين عنه، والمؤسسة التي يعملون بها": (العلاق، 2005، ص 84). أو أنها"عبارة عن تحليلات لكمية كبيرة من البيانات بغرض إيجاد قواعد، أمثلة ونماذج يمكن أن تدل أصحاب القرار، وتتنبأ بالسلوك المستقبلي": (Bazsalica ,2001, p.61). كما يمكن تعريفها بأنها: "تحليل لمجموعات كبيرة الحجم من البيانات المشاهدة للبحث عن علاقات محتملة وتلخيص للبيانات في أشكال جديدة لتكون مفهومة ومفيدة لمستخدمها": (Mannila, 2001,p.1).

ثانياً: وسائل التنقيب في البيانات

هناك عدة وسائل مختلفة للتنقيب في البيانات، ويعتمد اختيار الوسيلة المناسبة على طبيعة البيانات قيد الدراسة وعلى حجمها، ومن هذه الوسائل ما يأتي:

1. التفكير واستخلاص النتائج والقوانين من أمثلة حية Real-Case-Based reasoning.

2. الكشف عن قانون Rule Discovery: البحث عن منوال معين أو علاقة معينة في جزئية كبيرة من البيانات.

3. معالجة الإشارات Signal Processing: إيجاد الظواهر المتشابهة مع بعضها البعض.

4. شبكات عصبية Neural Networks: تطوير نماذج قابلة لتتبؤ النتائج. هذه النماذج تم تطويرها بناءً على أسس تم استنباطها من عقل الإنسان.

5. منحنيات غير ثابتة Fractals: تصغير البيانات الكبيرة من دون ضياع المعلومات: (نورين،2011، ص8).

ثالثاً: مراحل عملية التنقيب في البيانات

يمكن تلخيص مراحل، وخطوات عملية التنقيب في البيانات كما يأتي:

1. فهم طبيعة الأعمال (Business Understanding): يعتبر المطلب الأول لاكتشاف المعرفة هو فهم المشاكل، والمسائل التي تواجهها الأعمال، وبمعنى آخر، كيف يمكن تحقيق المنفعة الأعظم من التنقيب في البيانات، مما يتطلب توفر صيغة واضحة ومحددة لأهداف الأعمال.

2. فهم البيانات: (Data Understanding): تعتبر مسالة معرفة ماهية البيانات وطبيعتها عاملاً مهماً في نجاح عملية التنقيب في البيانات واكتشاف المعرفة، حيث إن معرفة البيانات بصورة جيدة تعني مساعدة المصممين على استخدام الخوارزميات، أو الأدوات المستخدمة للمسائل المحددة بدقة عالية. وهذا يقود إلى تعظيم فرص النجاح، بالإضافة إلى رفع الفعالية والكفاءة لنظام اكتشاف المعرفة.

ويمكن تلخيص الخطوات الضرورية لعملية فهم البيانات كالآتي:

أ. تجميع البيانات (Data Collection): وهي الخطوة الموجهة نحو تحديد مصدر البيانات في الدراسة بما في ذلك استخدام البيانات العامة الخارجية، مثل الضرائب وغيرها.

ب. توصيف البيانات (Data Description): وهي الخطوة التي تركز على توصيف محتويات الملف الواحد من الملفات أو الجداول.

ت. جودة البيانات، والتحقق منها (Data Quality and Verification): هذه الخطوة تحدد ما إذا كان تقليل أو إهمال بعض البيانات غير الضرورية أو كونها رديئة

الجودة، وقد لا تنفع في الدراسة، لأنَ النموذج الجيد يحتاج إلى بيانات جيدة مما يتطلب صحة البيانات ودقة مضمونها.

ث. التحليل الاسترشادي للبيانات (Exploratory Analysis of Data): تستخدم أساليب مثل الإظهار المرئي أو التصور، أو عملية التحليل المباشر(OLAP) التي تؤدي إلى إجراء التحليل الأولي للبيانات. وتعتبر هذه الخطوة مهمة، وضرورية لأنَها تركز على تطوير الفرضيات المتعلقة بالمشكلة قيد الدراسة.

رابعاً: تهيئة البيانات (Data Preparation): وتشمل الخطوات التالية:

1. الاختيار (Selection) ويعني اختيار المتغيرات المتوقعة وحجم العينة.
2. صياغة المتغيرات وتحويلها (Construction and Transformation Variables) حيث يجب دائماً أن تصاغ المتغيرات الجديدة لبناء النماذج الفعالة.
3. تكامل البيانات (Data Integration): حيث إن مجاميع البيانات في دراسة التنقيب عن البيانات من الممكن خزنها في قواعد بيانات متعددة الأغراض التي تكون بحاجة إلى توحيدها في قاعدة بيانية واحدة.
4. تصميم البيانات، وتنسيقها (Data Formatting)، حيث تتعلق هذه الخطوة بإعادة ترتيب حقول البيانات، كما يتطلب نموذج التنقيب في البيانات.
5. صياغة نماذج الحل وثبوتها (Model Building and Validation): إنَ بناء وصياغة نموذج الحل السليم والدقيق يتم من خلال عملية الخطأ والصواب، حيث تحتاج هذه العملية في غالب الأمر إلى مساعدة المختصين في التنقيب عن البيانات، بهدف اختبار وفحص مختلف البدائل، للحصول على أفضل نموذج لحل المشكلة قيد الدراسة.
6. التقييم وتحليل نتائج النموذج (Evaluation and Interpretation): حالما يتم صياغة النموذج والتحقق من ثباته وصدقه، تجرى عملية التحقق من ثبات حزمة البيانات التي يتم تغذيتها بواسطة النموذج. وبما أن نتائج هذه البيانات معروفة، لذا فإن النتائج المتوقعة تقارن مع النتائج الفعلية في ثبات حزمة البيانات قيد التشغيل. وتؤدي هذه المقارنة أو المفاضلة إلى التحقق من دقة النموذج.

7. نشر وتوزيع النموذج (Model Deployment): حيث تشتمل هذه الخطوة على نشر وتوزيع النموذج داخل المنظمة لمساعدة عملية صنع القرار. وأن النموذج الصالح يجب أيضاً أنَ يحقق الرضا لدى المستفيدين طالما أنَ اختيار النموذج لابد أن يتم من خلال الدراسة الاسترشادية أو نموذج مصغر من الدراسة الشاملة: (قنديلجي،2006، ص 100، 102).

خامساً: كيفية عمل نظام التنقيب عن البيانات

1. تعمل نظم التنقيب في البيانات باستخدام أساليب رياضية وإحصائية وطرق ذكية، على بناء التنبؤات المستقبلية واستكشاف السلوك والاتجاهات، مما يسمح بتقدير القرارات الصحيحة واتخاذها في الوقت المناسب.

2. تساعد على اتخاذ قرار سليم يؤدي إلى اقتصاد معافى، وذلك بعد دراسة وافية لحركة صادرات وواردات الصمغ العربي في السودان، وتستخدم هذه النظم في مجالات مختلفة، مثل التنبؤ بكميات صادر الصمغ العربي خلال العام القادم، ومنها بناء إستراتيجية اقتصادية سليمة كما في هذه الدراسة.

3. إن جوهر نظم دعم اتخاذ القرار هو التنبؤ والإنذار المبكر وصياغة السيناريوهات المبنية على نماذج المحاكاة، حيث تقوم نظم دعم القرار بعمل المزج بين البيانات المتاحة مع الرؤى الشخصية لمتخذ القرار، ويتم ذلك عبر مجموعة من النماذج الرياضية للتنبؤ، والمحاكاة.

4. نموذج التنبؤ: وتلعب نماذج التنبؤ دوراً مهماً في إمداد متخذي القرار بالتنبؤات والمعلومات المهمة بوقت كافٍ قبل وقوع الأزمات الناتجة عن المخاطر، سواء أكانت مخاطر ناتجة عن ظواهر طبيعية أم كانت بفعل الإنسان ولكنها غير متعمدة.

