

Mejoramiento del sistema de evaluación de desempeño en cultivo a través de la estandarización y análisis de datos en Florval S.A.S.

Cesar Yesith Rivera León

Director

Andrés Felipe Hernández Giraldo

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI
Especialización en Ciencia de Datos y Analítica

2026

Nota de Aceptación

Nombre Director de Trabajo de Grado

Jurado

Jurado

Dedicatoria

Dedico este trabajo principalmente a Dios, por brindarme la fortaleza, disciplina y perseverancia necesarias para culminar este proceso académico. A mi familia, por su apoyo constante, comprensión y motivación a lo largo de mi formación profesional, siendo un pilar fundamental en cada etapa de este camino.

De manera especial, dedico este proyecto a quienes han confiado en mi crecimiento personal y profesional, y han contribuido directa o indirectamente a que este logro sea posible.

Agradecimientos

Expreso mi agradecimiento a la Universidad Nacional Abierta y a Distancia (UNAD) por proporcionar las herramientas académicas y metodológicas necesarias para el desarrollo de este proyecto. De igual forma, agradezco al asesor Msc. Andrés Felipe Hernández Giraldo por su acompañamiento, orientación y retroalimentación durante el desarrollo del trabajo de grado. Agradezco a la empresa Florval S.A.S. por permitir el acceso a la información y el contexto organizacional requerido para la ejecución del proyecto, así como a los colaboradores y supervisores del área de cultivo que aportaron su conocimiento y experiencia al proceso investigativo.

Finalmente, agradezco a todas las personas que, de manera directa o indirecta, contribuyeron al desarrollo y culminación de este trabajo académico.

Resumen

El presente proyecto tiene como objetivo mejorar el Sistema de Evaluación de Desempeño en el área de cultivo de Florval S.A.S. mediante la aplicación de técnicas de analítica de datos. Actualmente, la organización genera información a partir de fuentes como Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos; sin embargo, estos datos no se integran ni analizan de forma sistemática, lo que limita la objetividad del proceso evaluativo.

La metodología empleada incluye un análisis exploratorio de datos para identificar patrones y variables críticas, la aplicación de modelos analíticos como regresión, clustering y árboles de decisión para segmentar colaboradores y determinar factores que influyen en el rendimiento, y la construcción de un dashboard analítico en Power BI para la visualización de indicadores clave de desempeño. Finalmente, se realiza una validación del modelo analítico comparando sus resultados con el sistema tradicional de evaluación utilizado por la empresa.

Los resultados evidencian que la analítica de datos mejora la precisión, trazabilidad y consistencia del proceso de evaluación de desempeño, aportando información relevante para la toma de decisiones operativas y estratégicas. Se concluye que la implementación de un enfoque basado en datos contribuye significativamente a la optimización del sistema de evaluación en el área de cultivo de Florval S.A.S..

Palabras clave: analítica de datos, evaluación de desempeño, ciencia de datos, Power BI, agricultura.

Abstract

This project aims to improve the Performance Evaluation System in the cultivation area of Florval S.A.S. through the application of data analytics techniques. Currently, the organization generates information from sources such as Frequent Aids (AF), Area Evaluation (EA), Clean Cut, and Pilots; however, these data are not systematically integrated or analyzed, which limits the objectivity of the evaluation process.

The methodology includes exploratory data analysis to identify patterns and critical variables, the application of analytical models such as regression, clustering, and decision trees to segment employees and determine factors affecting performance, and the development of an analytical dashboard in Power BI to visualize key performance indicators. Finally, the analytical model is validated by comparing its results with the traditional evaluation system used by the organization.

The results show that data analytics improves the accuracy, traceability, and consistency of the performance evaluation process, providing relevant information for operational and strategic decision-making. It is concluded that the implementation of a data-driven approach significantly contributes to the optimization of the performance evaluation system in Florval S.A.S.'s cultivation area.

Keywords: data analytics, performance evaluation, data science, Power BI, agriculture.

Tabla de Contenido

Introducción	13
Descripción del Problema	14
Planteamiento del Problema	15
Justificación	17
Objetivos	18
Objetivo General.....	18
Objetivos Específicos	18
Marco de Referencia	19
Estado del Arte	19
Marco Contextual	20
Marco Teórico	21
Marco Conceptual.....	23
Analítica de Datos.....	23
Ciencia de Datos	23
Análisis Exploratorio de Datos (EDA)	23
Modelos Analíticos	23
Modelo de Regresión	24
Modelo de Clustering	24
Modelo de Árboles de Decisión	25
Justificación de Exclusión de otros Modelos Analíticos	25
Evaluación de Desempeño.....	26
Dashboard Analítico	26

Metodología	27
Enfoque Metodológico	27
Fases de la Metodología	27
Recolección y Comprensión de los Datos	27
Preparación y Análisis Exploratorio de Datos (EDA)	28
Modelado Analítico	30
Validación y Visualización	32
Resultados	34
Desarrollo del Modelo Analítico	34
Estandarización y Consolidación de los Registros Históricos	39
Unificación de Estructuras y Criterios de Registro	40
Consolidación de Información Histórica y Trazabilidad	40
Soporte para el Análisis Exploratorio y el Modelado Analítico	41
Impacto en el Sistema de Evaluación de Desempeño	41
Valor Estratégico del Resultado Obtenido	41
Análisis Exploratorio del Desempeño (EDA)	45
Cobertura Temporal y Organizacional	46
Aporte del EDA al Modelo Analítico	47
Aplicación de Modelos Analíticos	48
Modelo de Regresión Análisis de Factores Influyentes en el Desempeño	48
Modelo de Clustering Segmentación del Desempeño de los Colaboradores	52
Modelo de Árboles de Decisión y Clasificación e Interpretación del Desempeño	54
Resultados Comparativos Frente al Sistema Tradicional de Evaluación	56

Generación de Métricas Analíticas y Clasificación del Desempeño	56
Integración de los Resultados en Herramientas de Visualización	57
Conclusiones.....	62
Recomendaciones	63
Referencias Bibliográficas	65

Lista de Tablas**Tabla 1** *Tamaño y Composición de los Conjuntos de Datos Utilizados en el Análisis*

<i>Exploratorio (EDA)</i>	45
---------------------------------	----

Lista de Figuras

Figura 1 <i>Recolección y Comprensión de los Datos del Área de Cultivo de Florval S.A.S.</i>	28
Figura 2 <i>Preparación y Análisis Exploratorio de Datos (EDA) para la Evaluación de Desempeño en el Área de Cultivo de Florval S.A.S.</i>	29
Figura 3 <i>Vista de los Datos de Cumplimiento por Labor en la Fase de Preparación y Análisis Exploratorio de Datos (EDA)</i>	30
Figura 4 <i>Esquema del Modelado Analítico para la Evaluación del Desempeño en el Área de Cultivo de Florval S.A.S.</i>	31
Figura 5 <i>Validación y Visualización de Resultados para la Evaluación del Desempeño en el Área de Cultivo de Florval S.A.S.</i>	32
Figura 6 <i>Impacto y Beneficios del Proyecto de Análisis de Desempeño en Florval S.A.S.</i>	33
Figura 7 <i>Desarrollo del Modelo Analítico para la Evaluación del Desempeño en Florval S.A.S.</i>	35
Figura 8 <i>Fragmento del Código del Modelo Analítico para la Composición y Validación de Datos por Fase</i>	36
Figura 9 <i>Fragmento del Código del Modelo Analítico para la Identificación y Normalización de la Fase Productiva</i>	37
Figura 10 <i>Cálculo del Indicador de Calidad Mediante el Modelo Analítico Desarrollado en Python</i>	38
Figura 11 <i>Estandarización y Consolidación de los Registros Históricos para la Evaluación del Desempeño en Florval S.A.S.</i>	42
Figura 12 <i>Vista de la Base de Datos Consolidada tras el Proceso de Estandarización de Registros Históricos</i>	43

Figura 13 <i>Resultados del Modelo de Regresión para el Análisis de Factores Influyentes en el Desempeño</i>	49
Figura 14 <i>Ejecución del Modelo de Regresión en Google Colab para el Análisis de Factores Influyentes en el Desempeño</i>	50
Figura 15 <i>Ejecución del Análisis Exploratorio y Modelado Analítico en Google Colab</i>	51
Figura 16 <i>Ejecución del Modelo de Clustering para la Segmentación del Desempeño de los Colaboradores en Google Colab</i>	53
Figura 17 <i>Ejecución del Modelo de Árboles de Decisión para la Clasificación e Interpretación del Desempeño en Google Colab</i>	55
Figura 18 <i>Dashboard del Indicador de Evaluación de Desempeño Semanal por Colaborador y Fase</i>	58
Figura 19 <i>Módulo de Análisis de Compensatorios y Calificación Histórica por Semanas</i>	59
Figura 20 <i>Mapa de Calor del Cumplimiento de Cantidad por Bloque y Labor</i>	60
Figura 21 <i>Ejecución del Script de Procesamiento y Consolidación de Datos en Python</i>	61

Introducción

La Evaluación de Desempeño en el área de cultivo de Florval S.A.S. constituye un proceso esencial para garantizar la productividad, la calidad del producto y la eficiencia operativa dentro de las actividades agrícolas. Actualmente, la empresa recolecta información mediante Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos; sin embargo, estos datos se utilizan únicamente como registros operativos sin un análisis profundo que permita comprender el comportamiento del desempeño, identificar patrones o anticipar problemáticas.

