

ANÁLISIS DE PRODUCCIÓN AVÍCOLA MEDIANTE TÉCNICAS DE INTELIGENCIA DE NEGOCIOS Y MINERÍA DE DATOS

ANALYSIS OF AVICULTURAL PRODUCTION THROUGH BUSINESS INTELLIGENCE AND DATA MINING TECHNIQUES

BERTHA MAZON-OLIVO¹, WILMER RIVAS-ASANZA², JOHNNY NOVILLO-VICUÑA³, CINTHIA FLORES-CABRERA⁴

1 Universidad Técnica de Machala, Ecuador. bmazon@utmachala.edu.ec

2 Universidad Técnica de Machala, Ecuador. wrvivas@utmachala.edu.ec

3 Universidad Técnica de Machala, Ecuador. jnovillo@utmachala.edu.ec

4 Universidad Técnica de Machala, Ecuador. cpflores_est@utmachala.edu.ec

RESUMEN

Las organizaciones siempre están buscando el mejoramiento u optimización de sus procesos para ofertar productos o servicios de calidad que satisfagan a sus clientes; y en consecuencia, lograr mayor rentabilidad y beneficio. En la actualidad, la Ciencia de Datos (DS), la Inteligencia de Negocios (BI) y la Minería de Datos (DM) son disciplinas que se centran en el análisis de datos para la toma de decisiones. El propósito de este trabajo es la implementación de una solución de BI y DM en ambiente web, enfocada en el análisis de los indicadores claves de desempeño (KPI's) de los procesos de producción y ventas de una empresa avícola ubicada en el cantón Pasaje - El Oro - Ecuador. La aplicación web denominada IncuAnalytic fue implementada según la metodología Cross Industry Standard Process for Data Mining (CRISP-DM) y mediante la utilización de las herramientas como Postgres SQL para el data warehouse (DW), Pentaho BI Server para el Procesamiento Analítico en Línea (OLAP) y el diseño de un tablero de control (dashboard), el lenguaje R y el framework Shiny para la implementación de los modelos predictivos con técnicas de regresión y series temporales. Los resultados se evidencian en IncuAnalytic, una herramienta que utiliza una data dinámica e histórica, y que transparenta la complejidad de las técnicas de análisis de datos al administrador o dueño de la empresa; permitiéndole obtener de manera oportuna información actualizada de sus KPI's o simplemente hacer predicciones que orienten la toma de decisiones.

PALABRAS CLAVE: Inteligencia de Negocios, Minería de Datos, regresión, series temporales, producción avícola.

ABSTRACT

The organizations are always looking for the improvement or optimization of their processes to offer products or services of quality that satisfy their clients, and consequently, to achieve greater profitability and profit. At present, Data Science (DS), Business Intelligence (BI) and Data Mining (DM) are disciplines that focus on data analysis for decision making. The purpose of this work is the implementation of a BI and DM solution in a web environment, focused on the analysis of the key performance indicators (KPIs) of the production and sales processes of a poultry company located in Pasaje - El Oro - Ecuador. The web application called IncuAnalytic was implemented according to the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology and through the use of tools such as Postgres SQL for the DW, Pentaho BI Server for Online Analytical Processing (OLAP) and the design of a dashboard, the R language and the Shiny framework for the implementation of predictive models with regression techniques and time series. The results are evident in IncuAnalytic, a tool that uses a dynamic and historical data, and that transforms the complexity of the techniques of data analysis to the administrator or owner of the company, allowing him to obtain in a timely manner updated information of his KPIs or simply to do predictions that guide decision making.

KEYWORDS: Business Intelligence, Data Mining, regression, time series, poultry production.

DOI: <http://dx.doi.org/10.23878/alternativas.v19i2.203>

RECIBIDO: 18/10/2017

ACEPTADO: 21/2/2018

INTRODUCCIÓN

La BI comprende un conjunto de estrategias y componentes que permiten transformar los datos obtenidos de Sistemas de Procesamiento Transaccionales (SPT) u otras fuentes de diferentes formatos (estructurados, semi-estructurados y no estructurados). Coelho et al. (2016) señalan que la arquitectura BI está constituida por tres entornos esenciales: fuentes de datos, almacén de datos y herramientas de análisis.

