



**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN**

**HERRAMIENTA DE APOYO AL ASEGURAMIENTO DE LA
CALIDAD DE DATOS EN SERIES DE TIEMPO PARA
INFORMACIÓN DE CARÁCTER FINANCIERO**

**TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN
TECNOLOGÍAS DE LA INFORMACIÓN**

CRISTIAN HORACIO MARIPANGUI GONZÁLEZ

**PROFESOR GUÍA
PABLO GONZÁLEZ JURE**

MIEMBROS DE LA COMISIÓN

José Pino Urtubia
José Benguria Donoso
Edmundo Leiva Lobos

**SANTIAGO DE CHILE
2022**

Resumen

La presente tesis tiene como objetivo la definición, diseño e implementación de una herramienta para dar aseguramiento de la calidad de datos en series de tiempo, para series de carácter financiero. La herramienta se ha anclado al procesamiento de información de datos estadísticos que realiza la División de Estadísticas Financieras de la Comisión para el Mercado Financiero (CMF) y contribuye a la transparencia y calidad de los datos de información de acceso público.

El procesamiento de datos de carácter financiero en series de tiempo que se trabaja en la División de Estadísticas Financieras, antes de ser información pública, pasa en forma regular y calendarizada, por las etapas de: Carga, Procesamiento y Cifras Preliminares, Análisis y Evaluación de la Calidad de la Información y Cifras Finales. Posteriormente, la información se divulga al público.

En la actualidad la información es generada a través de especialistas, que con su conocimiento experto desarrollan las distintas series de tiempo. Sin embargo, se requiere de un proceso estándar, transparente, independiente y objetivo. En ese contexto se crea el sistema Calidad de Datos de Series de Tiempo, integrado al Backoffice Institucional.

Este Sistema, al ser una herramienta estándar y validada por la Institución, aborda la homologación en la manera de analizar la información de las series de tiempo, mejorando los tiempos de la etapa de revisión y la calidad de la información, especialmente cuando el analista lo debe realizar para productos estadísticos no habituales. Por medio de distintos procedimientos estadísticos determina qué información es rechazada o aprobada y cuál puede ser publicada o no, además de determinar, de manera oportuna, las cifras que deben ser reprocesadas o consultadas a las instituciones que las proporcionan.

Agradecimientos

En primer lugar, agradecer al Profesor tutor Pablo González Jure que con sus conocimientos y disposición que me guiaron en este proyecto.

También quiero agradecer a la División de Estadísticas Financieras de la Comisión para el Mercado Financiero que confió en el proyecto, generó los espacio y el apoyo profesional para poder realizar su desarrollo e implementación.

Agradecer a mi querida familia que siempre me dieron palabras de apoyo para que continuara en el proyecto. Mención especial a mis padres por la infinita enseñanza.

TABLA DE CONTENIDO

1	Introducción	1
1.1	Contexto	1
1.2	Motivación	2
1.3	Oportunidad de Mejora	2
1.4	Objetivo General	3
1.5	Objetivos Específicos	3
1.6	Desarrollo y Alcance	3
1.7	Organización del Documento	4
2	Requerimientos del Sistema	5
2.1	Descripción General	5
2.1.1	Perspectiva del sistema	5
2.1.2	Tipos de Usuarios y Sus Características	5
2.1.3	Entorno operativo	6
2.1.4	Restricciones Institucionales de Diseño e Implementación	6
2.1.5	Suposiciones y Dependencias	6
2.2	Características del Sistema	7
2.2.1	Análisis de la Serie	7
2.2.2	Revisión de Análisis por Productor	8
2.2.3	Notificar a Productor por Analista	8
2.2.4	Administrar Sistema	9
2.3	Requisitos Funcionales	10
2.3.1	Listado de Casos de Uso	10
2.3.2	Casos de Uso	10
2.4	Requisitos No funcionales	25
2.4.1	Requisitos de desempeño	26
2.4.2	Requisitos de Seguridad	26
2.4.3	Requisitos de Interfaz	26
2.5	Atributos de Calidad del Software	26
3	Trazas de Eventos	28
3.1	Ingreso a la Aplicación	28
3.2	Estado de Series	28
3.3	Ingresar a serie desplegada	29

3.4 Cambiar de Estado y Agregar Comentarios a Serie	29
3.5 Ingresar a Mantenedor de Productos-Usuarios	30
3.6 Seleccionar a Usuarios de un Producto	30
3.7 Editar Asignación a Usuarios.....	31
3.8 Confirmar Cambios de Estados de Series.....	31
3.9 Enviar mensaje a usuario Productor	32
4 Medidas de Análisis de Series de Tiempo	33
4.1 Tipo de Series de Tiempo	33
4.1.1 Tendencia.....	33
4.1.2 Estacionalidad.....	34
4.1.3 Estacionalidad con Tendencia.....	34
4.1.4 Aleatoria	35
4.2 Medidas Aplicadas	35
4.2.1 Medidas de Orientación.....	36
4.2.2 Medidas de Preevaluación	37
5 Modelo de Datos del Sistema.....	38
5.1 Almacenamiento de las Series Analizadas.....	38
5.1.1 Tabla Medida de Contraste de la Serie.....	38
5.1.2 Tabla Medida de Preevaluación de la Serie	39
5.2 Almacenamiento de Proceso de Análisis y Administración del Sistema.....	40
5.2.1 Almacenamiento de Proceso de Análisis.....	41
5.2.2 Detalle Información Descriptiva Serie.....	43
5.2.3 Automatización de Definición de Tipo.....	44
6 Modelo de Arquitectura del Sistema	46
6.1 Capa de Datos	47
6.2 Capa de Negocios	48
6.3 Capa de Presentación	48
7 Vistas del Sistema de Calidad de Datos.....	49
7.1 Vista Principal Panel Evaluación de Series del Producto	49
7.2 Vista Revisión Sección del Producto.....	49
7.3 Vista Detalle Evaluación Serie – Medidas de Orientación.....	50
7.4 Vista Detalle Evaluación Serie – Medidas de Preevaluación.....	51
7.5 Vista Notificación a Productor	52

7.6 Vista Mensaje Término de Evaluación	52
8 Implementación y Evaluación del Sistema	53
8.1 Proceso de Implementación Sistema	53
8.2 Evaluación del Sistema.....	53
8.2.1 Evaluación de los Objetivos y Registros del Sistema.....	54
8.2.2 Evaluación de Usabilidad por parte de Usuarios	56
9 Conclusiones y Trabajo Futuro	57
9.1 Conclusiones.....	57
9.2 Trabajo Futuro	58
Bibliografía.....	59
ANEXOS	60
Anexo A. Procedimiento de Analista en el Sistema de Aseguramiento de Calidad	60
Anexo B. Manual Herramienta Calidad de Información.....	61
Anexo C. Procedimiento para Tratar Series con Tendencia y Estacionalidad	68
Anexo D. Pasos para Tratamiento de Series Tendencia Pura y Aleatorias.....	70

Índice de Tablas

Tabla 1 Actores y Casos de Uso del Sistema	10
Tabla 2 CU Ingreso Aplicación.....	12
Tabla 3 CU Ingreso Estado Series.....	13
Tabla 4 CU Despliega Listado Estado Series	15
Tabla 5 CU Cambio Estado	17
Tabla 6 CU Agregar Contenido.....	19
Tabla 7 CU Ingreso Mantenedor Series Usuario.....	20
Tabla 8 CU Selección Usuario	22
Tabla 9 CU Editar Asignación	24
Tabla 10 CU Confirmación Cambios.....	25
Tabla 11 Medidas Aplicadas por Tipo Serie.....	37
Tabla 12 Medida Contraste.....	39
Tabla 13 Medida Preevaluación.....	39
Tabla 14 Respuestas y Puntajes Encuesta Usabilidad	56

Índice de Figuras

Figura 1 Ingreso Aplicación.....	28
Figura 2 Estado Series	29
Figura 3 Serie Desplegada	29
Figura 4 Cambio Estado y Comentario	30
Figura 5 Ingreso Mantenedor Producto Usuario.....	30
Figura 6 Selección Productos Usuario	31
Figura 7 Edición Usuario.....	31
Figura 8 Confirmación Cambios.....	32
Figura 9 Notificación a Productor	32
Figura 10 Tipo de Serie Tendencia Creciente.....	33
Figura 11 Tipo de Serie Tendencia Decreciente	34
Figura 12 Tipo de Serie con Estacionalidad.....	34
Figura 13 Tipo de Serie con Tendencia y Estacionalidad.....	35
Figura 14 Tipo de Serie Aleatoria.....	35
Figura 15 Gráfica Medida Orientación – Tendencia	36
Figura 16 Modelo de Datos.....	40
Figura 17 Modelo de Arquitectura	46
Figura 18 Vista Principal	49
Figura 19 Revisión Sección	50
Figura 20 Detalle Evaluación Orientación	51
Figura 21 Detalle Evaluación Preevaluación.....	51
Figura 22 Notificación Productor.....	52
Figura 23 Término Evaluación	52
Figura 24 Total de Series Evaluadas - Aprobadas vs Rechazadas.....	54
Figura 25 Comparación Series Aprobadas vs Rechazadas por Periodo	55
Figura 26 Pasos Usuario - Uso de Sistema	60
Figura 27 Ingreso a Backoffice	61
Figura 28 Vista Principal Backoffice.....	61
Figura 29 Vista Productos.....	62
Figura 30 Vista Filtro y Selección de Productos.....	62
Figura 31 Vista Producto y sus Etapas	63
Figura 32 Vista Secciones Producto	63
Figura 33 Vista Series Sección	64
Figura 34 Vista Evaluación Serie Orientación	65
Figura 35 Vista Evaluación Serie Preevaluación.....	65
Figura 36 Vista Notificación a Productor	66
Figura 37 Vista Correo desde Sistema	66
Figura 38 Término de Evaluación	67

1 Introducción

1.1 Contexto

En el último tiempo se ha dado una gran importancia a los datos tanto para estudios, como para el análisis del comportamiento de los distintos hechos financieros que van ocurriendo en el tiempo en el país. Una de las preocupaciones y brechas mayores en el trabajo con datos es tener datos que cumplan con un grado de calidad confiable. Las series de tiempo de información financiera no quedan exentas de este enfoque, por lo mismo, dentro de los procesos para obtener y manejar la información es muy importante aplicar técnicas de calidad de datos.

En la Comisión para el Mercado Financiero (CMF) existen distintos niveles de calidad impuesta a las instituciones que entregan información, estos mecanismos se manejan como validaciones a la información proporcionada por las instituciones. Además, la División de Estadísticas Financieras aplica análisis y validaciones a la información agrupada y caracterizada en series de tiempo, esto implica trabajar con la información estadística proporcionada en el tiempo, y para realizarlo es necesario, como mínimo, tres años de información.

Cuando se observa una serie de tiempo en el cual tiene uno o más datos no acordes a su comportamiento son posibles outliers, identificar esos datos no acordes o anómalos en la serie de tiempo se refiere a la detección de outliers. La detección de outliers es un problema que matemáticamente se ha resuelto, para esto existen librerías y herramientas para generar una solución.

Para el caso de la detección de outliers en series de tiempo asociadas a información financiera, los programas más reconocidos y utilizados son TRAMO-SEATS [MTE_01] y X12-ARIMA (nueva versión renombrada a X-13-ARIMA-SEATS [MTE_02]). Ambos trabajan con algoritmos predictivos, como el Modelo autorregresivo integrado de promedio móvil (ARIMA), y utilizan técnicas de ajuste para suavizar la curva de la serie de tiempo. Estos programas son recomendados por entidades de reconocido prestigio en el sector financiero y estadístico, tales como el Banco de España y el Centro de Estadísticas Europeo (Eurostat).

Se puede observar que todos estos casos utilizan el modelo ARIMA ¹ para poder trabajar sobre los datos anómalos en las series de tiempo, tanto para su detección como para su ajuste, este modelo estadístico utiliza variaciones y regresiones de datos estadísticos con el fin de encontrar patrones para una predicción hacia el futuro.

1.2 Motivación

La naturaleza de las series financieras depende mucho de patrones o variables exógenas, por ejemplo, recesiones en el mercado, estaciones del año, pandemias, entre otros. Estos hitos pueden realizar cambios en los datos entregados generando distintos comportamientos de series en un determinado tramo de tiempo. El análisis de la información de una serie tomando como opción la predicción en base a información histórica se complica. Por lo mismo es de mayor importancia al final del proceso el criterio experto de un analista, o sea determinar que la serie con los nuevos datos es correcta o no. Por lo mismo, en el Sistema, se debe considerar el análisis experto y conservar las observaciones realizadas en el periodo, especialmente cuando se ha encontrado el dato fuera de rango normal en la serie, de esta manera en una etapa posterior tener historia de las evaluaciones.

1.3 Oportunidad de Mejora

Dentro de los desafíos de la División de Estadísticas Financieras de la CMF es poder homologar los procesos de calidad de información, para así, enfocar los esfuerzos y la experiencia de los analistas en mejorar las medidas de calidad, agregar más productos estadísticos y que estos queden a disposición pública (bancos, medios de información, consultoras, instituciones educacionales, entre otros). Además conservar la información histórica de análisis de calidad y que sea insumo para trabajo futuros.

¹ ARIMA (Autoregressive Integrated Moving Average) corresponde al modelo autoregresivo integrado de promedio móvil, el cual se utiliza para la predicción de valores futuros en series de tiempo.

Este sistema se agrega al abanico de innovaciones tecnológicas propuestas y desarrolladas por la División de Estadísticas Financieras de la CMF, aportando al mejoramiento continuo de las labores en materia de producción estadística.

1.4 Objetivo General

Se creará un sistema de apoyo al análisis denominado Calidad de Datos en Series de Tiempo, este se utilizará por los analistas que trabajen con datos estadísticos en series de tiempo de la División de Estadísticas Financieras. El sistema estandarizará el proceso de calidad de datos de las series abarcadas, almacenará la información entregada por las medidas utilizadas para la calidad de la serie y además la información proporcionada por el Analista.

1.5 Objetivos Específicos

- a) Incorporar los productos estadísticos de los analistas para su revisión y posterior uso de la herramienta.
- b) Incorporar medidas relevantes solicitadas por los analistas a las series específicas.
- c) Entregar la información descriptiva y gráfica de las series de tiempo incorporadas.
- d) Evaluación de impacto del uso de la herramienta y en los datos evaluados.

1.6 Desarrollo y Alcance

Este sistema se utiliza en los equipos de la Institución con sistema operativo Windows 10 y en la red interna, se debe incorporar al sistema Backoffice. El sistema solo se utilizará como guía para el analista, siempre el analista tendrá la responsabilidad de aprobar o rechazar la serie de datos. En la etapa de implementación se migrarán un subgrupo de productos estadísticos de la totalidad de la carga por analista.

1.7 Organización del Documento

En el capítulo uno se da una introducción y contexto al trabajo a realizar, además se definen los objetivos y alcances. En el capítulo dos se explica cómo se abordará la solución tomando en cuenta los requerimientos asociados. En el capítulo tres se muestran las trazas de los eventos principales generados por la aplicación. En el capítulo cuatro se definen las métricas utilizadas para las series de tiempo y como estas se relacionan dependiendo el comportamiento de la serie en el periodo de tiempo evaluado. En el capítulo cinco abordará el modelo de datos que contiene la solución. En el capítulo seis se detallará la arquitectura de la herramienta de calidad de datos. En el capítulo siete se mostrarán las vistas asociadas a la herramienta de calidad. En el capítulo ocho se hablará sobre la implementación y evaluación de impacto de la herramienta. En el capítulo nueve se darán las conclusiones y se comentan algunos trabajos futuros que podría tener la siguiente versión de esta herramienta.

2 Requerimientos del Sistema

2.1 Descripción General

Para la obtención de los requerimientos del sistema se generaron reuniones con los analistas y el jefe de la División Estadística. En este capítulo se detallan los requerimientos obtenidos, abordando la perspectiva del sistema, los tipos de usuario, el entorno donde se operativiza el sistema, restricciones, suposiciones y dependencias.

2.1.1 Perspectiva del sistema

Se crea un sistema para los usuarios que trabajen con datos estadísticos en Series de Tiempo de la División de Estadísticas Financieras de la CMF. El sistema mostrará la serie de datos y los resultados de las distintas métricas utilizadas, con esto el usuario podrá evaluar si es necesario realizar un análisis más profundo a los datos y si se aprueba o rechaza la información para publicación. La información entregada tanto por las medidas como por el analista se almacenarán. Esta aplicación se utiliza en los equipos de la Institución y en la red interna.

2.1.2 Tipos de Usuarios y Sus Características

Los tipos de usuarios son Administrador, Productor y Analista:

Administrador	Podrá utilizar además de los módulos para ver las métricas, los módulos de administración de Series a Usuarios. Los historiales de Estados de las series evaluadas
Productor	Será el usuario interesado en realizar la carga de datos para el análisis sobre las series. Además, está encargado en revisar el estado de la revisión de las series por el Analista.
Analista	Será el usuario interesado en realizar los análisis sobre las series asignadas en el módulo correspondiente, además establecer el estado de la Serie.

2.1.3 Entorno operativo

Se detallan a continuación los requerimientos respecto al entorno operativo del sistema (EO):

EO-1.- La aplicación debe estar disponible para laptop con Sistema Operativo Windows.

EO-2.- La aplicación debe estar disponible en los navegadores Mozilla, Chrome, IE, Opera y Safari.

2.1.4 Restricciones Institucionales de Diseño e Implementación

Las restricciones de diseño e implementación del sistema (RDI) son:

RDI-1.- El sistema debe ser desarrollado en lenguajes C#, AngularJs, Python (y/o R).