En un entorno empresarial donde la toma de decisiones basadas en datos se ha convertido en un factor estratégico, la ausencia de un modelo analítico limita la capacidad de la organización para obtener información valiosa a partir de los datos generados diariamente. La falta de integración, estandarización y análisis avanzado dificulta la objetividad de las evaluaciones, genera inconsistencias y limita la trazabilidad del proceso.

Por lo tanto, surge la necesidad de aplicar conceptos y técnicas de ciencia de datos que permitan transformar los registros operativos en información estructurada, analítica y visualmente comprensible. Este anteproyecto propone el diseño de un modelo analítico para mejorar el sistema de Evaluación de Desempeño en cultivo, integrando análisis exploratorio de datos (EDA), modelos estadísticos, técnicas de machine learning y herramientas de visualización como Power BI o Python. Con ello, se busca fortalecer la toma de decisiones, mejorar la eficiencia en los procesos operativos y promover una cultura organizacional basada en datos.

Descripción del Problema

La evaluación del desempeño en el área de cultivo de Florval S.A.S. es un proceso clave para garantizar la productividad, la calidad de las labores agrícolas y el cumplimiento de los objetivos operativos de la organización. Actualmente, este proceso se apoya en diferentes fuentes de información operativa, tales como Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos, las cuales generan un volumen considerable de datos relacionados con cantidad, calidad y cumplimiento de las labores realizadas por los colaboradores.

No obstante, dichos datos se gestionan de manera fragmentada, con estructuras heterogéneas y sin un proceso sistemático de integración, estandarización y análisis. En la práctica, esta situación limita el aprovechamiento de la información disponible, ya que los registros se utilizan principalmente como soporte operativo y no como insumo analítico para la toma de decisiones. Como consecuencia, el sistema de evaluación del desempeño presenta dificultades en términos de objetividad, trazabilidad y consistencia, lo que puede generar percepciones de subjetividad en los resultados y limitar la identificación oportuna de factores que influyen en el rendimiento de los colaboradores.

Adicionalmente, el sistema tradicional de evaluación no permite analizar de manera integral el comportamiento histórico del desempeño, ni identificar patrones, tendencias o segmentos de colaboradores con características similares. Esta limitación dificulta la comparación entre bloques, fases productivas y períodos de tiempo, así como la evaluación de la efectividad real de las decisiones tomadas a partir de los resultados obtenidos.

En un contexto organizacional donde la toma de decisiones basada en datos se ha convertido en un factor estratégico, la ausencia de un modelo analítico estructurado para la evaluación del desempeño representa una oportunidad de mejora para Florval S.A.S. Por ello,

surge la necesidad de implementar un enfoque basado en analítica de datos que permita integrar la información histórica, aplicar modelos analíticos y fortalecer el sistema de evaluación del desempeño en el área de cultivo.

Planteamiento del Problema

A pesar de que Florval S.A.S. dispone de múltiples fuentes de datos operativos relacionadas con la evaluación del desempeño en el área de cultivo, actualmente no cuenta con un modelo analítico que permita integrar, estandarizar y analizar dicha información de manera sistemática. El sistema de evaluación tradicional se basa principalmente en registros operativos aislados y en criterios que no siempre permiten identificar de forma objetiva los factores que inciden en el desempeño de los colaboradores.

Esta situación genera dificultades para segmentar a los colaboradores según su nivel de desempeño, identificar patrones de comportamiento, evaluar la influencia de variables como la fase productiva, el tipo de labor o la calidad alcanzada, y comparar los resultados obtenidos con un enfoque analítico frente al sistema tradicional de evaluación. Asimismo, la falta de análisis histórico limita la capacidad de la organización para anticipar desviaciones en el desempeño y para tomar decisiones oportunas orientadas a la mejora continua.

En consecuencia, Florval S.A.S. enfrenta el reto de fortalecer su sistema de evaluación del desempeño mediante la incorporación de herramientas de analítica de datos que permitan reducir la subjetividad, mejorar la trazabilidad de la información y apoyar la toma de decisiones basada en evidencia. En este sentido, se plantea la necesidad de responder a la siguiente pregunta de investigación:

¿Cómo la aplicación de modelos analíticos de regresión, clustering y árboles de decisión, basados en datos históricos del área de cultivo de Florval S.A.S., puede mejorar la evaluación del desempeño de los colaboradores en comparación con el sistema de evaluación tradicional?

Justificación

La evaluación del desempeño es un componente crítico en la gestión del talento humano y en la productividad del sector agrícola. Sin embargo, cuando este proceso se realiza sin un adecuado análisis de datos, surgen inconsistencias, subjetividad y dificultades para garantizar que las decisiones sean realmente objetivas y oportunas. En el caso de Florval S.A.S., aunque existen registros de AF, EA, Corte Limpio y Pilotos, estos datos no se integran ni se analizan mediante técnicas especializadas, lo cual impide obtener una visión clara y completa del rendimiento de los colaboradores del cultivo.

La implementación de un modelo analítico aporta una solución concreta a esta problemática, ya que permite estructurar y depurar los datos, identificar patrones de comportamiento, detectar desviaciones, clasificar perfiles de desempeño y determinar los factores que más influyen en la productividad. Además, el uso de técnicas como regresión, clustering y árboles de decisión fortalece la capacidad de la empresa para comprender el comportamiento del proceso y anticipar resultados.

Este proyecto es pertinente porque permitirá mejorar la precisión, confiabilidad y trazabilidad del sistema de evaluación; promoverá decisiones basadas en evidencia; reducirá errores operativos y sesgos; y facilitará la generación de dashboards interactivos para supervisores, auxiliares y gerencia. Asimismo, al estandarizar los registros y los procesos de análisis, la empresa podrá optimizar el monitoreo del desempeño en cultivo y fomentar un modelo de mejora continua orientado a datos

Objetivos

Objetivo General

Desarrollar un modelo analítico basado en el análisis de datos históricos del área de cultivo de Florval S.A.S., que permita identificar y definir indicadores clave de desempeño (KPI) para apoyar la toma de decisiones estratégicas y operativas.

Objetivos Específicos

Estandarizar los registros históricos de AF, EA, Corte Limpio y Pilotos para consolidar una base de datos limpia y estructurada.

Realizar un análisis exploratorio de datos (EDA) que permita identificar patrones, correlaciones y variables críticas asociadas al desempeño del personal.

Aplicar modelos analíticos de regresión, clustering y árboles de decisión para segmentar a los colaboradores, identificar los factores que inciden en su desempeño y evaluar la efectividad de estos modelos en comparación con el sistema de evaluación tradicional.

Construir un dashboard analítico en Power BI que integre indicadores clave de desempeño y facilite la toma de decisiones en el área de cultivo.

Marco de Referencia

Estado del Arte

El análisis del estado del arte evidencia que, en los últimos años, la analítica de datos se ha consolidado como una herramienta fundamental para la optimización de procesos organizacionales y el fortalecimiento de la toma de decisiones basada en datos. Diversos estudios coinciden en que el uso de información histórica reduce la subjetividad en la evaluación del desempeño y mejora la eficiencia operativa en distintos contextos empresariales.

Autores como Davenport y Harris (2007) destacan que las organizaciones que incorporan modelos analíticos para la gestión del desempeño logran ventajas competitivas sostenibles, al transformar los datos en conocimiento estratégico que respalda la toma de decisiones. De manera complementaria, Provost y Fawcett (2013) resaltan la importancia del análisis exploratorio de datos y de los modelos predictivos como soporte para la toma de decisiones gerenciales fundamentadas en evidencia.

En el ámbito de la evaluación del desempeño laboral, investigaciones como la desarrollada por Kwon y Sim (2013) señalan que el uso de enfoques basados en datos permite aplicar técnicas estadísticas y de aprendizaje automático —como regresión, clustering y árboles de decisión— para identificar indicadores clave de desempeño (KPI) más objetivos y consistentes, mejorando la transparencia y confiabilidad de los procesos evaluativos.

Por su parte, en el sector agroindustrial, estudios como Wolfert et al. (2017) y Kamilaris y Prenafeta-Boldú (2018) evidencian que la integración y análisis de grandes volúmenes de datos operativos facilitan la identificación de patrones de rendimiento, la optimización de recursos y el mejoramiento continuo de los procesos productivos. Estos trabajos destacan el valor del uso de datos históricos para apoyar la toma de decisiones en entornos agrícolas complejos.

No obstante, la literatura también señala que muchas organizaciones presentan limitaciones en la integración, estandarización y análisis sistemático de la información. LaValle et al. (2011) indican que la ausencia de modelos analíticos estructurados y de indicadores claramente definidos restringe el impacto de la analítica en la gestión del desempeño y en la generación de valor organizacional.

En este contexto, el presente proyecto se alinea con los aportes de la literatura existente y propone la aplicación de modelos analíticos de regresión, clustering y árboles de decisión para segmentar a los colaboradores, identificar los factores que inciden en su desempeño y evaluar la efectividad de estos modelos frente al sistema tradicional de evaluación, contribuyendo a una gestión del desempeño más objetiva y basada en datos en el área de cultivo de Florval S.A.S.