5. تلعب الأساليب الكمية (الرياضية والإحصائية) وتكنولوجيا الحاسب الآلي والاستشعار عن بعد دوراً أساسياً في عملية التنبؤ، حيث تتيح هذه التكنولوجيات إمكانية القياس

والمراقبة والرصد، وبالتالي إمداد فريق دعم القرار بالتحذيرات والتنبؤات بالأزمات التي يمكن حدوثها. ومن ثم يمكن تجنب الآثار السلبية أو أخذ الاحتياطات اللازمة للتخفيف من المخاطر: (نورين،2011، ص 2،3).

سادساً: السلاسل الزمنية:

السلاسل الزمنية عبارة عن قيم ظاهرة من الظواهر في سلسلة تواريخ متلاحقة أياماً أو أشهر أو سنوات، والهدف من وراء تحليل سلسلة زمنية هو دراسة التغيرات التي تكون قد طرأت على الظاهرة التي تمثلها خلال فترة من الزمن وتحليل أسبابها ونتائجها، أو التنبؤ اعتماداً على فكرة مدى تأثير حوادث الماضي على المستقبل. فالتنبؤ بالصادرات المستقبلية مثلاً يُمكِّن الإدارة من تخطيط احتياجاتها من عمالة ومواد أولية وموارد أخرى، بحيث تصبح في متناول اليد عند الاحتياج إليها (الصيرفي، 2012 ، ص 530 ،531).

ويمكن القول بشكل عام أن التغيرات التي تطرأ على ظاهرة ما خلال فترة من الزمن هي محصلة عدة عوامل، ولا يمكن أن نعزوها لعامل واحد من هذه العوامل، وإنَما يكون نتيجة هذه العوامل مجتمعة، وبتحليلنا للسلسلة الزمنية نتعرف على مقدار هذه التغيرات وإدراك طبيعتها واتجاهها، ويصبح في الإمكان القيام بالتقديرات والتنبؤات المستقبلية الضرورية. وهذه العوامل هي:

1. الاتجاه العام Secular Trend: يعد الاتجاه العام هو الجزء الرئيسي، فعلى الرغم من وجود تعرجات في المنحنى التاريخي للسلسلة الزمنية نجد أن هناك اتجاهاً خاصاً تتجه نحوه الظاهرة على مدى زمني طويل، وهذا الاتجاه قد يكون تصاعدياً أو تناقصياً، وقد يكون الاتجاه العام خطياً أو غير خطي.

2. التغيرات الموسميةSeasonal Variation: تشير التغيرات الموسمية إلى تلك التغيرات التي تحدث بصفة دورية في فترة زمنية مدتها أقل من سنة، وتجدر الإشارة إلى أن هذه التغيرات الموسمية التي تحدث في فترات زمنية سنوية تعتبر من أكثر التغيرات تعرضًا للدراسة. وترجع التغيرات الموسمية إلى عدد من العوامل، منها التغير في حالة

الجو، العادات، التقاليد، الأعياد والمواسم. وتعتبر التغيرات في حالة الجو من أهم العوامل التي تؤدي إلى تغيرات موسمية في الإنتاج الزراعي (زهدي ،2008، ص 19،21).

3. التغيرات الدورية والعرضية Cyclical Variation: تشير التغيرات الدورية إلى التغيرات الدورية المتكررة أعلى أو أسفل منحنى خط الاتجاه العام، وترجع هذه التغيرات إلى عوامل كثيرة منها، التغير في عرض السلع، والخدمات، وفى الطلب عليها، والسياسات الحكومية، والعلاقات الدولية (الصيرفي،2012، ص 500، 540).

4. التغيرات غير المنتظمة Irregular Variation: وهي التغيرات التي تحصل في الاتجاه العام للسلسلة، والتي لا يمكن اعتبارها موسمية أو دورية، مثل التغيرات الناتجة عن الحروب والكوارث الطبيعية أو غيرها (عبد الوهاب،2005، ص5، 7).

تعتبر تقنيات مخازن البيانات من التقنيات الحديثة في مجال قواعد البيانات العلائقية، وتستخدم في تخزين واسترجاع أنواع مختارة من البيانات المستخرجة من مجموعة مختلفة من قواعد بيانات نظم العمليات، ويتم إدراج هذه البيانات في بناء منظم لقاعدة بيانات المخزن. تكون هذه البيانات مختومة بالزمن، وتستخدم في ترشيد القرارات الإدارية في مختلف المستويات الإدارية، كلُ حسب اختصاصه، وذلك من خلال تقنيات المعالجة التحليلية الآنية، التنقيب عن البيانات، واستخدام لغة المعالجة الاستفسارية. تسمح هذه التقنيات بدخول قاعدة بيانات المخزن، ومن ثم إعداد التقارير العامة والخاصة لترشيد قرارات بحد ذاتها، وتكون على ثلاثة أنواع آنية، خاصة وتقارير مسبقة الإعداد (عثمان،2004، ص 1، 2). كما تم استخدام التنبؤ في تنقيب البيانات لترشيد القرارات في مجالات عديدة، نأخذ المصارف كمثال، حيث توجد في المصارف كمية كبيره من المعلومات في مخازن البيانات التي تخص القروض والزبائن، فعندما نستخدم تقنية التنقيب في البيانات لبناء نماذج تتنبأ بمخاطر القروض استناداً على حالات زبائن مختلفة، ذلك يساعد كثيراً في الحكم على طلبات الاقتراض، واستبعاد ماتنطبق عليها النماذج كونه زبوناً خطراً، والتركيز على مقترضين موثوق بهم في التسديد (الردايده،2012، ص2).

هيئة الجمارك السودانية هي المصدر الوحيد لبيانات الواردات والصادرات للعديد من الجهات الحكومية، مما جعل البيانات مصدراً مهماً لاتخاذ القرارات التي تخص اقتصاد السودان في مختلف بنوده، وتوفر البيانات لعدة سنوات يجعل منها بيئة ممتازة لإجراء البحوث، والخروج منها بكمية هائلة من النتائج التي يمكنها أن تحدد مسار الاقتصاد السوداني للسنوات المقبلة، وتوضيح نقاط القوة والضعف في السياسة الاقتصادية المتخذة في السودان (التقرير السنوي لهيئة الجمارك،2008، ص2،7).

يعتبر السودان أكبر دوله منتجة ومصدرة للصمغ العربي في العالم ومعظم إنتاج الصمغ العربي من شجرتي الهشاب والطلح، والصمغ العربي من أهم السلع المصدرة في السودان إلى جانب القطن والثروة الحيوانية والبترول، حيث يتم إنتاج الصمغ العربي بواسطة المنتجين الصغار في مناطق الزراعة المطرية التقليدية، خاصة في غرب ووسط السودان (تقرير البنك الدولي،2007، ص 2،3).

تقدم الورقة تحليل لبيانات من دورة الشهادة الجمركية لصادر ووارد للصمغ العربي في الفترة (من يناير 2006م إلى سبتمبر2015م شهرياً)، وهي الفترة التي وجدت لها بيانات متصلة لفترات طويلة في الصادر والوارد لدعم القرار بوزارة المالية والاقتصاد الوطني. تم استخراج بيانات السلعة من قاعدة البيانات الرئيسية للجمارك كقيمة مالية وكميات، وبما أن القيمة المالية تحدد تبعاً للكميات تم اعتماد القيمة المالية كمتغير تابع والكميات متغير مستقل، فاعتُمدت الكميات في التحليل. تحتاج رئاسة مجلس الوزراء لتقارير هذه السلعة من صادر ووارد بالكميات لدعم قراراتها، ولذلك أجريت هذه الدراسة التي تهدف إلى التنبؤ بكميات صادرات وواردات الصمغ العربي في السودان.

