



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Estudios de Postgrado

**MEJORA DE LA PRODUCTIVIDAD REALIZANDO SIMULACIÓN CON *DATA*
MINING EN PRONÓSTICOS DE VENTA**

Inga. Irma Jeannette Orozco García

Asesorado por el MA. Ing. Everest Darwin Medinilla Rodríguez

Guatemala, junio de 2015

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

MEJORA DE LA PRODUCTIVIDAD REALIZANDO SIMULACIÓN CON *DATA MINING* EN PRONÓSTICOS DE VENTA

PRESENTADO A LA ESCUELA DE ESTUDIOS DE POSTGRADO DE LA
FACULTAD DE INGENIERÍA

POR

INGA. IRMA JEANNETTE OROZCO GARCÍA

ASESORADO POR EL MA. ING. EVEREST DARWIN MEDINILLA RODRÍGUEZ

COMO REQUISITO PREVIO PARA OPTAR AL GRADO ACADÉMICO DE:

MAESTRO (*MAGISTER*) EN GESTIÓN INDUSTRIAL

GUATEMALA, JUNIO DE 2015

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Angel Roberto Sic García
VOCAL I	
VOCAL II	Ing. Pablo Christian de León Rodríguez
VOCAL III	Inga. Elvia Miriam Ruballos Samayoa
VOCAL IV	Br. Narda Lucía Pacay Barrientos
VOCAL V	Br. Walter Rafael Veliz Muñoz
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

**DIRECTOR DE LA ESCUELA DE ESTUDIOS DE POSTGRADO DE LA
FACULTAD DE INGENIERÍA**

MSc. Ing. Murphy Olympo Paiz Recinos

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

MEJORA DE LA PRODUCTIVIDAD REALIZANDO SIMULACIÓN CON *DATA* *MINING* EN PRONÓSTICOS DE VENTA

Tema que me fuera asignado por la Dirección de Escuela de Postgrado de la Facultad de Ingeniería, el 23 de enero de 2014.



Inga. Irma Jeannette Orozco García
abmj12@gmail.com
Carné: 88-12272



USAC
TRICENTENARIA
Universidad de San Carlos de Guatemala

Escuela de Estudios de Postgrado
Facultad de Ingeniería
Teléfono 2418-9142 / Ext. 86226



Ref. APT-2015-026

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Postgrado, al Trabajo de Tesis de la Maestría en Gestión Industrial titulado: **“MEJORA DE LA PRODUCTIVIDAD REALIZANDO SIMULACIÓN CON DATA MINING EN PRONÓSTICOS DE VENTA”**, presentado por la Ingeniera Industrial **Irma Jeannette Orozco García**, procede a la autorización para la impresión del mismo.

IMPRÍMASE.

A handwritten signature in black ink, appearing to read 'ARSIC'.

Ing. Ángel Roberto Sic García
DECANO

Guatemala, Junio de 2015.



USAC
TRICENTENARIA
Universidad de San Carlos de Guatemala



Escuela de Estudios de Postgrado
Facultad de Ingeniería
Teléfono 2418-9142 / Ext. 86226

APT-2015-026

El Director de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen y dar el visto bueno del revisor y la aprobación del área de Lingüística del trabajo de graduación titulado **"MEJORA DE LA PRODUCTIVIDAD REALIZANDO SIMULACIÓN CON DATA MINING EN PRONÓSTICOS DE VENTA"** presentado por la Ingeniera Industrial **Irma Jeannette Orozco García**, apruebo el presente y recomiendo la autorización del mismo.

"ID Y ENSEÑAD A TODOS"

MSc. Ing. *Murphy Olympo Paiz Recinos*
Director

Escuela de Estudios de Postgrado



Guatemala, Junio de 2015.

Cc: archivo
/la



USAC
TRICENTENARIA
Universidad de San Carlos de Guatemala

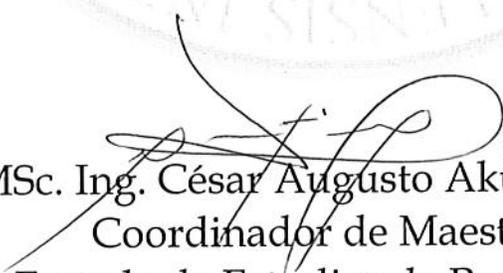


Escuela de Estudios de Postgrado
Facultad de Ingeniería
Teléfono 2418-9142 / Ext. 86226

APT-2015-026

Como Coordinador de la Maestría en Gestión Industrial y revisor del Trabajo de Tesis titulado **"MEJORA DE LA PRODUCTIVIDAD REALIZANDO SIMULACIÓN CON DATA MINING EN PRONÓSTICOS DE VENTA"**, presentado por la Ingeniera Industrial **Irma Jeannette Orozco García**, apruebo y recomiendo la autorización del mismo.

"ID Y ENSEÑAD A TODOS"


MSc. Ing. César Augusto Akú Castillo
Coordinador de Maestría
Escuela de Estudios de Postgrado



Guatemala, Junio de 2015.

Cc: archivo
/la



USAC
TRICENTENARIA
Universidad de San Carlos de Guatemala



Escuela de Estudios de Postgrado
Facultad de Ingeniería
Teléfono 2418-9142 / Ext. 86226

APT-2015-026

Como Revisor de la Maestría en Gestión Industrial del Trabajo de Tesis titulado **"MEJORA DE LA PRODUCTIVIDAD REALIZANDO SIMULACIÓN CON DATA MINING EN PRONÓSTICOS DE VENTA"**. Presentado por la Ingeniera Industrial **Irma Jeannette Orozco García**, apruebo el presente y recomiendo la autorización del mismo.

"ID Y ENSEÑAD A TODOS"

MSc. Inga. Rosa Amarilis Dubon Mazariegos

Revisor(a)

Escuela de Estudios de Postgrado



Guatemala, Junio de 2015.

Cc: archivo
/la

ACTO QUE DEDICO A:

- Dios** Por permitirme cumplir mis objetivos y bendecirme constantemente todos los días de mi vida.
- Mi madre** Por todo tu amor, por enseñarme que con esfuerzo y dedicación puedo llegar tan lejos como me lo proponga.
- Mi padre** Por tu amor y sobre todo por tu apoyo en mi desarrollo personal y profesional.
- Mis hijas** Por todo su amor, porque este logro es de las dos.

AGRADECIMIENTOS A:

La Universidad de San Carlos de Guatemala	Por ser mi casa de estudios, porque en esta institución tuve la oportunidad de formarme como toda una profesional.
Escuela de Postgrado	Por brindarme conocimientos y experiencia que permiten desarrollarme como una persona de éxito.
Mis padres	Por ser mi apoyo incondicional, por su amor y paciencia
Mis hijas	Por su apoyo, por estar juntas, compartir en familia los buenos y malos momentos.
Mi asesor	Por compartir sus conocimientos y colaborar al cumplimiento de este trabajo de graduación.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES

PLANTEAMIENTO DEL PROBLEMA

OBJETIVOS

MARCO METODOLÓGICO

RESUMEN.....	I
INTRODUCCIÓN	III

1. ANTECEDENTES	1
1.1. Marco teórico	3
1.1.1. Demanda	3
1.1.1.1. Importancia de los pronósticos	3
1.1.1.2. Satisfacción del cliente	5
1.1.1.3. Gestión de procesos	6
1.1.1.4. Ciclo de producción	7
1.1.2. Inventarios	7
1.1.2.1. Indicadores de inventarios	8
1.1.2.2. Gestión de inventarios	9
1.1.3. Cadena de valor	11
1.1.4. Productividad	12
1.1.5. <i>Data mining</i>	13
1.1.5.1. Bases de datos	15
1.1.5.2. Metodología de <i>Data mining</i>	16
1.1.5.3. Aplicaciones de <i>Oracle Data Mining</i> ..	18
1.1.5.4. Funciones de <i>Oracle Data Mining</i>	18
1.1.5.5. Pruebas de modelo supervisadas	20

	1.1.5.6.	Técnicas de <i>Data mining</i>	21
	1.1.5.7.	Preparación y análisis de datos	21
2.		DIAGNÓSTICO SITUACIONAL	23
	2.1.	Misión y visión	23
	2.2.	Productos	23
	2.3.	Clientes y alcance	24
	2.4.	Estructura organizacional	25
	2.5.	Identificación de problemas	28
		2.5.1. Flujo de información	28
		2.5.2. Planificación de compras.....	28
		2.5.3. Planificación de producción.....	29
	2.6.	Análisis de cadena de valor	29
		2.6.1. Proyección de ventas, compras y producción	30
		2.6.2. Compras de MP e insumos	32
		2.6.3. Abastecimiento de MP e insumos	34
		2.6.4. Producción	37
		2.6.5. Gestión del producto terminado	41
		2.6.6. Ventas	44
		2.6.7. Distribución.....	46
		2.6.8. Satisfacción al cliente como punto central de la cadena de valor	50
3.		SIMULACIÓN CON <i>DATA MINING</i>	53
	3.1.	Comprensión de datos	54
	3.2.	Fase 1. Preparación de datos	60
	3.3.	Fase 2. Construcción y evaluación de modelos	65
4.		EVALUACIÓN DE SIMULACIÓN	71

5.	PRESENTACIÓN DE RESULTADOS DE SIMULACIÓN.....	77
6.	DISCUSIÓN DE RESULTADOS	87
6.1.	Aumento de la productividad	87
6.1.1.	Proyección de ventas, compras y producción.....	87
6.1.2.	Compras de MP e insumos.....	88
6.1.3.	Abastecimiento de MP e insumos.....	89
6.1.4.	Producción.....	89
6.1.5.	Gestión del producto terminado.....	90
6.1.6.	Ventas	91
6.1.7.	Distribución.....	91
6.1.8.	Satisfacción del cliente	92
6.2.	Ventajas de herramienta ODM	92
6.3.	Beneficios o bondades de la implementación de ODM	93
	CONCLUSIONES	97
	RECOMENDACIONES.....	99
	BIBLIOGRAFÍA Y REFERENCIAS	101
	APÉNDICE.....	105
	Figura a. Creación del proyecto de simulación	105
	Figura b. Tabla que contiene datos preparados para la simulación.....	106
	Figura c. Datos preparados para la simulación.....	107
	Figura d. Agrupación de datos por id_producto	108
	Figura e. Creación de nodos para explorar los datos	109
	Figura f. Nodo de la creación de regresión	110
	Figura g. Creación de la división de prueba 80-20.....	111
	Figura h. Comparación de rendimiento para determinar el modelo	112
	Figura i. Comparación residual para determinar el mejor modelo.....	113
	Figura j. Detalles del modelo GLM	114
	Figura k. Coeficientes del modelo GLM	115

Figura l. Coeficientes del modelo GLM.....	116
Figura m. Columnas de salida del modelo GLM.....	117
Figura n. Flujo de trabajo con nodos de simulación del modelo GLM .	118
Figura o. Resultados de simulación del modelo GLM.....	119
Glosario	120

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Metodología de simulación	
2.	Sub funciones del sistema de abastecimiento	10
3.	Optimización de la cadena de valor	12
4.	Fases del proceso de ODM	17
5.	Preparación de datos	22
6.	Organigrama de la planta productora	26
7.	Cadena de valor de planta productora	29
8.	Análisis FODA proyección de ventas, compras y producción	31
9.	Análisis FODA compras	34
10.	Análisis FODA abastecimiento de MP e insumos	37
11.	Análisis FODA del área de producción	41
12.	Análisis FODA de gestión del producto terminado	43
13.	Análisis FODA de ventas	46
14.	Análisis FODA de distribución	49
15.	Análisis FODA del servicio al cliente	50
16.	Diagrama de flujo de la preparación de datos	61
17.	Construcción y evaluación de modelos	66
18.	Comparación de pruebas de simulación	68
19.	Datos de simulacion <i>versus</i> venta real Crema rizos 6 onzas	80
20.	Datos de simulacion <i>versus</i> venta real Crema rizos 8 onzas	82
21.	Datos de simulacion <i>versus</i> venta real Lágrimas 2 onzas	84

TABLAS

I.	Tipos de productos y presentación	
II.	Variables e indicadores	
III.	Tablas de datos	54
IV.	Datos de facturación	55
V.	Detalle de facturas	56
VI.	Inventario de productos	57
VII.	Datos de clientes	58
VIII.	Vendedores	59
IX.	Presentación del producto	59
X.	Departamento del cliente	60
XI.	Campos de archivo .xml	62
XII.	Fuentes preparadas para minar	63
XIII.	Datos preparados para simulación	64
XIV.	Coeficientes de prueba de simulación	69
XV.	Campos de tabla resultados_generales	71
XVI.	Detalle de simulación GLM	72
XVII.	Detalle de indicadores de regresión GLM	73
XVIII.	Columnas de datos de simulación	78
XIX.	Datos de simulación <i>versus</i> venta real Crema rizos 6 onzas	79
XX.	Datos de simulación <i>versus</i> venta real Crema rizos 8 onzas	81
XXI.	Datos de simulación <i>versus</i> venta real Lágrimas 2 onzas	83

PLANTEAMIENTO DEL PROBLEMA

- **Descripción del problema**

En el 2013, las ventas en la planta productora de productos para el cabello disminuyeron un 10 % anual (datos proporcionados en forma confidencial), el incumplimiento de la demanda del cliente afecta la operación de la empresa, puesto que se considera al cliente como el centro de la cadena de valor. Es evidente que no se está cumpliendo con la demanda, por lo que es necesario tomar medidas inmediatas para encontrar el pronóstico de ventas que ayude a planificar la producción.

La materia prima utilizada en la producción es importada y el desabastecimiento de esta causa paradas en la producción, puesto que el lapso de espera para ser liberado es de 2 a 3 meses y en algunas ocasiones la compra de materia prima local aumenta los costos de operación. Todos estos aspectos van amarrados con la falta de un pronóstico adecuado de ventas, el cual es importante para la programación de la producción, debe cubrir la demanda del consumidor final. El no tener productos presentes en los canales de distribución ocasiona que el consumidor final sustituya el producto por uno de la competencia, provocando la pérdida de fidelización en la marca.

Por estas razones, se debe buscar el mejor método de pronóstico de ventas. En la actualidad, el uso de herramientas tecnológicas como *Data Mining* contribuyen a disminuir el tiempo de implementación y exactitud de resultados.

- **Formulación del problema**

1. ¿Cuál es el mejor método y algoritmo de pronóstico de ventas que ayude a mejorar la productividad a través de la simulación con *Data Mining*?
2. ¿Cuál es la situación actual y el modelo de la empresa?
3. ¿Cuáles son los datos que ayudarán a determinar el nivel óptimo de ventas a través de la simulación?
4. ¿Qué cantidad de datos serán usados para que sea confiable el resultado de la simulación?

- **Delimitación**

La investigación se realizará en una pyme dedicada a la producción, distribución y comercialización de productos para el cuidado del cabello. Durante un lapso de 6 meses, tiempo en el cual se realizará la simulación con ODM (*Oracle Data Mining*).

OBJETIVOS

General:

Determinar el mejor método de pronóstico de ventas que ayude a mejorar la productividad, a través de la simulación con *Data Mining*.

Específicos:

1. Determinar la situación del negocio, con las fortalezas, debilidades, amenazas y oportunidades.
2. Identificar los datos que ayudarán a determinar el nivel de ventas por medio de la simulación con *Data Mining*.
3. Verificar que la cantidad de datos usados y la preparación de los mismos sean confiables en la simulación con *Data Mining*.

MARCO METODOLÓGICO

- **Tipo de estudio y diseño de investigación**

La investigación es de tipo retrospectivo, porque se indaga sobre los hechos ocurridos en el pasado utilizando datos pasados de ventas útiles para simular el pronóstico con el mejor modelo. Los datos utilizados fueron de cinco años, del 2009 al 2013.

El tipo de investigación es formulativa. Se realizó una investigación aplicada cuyo objetivo era encontrar el mejor modelo a través de la aplicación de la herramienta tecnológica, el cual se recomienda sea implementado por la planta productora aumentando con esto la productividad.

Por la profundidad de las variables, fue una investigación analítica.

- **Técnicas de investigación**

La técnica utilizada fue la de campo, a través del registro de datos históricos de ventas de todos los productos que fabrican y comercializan; de esta manera los pronósticos de ventas fueron más confiables. Se utilizó principalmente la base de datos de ventas del 2009 al 2013.

- **Tamaño de la muestra**

La simulación con *Data Mining* se realizó con todos los productos que produce la planta productora, realizando las combinaciones de producto,

número de presentaciones y cantidad de productos (véase tabla I). La muestra es de 27,125 registros.

Tabla I. **Tipos de producto y presentación**

Producto	Presentaciones								
	2 onzas	4 onzas	6 onzas	8 onzas	12 onzas	16 onzas	20 onzas	22 ml	44 ml
Crema rizos			X	X	X				
Crema para peinar			X						
Brisa para el cabello	X	X							
Lagrimas	X	X							
Gelatina para cabello			X						
Tratamientos para el cabello							X		
Ampollas para el cabello								X	X
Shampoo para cabello						X			
Acondicionador para cabello						X			

Fuente: elaboración propia.

- **Metodología de la simulación con *Data Mining***

La metodología propuesta por (Taft, Krishnan, Hornick, & Muhkin, 2005, p. 5-1) va de la mano con la metodología propuesta para determinar el mejor método de pronóstico de ventas.

Según (Taft *et al.*, 2005, p. 5-1) es importante tener presente que cada una de las fases utilizadas deben ser documentadas para recurrir a ellas si ocurriera algún error (véase figura 1). Estas se detallan a continuación:

1. **Comprensión del negocio:** comprender los objetivos y requisitos del estudio de investigación, convertir los datos como parte de la definición del problema.
2. **Comprensión de los datos:** inicialmente se deben recopilar los datos que cumplen con los objetivos los cuales nos ayudarán a encontrar el mejor método de pronóstico, familiarizarse con los datos identificando problemas de calidad de los mismos.
3. **Preparación de datos:** los cuales serán útiles para construir el conjunto final de datos a partir de los cuales se implementará la simulación.
4. **Modelado:** seleccionar el tipo de función a aplicar la cual puede ser supervisados y no supervisados. Los datos deben ser explorados con la herramienta. Para luego determinar que el modelado cumple con los objetivos de la investigación. Por lo general, existen varias técnicas para el mismo tipo de problema de *Data Mining*; en el caso al realizar la

simulación puede ocurrir que no cumplen con los objetivos establecidos en ese caso dar un paso atrás en la metodología es permitido.

5. Evaluación: evaluar el modelo implementado en la simulación revisando detalladamente las métricas o índices para asegurar que con la implementación de la simulación se logran los objetivos. En el caso alguna cuestión importante no ha sido considerada debe ser evaluada. Al finalizar esta fase se toma una decisión sobre el modelo que debe implementarse.
6. Presentación y discusión de Simulación: Después de determinar el mejor modelo, se debe implementar la simulación de pronóstico de ventas. Esta puede ser tan simple como generar un informe. A través de la implementación del pronóstico de ventas, se debe evaluar si se cumple con los objetivos.

Figura 1. Metodología de simulación



Fuente: elaboración propia.

Según (Taft *et al.*, 2005, p. 5-1) al ser *Data mining* una herramienta interactiva, después de ser implementada se le debe dar seguimiento a los resultados del pronóstico de ventas y solucionar problemas o situación que pueden ocurrir en los cálculos.

La técnica utilizada que proporcionó elementos clave para abordar los problemas que solucionan el problema de pronóstico de ventas, es el de Regresión basándose en los algoritmos estadísticos GLM (*Generalized Linear Models*) y SVM (*Support Vector Machine*).

- **Variables e indicadores de la simulación**

Los indicadores evaluados para determinar que el método simulado fue el mejor, se detallan en la tabla II.

Tabla II. **Variables e indicadores**

Variable	Indicadores
Coeficiente R^2	Valor muy cercano a 1, lo que indica que la línea de regresión se aproxima a las unidades reales vendidas por producto.
Error de media cuadrada	Valor muy cercano a 0 lo que indica exactitud en el modelo.
Bondad del modelo <i>F-value</i>	Valor alto lo que determina que el modelo es confiable.
GMSEP error estimado de la predicción	Valor muy cercano a 0 lo que determina que el modelo es confiable.

Fuente: elaboración propia.

RESUMEN

La aplicación de herramientas tecnológicas para optimizar los procesos ha crecido considerablemente, su implementación ha aportado beneficios a la industria como reducción de costos, aumento en la eficiencia, automatización de procesos, índices de patrones repetitivos y patrones de tendencia.

El trabajo de graduación detalla las fases y la metodología de la simulación con la herramienta ODM (*Oracle Data Mining*). La metodología utilizada dio como resultado el modelo que la planta productora debe implementar para obtener el mejor pronóstico de ventas y de esta manera mejorar la productividad. La simulación se aplicó a la línea completa de productos que produce y distribuye en el mercado nacional, en sus diversas presentaciones.

Se realizó la simulación con la implementación de tres fases:

- Fase 1. Adquisición de datos basada en datos históricos de ventas de 5 años.
- Fase 2. Construcción y evaluación de modelos de regresión GLM (*Generalized Linear Models*) y SVM (*Support Vector Machine*) simulando con 27 125 registros.
- Fase 3. Despliegue de resultados, evaluación y discusión; en esta fase se concluyó que el modelo que debe implementar la planta productora es GLM, obtuvo el mayor porcentaje de confianza, mejorando de esta manera la productividad de la planta productora.

INTRODUCCIÓN

El presente trabajo desarrolla un tema importante para toda empresa guatemalteca, el cumplimiento de la demanda a través de las ventas. Una industria que no tiene en sus manos el producto que el cliente necesita, provoca que se busquen productos sustitutos ocasionando pérdidas económicas para la empresa.

El cumplimiento de la demanda de las ventas puede optimizarse al producir con el mejor pronóstico de ventas, se puede lograr con mayor confiabilidad usando herramientas tecnológicas. Existe una gran cantidad de herramientas, entre las cuales se encuentra ODM (*Oracle Data Mining*), esta es usada en los negocios ayuda en el enfoque a clientes, fraudes, indicadores de desempeño y a encontrar información importante para la industria. Contribuye a las empresas en descubrir nueva información, construir aplicaciones que automaticen los procesos y a reducir costos.

La simulación se realizó con datos de una PYME dedicada a la producción, distribución y comercialización de insumos para el cabello, dado que durante el 2012 y el 2013 no se cumplió con la demanda.

Este estudio permitió a la planta productora encontrar el mejor modelo a través de la simulación con ODM, el cual se recomienda sea implementado para encontrar el mejor pronóstico de ventas; de esta manera se mejorará la productividad optimizando la cadena de valor desde la compra de materia prima, evitando almacenamiento innecesario, hasta la distribución adecuada en todos los canales de distribución.

La descripción de la situación actual de la planta productora, a través de la identificación de problemas en el flujo de información, la planificación de compras, de producción y de ventas ayudó a identificar el problema que se aborda en este trabajo de graduación. Se analizó la cadena de valor de la planta productora, la importancia desde el procedimiento de la selección del proveedor que influye en la entrada de materia prima en el tiempo exacto para cubrir la orden de producción; el procedimiento de ventas y el proceso de distribución para cubrir todos los canales de distribución y la importancia del cliente como centro de la cadena de valor.

Las fases realizadas en la simulación obtuvieron los siguientes resultados:

Fase 1. Adquisición de datos, se utilizó la base histórica de ventas de 5 años, que sirvió para preparar los datos. Esta información se almacenó en la tabla *Simulación*, los atributos utilizados fueron *cus_id*: identificador único por cada fila; *id_producto*: identificador por producto; *descripción_producto*: detalle del producto facturado; *id_presentacion*: identifica la presentación del producto; *fecha_factura*, *costo*, *precio_venta*, *docenas vendidas* y *unidades_vendidas*.

En la fase 2, se realizó la simulación del modelo de regresión con la tabla *Simulación*, corriendo los algoritmos SVM (*Support Vector Machine*) y GLM (*Generalized Linear Models*) con la muestra completa 27,125 registros, para efectos de prueba se dividió la tabla en 80 % para simular y 20 % para prueba.

En la fase 3, los resultados muestran que el modelo supervisado de regresión es el que la planta productora debe implementar, pues estos modelos se basan en predecir un valor objetivo específico. El modelo GLM (*Generalized Linear Models*) con porcentaje de confianza de 99.8627 % es el más alto. La simulación con este modelo de regresión proporcionó valores en

los indicadores, los cuales confirman que es un modelo y algoritmo confiable. La implementación de GLM resultó en el pronóstico de ventas de un año obteniendo como resultado la mejora de la productividad

1. ANTECEDENTES

Los países de América Latina y el Caribe han registrado avances en los indicadores económicos y sociales en los últimos 15 años. El crecimiento se ha incrementado en un 1.9 %, pero si se relaciona con otras economías Guatemala se ha quedado muy atrás.

El ingreso bajo disminuye la productividad y el crecimiento económico. Toda empresa busca crecimiento, es el objetivo de su existir. La planta productora que se estudiará registra en sus libros disminución en ventas de un 10 % por año; por lo que se realizó una evaluación del rendimiento de los productos con clientes mayoristas, así como la medición de la satisfacción del cliente y razón por la cual las compras han disminuido.

Los productos que distribuye la planta productora son fáciles de sustituirse por otros que sí se encuentran en góndolas de supermercados, por lo que es importante cubrir la demanda, entregar pedidos a los clientes mayoristas a tiempo; pero esto muchas veces no es posible porque en las bodegas de producto terminado no hay producto, la planificación de la producción no cubre la demanda.

Según (Torre, 1999, p. 14), la manera más usual de iniciar un plan de producción basado en la demanda del producto es elaborar un pronóstico. El pronóstico de la demanda se basa en la información referente al pasado. Por tal razón la planta productora debe tomar medidas inmediatas, identificar las variables que determinan las ventas e implementar un sistema que ayude a mejorar la productividad.

A nivel mundial se han usado herramientas tecnológicas para facilitar el trabajo del personal de producción y planificación, puesto que al determinar el pronóstico que es una predicción de lo que se cree sucederá en un futuro con las ventas, ayudará a mejorar la productividad.

Los procesos productivos generan información y se quedan almacenados sin darle el uso adecuado, la aplicación de *Data Mining* ayuda a explorar información identificando comportamientos escondidos. “El *Data Mining* abre una nueva dimensión en el diseño de estudios de mercado, levantamiento de datos y análisis” (Palma, Palma, & Pérez, 2009, p. 24).

La enorme cantidad de datos históricos, sirven de base para determinar la demanda, la cantidad de datos ocasionan un problema por lo que *Data Mining* proporciona una forma de obtener la información enterrada en los datos, crea modelos para encontrar patrones ocultos a través de métodos estadísticos.

“Actualmente nos encontramos en un mundo muy competitivo con una economía globalizada y ampliamente interconectada los flujos de información se han potenciado, dentro de la empresa y de fuentes externas. Los sistemas de información juegan un rol muy relevante “(Gonzalez R. , 2012, p. 1).

Data Mining es la herramienta que usará la empresa para conocer la demanda de los clientes y de esta manera reducir costos, disminuir deterioro en el almacenaje de productos y optimizar la cadena de valor. Esta herramienta permitirá mejorar las instancias de producción, reduciendo la posibilidad de fallas.

1.1. Marco teórico

La investigación incluye el conjunto de conocimientos sobre la demanda, inventarios, cadena de valor, productividad y *Data Mining*, que permiten una guía de referencia para el desarrollo de la simulación de ventas con la herramienta *Oracle Data Mining*.

1.1.1. Demanda

“Es la cantidad que están dispuestos a comprar de un producto un conjunto determinado de consumidores en un lugar y período de tiempo predeterminado, y en una situación del entorno y tras un esfuerzo comercial de la empresa y de los competidores” (Ruiz, Fuentes & Alcaraz, 2006, p. 42), el cumplimiento de la demanda es vital para la satisfacción del cliente

1.1.1.1. Importancia de los pronósticos

“Un buen pronóstico es de importancia crucial para todos los aspectos del negocio, el pronóstico es la única estimación de la demanda hasta que se conoce la demanda real” (Heizer & Render, 2004, p. 105)

Los pronósticos afectan la cadena de valor desde el ingreso de materia prima o insumos hasta que el producto o servicio están en las manos del cliente final.