En los últimos años, las empresas han formulado estrategias de negocios que les facilita la toma de decisiones basadas en la evaluación de datos históricos de sus procesos y mediante la identificación de sus KPI's. Los datos en crudo provenientes de varias fuentes son extraídos, transformados y cargados (ETL) en un nuevo DW; organizado a su vez, en temas (data marts) (Mazon-Olivo et al., 2017) para luego mediante técnicas de análisis acceder al DW y visualizar la información relevante, oportuna y en forma consolidada (Astriani & Trisminingsih, 2016; Lindberg, Tan, Yan, & Starfelt, 2015; Longo, Giacobelli, & Bochicchio, 2014; Meier, Lagemann, Morlock, & Rathmann, 2013).

Una aplicación puede estar conformada por varias técnicas de análisis descriptivo y predictivo de datos. Por ejemplo, el análisis multidimensional OLAP que sirve de base para la elaboración de paneles de control (dashboards), que a su vez son herramientas para la visualización estadística gráfica y descriptiva de los KPI's, formando un Sistema de Soporte de Decisiones (SSD) que permite un mayor nivel de comprensión de la información (Boulekruche, Jabeur, & Alimazighi, 2015; Gourbesville, Du, Zattero, & Ma, 2016; Gröger, Hillmann, Hahn, Mitschang, & Westkämper, 2013; Khan, Ehsan, Mirza, & Zahoor, 2012; Tokola, Gröger, Järvenpää, & Niemi, 2016; Tutunea, 2015; van Os, Herber, & Scholtens, 2014).

La DM también ofrece distintas estrategias para la obtención del conocimiento de bases de datos (KDD) aplicando técnicas de análisis para la detección de patrones ocultos, asociaciones, clasificaciones o diagnósticos; apoyándose en disciplinas como: Matemática, Estadística, Inteligencia Artificial y el aprendizaje automático (Machine Learning) (Femina & Sudheep, 2015; Hamrouni, Slimani, & Ben Charrada, 2015; Jothi, Rashid, & Husain, 2015). La DM también es considerada como el paso más importante en el proceso de descubrimiento de conocimiento potencialmente útil y comprensible basado en los datos, para establecer modelos tanto predictivos como descriptivos (Jothi et al., 2015; Kavakiotis et al., 2017).

Un modelo descriptivo generalmente no supervisado se enfoca en la búsqueda de patrones que describen los datos que pueden ser interpretados normalmente por los individuos. Por otra parte, los modelos predictivos aplican funciones de aprendizaje supervisado para predecir valores desconocidos o futuros a partir de variables de interés. Los algoritmos y modelos de DM pueden elaborarse en diferentes herramientas y lenguajes tales como: R, Python, Weka, Minitab, RapidMiner, etc. Sin embargo, existen aplicaciones BI que posibilitan la ejecución de dichas técnicas mediante la adaptación de pluggins; tal es el caso de Pentaho, Qlick, Power BI, Tableau, entre otras.

El objetivo de esta investigación es la implementación de una solución de BI y DM con interfaz web, que permita el análisis de los principales KPI's del proceso de producción y ventas de una empresa avícola, ubicada en el cantón Pasaje - El Oro - Ecuador. Cabe mencionar que en el Ecuador, el número de empresas dedicadas a la producción avícola se ha incrementado durante los últimos años. Es así como han obtenido cierto status competitivo respecto a las industrias de otra índole.

En la provincia de El Oro los productores avícolas se concentran en los exteriores de los cantones Balsas y Pasaje. No obstante, ninguna de estas empresas cuentan con un Sistema de Soporte de Decisiones (SSD) que les otorgue información respecto al avance actual y futuro del proceso productivo y de ventas; por lo que la toma de decisiones se torna complicada y con probabilidades de que las acciones ejecutadas conlleven al fracaso del negocio o amenacen su estabilidad y continuidad (Wessiani & Sarwoko, 2015).

Se utilizó la metodología CRISP-DM para el diseño de la aplicación BI denominada IncuAnalytic y fue implementada mediante las herramientas: Postgres SQL para el DW, Pentaho Data Integration para el proceso ETL, Pentaho BI Server para el análisis descriptivo multidimensional OLAP y un Dashboard que integra los KPI's.

Para efectos de análisis de datos predictivos, en este trabajo se han seleccionado las técnicas de regresión y series temporales. Se describe a la regresión como una técnica para la cuantificación de relaciones entre variables a través de una función matemática representada gráficamente por puntos dispersos cuya tendencia por lo general es de forma lineal, polinómica o logarítmica y requiere como complemento el desarrollo de un análisis de correlación para la

adecuada emisión de juicios de valor respecto a la solución planteada. Las series temporales son colecciones de observaciones obtenidas durante diferentes períodos de tiempo con lo que permite explicar el pasado y predecir el futuro respecto a las variables objeto de estudio.