هذه الدراسة تكتسب أهميتها في أنَ السودان يمتلك موارد طبيعية ضخمة، خاصة فيما يتعلق بالإنتاج الزراعي والحيواني والبترول، وله ميزة نسبية تميزه عن بقية أقطار العالم في إنتاج الصمغ العربي، وبالرغم من ذلك لا يوجد استغلال أمثل لهذه الموارد، لذلك ركزت الدراسة على تحليل معدلات النمو بالنسبة لواردات الصمغ العربي التي تنحصر في طلاء صمغي للأقمشة، مستخلص الصمغ الطبيعي، مستخلص الصمغ لاستخدامات

الطباعة، بدرة صمغ نقية، رماد الصمغ، مطحون الصمغ، مواد جيلاتينية من الصمغ، مواد مكثفه من الصمغ، مواد مخاطية ومكثفات من الصمغ، الصمغ العربي المنظف، لصادرات الصمغ العربي، حيث يصدر السودان خام الصمغ العربي، ويستورد مستخلصاته بعد أن يتم تنظيف الخام وإعادة تعبئته في أشكال مختلفة. وقد أخذت الدراسة الصمغ العربي كسلعة صادرة وواردة لمعرفة مدى استقرار الصادرات والواردات في المستقبل.

الطريقة والأدوات

تم إجراء هذا التحليل بهيئة الجمارك السودانية من قاعدة البيانات الرئيسية لهيئة الجمارك السودانية، في الفترة الزمنية من يناير 2006م إلى سبتمبر2015م شهرياً للتنبؤ بكميات الصادر والوارد للصمغ العربي في السودان، حيث تم تطبيق تقنيات تنقيب البيانات على البيانات التي تم الحصول عليها من دورة الشهادة الجمركية لصادر ووارد الصمغ العربي في السودان، وذلك باستخدام حزمة مايكروسوفت أوفيس، برنامج الإكسل، حزمة قواعد بيانات مايكروسوفت برنامج خادم لغة الاستفسار المهيكلة (Data Mining With Microsoft SQL Server and Excel 2007)

استخدم المنهج التحليلي، والمنهج الاستنباطي، والاستقرائي، وذلك بالاعتماد على أسلوب التحليل الوصفي، والاستعانة بالأسلوب الرياضي والإحصائي، والأشكال البيانية، لتحليل البيانات واستخراج النتائج، ومن ثم التنبؤ ببيانات صادر ووارد الصمغ العربي التي تم استخراجها من قاعدة بيانات هيئة الجمارك الرئيسية في الفترة من (من يناير 2006م إلى سبتمبر2015م شهرياً) وذلك باستخدام MICROSOT SQL SERVER DATAMINING WITH EXCEL 2007.

برنامج: Microsoft SQL SERVER 2008 DATAMINING WITH EXCEL 2007

يسمى برنامج خادم لغة الاستفسار المهيكلة الذي يمكن تعريفه (Structured Query Language Server 2008) بأنه نظام إدارة قاعدة البيانات للمساعدة في إدارة البيانات، فهو يقدم طرائق لمسح وتحديث وإدخال البيانات التي يمكن من خلالها

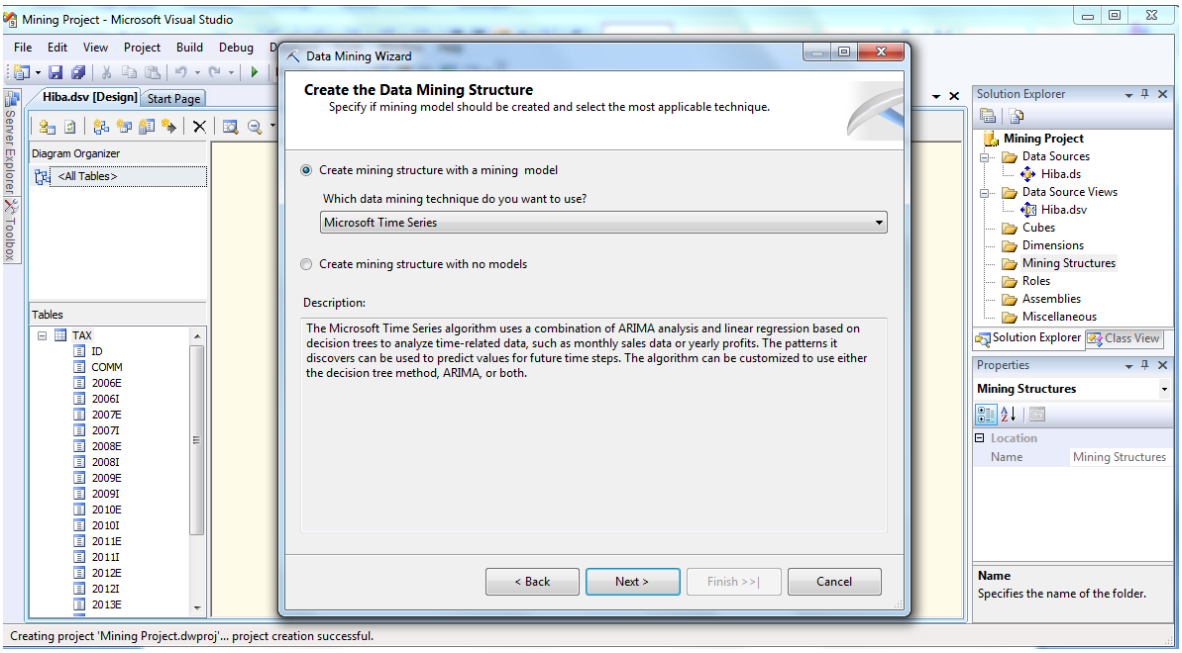
الحفاظ على قاعدة بيانات متجددة، يوفر البرنامج واجهة سهلة التعامل مع جميع خوارزميات تنقيب البيانات.

(Decision Trees ،Naïve Bayes ،Neural Nets،Clustering ،Sequence Clustering،Association Rules ،Time Series) .(Loría,2005,p.46.

يتيح النظام إمكانية استخراج التقارير، إدارة قواعد البيانات، بناء النماذج واستكشافها والتحليل (ZhaoHuiTang,2005,p.149).

تم تصميم البرنامج للتعامل المتقدم مع السلاسل الزمنية، حيث يمكنه التوقع بالطقس مشمس أو ممطر، توقع حركة موقع على الإنترنت، والتوقع من برامج الإدارة المستخدمة من قبل شركات المال والاقتصاد (DANIEL T. LAROSE, 2005, p.5-6-7)، تظهر واجهه البرنامج كما في الشكل رقم (1)، وهي واجهة رسومية تتعامل مع خوارزميات التنقيب في البيانات.

شكل (1) برنامج خادم لغة الاستفسار المهيكلة



النتائج

تم تقديم دراسة تحليلية عن الصادرات والواردات لسلعة الصمغ العربي في السودان في الفترة (من يناير 2006م إلى سبتمبر 2015م شهرياً)، وتقوم بالتنبؤ بواردات وصادرات الصمغ العربي في السنة القادمة.