La falta de cumplimiento en la demanda o ventas es un indicativo que los pronósticos no se calculan adecuadamente. Como consecuencia los siguientes objetivos no se cumplen:

- Fijar metas de ventas
- Cumplir con las metas de ventas
- Presupuesto de ventas

La frecuencia en el cálculo de los pronósticos se pueden clasificar por el periodo de tiempo en:

- Semanal: lo que nos ayuda a validar el plan de producción.
- Mensual: este cálculo realiza una proyección ajustada a los cambios que se puedan dar en el mercado.
- Anual: estos permiten proyectar las ventas del año siguiente, cuya finalidad es:
 - Financiera para cualificar flujos de efectivo al momento de importar materia prima.
 - Producción: para determinar la capacidad de la planta.
 - Comercial: compromiso en el cumplimiento de objetivos de ventas.

Para elaborar los pronósticos se pueden usar datos como:

- “información referente al pasado” (Torre, 1999, p. 14)
- Factores de ajuste
- Factores externos de influencia económica
- Factores internos de la administración de la empresa

El uso de pronósticos se da con menos frecuencia en empresas pequeñas, aunque no todos los pronósticos deben tener una orientación cuantitativa también pueden ser cualitativos. Las pymes conocen muy poco sobre el proceso de elaboración de pronósticos.

Según (Heizer & Render, 2004, p. 106) existen siete pasos para pronosticar:

- Objetivo del pronóstico
- Aspectos a pronosticar
- Determinar horizonte del pronóstico
- Seleccionar modelo
- Reunir datos para elaborar pronóstico
- Obtener datos
- Validar e implantar resultados

1.1.1.2. Satisfacción del cliente

Según (Velaso, 2010, p. 133), la gestión de la satisfacción del cliente se mide a través de satisfacer su auténtica necesidad, la percepción de ejecución y de sus expectativas. La satisfacción del cliente tiene varios componentes de calidad en el servicio, calidad del producto, precio, tiempo de espera.

La evaluación del cliente depende del cumplimiento de sus necesidades y expectativas, la insatisfacción estadísticamente causa mala referencia escuchamos con frecuencia según (Dutka, 1998, p. 16) que cien clientes satisfechos produce 25 nuevos clientes, por cada queja existen 20 clientes que opinan lo mismo.

“La medición de la satisfacción del cliente puede y debe ser vista como un proceso, es decir como una secuencia continua de actividades que tiene un producto” (Perez J. A., 1999, p. 205) no se puede aislar es necesario verla como un todo.

Generalmente se considera que todo termina con cumplir las expectativas del cliente, y no se toma en cuenta que “Complacer las expectativas del cliente trae como consecuencia la satisfacción y superarlas genera una fuerte lealtad de parte de ese cliente” (Dutka, 1998, p. 55), la deslealtad del cliente por productos sustitutos afecta la estabilidad de cualquier empresa.

Según (Zeithmal, 1990, p. 11), el servicio puede hacer la diferencia, puede ser útil para incrementar la productividad; este indicador es vital para los beneficios futuros. El modelo creado por Ziethmal ayuda a mejorar la calidad de servicio basándose en brechas de satisfacción o insatisfacción del cliente.

1.1.1.3. Gestión de procesos

"La gestión por procesos hace compatibles las necesidades organizativas internas con la satisfacción de los clientes" (Velaso, 2010, p. 15). A través de la maximización de la productividad generando máximo valor con menor cantidad de recursos.

"La finalidad última de la gestión de procesos es centrarse en lo crítico para generar valor, ahora y en el futuro" (Velaso, 2010, p. 15), planificando, ejecutando y controlando actividades desde el desarrollo de productos hasta la distribución del producto en los canales de distribución. Estas actividades son importantes para el buen desempeño del negocio asegurando calidad.

La maximización de la productividad se logra gestionando según (Velaso, 2010, p. 148), los procesos incorporando actividades de medición, análisis y mejora para la empresa es la generación del máximo valor con la menor cantidad de recursos y para el cliente o consumidor final es obtener el máximo beneficio del producto que está adquiriendo con la menor inversión.

1.1.1.4. Ciclo de producción

Según (Gonzalez M. , 2006, p. 109), para conocer el ciclo de producción, es necesario dividirlo en

- Función de repetición: tareas esporádicas, tareas regulares.
- Función del recurso que las realiza: tareas hechas por el hombre, tareas hechas por la máquina.
- En función de su duración: tareas constantes, tareas variables.

Las tareas que forman parte del ciclo productivo son las correspondientes a máquina parada y tiempo máquina.

1.1.2. Inventarios

Importantes para toda empresa, “Los inventarios son aquellos productos, independientemente de su finalidad o grado de acabos, utilizados u obtenidos en los procesos productivos” (Gonzalez M. , 2006, p. 70). La escases de los inventarios ocasionan pérdidas a las empresas, esta escases puede ser ocasionada por faltante de materia prima, envases, suministros, maquinaria averiada.

Según (Gonzalez M. , 2006, p. 70) los inventarios sirven como estabilizadores entre producción y demanda, es decir entre la capacidad productiva y requerimientos del cliente, cuando la empresa sigue una estrategia de mantenimiento de inventarios y no otra que prescindiera de ellos por ejemplo, el *just in time*.

El mantenimiento alto de *stock* en bodegas beneficia a las empresas en cumplimiento de entrega a cliente, aprovechamiento de economía de escala, cambios en comportamientos de demanda. Pero cuando el *stock* sobrepasa puede ocasionar costos de almacenamiento, desperdicio, fecha de vencimiento en productos con vigencia.

1.1.2.1. Indicadores de inventarios

En las empresas es necesario determinar y conocer métricas, las más utilizadas son los de volumen o cantidad, número de *ítems* almacenados, valor de existencias, días de venta almacenados, desempeño de inventario, porcentaje de venta que se pierde por falta de producto.

Según (Ramírez, 2006, p. 40), la rotación de inventarios es un indicador que ayuda a conocer el número de veces que la empresa cambia sus inventarios, el exceso o insuficiencia de inventarios incurre en costos que son perjudiciales para las finanzas de toda empresa.

El indicador de rotación de activos según (Ramírez, 2006, p. 42), mide la eficiencia de los activos para generar ventas, este ayuda a saber si las inversiones han tenido los resultados esperados.

Los indicadores son establecidos por cada empresa, con base a los factores que afectan la toma de decisiones del *stock* de productos en bodegas de producto terminado.

1.1.2.2. Gestión de inventarios

La administración de los inventarios, es un gran reto referente a la planificación y el control, a los inventarios se le considera como un amortiguador o colchón entre dos sistemas de oferta refiriéndonos a producción y el otro de demanda que son nuestros clientes.

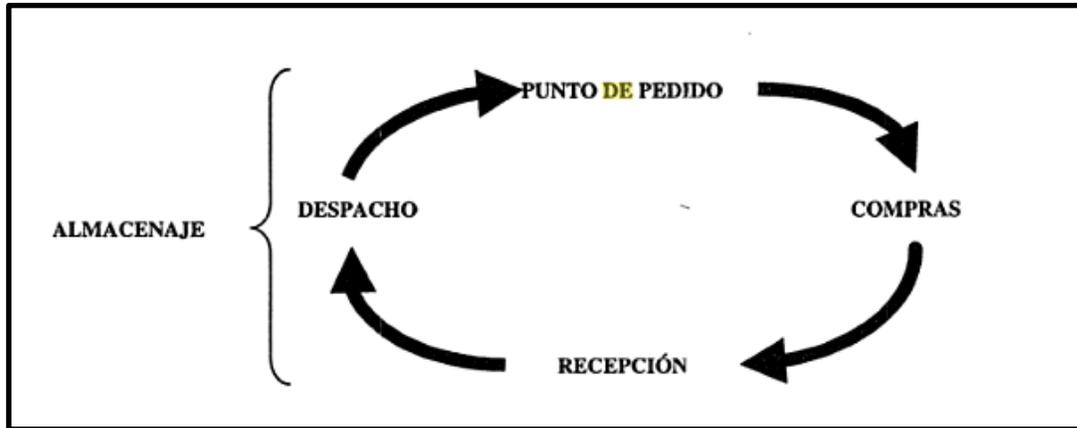
La mala gestión de inventarios ocasiona errores de planeación, desbalance en los inventarios, alta complejidad en la programación de la producción e inadecuada utilización de recursos de producción. Lo que puede ser atenuado al satisfacer la demanda del cliente.

Según (Heizer & Render, 2004, p. 452) la buena administración del inventario es crucial, dado que una empresa reduce sus costos disminuyendo el inventario, por otro lado, la falta de un artículo interrumpiría la producción y los clientes quedarían insatisfechos cuando hay faltantes.

“El inventario de seguridad es una protección contra la incertidumbre de la demanda, del tiempo de entrega y de suministros, son convenientes cuando los proveedores no entregan la cantidad deseada, en fecha convenida y con calidad aceptable” (Krajewski, Rizman, & Larry, 2000, p. 548) también dependerá de la variabilidad de la demanda.

En la figura 2, según (Boland, Carro, Stancatti, & Gismano, 2007, p. 126) el sistema de abastecimiento comprende tres sub funciones: la gestión de compras, recepción y almacenaje-administración de *stock*.

Figura 2. **Sub funciones del sistema de abastecimiento**



Fuente: Boland, L (2007). *Funciones de Administración* (p.126)

La optimización de los niveles de inventario es importante, dado que los costos de inventario son costos altos de muchas empresas.

Existen estrategias para lograr un buen balance de inventarios

- Para elegir el modelo matemático óptimo se deben seleccionar datos fiables para el cálculo de la demanda, la implementación de sistemas adecuados de pronósticos.
- Sistema de control de inventarios paralelos a los pronósticos.
- Ciclo de vida del producto, analizando los indicadores de eficiencia como nivel de servicio al cliente, grado de obsolescencia del inventario, indicadores financieros, rotación de inventario.

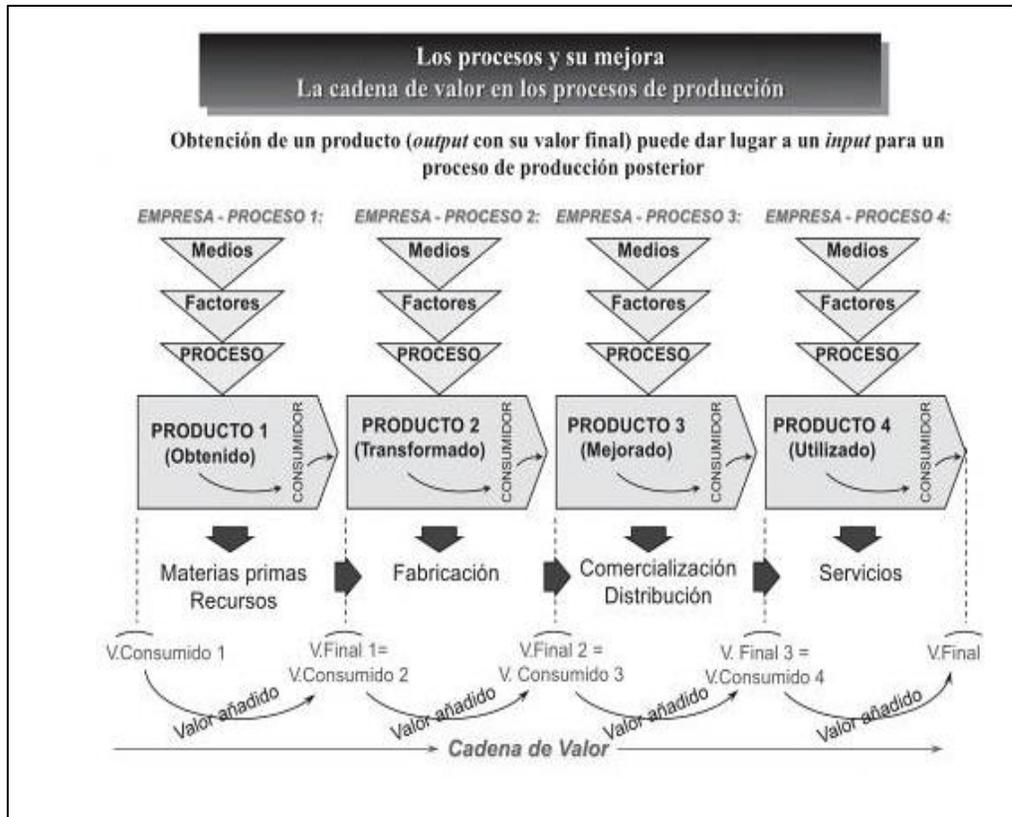
1.1.3. Cadena de valor

“La producción de cualquier bien o servicio se halla normalmente encadenado a la producción de otros, a través de la llamada cadena de valor (véase figura 3)” (Arbós, 2011, p. 14)

Según (Arbós, 2011, p. 14), el análisis de la cadena de valor, ayuda a optimizar el proceso productivo, puesto que se aprecia con detalle el funcionamiento; el objetivo primordial de toda empresa es reducir los costos y buscar la eficiencia al aprovechar todos los recursos. El estudio de la cadena de valor logra una ventaja estratégica.

“La optimización de la cadena de valor, se realiza tomando en cuenta desde el proveedor, el proceso de planificación y control en las actividades de abastecimiento, producción, logística y venta hasta el cliente. Con el auxilio de la informática y los recursos humanos.” (Singer & Donoso, 2007, p. 525)

Figura 3. Optimización de la cadena de valor



Fuente: Arbós, L. (2009) *Organización de la producción y dirección de operaciones*, (p. 15)

1.1.4. Productividad

“En un mundo como el actual, si queremos seguir siendo competitivos es imprescindible: mejorar de forma continua nuestra productividad, gestionar adecuadamente el conocimiento y estar adaptado a un mundo de constante cambio” (Fernández, 2010, p. 9), la importancia del mejoramiento de la productividad perfecciona los procesos y la operación de toda empresa, por lo que es un tema sumamente importante a implementar.

La mejora de la productividad puede gestionarse gracias a avances tecnológicos, según (Fernández, 2010, p. 66), las ventajas que conlleva no pueden lograrse plenamente si estas no cuentan con una fuerza de trabajo capaz de aprovechar sus potenciales. Por lo que es importante el talento humano, la disposición, el trabajo en equipo.

La reducción de costos, gastos de operación, eliminación de desperdicios, aumento en ventas, aumento en producción generan una mejora en la productividad. La cadena de valor debe ser gestionada con eficiencia para lograr el mejoramiento de la productividad.

1.1.5. *Data Mining*

“*Data Mining* se define como un proceso de encontrar nuevas y significativas relaciones al examinar grandes cantidades de datos” (López C.P, 2007, p. 1) proporciona una forma de obtener información enterrada en los datos, al crear modelos encuentra patrones ocultos que estadísticamente se escapan del análisis.

Según (López, 2007 C.P, p. 2), las herramientas de *Data Mining* permiten la extracción de patrones, comportamientos y regularidades las cuales describen información para predecir comportamientos futuros. La gran cantidad de datos deben tratarse antes y después de implementar algún modelo.

Los datos se almacenan en objetos (tablas) de base de datos, la preparación de datos puede llevarse a cabo utilizando las instalaciones de la base de datos.

“Con la información de las organizaciones y la aparición de aplicaciones de software operacionales sobre el sistema de información, brindan soporte a los procesos básicos de la organización (ventas, producción, personal” (López V.R, 2007, p. 1). Los modelos de *Data Mining* basados en un software operacional amigable han sido probados, validados, gestionados y desplegados en su entorno de dominio de aplicación adecuada.

Hacer todo el trabajo de la *Data Mining* de proceso de una manera reproducible y confiable es desafío, dado que puede implicar la automatización y las transferencias a través de servidores, bases de datos, aplicaciones, y herramientas. Por ejemplo, algunas herramientas de *Data Mining* requieren que los datos sean exportados de la base de datos corporativa y se convierte en el formato de la herramienta.

Las herramientas de ODM permiten extraer patrones, tendencias y regularidades para comprender los datos y predecir los comportamientos futuros; el almacén de los datos es el sistema central en todo el proceso orientada a integrar y ayudar en la toma de decisiones. Estos datos son históricos, internos, externos y descriptivos que están integrados para que permitan aplicar eficientemente las herramientas y analizar los datos para la toma de decisiones.

Según (Braga, 2009, p. 11) al aplicar *Data Mining* dirigida al cliente provee el conocimiento de las características y comportamiento de los mismos; la estimativa del ciclo de vida de un cliente es de alto retorno para la empresa. Las proyecciones arraigadas en el histórico real del negocio es parte de la metodología de *Data Mining* el conjunto de técnicas para describir y predecir a partir de grandes masas de datos por lo que *Data Mining* está asociada a

bases de datos, las cuales permiten la integración rápida de datos de diferentes fuentes.

“*Data Mining* es un tema práctico y consiste en aprender de una manera práctica, las técnicas para encontrar y describir patrones, es una herramienta para ayudar a explicar las predicciones de datos.” (Witte & Frank, 2005, p. 42) la información que maneja *Data Mining* proviene de tablas o vistas, que es una ventaja porque elimina la transferencia en herramientas independientes.

La extracción de información puede ser a partir de hojas electrónicas, lo que genera bondad en la herramienta

1.1.5.1. Bases de datos

La operación de empresas pequeñas y grandes gestionan información de clientes, productos, regiones, producción, ventas. Según (Coronel, Morris, & Rob, 2011, p. 5) las bases de datos acceden a sistemas para guardar, manipular y extraer datos rápidamente dado que son estructuras especializadas. Esta herramienta tecnológica ayuda a centralizar información con el propósito de asegurar la consistencia de datos

Es básicamente un sistema computarizado que lleva registros, información, es un contenedor de una colección de archivos de datos. La finalidad de estos es recuperar, actualizar información con base en peticiones.

1.1.5.2. Metodología de *Data Mining*

Según (Taft *et al.*, p. 5-1) las metodologías más usadas son: CRISP DM (*Cross Industry Standard Process for Data Mining*), SEMMA (*Sample, Explore, Modify, Model, Assess*) y Metodología de Berry y Linnoff. Fue creada en 1996 con gran éxito desarrollada con base a casos e implementaciones de proyectos de *Data Mining*. La metodología es amigable, plantea ideas que deben parametrizarse para cada entorno de ejecución.

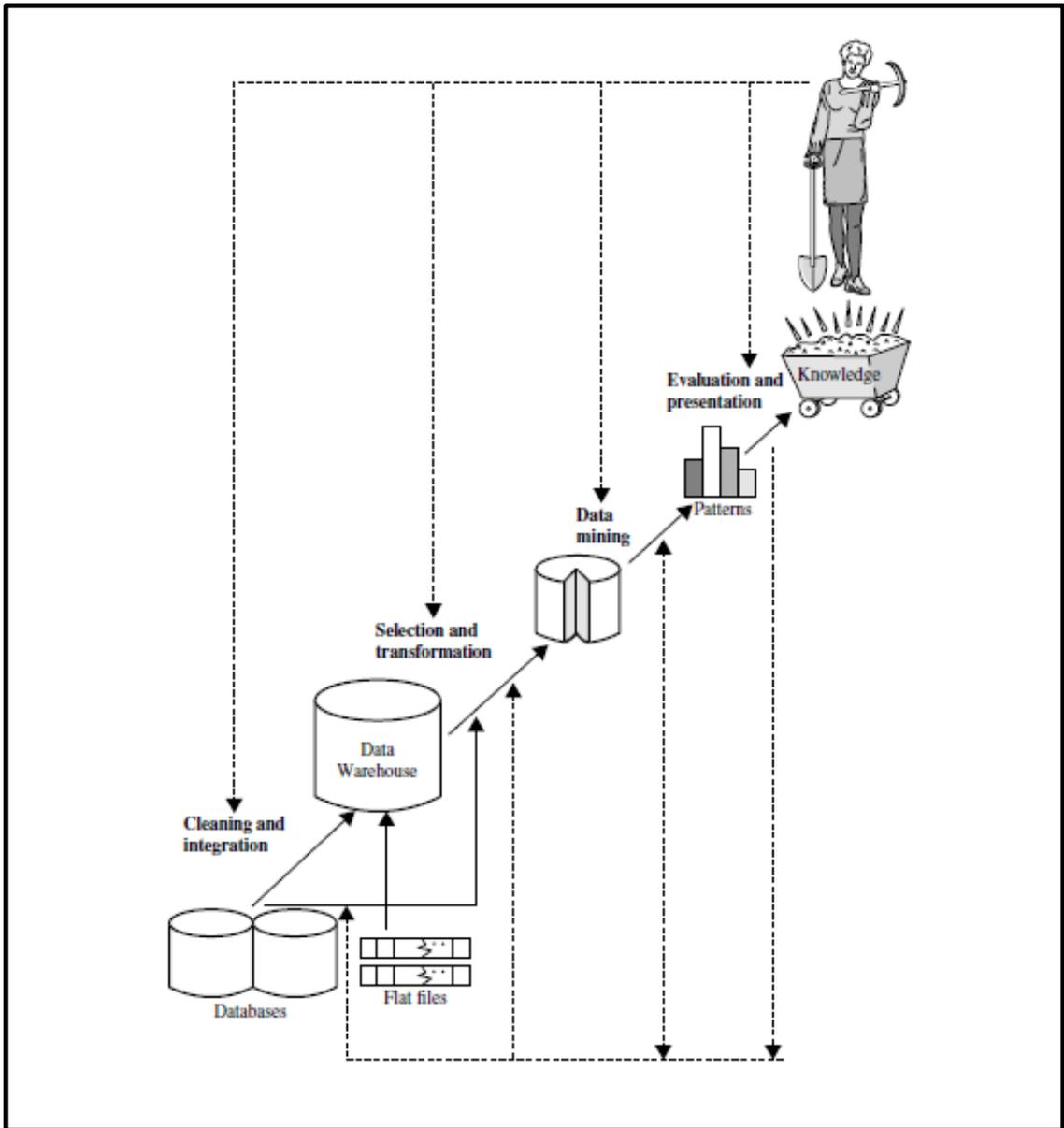
Las fases de la metodología (véase figura 4) generaliza la preparación de datos, análisis de datos, evaluación e interpretación de datos y la difusión del uso de modelos. Los detalles de cada fase son determinados por el ente que implementará esta herramienta.

La preparación de datos selecciona, explora, limpia y transforma la data que se utilizará en la simulación. La recopilación e integración de datos es relevante para una buena selección; la exploración de los datos describe la simetría, normalidad de datos; la limpieza elimina datos incorrectos y finalmente la transformación o desnormalización de datos permite tener datos confiables

El análisis de datos se realiza a través de la implementación de modelos supervisados o no supervisados. Al definir el modelo a simular se determina el más confiable con base a datos estadísticos.

La evaluación e interpretación de los datos nos permitirá concluir en la investigación, cumpliendo con los objetivos previamente establecidos. .

Figura 4. Fases del proceso de *Data Mining*



Fuente: Han J, Micheline, K, Pei J (2012). *Data Mining*, (p. 7)

1.1.5.3. Aplicaciones de Oracle Data Mining

Según (Witte & Frank, 2005, p. 35) el uso de datos de personas o datos confidenciales de empresas tiene graves implicaciones éticas y profesionales. Por lo que es importante que la persona que maneja la data sea una persona confiable. Todas las tareas de *Data Mining* a lo largo de las fases de extracción, transformación, evaluación y discusión

“Algunas de las aplicaciones de data mining están en el área de marketing y ventas. Estos son dominios en donde las compañías poseen grandes volumen de data” (Witte & Frank, 2005, p. 26). También es aplicada en el campo de la medicina, fraudes de clientes, terrorismo, diagnósticos, proyecciones, decisiones, La base de los modelos es estadística lo que permite comprobación de hipótesis.

1.1.5.4. Funciones de Oracle Data Mining

“Las funciones de *Data Mining* se pueden dividir en dos categorías: supervisados (dirigido) y sin supervisión (sin dirección).” (Taft, et al.,2005, p. 1-2). La diferencia radica en que los modelos supervisados son usados para predecir un valor y los no supervisados son usados para encontrar afinidad en los datos.

Según (López C.P, 2007, p. 8) las técnicas predictivas especifican el modelo para los datos esto en relación al conocimiento teórico previo.

Las funciones supervisadas se utilizan para predecir un valor; se requiere que el analista identifique un atributo objetivo o una variable dependiente. La predicción de nuevos valores se divide en clasificación y regresión.

Funciones sin supervisión se utilizan para encontrar la estructura intrínseca, las relaciones o afinidades en los datos no utilizan datos históricos.

La agrupación de los algoritmos se puede utilizar para encontrar grupos de forma natural en los datos. *Data Mining* también puede ser clasificado como predictivo o descriptivo. Los algoritmos son usados por compañías para diferentes propósitos en las cuales pueden encontrar ventajas y desventajas. La optimización de la implementación de esta herramienta ayuda a resolver problemas.

Oracle Data Mining admite las siguientes funciones, según (Taft *et al.*, 2005, p. 1-3)

- Supervisada: Identifican un valor objetivo (*target*)
 - Clasificación: este se enfoca en predecir agrupando elementos en clases discretas.
 - Regresión: pronostica o predice valores.
 - Importancia de atributos, de esta manera se identifican los valores que pueden afectar la predicción.
 - Detección de anomalías: cualquier elemento que muestre variabilidad en relación a la data.

- No supervisados: no tienen un valor objetivo (*target*)
 - *Clustering*: identifica y encuentra agrupaciones
 - Los modelos de asociación: muy conocido en el análisis de compra; encontrando la relación en el consumo de un producto en relación a otro.
 - Extracción de características: la unión de nuevos atributos resultantes de combinación de otros.

Oracle Data Mining se basa en la exploración de datos, es por ello que, cuando se habla de esta herramienta se relaciona con una montaña gigante de datos, los cuales necesitan ser minados para encontrar información importante que pueda solucionar la problemática previamente detallada.

1.1.5.5. Pruebas de modelo supervisadas

“Los modelos supervisados son probados para evaluar la exactitud de sus predicciones. Un modelo es probado mediante la aplicación a los datos con valores conocidos y la comparación”. (Taft *et al.*,2005, p. 3-1)

Los datos de prueba deben ser compatibles con los datos utilizados para construir el modelo y deben prepararse de la misma manera que la compilación de datos.

Resultados de las pruebas de modelos en el cálculo de métricas de prueba. Las métricas de prueba exacta calculada dependerán del tipo de modelo.

1.1.5.6. Técnicas de *Data Mining*

“Es una herramienta que realiza la comprobación metódica de hipótesis y modelos para alcanzar el conocimiento usando redes neuronales, técnicas estadísticas, algoritmos, simulación” (Lopez V, 2008, p. 3.2), a través de ellas podemos encontrar la mejor técnica que se adapta y soluciona el problema planteado. El modelado a través de la construcción de la situación donde se conoce la respuesta para posteriormente aplicarla a la situación en donde se desconoce respuesta.

Las técnicas utilizadas por *Data Mining* se basan en: agrupamiento, asociación, secuenciamiento, reconocimiento de patrones, previsión, simulación de escenarios, optimización, clasificación, redes neuronales, árboles de decisión, algoritmos genéticos, métodos del vecino más cercano y reglas de inducción. Basándose en algoritmos matemáticos y estadísticos.