Las técnicas de DM seleccionadas fueron implementadas utilizando el IDE RStudio con el lenguaje R, que incluye una consola, un editor y herramientas para la gestión del espacio de trabajo. Además, se utilizó el framework Shiny para la visualización de la información a través de un navegador web (Hermawati & Sitanggang, 2016). Shiny permite la construcción de aplicaciones web interactivas, consta de dos estructuras esenciales: un archivo server.R y un archivo ui.R, para instrucciones que construyen componentes R y la interfaz de usuario de la aplicación respectivamente (Hermawati & Sitanggang, 2016; Siknun & Sitanggang, 2016).

MATERIALES Y MÉTODOS

- **UBICACIÓN GEOGRÁFICA DONDE SE REALIZÓ LA INVESTIGACIÓN.** El estudio se desarrolló en las empresas productoras avícolas de la provincia de El Oro - Ecuador; concentrando la recogida de la data histórica en la empresa INCUPASAJE, que se encuentra ubicada en el cantón Pasaje. Se realizaron 1500 observaciones aproximadamente, referidas a los tres últimos años de los procesos de producción y ventas.
- **TIPO DE INVESTIGACIÓN.** Se trata de un estudio descriptivo, predictivo y aplicativo. Se empleó la metodología CRISP-DM para el diseño e implementación de un DW, la aplicación BI (Cubos OLAP y Dashboard) y DM (modelos predictivos de datos). Se utilizó esta metodología debido a su versatilidad de aplicación tanto en BI como en DM. Crisp-DM contiene fases que involucran: la comprensión del negocio, de los datos y su preparación; así como también, la modelación de supuestos posteriormente evaluados (Chalaris, Gritzalis, Maragoudakis, Sgouropoulou, & Tsolakidis, 2014).
- **MATERIALES Y HERRAMIENTAS.** Se utilizó la data cruda extraída de la base de datos del sistema de procesamiento transaccional (SPT) de la empresa objeto de estudio. Se realizó el proceso ETL mediante Pentaho Data Integration y se construyó la DW en el DBMS PostgreSQL. Se utilizó Pentaho BI Server combinando con HTML,

CSS y JavaScript para el diseño de cubos OLAP y construcción del Dashboard BI. Los modelos predictivos de MD fueron implementados utilizando RStudio, el lenguaje R y el framework Shiny.

A continuación, se describen las etapas aplicadas según la metodología CRISP-DM para el desarrollo de la solución BI y DM:

- a. **COMPRESIÓN DEL NEGOCIO.** La empresa INCUPASAJE tiene como finalidad obtener ingresos económicos mediante la producción de aves en estado inicial, garantizando así la solvencia del negocio. Para llevar el control de sus actividades, almacena registros en archivos de diferentes formatos (.xls, .csv) y posee un SPT que utiliza como base de datos a Microsoft Office Access.
- b. **COMPRESIÓN DE LOS DATOS.** Se recolectaron los datos necesarios y se establecieron relaciones entre ellos para el diseño y creación de un DW, cuya estructura fue de tipo copo de nieve; y en el que se especificaron cada una de las entidades inmersas en los procesos de producción de la empresa.

La figura 1 presenta el diseño de DW elaborado a partir de las entidades de datos que posee el negocio; el cual ha sido dividido en data marts que agrupa tablas de hechos y dimensiones.

Una tabla de hechos (th) agrupa las medidas o indicadores claves de desempeño (KPI's), y las dimensiones (dim) son las perspectivas o vistas de análisis de los KPI's. En la tabla 1 se recolectan los principales KPI's identificados en la empresa avícola.

TABLA 1. ESTRUCTURA DE UN DATA WAREHOUSE DISTRIBUIDO POR TEMAS DE LA EMPRESA DE PRODUCCIÓN AVÍCOLA

DATAMART	KPI'S	FÓRMULA
Incubación	Número de huevos incubados	Suma de huevos incubados
	Número de huevos perdidos	Suma de huevos perdidos
	Número de huevos producidos	(Número de huevos incubados - Número de huevos perdidos)
Crianza	Número de aves para crianza	Suma total de aves para crianza
	Número de aves perdidas	Suma total de aves perdidas
	Número de aves producidas	(Número de aves producidas - Número de aves perdidas)
Producción	Ingresos (aves, huevos)	Suma de ingresos (aves, huevos)
	Gastos por insumo	Suma de gastos por insumo
	Total de costo de producción	Suma de costos
Venta	Total de ventas	Suma de ventas realizadas
Utilidad	Utilidad	(Suma total de ventas - Suma Total de costo)
	Porcentaje de utilidad	(Utilidad / Ingresos)*100

- La regresión permite determinar el grado de dependencia o relación entre las variables X e Y, posibilitando su predicción en base a una función matemática generada a partir de los datos históricos de la empresa.
- Las series temporales permiten realizar los pronósticos en base al comportamiento pasado de la variable de interés en función del tiempo. Pueden ser de forma discreta o continua dependiendo de las observaciones que se realicen.