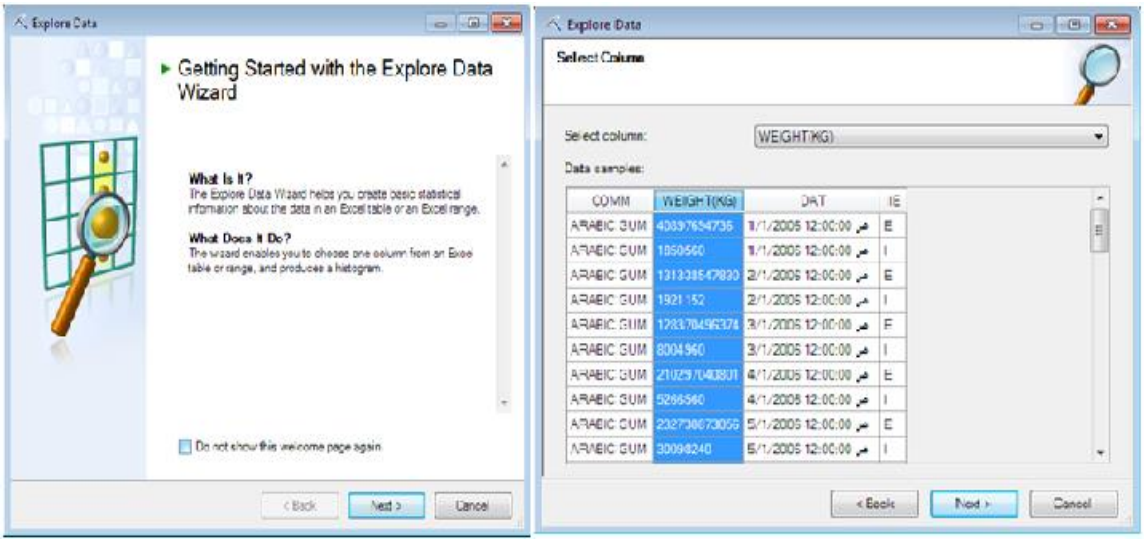
- جمع البيانات: تم جمع البيانات عن طريق استعلامات (SQL) من قاعدة بيانات أوراكل واستيرادها إلى ملف إكسل، كما في الجدول رقم (1) الذي يصف حقول البيانات كالآتي: اسم السلعة، وهي الصمغ العربي، وزن السلعة بالكيلوجرام في حالتي الصادر والوارد، تاريخ التصدير/ الاستيراد، وتحديد نوع العملية فيما إذا كان وارداً أو صادراً.

جدول (1): وصف حقول البيانات المستخدمة في الدراسة التحليلية

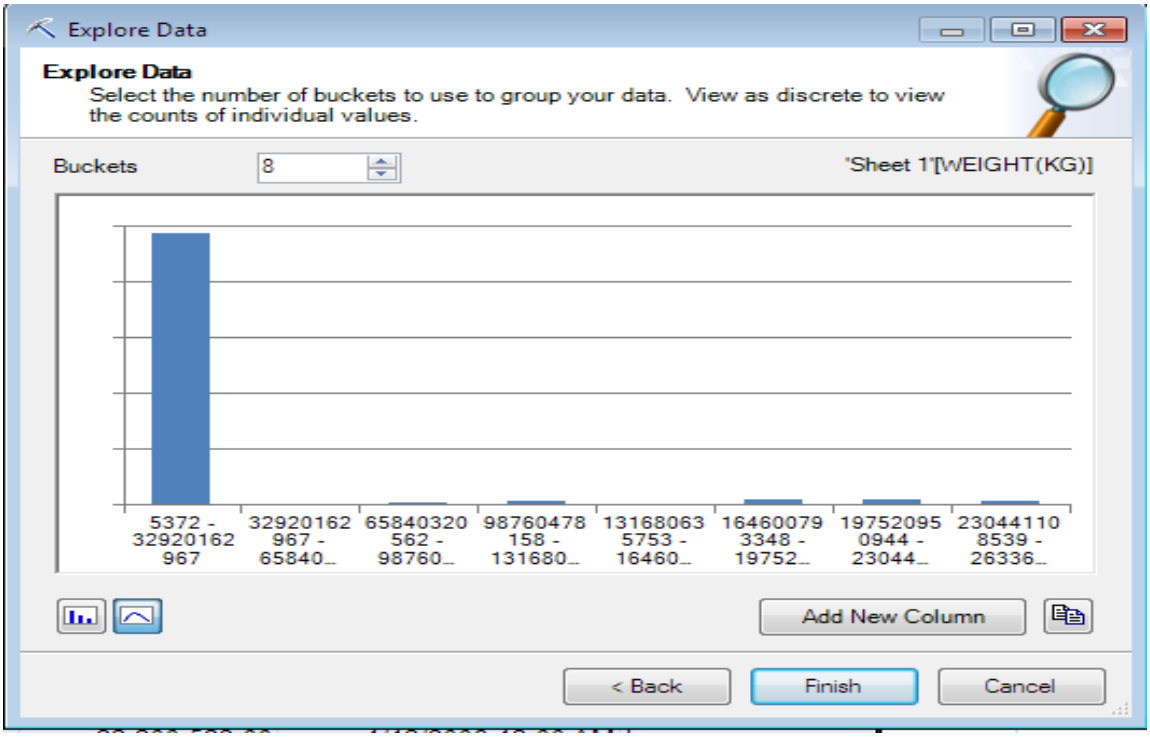
وصف الحقل	اسم الحقل	نوع الحقل
اسم السلعة	COMM	STRING
وزن الصادر/ الوارد بالكيلوجرامات	WEIGHTKG	DOUBLE
وارد/ صادر	IE	STRING
التاريخ	DAT	DATE

- تجهيز البيانات: عرض البيانات، يتم فتح ملف البيانات في البرنامج وعرضها، اخترنا حقل الوزن لعرضه باعتباره المتغير المستقل، كما يظهر في الشكل رقم (2).

شكل (2):عرض البيانات المستخدمة في الدراسة التحليلية لاختيار حقول التحليل



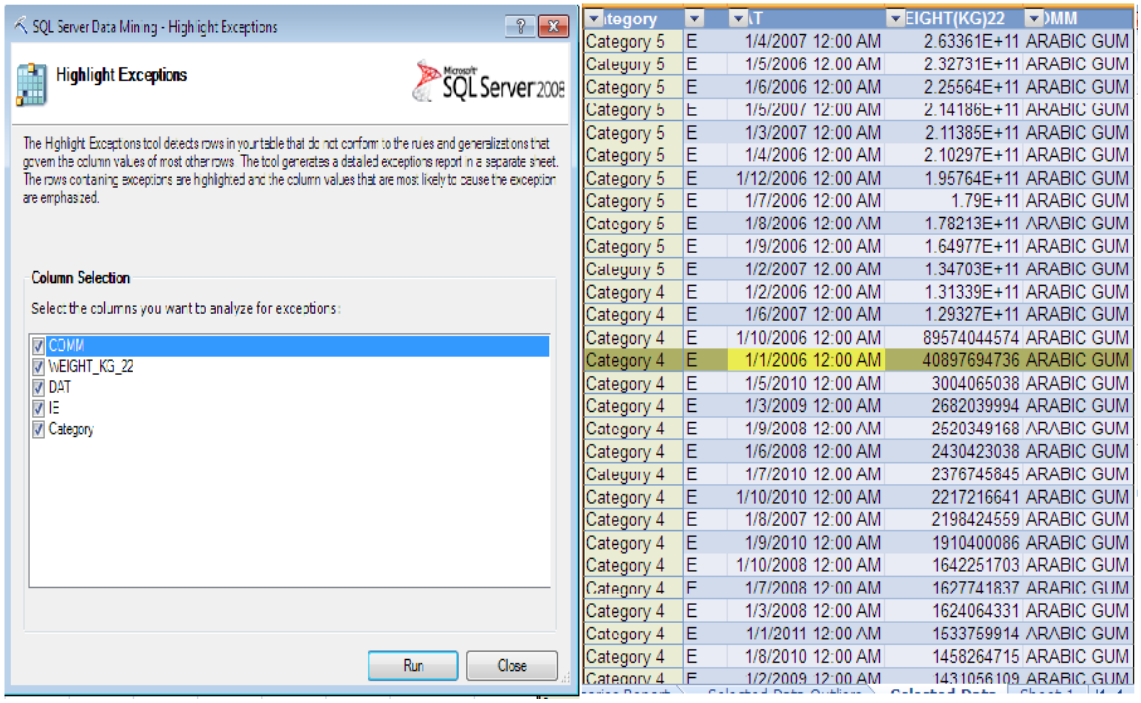
شكل (3): توزيع البيانات ويظهر غير طبيعي ومتقطع



تلاحظ أن التوزيع غير طبيعي كما في الشكل رقم (3)، وذلك لتمركز البيانات حول نقطه واحدة.