RDI-2.- La Base de datos del Sistema debe ser implementada en SQL Server.

RDI-3.- Se debe poder observar de manera gráfica la información de la Serie y sus métricas.

RDI-4.- Al terminar la aplicación se debe entregar un manual de usuario.

2.1.5 Suposiciones y Dependencias

Se observan las siguientes suposiciones (SU) o dependencias (DE) del sistema:

SU-1: Al comienzo la aplicación presentará poca utilización por los usuarios, ya que priorizarán el análisis de las herramientas creadas por ellos, hasta que se adapten.

SU -2: Con el tiempo disminuirán los casos outliers detectados por la aplicación, ya que, si fueron justificados y/o advertidos a las instituciones que entregan la información, estos no deberían aparecer en periodos posteriores

DE-1: El usuario debe utilizar el equipo de la institución.

DE-2: El usuario debe estar conectado a la red de la institución.

DE-3: El usuario debe estar registrado en la aplicación.

2.2 Características del Sistema

Las principales características obtenidas de los requerimientos del sistema son el análisis de la serie, revisión del análisis por el Productor, notificación hacia el Productor por el Analista y administración del sistema.

2.2.1 Análisis de la Serie

Corresponde al proceso donde un Analista puede determinar si una serie es aprobada o rechazada.

2.2.1.1 Descripción y prioridad

Al ingresar al módulo de Análisis de Calidad de Series se debe desplegar un listado con las series correspondientes el perfil del usuario. Al seleccionar una Serie se debe desplegar un gráfico que contiene las métricas evaluadas con sus resultados.

2.2.1.2 Secuencia de Estímulo – Respuesta

Estímulo: Usuario ingresa al sistema de Calidad de Series.

Respuesta: El sistema muestra las Series correspondientes a su carga, más los respectivos botones para aprobar o rechazar e ir al detalle de la Serie.

Estímulo: El usuario selecciona la serie a observar.

Respuesta: El sistema muestra las gráficas de las distintas métricas obtenidas para la serie, además los botones para poder asignar un estado de la serie (aprobado, rechazado) y agregar observaciones.

Estímulo: El usuario selecciona una de las opciones de aprobar o rechazar del módulo.

Respuesta: El sistema actualiza la información de la serie en el listado total.

Estímulo: El usuario cierra el detalle de la serie.

Respuesta: El sistema vuelve al listado de Series con el cambio de estado en pantalla.

Estímulo: El usuario presiona el botón confirmar revisión.

Respuesta: El sistema envía el cambio de estado de las series al usuario productor.

2.2.2 Revisión de Análisis por Productor

Este proceso corresponde a la revisión de la serie, donde se debe desplegar la información complementaria y preevaluación generada por el sistema.

2.2.2.1 Descripción y prioridad

El Productor al ingresar al módulo de Análisis de Calidad de Series se debe desplegar un listado con las series correspondientes a su carga. Al seleccionar una Serie se debe desplazar un gráfico que contiene las métricas evaluadas con sus resultados.

2.2.2.2 Secuencia de Estímulo – Respuesta

Estímulo: Productor ingresa al módulo de Calidad de Series.

Respuesta: El sistema muestra las Series correspondientes a su carga, más los resultados de la revisión del Analista.

Estímulo: El usuario selecciona la serie a observar.

Respuesta: El sistema muestra las gráficas de las distintas métricas obtenidas para la Serie, además el estado de la serie (aprobado, rechazado, observaciones).

Estímulo: El usuario cierra el detalle de la serie.

Respuesta: El sistema vuelve al listado de Series.

2.2.3 Notificar a Productor por Analista

Esta característica se genera para que el analista tenga la opción de escribir al Productor por medio del sistema, dejando un registro de este evento y dando formalidad al proceso.

2.2.3.1 Descripción y prioridad

El Analista al ingresar al módulo de Análisis de Calidad de Series debe presionar el botón “Notificar a Productor”, con esto se despliega panel donde el analista ingresa el asunto y el mensaje a enviar, más el botón de enviar o cancelar la operación.

2.2.3.2 Secuencia de Estímulo – Respuesta

Estímulo: El usuario ingresa al módulo de Calidad de Series

Respuesta: El sistema muestra el módulo con las distintas funcionalidades

Estímulo: El usuario selecciona la opción de notificar a productor.

Respuesta: El sistema muestra panel donde el analista puede ingresar la información a enviar.

Estímulo: El usuario ingresa los datos requeridos, luego presiona el botón para enviar

Respuesta: El sistema realiza y confirma el envío del correo al usuario productor respectivo.

2.2.4 Administrar Sistema

Para la administración del sistema se orienta a administración de métricas por tipo de serie, roles para usuarios y series a revisar por usuario.

2.2.4.1 Descripción y prioridad

El usuario administrador al ingresar al módulo administración de Calidad de Series debe ver las opciones de editar las medidas, editar las tipologías y los usuarios por productos.

2.2.4.2 Secuencia de Estímulo – Respuesta

Estímulo: El usuario ingresa al módulo administración de Calidad de Series.

Respuesta: El sistema muestra el módulo de administración con las distintas funcionalidades.

Estímulo: El usuario modifica las métricas de un tipo de series.

Respuesta: El sistema modifica las métricas para el tipo de series y muestra esa modificación en el panel.

Estímulo: El usuario modifica los permisos y roles de usuarios de un producto.

Respuesta: El sistema modifica los usuarios para el tipo de series y muestra esa modificación en el panel.

2.3 Requisitos Funcionales

Para la descripción de los requerimientos funcionales se han utilizado casos de uso, de esta manera se identifican los distintos pasos para la obtención del resultado deseado por los usuarios del sistema.

2.3.1 Listado de Casos de Uso

Los principales casos de uso del sistema se listan en la tabla Actores y Casos de Uso del Sistema.

Actores Principales	Casos de Uso
Usuario Productor Usuario Analista Usuario Administrador	1.- Ingresar a la Aplicación 2.- Ingresar a “Estado de Series” 3.- Ingresar a una Serie de las desplegadas en el listado del módulo “Estado de Series” 4.- Cambiar de estado pendiente a aprobada o rechazada. 5.- Agregar comentarios a la revisión de la Serie. 6.- Ingresar a “Mantenedor de Productos-Usuarios” 7.- Seleccionar a usuario de listado 8.- Editar asignación a Usuarios 9.- Confirmar cambios realizados al grupo de Series.

Tabla 1 Actores y Casos de Uso del Sistema

2.3.2 Casos de Uso

Se detallan los distintos casos de uso del listado a continuación:

ID Caso de Uso:	1		
Nombre Caso de Uso:	Ingresar a la Aplicación		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Productor Usuario Analista Usuario Administrador		
Descripción:	El usuario debe ingresar al enlace de la aplicación. Al realizar esto el sitio enviará al usuario a la aplicación Security, en esa aplicación el usuario debe ingresar su nombre y clave de usuario interno.		
Como desencadena:	Usuario se entera que existe la aplicación, o le piden que la use, y que puede utilizarla en su laptop.		
Precondiciones:	1.- El usuario debe utilizar el equipo de la institución. 2.- El usuario debe estar conectado a la red de la institución. 3.- El usuario debe estar registrado en la aplicación Security.		
Postcondiciones:	1. Disponibilidad de la aplicación Security y el registro del usuario en esta.		
Flujo Normal:	1.- El usuario ingresa al enlace de la aplicación de calidad de datos. 2.- La aplicación lo redirige a Security. 3.- El usuario ingresa su nombre y clave interno (intranet). 4.- Security lo redirige a la aplicación de calidad de datos.		
Flujo Alternativo:	-		
Excepciones:	-		

Incluir:	Ninguno
Prioridad:	Alta
Frecuencia de Uso:	El 100% de los usuarios que deben usar la aplicación utilizarán este caso
Reglas de Negocio:	El usuario logra ingresar a la aplicación para utilizar los módulos que tiene permisos desde su laptop.
Requerimientos Especiales:	-
Suposiciones:	Todo usuario que deba usar esta aplicación sabe utilizar un computador nivel intermedio.
Notas y Casos:	Determinar si las instrucciones dadas son lo suficientemente claras para los usuarios.

Tabla 2 CU Ingreso Aplicación

ID Caso de Uso:	2		
Nombre Caso de Uso:	Ingresar a "Estado de Series"		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Productor Usuario Analista		
Descripción:	El usuario debe ingresar al módulo "Estado de Series", este mostrará un listado de las series asignadas a su carga y el estado de estas (pendiente, aprobada, rechazada).		
Como desencadena:	Usuario ingresa al módulo de la barra de menú.		

Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1.
Postcondiciones:	Disponibilidad de las series asignadas al usuario.
Flujo Normal:	1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de Series Asignadas.
Flujo Alternativo:	1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de Series Asignadas.
Excepciones:	1.- Si el usuario no tiene series asignadas no se desprenderá ningún listado.
Incluir:	Ninguno
Prioridad:	Alta
Frecuencia de Uso:	Todos los usuarios Analistas y Productor deben usar el módulo.
Reglas de Negocio:	El usuario ve el listado de Series asignadas y el estado de estas.
Requerimientos Especiales:	-
Suposiciones:	Todo usuario que deba usar este módulo entiende las series asignadas.
Notas y Casos:	Determinar si las instrucciones dadas son lo suficientemente claras para los usuarios.

Tabla 3 CU Ingreso Estado Series

ID Caso de Uso:	3		
Nombre Caso de Uso:	Ingresar a una Serie de las desplegadas en el listado del módulo "Estado de Series"		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Productor Usuario Analista		
Descripción:	El usuario debe seleccionar una serie del listado de las series asignadas en el módulo "Estado de Series", al realizar esto se mostrará el detalle de la Serie: Resultado de las Métricas, observaciones de la Serie, estado de la Serie, posibles cambios a realizar.		
Como desencadena:	Usuario ingresa a la Serie desde el listado de módulo.		
Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1.		
Postcondiciones:	Disponibilidad del detalle de la serie seleccionada.		
Flujo Normal:	1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario selecciona el detalle de una serie del listado.		
Flujo Alternativo:	1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario selecciona el detalle de una serie del listado.		
Excepciones:	1.- Si el usuario no tiene series asignadas no se desprenderá ningún listado.		

Incluir:	Ninguno
Prioridad:	Alta
Frecuencia de Uso:	Todos los usuarios Analistas y Productor deben usar el módulo.
Reglas de Negocio:	El usuario ve el detalle de la Serie seleccionada.
Requerimientos Especiales:	Los resultados de las métricas se deben mostrar con gráficos y deben ser descargables en un archivo (Excel y/o CSV)
Suposiciones:	Todo usuario que deba usar este módulo debe saber interpretar las métricas utilizadas para la Serie e identificar si los outliers afectan para la serie.
Notas y Casos:	Determinar si las instrucciones, métricas y gráficos dados son lo suficientemente claros para los usuarios.

Tabla 4 CU Despliega Listado Estado Series

ID Caso de Uso:	4		
Nombre Caso de Uso:	Cambiar de estado pendiente a aprobada o rechazada.		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Productor Usuario Analista		
Descripción:	El usuario debe seleccionar una serie del listado de las series asignadas en el módulo "Estado de Series", al realizar esto se		

	mostrará el detalle de la Serie: Resultado de las Métricas, observaciones de la Serie, estado de la Serie, posibles cambios a realizar. Dentro de los cambios el usuario cambia de estado pendiente a rechazada o aprobada.
Como desencadena:	Usuario ingresa a la Serie desde el listado de módulo.
Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1.
Postcondiciones:	Disponibilidad del detalle de la serie seleccionada.
Flujo Normal:	<ol style="list-style-type: none"> 1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario selecciona el detalle de una serie del listado. 4.- El usuario cambia de estado pendiente a rechazada o aprobada.
Flujo Alternativo:	<ol style="list-style-type: none"> 1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario selecciona el detalle de una serie del listado. 4.- El usuario cambia de estado pendiente a rechazada o aprobada.
Excepciones:	1.- Si el usuario no tiene series asignadas no se desprenderá ningún listado.
Incluir:	Ninguno
Prioridad:	Alta
Frecuencia de Uso:	Todos los usuarios Analistas y Productor deben usar el módulo.

Reglas de Negocio:	El usuario ve el detalle de la Serie seleccionada.
Requerimientos Especiales:	Los resultados de las métricas se deben mostrar con gráficos.
Suposiciones:	Todo usuario que deba usar este módulo debe saber interpretar las métricas utilizadas para la Serie e identificar si los outliers afectan para la serie.
Notas y Casos:	Determinar si las instrucciones, métricas y gráficos dados son lo suficientemente claros para los usuarios.

Tabla 5 CU Cambio Estado

ID Caso de Uso:	5		
Nombre Caso de Uso:	Agregar comentarios a la revisión de la Serie.		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Productor Usuario Analista		
Descripción:	El usuario debe seleccionar una serie del listado de las series asignadas en el módulo "Estado de Series", al realizar esto se mostrará el detalle de la Serie: Resultado de las Métricas, observaciones de la Serie, estado de la Serie, posibles cambios a realizar. Dentro de los cambios el usuario agrega un comentario y luego lo guarda.		

Como desencadena:	Usuario ingresa a la Serie desde el listado de módulo.
Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1.
Postcondiciones:	Disponibilidad del detalle de la serie seleccionada.
Flujo Normal:	<ol style="list-style-type: none"> 1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario selecciona el detalle de una serie del listado. 4.- El usuario agrega un comentario y luego lo guarda.
Flujo Alternativo:	<ol style="list-style-type: none"> 1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario selecciona el detalle de una serie del listado. 4.- El usuario agrega un comentario y luego lo guarda.
Excepciones:	1.- Si el usuario no tiene series asignadas no se desprenderá ningún listado.
Incluir:	Ninguno
Prioridad:	Alta
Frecuencia de Uso:	Todos los usuarios Analistas y Productor deben usar el módulo.
Reglas de Negocio:	El usuario ve el detalle de la Serie seleccionada.
Requerimientos Especiales:	Los resultados de las métricas se deben mostrar con gráficos.

Suposiciones:	Todo usuario que deba usar este módulo debe saber interpretar las métricas utilizadas para la Serie e identificar si los outliers afectan para la serie.
Notas y Casos:	Determinar si las instrucciones, métricas y gráficos dados son lo suficientemente claros para los usuarios.

Tabla 6 CU Agregar Contenido

ID Caso de Uso:	6		
Nombre Caso de Uso:	Ingresar a “Mantenedor de Series-Usuarios”		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Administrador		
Descripción:	El usuario debe ingresar al módulo “Mantenedor de Series-Usuarios”, este mostrara un listado de los usuarios y el o los perfiles asignados.		
Como desencadena:	Usuario ingresa al módulo de la barra de menú.		
Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1.		
Postcondiciones:	Disponibilidad de los usuarios con permiso de ingresar al módulo.		
Flujo Normal:	1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de usuarios de la aplicación.		

Flujo Alternativo:	1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de usuarios de la aplicación
Excepciones:	-
Incluir:	Ninguno
Prioridad:	Alta
Frecuencia de Uso:	La principal función el administrador se debe realizar en este módulo.
Reglas de Negocio:	El usuario ve el listado de usuarios de la aplicación y las Series asignadas de estos.
Requerimientos Especiales:	-
Suposiciones:	-
Notas y Casos:	Determinar si las instrucciones dadas son lo suficientemente claras para los usuarios.

Tabla 7 CU Ingreso Mantenedor Series Usuario

ID Caso de Uso:	7		
Nombre Caso de Uso:	Seleccionar a usuario de listado		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Administrador		

Descripción:	El usuario debe ingresar al módulo “Mantenedor de Series-Usuarios”, este mostrara un listado de los usuarios y el o los perfiles asignados. El usuario Administrador selecciona un usuario del listado, al realizar esto se ve el detalle de las series asignadas al usuario seleccionado, si es usuario Analista muestra el listado de los usuarios productores asignados.
Como desencadena:	Usuario ingresa al módulo de la barra de menú y selecciona un analista o productor del listado desprendido.
Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1. 2.- El usuario debe haber ingresado al módulo “Mantenedor de Series-Usuarios”
Postcondiciones:	Disponibilidad de los usuarios con permiso de ingresar al módulo.
Flujo Normal:	1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de usuarios de la aplicación. 3.- Selecciona un analista o productor del listado.
Flujo Alternativo:	1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de usuarios de la aplicación 3.- Selecciona un analista o productor del listado.
Excepciones:	-
Incluir:	Ninguno
Prioridad:	Alta

Frecuencia de Uso:	La principal función el administrador se debe realizar en este módulo.
Reglas de Negocio:	El usuario ve el listado de usuarios de la aplicación y las Series asignadas de estos.
Requerimientos Especiales:	-
Suposiciones:	-
Notas y Casos:	Determinar si las instrucciones dadas son lo suficientemente claras para los usuarios.