Marco Contextual

El proyecto se desarrolla en la empresa Florval S.A.S., específicamente en el área de cultivo, donde se lleva a cabo la evaluación de desempeño de los colaboradores mediante herramientas como Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos. Estas evaluaciones generan información relevante sobre productividad, calidad y cumplimiento de labores; sin embargo, los datos se gestionan de forma fragmentada, lo que limita su análisis integral.

El contexto organizacional evidencia la necesidad de implementar soluciones basadas en ciencia de datos que permitan integrar la información disponible, mejorar la objetividad del proceso evaluativo y fortalecer la toma de decisiones operativas y estratégicas en el área de cultivo.

Marco Teórico

La analítica de datos se ha consolidado como un componente estratégico para la gestión organizacional moderna, especialmente en los procesos de evaluación del desempeño y apoyo a la toma de decisiones. En este contexto, Davenport y Harris (2007) plantean que las organizaciones que incorporan modelos analíticos en sus procesos de gestión logran ventajas competitivas sostenibles, al transformar los datos históricos en conocimiento estratégico. Según los autores, este enfoque reduce la dependencia de la intuición y aumenta la objetividad y coherencia en los procesos de gestión del desempeño.

Desde la perspectiva de la ciencia de datos aplicada a los negocios, Provost y Fawcett (2013) establecen un marco conceptual que resalta la importancia del análisis exploratorio de datos, los modelos estadísticos y los modelos predictivos como herramientas fundamentales para la toma de decisiones gerenciales. Los autores señalan que el análisis sistemático de información histórica permite identificar patrones, tendencias y relaciones relevantes que apoyan la planificación, el control y la mejora continua de los procesos organizacionales.

En el ámbito específico de la evaluación del desempeño laboral, Kwon y Sim (2013) proponen un enfoque basado en datos que integra técnicas estadísticas y de aprendizaje automático, tales como regresión, clustering y árboles de decisión. Los autores argumentan que la aplicación de estos modelos contribuye a la identificación de indicadores clave de desempeño (KPI) más objetivos y consistentes, mejorando la transparencia y confiabilidad de los procesos evaluativos y reduciendo la subjetividad en la medición del desempeño laboral.

Por su parte, en el sector agroindustrial, Wolfert et al. (2017) analizan el impacto de la analítica de grandes volúmenes de datos en la agricultura, destacando su utilidad para integrar información operativa proveniente de distintas etapas del proceso productivo. El estudio

evidencia que el uso de datos históricos facilita la identificación de patrones de rendimiento, la optimización de recursos y el mejoramiento continuo de la productividad en entornos agrícolas complejos.

De manera complementaria, Kamilaris y Prenafeta-Boldú (2018) presentan un enfoque de toma de decisiones basada en datos aplicado a la agricultura inteligente. Los autores destacan que la analítica de datos permite apoyar la gestión de los procesos productivos mediante el análisis de información histórica, contribuyendo a una mayor eficiencia operativa y a decisiones más precisas en escenarios caracterizados por alta variabilidad.

Finalmente, LaValle et al. (2011) señalan que, a pesar del crecimiento exponencial en la generación de datos, muchas organizaciones enfrentan limitaciones relacionadas con la integración, estandarización y análisis sistemático de la información. Según los autores, la ausencia de modelos analíticos estructurados y de indicadores claramente definidos reduce el impacto de la analítica en la gestión del desempeño y en la generación de valor organizacional.

A partir de estos fundamentos teóricos, el presente proyecto se sustenta en los principios de la analítica de datos y la evaluación del desempeño basada en indicadores, proponiendo la aplicación de modelos analíticos de regresión, clustering y árboles de decisión para segmentar a los colaboradores, identificar los factores que inciden en su desempeño y evaluar la efectividad de estos modelos en comparación con el sistema de evaluación tradicional, con el fin de fortalecer la toma de decisiones, mejorar la evaluación del desempeño y disminuir la subjetividad en los procesos organizacionales del área de cultivo de Florval S.A.S.

Marco Conceptual

El desarrollo del presente proyecto se fundamenta en un conjunto de conceptos clave que permiten comprender la aplicación de la ciencia de datos en la optimización de procesos organizacionales, específicamente en la evaluación de desempeño en el sector agrícola.

Analítica de Datos

Se entiende como el proceso sistemático de inspección, limpieza, transformación y modelado de datos con el objetivo de descubrir información útil, generar conclusiones y apoyar la toma de decisiones. En el contexto de Florval S.A.S., la analítica de datos permite transformar registros operativos dispersos en información estructurada y accionable.

Ciencia de Datos

La ciencia de datos integra estadística, programación, aprendizaje automático y visualización para extraer conocimiento a partir de grandes volúmenes de datos. Su aplicación en procesos productivos agrícolas posibilita una evaluación objetiva del desempeño y la identificación de patrones que no son evidentes mediante métodos tradicionales.

Análisis Exploratorio de Datos (EDA)

El EDA corresponde a un conjunto de técnicas estadísticas y gráficas utilizadas para comprender la estructura de los datos, identificar valores atípicos, analizar distribuciones y detectar relaciones entre variables. En este proyecto, el EDA permitió reconocer variables críticas que influyen en el desempeño de los colaboradores del cultivo.

Modelos Analíticos

Los modelos analíticos son representaciones matemáticas y computacionales que permiten analizar datos históricos con el fin de identificar patrones, explicar relaciones entre variables y apoyar la toma de decisiones basada en datos. En el presente proyecto se emplean

modelos analíticos descriptivos y predictivos, seleccionados de acuerdo con la naturaleza de los datos del área de cultivo de Florval S.A.S., el tamaño del conjunto de datos y los objetivos del estudio.

Modelo de Regresión

El modelo de regresión se utiliza para analizar la relación entre una variable dependiente asociada al desempeño y una o varias variables independientes relacionadas con factores operativos del área de cultivo. Este modelo permite cuantificar el impacto de cada variable sobre el rendimiento, facilitando la identificación de los factores que influyen de manera significativa en el desempeño de los colaboradores.

La regresión fue seleccionada por su capacidad para explicar relaciones causa–efecto, su fácil interpretación y su adecuación al análisis de datos históricos estructurados. No se utilizaron modelos más complejos, como redes neuronales, debido a que estos requieren mayores volúmenes de datos y presentan menor interpretabilidad, lo cual no es adecuado para procesos de evaluación del desempeño donde la comprensión de los resultados es prioritaria.

Modelo de Clustering

El modelo de clustering se emplea como técnica de análisis descriptivo para la segmentación de colaboradores según características similares de desempeño. Este modelo permite identificar grupos homogéneos y patrones de comportamiento que apoyan una evaluación más objetiva y diferenciada del rendimiento.

Se seleccionó el clustering por su utilidad en el análisis exploratorio y su capacidad para descubrir estructuras internas en los datos sin necesidad de variables objetivo. No se utilizaron técnicas de clasificación supervisada en esta etapa, debido a la ausencia de etiquetas previamente definidas y al enfoque exploratorio del análisis.

Modelo de Árboles de Decisión

El modelo de árboles de decisión se utiliza para clasificar y analizar el desempeño a partir de reglas lógicas derivadas de los datos históricos. Este modelo facilita la identificación de los factores determinantes del rendimiento y las condiciones bajo las cuales se alcanzan determinados niveles de desempeño.

Los árboles de decisión fueron seleccionados por su alta interpretabilidad, facilidad de visualización y alineación con la toma de decisiones operativas. No se emplearon modelos de ensamble más avanzados, como random forest o gradient boosting, debido a que, aunque ofrecen mayor precisión, reducen la transparencia del modelo y aumentan la complejidad computacional, lo cual no es prioritario para los objetivos del proyecto.

Justificación de Exclusión de otros Modelos Analíticos

No se utilizaron modelos como redes neuronales, máquinas de soporte vectorial (SVM) o modelos de deep learning, debido a que el proyecto prioriza la interpretabilidad de los resultados, la facilidad de implementación y la alineación con procesos de gestión organizacional.

Asimismo, estos modelos requieren mayores volúmenes de datos, mayor complejidad técnica y un proceso de calibración más avanzado, lo cual no se ajusta al alcance ni a los objetivos del presente estudio.

En conjunto, los modelos de regresión, clustering y árboles de decisión permiten abordar el análisis del desempeño desde una perspectiva explicativa, descriptiva y predictiva, garantizando resultados comprensibles, aplicables y alineados con la identificación de indicadores clave de desempeño (KPI) para el área de cultivo de Florval S.A.S.

Evaluación de Desempeño

Es un proceso sistemático mediante el cual se mide el rendimiento de los colaboradores con base en criterios de productividad, calidad y cumplimiento. La incorporación de analítica de datos permite reducir la subjetividad, mejorar la trazabilidad y fortalecer la objetividad del proceso evaluativo.

Dashboard Analítico

Herramienta de visualización que integra indicadores clave de desempeño (KPIs) y facilita el seguimiento, análisis comparativo y toma de decisiones de manera interactiva.

Metodología

La metodología adoptada para el desarrollo del proyecto se fundamenta en un enfoque cuantitativo, orientado al análisis de datos operativos históricos generados en el área de cultivo de Florval S.A.S. Este enfoque permite trabajar con información numérica proveniente de registros reales del proceso productivo, facilitando la aplicación de técnicas estadísticas y modelos analíticos para el estudio del desempeño de los colaboradores.