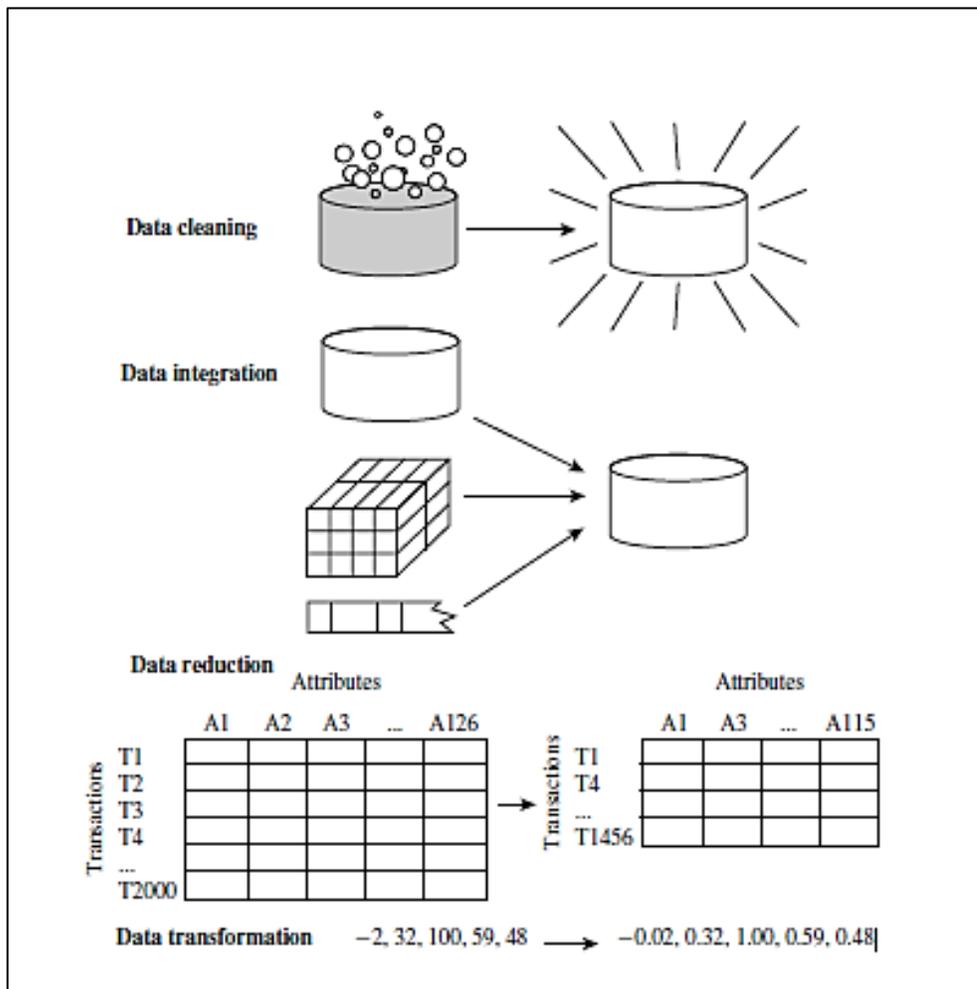
Según (Lopez V, 2008, p. 3.2) estas técnicas son aplicadas en gestión de proyectos, recursos humanos, medicina, salud industria. Sacando al máximo de utilidad de la información a un costo razonable.

1.1.5.7. Preparación y análisis de datos

La fuente de datos según (Han, Micheline, & Pei, 2012, p. 84) deben prepararse y limpiarse para cumplir con los valores, identificando valores atípicos los cuales pueden afectar la simulación. Al seleccionar la data con información fiable evitará la confusión en los resultados

Las tareas que se deben realizar para conseguir este objetivo se observan en la figura 5.

Figura 5. Preparación de datos



Fuente: Han J, Micheline, K, Pei J (2012). *Data Mining*, (p. 87)

La reducción de los datos se puede realizar a través de la desnormalización, se tomó en cuenta el objetivo de la implementación.

2. DIAGNÓSTICO SITUACIONAL

La planta productora es una empresa PYME que inició sus operaciones en el año 1999 teniendo al 2015, 16 años en el mercado, se dedica desde sus inicios a la manufacturación y distribución de productos para el cabello.

2.1. Misión y visión

- Misión. Fabricar con excelencia, comercializar con eficiencia y satisfacer las necesidades de nuestros clientes con productos innovadores de alta calidad y a precios accesibles en el área de la República de Guatemala. (fuente: datos proporcionados por planta productora)
- Visión. Ser una empresa líder en el territorio guatemalteco y centroamericano en la producción y venta de productos innovadores para el cuidado del cabello e higiene del hogar, ofreciendo productos de alta calidad y a precios accesibles. (fuente: datos proporcionados por planta productora)

2.2. Productos

Dentro de la variedad de productos que manufactura y distribuye actualmente, están: tratamientos humectantes para el cabello, lágrimas alisadoras, ampollas humectantes y restauradoras, gelatinas estilizadoras, cremas para peinar, shampoo, acondicionadores.

2.3. Clientes y alcance

Los clientes y canales de distribución corresponden a cadenas de supermercados, salas de belleza y mayoristas ubicados en toda la República de Guatemala:

- Cadena de supermercados
 - Supermercado 1: la planta productora ingresó a esta cadena bajo el programa “una mano para crecer” en donde apoyan empresas nacionales.
 - Supermercado 2: en el 2012 la planta productora ocupó el cuarto lugar en el *top 20* de empresas que proveen productos de cuidado personal. Cubre la cadena completa de supermercados

- Salas de belleza

El producto es de consumo, ya que los costos son accesibles beneficiando al dueño del negocio. Los productos son usados para los servicios ofrecidos como: planchado, cepillado, mechas, tintes y peinados.

- Mayoristas

Distribuyen los productos por mayor, son generalmente empresas grandes con capital y recursos elevados, por lo que pueden mantener gran cantidad de mercancía en su almacén. Tienen su propia fuerza de ventas y publicidad con promociones. Trabajan pedidos en base a escalas, las cuales les benefician en el costo del producto.

El alcance de distribución es en toda la República. En el interior está regionalizado de la siguiente forma:

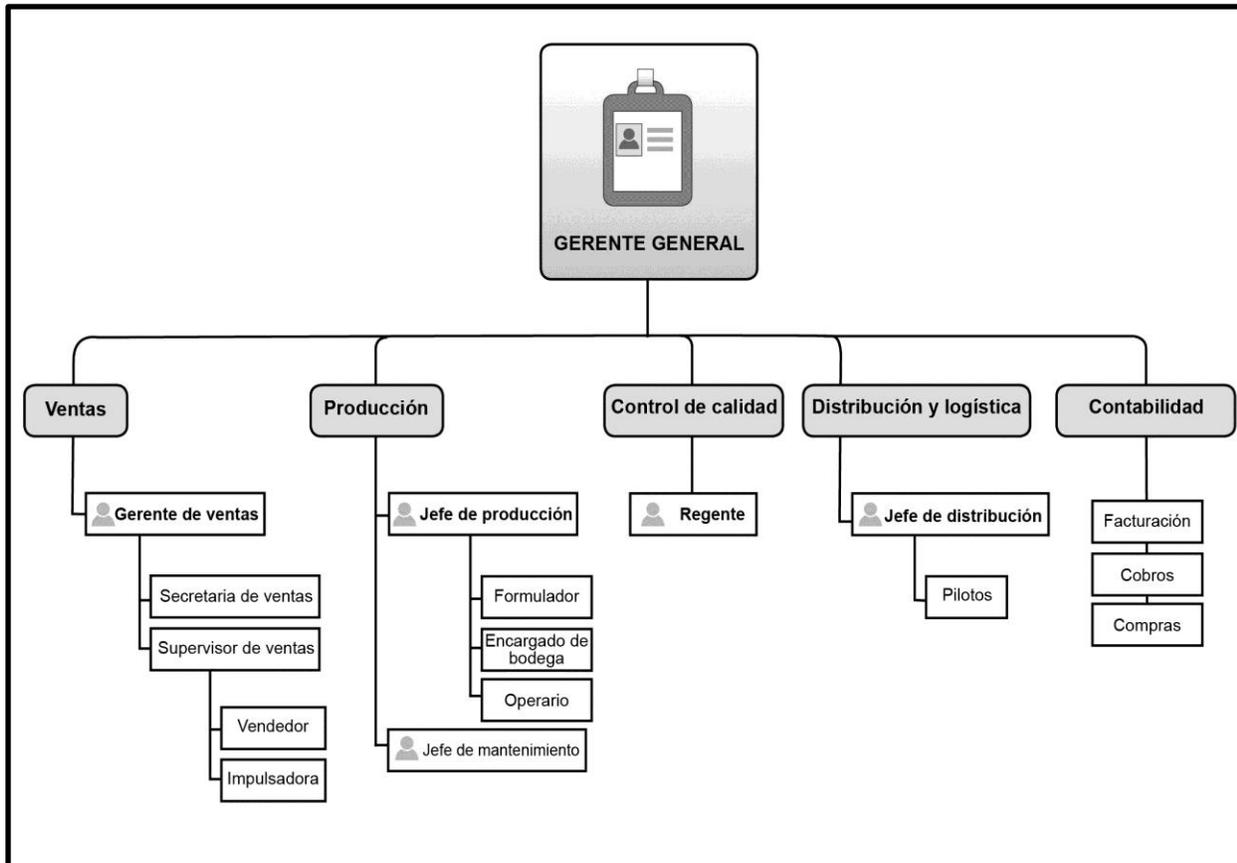
- Nororiente
- Costa chica
- Costa grande
- Occidente grande
- Occidente chico
- Petén e Izabal

2.4. Estructura organizacional

La relación de autoridad es funcional y se representa por medio de líneas cortas, este tipo de organización beneficia la comunicación directa de cada departamento, puesto que realiza actividades específicas.

El organigrama (véase figura 6) muestra la autoridad y relación mencionada.

Figura 6. Organigrama de planta productora



Fuente: proporcionado por empresa.

El gerente general es el responsable de las gestiones de la planta productora, planifica, programa y dirige las actividades con el objetivo de cumplir con la misión y visión de la empresa.

Las jefaturas de ventas, producción, distribución y logística gestionan la ejecución y control de actividades específicas de cada área. Cumpliendo con las especificaciones previamente definidas por el gerente general de la planta productora.

El jefe de producción es el responsable del desarrollo del producto desde la entrada de materia prima, transformación de la materia prima en productos cosméticos y la correcta gestión de almacenaje de materia prima y producto terminado, cumpliendo con las Buenas Prácticas de Manufactura (BPM) para laboratorios de productos cosméticos, las cuales son requeridas por el Ministerio de Salud de Guatemala, son de carácter obligatorio para continuar en operación.

El gerente de ventas es el encargado de desarrollar la planificación de ventas, visita a clientes, generación de nuevos negocios y las estrategias de ventas en los canales de distribución. La planificación de ventas, escalas, productos de temporada es parte importante de sus funciones.

El jefe de distribución y logística es el responsable que los productos de cuidado personal lleguen en tiempo a los canales de distribución.

El área contable es la encargada de facturar, cobrar y de monitorear que la base de datos de clientes, productos, vendedores esté actualizada. También es el encargado de correr los pronósticos de ventas en relación a datos actuales.

2.5. Identificación de problemas

Las áreas problemáticas identificadas por la alta gerencia se detallan a continuación, las cuales se basan en el diagnóstico del flujo de información, planificación de compras y la planificación de la producción.

2.5.1. Flujo de información

La falta de comunicación está provocando problemas en el área de ventas, pues los vendedores no comunican en tiempo los pedidos de clientes y la urgencia de muchos de esos pedidos. No existe comunicación de parte de producción y de los responsables de la bodega de producto terminado frecuentemente no hay existencia de productos para cumplir con los requerimientos del cliente.

2.5.2. Planificación de compras

Como parte del problema del flujo de información, el Departamento de Producción no tiene una planificación adecuada de compra de MP, la cual es indispensable para la producción. Esto causa retardo en la producción, debido a que los proveedores de MP no cumplen inmediatamente con los requerimientos.

El ingreso de MP fuera de tiempo provoca el pago de horas extras a los empleados y largas jornadas de trabajo, las cuales deben ser cumplidas para cubrir los pedidos facturados. Los lotes de producción que se manufacturan de esta manera no cumplen con el tiempo de cuarentena establecido por el MSP.

2.5.3. Planificación de producción

La planificación de la producción actual, no satisface la demanda, por lo que el propósito de reducir la incertidumbre no se cumple.

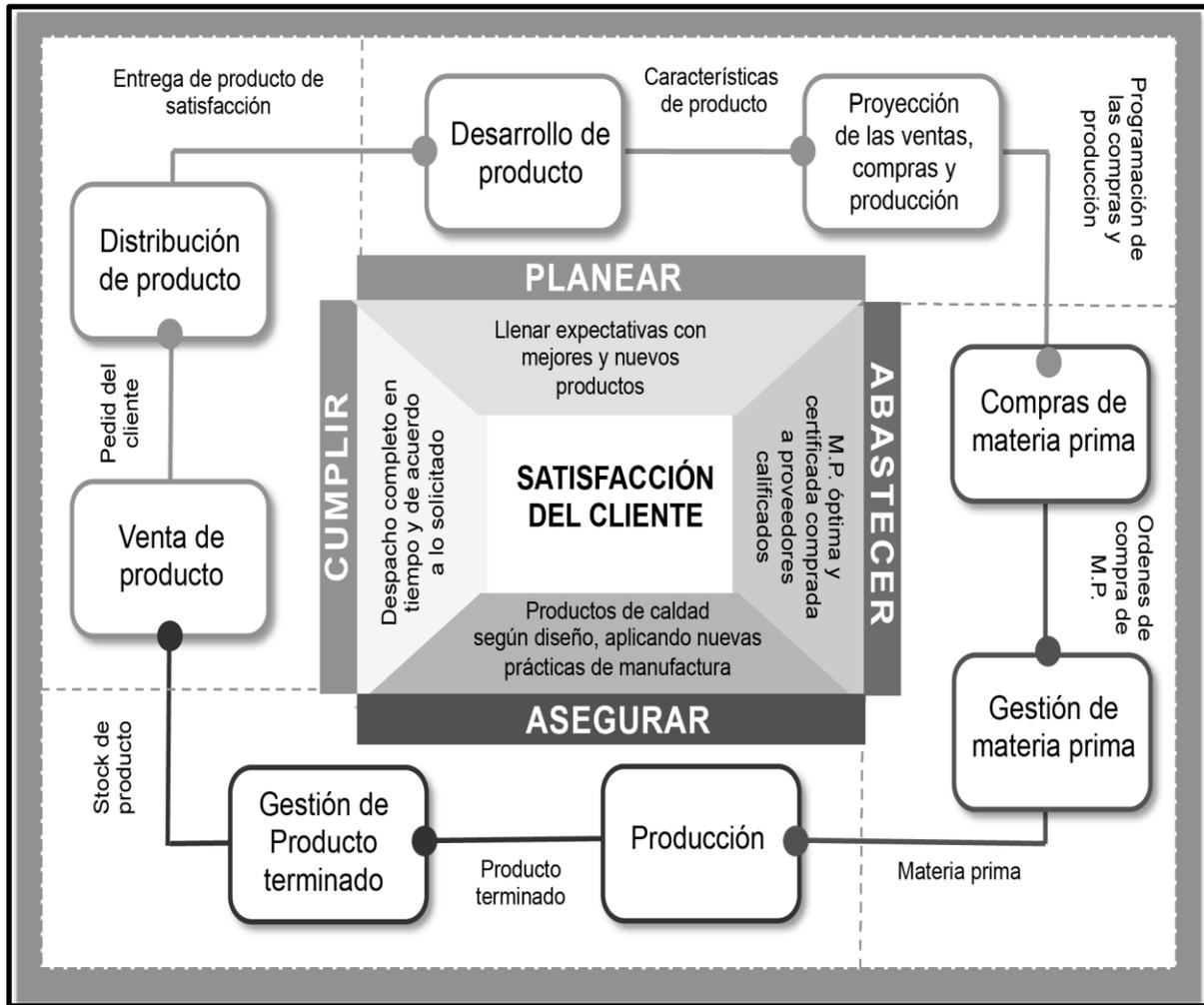
Cuando la producción es insuficiente para satisfacer la demanda diaria, se opta como recurso:

- Envío de producción que no ha cumplido con el tiempo de cuarentena reglamentado por el Ministerio de Salud Pública; que es un riesgo para la planta productora.
- Comunicación con el cliente informando del faltante, pero en muchas ocasiones no aceptan el envío de productos 2 o 3 días después de lo requerido o pactado con el vendedor.

2.6. Análisis de cadena de valor

El análisis de la situación actual de la cadena de valor de la planta productora, ayuda a definir los procesos involucrados en la planificación de compras y de producción detallados anteriormente. Para el análisis, se tomó en cuenta toda la información disponible y proporcionada por las jefaturas y gerencias. La figura 7 muestra gráficamente la cadena de valor.

Figura 7. Cadena de valor de planta productora



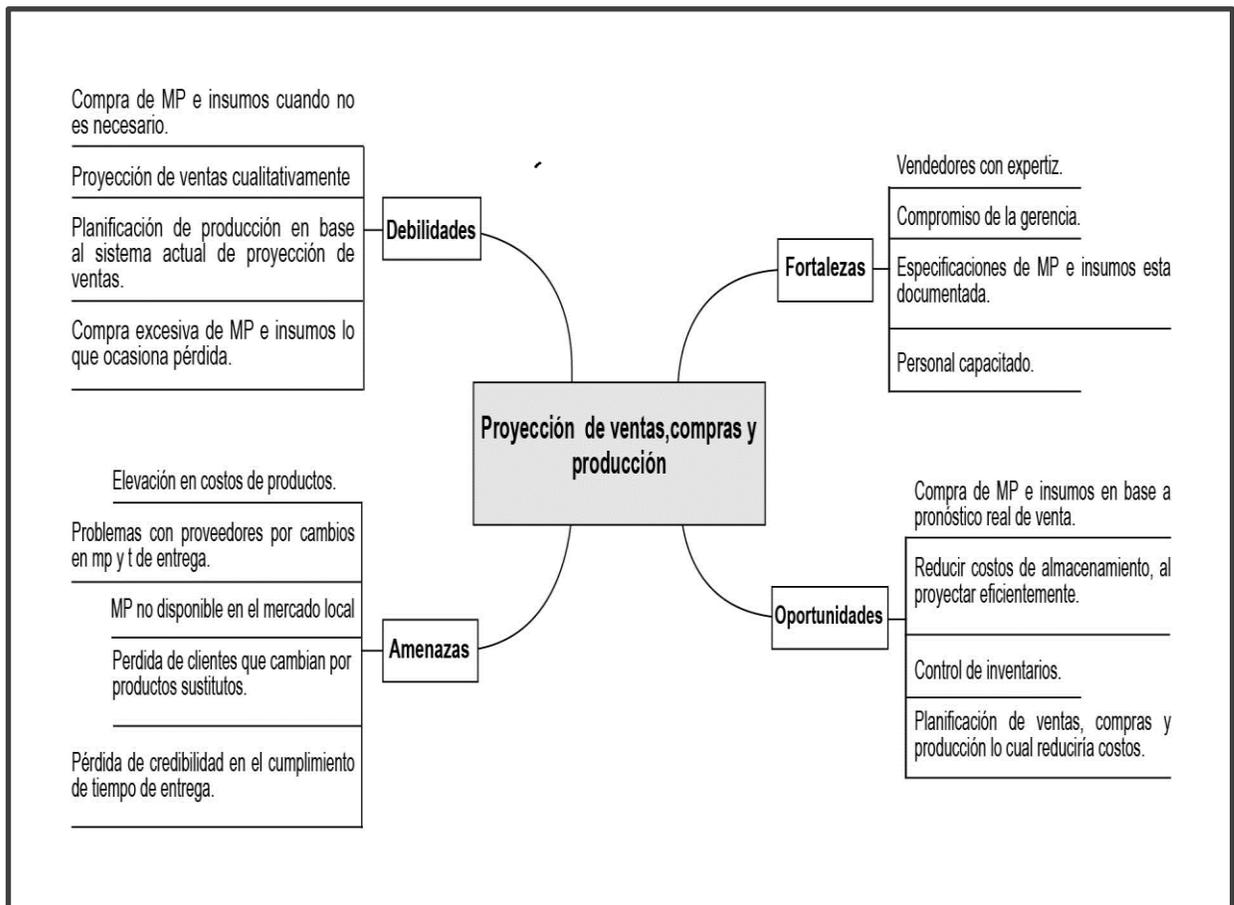
Fuente: elaboración propia.

2.6.1. Proyección de ventas, compras y producción

La proyección de ventas es parte clave de la cadena de valor puesto que a través del mismo se estiman las ventas futuras. Los pedidos son recibidos pero no se conoce con exactitud los productos requeridos por los clientes y el tiempo de entrega.

El método de predicción de venta utilizado por la planta productora es cualitativo, se basan en intuición referida a los datos históricos de ventas y cuantitativa de forma empírica. Esta información es recolectada y proporcionada por el equipo de ventas los cuales se apoyan en la estimación de ventas e historial del cliente. La figura 8 muestra el análisis FODA de la proyección de ventas, compras y producción

Figura 8. **Análisis FODA proyección de ventas, compras y producción**



Fuente: elaboración propia.

2.6.2. Compras de MP e insumos

Esta etapa consiste en la emisión de orden de compra de materia prima nacional o importada e insumos para la producción, tomando en cuenta el pronóstico de ventas.

La evaluación de los proveedores previamente calificados es importante para el cumplimiento del reglamento establecido por el Ministerio de Salud. El responsable de velar por el cumplimiento del desempeño del proveedor es el Departamento de Compras y el Departamento de Control de Calidad, los cuales tienen a su cargo la compra de materia prima e insumos para el envasado y empaque de producto terminado.

Los objetivos de este departamento son:

- Aseguramiento de materiales requeridos en tiempo.
- Reducción de costos sin perder la calidad de los mismos.
- Validar los requerimientos que cumplen con la calidad de los materiales.
- Tener un listado de posibles proveedores en casos emergentes.

Los productos resultantes de esta etapa son:

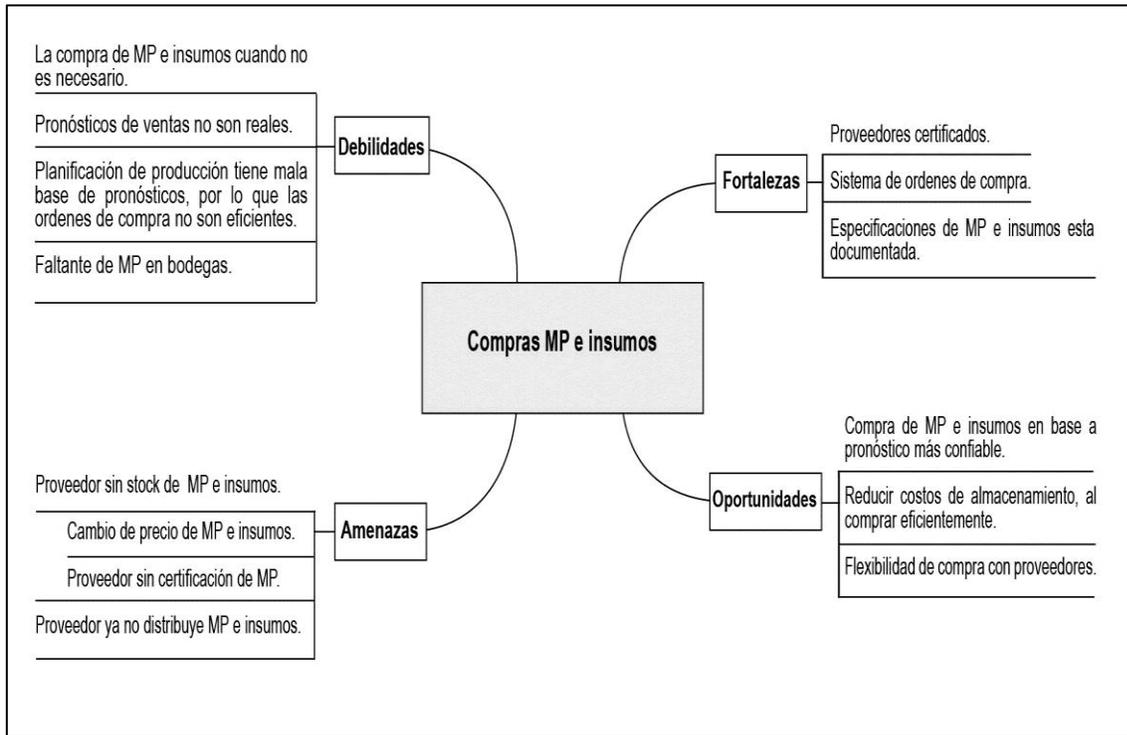
- Órdenes compra de materia prima: Esta solicitud escrita al proveedor, determina tipo, cantidad y precio; autoriza al proveedor la entrega del producto descrito evitando confusiones en el momento de entrega dado que se debe verificar orden de compra con producto físico. Y optimiza la comunicación con el proveedor.
- Registros de evaluación de proveedores, según documentación interna de la planta productora.

Algunos aspectos que contribuyen en la satisfacción del cliente son:

- Proveedores previamente calificados en cumplimiento de reglamento del Ministerio de Salud.
- Información que los proveedores deben proporcionar para asegurar que la MP es óptima para la fabricación y que no atentará contra la salud del consumidor.
- Las especificaciones exactas sobre el nombre del producto, nombre del proveedor, la fecha de manufactura, fecha de expiración, peligros y limitaciones, número de lote de fabricación.
- El certificado de análisis de toda materia prima que ingrese deberá determinar las características fisicoquímicas de los productos recibidos. Los dos últimos puntos aseguran la calidad de materia prima, la cual es importante para la fabricación del producto cosmético.

La figura 9 muestra el análisis FODA del área de compras.

Figura 9. **Análisis FODA compras**



Fuente: elaboración propia.

2.6.3. **Abastecimiento de MP e insumos**

Esta etapa consiste en el resguardo de la materia prima, en condiciones óptimas, desde su recepción hasta su despacho. Incluye los aspectos relacionados con la recepción según órdenes de compra de parte de los proveedores y el control de calidad. También el adecuado almacenamiento y el despacho en tiempo a producción.

Las áreas de almacenamiento están separadas de las otras; de tamaño adecuado, ventilado y equipado. Estas áreas están delimitadas para:

- Recepción
- Cuarentena
- Aprobación
- Rechazo

Las áreas de almacenamiento de sustancias inflamables, se encuentran separadas de toda materia prima, la cual está colocada en una tarima, y tiene registros de control de humedad y temperatura.

El responsable es el almacén de materia prima, regente y Departamento de Control de Calidad. Para ser recibida por el encargado de almacén toda materia prima tiene registros de orden de compra anticipada, la cual tiene la siguiente información:

- Número de orden
- Fecha de elaboración de la orden
- Nombre del proveedor
- Nombre y dirección de la empresa
- Términos de la compra
- Código del producto
- Cantidad del producto requerido
- Precio unitario del producto requerido

La materia prima y o materiales en el momento del ingreso se verifican los rubros especificados en la orden con la factura emitida por el proveedor, toda materia prima debe ingresar con certificado de análisis, que detalla las características fisicoquímicas del producto, sin este no debe ser recibido.

El producto resultante de esta etapa es la entrega de materia prima a producción, según orden de producción. Las materias primas afectan la calidad de los productos por lo que principalmente deben ser verificadas.

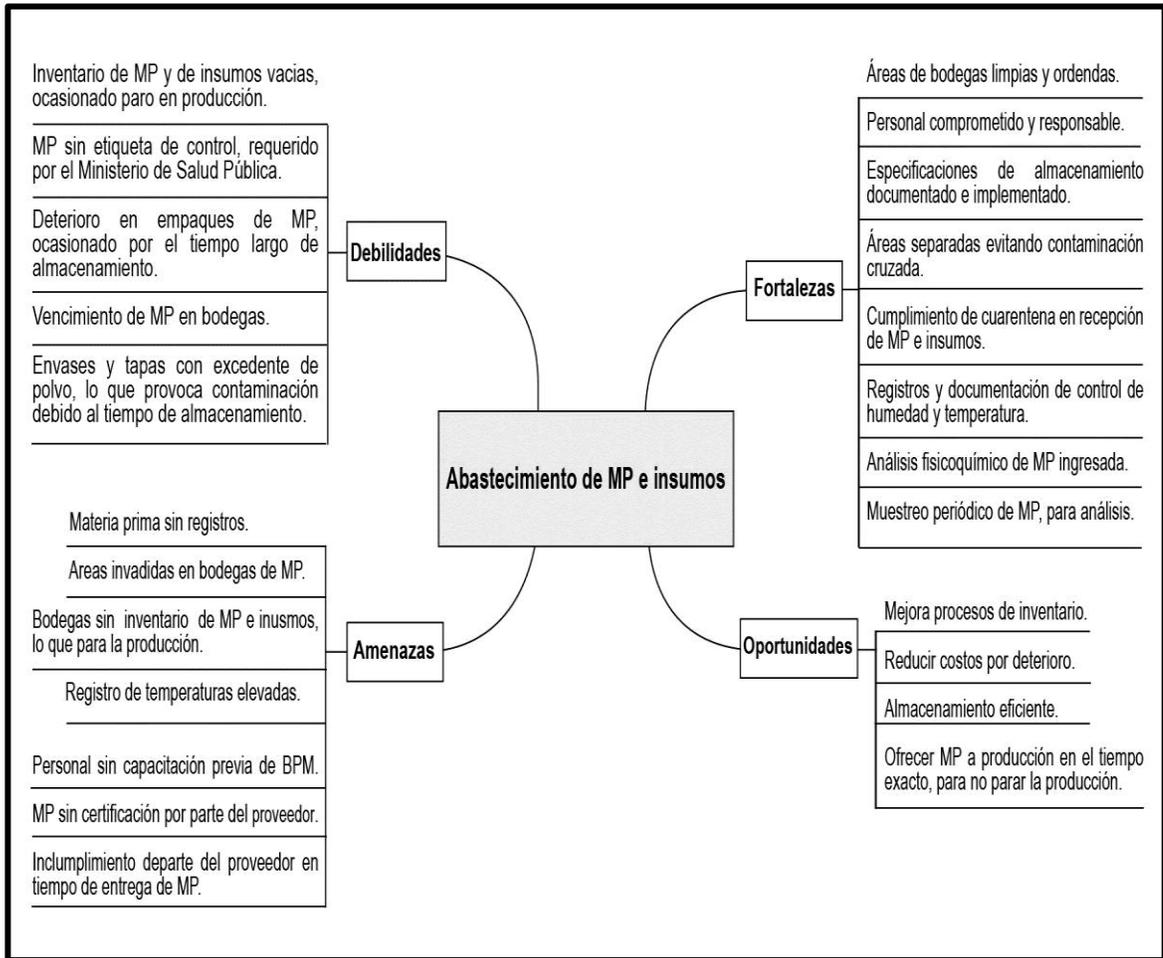
- Las características tangibles como; peso, tamaño, forma, color, textura, compatibilidad química, vida útil.
- Calidad de materiales proveídos por proveedores certificados.
- Los lotes deben analizarse, con certificado de calidad proporcionado por el proveedor y validar las variaciones.
- Cumplimiento de especificaciones atreves del muestreo, periódico realizado por el regente de la planta.
- Información de las etiquetas de todos los productos que llegan a la planta.