CONSTRUCCIÓN DEL MODELO

- **MODELOS DE REGRESIÓN.** Se evaluaron diferentes modelos de regresión y para determinar el óptimo, se consideró el coeficiente de determinación (correlación) R² que indica la proporción de variación de la variable dependiente explicada por la variable independiente. Para que el coeficiente sea estadísticamente significativo, la probabilidad (P-valor) debe ser menor a 0.05. Dichos modelos se especifican en la tabla 2.

TABLA 2. MODELOS DE REGRESIÓN

TIPO DE REGRESIÓN	FÓRMULA
Lineal	$y=a+bx$
Cuadrática	$y=a+bx+cx^2$
Cúbica	$y=a+bx+cx^2+dx^3$
Logarítmica	$y=a+b \ln(x)$

Los criterios de evaluación considerados para R² son los que se especifican en la tabla 3.

TABLA 3. CRITERIOS DE EVALUACIÓN PARA R²

R ² ≥0.95	Eficiente
R ² ≥0.85 & R ² ≥0.94	Bueno
R ² ≥0.70 & R ² ≥0.84	Regular
R ² ≤0.7	No aceptable

Por otra parte, para la evaluación de P-valor se consideró un nivel de confianza del 95% (0.95) y un grado de error del 5% (0.05).

- **SERIES TEMPORALES.** Es una secuencia de datos que se encuentra distribuida cronológicamente. Los intervalos de separación entre datos pueden ser iguales o desiguales. En el análisis de series temporales se aplican métodos para interpretar y extraer información representativa de las relaciones entre los datos

de la serie con la finalidad de extrapolar o interpolar, y de esta manera predecir el comportamiento de en un momento no observado pasado o futuro. En el análisis de la serie se identifican componentes como, tendencia y variación (estacional, cíclica, aleatoria, etc.). La estimación de la tendencia puede realizarse mediante los métodos de cuadrados mínimos o promedios móviles simples. La estimación de las variaciones mediante el análisis espectral o el análisis de ondas. Los tipos de series temporales pueden ser aditivas, multiplicativas o mixtas. A continuación, se explican los modelos empleados.

- **MODELO PREDICTIVO PARA INGRESOS SEGÚN NÚMERO DE VENTAS.** El modelo predictivo de ingresos consistió en considerar el número de ventas realizadas desde el año 2014 hasta el 2016, con el fin de estimar los ingresos a generarse según el número de ventas. Para la construcción del primer modelo se emplearon las técnicas de regresión para lo que se estableció una relación entre la variable dependiente Total de ingresos (Y) y la independiente Número de ventas (X).
- **MODELO PREDICTIVO PARA PRODUCCIÓN DE AVES EN ESTADO INICIAL SEGÚN EL PERÍODO.** Para el modelo predictivo de producción de aves en estado inicial se consideró el total de aves de primera, producidas desde el año 2014 hasta el mes de junio del 2016. Esto permitió que, a través de los datos históricos, se pudieran estimar las producciones futuras en base a la aplicación de la técnica de serie temporal. Para este modelo de predicción la variable independiente (X) está determinada por los años de producción y la variable dependiente (Y) es el número de aves en estado inicial producidas.
- **MODELO PREDICTIVO PARA VENTAS SEGÚN PERÍODO.** El modelo predictivo de ventas consistió en considerar las ventas totales realizadas desde el año de 2014 hasta el mes de junio del 2016, con el fin de estimar las ventas a realizarse en los próximos años. Para la construcción del modelo se aplicó la técnica de series temporales y se consideraron las siguientes variables: la variable independiente (X) representa el período de ventas (año o meses) y la variable dependiente (Y) el Total de ventas.

**DISCUSIÓN DE RESULTADOS
RESULTADOS**

En la figura 3 se presenta el dashboard perteneciente al datamart de crianza elaborado mediante una plataforma BI, a través del cual, pueden visualizarse la cantidad inicial de aves de crianza, el número de aves de pérdida y el total de aves producidas.

En lo referente al modelo predictivo para ingresos según número de ventas mediante técnicas de regresión, su evaluación arrojó los valores que se detallan a continuación en la tabla 4.