El proyecto emplea analítica de datos descriptiva, exploratoria y predictiva, apoyada en modelos de regresión, clustering y árboles de decisión, con el objetivo de identificar patrones de comportamiento, analizar los factores que influyen en el desempeño y fortalecer la toma de decisiones basada en datos. La metodología se estructura en fases secuenciales que garantizan trazabilidad, coherencia y alineación con los objetivos planteados.

Enfoque Metodológico

El estudio adopta un enfoque cuantitativo debido a que se basa en el análisis de datos numéricos provenientes de los sistemas operativos del área de cultivo. Asimismo, presenta un alcance descriptivo, al caracterizar el comportamiento del desempeño de los colaboradores, y explicativo, al analizar la relación entre variables operativas y resultados de desempeño mediante modelos analíticos.

Este enfoque metodológico permite reducir la subjetividad en la evaluación del desempeño, proporcionando un soporte objetivo y medible para la gestión organizacional.

Fases de la Metodología

Recolección y Comprensión de los Datos

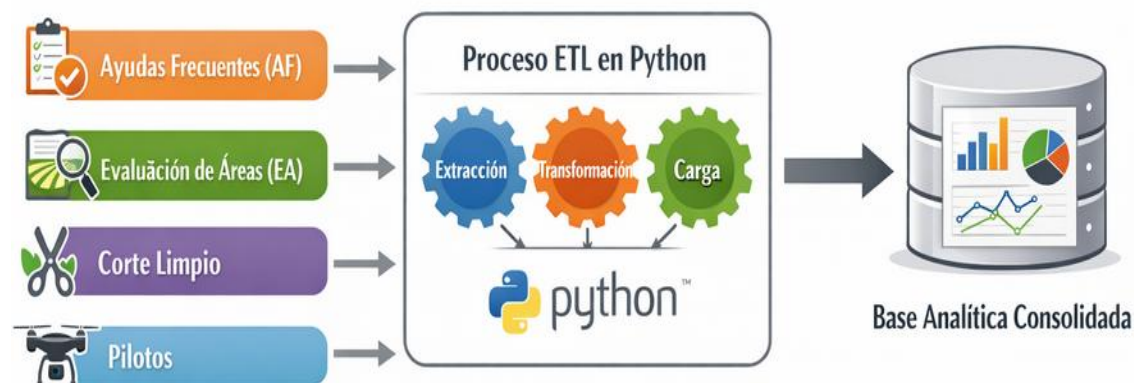
En esta fase se realiza la identificación y recopilación de datos históricos provenientes de las fuentes operativas del área de cultivo de Florval S.A.S., específicamente Ayudas Frecuentes

(AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos. Estas fuentes contienen registros asociados a productividad, calidad y cumplimiento de labores por colaborador y por período.

La información recolectada es integrada mediante un proceso de extracción, transformación y carga (ETL) desarrollado en Python, lo que permite consolidar los datos en una única base analítica estructurada. Durante esta etapa se busca comprender la estructura de los datos, las variables disponibles y su relación con el proceso de evaluación del desempeño.

Figura 1

Recolección y Comprensión de los Datos del Área de Cultivo de Florval S.A.S.



Nota. La figura representa el proceso de recolección e integración de datos provenientes de Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos, los cuales son consolidados mediante un proceso de extracción, transformación y carga (ETL) desarrollado en Python, generando una base analítica consolidada para el análisis del desempeño.

Preparación y Análisis Exploratorio de Datos (EDA)

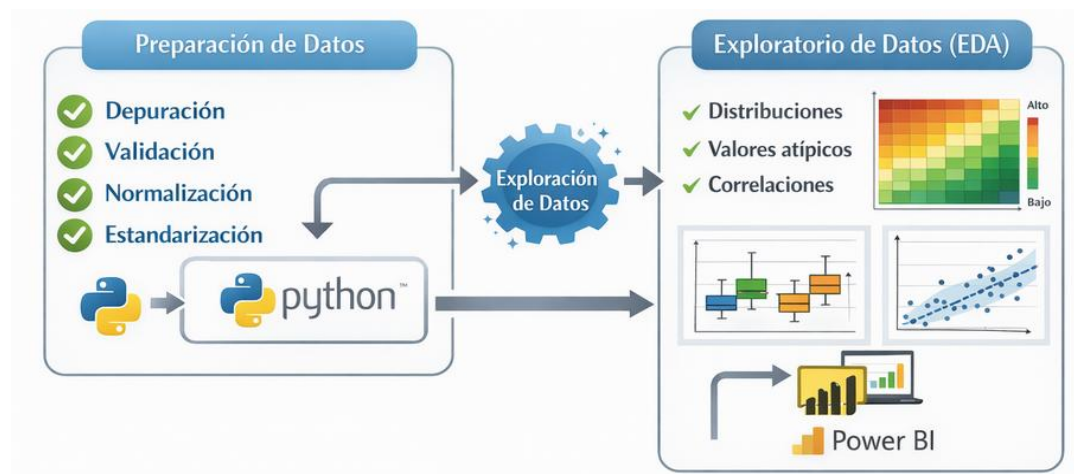
Una vez integrados los datos, se realiza la depuración, validación, normalización y estandarización de la información, con el fin de garantizar su calidad y consistencia antes de la aplicación de los modelos analíticos. Estas actividades se ejecutan principalmente en Python,

permitiendo automatizar el tratamiento de valores faltantes, inconsistencias y formatos heterogéneos.

Posteriormente, se lleva a cabo el análisis exploratorio de datos (EDA), cuyo propósito es comprender el comportamiento general de las variables de desempeño, identificar distribuciones, detectar valores atípicos y analizar relaciones preliminares entre métricas de cantidad y calidad. Como apoyo visual, se utilizan dashboards exploratorios en Power BI, que facilitan la comprensión de grandes volúmenes de datos y permiten observar patrones de forma intuitiva, sin emitir conclusiones definitivas en esta etapa.

Figura 2

Preparación y Análisis Exploratorio de Datos (EDA) para la Evaluación de Desempeño en el Área de Cultivo de Florval S.A.S.



Nota. La figura representa el proceso metodológico de preparación y análisis exploratorio de datos, que incluye actividades de depuración, validación, normalización y estandarización realizadas en Python, así como la exploración de distribuciones, valores atípicos y correlaciones mediante visualizaciones analíticas apoyadas en Power BI.

Figura 3

Vista de los Datos de Cumplimiento por Labor en la Fase de Preparación y Análisis

Exploratorio de Datos (EDA)

do	Finca	Modified	ID.1	Created	AuthorId	EditorId	OData_UIVersionString	Attachments	GUID
9		28/05/2025 10:57:22	3854	28/05/2025 10:57:22	33	33	1.0	FALSE	023b1126-0a44-4155-aa
9		28/05/2025 10:57:22	3855	28/05/2025 10:57:22	33	33	1.0	FALSE	d984508f-68af-4c9e-b09
9		28/05/2025 10:57:24	3856	28/05/2025 10:57:24	33	33	1.0	FALSE	20f93c96-6238-40c5-9af
9		05/06/2025 15:21:37	4193	05/06/2025 15:21:37	33	33	1.0	FALSE	3d918738-82a1-4d9b-b4
9		05/06/2025 15:21:38	4194	05/06/2025 15:21:38	33	33	1.0	FALSE	f4a34790-d9c7-4f27-9c1
9		05/06/2025 15:21:38	4195	05/06/2025 15:21:38	33	33	1.0	FALSE	bc4f3db6-acbe-49f5-b14
9		16/06/2025 16:52:40	4514	16/06/2025 16:52:40	33	33	1.0	FALSE	c27434de-41f4-4d9e-87f
9		16/06/2025 16:52:41	4515	16/06/2025 16:52:41	33	33	1.0	FALSE	f3722c0c-ff85-48e9-a0b4
9		16/06/2025 16:52:41	4516	16/06/2025 16:52:41	33	33	1.0	FALSE	51d2e1e1-4b0c-403d-86
9		27/06/2025 19:20:31	4843	27/06/2025 19:20:31	33	33	1.0	FALSE	22c270e0-1f7d-4285-bd
9		27/06/2025 19:20:31	4844	27/06/2025 19:20:31	33	33	1.0	FALSE	27d2121d-7453-470f-8b
9		27/06/2025 19:20:36	4851	27/06/2025 19:20:36	33	33	1.0	FALSE	3e521654-613c-403c-9fa
9		27/06/2025 19:20:37	4852	27/06/2025 19:20:37	33	33	1.0	FALSE	06b4b11e-351a-41d4-b3
9		27/06/2025 19:20:37	4853	27/06/2025 19:20:37	33	33	1.0	FALSE	97a8e01e-830b-452c-bd
9		03/07/2025 15:35:43	4996	03/07/2025 15:35:43	33	33	1.0	FALSE	4e9e72bb-b21d-4f70-bf
9		11/07/2025 15:39:37	5173	11/07/2025 15:39:37	33	33	1.0	FALSE	e1d1ed00-c749-4b97-b5
9		11/07/2025 15:39:37	5174	11/07/2025 15:39:37	33	33	1.0	FALSE	b46dcdb1f-b60f-4300-97e
9		11/07/2025 15:39:38	5175	11/07/2025 15:39:38	33	33	1.0	FALSE	2319e9bc-a377-4f59-9f0
9		11/07/2025 15:39:39	5176	11/07/2025 15:39:39	33	33	1.0	FALSE	f916e32e-5dd7-4edb-8e
9		11/07/2025 15:39:39	5177	11/07/2025 15:39:39	33	33	1.0	FALSE	8f9b518f-2c66-4c09-89a
9		11/07/2025 15:39:40	5178	11/07/2025 15:39:40	33	33	1.0	FALSE	1b82448f-43bf-4c0e-bda
9		11/07/2025 15:39:41	5179	11/07/2025 15:39:41	33	33	1.0	FALSE	8a2e0cce-4fb7-4980-b9c
9		14/07/2025 16:08:24	5255	14/07/2025 16:08:24	33	33	1.0	FALSE	3ce41918-e14c-412c-9fb
9		14/07/2025 16:08:52	5256	14/07/2025 16:08:52	33	33	1.0	FALSE	c10aea24-adbd-44fa-80e

Nota. La figura presenta un extracto de la base de datos utilizada en la fase de preparación y análisis exploratorio de datos, donde se visualizan los valores de cumplimiento por tipo de labor (descabece, empiole, encanaste, enfibre y limpieza y programación), empleados para la validación de consistencia y estructura de la información previa al modelado analítico.