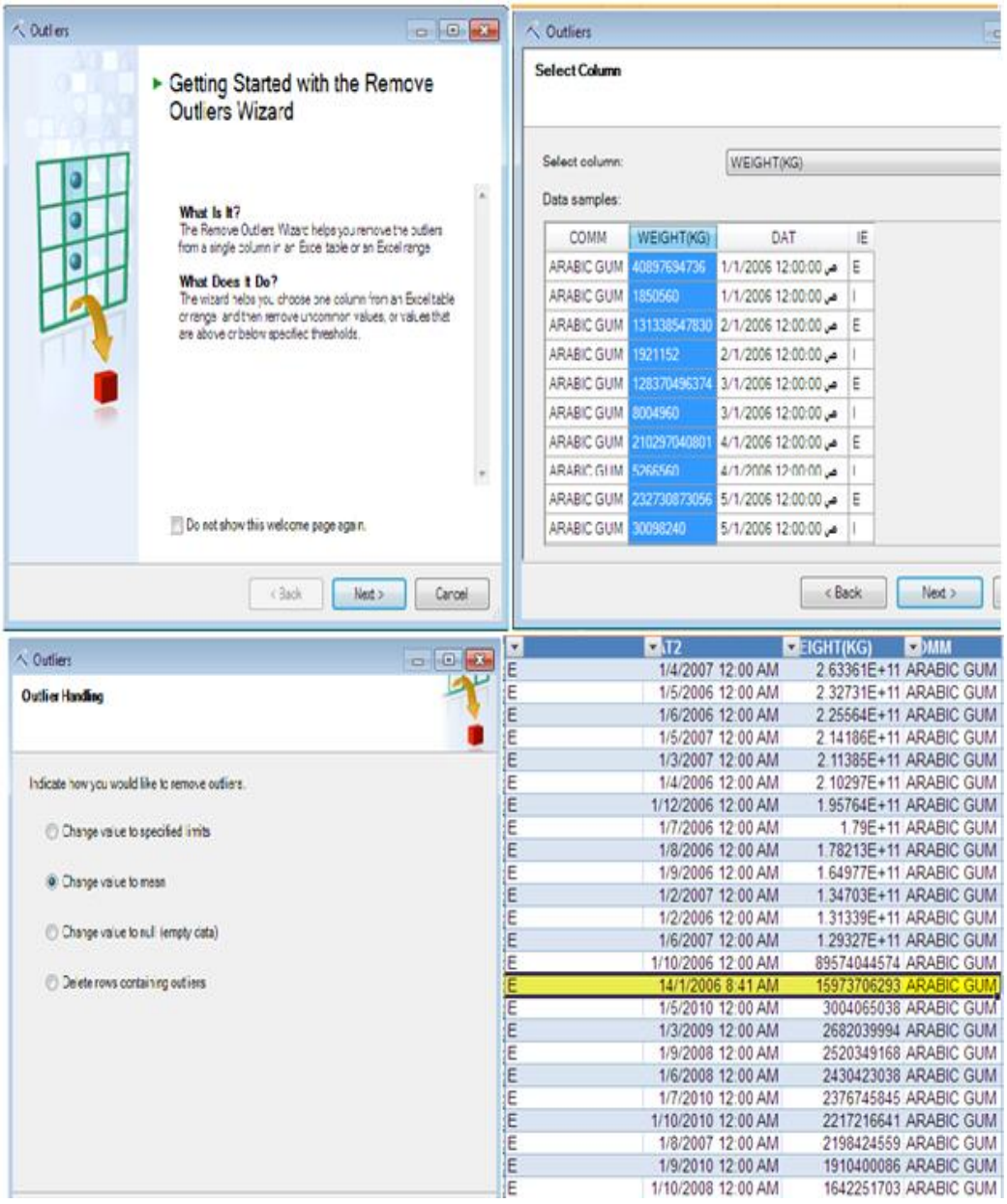
تم تحديد القيم الشاذة، وهي قيمة كبيرة جداً بالنسبة لسلوك البيانات في شهر يناير من كل عام، حيث تكون حركة الصادر والوارد ضعيفة في مثل هذا الوقت من العام، كما يظهر في الشكل رقم (4):

شكل (4): عملية تحديد القيم الشاذة بالبيانات المستخدمة في الدراسة التحليلية



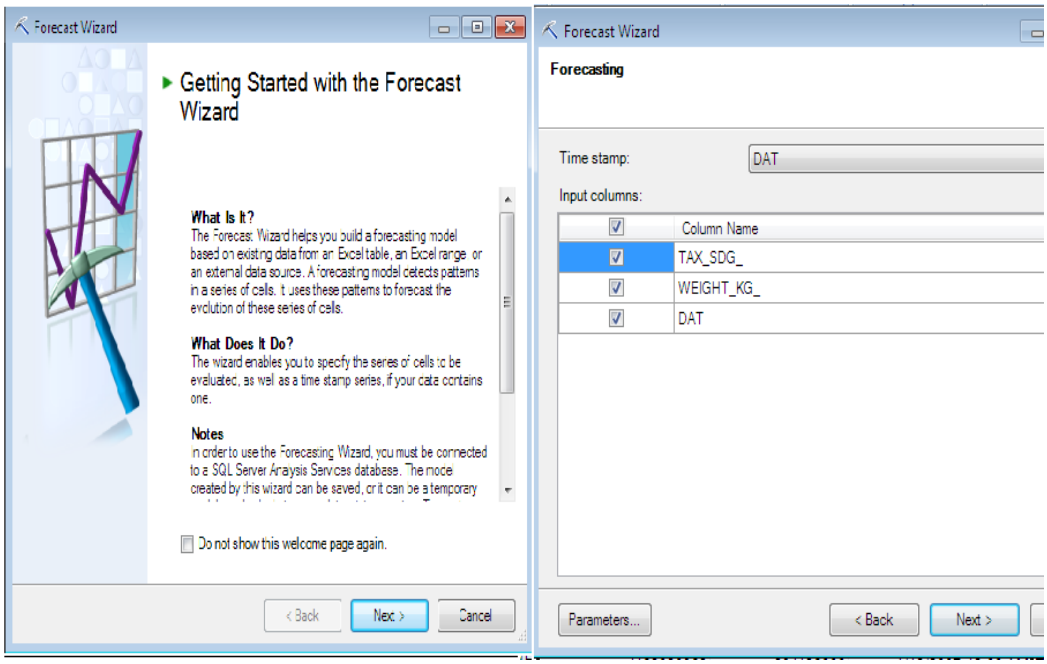
معالجة القيم الشاذة، تمت معالجة القيم الشاذة، وذلك بإضافة قيمة الوسيط، كما يظهر في الشكل رقم (5):