En caso el análisis indique que un producto no cumple con las especificaciones requeridas y el proveedor suministre continuamente productos de mala calidad, influirá en la calificación de dicho proveedor. Para readmitir a un proveedor este debe mostrar que ha tomado las medidas correctivas.

Algunos aspectos que contribuyen en la satisfacción del cliente son: el control de calidad, cumplimiento de normas de salud y despacho en tiempo a producción.

La figura 10 detalla el análisis FODA del abastecimiento de MP e insumos como parte de la cadena de valor.

Figura 10. **Análisis FODA abastecimiento de MP e insumos**



Fuente: elaboración propia.

2.6.4. Producción

La producción en una planta productora que se basa en las Buenas Prácticas de Manufactura regidas por el Ministerio de Salud Pública este punto es importante en la cadena de valor. La validación de las especificaciones de productos, normas de higiene y salud de los operarios y seguridad industrial;

son vitales para la renovación de Licencia Sanitaria en una planta productora de productos cosméticos.

El Departamento de Producción y el regente externo tienen la responsabilidad de velar por el cumplimiento del normativo RTAC 71.03.49:07 para laboratorios fabricantes de productos cosméticos.

La empresa debe:

- Contratar los servicios de supervisores de área: producción, control de calidad.
- Velar por la capacitación, experiencia y habilidades del puesto asignado.
- Delegar la responsabilidad de supervisión en los jefes de áreas.

Referente a la higiene y salud del personal:

- Deben presentar tarjetas de pulmones, manipulación de alimentos aunque no es una planta de alimentos, los supermercados exigen esta información como parte de la papelería que deben presentar vendedores.
- En caso de tener lesiones debe informar al jefe superior para determinar de esta manera si puede continuar con sus funciones.
- El uso del equipo de seguridad e higiene es indispensable, evitando el contacto directo con materias primas. Esto es por seguridad del colaborador y del aseguramiento de la calidad del producto terminado.
- Está prohibido el uso de maquillaje, joyas, relojes, celulares.
- Está prohibido el ingreso de alimentos a la planta productora.
- Es obligación del personal lavarse las manos antes de ingresar a las áreas de manufactura después de utilizar los servicios sanitarios y después de comer.

- Está prohibido el ingreso de personas al área de acceso restringido.
- Se prohíbe, comer, beber, fumar, masticar, así como guardar comida, bebidas, cigarrillos, medicamentos personales en las áreas de producción y almacenamiento.

El responsable es el Departamento de Producción, el cual lleva registros de los lotes de producción, estos entran al área de cuarentena. Se eligen dos muestras una para retención de producto terminado y otra que se envía al laboratorio para análisis fisicoquímico y bacteriológico. La muestra de retención es exigida por el Ministerio de Salud, dado que al realizar auditorías de Buenas Prácticas de Manufactura son evaluados, esta muestra debe desecharse cuando la fecha de producción es mayor a 5 años.

El producto resultante de esta etapa es el producto terminado de acuerdo a especificaciones y normas según documentación. Cuando el análisis de producto terminado regresa a planta, con el dictamen de cumplimiento entonces es liberado del área de cuarentena y enviado a almacén. Algunos aspectos que contribuyen en la satisfacción del cliente son:

- Producto confiable para ser usado, debido a la aplicación de las BPM exigidas por el Ministerio de Salud especificaciones las cuales velan por la calidad del producto terminado.
- Producto aprobado por control de calidad que asegura consistencia, olor, color requerido.

El proceso de producción se realiza por inventario, tomando como base datos históricos de ventas por canal de distribución. Actualmente, hay muchos faltantes en bodegas, la producción no está cumpliendo con la demanda.

Este proceso debe cumplir en todas sus etapas con el Reglamento Técnico Centro Americano RTCA 71.03.49:07 para laboratorios fabricantes de productos cosméticos; en donde se le da suma importancia al cliente, el producto a consumir no debe atentar contra la salud.

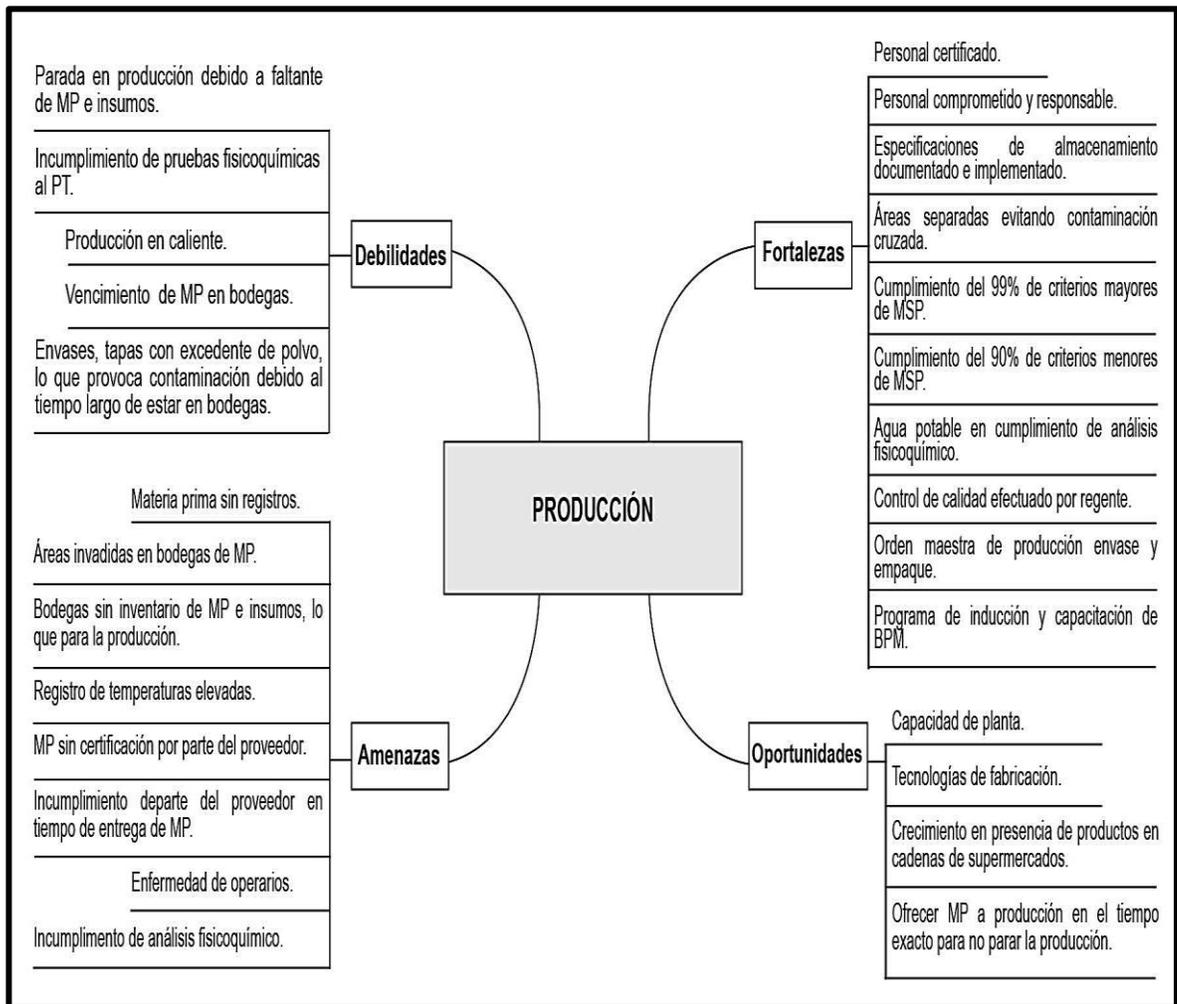
Las órdenes de producción se desarrollan con base al establecimiento de la proyección de ventas.

Las proyecciones son importantes dado que la compra de materia prima e insumos de envasado tienen las siguientes restricciones:

- La tapa importada, debe ser requerida al proveedor con un tiempo aproximado de 6 meses.
- La compra de materia prima local, y material de envasado y empaque. Se debe tomar en cuenta la normativa de salud con respecto a la planeación de adquisición de materia prima, el tiempo de vida útil y los certificados de calidad del proveedor.

En la figura 11 se detalla el análisis FODA de la producción.

Figura 11. Análisis FODA del área de producción



Fuente: elaboración propia.

2.6.5. Gestión del producto terminado

Esta etapa consiste en el resguardo del producto terminado, en condiciones óptimas, desde su recepción hasta su despacho. Incluye los aspectos relacionados con la recepción de parte de producción, el adecuado almacenamiento y el despacho en tiempo para distribución.

Las áreas de almacenamiento de producto terminado, cumplen con:

- Orden, limpieza y cumplimiento de buenas condiciones de mantenimiento.
- Se identifican para recepción, cuarentena, aprobación y rechazo.
- Todo producto terminado, debe ser colocado en tarimas y cumplir con la estiba establecida.

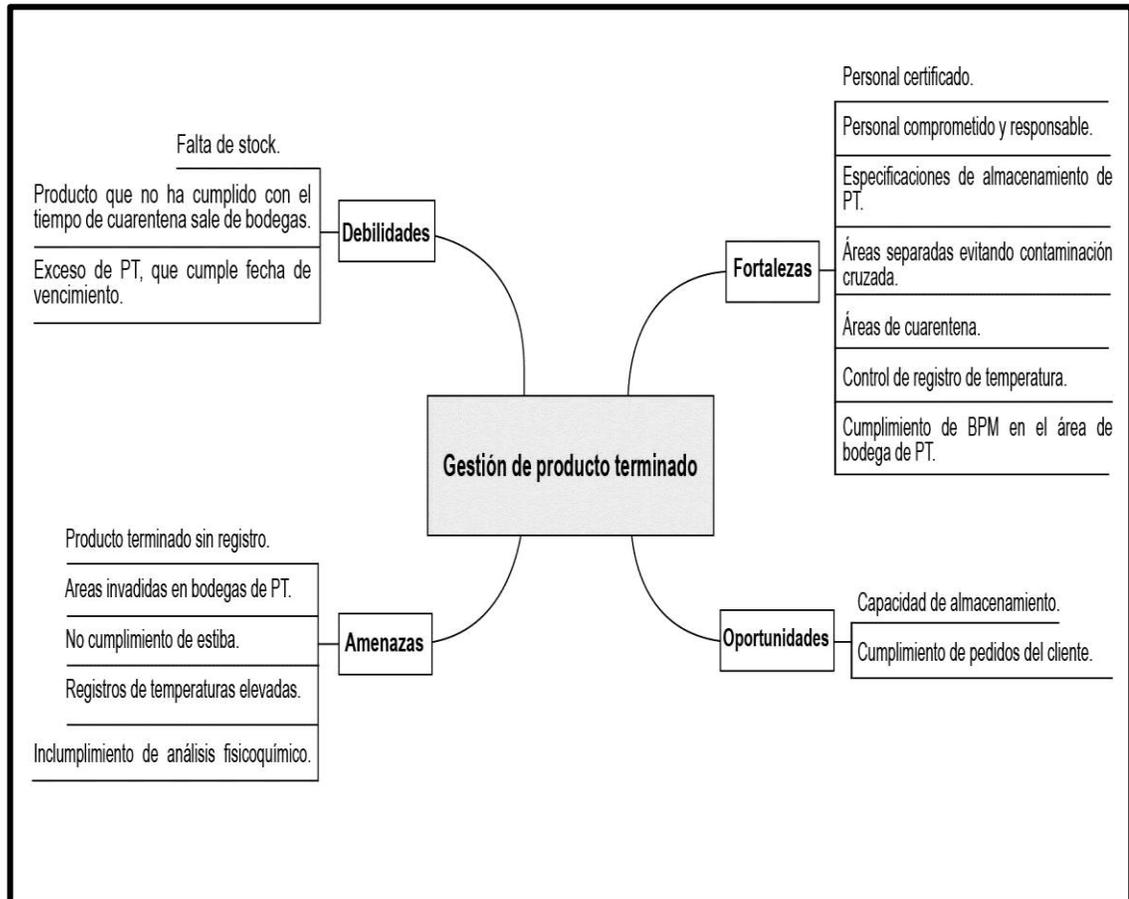
El responsable es el almacén de producto terminado, el cual vela por el cumplimiento de las condiciones de bodega, según el reglamento del Ministerio de Salud.

El producto resultante de esta etapa es el despacho de producto terminado, según requerimientos de órdenes de compra, emitidas por los clientes.

Algunos aspectos que contribuyen en la satisfacción del cliente son: control de calidad; cumplimiento de normas de salud; despacho en tiempo a los canales de distribución.

El análisis FODA se detalla en la figura 12.

Figura 12. FODA de gestión del producto terminado



Fuente: elaboración propia.

2.6.6. Ventas

Esta etapa consiste en la recepción de órdenes de pedido, por los canales de distribución.

- Canal de distribución de supermercados.
 - Supermercado 1: los pedidos se emiten por la herramienta SRE1 donde se genera el pedido con base en las ventas por tienda. La información contenida en la orden es: número de proveedor, número de tienda, ítem de producto, código de producto, número de cajas por producto, unidades por caja, descripción del producto, cantidad requerida, costo (sin IVA), fecha de emisión y fecha de cancelación. Esta plataforma permite trabajar con el centro de distribución, especificando los medios para conectar, entregar y validar la entrega de pedidos a través de internet.
 - Supermercado 2: órdenes basados en la herramienta SRE2, la cual proporciona inventario actual y pedido sugerido.
- Mayoristas: pedidos tomados por vendedores vía telefónica e internet.
- Salas de belleza: en este canal, los pedidos son tomados vía telefónica. El vendedor visita la sala de belleza solamente cuando es requerido por el cliente.

El responsable es el Departamento de Ventas que está formado por: supervisor de ventas mayoristas; supervisor de impulsadoras de

supermercados; impulsadoras de supermercados; vendedores departamentales; vendedor mayorista.

El producto resultante de esta etapa es el pedido del cliente. Este pedido es trasladado a facturación, quien elabora factura contable y otros documentos que dependen del canal de distribución:

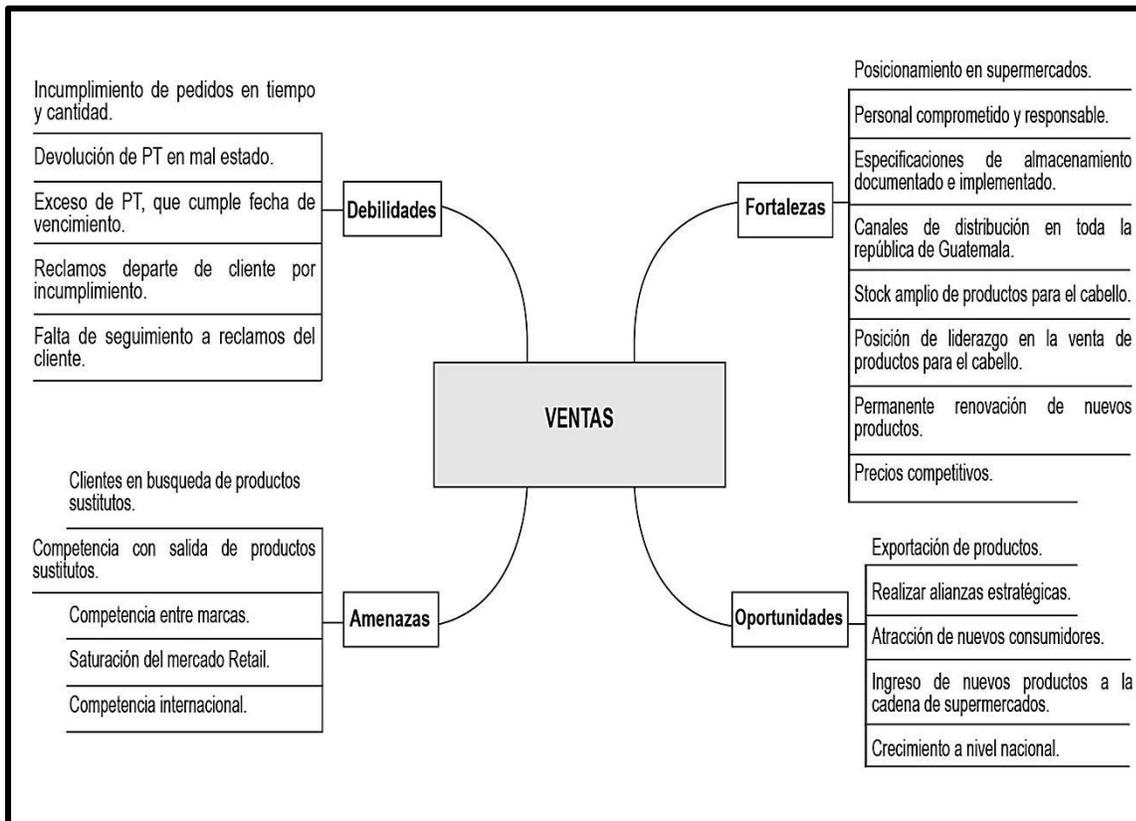
- Mayoristas: envío que incluye muestras de producto, para los mayoristas del interior se adjunta guía de transporte que hace constar cantidad de cajas de producto.
- Supermercados: orden de compra generada.
- Salas de belleza: guía de transporte contratado.

Algunos aspectos que contribuyen en la satisfacción del cliente son:

- Condiciones de la venta
- *Stock* de productos que asegura el envío de orden completa
- Características del producto

El análisis FODA del proceso de ventas se muestra en la figura 13.

Figura 13. **Análisis FODA de ventas**



Fuente: elaboración propia.

2.6.7. Distribución

La distribución se basa en tres procesos:

- Preparación de pedido: la preparación de pedido es efectuada por el encargado de distribución, utilizando una herramienta informática propia, que incluye la recepción automática de pedidos, la preparación del mismo y la emisión de la factura. (Este fue detallado en la fase de ventas).

- Traslado y entrega: el encargado de distribución coordina y programa las rutas de distribución de acuerdo a los pedidos efectuados. De igual forma asigna al personal que acompaña el reparto de acuerdo al número de pedidos y las consideraciones de temporada, porque si esta es alta se requiere mayor tiempo para realizar la entrega (navidad, semana santa, día de la madre).

Se utiliza el método de abasto, en el cual los artículos se almacenan en el centro de distribución, son compras con una distribución. Modalidad surte a las tiendas en el próximo despacho.

Algunos beneficios del método, es la reducción de los tiempos de espera en las tiendas aumentando la disponibilidad del producto, minimizando necesidades de almacenaje y reduciendo los costos de inventario y operación.

Para la entrega en el centro de distribución, se toman en cuenta los siguientes aspectos:

- El pedido y la solicitud de cita, ésta última para el día siguiente de la fecha de recepción de la solicitud de pedido. Esta es importante porque el centro de distribución recibe solamente pedidos reportados durante la cita, pues no recibe productos sin ésta.
- El producto se debe entregar completo (no se puede hacer entrega parcial de un mismo pedido) en el centro de distribución en la fecha de cancelación.
- No se puede entregar producto de una misma factura en dos camiones, al menos que ambos camiones lleguen juntos al centro de distribución y entreguen uno tras otro.

La emisión de las solicitudes de pedido se realiza por el encargado de operar SRE1, el cual monitorea inventarios en tiendas, productos con mayor demanda, ventas, etc. Así, posteriormente, el encargado de distribución llena el formulario de entrega (POD) y solicita cita de entrega en el centro de distribución.

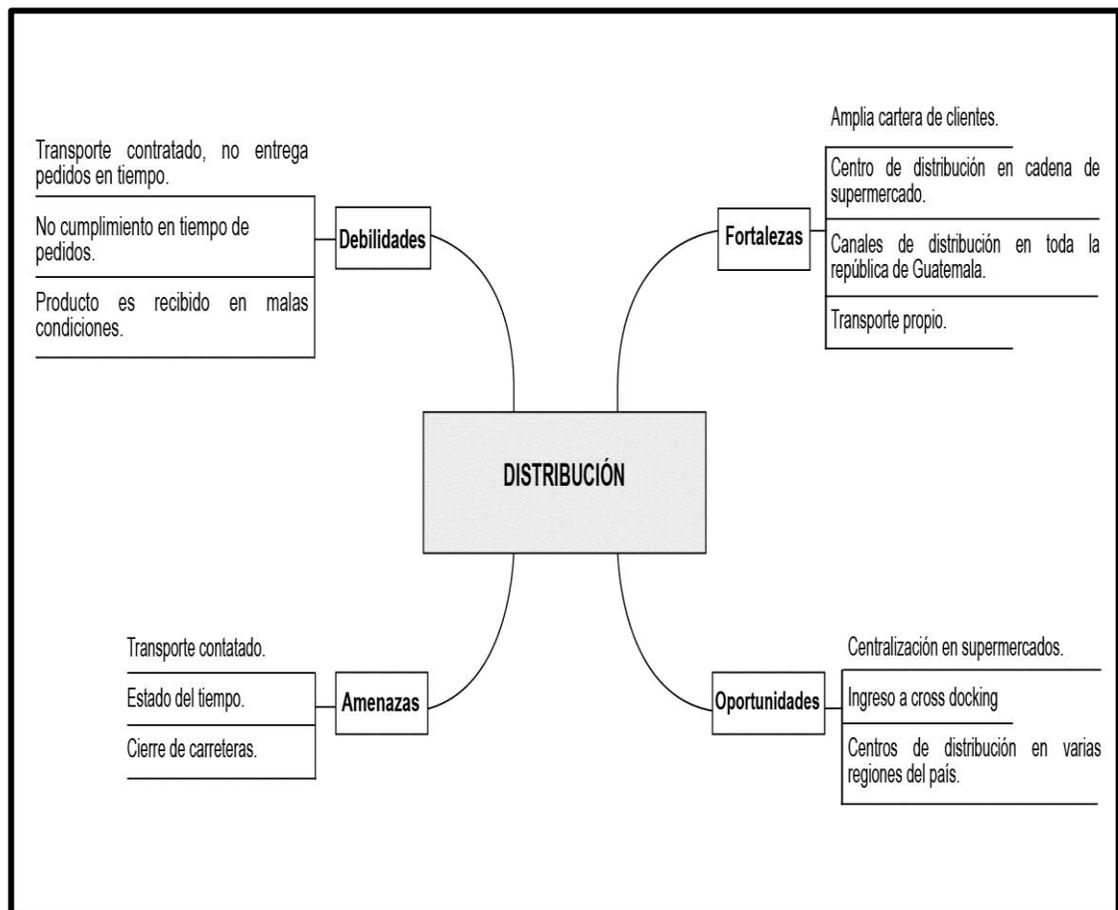
El traslado y entrega se desarrolla de la siguiente manera:

- Supermercados: a través de un centro de distribución. Los vehículos con los productos se trasladan al centro de distribución con su respectivo formulario de entrega (POD), en el caso del supermercado 1, y con la solicitud de pedido, en el caso del supermercado 2, con su respectiva factura. Entrega el producto, el piloto recibe comprobante de entrega el cual servirá para hacer el cobro posteriormente.
- Mayoristas y salas de belleza: para mayoristas y salas de belleza del interior del país cubiertas por transporte interno dos veces al mes. Las rutas se asignan según necesidad. También pueden ser cubiertas por transporte contratado, con envíos todos los días según necesidad. Los mayoristas y salas de belleza de la ciudad capital se trabajan cubriendo grupos de zonas:
 - Ruta 1 (zonas 6, 1, 2 ,6, 18);
 - Ruta 2 (zonas 9, 10, 14,carretera al El Salvador, Fraijanes);
 - Ruta 3 (zonas 7,11 carretera Roosevelt);
 - Ruta 4 (zonas 7, 19, El Milagro, San Juan Sacatepéquez, calzada San Juan);
 - Ruta 5 (Villa Nueva, Villa Canales).

- Cobro. en el caso de mayoristas y salas de belleza el cobro lo realiza el vendedor cuando realiza la siguiente gira a los clientes, validando que el producto llego en buenas condiciones. En el caso de supermercados, de acuerdo al sistema aceptado con ellos.

La figura 14 detalla el análisis FODA de la distribución.

Figura 14. **Análisis FODA de distribución**



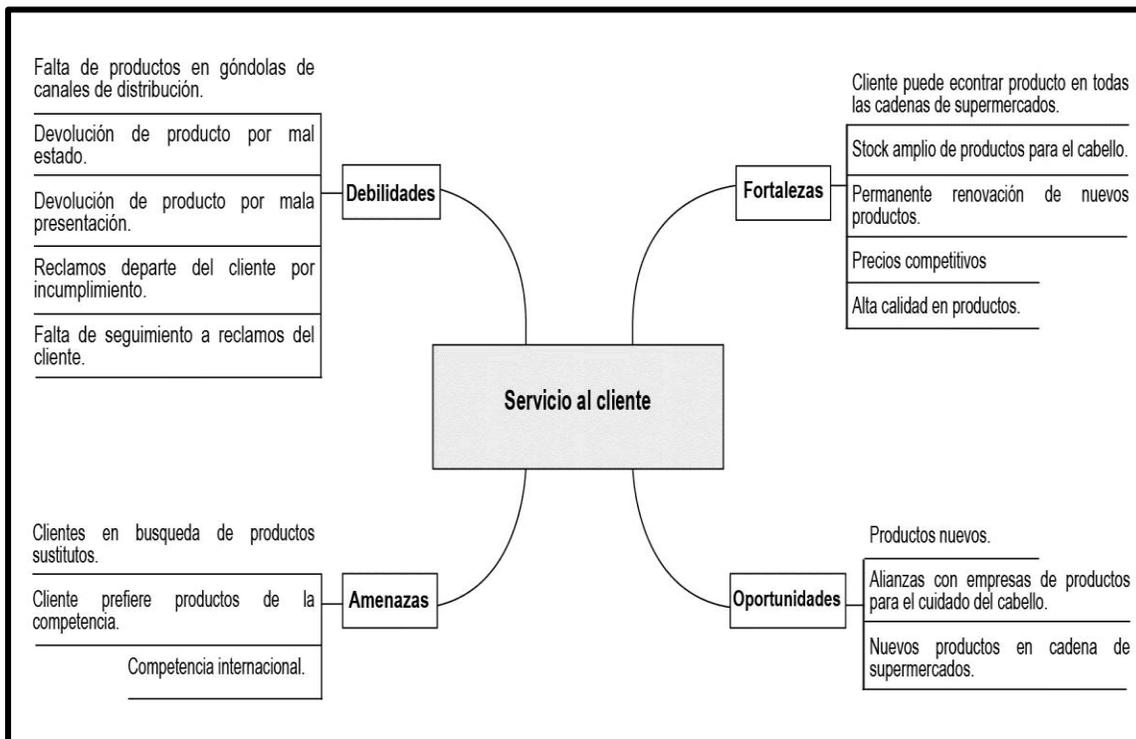
Fuente: elaboración propia.

2.6.8. Satisfacción al cliente como punto central de la cadena de valor

La figura 14 muestra el punto central de la cadena de suministro, la cual debe ser la satisfacción del cliente llenando expectativas con mejores y nuevos productos asegurando productos de calidad según diseño y aplicando BPM.