TABLA 4. VALORES OBTENIDOS DE LA EVALUACIÓN

TIPO DE REGRESIÓN	R2	P_VALOR
Lineal	0.887	8.03 e-14
Cuadrática	0.892	8.104 e-13
Cúbica	0.893	8.776 e-12
Logarítmica	0.7968	1.734 e-10

El modelo más representativo para la predicción de ingresos de producción a partir del número de ventas es el de regresión cúbica con un valor de correlación del R2=0.893. Además, tiene un P_valor < 0.05. Estableciendo el modelo de predicción tal y como se presenta en la ecuación (1):

$$y = -184.9x^3 + 20175.2x^2 + 2636.1x + 12285.2$$

En figura 4 se presentan las gráficas de cada uno de los modelos evaluados con su respectivo factor de correlación (R2) y su línea de tendencia.

En cuanto al modelo predictivo para la producción de aves en estado inicial según el período, los resultados obtenidos se muestran en las figuras 5 y 6. Se presenta un crecimiento mínimo en los niveles de producción, lo que debe ser analizado por el administrador para tomar las decisiones del caso.

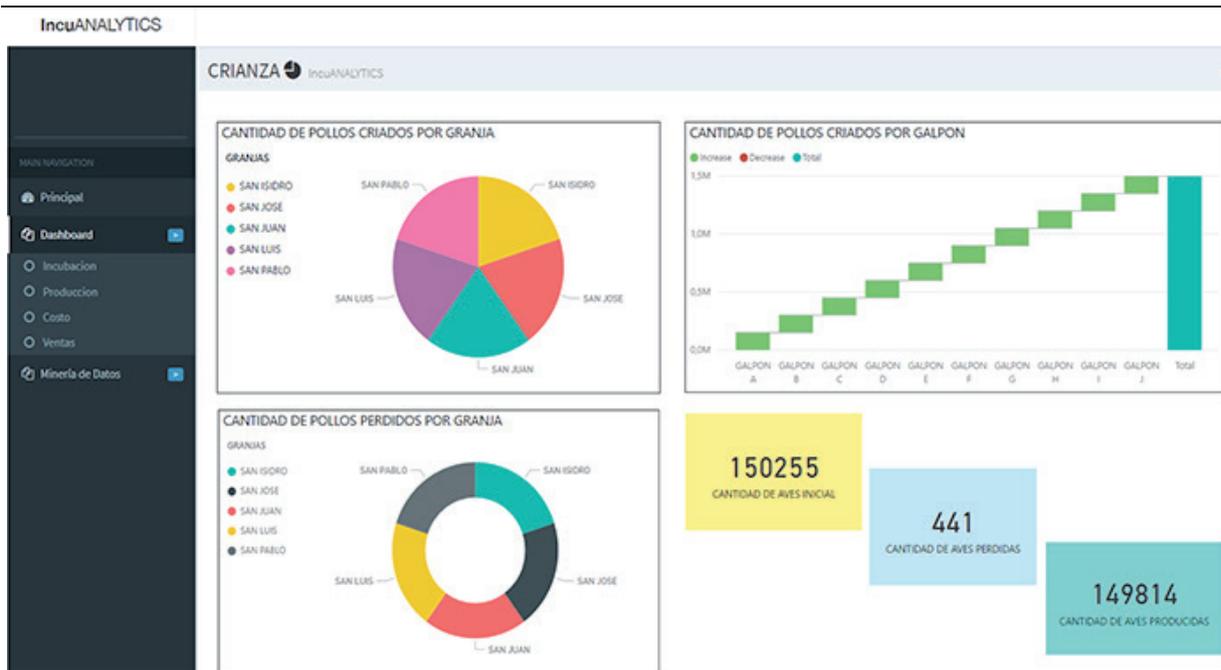


Figura 3. Dashboard perteneciente al datamart de Crianza

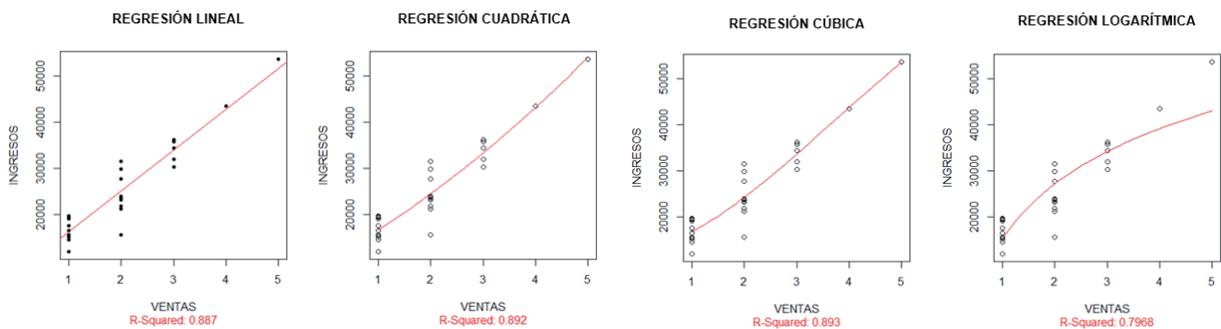


Figura 4. Modelos de Regresión Evaluados

PREDICION DE PRODUCCION DE AVES EN ESTADO INICIAL

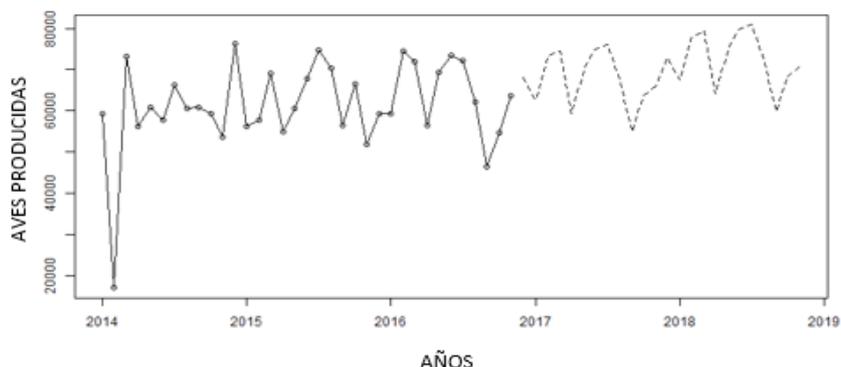


Figura 5. Serie temporal de aves producidas y pronóstico

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2016												
2017	62731.02	73094.27	74443.08	59406.42	69902.32	74968.55	75961.80	67881.91	54967.98	63554.42	65956.62	68004.15
2018	67581.11	77944.36	79293.17	64256.52	74752.41	79818.64	80811.89	72732.01	59818.07	68404.51	70806.71	72854.24