شكل (5): معالجة القيم الشاذة باستخدام الوسيط

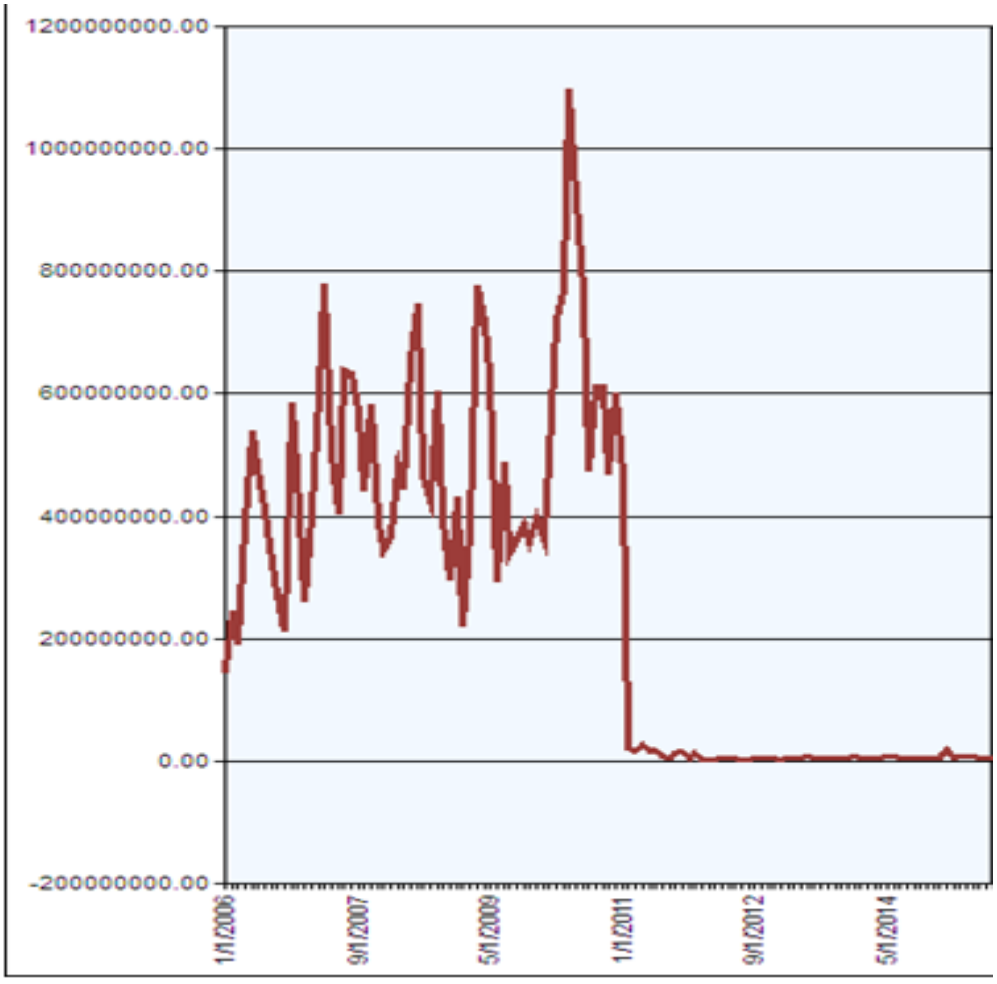


- بناء نموذج التنبؤ بحجم الصادر والوارد المتوقع للصمغ العربي في السودان لعام قادم، تم تحديد حقول الاختبار، وهي الكمية (الوزن بالكيلو جرام)، حيث إن الكمية (الوزن بالكيلوجرام) يتم التنبؤ به تبعاً للتاريخ، كما يظهر في الشكل رقم (6):

شكل (6): بناء نموذج التنبؤ بحجم الصادر والوارد المتوقع للصمغ العربي في السودان

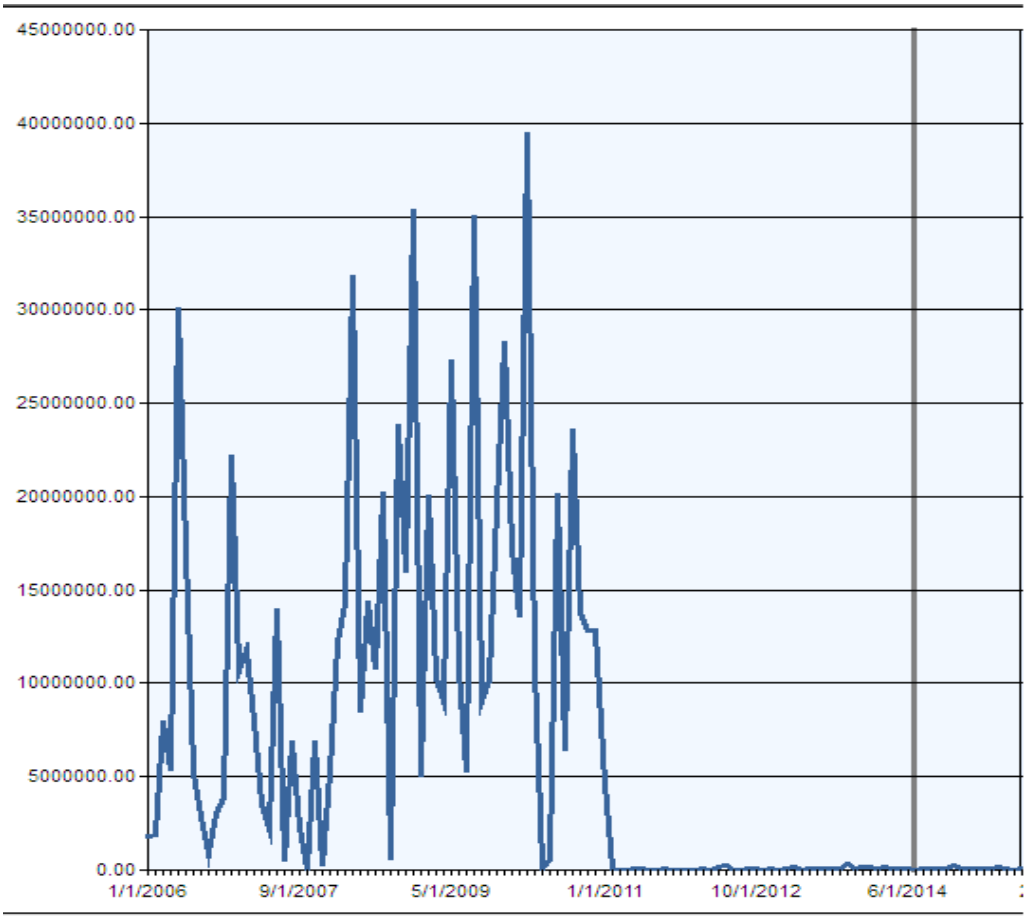


شكل (7): حجم صادرات الصمغ العربي الفعلية في السودان



الشكل رقم (7) يوضح حجم صادرات الصمغ العربي الفعلية في السودان في الفترة الزمنية من يناير 2006 وحتى سبتمبر 2015م، حيث بلغ الصادر الفعلي أعلى كمية له (1096547008) كجم في مايو 2010، وبلغ أدنى كمية له (1472340) كجم في يناير 2012. الملاحظ أنه لا يوجد اتجاه واضح لمسار الصادرات، فهي في تصاعد وانخفاض مستمر على حسب الموسم.