Tabla 8 CU Selección Usuario

ID Caso de Uso:	8		
Nombre Caso de Uso:	Editar asignación a Usuarios		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Administrador		
Descripción:	<p>El usuario debe ingresar al módulo “Mantenedor de Series-Usuarios”, este mostrara un listado de los usuarios y el o los perfiles asignados. El usuario Administrador selecciona un usuario del listado, al realizar esto se ve el detalle de las series asignadas al usuario seleccionado, si es usuario Analista muestra el listado de los usuarios productores asignados. El usuario Administrador puede modificar tanto el listado de series asignadas al analista como el perfil de este, luego</p>		

	guarda esta información. Para el caso que el usuario seleccionado tenga productores asignados, el administrador puede modificar este listado.
Como desencadena:	Usuario ingresa al módulo de la barra de menú y selecciona un analista o productor del listado desprendido.
Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1. 2.- El usuario debe haber ingresado al módulo "Mantenedor de Series-Usuarios"
Postcondiciones:	Disponibilidad de los usuarios con permiso de ingresar al módulo.
Flujo Normal:	1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de usuarios de la aplicación. 3.- Selecciona un analista o productor del listado. 4.- Modifica los datos del analista o productor, series asignadas, y/o perfil. Para el caso que el usuario seleccionado tenga productores asignados, el administrador puede modificar este listado.
Flujo Alternativo:	1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de usuarios de la aplicación 3.- Selecciona un analista o productor del listado. 4.- Modifica los datos del analista o productor, series asignadas, y/o perfil. Para el caso que el usuario seleccionado tenga productores asignados, el administrador puede modificar este listado.

Excepciones:	-
Incluir:	Ninguno
Prioridad:	Media
Frecuencia de Uso:	Una de las funciones del administrador se debe realizar en este módulo.
Reglas de Negocio:	El usuario ve el listado de usuarios de la aplicación, las Series asignadas de estos y los productores asignados.
Requerimientos Especiales:	-
Suposiciones:	-
Notas y Casos:	Determinar si las instrucciones dadas son lo suficientemente claras para los usuarios.

Tabla 9 CU Editar Asignación

ID Caso de Uso:	9		
Nombre Caso de Uso:	Confirmar cambios realizados al grupo de Series		
Creado por:	Cristian Maripangui	Última actualización por:	Cristian Maripangui
Fecha Creación (aaaa-mm-dd):	2020-07-06	Fecha última actualización:	2020-07-06
Actores:	Usuario Productor Usuario Analista		
Descripción:	El usuario debe seleccionar el módulo "Estado de Series", al realizar esto se mostrará el listado de las series con sus respectivos estados, el usuario debe presionar confirmar		

	estados. Al confirmar estados si es un usuario productor se le notificará al usuario Analista de su confirmación.
Como desencadena:	Usuario ingresa al módulo “Estado de Series” y confirma los estados de las series.
Precondiciones:	1.- El usuario debe haber pasado por el Caso de Uso 1.
Postcondiciones:	Disponibilidad del listado de series asignadas con sus estados.
Flujo Normal:	<ol style="list-style-type: none"> 1.- El usuario ya ingresado en la aplicación ingresa al módulo que está en su menú. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario confirma los estados de las series. 4.- La aplicación notifica a los usuarios analistas correspondientes de la confirmación.
Flujo Alternativo:	<ol style="list-style-type: none"> 1.- El usuario ya ingresado en la aplicación ingresa por URL al módulo. 2.- La aplicación desprende el listado de Series Asignadas. 3.- El usuario confirma los estados de las series. 4.- La aplicación notifica a los usuarios analistas correspondientes de la confirmación.
Excepciones:	1.- Si el usuario no tiene series asignadas no se desprenderá ningún listado.
Incluir:	Ninguno

Tabla 10 CU Confirmación Cambios

2.4 Requisitos No funcionales

De la obtención de requerimientos se han definido distintos requisitos no funcionales para el óptimo funcionamiento del sistema. Estos se pueden separar en Requisitos de desempeño, seguridad e Interfaz.

2.4.1 Requisitos de desempeño

Para los requisitos de desempeño se deben considerar:

- El sistema no debe presentar latencia en los periodos de análisis de datos.
- El sistema debe tener una escalabilidad de almacenamiento acorde a la alta demanda de series para analizar.

2.4.2 Requisitos de Seguridad

Debido a la utilización de información que se debe resguardar, además que hay procesos asociados a un usuario, es importante considerar los requisitos de seguridad.

La privacidad de la información generada por el sistema deberá ser resguardada y protegida para cualquier uso indebido o no consentido.

2.4.3 Requisitos de Interfaz

Debido a que este sistema se debe incorporar al sistema centralizado del proceso de producción de estadísticas entonces se debe utilizar los temas y colores del sistema Backoffice en los componentes de la interfaz para facilitar su uso para los usuarios.

2.5 Atributos de Calidad del Software

Se definen los siguientes atributos de calidad del software:

Interoperabilidad-1 El sistema debe ser compatible con versiones específicas de navegadores y SO con el fin de llegar a la mayor cantidad de usuarios.

Usabilidad-1. El sistema debe tener un diseño que sea comprensible por los usuarios, de esta manera que apoye a la rápida implementación y disminuya las resistencias de uso de los usuarios dentro del proceso de análisis y posterior operación.

Mantenibilidad-1 Se proyecta que la aplicación perdure en el tiempo teniendo sucesivos avances en términos de mejor usabilidad. Para ello,

su desarrollo y pruebas de usuario deben ser de manera continua para detectar posibles mejoras.

Mantenibilidad-2 Se planifica aumentar la batería de indicadores de posibles outliers.

Mantenibilidad-3 Se planifica aumentar la cantidad de series a evaluar en el sistema.

3 Trazas de Eventos

En el proceso de análisis de requerimientos y diseño de la aplicación se han utilizado las trazas de eventos, para observar los distintos participantes de cada requerimiento, las clases y los procesos.

3.1 Ingreso a la Aplicación

Para el ingreso a la aplicación cada usuario ingresa su información de autenticación, este se comunica con sistema Security, el cual mantiene las cuentas de usuario y su información. El sistema Security envía un token de verificación a la aplicación del sistema de calidad con esto se confirma el inicio de sesión y luego el usuario ingresar. Ver “Figura 1 Ingreso Aplicación”.

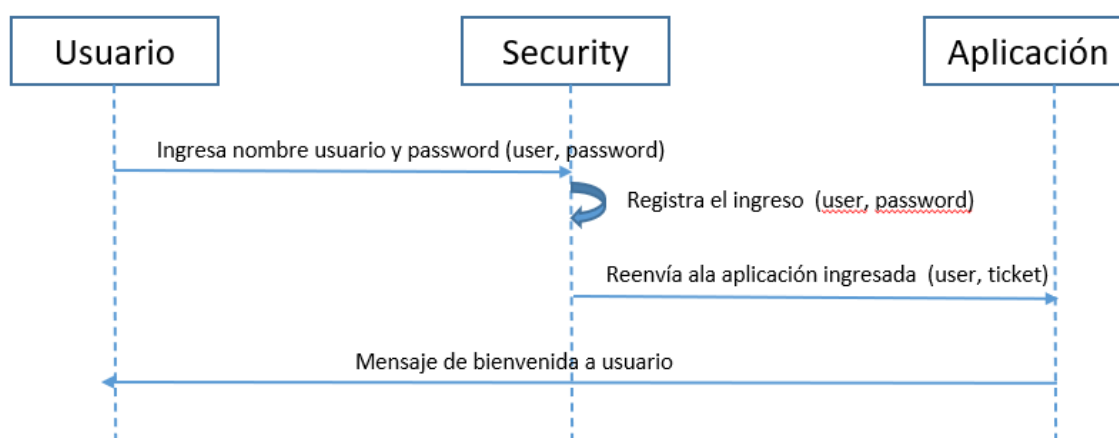


Figura 1 Ingreso Aplicación

3.2 Estado de Series

En la obtención de Estado de series se describe cuando el usuario ingresa al sistema de calidad y luego aparece en su pantalla el listado de las series del producto estadístico. Para esto el usuario desde el menú seleccione a la opción de ingresar al sistema de calidad de información, luego si los permisos corresponden el usuario puede ver el listado de estado de series. Ver “Figura 2 Estado Series”

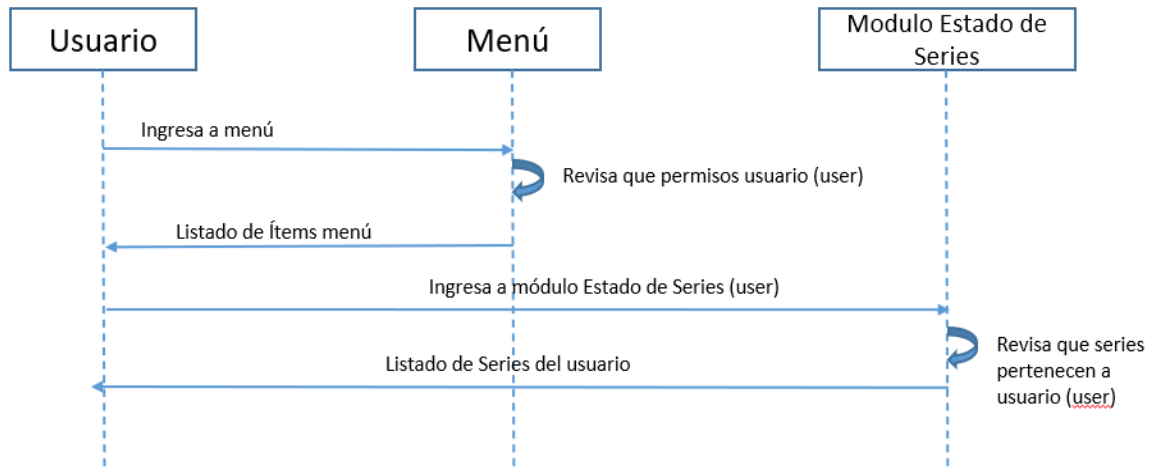


Figura 2 Estado Series

3.3 Ingresar a serie desplegada

Después que el usuario ingresa al listado de estado de series puede seleccionar un caso para poder observar su detalle. Ver “Figura 3 Serie Desplegada”.

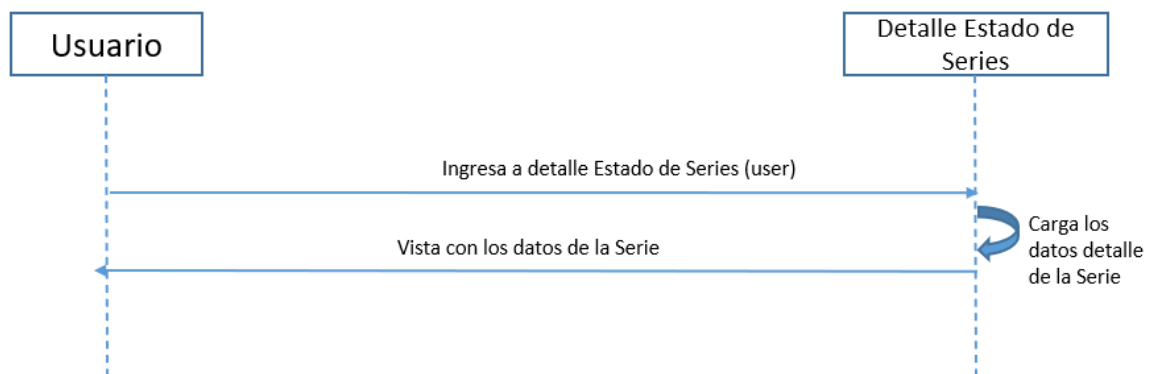


Figura 3 Serie Desplegada

3.4 Cambiar de Estado y Agregar Comentarios a Serie

Al observar la serie el usuario tiene la opción e ingresar un nuevo estado para esa serie, además se agrega una observación, el módulo confirma la edición. Ver “Figura 4 Cambio Estado y Comentario”.

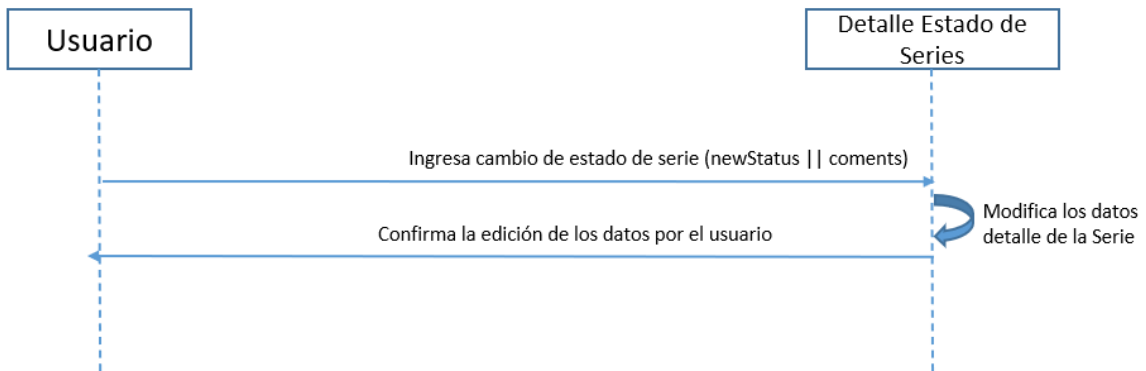


Figura 4 Cambio Estado y Comentario

3.5 Ingresar a Mantenedor de Productos-Usuarios

El usuario administrador al ingresar al menú del sistema, en ese menú aparece la opción de mantenedor de productos usuarios, al ingresar se observa el listado del producto estadístico y los usuarios por producto. Ver “Figura 5 Ingreso Mantenedor Producto Usuario”.

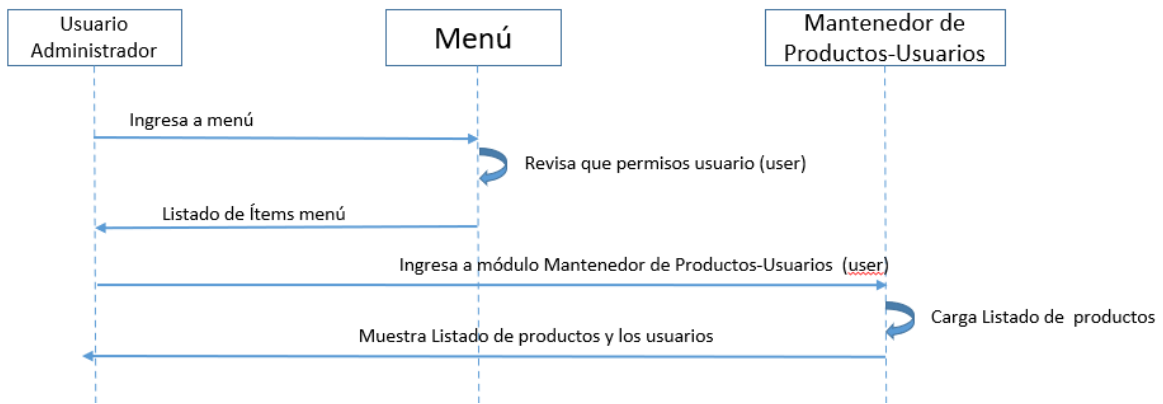


Figura 5 Ingreso Mantenedor Producto Usuario

3.6 Seleccionar a Usuarios de un Producto

Al seleccionar el producto con sus usuarios respectivos, se observa el listado de usuarios con sus respectivos perfiles para el producto. Ver “Figura 6 Selección Productos Usuario”

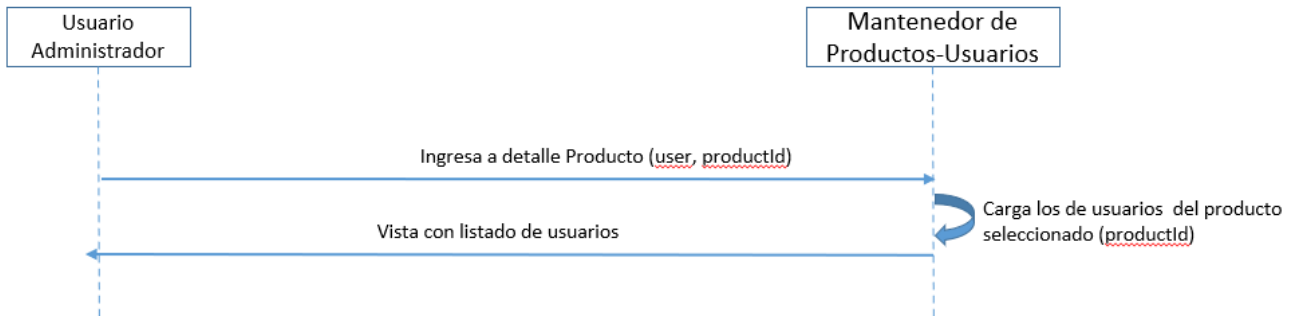


Figura 6 Selección Productos Usuario

3.7 Editar Asignación a Usuarios

Desde el mantenedor de productos usuario se editan los usuarios del producto, para esto se debe seleccionar y luego generar la edición. Ver “Figura 7 Edición Usuario”

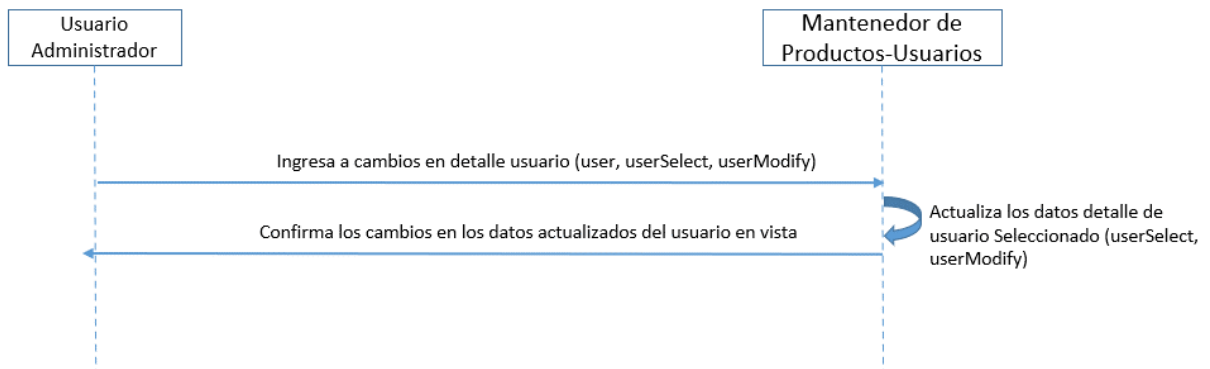


Figura 7 Edición Usuario

3.8 Confirmar Cambios de Estados de Series

En el módulo donde el usuario analista modifica las series, luego de realizar los cambios de estado (aprobado o rechazado) tiene la opción de confirmar los cambios, esto actualiza el estado de cambios en el sistema y cierra la etapa de análisis de calidad del producto para esa fecha. Ver “Figura 8 Confirmación Cambios”