Figura 6. Datos pronosticados de producción de aves en su etapa inicial para dos años

PREDICION DE VENTAS EN LOS PROXIMOS DOS AÑOS

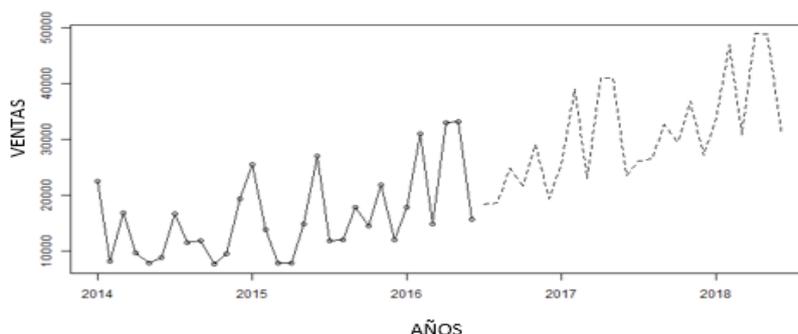


Figura 7. Representación de los datos pronosticados

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2016							18370.27	18714.74	24819.93	21637.07	29051.28	19372.75
2017	25605.56	39037.08	23046.18	41039.37	41028.41	23598.41	26218.67	26563.15	32668.34	29485.48	36899.69	27221.16
2018	33453.97	46885.48	30894.59	48887.77	48876.82	31446.82						

Figura 8. Datos pronosticados de las ventas para dos años.

En lo que respecta al modelo predictivo para ventas según período, en las figuras 7 y 8 se muestran las ventas reales y el pronóstico obtenido al predecir las ventas en dos años (2017 y 2018) en base a la técnica de series temporales. En la gráfica las líneas continuas representan datos reales históricos y las líneas entre cortadas representan la predicción.

Según la predicción, la empresa tendría un incremento en sus ventas en los próximos dos años. Cabe recalcar, que la predicción inicia en el mes de julio del 2016 debido a que los datos que se recopilaban son hasta el mes de junio de 2016.

DISCUSIÓN

En diferentes artículos se emplea como solución de BI, la creación de tableros de control. Tal es el caso de Erkollar & Oberer (2016) que se basaron en un modelo scorecard junto con la creación y diseño de un dashboard integral definiendo dimensiones, KPI, métricas y marcos de tiempo que le posibilitaran a los dueños de la empresa una mejor visualización de sus datos.