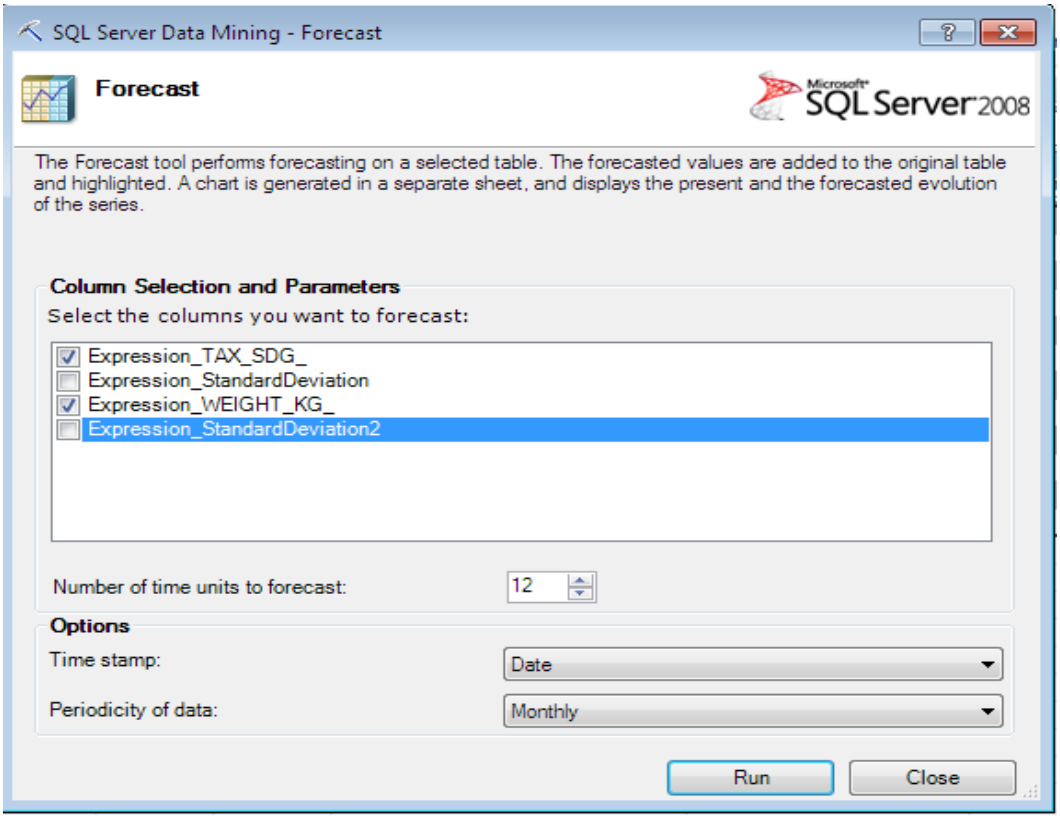
شكل (8): حجم واردات الصمغ العربي الفعلية في السودان



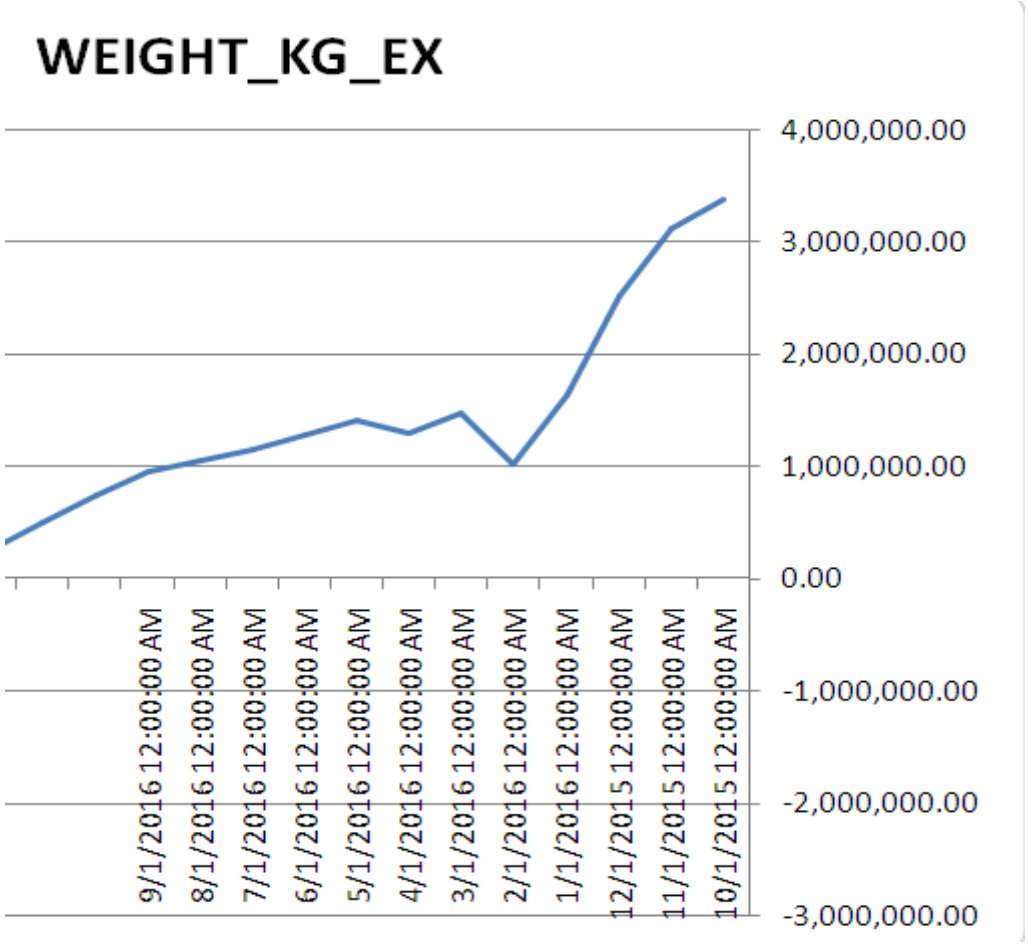
الشكل (8) يوضح واردات الصمغ العربي الفعلية في السودان شهرياً في الفترة من يناير 2011 وحتى سبتمبر 2015، حيث بلغت أعلى كمية لها (4592856) كجم في يناير 2011 وأدناها (5372) كجم في أبريل 2012، الملاحظ أنه لا يوجد اتجاه واضح لمسار الواردات، فهي في تصاعد وانخفاض مستمر على حسب الموسم.

- تحليل نموذج التنبؤ بحجم صادرات وواردات الصمغ العربي في السودان، حيث يتم تحديد الحقول المطلوب التنبؤ بها (الكمية)، وعدد الأشهر المطلوب التنبؤ بها، كما يظهر في الشكل رقم (9):

شكل (9): تحليل نموذج التنبؤ بحجم صادرات وواردات الصمغ العربي في السودان لعام مقبل

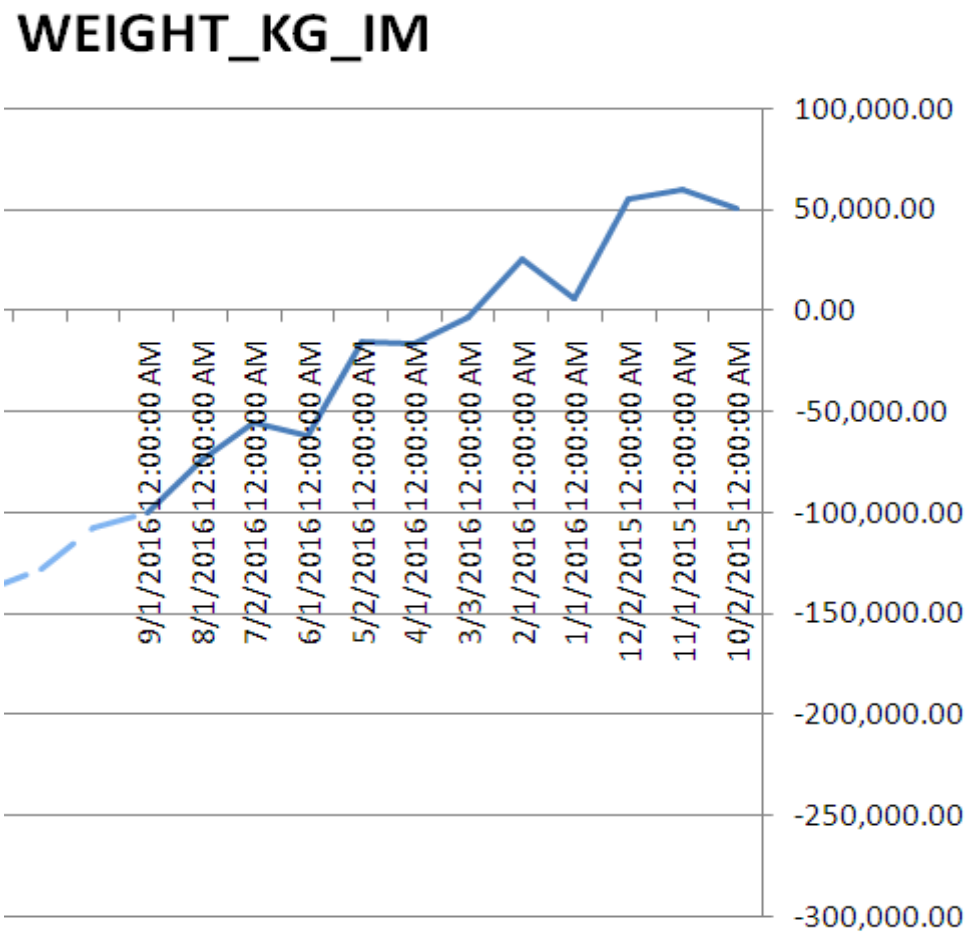


شكل (10): التنبؤ بصادرات الصمغ العربي في السودان



شكل (10) يوضح التنبؤ بالسنة القادمة لصادرات الصمغ العربي في الفترة من أكتوبر 2015 وحتى سبتمبر 2016، حيث من المتوقع أن تبلغ أعلى كيمة لها في يناير 2015، وهي 3367881 كجم، وأدنى كمية لها في نوفمبر 2016، وهي 941619 كجم.