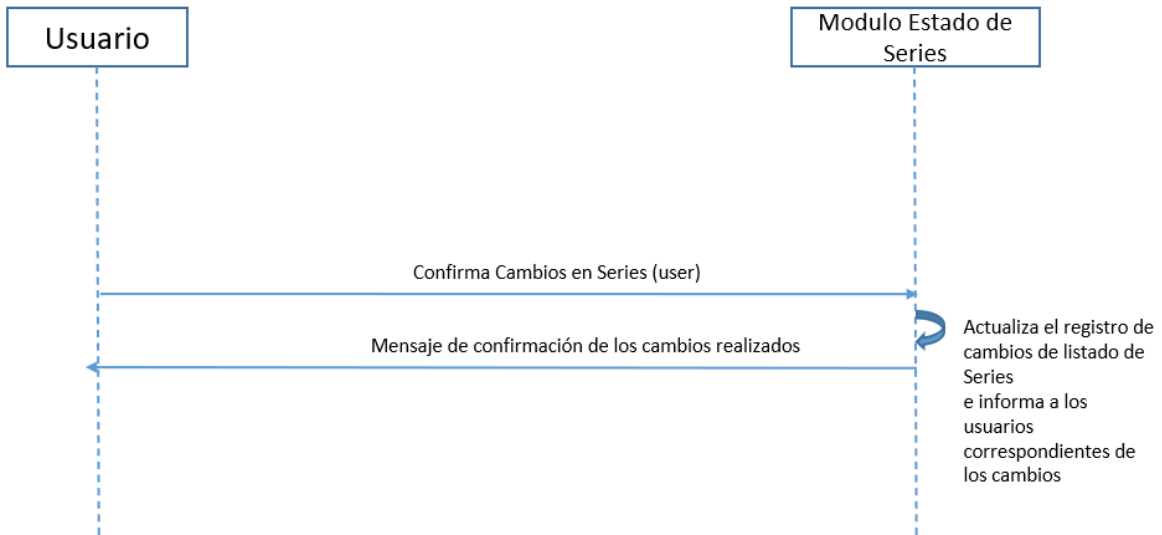


Figura 8 Confirmación Cambios

3.9 Enviar mensaje a usuario Productor

En el módulo donde el usuario analista modifica las series, puede enviar mensaje al productor con algún tipo, este mensaje lo crea el sistema para luego enviar a la cuenta del usuario. Ver “Figura 9 Notificación a Productor”.

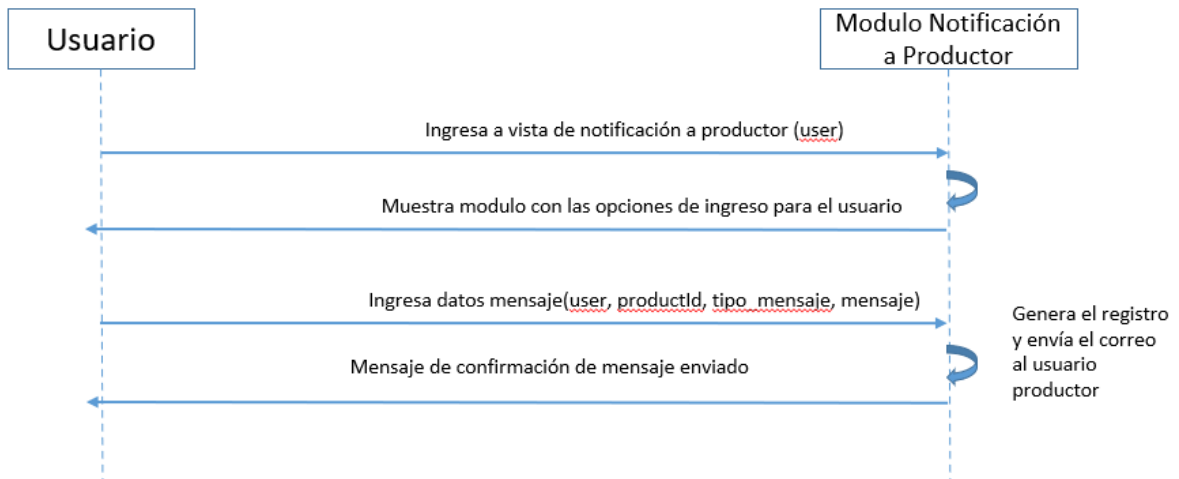


Figura 9 Notificación a Productor

4 Medidas de Análisis de Series de Tiempo

Estas medidas son algoritmos creados para este sistema los cuales se aplican para cada serie de tiempo. La opción de medida depende del comportamiento de la serie, ósea cuál es su tipo, para esto la serie es clasificada en un tipo por su tendencia, estacionalidad o aleatoriedad. Además, las medidas que se aplican por cada serie se separan en dos grupos, las de Orientación y las de Preevaluación.

4.1 Tipo de Series de Tiempo

Se considera el Tipo de Series de Tiempo a la clasificación de la serie por el comportamiento de sus datos en la muestra de datos obtenida.

4.1.1 Tendencia

Es cuando el comportamiento de la serie en el tiempo evaluado tiene un patrón gradual creciente o decreciente. Ver “Figura 10 Tipo de Serie Tendencia Creciente” y “Figura 11 Tipo de Serie Tendencia Decreciente”.

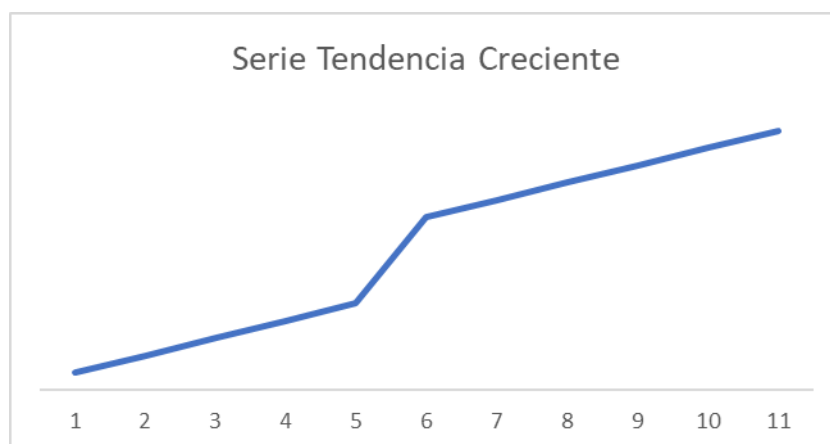


Figura 10 Tipo de Serie Tendencia Creciente

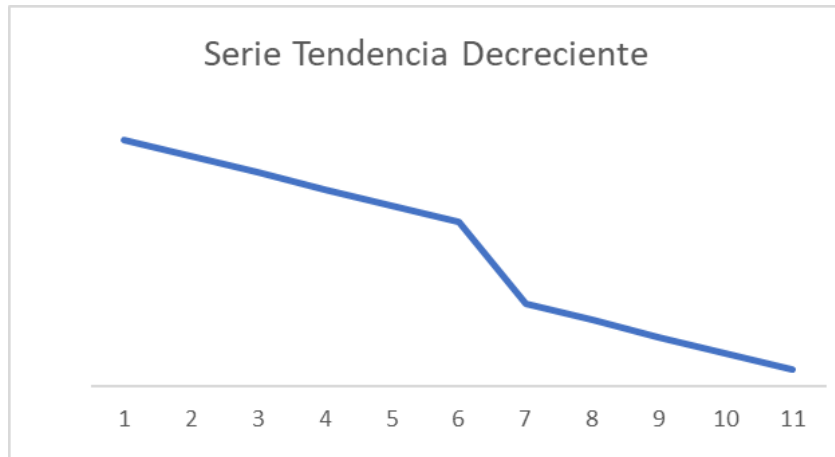


Figura 11 Tipo de Serie Tendencia Decreciente

4.1.2 Estacionalidad

Es cuando una serie en el tiempo evaluado de un año tiene variaciones y estas tienen un comportamiento predecible con respecto a los años anteriores. Ver “Figura 12 Tipo de Serie con Estacionalidad”.



Figura 12 Tipo de Serie con Estacionalidad

4.1.3 Estacionalidad con Tendencia

Es cuando una serie tiene variaciones en un año y estas tiene un comportamiento predecible respecto a años anteriores y además es creciente o decreciente. Ver “Figura 13 Tipo de Serie con Tendencia y Estacionalidad”.



Figura 13 Tipo de Serie con Tendencia y Estacionalidad

4.1.4 Aleatoria

Es cuando la serie en el tiempo evaluado no tiene un patrón identificable. Ver “Figura 14 Tipo de Serie Aleatoria”.



Figura 14 Tipo de Serie Aleatoria

4.2 Medidas Aplicadas

Consiste en el grupo de medidas generadas en el proceso de evaluación de la serie. Se separan en dos enfoques, las medidas de orientación y las medidas preevaluación. Sus funciones son apoyar el análisis mediante información de comportamiento, gráficos, valor dentro de rangos y/o que su valor aprueba algún “test” estadístico.

4.2.1 Medidas de Orientación

Consiste en las medidas obtenidas para cada serie en donde se pueda observar el comportamiento en el tiempo. Estas medidas son promedio, variación, tendencia, estandarización y distancia a la mediana de los tramos de tres, seis, doce y veinticuatro meses de la serie. Ver “Figura 15 Gráfica Medida Orientación – Tendencia”.

Para el caso del promedio por distintos tramos se puede comparar de manera visual el comportamiento de la serie en el periodo evaluado, por ejemplo, comparar el promedio del primer año ver sus segundo y tercero.

La variación contempla que para un grupo de datos se compara su variación respecto a los periodos anteriores, por ejemplo, compara la variación de grupo de tres meses seguidos con el siguiente grupo de tres meses, lo mismo para seis, doce y veinticuatro meses.

La medida de tendencia determina que la serie de tiempo es creciente, decreciente o sin tendencia.

La estandarización muestra la serie sin su tendencia y quitando la varianza.

Para el caso de la distancia media entre tramos determina la diferencia absoluta de un tramo u otro, por ejemplo, las diferencia entre el primer año con el segundo y el tercero.

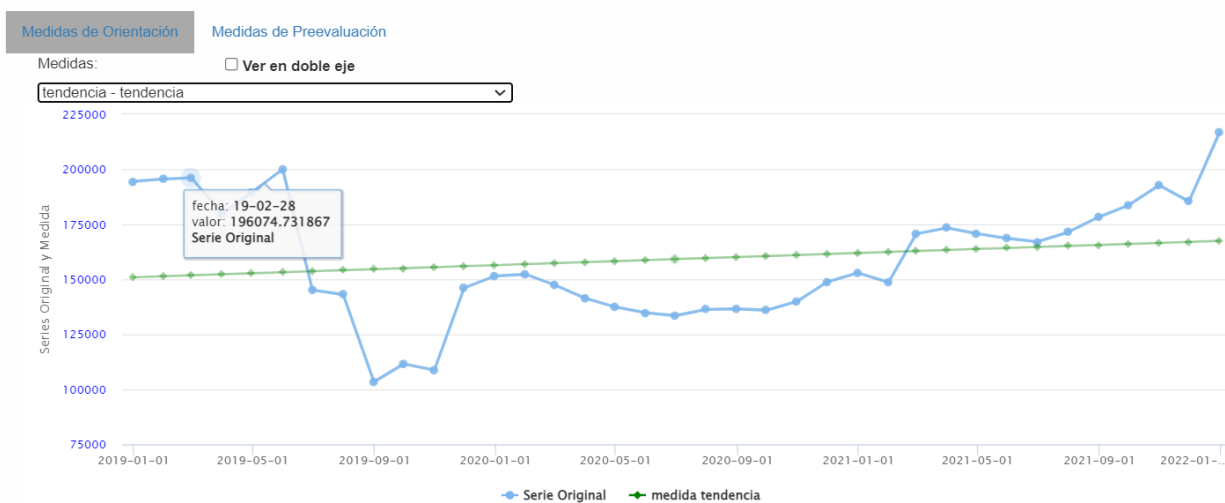


Figura 15 Gráfica Medida Orientación – Tendencia

4.2.2 Medidas de Preevaluación

Las medidas de preevaluación son las que se ejecutan en el sistema para establecer si la serie se aprueba o se rechaza. Estos distintos criterios estadísticos dependen del tipo de serie definido, estos se muestran para el usuario como el límite inferior y superior de una serie para el periodo evaluado y/o que el valor evaluado apruebe un “test” estadístico. Ver “Tabla 11 Medidas Aplicadas por Tipo Serie”.

		Tipo de Series			
		Estacionalidad (E)	Estacionalidad con Tendencia (TE)	Tendencia (T)	Aleatorias (A)
Medida	Límite Inferior	•	•		
	Límite Superior	•	•		
	Banda Re-muestreo			•	•
	Banda Inter cuartil			•	•
	Test Chauvenet 36			•	•
	Test Chauvenet 24			•	•
	Test Desviación Mediana Absoluta (MAD)			•	•
	Intervalo de Confianza 95			•	•

Tabla 11 Medidas Aplicadas por Tipo Serie

Las fórmulas matemáticas para crear las medidas de preevaluación implementados en el sistema se encuentran detalladas en “Anexo C. Procedimiento para Tratar Series con Tendencia y Estacionalidad” y “Anexo D. Pasos para Tratamiento de Series Aleatorias”.

5 Modelo de Datos del Sistema

El Modelo de Datos del sistema de Calidad de información consiste en las tablas, columnas y sus relaciones para poder dar solución al almacenamiento de la información analizada, las medidas obtenidas de los procesos de cálculo, registros de los procesos de análisis y la administración de las distintas características de la solución. Debido al escalamiento que puede tener la cantidad de series analizadas en el sistema se ha separado en dos bases de datos en el servidor, la base de datos que almacena las series evaluadas y la base de datos que contiene el modelo de datos que gestiona el sistema.

5.1 Almacenamiento de las Series Analizadas

Para el caso de la información de las series analizada se ha creado la base de datos EEFF_Metrics, esta contiene la información las series de tiempo con los resultados de sus medidas. El sistema por cada serie analizada genera dos tablas ambas con el nombre del id de la serie, pero con distinto esquema en la base de datos, la primera contiene la información de la serie con sus medidas de orientación, la segunda contiene las medidas de preevaluación.

5.1.1 Tabla Medida de Contraste de la Serie

Las columnas de una tabla con las medidas de contraste son:

Columna	tipo	
Fecha	int	NOT NULL
Valor	float	NULL
Tendencia	float	NULL
valor_sin_tendencia	float	NULL
metric1	float	NULL
metric2	float	NULL
metric3	float	NULL

metric4	float	NULL
metric5	float	NULL
metric6	float	NULL
metric7	Float	NULL
metric8	float	NULL
metric9	float	NULL
metric10	float	NULL

Tabla 12 Medida Contraste

- fecha: contiene la fecha en formato AAAAMMDD, tipo int.
- valor: contiene el valor de la serie, tipo float.
- tendencia: contiene el valor de la tendencia de la serie, tipo float.
- valor_sin_tendencia: contiene el valor quitando la tendencia, tipo float.
- De metric1 a metric10: contiene los valores generados en el cálculo de medidas, tipo float.

5.1.2 Tabla Medida de Preevaluación de la Serie

Las columnas de una tabla con las medidas de preevaluación son:

columna	tipo	
fecha	int	NOT NULL
codMetric	int	NOT NULL
valorMetric	float	NULL
estadoMetric	int	NULL

Tabla 13 Medida Preevaluación

- fecha: contiene la fecha en formato AAAAMMDD, tipo int.
- codMetric: contiene el id de la medida de preevaluación, tipo int.
- valorMetric: contiene el valor de la medida de preevaluación para el registro de la serie, tipo float.

- estadoMetric: contiene el estado de la preevaluación (aprobado, rechazado, pendiente), tipo int.

5.2 Almacenamiento de Proceso de Análisis y Administración del Sistema

La base de datos EEFF_Metrics_Gestion es donde se implementa el siguiente modelo de datos. Ver “Figura 16 Modelo de Datos”.