La planta productora tiene varias amenazas y debilidades que están ocasionando el incumplimiento de satisfacción en el cliente. En la figura 15 se muestra el análisis FODA.

Figura 15. Análisis FODA del servicio al cliente



Fuente: elaboración propia.

Derivado del análisis FODA de cada fase de la cadena de valor, se observó que las debilidades del sistema se fundamentan en el incumplimiento de pedidos para los canales de distribución debido a varias razones:

- Producto terminado en mal estado.
- Faltante de producto terminado en bodegas.
- Faltante de materia prima ocasionando paradas en la producción.
- Faltante de envase, el cual es requerido con 60 días de anticipación.
- Producción en caliente, lo que ocasiona producto en mal estado cuando llega a manos del consumidor final.

Como parte inicial de la cadena de valor, la proyección de ventas no es confiable dado que se lleva a cabo de una manera empírica. Es necesario implementar pronósticos de ventas confiables, por lo que la simulación con la herramienta *Data Mining* ayudará a determinar el mejor método, con el mayor porcentaje de confiabilidad.

3. SIMULACIÓN CON *DATA MINING*

Oracle Data Mining (ODM) maneja técnicas con aprendizaje supervisado y no supervisado. Los algoritmos de aprendizaje supervisados que predicen un valor objetivo, se encuentran en los siguientes modelos:

- Clasificación
- Regresión
- Atributo de importancia
- Detección de anomalía

El modelo utilizado es el de regresión, ya que predice un valor objetivo específico para cada caso de entre un número infinito de valores. Los valores existentes son las ventas históricas de 5 años de la planta productora lo cual facilitará la simulación.

Con la regresión se utilizaron los algoritmos estadísticos estándar:

- GLM: modelo lineal generalizado, la cual es la técnica estadística de modelos lineales.
- SVM: potente algoritmo para la regresión lineal y no lineal.

En la simulación se realizaron las fases descritas a continuación.

3.1. Comprensión de datos

El objetivo de la comprensión de datos es asegurar la entrada de datos falsos o erróneos que perjudiquen la simulación con *Data Mining*. La recolección de datos a través de la elección de tablas, los campos, el tipo de campo y la descripción del mismo. En la tabla III se detallan las tablas utilizadas. Esta información fue obtenida de la base de datos de facturación de la planta productora.

Tabla III. **Tablas de datos**

Nombre tabla	Número de atributos
Datos de facturación	5
Detalle de factura	5
Inventario de productos	7
Datos de clientes	6
Vendedores	3
Presentación del producto	2
Departamento del cliente	2

Fuente: elaboración propia.

La tabla IV, describe los datos utilizados para realizar la facturación, cada factura se identifica por un número único el cual es un número secuencial. El *id_cliente* y el *id_vendedor* son identificadores del cliente que solicitó la venta y

del vendedor que realizó la misma; los datos específicos de los mismos se detallan en las tablas de cliente (tabla VII) y vendedor (tabla VIII).

Tabla IV. Datos de facturación

Campo	Tipo	Descripción del campo
id_factura	número (30)	Identificador único de la factura, con valor numérico.
fecha_factura	Date	En este campo se almacena la fecha de facturación.
Id_cliente	número (20)	Código del cliente al cual se le factura.
tipo_venta	char2(10)	Este campo contiene la condición de venta, la cual puede ser crédito, contado.
Id_vendedor	número (4)	Identificador del vendedor al cual se le dará comisión por venta.

Fuente: elaboración propia

La tabla V, detalla los datos de los productos que el cliente solicitó en su compra, los campos identificador del producto (id_producto) y de la presentación (id_presentacion) se relacionan para verificar la existencia de producto en bodegas y disponibilidad de facturación.

Tabla V. **Detalle de facturas**

Campo	Tipo	Descripción del campo
id_factura	número (30)	Número de factura, la cual consta de una letra y 6 dígitos.
id_producto	número (20)	Código del producto que se facturará.
id_presentación	número	Código de la presentación del producto.
cantidad_producto	número (6)	Cantidad del producto solicitado por el cliente, el cual será facturado.
precio_total	número (12,2)	Precio total que es la multiplicación de la cantidad de producto a facturar por el precio unitario del producto.

Fuente: elaboración propia.

En la tabla VI se muestran los campos del inventario de los productos para el cuidado del cabello, el id_producto posee un código único de producto para identificarlo. El dato inventario de producto contiene la cantidad de producto terminado en bodega, los datos de esta tabla permiten el control de bodega de producto terminado y el punto de reorden para gestionar una orden de reabastecimiento.

Los datos del inventario de productos tienen estrecha relación con los datos del detalle de la facturación.

Tabla VI. **Inventario de productos**

Campo	Tipo	Descripción del campo
id_producto	número(20)(PK)	Código único del producto que se facturará.
descripcion_producto	número 2 (60)	Nombre del producto a facturar.
id_presentación	número (6)	Código de la presentación del producto.
inventario_producto	número (12)	Cantidad del producto terminado, el cual se encuentra en bodegas.
precio_unitario	número (6,2)	Precio de producto a facturar, el cual incluye IVA
costo	número (6,2)	Costo del producto a facturar.
punto_reorden	número (12)	Cantidad de producto en bodega, que nos indica que es necesario realizar una orden de reabastecimiento.

Fuente: elaboración propia.

Los datos de los clientes de todos los canales de distribución, contienen el nombre del cliente, número de identificación tributaria, dirección y departamento en el cual se encuentra el domicilio; esta información se encuentra en la tabla VII datos del cliente.

Tabla VII. **Datos de clientes**

Campo	Tipo	Descripción del campo
id_cliente	número(20)(PK)	Código del cliente.
nombre_cliente	Varchar2(60)	Nombre del cliente, el cual incluye nombre y apellido.
nit_cliente	Varchar2(15)	Número de identificación tributaria del cliente.
dirección_cliente	Varchar2(60)	Dirección domiciliar del cliente.
departamento_cliente	Varchar2(30)	Departamento de la república de Guatemala, del domicilio del cliente.
teléfono_cliente	número (10)	Número telefónico del cliente.

Fuente: elaboración propia.

Los datos del vendedor que realizó la venta, se encuentran en la Tabla VIII, en donde cada vendedor tiene un identificador único, el nombre del vendedor y la comisión correspondiente.

Tabla VIII. **Vendedores**

Campo	Tipo	Descripción del campo
id_vendedor	número(4)	Código del vendedor, que realizó la venta.
nombre_vendedor	Varchar2(40)	Nombre del vendedor, que realizó la venta.
comision_vendedor	número	Porcentaje de comisión del vendedor.

Fuente: elaboración propia.

Los productos se clasifican por la presentación, la cual está definida por el tipo de envase y la cantidad de medida del producto terminado a envasar. (Véase Tabla IX). Los datos de esta tabla tienen estrecha relación con la tabla VI Inventario de productos.

Tabla IX. **Presentación del producto**

Campo	Tipo	Descripción del campo
id_presentacion	Numérico	Código de la presentación del producto.
tipo_presentacion	Varchar2(20)	Detalla el tipo de presentación.

Fuente: elaboración propia.

Los canales de distribución están localizados en toda la República de Guatemala los datos importantes de estos se muestran en la tabla X; este dato se considera importante ya que un cliente puede tener varias sucursales.

Tabla X. **Departamento del cliente**

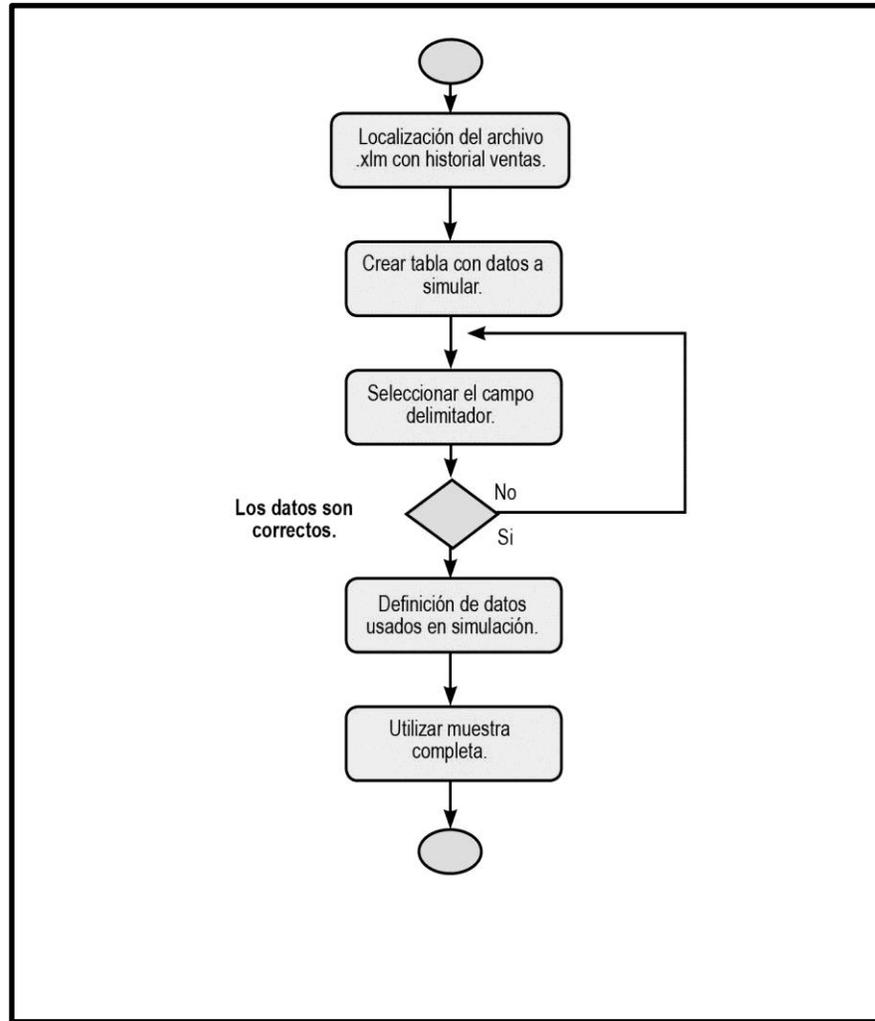
Campo	Tipo	Descripción del campo
id_departamento	Numérico	Código del departamento de Guatemala.
nombre_departamento	Varchar2(25)	Nombre del departamento del domicilio del cliente.

Fuente: elaboración propia.

3.2. Fase 1. Preparación de datos

En la preparación de los datos se utilizó la adquisición de datos históricos de ventas de la planta productora de un periodo de 5 años, los cuales se obtuvieron de un archivo .xml que contienen la base de datos de facturación de la planta productora. En la figura 16 se muestran los pasos para realizar la preparación de datos.

Figura 16. Diagrama de flujo de la preparación de datos



Fuente: elaboración propia.

El archivo .xml contiene los datos proporcionados por la planta productora véase tabla XI.

Tabla XI. **Campos de archivo .xml**

CAMPO	DESCRIPCION
fecha_factura	fecha
id_producto	Identificador del producto
id_presentacion	Identificador de la presentación del producto
descripcion_producto	Detalle del producto.
tipo_presentacion	Contenido de la presentación.
precio_unitario	Precio del producto.
costo	Costo del producto.
ventas	Número de unidades vendidas.
inventario_producto	Número de unidades en bodega.
punto_reorden	Número de unidades mínimas en bodega.
tipo_venta	Venta al crédito o al contado.
id_cliente	Identificador del cliente.
nombre_cliente	Nombres y apellidos del cliente.
nit_cliente	Número de identificación tributaria del cliente.
direccion_cliente	Dirección del cliente.
telefono_cliente	Teléfono del cliente.
id_departamento	Identificador del departamento del cliente.
nombre_departamento	Descripción de los departamentos de Guatemala.
id_vendedor	Identificador del vendedor.
nombre_vendedor	Nombre completo del vendedor que realizó la venta.
comision_vendedor	Porcentaje de comisión del vendedor.

Fuente: elaboración propia.

Los atributos seleccionados (véase tabla XII) están relacionados con el problema del pronóstico de ventas; los datos de la nueva tabla que contiene las fuentes preparadas para minar son el resultado de la depuración de fuentes de datos preparadas; estas corresponden a los 5 años de datos históricos de la planta productora un total de 27,125 registros.

Tabla XII. **Fuentes preparadas para minar**

Nombre	Tipo de dato
docenas_ventat	numérico
costot	numérico
unidades_ventat	numérico
precio_ventat	numérico
id_presentaciont	numérico
fecha_facturat	date
cus_idt	varchar2
tipo_presentaciont	varchar2
descripcion_productot	varchar2
id_productot	numérico r

Fuente: elaboración propia.

Estos datos fueron seleccionados de la tabla inicial *Simulacionshal* (tabla XI) y se colocaron en una nueva tabla llamada *Simulacion* véase tabla XII. En la tabla XIII se muestra un segmento de datos utilizados en la simulación con la herramienta *Data Mining*.

Tabla XIII. Datos preparados para simulación

CUS_IDT	ID_PRODUCTOT	DESCRIPCION_PRODUCTOT	ID_PRESENTACION	ID_PRESENTACION	FECHA_FACTURAT	PRECIO_VENTAT	COSTOT	DOCENAS_VENTAT	UNIDADES_VENTAT
14/11/1113	1	Crema rizos	3	6 onzas	14/11/11	Q 24.50	Q 19.40	11	132
14/11/1114	1	Crema rizos	4	8 onzas	14/11/11	Q 32.00	Q 25.40	7	84
14/11/1115	1	Crema rizos	5	12 onzas	14/11/11	Q 40.25	Q 34.40	5	60
14/11/1151	5	Brisa	1	2 onzas	14/11/11	Q 30.00	Q 19.50	5	60
14/11/1152	5	Brisa	2	4 onzas	14/11/11	Q 45.00	Q 29.95	3	36
14/11/1171	7	Lagrimas	1	2 onzas	14/11/11	Q 32.50	Q 19.50	4	48
14/11/1172	7	Lagrimas	2	4 onzas	14/11/11	Q 47.50	Q 29.99	2	24
14/11/1123	2	Crema para peinar	3	6 onzas	14/11/11	Q 24.50	Q 19.90	2	24
14/11/1193	9	Gelatina para peinar	3	6 onzas	14/11/11	Q 20.40	Q 11.40	1	12
14/11/1167	16	Ampolla crecimiento	7	22 ml	14/11/11	Q 8.80	Q 5.40	12	144
14/11/1157	15	Ampolla verde	7	22 ml	14/11/11	Q 8.80	Q 5.40	7	84
14/11/1177	17	Ampolla liso	7	22 ml	14/11/11	Q 8.80	Q 5.40	4	48
14/11/1187	18	Ampolla nutritiva	7	22 ml	14/11/11	Q 8.80	Q 5.40	3	36
14/11/1197	19	Ampolla miel	7	22 ml	14/11/11	Q 8.80	Q 5.40	2	24
14/11/11279	27	Shampoo crecimiento	9	16 onzas	14/11/11	Q 22.50	Q 12.40	2	24
14/11/11289	28	Shampoo nutritivo	9	16 onzas	14/11/11	Q 22.50	Q 12.40	1	12
14/11/11299	29	Shampoo negro	9	16 onzas	14/11/11	Q 22.50	Q 12.40	1	12
14/11/11319	31	Acondicionador nutritivo	9	16 onzas	14/11/11	Q 22.50	Q 12.40	1	12
14/11/11329	32	Acondicionador negro	9	16 onzas	14/11/11	Q 22.50	Q 12.40	1	12
14/11/11106	10	Tratamiento crecimiento	6	20 onzas	14/11/11	Q 39.90	Q 21.40	4	48
14/11/11116	11	Tratamiento liso	6	20 onzas	14/11/11	Q 39.90	Q 21.40	3	36
14/11/11136	13	Tratamiento verde	6	20 onzas	14/11/11	Q 39.90	Q 21.40	3	36
14/11/11126	12	Tratamiento chocol	6	20 onzas	14/11/11	Q 39.90	Q 21.40	2	24
14/11/11168	16	Ampolla crecimiento	8	44 ml	14/11/11	Q 16.80	Q 11.40	9	108
14/11/11158	15	Ampolla verde	8	44 ml	14/11/11	Q 16.80	Q 11.40	6	72
14/11/11178	17	Ampolla liso	8	44 ml	14/11/11	Q 16.80	Q 11.40	3	36
14/11/11188	18	Ampolla nutritiva	8	44 ml	14/11/11	Q 16.80	Q 11.40	5	60
14/11/11198	19	Ampolla miel	8	44 ml	14/11/11	Q 16.80	Q 11.40	2	24

Fuente: elaboración propia.

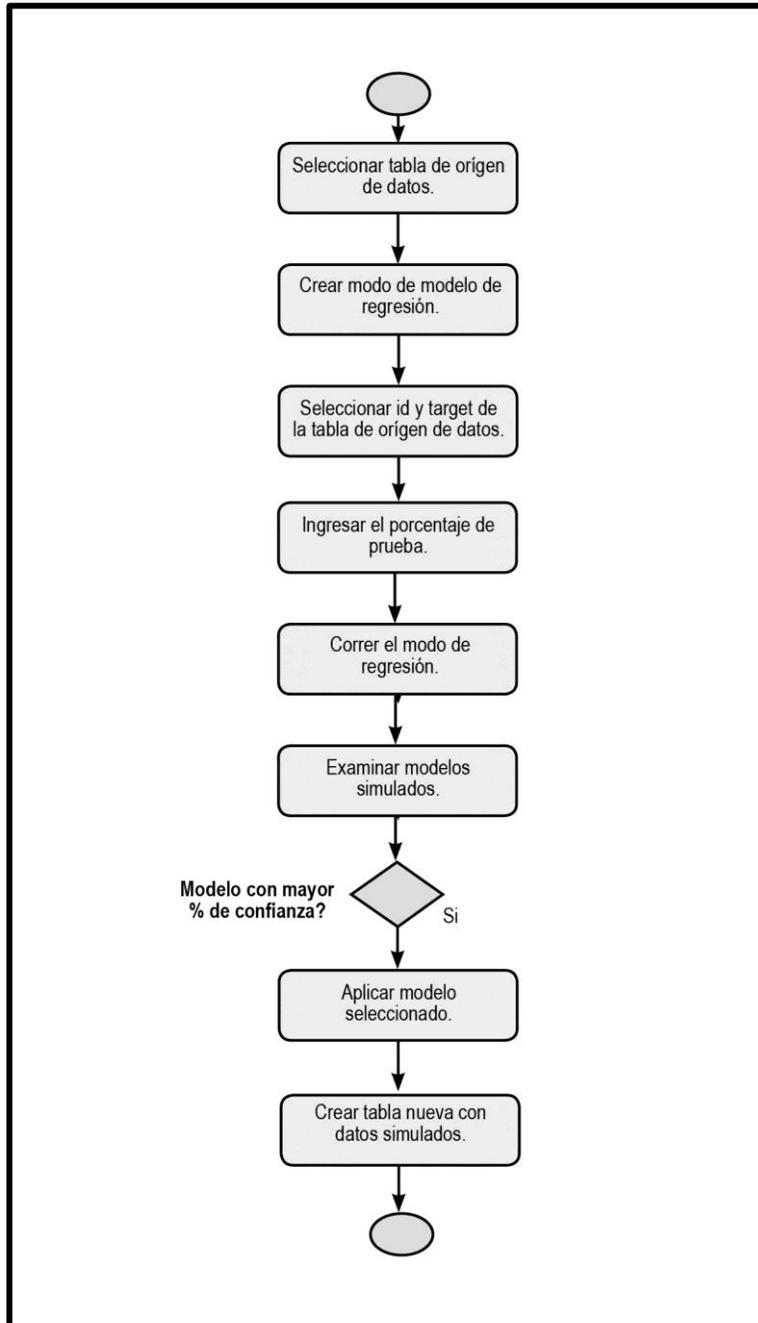
Para la exploración de datos se realizaron los siguientes pasos (gráficamente ver anexos):

- El muestreo para la selección de los datos para la exploración fueron los datos completos de la tabla *Simulacion* (27,125 registros) (anexo figura b).
- Los atributos seleccionados fueron los descritos en la tabla XII (anexo figura c).
- La agrupación se realizó por el atributo *id_productot* (anexo figura d).

3.3. Fase 2. Construcción y evaluación de modelos

Los pasos realizados para la construcción y evaluación de los modelos se muestran en el diagrama de flujo (véase figura 17).

Figura 17. **Construcción y evaluación de modelos**



Fuente: elaboración propia.

Para la creación del nodo de regresión se realizaron los siguientes pasos (gráficamente ver anexos):

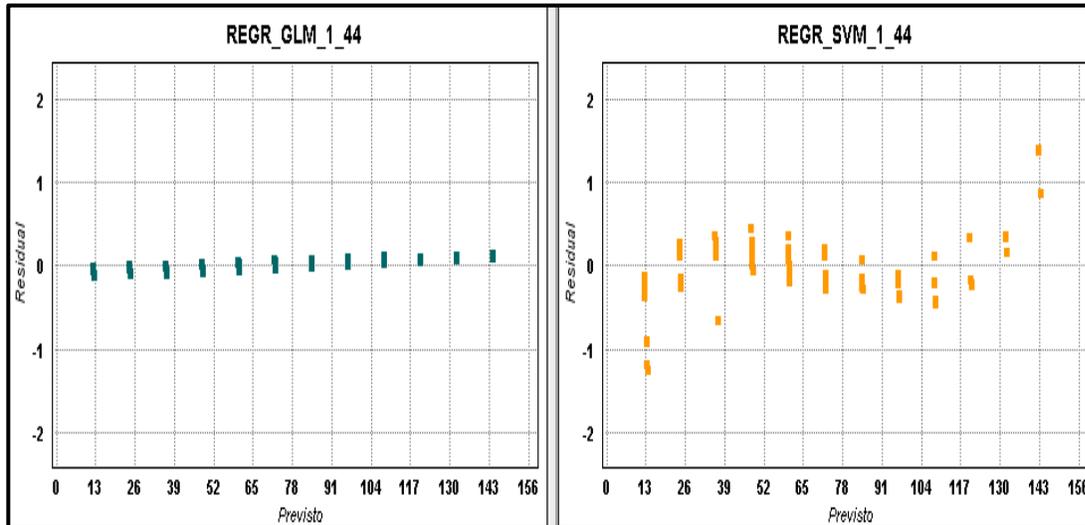
- Elección del atributo destino o variable objetivo: *unidades_venta_t* (anexo figura f).
- Elección del ID de caso o llave : *cus_idt* (anexo figura f)
- Elección de datos para simulación: usar todos los datos de creación de minería para la prueba (27,125 registros).
- El porcentaje de prueba es 80-20. Simulación con el 80 % de los datos (21,688 registros) y prueba con el 20 % restante (anexo figura g).

El tipo de técnica supervisada simulada es el método de regresión, en la cual se simularon dos algoritmos:

- *Support Vector Machines* (SVM)
- *Linear Regression* – (GLM)

La comparación de los algoritmos SVM y GLM en la prueba de simulación se pueden observar en la gráfica residual (figura 18) . Esta comparación fue determinante al igual que el resultado de los coeficientes para elegir el algoritmo a simular con los datos previamente preparados.

Figura 18. Comparación de pruebas de simulación



Fuente: elaboración propia.

Comparando los algoritmos GLM y SVM, se puede observar que el algoritmo GLM no presenta mayor variabilidad en los resultados de la prueba de simulación, por lo que es el modelo que se simulará.

Los coeficientes resultantes de esta fase se muestran en la tabla XIV (gráficamente ver anexos figura h y figura i). El algoritmo a simular, es el modelo GLM ya que es el más confiable debido al mayor porcentaje de confianza (99.8627 %) y el menor error promedio absoluto (0.0283).

Tabla XIV. **Coeficientes de prueba de simulación**

Modelo	% de confianza	Error de raiz cuadrada media	Error de promedio absoluto	Valor promedio previsto	Valor promedio real
SVM	99.0235	0.28	0.2316	35.8137	35.8185
GLM	99.8627	0.039	0.0283	35.8185	35.8185

Fuente: elaboración propia.

4. EVALUACIÓN DE SIMULACIÓN

El despliegue de resultados y evaluación de la simulación correspondiente a la fase 3, se basa en el modelo de regresión y algoritmo GLM (modelo lineal generalizado); los datos usados para la simulación fue de 27,125 registros de ventas con un total de 30 productos. El porcentaje de test para prueba se realizó usando el 80 % de los datos equivalente a 21,688 registros.

Se aplicó el nodo de regresión con el algoritmo GLM (*Generalized Linear Models*) con un nivel de confianza de 0.95; el atributo destino de la simulación son las unidades de venta, ya que el objetivo del pronóstico es encontrar las unidades que se venderán por producto.

El resultado de la simulación se registrará en un tabla nueva con los atributos detallados en la tabla XV.

Tabla XV. **Campos de tabla Resultados_generales**

NOMBRE	TIPO DE DATOS
fecha_facturat	date
id_productot	number
descripcion_productot	varchar2
id_presentaciont	number
tipo_presentaciont	varchar2
unidades_ventat	number

Fuente: elaboración propia.

Los resultados de la simulación del modelo de regresión GLM (*Generalized Linear Models*), muestra un porcentaje de confianza mayor al resultantes del algoritmo SVM; por lo que el mejor algoritmo a implementar es el GLM, en la tabla XVI se puede observar los valores resultantes de los indicadores que confirman que la simulación con este modelo es confiable.

Tabla XVI. **Detalle de simulacion GLM**

NOMBRE	VALOR
<i>ADJUSTED_R_SQUARE</i>	0.99999818
<i>AIC</i>	-35,348.33658
<i>COEFF_VAR</i>	0.10859182
<i>CORRECTED_TOTAL_DF</i>	5,436
<i>CORRECTED_TOT_SS</i>	4,458,528.9
<i>DEPENDENT_MEAN</i>	35.5696156
<i>ERROR_DV</i>	5,403
<i>ERROR_MEAN_SQUARE</i>	0.00149194
<i>ERROR_SUM_SQUARE</i>	8.06097459
<i>F_VALUE</i>	90,377,008.55
<i>GIMSEP</i>	0.00150133
<i>HOCKING_SP</i>	0.00000028
<i>J_P</i>	0.00150127
<i>MODEL_CONVERGED</i>	si
<i>MODEL_DF</i>	33
<i>MODEL_F_P_VALUE</i>	0
<i>MODEL_MEAN_SQUARE</i>	134,837.4566
<i>MODEL_SUM_SQUARES</i>	449,636.069
<i>NUM_PARAMS</i>	34
<i>NUM_ROWS</i>	5,437
<i>ROOTS_MEAN_SQ</i>	0.03862569
<i>R_SQ</i>	0.999999819
<i>SBIC</i>	-35,123.90316
<i>VAUD_COVARIANCE_MATRIX</i>	no

Fuente: elaboración propia.

La tabla XVII detalla el significado de estos indicadores.

Tabla XVII. **Detalle de indicadores de regresión GLM**

Nombre del indicador	Descripción
<i>ADJUSTED_R_SQUARE</i>	<i>Adjusted R-Square</i>
<i>AIC</i>	<i>Akaike's information criterion</i>
<i>COEFF_VAR</i>	<i>Coefficient of variation</i>
<i>CORRECTED_TOTAL_DF</i>	<i>Corrected total degrees of freedom</i>
<i>CORRECTED_TOT_SS</i>	<i>Corrected total sum of squares</i>
<i>DEPENDENT_MEAN</i>	<i>Dependent mean</i>
<i>ERROR_DF</i>	<i>Error degrees of freedom</i>
<i>ERROR_SUM_SQUARES</i>	<i>Error sum of square</i>
<i>ERROR_MEAN_SQUARE</i>	<i>Error mean square</i>
<i>F_VALUE</i>	<i>Model F value statistic</i>
<i>GMSEP</i>	<i>Estimated mean square error of the prediction</i>
<i>HOCKING_SP</i>	<i>Hocking Sp statistic</i>
<i>J_P</i>	<i>JP statistic (the final prediction error)</i>
<i>ROOT_MEAN_SQ</i>	<i>Root mean square error</i>
<i>R_SQ</i>	<i>R_square</i>
<i>SBIC</i>	<i>Schwarz's Bayesian information criterion</i>

Fuente: documentación Oracle.