شكل (11): التنبؤ بواردات الصمغ العربي في السودان لعام مقبل



شكل(11) يوضح التنبؤ للعام القادم لواردات الصمغ العربي في الفترة من أكتوبر 2015 وحتي سبتمبر 2016، حيث من المتوقع أن تبلغ أعلى كمية لها في نوفمبر 2015، وهي 59584 كجم، وأدنى كمية لها في نوفمبر 2016، وهي -100218 كجم.

المناقشة:

بعد دراسة صادر ووارد الصمغ العربي في السودان في الفترة (من يناير 2006 إلى سبتمبر 2015م شهرياً)، وكما يظهر في الرسوم البيانية، توصلت الدراسة للآتي:

1. حجم صادرات الصمغ العربي أعلى بنسبة كبيرة من حجم واردات الصمغ العربي، كما يظهر في الشكلين (7،8).
2. حجم صادرات الصمغ العربي متذبذبة مع الزمن، وتختلف من موسم لآخر، وليس لها سلوك محدد، وإن حجم واردات الصمغ العربي أيضاً متذبذبة مع الزمن، وتختلف من موسم لآخر وليس لها سلوك محدد، كما يظهر في الشكلين (8،7).
3. من الممكن التنبؤ بكميات الصادر والوارد للصمغ العربي في المستقبل، مما يساعد على تفادي الانهيار الاقتصادي، وأيضاً يساعد في معرفة السلع الأنسب للاستثمار، كما يظهر في الشكلين (10،11).
4. تتنبأ الدراسة بتدني صادرات وواردات الصمغ العربي في السودان، ولكن بصورة أعلى في حالة الواردات، وهو مؤشر جيد في حالة الوارد، ولكن تدني الصادر مؤشر غير جيد، زيادة الصادر تقابلها نقص في الوارد كما في الشكلين (11، 10).

الخلاصة

خلصت الدراسة إلى ضرورة الاهتمام بتطوير القطاع الصناعي المرتبط بزراعة الصمغ العربي للاستغناء عن استيراد منتجات الصمغ العربي، وتصدير متبقي خام الصمغ العربي، وهذا يؤدي إلى تطوير السوق المحلي، وزيادة فرص سوق العمل، وبالتالي تحسين الاقتصاد السوداني.

تنبأت الدراسة بتدني صادرات وواردات الصمغ العربي في السودان، ولكن بصورة أعلى في حالة الواردات، وهو مؤشر جيد في حالة الوارد، أما تدني صادر الصمغ العربي، فهو مؤشر غير جيد، لذا وجب الاهتمام بهذا القطاع.

التوصيات

1. توصي الدراسة بالاهتمام بزيادة الإنتاج الزراعي للصمغ العربي التي تؤدي بدورها لزيادة الصادرات الزراعية، وبالتالي زيادة نسبة مساهمتها في الصادرات الكلية.
2. زيادة الاهتمام بالصناعة المحلية القائمة على الصمغ العربي، وبالتالي زيادة صادر الصمغ العربي، كما يتم تصديره في شكل منتجات مصنعة وليس في شكل خام، وبالتالي ضخ العملة الصعبة في السوق السوداني مع فتح أسواق جديدة للعمل.
3. الاستفادة من البيانات الموجودة في قواعد بيانات هيئة الجمارك السودانية، وذلك لأنها المصدر الأول لحركة الاقتصاد السوداني، ومصدر موثوق لحجم التعاملات مع الدول الأخرى ولحركة التجارة الخارجية، منها نستطيع وضع سياسة اقتصادية قوية وتوقع الوضع لسنوات قادمة.
4. تطوير البيانات الموجودة بتحديث وسائل الحفظ والاسترجاع، باعتبارها ثروة وطنية مهمة في دعم الاقتصاد الوطني.
5. التدقيق في تفاصيل الإنتاج المحلي للمصانع التي تحتاج إلى استيراد الصمغ العربي (طلاء صمغي للأقمشة، مستخلص الصمغ الطبيعي، مستخلص الصمغ لاستخدامات الطباعة، بدرة صمغ نقية، رماد الصمغ، مطحون الصمغ، مواد جيلاتينية من الصمغ، مواد مكثفة من الصمغ، مواد مخاطية ومكثفات من الصمغ، الصمغ العربي المنظف) كمدخلات إنتاج، ودراسة كيفية تقليل واردات الصمغ العربي حتى تصل إلى الصفر.

المراجع

1. التقرير السنوي لهيئة الجمارك للعام 2008م، مكتبة هيئة الجمارك رئاسة الهيئة الطابق الثاني.

2. تقرير البنك الدولي، 2007م.

http://Siteresources.worldbank.org/INTAFRMDTF/Resources/Gum_Arabic_Policy_Note_AR.doc

3. الردايده، مراد خالد، عبدالرازق الشحادة، 2012م، تقنيات التنقيب عن البيانات وأهميتها في إدارة العمليات المصرفية والمحاسبية في البنوك الأردنية، عمان، جامعة الزيتونة الأردنية.

4. زهدي، خالد خواجة، 2008م، السلاسل الزمنية، بغداد، المعهد العربي للتدريب والبحوث الإحصائية.

5. الصيرفي، محمد، 2012م، الحاسوب والإحصاء الاجتماعي، الإسكندرية، دار الوفاء لدنيا الطباعة والنشر.

6. عبد الوهاب، إبراهيم طه،2005م، إستراتيجيات الأعمال في مواجهة تحديات العولمة، فيلادلفيا، جامعه فيلادلفيا.

7. عثمان، حسان، 2004م، ترشيد قرارات الحالات غير المتوقعة باستخدام تقنيات مخازن البيانات، الرياض، مدينة الملك عبد العزيز للبحوث والتقنية.

8. العلاق، بشير عباس، 2005م، الإدارة الرقمية: المجالات والتطبيقات، أبوظبي، مركز الإمارات للدراسات والبحوث الإستراتيجية.

9. قنديلجي، عامر إبراهيم، عبد الستار العلي، غسان العمري، 2006م، المدخل إلى إدارة المعرفة، عمان، دار المسيرة للنشر والتوزيع والطباعة.

10. نورين، خالد أحمد عبدالله، 2011م، قواعد البيانات، الخرطوم، جامعة العلوم والتقانة.

11. Bazsalica M., Naim P., 2001, Data mining pour le Web, Paris, éd. Eyrolles.

12. Hand D., Mannila H., Smyth R., 2001, Principles of data mining, London, MIT Press.
13. Loría. J, 2005, Building business intelligence and data mining applications with Microsoft SQL Server, USA, Wiley Publishing.
14. Tang, Z.H, and MacLennan, J., 2005, Data mining with SQL Server 2005, USA, Wiley Publishing.
15. Larose, D., 2005, Discovering knowledge in data an introduction to data mining, USA, Wiley Publishing.