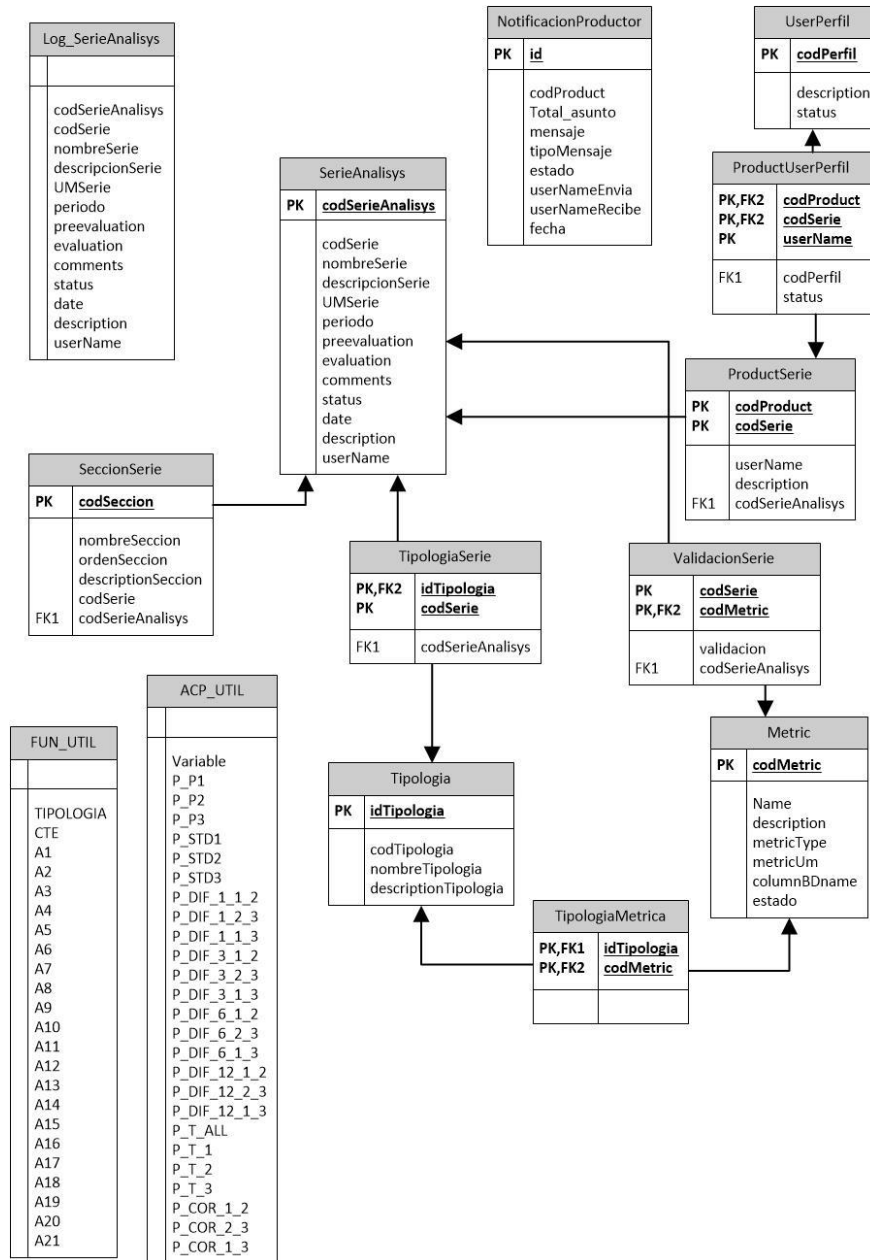


Figura 16 Modelo de Datos

Las tablas se pueden clasificar por tres funcionalidades:

- Las tablas que almacenan el proceso de análisis.
- Las que contiene el detalle de la información descriptiva de la serie como su sección, tipo y medidas.
- Las tablas que almacenan información para la automatización de la definición de tipo de la serie.

Este modelo responde a lo especificado en la toma de requerimientos de cómo se debería gestionar el análisis con respecto a sus usuarios, productos, series, medidas y tipologías.

5.2.1 Almacenamiento de Proceso de Análisis

Estas contienen la información descriptiva generada en el proceso de análisis de cada serie, las evaluaciones generadas por los algoritmos del sistema, las evaluaciones del analista, sus comentarios, mensajes enviados al productor y las medidas aplicadas.

Un proceso de Análisis va registrando en su Log cada proceso ejecutado y las acciones y revisiones del analista.

Un proceso de análisis contiene varias series evaluadas, validaciones y medidas analizadas, este proceso además puede conllevar a varias notificaciones a productor.

Tabla SerieAnalysis

Contiene la información de la serie analizada en el periodo solicitado, el resultado de la evaluación y comentarios que ingresa el analista, además el nombre de usuario y la fecha de la operación.

Esta tabla se relaciona con:

- Tabla SeccionSerie, obtiene las secciones asociadas a la serie.
- Tabla TipologiaSerie, obtiene los tipos de la serie.
- Tabla ProductSerie, obtiene el producto de la serie.

- Tabla ValidacionSerie, contiene las validaciones de la serie.

Cuando se modifica una evaluación este registro se actualiza, para almacenar el historial de cambios se utiliza la tabla Log_SerieAnalysis.

Tabla Log_SerieAnalysis

Contiene el historial de registros de la tabla SerieAnalysis, esta es alimentada por un trigger a esa tabla para cuando se inserta, elimina o actualiza un registro.

Tabla NotificacionProductor

Contiene el registro de las notificaciones enviadas por parte del analista al productor, los usuarios participantes y la fecha de la operación.

Tabla ValidacionSerie

Corresponde a la relación de la serie con las medidas de evaluación que se le aplican en el proceso de análisis.

Está relacionada con:

- Tabla SerieAnalysis, entrega las validaciones al proceso de análisis ejecutado.
- Tabla Metric, obtiene las medidas que se aplican a cada validador.

Tabla Metric

Detalla las medidas que existen de orientación y preevaluación, su descripción, unidad de medida.

Está relacionada con:

- Tabla ValidacionSerie, entrega las medidas aplicadas en la validación.
- Tabla TipologiaMetrica, obtiene la relación de las tipologías y las métricas para saber cual o cuales tipologías de series se le aplican las métricas declaradas.

5.2.2 Detalle Información Descriptiva Serie

Las tablas con la información descriptiva de la serie guardan la información de clasificación del producto que pertenece la serie, los usuarios y sus perfiles, la sección y tipología que pertenece la serie, además las medidas.

Un producto puede tener varias secciones y cada sección un grupo de series. Además, puede haber más de un usuario habilitado para trabajar en el producto. Una serie de tiempo pertenece a solo una tipología según su comportamiento y esta tipología tiene distintas medidas asociadas para las series.

Tabla ProductSerie

Contiene la información del producto estadístico al que pertenece la serie, además del usuario a cargo del análisis.

Está relacionada con:

- Tabla SerieAnalysis, entrega la información de las series que pertenecen al producto estadístico.
- Tabla ProductUserPerfil, obtiene los usuarios que participan en la generación de calidad del producto y sus perfiles.

Tabla ProductUserPerfil

Almacena la información de la relación entre el usuario, el producto estadístico y la serie. Se agrega un estado para saber si el usuario está habilitado para realizar operaciones a ese producto y serie.

Está relacionada con:

- Tabla UserPerfil, obtiene la información de los perfiles del sistema.
- Tabla ProductSerie, entrega la información de los usuarios y perfiles para cada producto estadístico y sus series.

Tabla UserPerfil

Detalla la información de los perfiles de usuario.

Está relacionada con:

- Tabla ProductUserPerfil, entrega los perfiles para los productos estadísticos.

Tabla SeccionSerie

Contiene el detalle de la sección a la cual pertenece la serie.

Está relacionada con:

- Tabla SerieAnalysis, entrega las secciones para las series evaluadas del producto.

Tabla TipologiaSerie

Contiene la relación de la serie con las tipologías.

Está relacionada con:

- Tabla SerieAnalysis, entrega los tipos de las series evaluadas.
- Tabla Tipologia, obtiene la información descriptiva de las tipologías definidas.

Tabla Tipologia

Detalla las descripciones de las tipologías.

Está relacionada con:

- Tabla TipologíaSerie, entrega la información descriptiva.
- Tabla TipologiaMetrica, obtiene las métricas asociadas a la tipología.

TipologiaMetrica:

Contiene la relación de tipología y las medidas asociadas.

Está relacionada con:

- Tabla Tipología, entrega las medidas asociadas a las tipologías.
- Tabla Metric, obtiene la información descriptiva de las métricas.

5.2.3 Automatización de Definición de Tipo

Las tablas de automatización de definiciones de tipo se utilizan para realizar la calibración de las series con respecto a qué tipología pertenecen. Estas tablas solo se utilizan para almacenar resultados de cálculos realizados por equipo estadístico de la división, para efectos de este trabajo no se implementa ese algoritmo de cálculo.

Tabla FUN_UTIL:

Contiene valores aplicados para la clasificación de tipología de una serie.

Tabla ACP_UTIL:

Contiene medidas estadísticas por cada variable de clasificación de FUN_UTIL para determinar la tipología de la serie.

6 Modelo de Arquitectura del Sistema

El modelo de arquitectura del sistema es un modelo de tres capas, separando la funcionalidad por capa y para su comunicación se utiliza llamadas API REST. Esto fortalece la estructura jerárquica, además mejora la modularidad de la aplicación, ya que se implementa en el Backoffice de la División, pero no se descarta en un futuro poder trabajar en un ambiente dedicado solo al objetivo de calidad. Cabe mencionar que los componentes de seguridad de usuario (por ejemplo, login del usuario) son parte de sistema Backoffice, estos se muestran en el modelo para que se comprenda la interacción entre los sistemas, al igual que la base de datos EEFF_Analytics. Ver “Figura 17 Modelo de Arquitectura”.

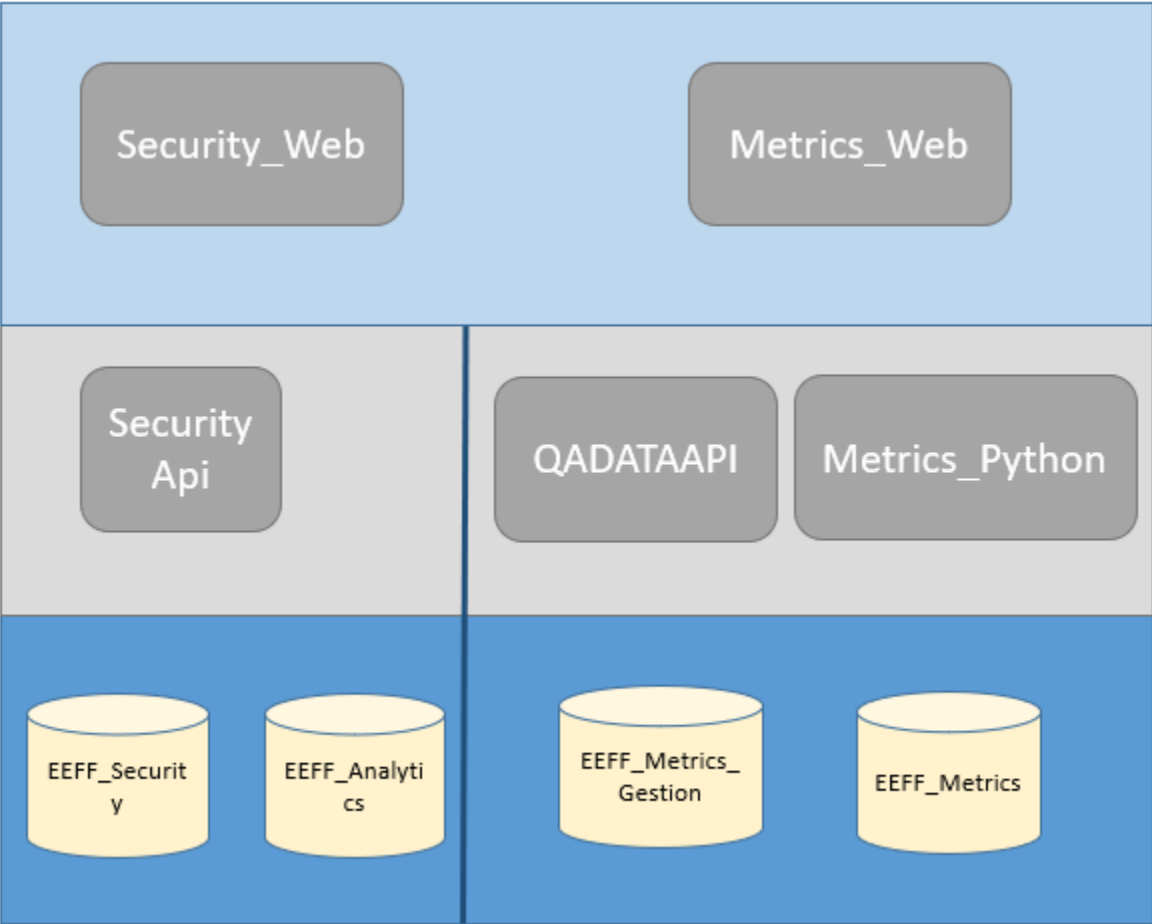


Figura 17 Modelo de Arquitectura

Al tener la separación de servidores del Sistema de Calidad de Datos con sistema Backoffice los procesos de mayor uso de memoria no interrumpen el uso del sistema.

Como la base de datos de Gestión del sistema y la base de datos de medidas de series se almacenan de manera separada, es posible escalar y mejorar el orden de los archivos de esta última base. Esto es importante considerarlo porque en el caso de que se aumente su escalabilidad no afectará al rendimiento.

Los algoritmos de ejecución de cálculo de métricas se calendarizan y se ejecutan en tiempos distintos de la revisión, de esta manera no interrumpen el uso del servidor, pero se debe pensar más adelante en un servidor dedicado para cálculos de alto computo.

6.1 Capa de Datos

La capa de los datos donde se encuentran las bases de datos y procedimientos almacenados para almacenar y gestionar la información de la solución en los servidores involucrados, estos son:

- EEFF_Security. Es la Base de Datos donde se almacena la información de los usuarios, claves y perfiles de los distintos sistemas de la División. Esta base de datos está en servidor Puestad.
- EEFF_Analytics. Es la Base de datos que contiene la información estadística desde el dato de origen hasta el dato publicado, en esta base de datos se encuentran las series de tiempo que se deben analizar. Esta base de datos está en servidor Puestad.
- EEFF_Metrics_Gestión. Es la base de datos que contiene el modelo de datos del sistema de Calidad de Datos en Series de Tiempo. Esta base de datos está en servidor Questad.
- EEFF_Metrics. Es la base de datos que contiene las series de tiempo evaluadas más las medidas obtenidas por los algoritmos creados para la automatización del análisis de calidad. Esta base de datos está en servidor Questad.

Servidor SQL Puestad están montados en Windows Server Enterprise 2016 y Questad está Windows Server Standard 2016.

6.2 Capa de Negocios

La capa de negocios es el Backend del sistema que contiene las aplicaciones que consumen los mensajes capturados en la capa de presentación por medio de patrones API REST y realizan las modificaciones en las bases de datos respectivas:

- Security Api. Es el encargado coordinar los mensajes de la aplicación de seguridad. Este se utiliza para el login de los usuarios que participan en el Sistema. Aplicación está en montada en servidor Duestad.
- QADATAAPI. Backend encargado de la transmisión de mensajes desde el Frontend del sistema de Calidad de Datos a sus Bases de datos respectivas.
- Metrics Python. Aplicación creada para ejecutar las métricas de calidad de las series de tiempo solicitadas por los usuarios. Aplicación está creada en Python y montada en servidor Questad. Este algoritmo está basado en las fórmulas matemáticas mostradas en “Anexo C. Procedimiento para Tratar Series con Tendencia y Estacionalidad” y “Anexo D. Pasos para Tratamiento de Series Aleatorias”.

Las aplicaciones Security Api y QADATAAPI están creadas en C# .Net

6.3 Capa de Presentación

La capa de presentación contiene el Frontend del sistema, esta es la interfaz con la que el usuario interactúa.

- Security_Web. Es donde el usuario realiza el login a los sistemas.
- Metrics_Web. Es la parte donde se implementa la interfaz del Sistema de Calidad de Datos.

Ambas se encuentran levantadas en servidor Duestad.

7 Vistas del Sistema de Calidad de Datos

7.1 Vista Principal Panel Evaluación de Series del Producto

Corresponde a la vista que ve el usuario al ingresar al sistema. Se puede observar un resumen del estado de las series del producto, el listado de secciones de series, las opciones de notificar a productor y Terminar evaluación. Ver “Figura 18 Vista Principal”.

Backoffice Estadístico Hola Cristian Maripanguil

CARTCOMFLU - Cartera Comercial Flujos

Flujos de Créditos Comerciales por Tamaño de Empresa

Flujo de Tareas Origenes ETLs Tablas Procedimientos/funciones **Calidad de Información** Cargas BEST Reportes Reportes BEST

Totales Preevaluacion Rechazada : 22
Totales Preevaluacion Aceptada : 63

Totales Evaluacion Rechazada : 0
Totales Evaluacion Aceptada : 0
Totales Evaluacion Pendiente : 85

Sección	Descripcion	Evaluación	Acciones
Operaciones en cartera comercial segregadas por tamaño según deuda, flujos, plazo menor o igual a un año	(número)	En Proceso	
Cartera comercial segregada por tamaño según deuda, flujos, plazo menor o igual a un año	(unidades de fomento)	En Proceso	
Operaciones en cartera comercial segregadas por tamaño según deuda, flujos, plazo mayor a dos años	(número)	En Proceso	
Cartera comercial segregada por tamaño según deuda, flujos, plazo entre uno e igual o menor a dos años	(unidades de fomento)	En Proceso	
Operaciones en cartera comercial segregadas por tamaño según deuda, flujos, plazo mayor a dos años	(número)	En Proceso	
Cartera comercial segregada por tamaño según deuda, flujos, plazo mayor a dos años	(unidades de fomento)	En Proceso	
Operaciones en cartera comercial segregadas por tamaño según deuda, flujos, moneda nacional no reajutable	(número)	En Proceso	
Cartera comercial segregada por tamaño de deuda, flujos, moneda nacional no reajutable	Tasa de interés promedio ponderada	En Proceso	
Cartera comercial segregada por tamaño de deuda, flujos, moneda nacional no reajutable	(unidades de fomento)	En Proceso	
Operaciones en cartera comercial segregadas por tamaño según deuda, flujos, moneda nacional reajutable	(número)	En Proceso	

[Notificar a Productor](#)
[Terminar Evaluación Producto](#)

Figura 18 Vista Principal

7.2 Vista Revisión Sección del Producto

Al seleccionar una sección del listado se observan las series correspondientes a ésta, además la información de la serie, el tipo de serie, un gráfico pequeño que muestra los últimos 12 meses, el estado de preevaluación, la evaluación realizada por el analista y las observaciones ingresadas. Además, está el botón de aprobar o rechazar para la sección o para la serie. Ver “Figura 19 Revisión Sección”.