Los criterios de valoración del modelo de regresión GLM y en específico el coeficiente de determinación o R^2 según (Puente, 1995, p. 528) un valor muy cercano a 1 significa perfecta predicción, el valor del coeficiente R^2 , si todas las observaciones están en la línea de regresión, el valor es 1 y si no hay relación lineal entre las variables el valor de R^2 es 0.

El modelo seleccionado para realizar la simulación tiene un coeficiente R^2 y coeficiente R^2 ajustado con un valor muy cercano a 1, indica que la línea de regresión se aproxima perfectamente a los datos reales de unidades vendidas por producto, la diferencia en el cálculo de ambas es la penalización como variables adicionales que se incluyen en el modelo.

Según (Arnau, 1996, p. 104) los criterios de información muestran la evaluación de bondad de los modelos, por lo que al presentarse el valor reducido del criterio de evaluación de información de Akaike (AIC) indicará un buen desempeño del modelo. El número resultante en la simulación de este modelo es muy pequeño (-35,348) por lo que la ventaja de este modelo es útil para el desempeño de la predicción de ventas dado que este número evalúa el ajuste del modelo a los datos.

El coeficiente de variación (*coeff_var*) es menor del 20 %, por lo que se observa que los datos utilizados para la simulación son homogéneos, tienen poca variabilidad y por lo tanto una alta precisión. Según (Brown, 1998, p. 19.7) el coeficiente de variación superior al 30 % es un indicativo de problemas en los datos o bien un indicativo de que un experimento está fuera de control.

La bondad del modelo es alta esto lo muestra el valor de *f_value* el cual es un valor muy alto.

El criterio de información Schwarz (SBIC) según (Hanke & Wichern, 2006, p. 413) es un método popular y conocido el cual determina el algoritmo simulador como bueno y muy útil. Indican que mientras mas pequeño sea su valor, mejor es el modelo. Al igual el criterio de información de Akaike (AIC).

El criterio de Hocking S_p , el GMSEP (error estimado de la predicción), y J_P (error de la predicción final) según (Moreno, 2010, p. 308) cuanto mayor sea el error peor será la bondad del ajustado. En este caso estos valores son tendientes a 0.

Referente a los errores mostrados en la tabla XIV; el error de la media cuadrada (*error_mean_square*) y el error de la media cuadrática (*root_mean_sq*) según (Moreno, 2010, p. 308) estos errores son los principales y los utilizados comunmente debido a la facilidad de manipulación matemática. En este caso los valores son casi 0 lo que indica que el modelo es bastante confiable.

El análisis de estas variables comprueban que el modelo elegido para realizar la simulación es confiable puesto que los valores se encuentran dentro de los rangos establecidos por fuentes confiables.

5. PRESENTACIÓN DE RESULTADOS DE SIMULACIÓN

El uso de la herramienta ODM, descubre información que no se espera obtener. Los modelos son confiables, en la etapas de prueba usan técnicas estadísticas las que luego son validadas, por la misma herramienta.

Las pruebas de simulación se realizaron inicialmente con todos los datos proporcionados por la planta productora (59,000 registros) lo que generaba un porcentaje de confianza tendiente a 0; por lo que la construcción de datos se realizó varias veces hasta que la selección de datos fuera la óptima.

Finalmente, la tabla elegida para la simulación, fue la agrupada por producto por día.

Los resultados de la simulación se generaron en una tabla nueva llamada Resultados_generales (gráficamente ver anexos figura n) en la tabla XVIII se muestra un segmento de esta tabla, se puede observar que los valores de la columna regr_glm_1_44_pred son 99.86 % de los valores reales de ventas de la columna unidades_ventat.

Tabla XVIII. Columnas de datos de simulación

REGR_GLM_1_4_4_PRED	FECHA_FACTURAT	ID_PRODUCTOT	DESCRIPCION PRODUC-TOT	ID_PRESENTACION	ID_PRESENTACION	UNIDADES-VENTAT
131.8837572	12/01/09	1	Crema rizos	3	6 onzas	132
107.8755024	12/01/09	1	Crema rizos	4	8 onzas	108
59.94539443	12/01/09	1	Crema rizos	5	12 onzas	60
71.98604448	12/01/09	5	Brisa	1	2 onzas	72
47.94950472	12/01/09	5	Brisa	2	4 onzas	48
95.97664974	12/01/09	7	Lagrimas	1	2 onzas	96
36.00535306	12/01/09	7	Lagrimas	2	4 onzas	36
35.98865243	12/01/09	2	Crema para peinar	3	6 onzas	36
23.99972002	12/01/09	9	Gelatina para peinar	3	6 onzas	24
107.9214058	12/01/09	16	Ampolla crecimiento	7	22 ml	108
83.91829693	12/01/09	15	Ampolla verde	7	22 ml	84
47.95668291	12/01/09	17	Ampolla liso	7	22 ml	48
24.01237221	12/01/09	18	Ampolla nutritiva	7	22 ml	24
12.02352483	12/01/09	19	Ampolla miel	7	22 ml	12
23.99947104	12/01/09	27	Shampoo crecimiento	9	16 onzas	24
23.98587756	12/01/09	28	Shampoo nutritivo	9	16 onzas	24
12.00001756	12/01/09	29	Shampoo negro	9	16 onzas	12
12.00001527	12/01/09	31	Acondicionador nutritivo	9	16 onzas	12
95.94675364	12/01/09	16	Ampolla crecimiento	8	44 ml	96
71.94364473	12/01/09	15	Ampolla verde	8	44 ml	72
35.98203071	12/01/09	17	Ampolla liso	8	44 ml	36
47.97245625	12/01/09	18	Ampolla nutritiva	8	44 ml	48
47.96185428	12/01/09	19	Ampolla miel	8	44 ml	48
59.95700752	12/01/09	10	Tratamiento crecimiento	6	20 onzas	60
35.9798486	12/01/09	11	Tratamiento liso	6	20 onzas	36
35.96999389	12/01/09	13	Tratamiento verde	6	20 onzas	36
12.02188011	12/01/09	12	Tratamiento chocol	6	20 onzas	12

Fuente: elaboración propia.

A continuación se presentan tablas y gráficas comparativas de los productos líderes en la planta productora, como resultantes de la simulación de datos.

La tabla XIX muestra la simulación de ventas en relación a las ventas reales del período junio 2012 a diciembre 2013 correspondientes al producto Crema rizos en la presentación de 6 onzas; la diferencia de las columnas simulación y ventas reales es por décimas.

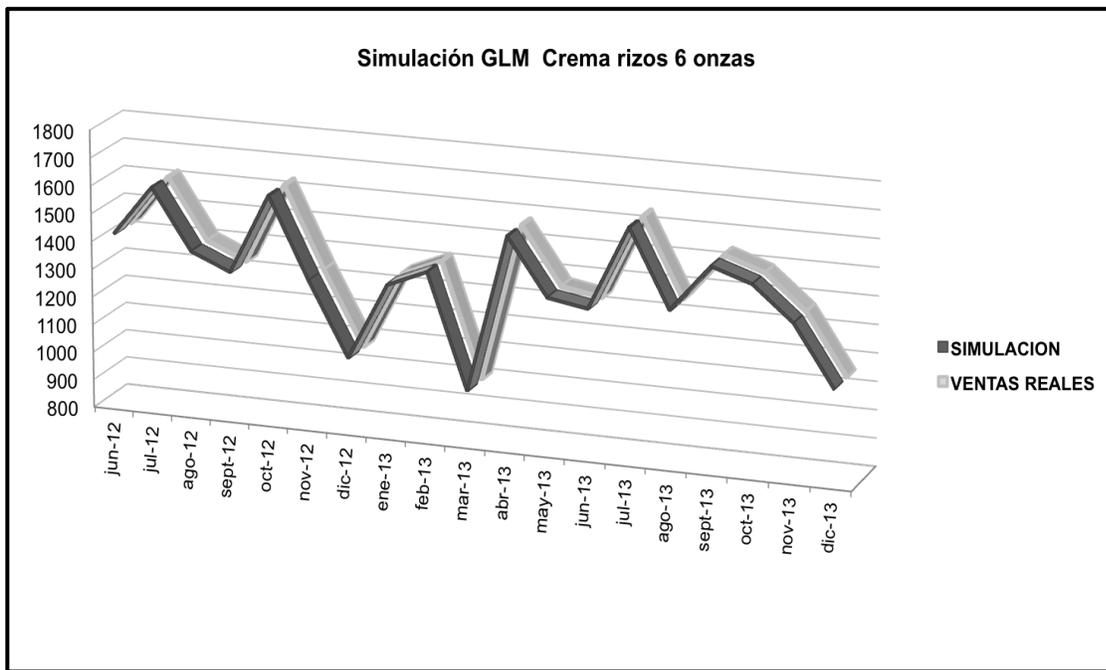
Tabla XIX. **Simulación versus venta real Crema rizos 6 onzas**

Simulación de regresión GLM		
id_producto 1	Crema para rizos	id_presentacion 3 6onzas
FECHA_FACTURAT	SIMULACION (UNIDADES)	VENTAS REALES (UNIDADES)
jun-12	1,427.99	1,428
jul-12	1,607.79	1,608
ago-12	1,392.18	1,392
sept-12	1,332.04	1,332
oct-12	1,619.89	1,620
nov-12	1,332.17	1,332
dic-12	1,068.28	1,068
ene-13	1,343.90	1,344
feb-13	1,403.92	1,404
mar-13	996.16	996
abr-13	1,559.88	1,560
may-13	1,356.25	1,356
jun-13	1,332.05	1,332
jul-13	1,631.87	1,632
ago-13	1,356.13	1,356
sept-13	1,523.82	1,524
oct-13	1,476.15	1,476
nov-13	1,356.00	1,356
dic-13	1,140.15	1,140

Fuente: elaboración propia.

La figura 19 muestra la gráfica de los datos anteriores, se puede observar que la diferencia entre los datos de simulación y ventas reales es relativamente pequeña.

Figura 19. **Simulación (unidades) versus venta real (unidades)**
Crema rizos 6 onzas



Fuente: elaboración propia.

. La tabla XX muestra la simulación de ventas en relación a las ventas reales del período junio 2012 a diciembre 2013 correspondientes al producto crema rizos en la presentación de 8 onzas; la diferencia de las columnas simulación y ventas reales es por décimas.

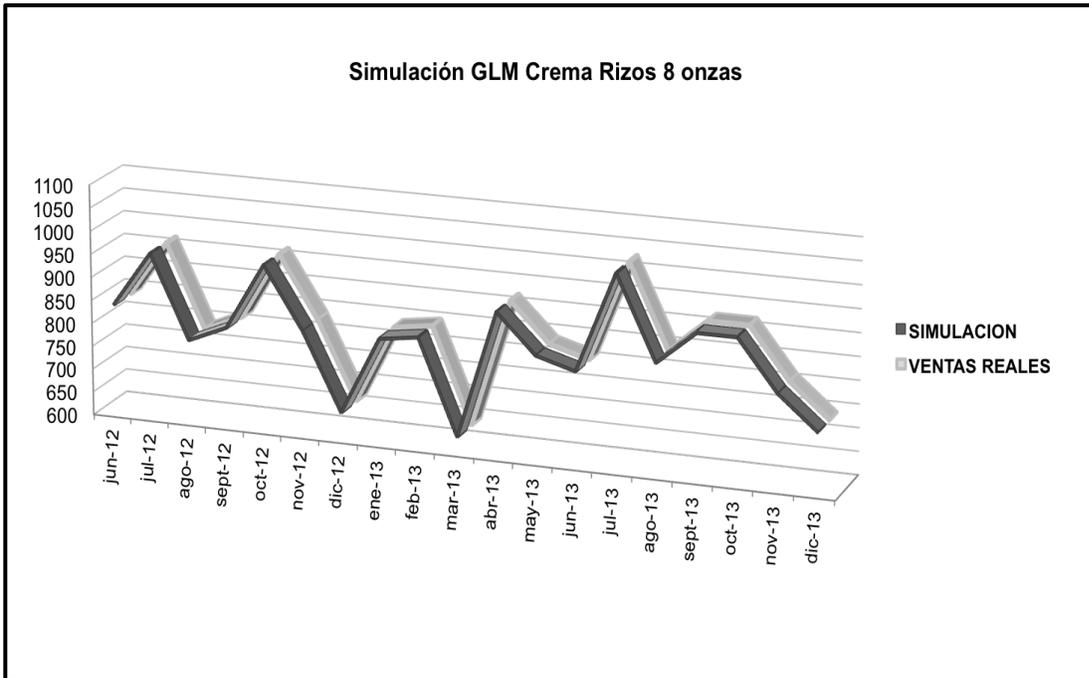
Tabla XX. **Simulación versus Venta real Crema rizos 8 onzas**

Simulación de regresión GLM id_producto 1 Crema para rizos id_presentacion 8 onzas		
FECHA_FACTURAT	SIMULACION (UNIDADES)	VENTAS REALES (UNIDADES)
jun-12	839.97	840
jul-12	959.83	960
ago-12	780.15	780
sept-12	815.95	816
oct-12	959.90	960
nov-12	828.00	828
dic-12	660.09	660
ene-13	827.85	828
feb-13	839.90	840
mar-13	635.99	636
abr-13	911.92	912
may-13	828.07	828
jun-13	803.97	804
jul-13	1,019.79	1,020
ago-13	839.97	840
sept-13	911.84	912
oct-13	911.99	912
nov-13	803.97	804
dic-13	731.96	732

Fuente: elaboración propia.

La figura 20 muestra la gráfica de los datos anteriores, se puede observar que la diferencia entre los datos de simulación y ventas reales es relativamente pequeña.

Figura 20. **Simulación (unidades) versus Venta real (unidades)**
Crema rizos 8 onzas



Fuente: elaboración propia.

La tabla XXI muestra la simulación de ventas en relación a las ventas reales del período junio 2012 a diciembre 2013 correspondientes al producto Lágrimas en la presentación de 2 onzas; la diferencia de las columnas simulación y ventas reales es por décimas.

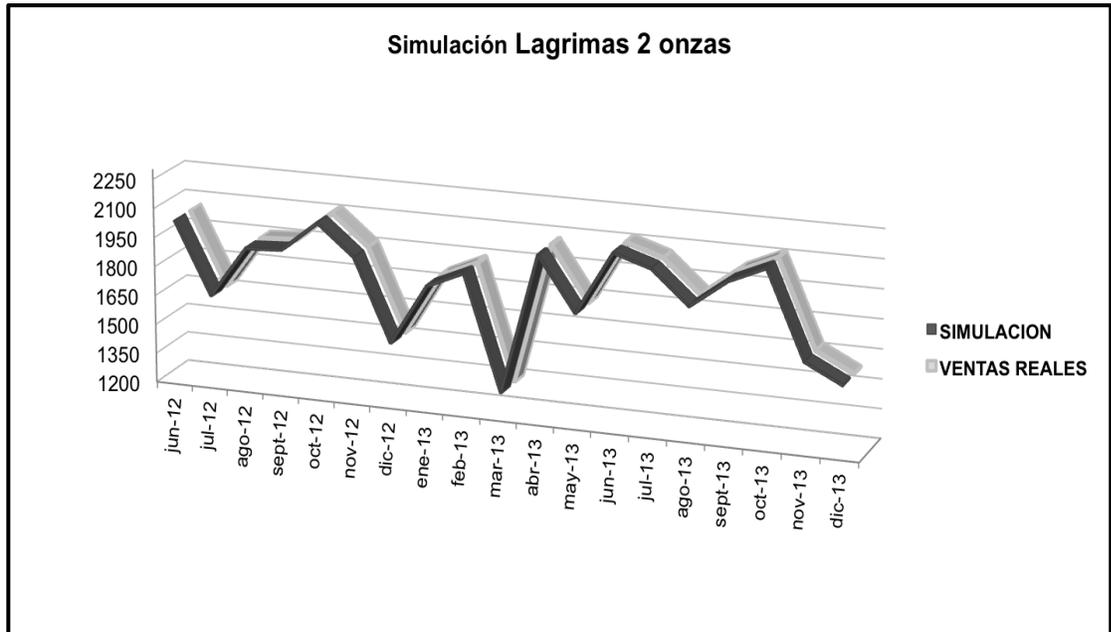
Tabla XXI. **Simulación versus Venta real Lágrimas 2 onzas**

Simulación de regresión GLM		
id_producto 7	Lágrimas	id_presentacion 1 - 2
onzas		
FECHA_FACTURAT	SIMULACIÓN (UNIDADES)	VENTAS REALES (UNIDADES)
jun-12	2,039.47	2,040
jul-12	1,668.29	1,668
ago-12	1,931.81	1,932
sept-12	1,943.49	1,944
oct-12	2,099.66	2,100
nov-12	1,943.64	1,944
dic-12	1,523.95	1,524
ene-13	1,835.53	1,836
feb-13	1,919.53	1,920
mar-13	1,331.99	1,332
abr-13	2,051.59	2,052
may-13	1,776.09	1,776
jun-13	2,099.21	2,100
jul-13	2,039.77	2,040
ago-13	1,871.77	1,872
sept-13	2,015.51	2,016
oct-13	2,099.66	2,100
nov-13	1,656.01	1,656
dic-13	1,559.88	1,560

Fuente: elaboración propia.

La figura 21 muestra la gráfica de los datos anteriores, se puede observar que la diferencia entre los datos de simulación y ventas reales es relativamente pequeña.

Figura 21. **Simulación (unidades) versus Venta real (unidades)**
Lágrimas 2 onzas



Fuente: elaboración propia.

Para los productos simulados con el modelo de regresión GLM existe similitud en las gráficas de datos de simulación de ventas *versus* datos de venta real.

Después de obtener como resultado de la simulación que el mejor método de regresión es el algoritmo GLM, se propone: implementar el modelo de regresión y algoritmo GLM para pronosticar las ventas de la planta productora.

En relación a la mejora de la productividad, se observan cambios notables en los siguientes factores (por ser datos confidenciales no se detallan cálculos y datos numéricos exactos):

- Aumento en unidades producidas.
- Reducción en devolución de producto por estar en malas condiciones o no cumplir con análisis fisicoquímicos.
- Eliminación gradual en el pago de horas extras.
- Mayor producción por operario.
- Cumplimiento en entrega de producto terminado en tiempo y cantidad solicitada por el cliente.
- Aumento en ventas del 10 %.
- Reducción gradual en energía eléctrica, agua por ser insumos de producción.
- Eliminación de tiempos muertos en la producción.
- Reducción del 8 % en el costo del producto.
- Indicador de satisfacción del cliente con un valor óptimo de 2.

El retorno sobre la inversión (ROI) es mayor en relación a otras herramientas de pronósticos como *Forecast Pro*, *R*, entre otros; debido al soporte que se adquiere con la adquisición del mismo.

6. DISCUSIÓN DE RESULTADOS

La implementación de la herramienta *Oracle Data Mining* a través del cálculo del pronóstico de ventas ayudará a aumentar la productividad desde el inicio de la cadena de valor. A continuación se detalla de qué manera aumentará la productividad en la planta productora y los beneficios de la implementación de la herramienta tecnológica.

6.1. Aumento de la productividad

En base al análisis FODA de la cadena de valor descrita en el capítulo 2 Diagnóstico Situacional y de la implementación del método de pronóstico GLM se percibe un aumento en la productividad en los puntos de la cadena de valor.

6.1.1. Proyección de ventas, compras y producción

La proyección de ventas proporcionó a la empresa datos certeros en relación a la compra de materia prima, insumos, planificación de la producción, esto permitió tener holgura, estabilizando la planta productora, generando tiempo productivo y sin recargar las instalaciones, maquinaria y el recurso humano.

El proceso expedido, sin paros, ni retrasos, de producción, porque se cuenta con la materia prima exacta y correcta, permitió reducir los periodos de tiempo ociosos, provocados por la falta de materia prima, y los periodos de tiempo con trabajo excesivo, provocados por horarios de producción

extendidos. Además de reducir o eliminar el pago no planeado, que resulta de las horas extras.

6.1.2. Compras de MP e insumos

La compra de materia prima es un factor externo que influye en la productividad. Determinar el mejor método de pronóstico de ventas como la entrada para el buen manejo de materia prima e insumos ayudó a disminuir las paradas de producción haciendo a la planta productora más productiva.

Al implementar el modelo GLM para pronóstico de ventas, se realizó el cálculo de materia prima a partir de octubre de 2014.

La importancia de la materia prima radica en que el 70 % del costo del producto terminado está en el valor de la misma. El exceso y la falta de materia prima causaban impacto financiero; por un lado, tener materia prima en grandes cantidades aumentaban los costos de almacenaje y la falta afectaba en el no cumplimiento de entrega de producto terminado, en algunos casos los clientes castigan con multa por el no cumplimiento (% en relación a la orden de compra).

Otro aspecto importante son los certificados de análisis fisicoquímicos de materia prima los cuales están antes de ser usados en la producción, esto beneficia a la planta productora en el cumplimiento con los requerimientos del Ministerio de Salud y en la calidad del producto terminado.

A futuro la compra de materia prima se realizará con mejores negociaciones pues al tener una proyección certera, los costos proporcionados por los proveedores serán más económicos. Al tener un mejor costo con los

proveedores reducirá el costo del producto terminado. Durante los meses de enero, febrero y marzo del 2015 hubo una reducción del 8 % en los costos de producto, el cual debe aumentar en forma gradual.

6.1.3. Abastecimiento de MP e insumos

El resguardo de la materia prima es óptimo con la implementación del pronóstico de ventas a través de la simulación, hay flujo constante de materia prima.

Los insumos, envases y tapas utilizados para empacar los productos están abastecidos en relación a los requerimientos basados en el pronóstico de ventas. Se redujo el espacio en el almacenamiento de envases, tapas eliminando de esta manera la contaminación por polvo, moho; reduciendo el pago de almacenaje o embodegado.

6.1.4. Producción

Las paradas en producción disminuyeron en relación a las paradas del año 2014, aumentando de esta manera la productividad en relación a:

- Ahorro en compra de insumos.
- Reducción en el pago de horas extras a personal.
- Costo de operación.
- Disminución en devolución de producto en malas condiciones: derrames, tapas y envases quebrados.
- Eliminación de producto terminado despachado en caliente.

La eliminación gradual de horas extras hasta llegar a 0, eliminación de sobrecarga o forzado de maquinaria, reducción de recursos de energía eléctrica y agua son factores que influyen en el mejoramiento de la productividad. A medida que se fortalece el plan de producción disminuirán estos rubros.

Los controles de calidad en producción aumentaron, actualmente se realiza al 100 % los controles en producción asegurando la calidad en las mejoras de los procesos y operaciones. El cumplimiento de estos controles asegura la renovación de licencia sanitaria extendida por el Ministerio de Salud y el cumplimiento de criterios mayores en la certificación de Buenas Prácticas de Manufactura.

6.1.5. Gestión de producto terminado

El costo financiero por tener producto terminado en bodegas disminuyó en relación al 2014 dado que el exceso en los productos con menor demanda estaba ocasionando pérdidas.

La productividad aumentó en relación a la gestión de producto terminado disminuyendo los costos por producto en bodegas, costos por producto obsoleto aprovechando el espacio en bodegas para producto terminado con mayor demanda.

La aplicación del pronóstico de ventas aumentó la productividad en relación al desperdicio que existía de desecho de producto terminado debido al no cumplimiento de análisis fisicoquímicos como parte de la cuarentena exigida por el Ministerio de Salud.

6.1.6. Ventas

Las ventas de la planta productora aumentaron en un 10 %, estos datos se basaron en la comparación de ventas de octubre del 2014 a marzo del 2015 tiempo en el cual está implementado el pronóstico después de la simulación.

Lo ideal es cumplir un año de pronóstico de ventas y verificar el crecimiento mensual de las mismas, éste debe aumentar gradualmente a partir del 10 %. Se espera que el incremento sea mayor, en relación a la capacidad de respuesta de la empresa porque puede negociar con mayor margen de confianza.

Otro factor influyente en el aumento de ventas es la credibilidad del vendedor, puesto que en el caso se ofrezca el producto y cumpla con la demanda la negociación se dará y generará confianza en el cliente.

6.1.7. Distribución

Los tiempos de distribución de parte de la empresa externa que distribuye en el interior de la república disminuyeron, el producto está justo a tiempo para darles salida en base a los pedidos de clientes. Cuando no existía producto en bodegas para despachar el producto era producido en caliente ocasionando retrasos en el tiempo de espera de los transportistas.

El producto actualmente está en tiempo con los distribuidores y puntos de venta a nivel nacional.

6.1.8. Satisfacción al cliente

La productividad en relación a la satisfacción del cliente se cuantificó con base a dos puntos:

- Entrega en tiempo de lo requerido por el cliente 1 punto
- Entrega en cantidad de lo requerido por el cliente 1 punto

El indicador de que la productividad aumentó en relación a la satisfacción del cliente será la suma de ambos lo que dará 2 puntos por producto por presentación. En los meses de enero, febrero y marzo del 2015 estos indicadores están con el valor de 2.

El rechazo de productos por parte del cliente debido a productos en mal estado ha disminuido, ya que se cumplen con los análisis fisicoquímicos, homogeneidad, color, olor de los productos en base a las especificaciones establecidas por el Regente externo. Esto genera satisfacción en el cliente dado que recibe el producto esperado.

6.2. Ventajas de la herramienta ODM

La PYME actualmente utiliza una base de datos que contiene los datos de productos, presentaciones, clientes, vendedores, facturación. Después del análisis del retorno sobre la inversión con base en el costo total de la propiedad del proyecto (servidor, base de datos, herramienta de minería de datos, alta disponibilidad, instalación, configuración, programación del modelo, capacitación, soporte técnico y asesoría, derechos actualizaciones, productividad, facilidad para realizar cambios) se decidió utilizar ODM, porque

presenta un ROI mayor en relación a otros. Un efecto colateral que no se había predicho fue el control y la mejora en la eficiencia de los procesos.

El aporte de la implementación de ODM genera valor en los procesos del negocio, ya que al implementar el mejor método de pronóstico GLM de ventas disminuirán las debilidades de la cadena de valor descritas en el capítulo 2 Diagnóstico situacional, esto es, desde la proyección de ventas hasta la distribución del producto terminado.

La influencia de tener un adecuado pronóstico de ventas en los procesos es determinante, puesto que al no tenerlo, entre los efectos negativos está que perjudicará el proceso central de la cadena de valor “satisfacción del cliente”. El valor de la herramienta depende de los beneficios que aporta a la cadena de valor.

6.3. Beneficios o bondades de la implementación de ODM

- La capacitación no es necesaria dado que el administrador de la base de datos tiene los conocimientos para operar la herramienta.
- El costo del proyecto es \$15000.00, actualmente la planta productora tiene implementado *software de Oracle* para gestionar la base de datos de productos, vendedores, departamentos, facturación.
- El tiempo de detener la operación de la planta es 0, porque es una herramienta conocida por el implementador.
- Los costos de los productos no se ven afectados debido a la implementación de esta herramienta.

- Facilidad de actualización.
- Software confiable.
- Otro beneficio de ODM es el aprovechamiento de generar.
- Análisis de compras de clientes por:
 - Departamento
 - Región
 - Producto
 - Presentación
 - Zona
 - Vendedor
- Fidelidad y fraude de clientes.
- Identificar a los mejores clientes.
- Descubrir atributos que afecten KPI.
- Determina probabilidades de clientes.
- Segmentación de clientes.
- Amigable debido a la interfaz gráfica.

Con el aumento en ingresos y reducción de costos, la inversión en el proyecto se recuperará en 6 meses a partir de la implementación que se realizó en octubre de 2014.

En el caso de la planta productora, la implementación del pronóstico de ventas a través de ODM es lo más aconsejable, ya que aunque es un *software* certificado de paga tiene otras ventajas como soporte técnico, actualizaciones, capacitación, actualmente está implementada la base de datos en *Oracle* edición *Express*.

CONCLUSIONES

1. La simulación con *Data Mining* determinó que el mejor método para implementar y encontrar el pronóstico de ventas que mejore la productividad es el de regresión y algoritmo GLM (*General Linear Models*).
2. El análisis FODA de la cadena de valor determinó que el factor común en las amenazas y debilidades es por la falta de un pronóstico de ventas, resultando como el incumplimiento de la satisfacción del cliente. En el caso de las fortalezas y oportunidades existe compromiso de parte del recurso humano y de la Gerencia General.
3. Los datos ingresados para realizar la simulación fueron los relacionados con el producto, presentación, costo, precio de venta, docenas vendidas, unidades vendidas.
4. La cantidad de datos simulados fue de 27,125 registros se concluyeron después de la des normalización de datos y la preparación previa.
5. Los indicadores obtenidos en la simulación aseguran que el algoritmo GLM es el más confiable para aplicarse con un porcentaje de confianza del 99.8627 % y un error de promedio 0.0283.
6. La mejora de la productividad después de implementar el pronóstico de ventas, se generó a través del conjunto de resultados positivos en factores productivos, indicador de satisfacción del cliente con valor 2 por

cumplimiento en tiempo y cantidad solicitada, aumento de 10 % en ventas y reducción del 8 % en el costo del producto.