Modelado Analítico

En esta fase se aplican los modelos analíticos definidos en el marco conceptual, utilizando la base analítica consolidada y preparada. El entrenamiento y ajuste de los modelos se realiza en Python, empleando técnicas estadísticas y de aprendizaje automático acordes con la naturaleza de los datos.

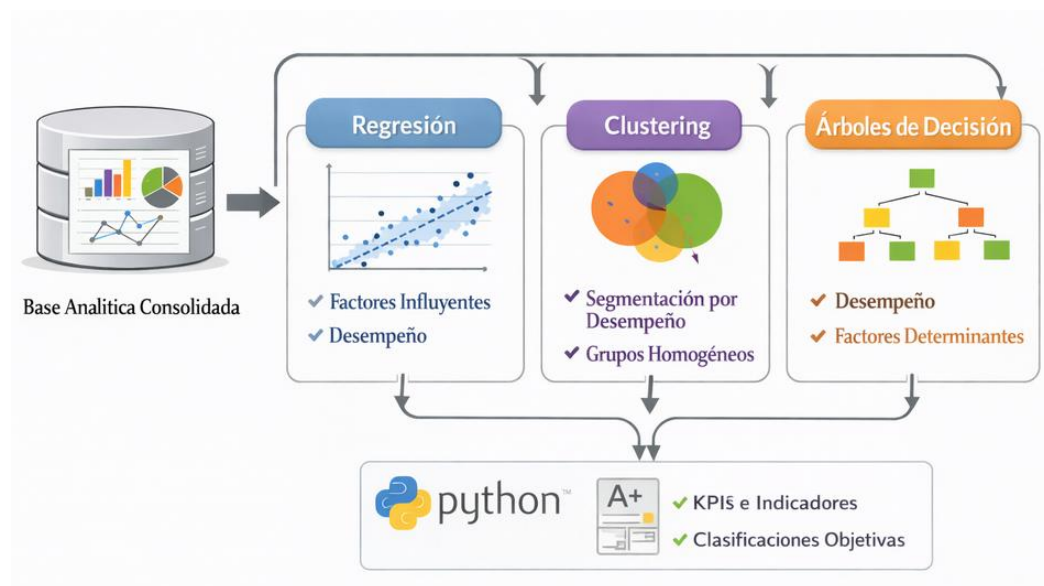
El modelo de regresión se utiliza para analizar la relación entre variables operativas y los resultados de desempeño, permitiendo cuantificar el impacto de cada factor. El modelo de

clustering se aplica para segmentar a los colaboradores en grupos homogéneos según características similares de desempeño. Por su parte, el modelo de árboles de decisión se emplea para clasificar el desempeño y establecer reglas claras e interpretables que faciliten la identificación de factores determinantes.

Esta fase permite generar métricas analíticas objetivas que sirven como insumo para la definición de indicadores clave de desempeño (KPI).

Figura 4

Esquema del Modelado Analítico para la Evaluación del Desempeño en el Área de Cultivo de Florval S.A.S.



Nota. La figura representa el proceso de modelado analítico aplicado a partir de una base analítica consolidada, en el cual se utilizan modelos de regresión para el análisis de relaciones entre variables, técnicas de clustering para la segmentación de colaboradores y árboles de decisión para la clasificación del desempeño e identificación de factores determinantes. El proceso es desarrollado en Python y genera métricas analíticas utilizadas para la definición de indicadores clave de desempeño (KPI)

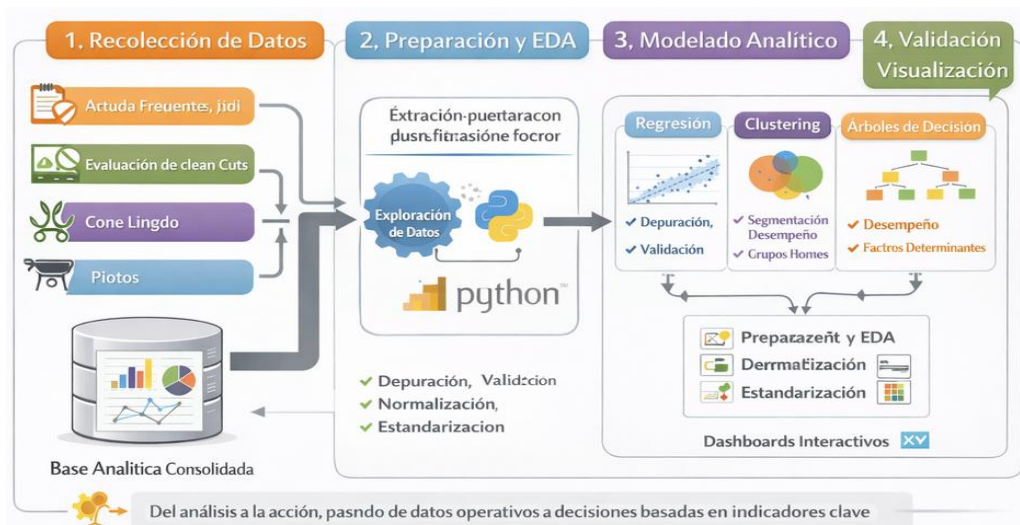
Validación y Visualización

En la fase final se realiza la validación del modelo analítico, comparando los resultados obtenidos con el sistema tradicional de evaluación del desempeño utilizado por la organización. Esta validación permite analizar la coherencia entre ambos enfoques y evaluar el aporte del modelo analítico en términos de objetividad, trazabilidad y soporte a la toma de decisiones.

Los resultados del modelo y los indicadores clave de desempeño (KPI) se presentan mediante dashboards interactivos desarrollados en Power BI, los cuales facilitan la interpretación de la información por parte de los responsables del proceso y apoyan la gestión operativa y estratégica del área de cultivo.

Figura 5

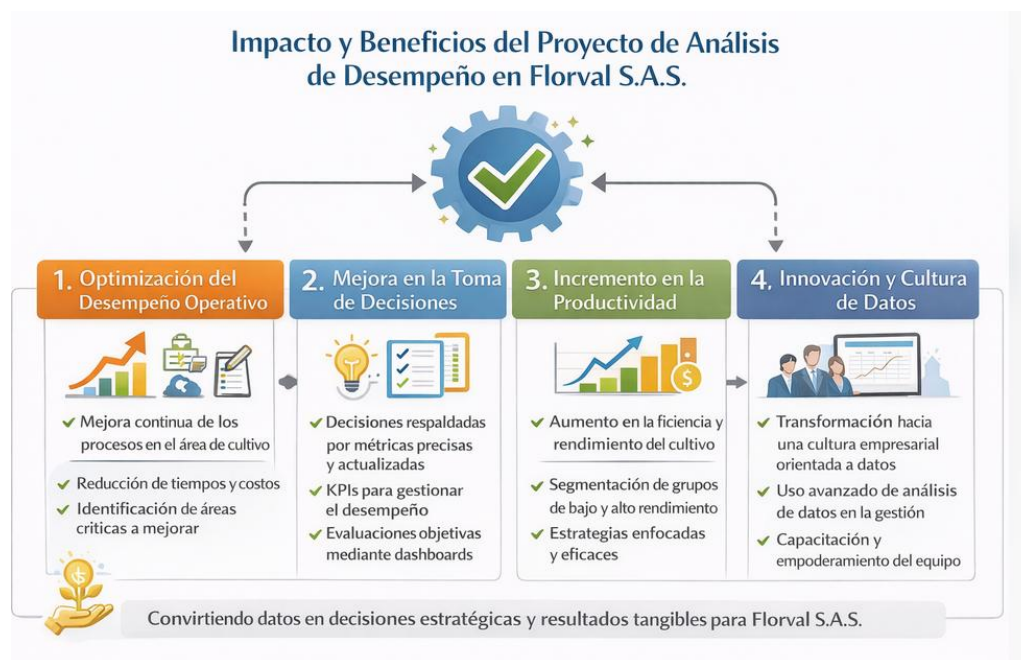
Validación y Visualización de Resultados para la Evaluación del Desempeño en el Área de Cultivo de Florval S.A.S.