Otra de las opciones usualmente utilizadas es la aplicación de DM teniendo como ejemplo al trabajo desarrollado por Reuter & Brambring (2016) que emplearon métodos Bayesianos y de árboles de decisión para el aumento de la

consistencia en los controles de producción; obteniendo como resultado una alta precisión en el pronóstico de múltiples opciones para una misma variable objetivo. Por otra parte, Ng, Bandaru, & Frantzén (2016) implementaron un sistema de producción con técnicas de optimización multiobjetivo mediante la utilización de métodos descriptivos, tales como gráficos, Heatmaps y clúster de colores que permitieron una transmisión de conocimiento visual hacia los directivos de la entidad.

Mediante el análisis de los proyectos anteriores, se constató que en cada uno de ellos se aplica una sola herramienta de BI como solución a un problema específico. Es por ello que en el presente proyecto, se integraron en una aplicación web, herramientas tanto de BI como de DM. Estas, comparten un mismo almacén de datos para análisis con técnicas multidimensional OLAP y métodos predictivos que facilitan a los usuarios del sector estratégico de la empresa, tener acceso a información consolidada de forma oportuna para la toma de decisiones en pro del incremento de la producción y ventas de sus productos; y por ende, en sus ingresos para una mayor rentabilidad y beneficio.

CONCLUSIONES

La aplicación llamada IncuAnalytic fue desarrollada según la metodología CRISP-DM que sirvió de guía en cada fase del proceso. Las funcionalidades de esta aplicación se enfocan en el análisis de datos aplicando tanto BI como DM. En lo referente a BI se implementó un dashboard con interfaz web para la visualización consolidada y estadística de los KPI's y análisis multidimensional OLAP de información de producción y ventas de una empresa avícola del cantón Pasaje. Se aprovechó la DM para crear modelos predictivos en base a las técnicas de regresión y series temporales, permitiendo al administrador o usuario estratégico del negocio contar con una herramienta que le proporciona patrones basados en información histórica y tendencias futuras.

Las herramientas utilizadas son open access (libre acceso), entre ellas están Pentaho, PostgreSQL y R. La aplicación IncuAnalytic trabaja con un almacén de datos dinámico que se alimenta periódicamente mediante un proceso ETL. Esta solución de BI y DM puede ser implementada en cualquier empresa avícola con procesos de producción y venta similares con mínimos cambios de reajuste de requerimientos particulares.

REFERENCIAS BIBLIOGRÁFICAS

- Astriani, W., & Trisminingsih, R. (2016). Extraction, Transformation, and Loading (ETL) Module for Hotspot Spatial Data Warehouse Using Geokettle. *Procedia Environmental Sciences*, 33, 626-634. <https://doi.org/10.1016/j.proenv.2016.03.117>
- Boulekrouche, B., Jabeur, N., & Alimazighi, Z. (2015). An intelligent ETL grid-based solution to enable spatial data warehouse deployment in cyber physical system context. *Procedia Computer Science*, 56(1), 111-118. <https://doi.org/10.1016/j.procs.2015.07.176>
- Chalaris, M., Gritzalis, S., Maragoudakis, M., Sgouropoulou, C., & Tsolakidis, A. (2014). ScienceDirect ICININFO Improving Quality of Educational Processes Providing New Knowledge using Data Mining Techniques. *Procedia - Social and Behavioral Sciences*, 147, 390-397. <https://doi.org/10.1016/j.sbspro.2014.07.117>
- Coelho, D., Miranda, J., Portela, F., Machado, J., Santos, M. F., & Abelha, A. (2016). Towards of a Business Intelligence Platform to Portuguese Miseriçórdias. *Procedia Computer Science*, 100, 762-767. <https://doi.org/10.1016/j.procs.2016.09.222>
- Erkollar, A., & Oberer, B. (2016). Multidimensional Dashboards for Evaluating Strategic brand Management Processes for Multi-brand Companies. *Procedia - Social and Behavioral Sciences*, 235(October), 505-513. <https://doi.org/10.1016/j.sbspro.2016.11.062>
- Femina, B. T., & Sudheep, E. M. (2015). An efficient CRM-data mining framework for the prediction of customer behaviour. *Procedia Computer Science*, 46, 725-731. <https://doi.org/10.1016/j.procs.2015.02.136>
- Gourbesville, P., Du, M., Zavattoni, E., & Ma, Q. (2016). DSS Architecture for Water Uses Management. *Procedia Engineering*, 154, 928-935. <https://doi.org/10.1016/j.proeng.2016.07.512>
- Gröger, C., Hillmann, M., Hahn, F., Mitschang, B., & Westkämper, E. (2013). The operational process dashboard for manufacturing. *Procedia CIRP*, 7, 205-210. <https://doi.org/10.1016/j.procir.2013.05.035>
- Hamrouni, T., Slimani, S., & Ben Charrada, F. (2015). A critical survey of data grid Replication strategies based on data mining techniques. *Procedia Computer Science*, 51(1), 2779-2788. <https://doi.org/10.1016/j.procs.2015.05.434>
- Hermawati, R., & Sitanggang, I. S. (2016). Web-Based Clustering Application Using Shiny Framework and DBSCAN Algorithm for