7. Los factores económicos financieros, gestión de procesos y gestión de recursos humanos son indispensables para medir la productividad. El factor humano no fue analizado en este trabajo de graduación.

RECOMENDACIONES

1. Implementar el modelo de regresión y algoritmo GLM (*Generalized Linear Models*) para encontrar el mejor pronóstico de ventas, lo cual contribuirá en la mejora de la productividad en la planta productora.
2. La mejora continua de la implementación del algoritmo GLM, debe ser realizada en la cadena de valor cada año, analizando las fortalezas, amenazas, oportunidades y debilidades, debido a que factores externos podrían afectar el pronóstico.
3. Implementar la estrategia FO, ya que el compromiso del personal, proveedores y gerencia incrementarán la oportunidad de mejorar los procesos de inventarios, a través de la implementación de ODM.
4. Implementar la táctica defensiva DA, a través de la reducción de una falta de proyección de ventas pronosticando y evitando de esta manera la pérdida de clientes potenciales
5. Aprovechar la estrategia FA, disminuyendo la fidelidad de cliente, a través la fortaleza del compromiso de parte de gerencia general y talento humano de la planta productora.
6. Aprovechar el análisis DO, atacando la debilidad de la falta de proyección de ventas a través de la implementación de ODM y de esta manera mejorar la productividad y la proyección de ventas.

7. El cálculo de pronósticos es recomendable elaborarlos por año para evaluar flujos de efectivo al momento de importar materia prima, para determinar la capacidad de la planta y el compromiso comercial. Se sugiere realizarlo mensualmente para hacer una proyección ajustada a los cambios que se pueden dar en el mercado, y semanal para validar el plan de producción.
8. La versatilidad de esta herramienta ODM, permite pronosticar ventas por cliente, región, vendedor por lo que se recomienda hacer uso de ODM para dar un seguimiento personalizado y puntual. Sacar provecho de otras aplicaciones de ODM como fraude de clientes, fidelización del cliente, entre otros.
9. En la evaluación de la productividad es importante tomar en cuenta el recurso o talento humano, factor no fue analizado en su totalidad en este trabajo de graduación se recomienda a la empresa implementarlo a través de la evaluación del desempeño, unidades producidas por operario. Analizar indicadores de rotación de personal, accidentes en planta, entre otros.

BIBLIOGRAFÍA Y REFERENCIAS

1. Arbós, L. C. (2011). *Organización de la producción y dirección de operaciones*. Madrid: Ediciones Diaz.
2. Arnau, J. (1996). *Métodos y técnicas avanzadas de análisis de datos en ciencias del comportamiento*. Barcelona: Ediciones Universitat de Barcelona.
3. Boland, L., Carro, F., Stancatti, M., & Gismano, Y. (2007). *Funciones de la Administración, Teoría y práctica*. Buenos Aires, Argentina: UNS.
4. Braga, L. P. (2009). *Introducción a la Minería de Datos*. Brasil: Epapers.
5. Brown, C. (1998). *Applied Multivariate Statistics in Geohydrology and Related Sciences*. Alemania: Springer Berlin Heidelberg.
6. Coronel, C., Morris, S., & Rob, P. (2011). *Bases de Datos diseño, implementación y administración*. México, D.F.: Cengage Learning.
7. Dutka, M. A. (1998). *Manual de AMA para la satisfacción del cliente*. Buenos Aires, Argentina: Ediciones Granica, S.A.
8. Fernández, R. (2010). *La mejora de la productividad en la pequeña y mediana empresa*. España: Club Universitario.

9. Gonzalez, M. (2006). *Gestión de la producción, cómo planificar y controlar la producción industrial*. España: Ideaspropias.
10. Gonzalez, R. (2012). *Impacto de la Data Warehouse e Inteligencia de Negocios en el desempeño de las Empresas (Tesis de Doctorado)*. Peru: Universidad Ramon Lull.
11. Han, J., Micheline, K., & Pei, J. (2012). *Data Mining Concepts and Techniques*. United States of America: Elsevier.
12. Hanke, J., & Wichern, D. (2006). *Pronósticos en los negocios*. México: Pearson.
13. Heizer, J., & Render, B. (2004). *Principios de Administración de Operaciones*. México: Pearson.
14. Ian H. Witten, Eibe Frank, Mark A. Hall. (2011). *Data Mining Practical Machine Learning Tools and Techniques*. USA: Morgan Kauffman.
15. Krajewski, L., Rizman, J., & Larry, P. (2000). *Administración de operaciones: estrategia y análisis*. México: Pearson Educación.
16. López, C. P. (2007). *Minería de Datos*. España: Editorial Area Universitaria.
17. Lopez, V. R. (2008). *Gestión eficaz de los procesos productivos*. España: Especial Directivos.

18. Moreno, J. S. (2010). *Procedimientos estadísticos*. Madrid: ESIC.
19. Palma, C., Palma, W., & Pérez, R. (2009). *Data Mining, El arte de anticipar*. Santiago: RIL editores.
20. Perez, J. A. (1999). *Gestión de la calidad orientada a los procesos*. Madrid, España: Esic Editorial.
21. Puente, C. V. (1995). *Un guía para la investigación*. Madrid: Complutense.
22. Rodriguez, A. M. (2007). *Tools and Functionalities of Oracle Data Mining*. *Revista Técnica de la Empresa de Telecomunicaciones de Cuba, S.A.*, 110-115.
23. Ruiz, G. E., Fuentes, M. M., & Alcaraz, J. (2006). *Políticas de Marketing*. Madrid, España: Paraninfo.
24. Singer, M., & Donoso, P. (2007). *Internal supply chain management in the Chilean sawmill industry*. Chile: Pontificia Universidad Católica de Chile.
25. Taft, M., Krishnan, R., Hornick, M., & Muhkin, D. (2005). *Oracle Data Mining Concepts*. United States: Oracle.
26. Torre, J. O. (1999). *Conceptos generales de productividad, sistemas, normalización y competitividad para la pequeña y mediana empresa*. Mexico: Universidad Iberoamericana.

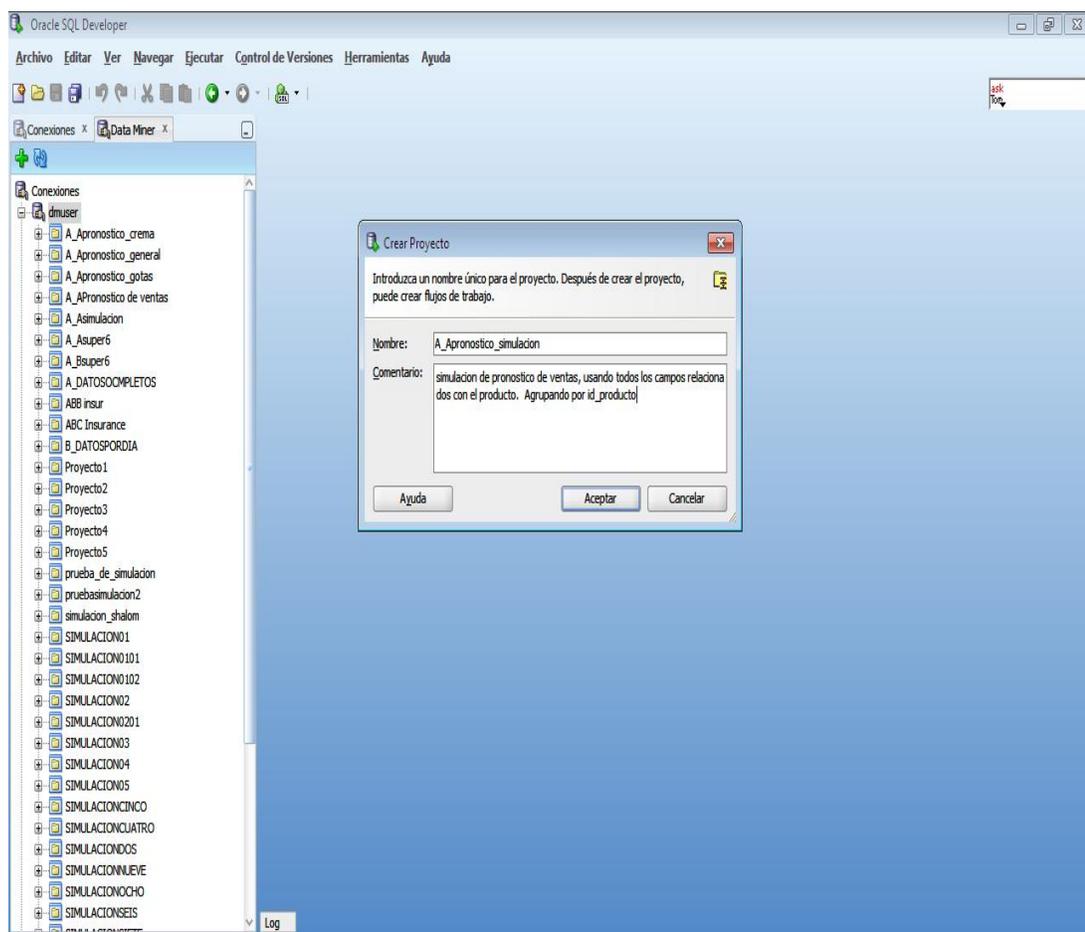
27. Velaso, J. A. (2010). *Gestión por procesos*. España: Graficas Dehon.

28. Witte, I., & Frank, E. (2005). *Data Mining Practical Machine Learning Tools and Techniques*. United States of America: Elsevier.

29. Zeithmal, V. (1990). *Delivering Quality Service*. USA: MacMillan.

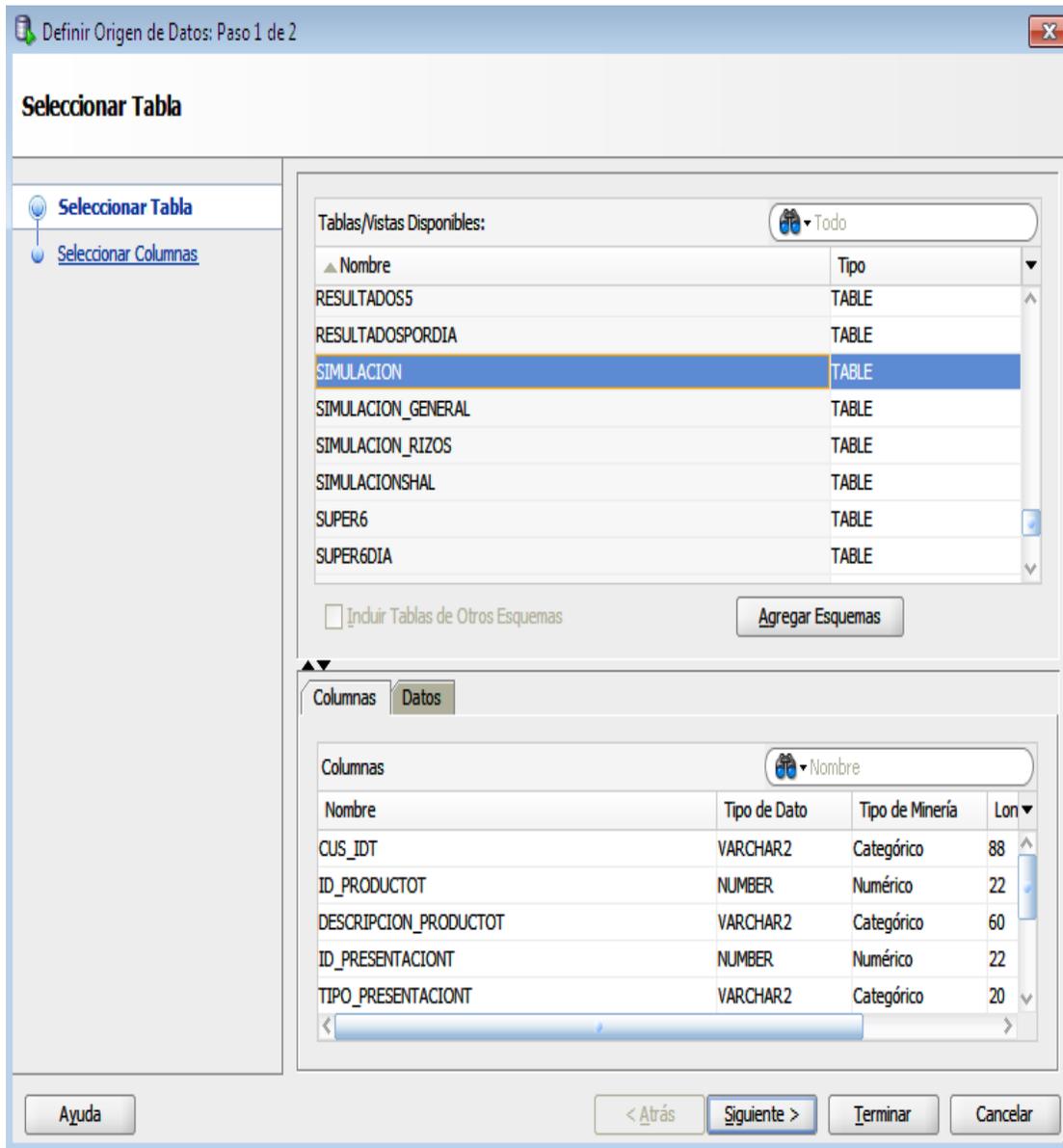
APÉNDICES

Figura a. Creación del proyecto de simulación



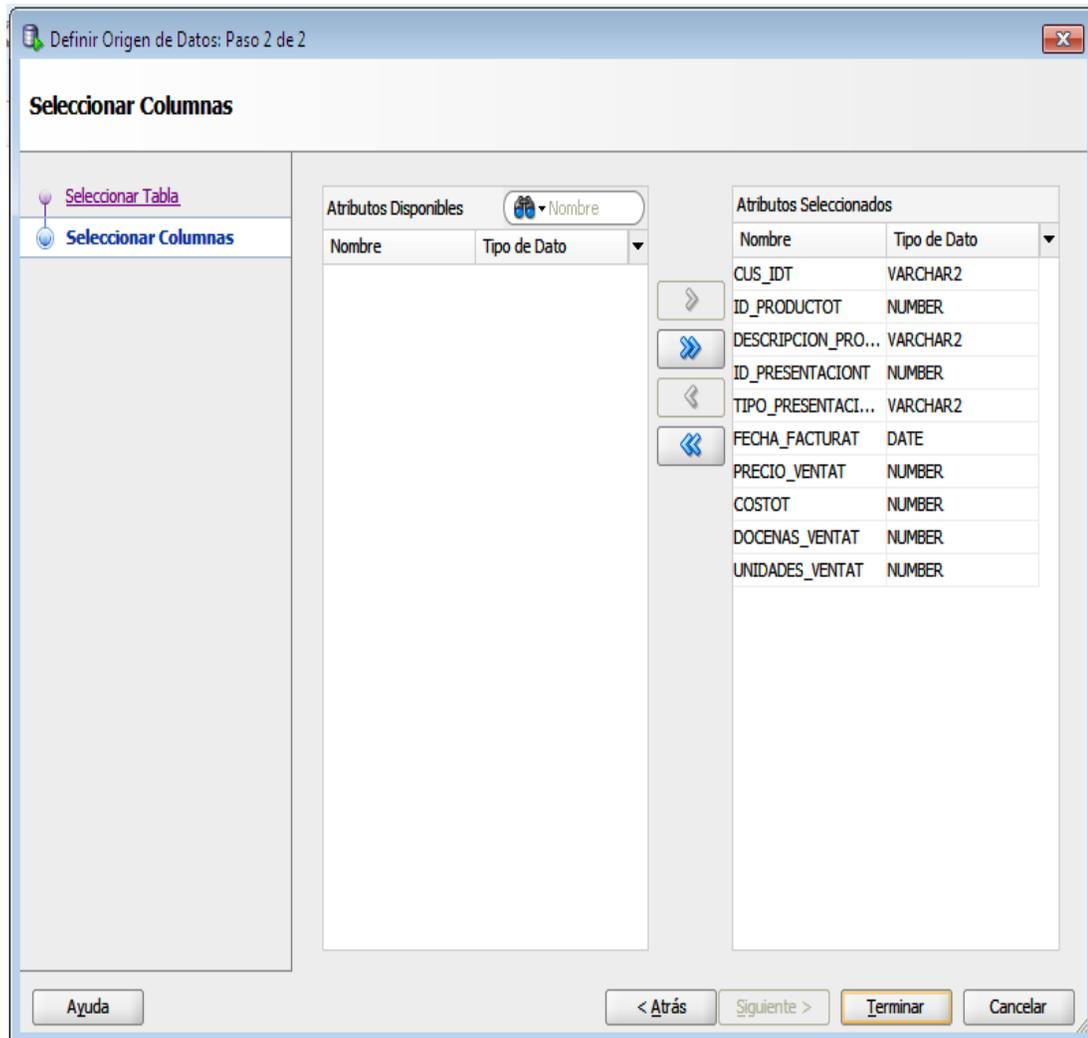
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura b. **Tabla que contiene datos preparados para la simulación**



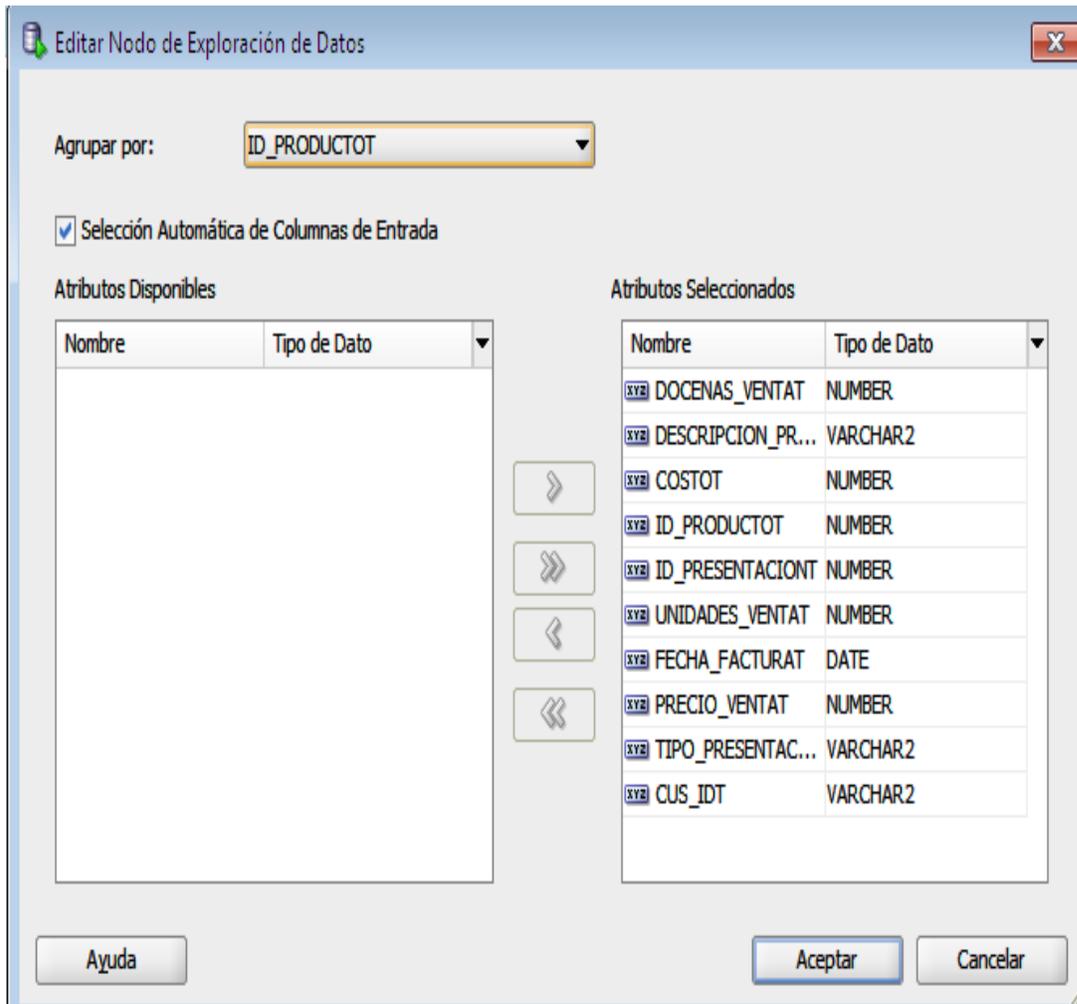
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura c. Datos preparados para la simulación



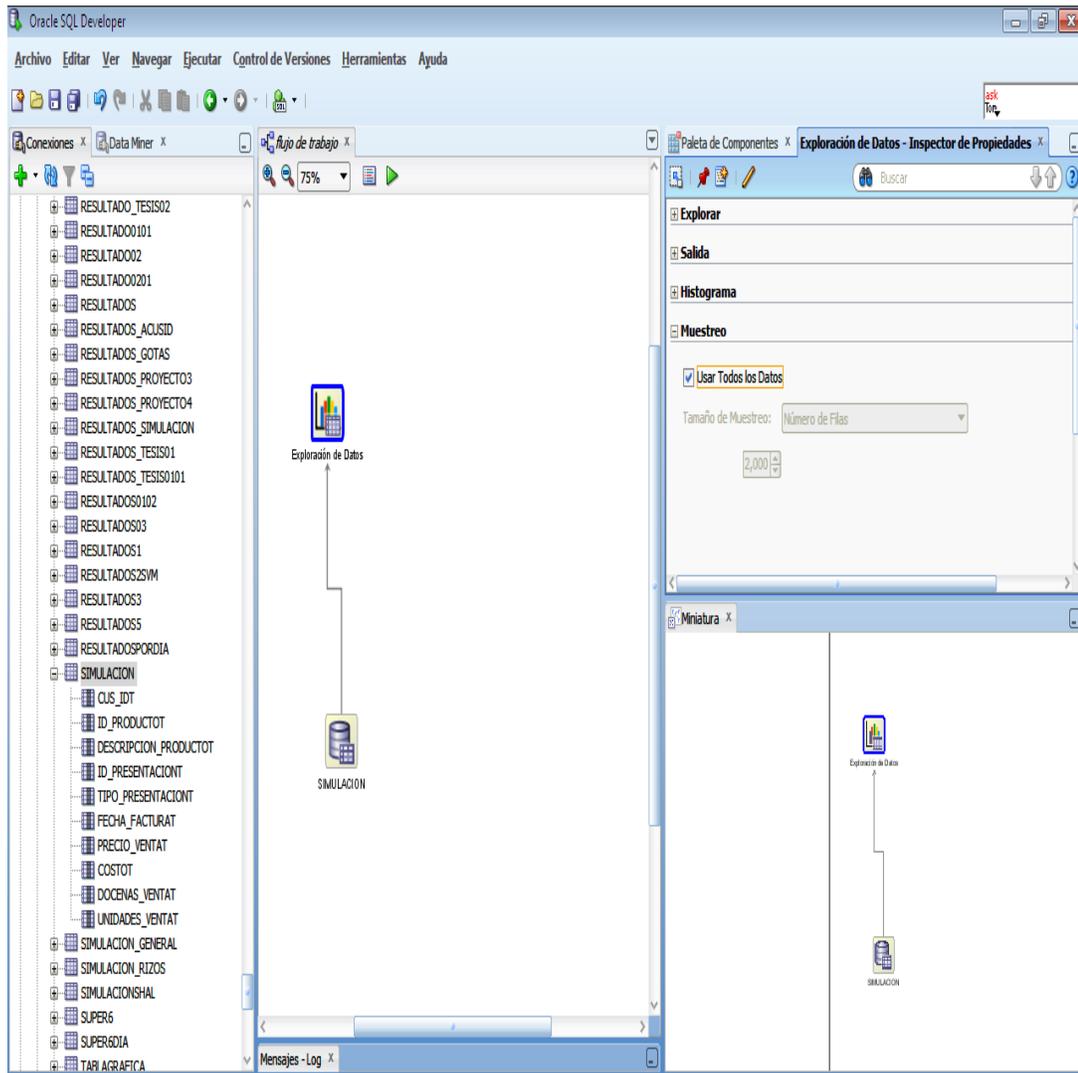
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura d. Agrupacion de datos por id_producto



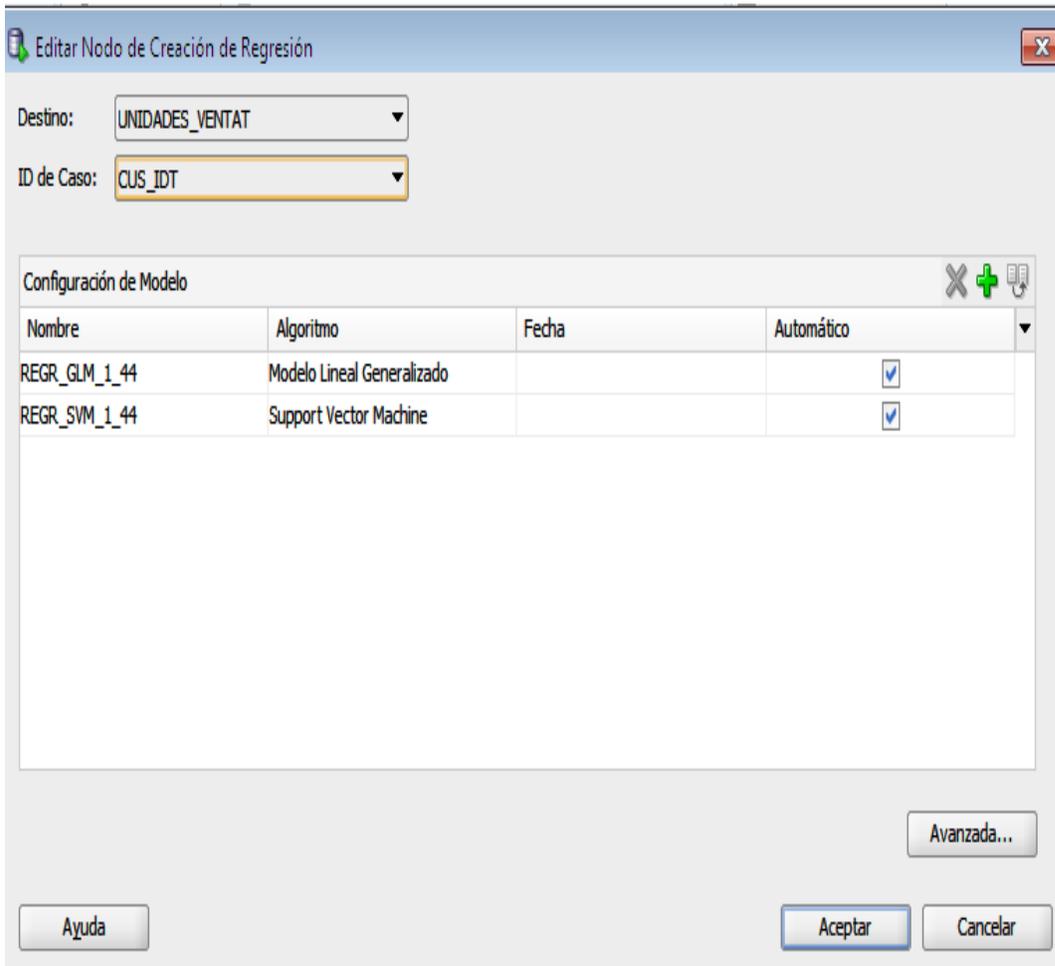
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura e. Creación de nodos para explorar todos los datos