Nota. La figura ilustra el proceso de validación del modelo analítico mediante la comparación entre el sistema tradicional de evaluación del desempeño y los resultados analíticos obtenidos. Asimismo, muestra la visualización de indicadores clave de desempeño (KPI) a través de dashboards interactivos desarrollados en Power BI, orientados a apoyar la gestión y la toma de decisiones.

Figura 6

Impacto y Beneficios del Proyecto de Análisis de Desempeño en Florval S.A.S.



Nota. La figura presenta los principales impactos del modelo analítico aplicado al área de cultivo de Florval S.A.S., destacando la optimización del desempeño operativo, la mejora en la toma de decisiones mediante indicadores clave de desempeño (KPI), el incremento en la productividad y el fortalecimiento de una cultura organizacional orientada al uso de datos.

Resultados

Los resultados obtenidos a partir de la implementación del modelo analítico desarrollado para la evaluación del desempeño en el área de cultivo de Florval S.A.S. Los resultados se organizan en función de los componentes principales del proyecto, permitiendo evidenciar el impacto de la analítica de datos en la mejora de los procesos de evaluación, control y toma de decisiones.

El análisis realizado permitió transformar datos operativos históricos en información estructurada y visualmente interpretable, superando las limitaciones del sistema tradicional de evaluación. A lo largo del capítulo se describen los principales hallazgos derivados de la integración de datos, el análisis exploratorio, la aplicación de modelos analíticos y la visualización de indicadores clave de desempeño.

Desarrollo del Modelo Analítico

El desarrollo del modelo analítico constituyó el eje central del proyecto, ya que permitió estructurar un proceso sistemático para el análisis del desempeño del personal del área de cultivo. A partir de la integración de múltiples fuentes de datos operativos, se construyó un modelo capaz de transformar información dispersa en conocimiento útil para la organización.

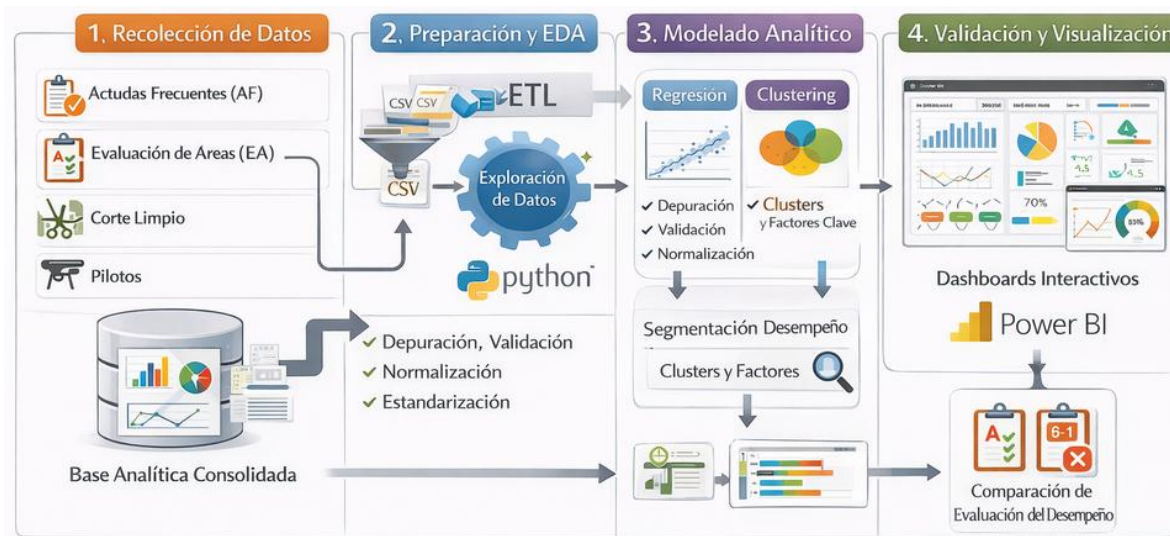
El modelo permitió consolidar datos provenientes de Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos, los cuales anteriormente se encontraban aislados y con diferentes estructuras. Esta integración facilitó el análisis histórico del desempeño y la comparación de resultados entre semanas, bloques y fases productivas.

Adicionalmente, el modelo analítico fortaleció la objetividad del proceso evaluativo, al basarse en métricas cuantificables y reproducibles. La trazabilidad de los datos permitió

identificar el origen de cada indicador y su relación con los resultados finales, lo que representa una mejora sustancial frente al sistema tradicional de evaluación.

Figura 7

Desarrollo del Modelo Analítico para la Evaluación del Desempeño en Florval S.A.S.



Nota. Flujo general del modelo analítico desde la integración de datos hasta la generación de indicadores clave de desempeño.

La Figura 7 presenta el desarrollo integral del modelo analítico aplicado a la evaluación del desempeño en el área de cultivo de Florval S.A.S. En ella se observa el flujo secuencial del proceso, iniciando con la recolección de datos provenientes de Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos, seguido de la preparación y análisis exploratorio de datos (EDA). Posteriormente, se aplican modelos analíticos de regresión y clustering para identificar factores influyentes y segmentar el desempeño, culminando con la validación y visualización de resultados mediante dashboards interactivos en Power BI. Este esquema evidencia la integración estructurada de técnicas de ciencia de datos para fortalecer la objetividad y trazabilidad del proceso evaluativo.

Figura 8

Fragmento del Código del Modelo Analítico para la Composición y Validación de Datos por Fase

```
import pandas as pd
import ipywidgets as widgets
from IPython.display import display

print("[INFO] Ejecutando PASO 7 - Composición por FASE...")

# -----
# Validar DataFrames requeridos
# -----
required = ["cant_largo", "calidad_largo"]
for r in required:
    if r not in globals():
        raise RuntimeError(f"[ERROR] No se encontró el DataFrame '{r}'. Ejecuta los pasos 5 y 6 antes del Paso 7.")

# Helper para buscar columnas
def pick(df, prefer, options):
    if prefer in df.columns:
        return prefer
    for c in options:
        if c in df.columns:
            return c
    return None

# Mapa de equivalencias de labores
labor_equiv = {
    "ENFIBRE Y TENSADO": "ENFIBRE",
    "EMPÍOLE": "EMPÍOLE",
    "EMPÍOLE": "EMPÍOLE",
}

# Copias de trabajo
cant = cant_largo.copy()
cal = calidad_largo.copy()

print(f"[INFO] Registros CANTIDAD (agrupados): {len(cant):,}")
print(f"[INFO] Registros CALIDAD (agrupados): {len(cal):,}")

# Normalizar columnas clave en CANTIDAD
for df in [cant]:
```

Nota. La figura presenta un fragmento del script desarrollado en Python utilizado en el modelo analítico, el cual permite validar la existencia de los conjuntos de datos requeridos, normalizar columnas clave y preparar la información para el análisis del desempeño por fase del cultivo.

La Figura 8 presenta un fragmento del código desarrollado en Python para la composición y validación de los datos por fase productiva dentro del modelo analítico. En este script se realizan procesos de verificación de existencia de conjuntos de datos, normalización de columnas clave y unificación de nomenclaturas de labores, lo que garantiza la consistencia y calidad de la información antes de su análisis. Este procedimiento fue fundamental para asegurar

que los datos utilizados en el modelado analítico estuvieran correctamente estructurados y preparados para la generación de indicadores de desempeño.

Figura 9

Fragmento del Código del Modelo Analítico para la Identificación y Normalización de la Fase Productiva

```
# -----
# Obtener FASE por colaborador/semana
# usando 'ev' si existe, de lo contrario 'af'
# -----
fase_source = None
if "ev" in globals():
    fase_source = "ev"
    base_fase = ev.copy()
elif "af" in globals():
    fase_source = "af"
    base_fase = af.copy()
else:
    base_fase = None

if base_fase is None:
    print("[AVISO] No se encontró 'ev' ni 'af'. 'ed_largo' se generará sin columna 'fase'.")
    ed["fase"] = pd.NA
else:
    print(f"[INFO] La FASE se obtendrá desde el DataFrame '{fase_source}'.")

col_colab_f = pick(base_fase, "colaborador_std", ["Operario", "Title", "Responsable"])
col_sem_f = pick(base_fase, "semana_id", ["Semana", "semana"])
col_fase_f = pick(base_fase, "fase_ev", ["Fase", "fase", "FASE_EV"])

if None in [col_colab_f, col_sem_f, col_fase_f]:
    print("[AVISO] No se pudieron mapear columnas de FASE. 'ed_largo' se generará sin fase.")
    ed["fase"] = pd.NA
else:
    fase_df = base_fase[[col_colab_f, col_sem_f, col_fase_f]].copy()
    fase_df[col_colab_f] = fase_df[col_colab_f].astype(str).str.strip()
    fase_df[col_sem_f] = fase_df[col_sem_f].astype(str).str.strip()
    fase_df["fase_tmp"] = (
        fase_df[col_fase_f]
        .astype(str).str.upper()
        .str.replace(r"\s+", " ", regex=True)
        .str.strip()
    )

# Moda de FASE por colaborador/semana (una sola fase por semana)
fase_map = (
    fase_df
    .dropna(subset=["fase_tmp"])
    .groupby([col_colab_f, col_sem_f])["fase_tmp"]
    .agg(lambda x: x.value_counts().index[0])
    .reset_index()
    .rename(columns={
        col_colab_f: "colaborador_std",
        col_sem_f: "semana_id",
        "fase_tmp": "fase"
    })
)

# Normalizar claves de unión
fase_map["colaborador_std"] = fase_map["colaborador_std"].astype(str).str.strip()
fase_map["semana_id"] = fase_map["semana_id"].astype(str).str.strip()

# Unir fase al DataFrame ed
```

Nota. La figura muestra parte del script encargado de la validación de columnas y normalización de valores relacionados con la fase productiva, asegurando su correcta integración en el modelo analítico.