Backoffice Estadístico Hola Cristian Maripanguil v. 5.0.0

Totales Preevaluacion Rechazada : 22
 Totales Preevaluacion Aceptada : 63

Totales Evaluacion Rechazada : 0
 Totales Evaluacion Aceptada : 0
 Totales Evaluacion Pendiente : 85

Sección	Descripción	Evaluación	Acciones
Operaciones en cartera comercial segregadas por tamaño según deuda, flujos, plazo menor o igual a un año	(número)	En Proceso	
Cartera comercial segregada por tamaño según deuda, flujos, plazo menor o igual a un año	(unidades de fomento)	En Proceso	<input type="button" value="Rechazar Sección"/> <input type="button" value="Aprobar Sección"/>

Serie(Click para ver/editar detalles)	Tipo	Gráfico (últimos periodos)	Pre Evaluación	Evaluación	Comentarios	Acciones
Mediano	Estacionalidad		Rechazado	pendiente		<input type="button" value="Rechazar"/> <input type="button" value="Aprobar"/>
Micro	Estacionalidad		Rechazado	pendiente		<input type="button" value="Rechazar"/> <input type="button" value="Aprobar"/>
Pequeño	Estacionalidad		Rechazado	pendiente		<input type="button" value="Rechazar"/> <input type="button" value="Aprobar"/>
Grande	Tendencia		Aprobado	pendiente		<input type="button" value="Rechazar"/> <input type="button" value="Aprobar"/>

Figura 19 Revisión Sección

7.3 Vista Detalle Evaluación Serie – Medidas de Orientación

Esta corresponde a la vista del detalle de la serie seleccionada. Contiene información de la serie, situación de análisis, gráficos lineales de las medidas de orientación y las opciones del analista para aprobar o rechazar y agregar comentarios. Ver “Figura 20 Detalle Evaluación Orientación”

Información Serie

Nombre Serie : Micro
 Código Serie : SBIF_CART_CCO_FLU_TAMDEU_MIC_OA1_UF_MONT
 Descripción Serie : Cartera comercial segregada por tamaño de deuda, flujos, deudores micro, plazo menor o igual a un año
 Unidad de Medida Serie : Sin medida representación

Comentarios: Valor no se encuentra a los rangos esperados, posible caso de rectificación de cifras
 Máx. 512 caracteres

Rechazar
 Aprobar

Análisis

Medidas de Orientación

Medidas de Preevaluación

Medidas: Ver en doble eje
 distMed12 - distancia absoluta entre la mediana desde menos 12 m

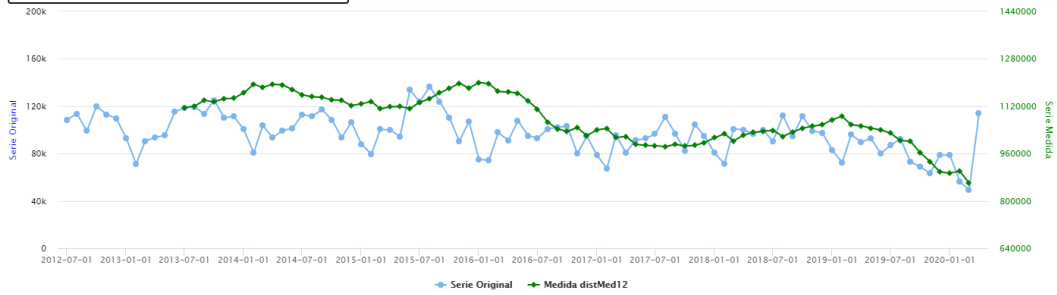


Figura 20 Detalle Evaluación Orientación

7.4 Vista Detalle Evaluación Serie – Medidas de Preevaluación

En la vista de detalle de la serie además está la información de la preevaluación que corresponde a las métricas aplicadas a la serie por su tipología. Ver “Figura 21 Detalle Evaluación Preevaluación”

Información Serie

Nombre Serie : Micro
 Código Serie : SBIF_CART_CCO_FLU_TAMDEU_MIC_OA1_UF_MONT
 Descripción Serie : Cartera comercial segregada por tamaño de deuda, flujos, deudores micro, plazo menor o igual a un año
 Unidad de Medida Serie : Sin medida representación

Comentarios: Valor no se encuentra a los rangos esperados, posible caso de rectificación de cifras
 Máx. 512 caracteres

Rechazar
 Aprobar

Análisis

Medidas de Orientación

Medidas de Preevaluación

periodo Evaluado	Nombre	Descripción	Valor	Condición	Estado
20200401	BANDA	Contienen banda como medida de preevaluación	112508.05124494468	Lim_Inf <= @valor <= Lim_Sup	Rechazado
20200401	BANDA	Contienen banda como medida de preevaluación	67895.03555505534	Lim_Inf <= @valor <= Lim_Sup	Rechazado

Figura 21 Detalle Evaluación Preevaluación

7.5 Vista Notificación a Productor

En la vista el usuario analista puede enviar mensajes al productor, por ejemplo, para solicitar una rectificación de los datos o actualización de las métricas. Ver “Figura 22 Notificación Productor”.

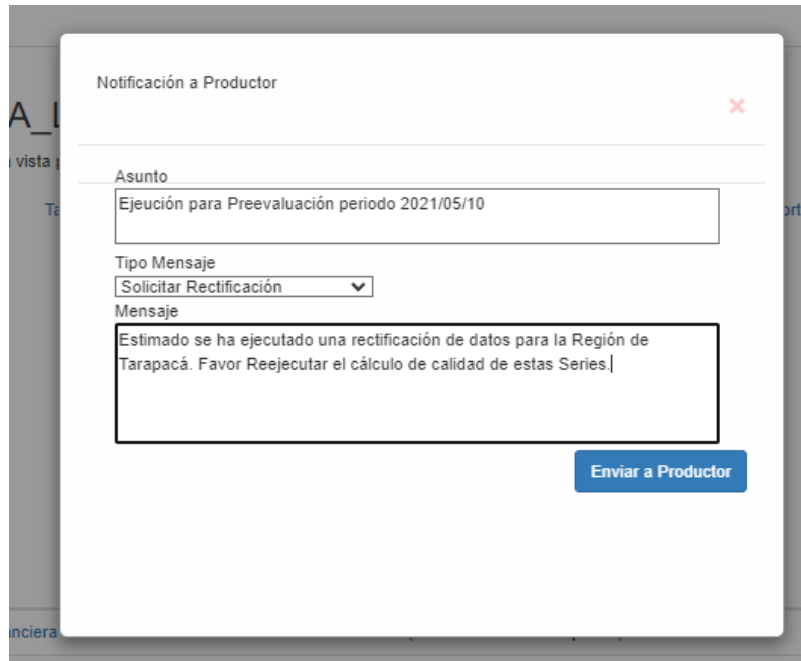


Figura 22 Notificación Productor

7.6 Vista Mensaje Término de Evaluación

En la vista principal al seleccionar el botón para terminar la evaluación muestra la opción de confirmar el término o cancelar. Ver “Figura 23 Término Evaluación”.

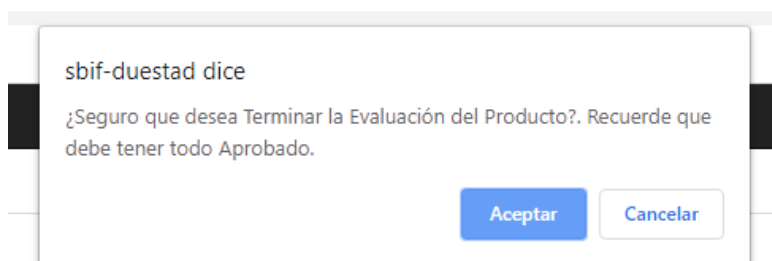


Figura 23 Término Evaluación

8 Implementación y Evaluación del Sistema

8.1 Proceso de Implementación Sistema

Para poder realizar la implementación del sistema se realizaron los siguientes pasos:

Primero, se trabajó con ambientes de prueba, en el cual se agregó un producto por analista, que pudiera revisar las series del producto y saber si están bien ajustadas las medidas obtenidas por los algoritmos de calidad de información. Para esto se dio el plazo de un mes ya que debían comparar con los análisis de las planillas de cada analista, además de que esta revisión no atrasara las entregas de información habitual. Para este paso sólo se ha trabajado con series de tipo estacional.

Segundo, luego que este sistema fuera evaluado a nivel de usuarios con sus respectivos productos estadísticos se agregaron las mejoras solicitadas, y se habilitó el ambiente de producción.

Tercero, ya validado el sistema anclado a Backoffice de la División Estadística se implementó un proceso de marcha blanca para un grupo de productos estadísticos. Para este grupo de productos actualmente se controla su calidad de información por el sistema.

Cuarto, se agregan medidas para tipos de series con tendencia pura y aleatorias, con esto se observa un aumento en la cantidad de Outliers.

8.2 Evaluación del Sistema

Es necesario para la evaluación del sistema capturar a través de medidas la percepción del usuario y usabilidad del sistema. Además, que esta sea acorde a los registros que se pueden obtener desde la misma base de datos, por ejemplo, la cantidad de periodos de calidad ejecutados, los productos implementados, la respuesta de las medidas de preevaluación y si estas son acordes al análisis del usuario.

8.2.1 Evaluación de los Objetivos y Registros del Sistema

Dentro de los objetivos establecidos en un principio, el incorporar los productos estadísticos de los analistas en la herramienta ha sido el que más ha costado, ya que del total de cuarenta productos solo se han incorporado cinco, lo que equivale al 12,5%.

Por otro lado, en los algoritmos de cálculo se han incorporado el total de medidas solicitadas por los analistas, considerando los tipos de series especificados. Además, la información entregada por el sistema responde a la información requerida para el correcto análisis.

Para poder evaluar la herramienta se utilizaron los registros de los análisis de series realizados desde el principio de la marcha blanca hasta la fecha de inicio de escritura de este documento, observándose lo siguiente:

El total de series evaluadas con la herramienta fueron 8.071 de los cuales 643 casos corresponden a datos rechazados por los algoritmos de calidad (8%). Ver “Figura 24 Total de Series Evaluadas - Aprobadas vs Rechazadas”.

Además, solo han existido 122 casos en que el analista ha rechazado una serie que fue aprobada por el algoritmo, esto equivale al 1,6% de las series aprobadas.

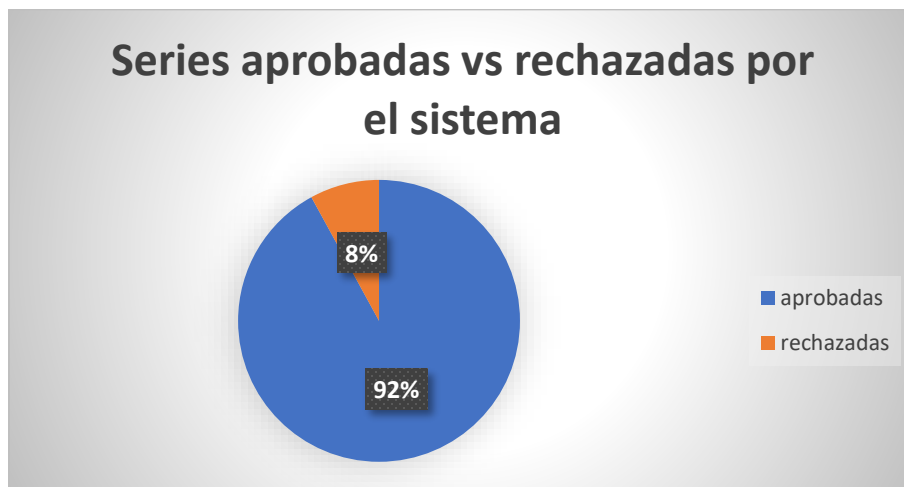


Figura 24 Total de Series Evaluadas - Aprobadas vs Rechazadas

Al revisar la cantidad de series evaluadas por periodo, ver “Figura 25 Comparación Series Aprobadas vs Rechazadas por Periodo”, se observa que en un principio los porcentajes de outliers encontrados eran mínimos, pero con el tiempo empezaron a aumentar, esto porque, además de los acontecimientos exógenos que pudieron modificar la cantidad de series rechazadas, como por ejemplo pandemia o cierre de sucursales y cajeros bancarios, además se agregaron las medidas de otras tipologías de series, las cuales se explicaron en el capítulo 4.

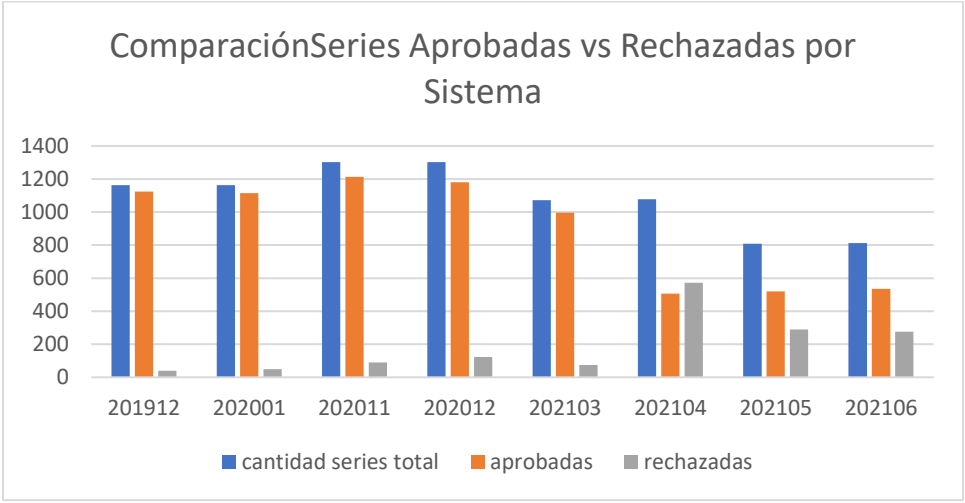


Figura 25 Comparación Series Aprobadas vs Rechazadas por Periodo

Se observa del gráfico anterior que por periodo se evalúan entre 1.000 y 1.400 series de los distintos productos hasta mayo 2021 el cual presenta una disminución, esto se puede asociar a la resistencia al cambio de las nuevas metodologías implementadas para la evaluación de calidad de la serie, ya que al integrar estas medidas la cantidad de outliers aumentaron de un 6% en promedio a un 40% aproximadamente.

8.2.2 Evaluación de Usabilidad por parte de Usuarios

Para poder dar respuesta a la percepción de los usuarios de la División y poder establecer una medida se genera una encuesta de diez preguntas orientados a este sistema. De las cuales las preguntas impares son orientadas a factores positivos del sistema y las pares a aspectos negativos. Se detallan resultados en “Tabla 14 Respuestas y Puntajes Encuesta Usabilidad”.

Las respuestas de estas preguntas son cinco, de las cuales cada una se le asigna un puntaje, se detalla a continuación:

- Respuesta 1, Totalmente en desacuerdo, esta tiene un valor de 1.
- Respuesta 2, En desacuerdo, esta tiene un valor de 2.
- Respuesta 3, Ni de acuerdo ni en desacuerdo, esta tiene un valor de 3.
- Respuesta 4, De acuerdo, esta tiene un valor de 4.
- Respuesta 5, Totalmente de acuerdo, esta tiene un valor de 5.

El promedio de resultados de la encuesta es 61,25%, donde si es mayor del 50% esta en el rango aceptable. La pregunta con mayor aceptación es si se desea utilizar este sistema más frecuentemente, donde la opción mayormente seleccionada fue “De acuerdo”. La pregunta con menor aceptación, ósea mayores casos de “En desacuerdo” o “Totalmente en desacuerdo” es si era necesario ayuda de soporte técnico para el uso del sistema.

Pregunta	Respuestas					
	usuario 1	usuario 2	usuario 3	usuario 4	usuario 5	usuario 6
1	4	5	4	4	4	4
2	3	3	3	3	2	2
3	3	4	2	4	3	4
4	3	2	1	4	2	1
5	3	3	2	3	3	4
6	3	2	4	2	2	2
7	4	4	2	2	4	5
8	3	2	3	5	3	2
9	3	4	4	2	4	4
10	4	2	4	3	2	1

Tabla 14 Respuestas y Puntajes Encuesta Usabilidad

9 Conclusiones y Trabajo Futuro

9.1 Conclusiones

Desde que se empezó a utilizar la herramienta se han evaluado 8.071 series distintas, de estas, el 8% de casos se identificaron como outliers de los cuales 3 han terminado como rectificadas de información por parte de la institución que la proporciona, el resto es por un comportamiento distinto en la serie a su tipología.

Los casos nombrados anteriormente apoyan al objetivo de la no publicación de información errónea para disminuir el riesgo constante en las instituciones que trabajan con datos públicos.

El nivel de aceptación de la herramienta por parte de los usuarios es de un 61,25%, esto respalda a que la poca usabilidad que tuvo ha sido aceptable por los usuarios y que se debe seguir trabajando para poder mejorar esta medida.