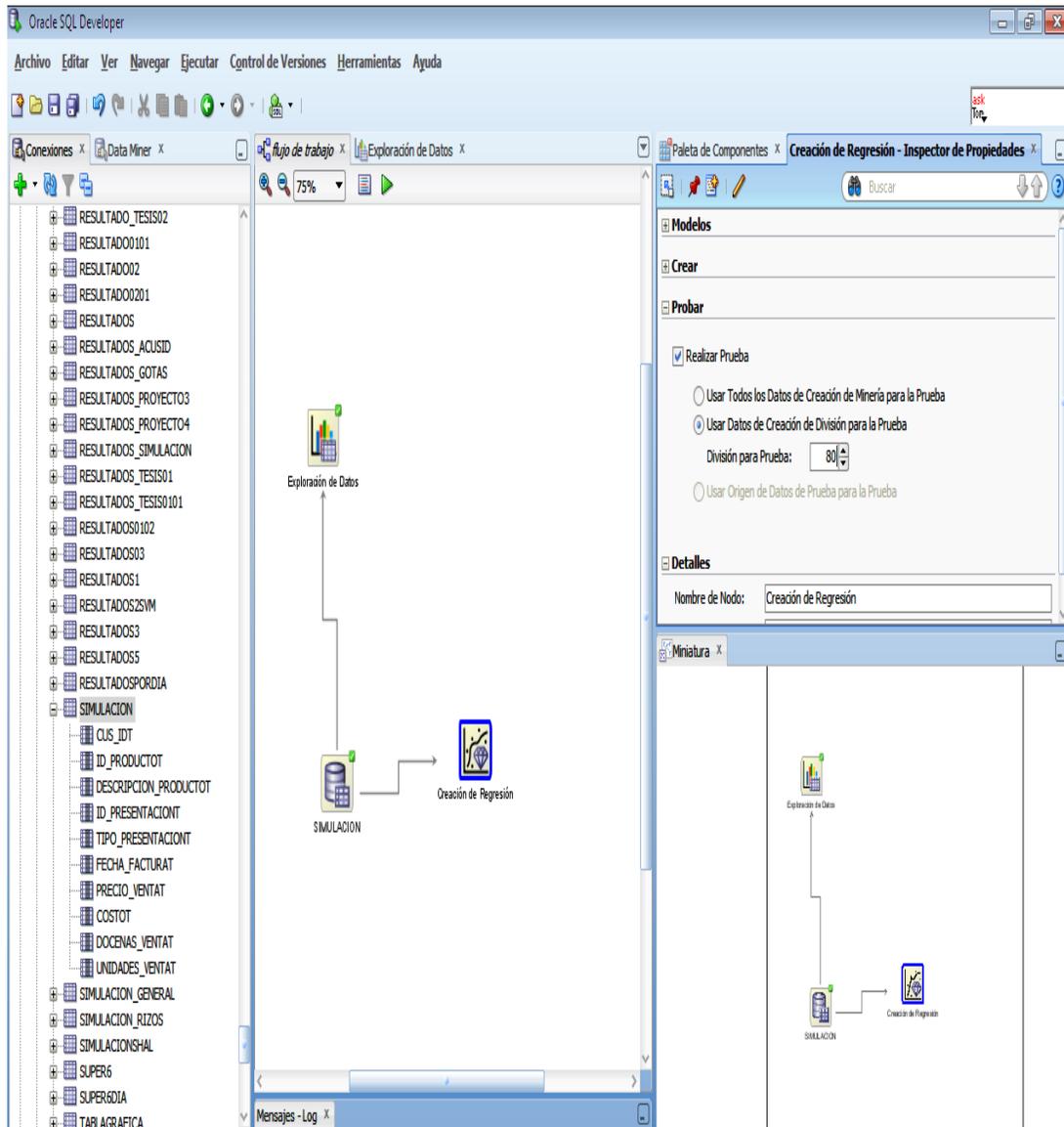
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura f. **Nodo de la creación de regresión**



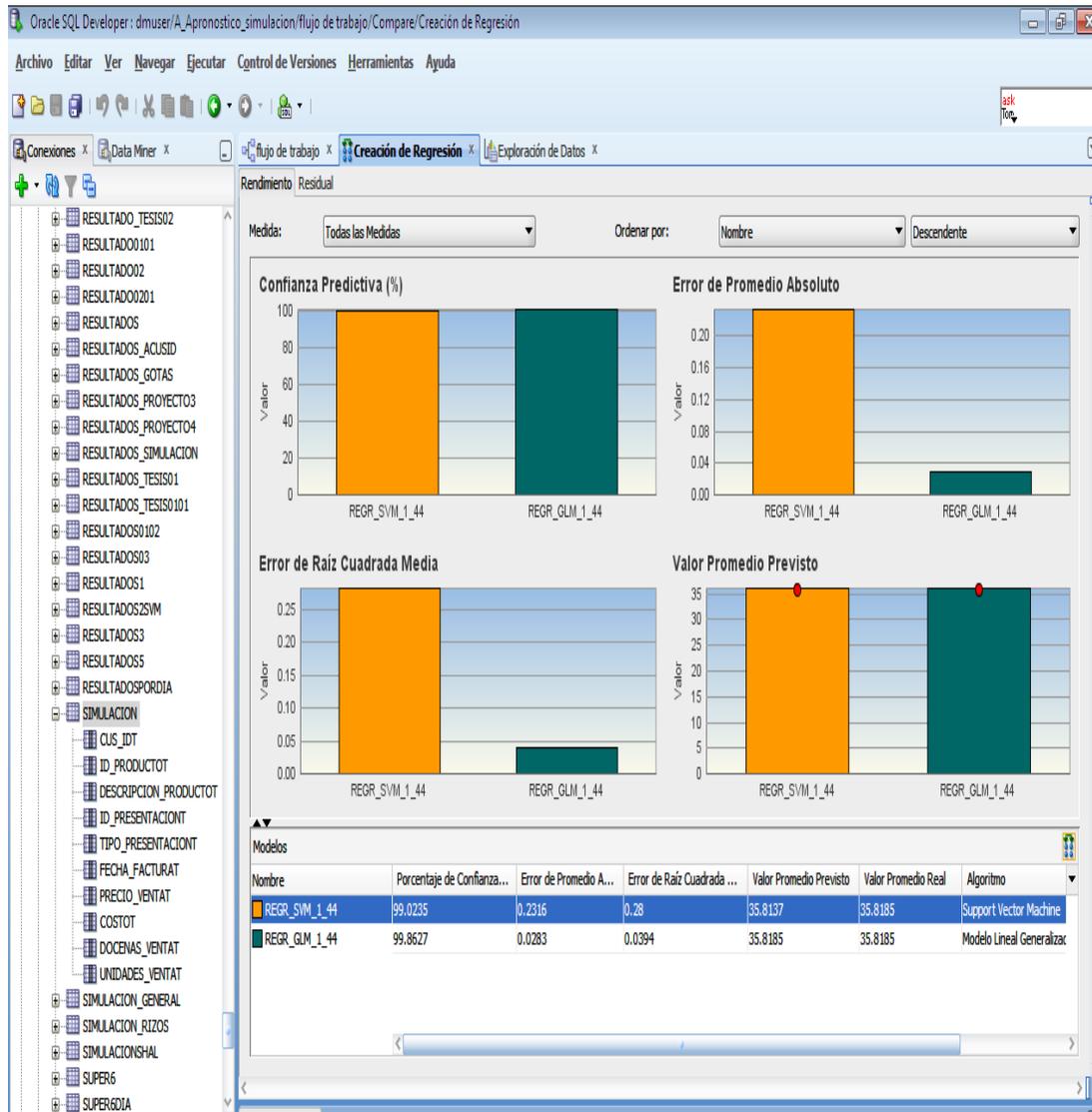
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura g. Creación de la división de prueba 80-20



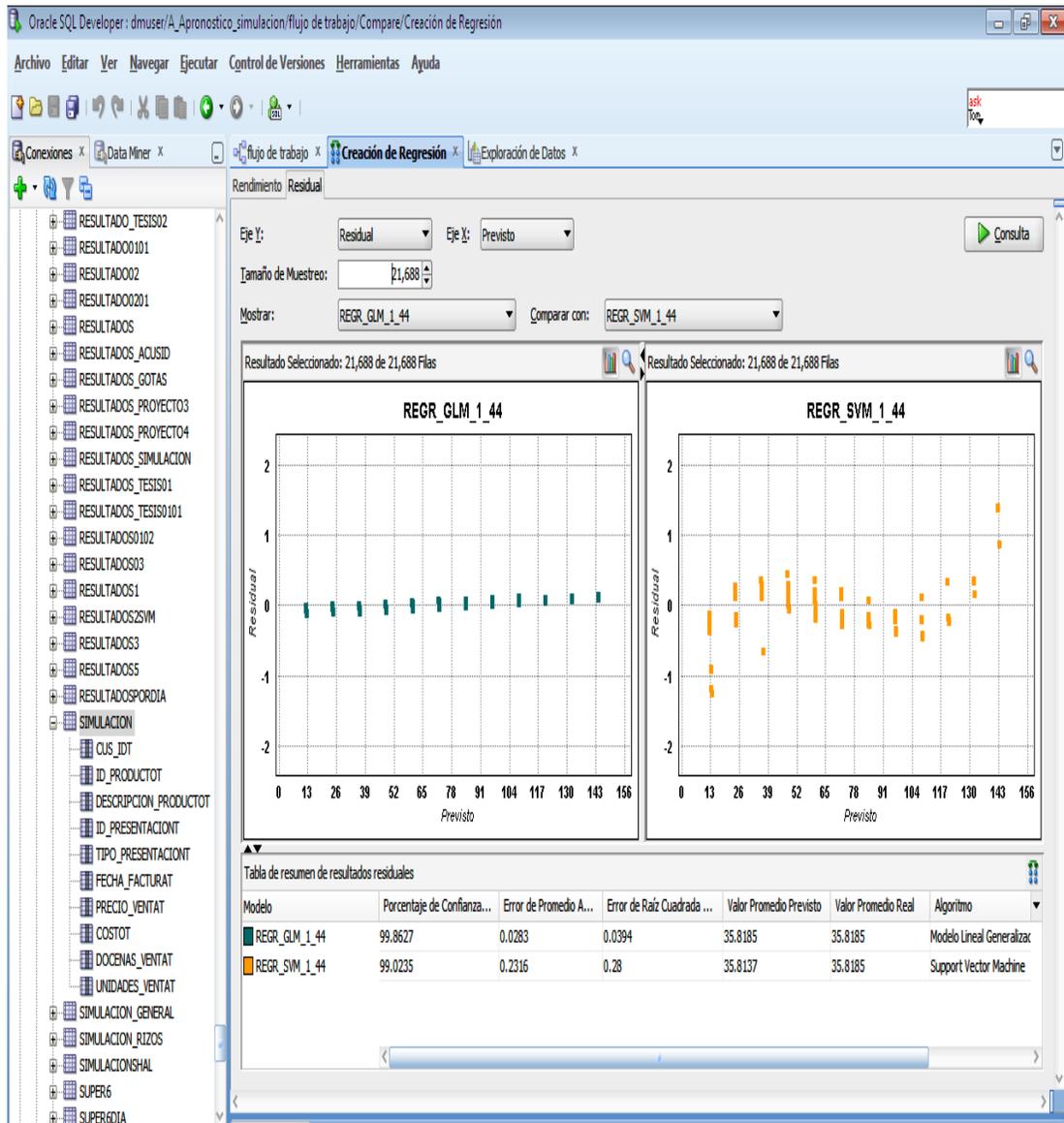
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura h. **Comparación de rendimiento para determinar el mejor modelo**



Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura i. Comparación residual para determinar el mejor modelo



Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura j. Detalles del modelo GLM

Oracle SQL Developer: dmuser/A_Apronostico_simulacion/flujo de trabajo/Creación de Regresión/REGR_GLM_1_44

Archivo Editar Ver Navegar Ejecutar Control de Versiones Herramientas Ayuda

Conexiones x Data Miner x flujo de trabajo x REGR_GLM_1_44 x Creación de Regresión x

Detalles Coeficientes Configuración

Nombre	Valor
ADJUSTED_R_SQUARE	0.99999818
AIC	-35348.33657508
COEFF_VAR	0.10859182
CORRECTED_TOTAL_DF	5436
CORRECTED_TOT_SS	4458528.9004966
DEPENDENT_MEAN	35.5696156
ERROR_DF	5403
ERROR_MEAN_SQUARE	0.00149194
ERROR_SUM_SQUARES	8.06097469
F_VALUE	90377008.5460391
GMSEP	0.00150133
HOCKING_SP	0.00000028
J_P	0.00150127
MODEL_CONVERGED	Sí
MODEL_DF	33
MODEL_F_P_VALUE	0
MODEL_MEAN_SQUARE	134837.45664387
MODEL_SUM_SQUARES	4449636.06924778
NUM_PARAMS	34
NUM_ROWS	5437
ROOT_MEAN_SQ	0.03862569
R_SQ	0.99999819
SBIC	-35123.9031627
VALID_COVARIANCE_MATRIX	No

Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura k. Coeficientes del modelo GLM

Oracle SQL Developer: dmuser/A_Apronostico_simulacion/flujo de trabajo/Creación de Regresión/REGR_GLM_1_44

Archivo Editar Ver Navegar Ejecutar Control de Versiones Herramientas Ayuda

Conexiones x Data Miner x Flujo de trabajo x REGR_GLM_1_44 x Creación de Regresión x

Detalles Coeficientes Configuración

Ordenar por Valor Absoluto Tamaño de Recuper... 22,000 Consulta

Coeficientes: 34 de 34

Atributo	Valor	Coefficiente	Coefficiente Estandarizado
DOCENAS_VENTAT		11.97824541	998.18712
<Interceptar >		.11969634	0
DESCRIPCION_PRODUCTOT	Gel para peinar	-.07575988	-.00055093
DESCRIPCION_PRODUCTOT	Crema para peinar	-.06770698	-.00048951
DESCRIPCION_PRODUCTOT	Ampolla berro	.04924984	.00043689
TIPO_PRESENTACIONT	2 onzas	.03829982	.00038405
TIPO_PRESENTACIONT	4 onzas	-.0296362	-.0002996
TIPO_PRESENTACIONT	6 onzas	.02869265	.00034121
DESCRIPCION_PRODUCTOT	Tratamiento miel	-.02533299	-.00014996
TIPO_PRESENTACIONT	12 onzas	-.02319587	-.00012604
DESCRIPCION_PRODUCTOT	Ampolla Keratina	-.02307011	-.00017865
DESCRIPCION_PRODUCTOT	Drops	.02283953	.00022881
DESCRIPCION_PRODUCTOT	Tratamiento sabila	-.01924601	-.00009914
DESCRIPCION_PRODUCTOT	Ampolla miel	-.01914015	-.00016875
DESCRIPCION_PRODUCTOT	Shampoo horse	-.01788556	-.00008392
DESCRIPCION_PRODUCTOT	Acondicionador berro	-.01671082	-.00008031
DESCRIPCION_PRODUCTOT	Acondicionador placenta	-.01553603	-.00007035
TIPO_PRESENTACIONT	8 onzas	-.01545617	-.00010566
DESCRIPCION_PRODUCTOT	Acondicionador horse	-.01436132	-.00007092
DESCRIPCION_PRODUCTOT	Gotas	-.01435867	-.00014529
TIPO_PRESENTACIONT	20 onzas	-.01326007	-.00016429
TIPO_PRESENTACIONT	16 onzas	-.01209418	-.00014741
DESCRIPCION_PRODUCTOT	Tratamiento chocolate	-.01204488	-.00007744
DESCRIPCION_PRODUCTOT	Tratamiento keratina	-.01174312	-.00007014
DESCRIPCION_PRODUCTOT	Shampoo placenta	-.01144687	-.00006694

Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura I. Coeficientes del modelo GLM

Oracle SQL Developer: dmuser/A_Apronostico_simulacion/flujo de trabajo/Creación de Regresión/REGR_GLM_1_44

Archivo Editar Ver Navegar Ejecutar Control de Versiones Herramientas Ayuda

Conexiones x Data Miner x flujo de trabajo x REGR_GLM_1_44 x Creación de Regresión x

Detalles Coeficientes Configuración

Ordenar por Valor Absoluto Tamaño de Recuper... 22,000 Consulta

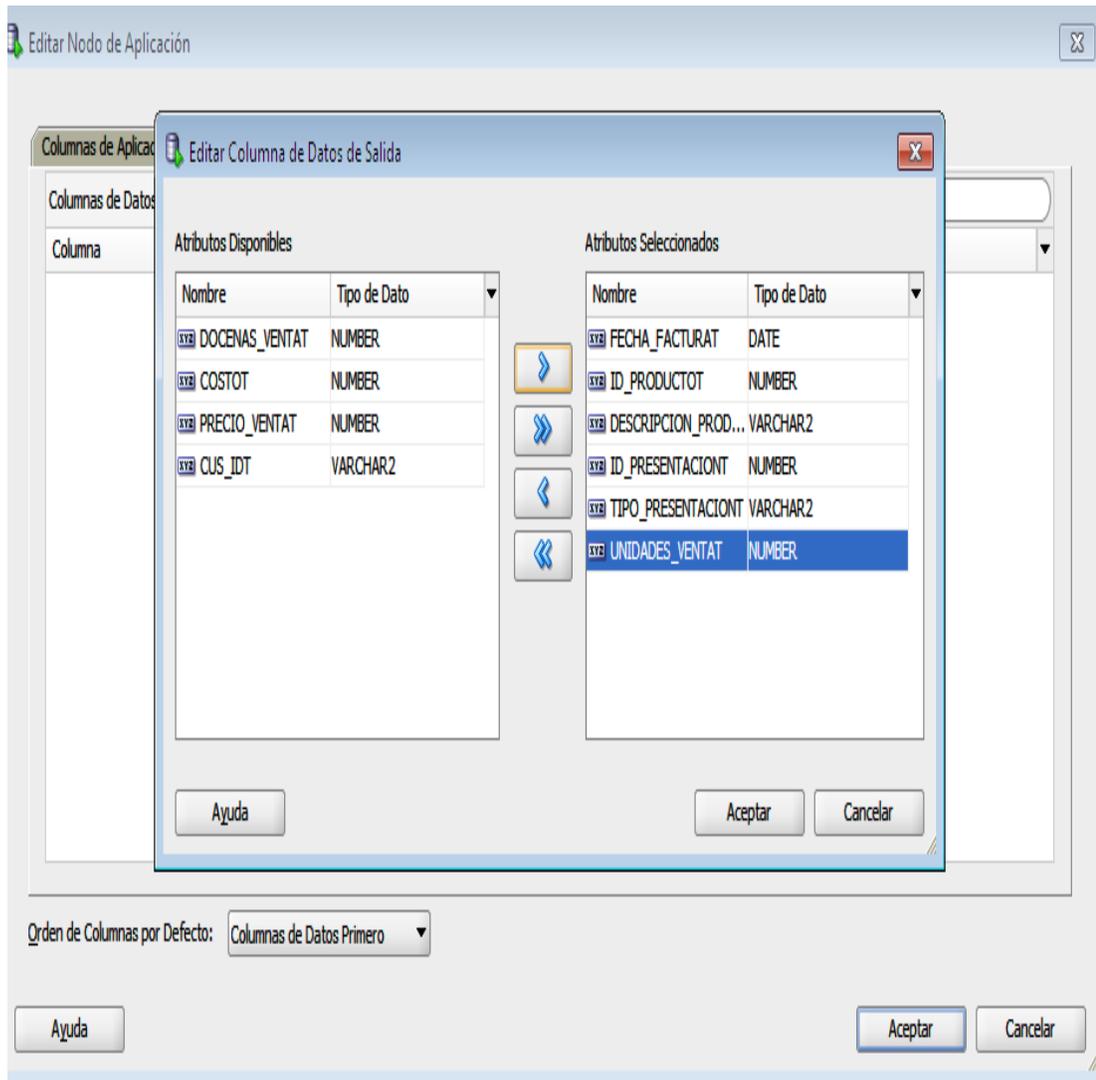
Coeficientes: 34 de 34

Atributo	Valor	Coeficiente	Coeficiente Estandarizado
TIPO_PRESENTACIONT	12 onzas	-0.02319587	-0.0012604
DESCRIPCION_PRODUCTOT	Ampolla Keratina	-0.02307011	-0.00127865
DESCRIPCION_PRODUCTOT	Drops	.02283953	.0002281
DESCRIPCION_PRODUCTOT	Tratamiento sabila	-0.01924601	-0.0009914
DESCRIPCION_PRODUCTOT	Ampolla miel	-0.01914015	-0.0016875
DESCRIPCION_PRODUCTOT	Shampoo horse	-0.01788556	-0.0008392
DESCRIPCION_PRODUCTOT	Acondicionador berro	-0.01671082	-0.0008031
DESCRIPCION_PRODUCTOT	Acondicionador placenta	-0.01553603	-0.0007035
TIPO_PRESENTACIONT	8 onzas	-0.01545617	-0.0010366
DESCRIPCION_PRODUCTOT	Acondicionador horse	-0.01436132	-0.0007092
DESCRIPCION_PRODUCTOT	Gotas	-0.01435867	-0.0014529
TIPO_PRESENTACIONT	20 onzas	-0.01326007	-0.0016429
TIPO_PRESENTACIONT	16 onzas	-0.01209418	-0.0014741
DESCRIPCION_PRODUCTOT	Tratamiento chocolate	-0.01204488	-0.0007744
DESCRIPCION_PRODUCTOT	Tratamiento keratina	-0.01174312	-0.0007014
DESCRIPCION_PRODUCTOT	Shampoo placenta	-0.01144687	-0.0006694
TIPO_PRESENTACIONT	22 ml	-0.01135481	-0.0014644
DESCRIPCION_PRODUCTOT	Ampolla placenta	-0.00971408	-0.0008596
DESCRIPCION_PRODUCTOT	Tratamiento berro	.00774907	.00004656
ID_PRESENTACIONT		-0.00231571	-0.00021503
DESCRIPCION_PRODUCTOT	Ampolla Sabila	.00145585	.00001298
ID_PRODUCTOT		-0.00117591	-0.00035422
DESCRIPCION_PRODUCTOT	Shampoo berro	.0009707	.00000647
COSTOT		-0.00051728	-0.00014114
PRECIO_VENTAT		-0.00029278	-0.00012143

Mensajes -Log x

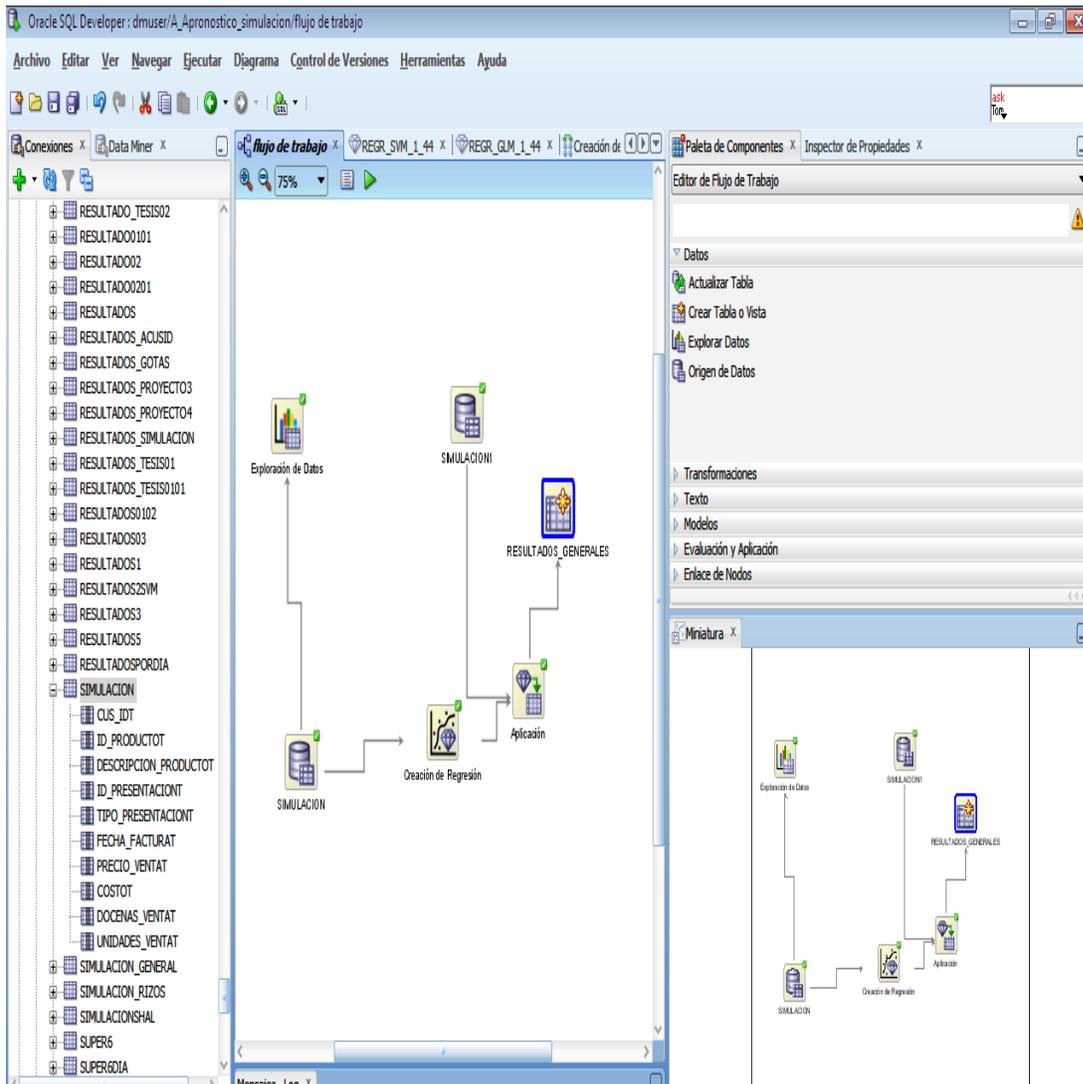
Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura m. **Columnas de salida del modelo GLM**



Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura n. Flujo de trabajo con nodos de simulacion del modelo GLM



Fuente: elaboración propia, con base en la herramienta *Oracle Data Mining*.

Figura o. Resultados de simulacion del modelo GLM

REGR_GLM_1_44_PRED	FECHA_F...	ID_PRODUCTOT	DESCRIPCION_PRODUCTOT	ID_PRESENTACIONT	TIPO_PRESENTACIONT	UNIDADES_VENTAT
1	71.986	12/01/09...	5 Gotas		1.2 onzas	72
2	59.9454	12/01/09...	1 Crema para rizos		5.12 onzas	60
3	107.8755	12/01/09...	1 Crema para rizos		4.8 onzas	108
4	131.8838	12/01/09...	1 Crema para rizos		3.6 onzas	132
5	23.9997	12/01/09...	9 Gel para peinar		3.6 onzas	24
6	95.9766	12/01/09...	7 Drops		1.2 onzas	96
7	36.0054	12/01/09...	7 Drops		2.4 onzas	36
8	35.9887	12/01/09...	2 Crema para peinar		3.6 onzas	36
9	47.9495	12/01/09...	5 Gotas		2.4 onzas	48
10	24.0124	12/01/09...	18 Ampolla placentia		7.22 ml	24
11	47.9567	12/01/09...	17 Ampolla Keratina		7.22 ml	48
12	83.9183	12/01/09...	15 Ampolla Sabila		7.22 ml	84
13	107.9214	12/01/09...	16 Ampolla berro		7.22 ml	108
14	12.0235	12/01/09...	19 Ampolla miel		7.22 ml	12
15	71.9436	12/01/09...	15 Ampolla Sabila		8.44 ml	72
16	47.9619	12/01/09...	19 Ampolla miel		8.44 ml	48
17	35.982	12/01/09...	17 Ampolla Keratina		8.44 ml	36
18	95.9468	12/01/09...	16 Ampolla berro		8.44 ml	96
19	47.9725	12/01/09...	18 Ampolla placentia		8.44 ml	48
20	12.0219	12/01/09...	12 Tratamiento chocolate		6.20 onzas	12
21	35.9798	12/01/09...	11 Tratamiento keratina		6.20 onzas	36
22	59.957	12/01/09...	10 Tratamiento berro		6.20 onzas	60
23	35.97	12/01/09...	13 Tratamiento sabila		6.20 onzas	36
24	12	12/01/09...	29 Shampoo horse		9.16 onzas	12
25	23.9859	12/01/09...	28 Shampoo placentia		9.16 onzas	24

Fuente: elaboración propia, con base en la herramienta Oracle Data Mining.

GLOSARIO

Demanda	Cantidad y calidad de bienes y servicios que pueden ser adquiridos a los diferentes precios del mercado por un consumidor.
JAVA	Lenguaje de PPO, usado en la plataforma Java la cual es capaz de ejecutar aplicaciones.
ODM	<i>Oracle Data Mining</i> , herramienta desarrollada por Oracle.
POO	Programación orientada a objetos.
Simulación	Técnica numérica para conducir experimentos en una computadora.