La Figura 9 presenta un fragmento del código desarrollado en Python para la identificación y normalización de la fase productiva dentro del modelo analítico. En este proceso se valida la existencia de columnas clave, se unifican las denominaciones de las fases y se estandarizan los valores provenientes de distintas fuentes de datos. Este procedimiento garantiza la consistencia estructural de la información antes de su integración en la base analítica consolidada, reduciendo errores derivados de inconsistencias en nomenclaturas y fortaleciendo la confiabilidad del análisis del desempeño por fase.

Figura 10

Cálculo del Indicador de Calidad Mediante el Modelo Analítico Desarrollado en Python

6) Cálculo de CALIDAD (EO-13.09, renormalización de componentes Productivo)

Mostrar código

```

--- [INFO] Ejecutando PASO 6 - Cálculo de calidad...
[INFO] Pilotos cargado desde variable 'pilotos' con 11,968 filas.
[INFO] AF disponible con 24,084 filas.
[INFO] Columnas clave AF: {'colaborador_std': 'colaborador_std', 'semana': 'semana_id', 'labor': 'labor_std'}
[INFO] Columnas métricas AF: {'calidad': 'calidad', 'producto': 'producto', 'proceso': 'proceso', 'rendimiento': 'rendimiento'}
[INFO] Registros de CALIDAD desde AF (labores estándar): 4,741
[INFO] Registros de CALIDAD desde AF (CORTE): 1,815
[INFO] Columnas Pilotos usadas: {'colaborador': 'operario', 'semana': 'semana', 'calificacion': 'calificacion'}
[INFO] Registros de CALIDAD desde Pilotos: 9,593
[OK] calidad_largo generado con 16,149 filas.

```

colab	colaborador_std	semana_id	labor_std	calidad_labor
6556	ABRIL RONDON ADRIANA MARIA	2025-48	PILOTO	1.000000
6557	ABRIL RONDON ADRIANA MARIA	2025-49	PILOTO	0.835000
6558	ABRIL RONDON ADRIANA MARIA	2025-50	PILOTO	0.945000
6559	ABRIL RONDON ADRIANA MARIA	2025-51	PILOTO	0.890000
6560	ABRIL RONDON ADRIANA MARIA	2025-52	PILOTO	1.000000
6561	ABRIL RONDON ADRIANA MARIA	2025-53	PILOTO	0.780000
4741	ABRIL RONDON ADRIANA MARIA	2026-02	CORTE	1.000000
0	ABRIL RONDON ADRIANA MARIA	2026-02	CORTE LIMPIO	0.981132
1	ABRIL RONDON ADRIANA MARIA	2026-02	ENCANASTE	0.833333
6562	ABRIL RONDON ADRIANA MARIA	2026-02	PILOTO	0.780000
4742	ABRIL RONDON ADRIANA MARIA	2026-03	CORTE	1.000000
2	ABRIL RONDON ADRIANA MARIA	2026-03	CORTE LIMPIO	0.935780
3	ABRIL RONDON ADRIANA MARIA	2026-03	EMPIOLE	0.666667
6563	ABRIL RONDON ADRIANA MARIA	2026-03	PILOTO	0.780000
4743	ABRIL RONDON ADRIANA MARIA	2026-04	CORTE	0.982857
4	ABRIL RONDON ADRIANA MARIA	2026-04	CORTE LIMPIO	0.955556
5	ABRIL RONDON ADRIANA MARIA	2026-04	DESCABECE	1.000000
6	ABRIL RONDON ADRIANA MARIA	2026-04	ENCANASTE	0.879630
6564	ABRIL RONDON ADRIANA MARIA	2026-04	PILOTO	1.000000
4744	ABRIL RONDON ADRIANA MARIA	2026-05	CORTE	0.986190
7	ABRIL RONDON ADRIANA MARIA	2026-05	CORTE LIMPIO	0.965517
8	ABRIL RONDON ADRIANA MARIA	2026-05	DESCABECE	0.961538

Nota. La figura representa la ejecución del módulo de cálculo de calidad del modelo analítico, donde se integran registros de distintas fuentes y se aplican procesos de normalización conforme al procedimiento definido.

La Figura 10 muestra el proceso de cálculo del indicador de calidad dentro del modelo analítico, donde se integran registros provenientes de distintas fuentes operativas para generar una métrica consolidada del desempeño. En esta etapa se aplican procedimientos de normalización y ponderación de componentes productivos, permitiendo obtener un valor estandarizado comparable entre colaboradores, bloques y fases. Este cálculo constituye un elemento central del modelo, ya que transforma datos operativos en un indicador cuantitativo que respalda la toma de decisiones basada en evidencia.

Estandarización y Consolidación de los Registros Históricos

La estandarización y consolidación de los registros históricos constituyó uno de los resultados más relevantes del proyecto, dado que permitió resolver una de las principales limitaciones del sistema de Evaluación de Desempeño utilizado previamente en el área de cultivo de Florval S.A.S.: la fragmentación, heterogeneidad y falta de consistencia de la información operativa.

Antes de la implementación del modelo analítico, los datos provenientes de Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos se encontraban distribuidos en diferentes archivos y formatos, con estructuras variables, criterios de registro no homogéneos y nomenclaturas inconsistentes. Esta situación dificultaba la comparación de resultados, la trazabilidad de la información y el análisis histórico del desempeño del personal.

Como resultado del proceso de estandarización, se logró integrar y unificar los registros históricos en una base de datos analítica consolidada, permitiendo transformar información dispersa en un conjunto de datos estructurado, coherente y reutilizable. Este logro representó un cambio sustancial frente al manejo tradicional de la información, al establecer un único repositorio analítico con reglas claras de validación y normalización.

Unificación de Estructuras y Criterios de Registro

Uno de los principales resultados alcanzados fue la homogeneización de la estructura de los datos, independientemente de su fuente de origen. Cada conjunto de registros presentaba inicialmente diferencias en la forma de representar variables clave como colaborador, semana, labor, fase productiva, cantidad y calidad. A través del proceso de estandarización, se definieron campos comunes, tipos de datos consistentes y criterios uniformes de identificación, lo que permitió integrar la información sin pérdida de significado.

Esta unificación estructural facilitó la consolidación de los registros históricos y permitió que datos provenientes de distintas fuentes pudieran ser comparados de manera directa. Como resultado, fue posible analizar el desempeño del personal de forma transversal, considerando múltiples dimensiones del proceso productivo, algo que no era viable con la estructura previa de los datos.

Consolidación de Información Histórica y Trazabilidad

Un resultado clave de la estandarización fue la consolidación efectiva de la información histórica, lo que permitió analizar el desempeño del personal de manera longitudinal. La base de datos consolidada incluye registros por colaborador, semana, labor y fase productiva, lo que facilita el seguimiento del desempeño a lo largo del tiempo.

Esta consolidación permitió mejorar la trazabilidad de la información, ya que cada registro puede ser rastreado desde su fuente original hasta su impacto en los indicadores finales de desempeño. Este aspecto es especialmente relevante para la gestión organizacional, dado que fortalece la transparencia del proceso evaluativo y facilita la auditoría de los resultados.

Soporte para el Análisis Exploratorio y el Modelado Analítico

La base de datos estandarizada y consolidada se convirtió en el insumo fundamental para las etapas posteriores del proyecto, particularmente el análisis exploratorio de datos y la aplicación de modelos analíticos. Gracias a la calidad y consistencia de la información, fue posible identificar patrones, variaciones y relaciones entre variables de desempeño de forma más precisa.

Este resultado demuestra que la estandarización de los datos no solo cumple una función técnica, sino que constituye un habilitador clave del análisis avanzado, permitiendo que los modelos de regresión, clustering y árboles de decisión se alimenten de información confiable y estructurada.

Impacto en el Sistema de Evaluación de Desempeño

Desde una perspectiva organizacional, la estandarización y consolidación de los registros históricos tuvo un impacto directo en la mejora del sistema de Evaluación de Desempeño del área de cultivo. Al disponer de una base de datos única y estructurada, la organización puede evaluar el desempeño de manera más objetiva, consistente y comparable entre colaboradores, bloques y fases productivas.

Este resultado permitió reducir la dependencia de evaluaciones subjetivas y facilitar la toma de decisiones basada en datos, alineando el proceso de evaluación con los principios de mejora continua y gestión por indicadores.