- Hotspots Data in Peatland in Sumatra. *Procedia Environmental Sciences*, 33, 317-323. <https://doi.org/10.1016/j.proenv.2016.03.082>
- Jothi, N., Rashid, N. A., & Husain, W. (2015). Data Mining in Healthcare - A Review. *Procedia Computer Science*, 72, 306-313. <https://doi.org/10.1016/j.procs.2015.12.145>
- Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine Learning and Data Mining Methods in Diabetes Research. *Computational and Structural Biotechnology Journal*, 15, 104-116. <https://doi.org/10.1016/j.csbj.2016.12.005>
- Khan, A., Ehsan, N., Mirza, E., & Zahoor, S. (2012). Integration between Customer Relationship Management (CRM) and Data Warehousing, *Procedia Technology* 1, 239-249. <https://doi.org/10.1016/j.protcy.2012.02.050>
- Lindberg, C. F., Tan, S., Yan, J., & Starfelt, F. (2015). Key Performance Indicators Improve Industrial Performance. *Energy Procedia*, 75, 1785-1790. <https://doi.org/10.1016/j.egypro.2015.07.474>
- Longo, A., Giacobelli, S., & Boichichio, M. A. (2014). Fact - Centered ETL: A Proposal for Speeding Business Analytics up. *Procedia Technology*, 16, 471-480. <https://doi.org/10.1016/j.protcy.2014.10.114>
- Mazon-Olivo, B., Rivas, W., Pinta, M., Mosquera, A., Astudillo, L., & Gallegos, H. (2017). Dashboard para el soporte de decisiones en una empresa del sector minero. *Conference Proceedings - Universidad Técnica de Machala*, 1, 1218-1229.
- Meier, H., Lagemann, H., Morlock, F., & Rathmann, C. (2013). Key performance indicators for assessing the planning and delivery of industrial services. *Procedia CIRP*, 11, 99-104. <https://doi.org/10.1016/j.procir.2013.07.056>
- Ng, A. H. C., Bandaru, S., & Frantzén, M. (2016). Innovative Design and Analysis of Production Systems by Multi-objective Optimization and Data Mining. *Procedia CIRP* (50), 665-671. <https://doi.org/10.1016/j.procir.2016.04.159>
- Reuter, C., & Brambring, F. (2016). Improving Data Consistency in Production Control. *Procedia CIRP*, 41, 51-56. <https://doi.org/10.1016/j.procir.2015.12.116>
- Siknun, G. P., & Sitanggang, I. S. (2016). Web-based Classification Application for Forest Fire Data Using the Shiny Framework and the C5.0 Algorithm. *Procedia Environmental Sciences*, 33, 332-339. <https://doi.org/10.1016/j.proenv.2016.03.084>
- Tokola, H., Gröger, C., Järvenpää, E., & Niemi, E. (2016). ScienceDirect Designing manufacturing dashboards on the basis of a Key Performance Indicator survey. *Procedia CIRP*, 0, 619-624. <https://doi.org/10.1016/j.procir.2016.11.107>
- Tutunea, M. F. (2015). Business Intelligence Solutions for Mobile Devices - An Overview. *Procedia Economics and Finance*, 27(15), 160-169. [https://doi.org/10.1016/S2212-5671\(15\)00985-5](https://doi.org/10.1016/S2212-5671(15)00985-5)
- van Os, H. W. A., Herber, R., & Scholtens, B. (2014). Designing a Decision Support System for Subsurface Activities. *Procedia Environmental Sciences*, 22, 12-19. <https://doi.org/10.1016/j.proenv.2014.11.002>
- Wessiani, N. A., & Sarwoko, S. O. (2015). Risk Analysis of Poultry Feed Production Using Fuzzy FMEA. *Procedia Manufacturing*, 4(IESS), 270-281. <https://doi.org/10.1016/j.promfg.2015.11.041>