Esta herramienta fue una idea propuesta por el alumno para la División de Estadísticas Financieras la que fue bien acogida, por el valor agregado que aporta, además se enriqueció con los distintos procesos del proyecto que se realizaron con el aporte del equipo de trabajo, desde la captura de requerimientos, utilizando casos de uso para la identificación de cada necesidad, además integrando en la solución requerimientos no funcionales como mantenibilidad y escalabilidad, utilizando buenas prácticas de programación e implementando a un ciclo de integración continua en ambiente de desarrollo, pruebas y producción. Este último punto fue fundamental para ir entregando y mejorando la aplicación a los usuarios, y tener dentro de las otras tareas de la división un espacio para avanzar en este proyecto. Las conclusiones que se desprenden de la experiencia al usar la herramienta por los analistas son que el proyecto cubre con éxito las necesidades de integración, estandarización, documentación y registro del proceso de análisis de calidad estadística de series para grandes volúmenes de datos. Además, como esta se integró al sistema Backoffice como una etapa más del proceso estadístico se incorpora de manera natural al ciclo de producción estadística, permitiendo la documentación de las acciones tomadas en términos del proceso de calidad estadística. Antes de esto quedaban en un documento en la carpeta del analista y/o

espacio compartido de la división, lo que disminuye el riesgo de pérdida de información procedimental. Robustece y formaliza la tarea de análisis de series de tiempo de los productos estadísticos, agregando además la variable de tiempo de análisis por producto estadístico y un historial de los análisis realizados por el analista.

Como fue una herramienta creada desde la División de Estadísticas Financieras para sus tareas habituales, la metodología y aplicación propuestas contribuyen al uso eficiente de los recursos Institucionales. Por otra parte, la incorporación de esta herramienta ha contribuido a la transparencia y al acceso a la información de calidad, principios fundamentales en la función pública.

9.2 Trabajo Futuro

Además de ir implementando nuevos productos a la herramienta e ir mejorando, incorporando nuevas medidas de detección de outliers, se han abierto las miradas a una siguiente versión de la herramienta. Los desafíos que se ven en esta nueva versión son orientados a cómo enfrentar los distintos escenarios, como el aumento de frecuencia de información, el aumento del volumen de información, investigar e incorporar algoritmos de machine learning para la detección de series con datos anómalos, como esto afectaría la escalabilidad y los recursos en máquina para el procesamiento de datos, además evaluar soluciones como la nube.

Bibliografía

[LGB_01] Ley General de Bancos. <https://www.leychile.cl/Navegar?idNorma=83135>. Fecha de última visita: Octubre de 2021.

[MSI_04] Manual Sistema de Información Bancos y Financieras, sitio <https://www.sbif.cl/sbifweb/servlet/LeyNorma?indice=3.2&idCategoria=6>. Fecha de última visita: Octubre de 2021.

[MTE_01] "Statistical and Econometrics Software: TRAMO and SEATS," Statistical and Econometrics Software tramoseats, Banco de España; Statistical and Econometrics Software Homepage, revised 2015. Agustín Maravall & Victor Gómez & Gianluca Caporello, 1996. <https://ideas.repec.org/c/bde/sfware/tramoseats.html>. Fecha de última visita: Octubre de 2021.

[MTE_02] <https://www.census.gov/data/software/x13as.html>. Fecha de última visita: Octubre de 2021.

[MTE_03] https://ec.europa.eu/eurostat/cros/content/software-jdemetra_en. Fecha de última visita: Octubre de 2021.

[ORG_02] Organigrama Comisión para el Mercado Financiero <http://www.cmfchile.cl/portal/principal/605/w3-propertyvalue-25543.html>. Fecha de última visita: Octubre de 2021.

[RAN_03] Recopilación Actualizada de Normas Bancos (RAN), sitio <https://www.sbif.cl/sbifweb/servlet/LeyNorma?indice=3.1.2&LNAN=1>. Fecha de última visita: Octubre de 2021.

ANEXOS

Anexo A. Procedimiento de Analista en el Sistema de Aseguramiento de Calidad

El analista debe seguir los siguientes pasos en el sistema (Ver “Figura 26 Pasos Usuario - Uso de Sistema”):

1. Ingreso a Sistema
2. Observar las series evaluadas por algoritmo
3. Analizar si la serie debe ser aprobada o no
4. En el caso de que no sea aprobada, solicitar la información respectiva al proveedor de información, para ver si es necesario actualizar la información
5. En caso de que la serie es aprobada, se debe continuar y publicar

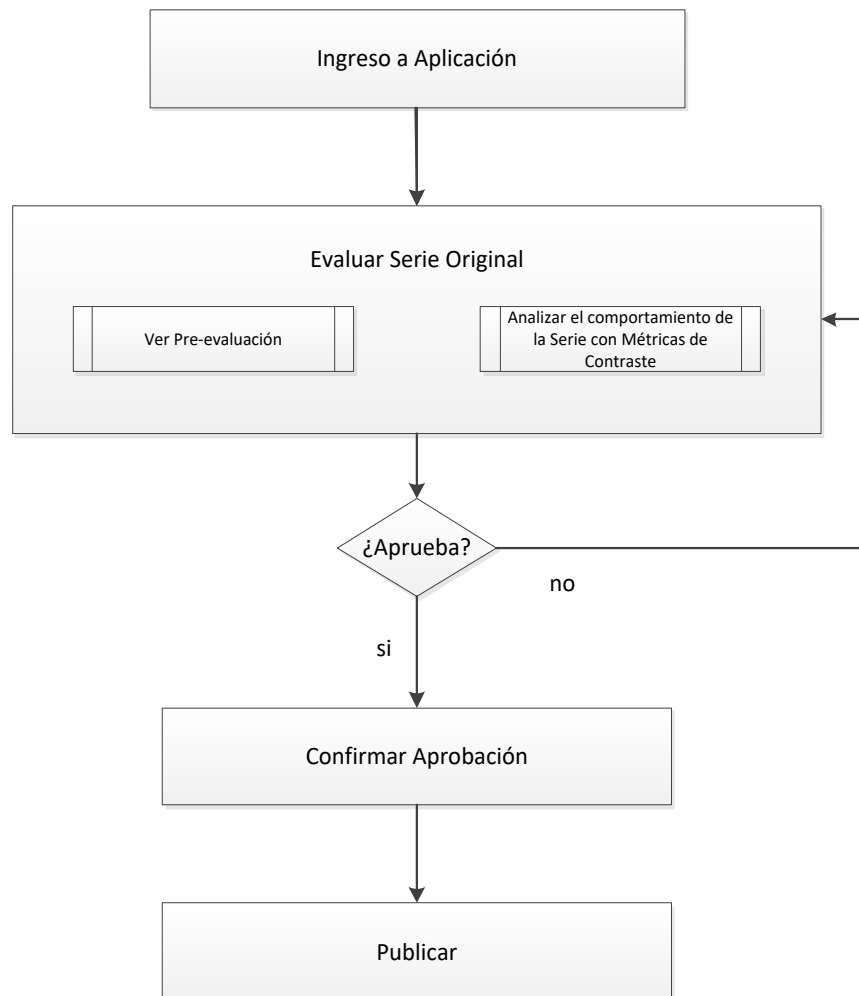


Figura 26 Pasos Usuario - Uso de Sistema

Anexo B. Manual Herramienta Calidad de Información

Ingrese al sistema Backoffice con la siguiente dirección <http://puestad/apps/backoffice/#/>

En Backoffice ingrese su usuario y contraseña ver “Figura 27 Ingreso a Backoffice”

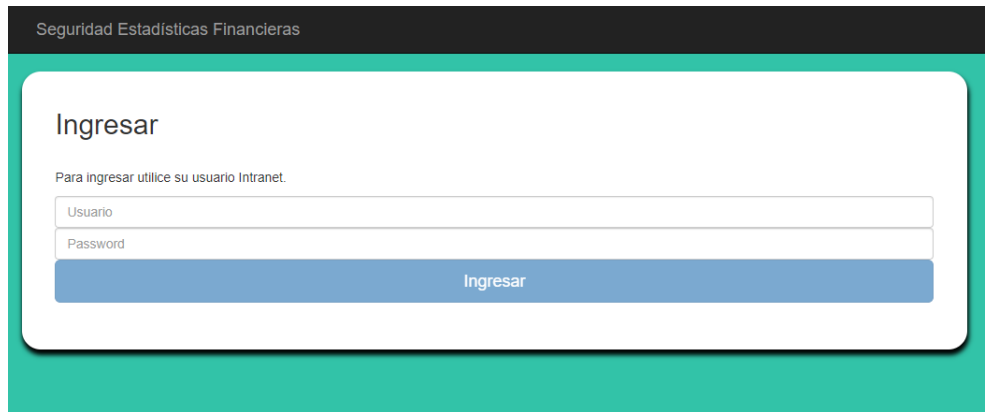
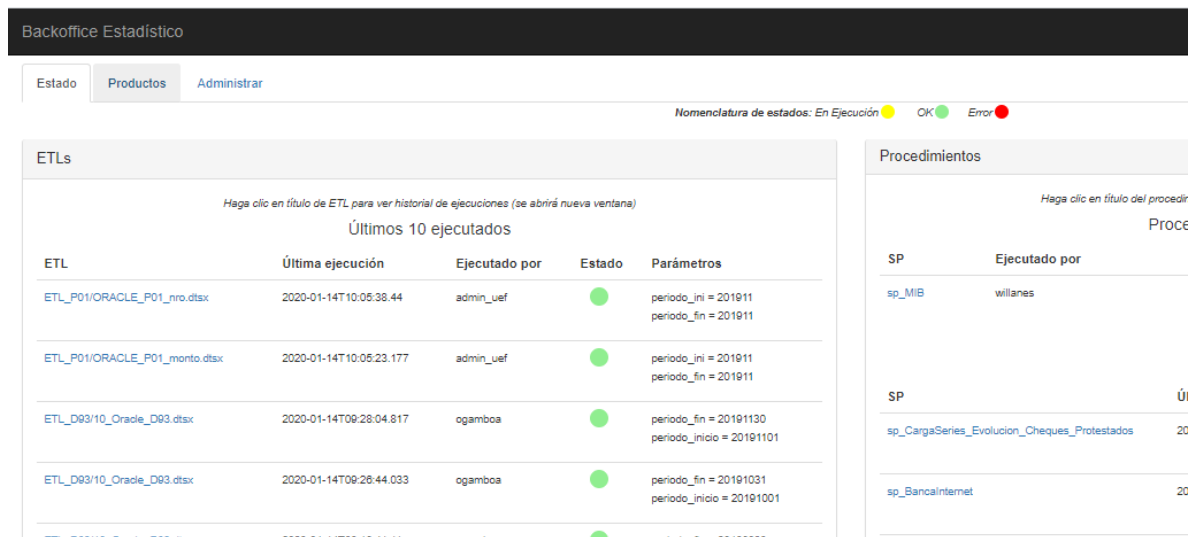


Figura 27 Ingreso a Backoffice

Al ingresar se observará el listado actual de las ejecuciones en Backoffice. Además, la primera vista de sistema backoffice muestra el estado de las etapas actuales del proceso estadístico. Además, muestra una pestaña de los productos estadísticos y la opción de administración.



ETL	Última ejecución	Ejecutado por	Estado	Parámetros
ETL_P01/ORACLE_P01_nro.dtsx	2020-01-14T10:05:38.44	admin_uf	OK	periodo_ini = 201911 periodo_fin = 201911
ETL_P01/ORACLE_P01_monto.dtsx	2020-01-14T10:05:23.177	admin_uf	OK	periodo_ini = 201911 periodo_fin = 201911
ETL_D93/10_Oracle_D93.dtsx	2020-01-14T09:28:04.817	ogambo	OK	periodo_fin = 20191130 periodo_inicio = 20191101
ETL_D93/10_Oracle_D93.dtsx	2020-01-14T09:26:44.033	ogambo	OK	periodo_fin = 20191031 periodo_inicio = 20191001
ETL_D93/10_Oracle_D93.dtsx	2020-01-14T09:19:41.44	ogambo	OK	periodo_fin = 20191030

SP	Ejecutado por	Última ejecución
sp_MIB	willanes	
sp_CargaSeries_Evolucion_Cheques_Protestados		20
sp_BancalInternet		20

Figura 28 Vista Principal Backoffice

Al seleccionar la pestaña “Productos” puede filtrar por el producto a evaluar.

Codigo	Nombre	Descripción	Detalles
ACREEDORES	Acreedores financieros por género	Estadísticas sobre tenedores de productos de ahorro o administración del efectivo, para bancos y cooperativas.	Ver... Editar Eliminar
AHORROVIVIENDA	Ahorro Vivienda	Estadísticas de ahorro vivienda	Ver... Editar Eliminar
APV	Ahorro Previsional Voluntario	Estadísticas de Ahorro Previsional Voluntario	Ver... Editar Eliminar
BANCAINTERNET	Banca Internet	Estadísticas Banca Internet	Ver... Editar Eliminar
CAJEROS	Cajeros	Numero cajeros, numero y monto de operaciones en cajeros por región e institución financiera	Ver... Editar Eliminar
CARTCOM	Cartera Comercial	Cartera Comercial por tamaño de deuda y promedio instituciones	Ver... Editar Eliminar
CARTCOMFLU	Cartera Comercial Flujos	Flujos de Créditos Comerciales por Tamaño de Empresa	Ver... Editar Eliminar
CCDEPCA	Cuentas Corrientes, Depósitos y Cuentas de ahorro	Informe de cuentas corrientes, depósitos y cuentas de ahorro por región	Ver... Editar Eliminar
CCSEGGEN	Condiciones de crédito por género	Condiciones de monto, tasa y plaza por tipo de crédito y género	Ver... Editar Eliminar
CEF	CEF	Estadísticas para el consejo de estabilidad financiera	Ver... Editar Eliminar
CHEQUES	Cheques	Evolución mensual de cheques protestados	Ver... Editar Eliminar

Figura 29 Vista Productos

Codigo	Nombre	Descripción	Detalles
CHEQUES	Cheques	Evolución mensual de cheques protestados	Ver... Editar Eliminar
TESTCALIDAD	TEST CALIDAD	Producto para realizar las pruebas de módulo de calidad en backoffice	Ver... Editar Eliminar

Figura 30 Vista Filtro y Selección de Productos

Al ingresar al producto se muestran las distintas etapas del proceso estadístico con sus elementos correspondientes.

Se enfocará en la etapa de Calidad de la Información.

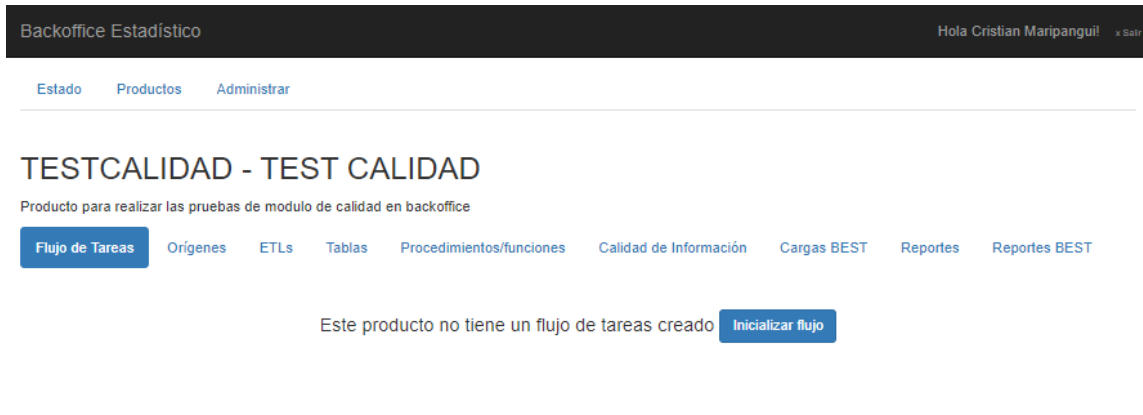


Figura 31 Vista Producto y sus Etapas

Al ingresar al módulo de Calidad de Información se observa:

- Resumen de resultados de la evaluación
- Listado de secciones a evaluar
- Notificación a productor
- Terminación de Evaluación

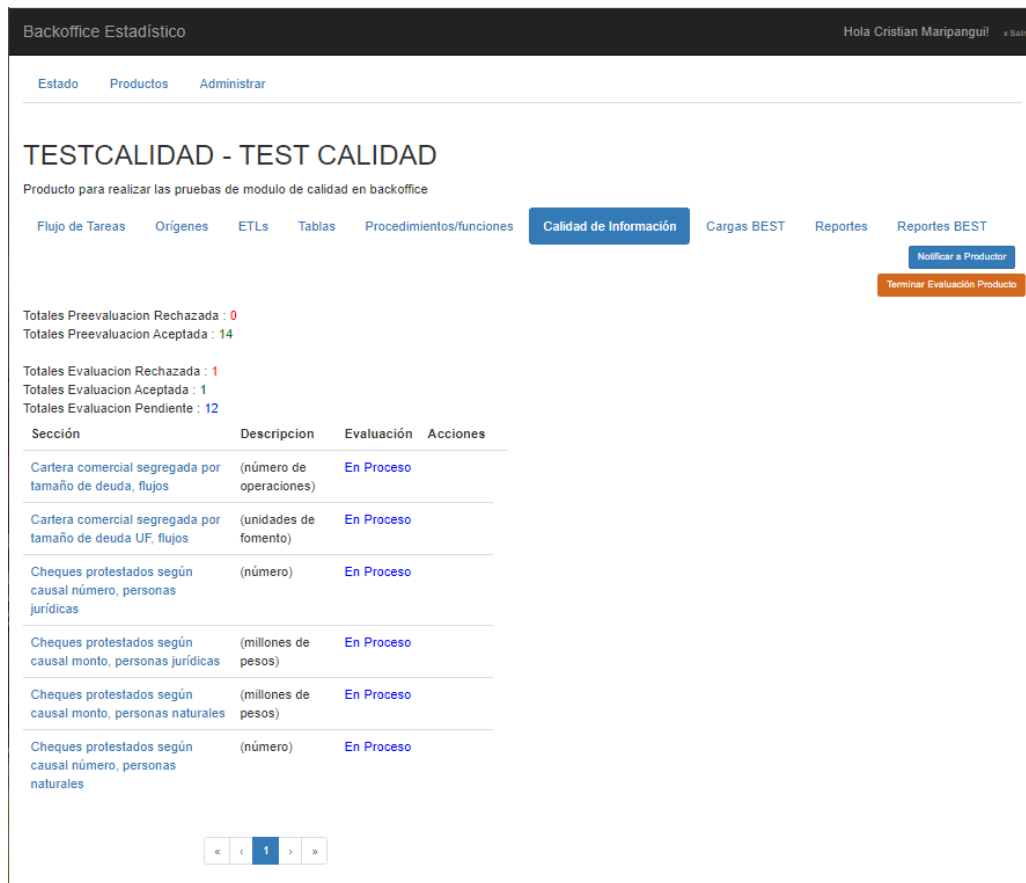


Figura 32 Vista Secciones Producto

Al ampliar una sección de series ve el listado de series de la sección, además de las acciones de rechazar o aprobar la sección y cada serie.

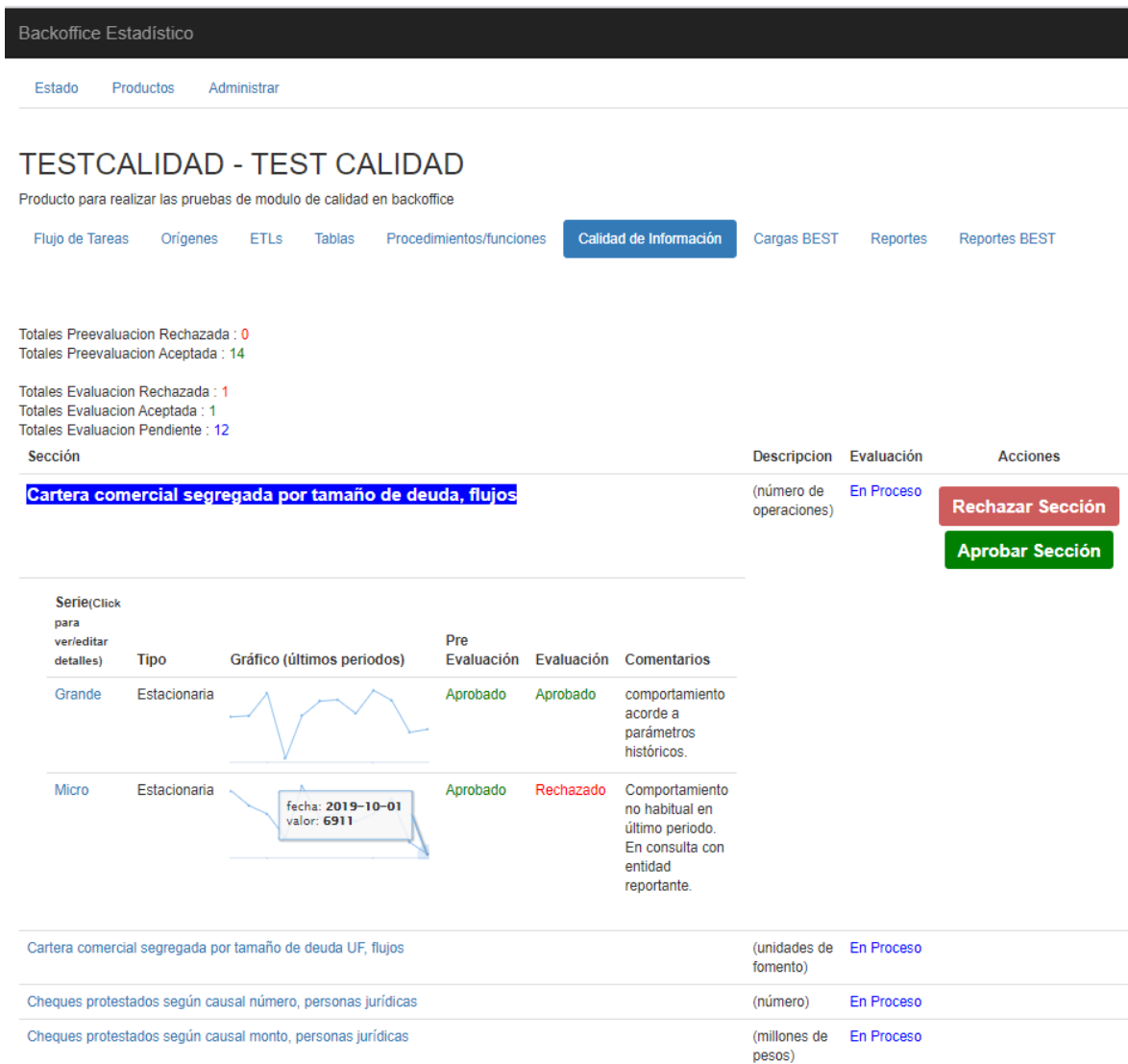


Figura 33 Vista Series Sección

Al seleccionar la serie se obtiene:

- Información descriptiva de la serie
- Representaciones graficas de la serie y el listado de sus medidas
- Medidas de aprobación por parte de los algoritmos
- Opciones de comentar y evaluar la serie

Información Serie

Nombre Serie : Grande
 Descripción Serie : Cartera comercial segregada por tamaño de deuda, flujos, deudores grandes
 Unidad de Medida Serie : Sin medida representación

Análisis



Figura 34 Vista Evaluación Serie Orientación

Información Serie

Nombre Serie : Grande
 Descripción Serie : Cartera comercial segregada por tamaño de deuda, flujos, deudores grandes
 Unidad de Medida Serie : Sin medida representación

Análisis

Medidas de Orientación

Medidas de Preevaluación

periodo Evaluado	Nombre	Descripción	Valor	Condición	Estado
20191001	ZSCORE	descripcion semaforo 1	-0.650491891786	@valor < 4	Aprobado
20191001	ZSCORE	descripcion semaforo 2	-0.650491891786	@valor > -4	Aprobado
20191001	BANDA	Contienen banda como medida de preevaluación	6679.24548742	Lim_Inf <= @valor <= Lim_Sup	Aprobado
20191001	BANDA	Contienen banda como medida de preevaluación	8381.42117925	Lim_Inf <= @valor <= Lim_Sup	Aprobado

Comentarios: comportamiento acorde a parámetros históricos.

Rechazar

Aprobar

Máx. 512 caracteres

Figura 35 Vista Evaluación Serie Preevaluación

En la pestaña principal del sistema de calidad de información al seleccionar el botón de notificar a productor aparece la ventana donde se agrega el mensaje hacia el productor. Este mensaje llega al correo del usuario productor asignado.

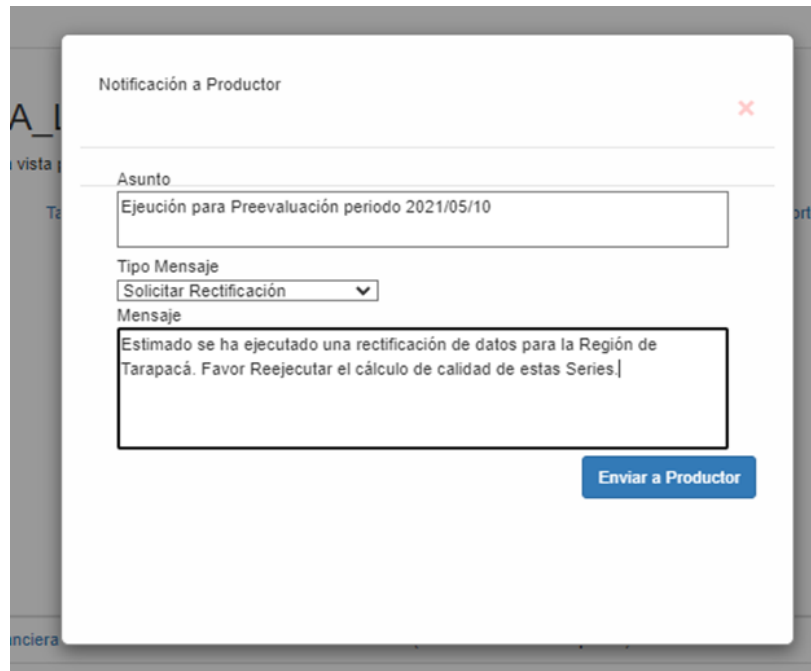


Figura 36 Vista Notificación a Productor

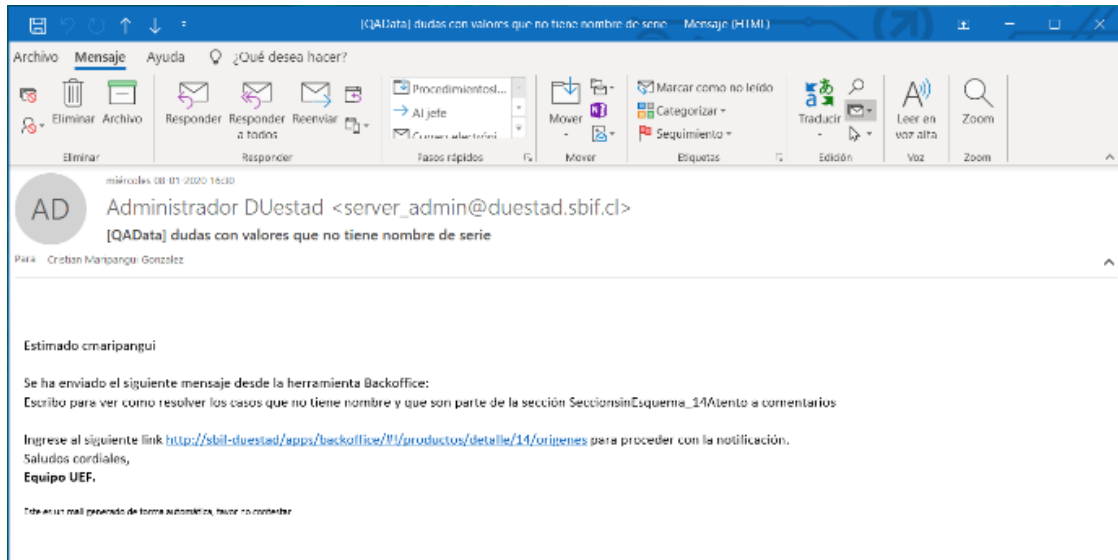


Figura 37 Vista Correo desde Sistema

Además, aparece el cierre de evaluación el producto el cual solicita su confirmación.

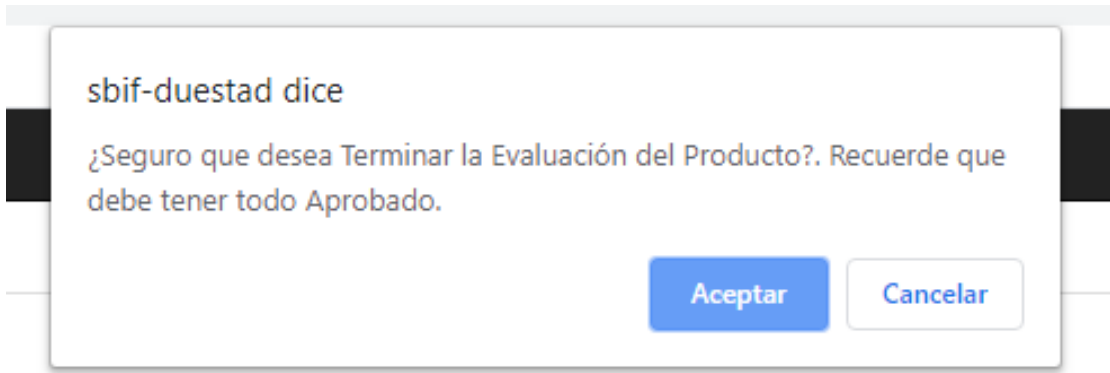


Figura 38 Término de Evaluación

Anexo C. Procedimiento para Tratar Series con Tendencia y Estacionalidad

Para la obtención de las medidas para las series de tipo estacional con tendencia se definen los siguientes pasos para quitar la tendencia:

1. Obtener los datos de la serie respecto al tramo de análisis (3 años).
2. Se obtienen los componentes de regresión lineal para en un paso siguiente estimar la tendencia. Donde “x” es el valor de cada dato de la serie, “y” es la posición “n” es total de datos.

a. $\sum x$

b. $\sum x^2$

c. $\sum y$

d. $\sum xy$

e. $\hat{\beta}_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$

f. $\hat{\beta}_0 = \frac{\sum y - \hat{\beta}_1 \sum x}{n}$

3. Se estima la tendencia de los 36 periodos a través de la siguiente formula, donde \hat{Y} es la tendencia.

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1$$

4. Se resta la tendencia:

$$\tilde{Y} = Y_{original} + \hat{Y}$$

5. Se estandarizan las cifras sin tendencia

$$Z = \frac{\tilde{Y} - \bar{\tilde{Y}}}{\tilde{S}}$$

6. Se identifican las observaciones más lejanas a ± 3 de la media o que sean mayores en 2 unidades a la observación siguiente, suavizándolas de la siguiente manera:

- a. Si la observación 1 es detectada, entonces se reemplaza por el promedio de las observaciones 2 y 3.
 - b. Si la observación está entre la posición 2 a la 35, entonces esta se reemplazará por el promedio de los valores en la posición anterior y la posición siguiente.
 - c. Si la observación está en la posición 36 se reemplaza por el promedio de los valores en la posición 34 y 35.
7. Los datos resultantes ya no presentan tendencia y sólo presentan estacionalidad, por lo cual se puede aplicar el algoritmo ya creado para series estacionales.

Anexo D. Pasos para Tratamiento de Series Tendencia Pura y Aleatorias

Para el caso de la obtención de series con Tendencia Pura y Aleatorias se aplican las mismas medidas, en el caso de la serie con Tendencia Pura se le quita la tendencia.

1. Se obtiene la media de los 36 periodos (Me) y la desviación estándar de los 36 periodos (Sd).
2. A cada valor de la serie se le resta la Me y se divide por la Sd de la siguiente manera:

$$z = \frac{x - Me}{Sd}$$

3. Medidas de análisis

a. Bandas con re-muestreo

- i. Se obtienen 200 muestras aleatorias de tamaño 10 (re-muestreo)
- ii. Se obtienen la media para cada muestra
- iii. Se obtiene la media ($m1$) y desviación estándar ($s1$) de la muestra en el punto ii.
- iv. Se crea una banda de la siguiente manera:

$$\text{Banda} = [m1 - 6 * d1 ; m1 + 6 * d1]$$

- v. Se evalúa el nuevo valor (observación 37 más actual) en la banda, si cae fuera de la banda es una alerta.

b. Banda intercuartíl

- i. Se obtiene el cuartil 1 ($Q1$) y cuartil 3 ($Q3$) y se obtiene su diferencia que corresponde al Rango Inter-cuartílico (RIC)
- ii. Se obtiene la banda de tolerancia de la siguiente manera:

$$\text{Banda} = [Q1 - 1.5 * RIC ; Q3 + 1.5 * RIC]$$

- iii. Se evalúa el nuevo valor (observación 37 más actual) en la banda, si cae fuera de la banda es una alerta.

c. Test de Chauvenet (36)

- i. Se obtiene la media ($m2$) y desviación estándar ($d2$) de la muestra de 36 periodos
- ii. Se obtiene el siguiente valor:

$$Valor = \frac{ABS(obs37 - m2)}{d2}$$

- iii. Ahora se evalúa este valor, y si es mayor a 2.46, entonces es alerta.

d. Test de Chauvenet (24)

- i. Se obtiene la media (m2) y desviación estándar (d2) de la muestra de 24 periodos
- ii. Se obtiene el siguiente valor:
- iii.

$$Valor = \frac{ABS(obs37 - m2)}{d2}$$

- iv. Ahora se evalúa este valor, y si es mayor a 2.46, entonces es alerta.

e. Test MAD (desviación de la mediana absoluta)

- i. Obtener la mediana de los 36 periodos (medi)
- ii. Para cada observación de los 36 periodos obtener la diferencia absoluta del valor respecto a la mediana.

$$Valor_i = ABS(obs_i - medi)$$

- iii. El MAD será la mediana de las diferencias en el punto ii.
- iv. Para el punto 37 (nueva observación) se obtiene:

$$Valor\ Crítico\ (VC) = \frac{ABS(obs37 - medi)}{MAD}$$

- v. Evaluar este VC, si es mayor a 3, es alerta.

f. Intervalo de Confianza (95%)

- i. Se obtiene la media (m3) y desviación estándar (d3) de los 36 periodos.
- ii. Se establece una significancia del 5%, por lo cual se obtiene el valor Z de la curva normal a 0.975 (97.5%), donde el valor Z es 1.95996.
- iii. Se obtiene el siguiente intervalo:

$$Banda = [m3 - 1.95996 * d3 ; m3 + 1.95996 * d3]$$

- iv. Se evalúa el nuevo valor (observación 37 más actual) en la banda, si cae fuera de la banda es una alerta.