Valor Estratégico del Resultado Obtenido

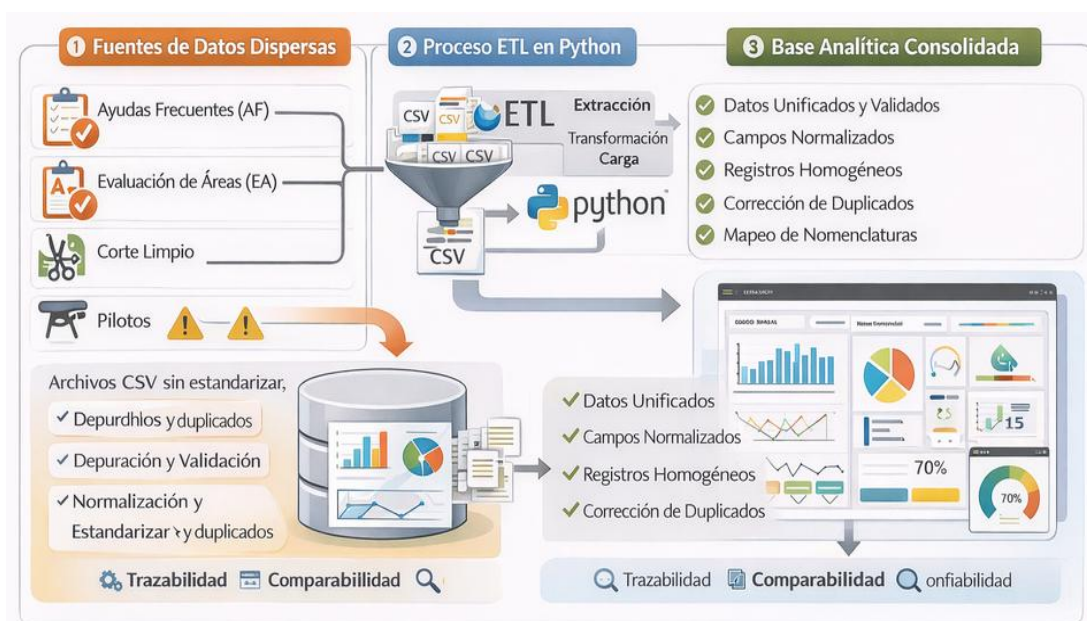
Finalmente, la estandarización y consolidación de los registros históricos generó un valor estratégico para Florval S.A.S., al sentar las bases para la implementación de futuras iniciativas de analítica de datos. La base de datos consolidada puede ser reutilizada y ampliada para nuevos

análisis, incorporando indicadores adicionales o modelos más avanzados sin necesidad de reconstruir el proceso desde cero.

Este resultado posiciona a la organización en un escenario favorable para fortalecer su cultura analítica, mejorar la gestión del desempeño y avanzar hacia una toma de decisiones cada vez más informada y sustentada en datos.

Figura 11

Estandarización y Consolidación de los Registros Históricos para la Evaluación del Desempeño en Florval S.A.S.



Nota. La figura ilustra el flujo de trabajo utilizado para integrar y unificar registros históricos mediante procesos de extracción, transformación y carga (ETL) desarrollados en Python.

La Figura 11 muestra el proceso de estandarización y consolidación de los registros históricos utilizados en la evaluación del desempeño en el área de cultivo. En este flujo se integran datos provenientes de Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio y Pilotos mediante un proceso ETL desarrollado en Python, que incluye actividades de

depuración, validación, transformación y unificación de nomenclaturas. Como resultado, se obtiene una base analítica consolidada con datos homogéneos y estructurados, lo que garantiza trazabilidad, comparabilidad y confiabilidad en el análisis del desempeño.

Figura 12

Vista de la Base de Datos Consolidada tras el Proceso de Estandarización de Registros

Históricos

Bloque	Fase	Responsable	Modificado	ID	Año	Semana	AñoSemana
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69309	2026	3	2603
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69310	2026	3	2603
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69312	2026	3	2603
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69316	2026	3	2603
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69317	2026	3	2603
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69318	2026	3	2603
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69319	2026	3	2603
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69323	2026	3	2603
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69304	2026	3	2603
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69305	2026	3	2603
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	17/01/2026 11:38:00	69308	2026	3	2603
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68858	2026	2	2602
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68823	2026	2	2602
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68824	2026	2	2602
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68826	2026	2	2602
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68827	2026	2	2602
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68831	2026	2	2602
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68832	2026	2	2602
1	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68833	2026	2	2602
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68838	2026	2	2602
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68839	2026	2	2602
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68840	2026	2	2602
3	PRODUCCIÓN	Doris Consuelo Contreras Trivino	10/01/2026 11:25:00	68841	2026	2	2602

Nota. La figura muestra un extracto de la base analítica consolidada utilizada en el proyecto, resultado del proceso de normalización y consolidación de datos históricos del área de cultivo de Florval S.A.S.

La Figura 12 presenta un extracto de la base de datos analítica consolidada obtenida después del proceso de estandarización e integración de los registros históricos. En esta estructura se evidencian campos normalizados como bloque, fase productiva, responsable, identificador del registro, año y semana, los cuales permiten realizar análisis comparativos y

seguimiento temporal del desempeño. Esta base consolidada constituye el insumo principal para la aplicación de modelos analíticos y la generación de indicadores clave que fortalecen la toma de decisiones en el área de cultivo.

Análisis Exploratorio del Desempeño (EDA)

El análisis exploratorio de datos (EDA) se realizó sobre la base de datos analítica consolidada, resultado del proceso de estandarización e integración de los registros históricos provenientes de las fuentes operativas del área de cultivo de Florval S.A.S. Este análisis tuvo como objetivo comprender la estructura de la información, evaluar su calidad y caracterizar el comportamiento general del desempeño antes de la aplicación de modelos analíticos.

Tamaño y composición de los conjuntos de datos:

Como resultado del proceso de integración, se consolidaron los siguientes volúmenes de información:

Tabla 1

Tamaño y Composición de los Conjuntos de Datos Utilizados en el Análisis Exploratorio (EDA)

Conjunto de datos	Número de registros	Descripción
Ayudas Frecuentes (AF)	24.084	Registros históricos asociados a labores estándar del área de cultivo, que contienen información de cantidad, calidad y cumplimiento por colaborador y semana.
Corte Limpio	1.815	Registros correspondientes a evaluaciones específicas de la labor de corte limpio, integrados como parte del análisis del indicador de calidad.
Pilotos	9.593	Registros asociados a labores piloto, utilizados principalmente para el cálculo, validación y ajuste del indicador de calidad.

Conjunto de datos	Número de registros	Descripción
Base consolidada de calidad (calidad_largo)	16.149	Registros finales obtenidos tras la integración, depuración y renormalización de los datos provenientes de Ayudas Frecuentes, Corte Limpio y Pilotos.
Base consolidada de cantidad (cant_largo)	8.932	Dataset estructurado por colaborador, semana y labor, utilizado para el análisis del cumplimiento de cantidad y su posterior ponderación según la fase productiva.

Nota. La tabla resume los volúmenes de registros correspondientes a las fuentes originales y a las bases de datos consolidadas empleadas en la fase de análisis exploratorio del desempeño en el área de cultivo de Florval S.A.S.

La Tabla 1 presenta el tamaño y la composición de los conjuntos de datos empleados en la fase de análisis exploratorio del desempeño. Se observa la cantidad total de registros provenientes de las diferentes fuentes operativas, así como el volumen de datos consolidados utilizados para el modelado analítico. Esta información permitió dimensionar la magnitud del proceso de integración y validar que la base de datos contara con suficiente representatividad histórica para garantizar la solidez del análisis posterior.

Estos conjuntos de datos constituyeron el insumo principal del análisis exploratorio y permitieron realizar un estudio representativo del desempeño del personal del área de cultivo durante el período analizado.

Cobertura Temporal y Organizacional

El análisis exploratorio abarcó registros correspondientes a los años 2025 y 2026, organizados por semana, bloque, fase productiva y tipo de labor. La estructura temporal de los

datos permitió evaluar el desempeño de forma histórica y comparativa, facilitando la identificación de tendencias, variaciones y comportamientos recurrentes.

Desde el punto de vista organizacional, los registros incluyen información asociada a múltiples bloques del cultivo y a diferentes responsables, lo que permitió analizar el desempeño desde una perspectiva transversal y no limitada a un solo frente operativo.

Identificación de patrones generales de desempeño

A partir del análisis exploratorio se identificaron patrones de comportamiento diferenciados en el desempeño del personal según bloque, tipo de labor y fase productiva. Las visualizaciones exploratorias evidenciaron variaciones en los niveles de cumplimiento, así como diferencias en la dispersión de los resultados entre fases vegetativas y productivas.

Aporte del EDA al Modelo Analítico

El análisis exploratorio del desempeño cumplió un papel fundamental como etapa previa al modelado, ya que permitió comprender la estructura real de los datos, validar su calidad y definir los criterios para la construcción de indicadores clave de desempeño (KPI).

- Los resultados del EDA sirvieron como base para:
- La selección de variables relevantes para el modelado.
- La segmentación del desempeño por fase productiva.
- La definición de reglas de ponderación de cantidad y calidad.
- La identificación de comportamientos atípicos que requerían análisis específico.

En conjunto, el EDA permitió asegurar que los modelos analíticos se alimentaran de información confiable, representativa y estructurada, fortaleciendo la validez de los resultados obtenidos en las etapas posteriores del proyecto.

Aplicación de Modelos Analíticos

La aplicación de modelos analíticos constituyó una de las fases centrales del proyecto, ya que permitió pasar del análisis descriptivo del desempeño a un enfoque explicativo y clasificatorio basado en datos. A partir de la base analítica consolidada, se implementaron modelos de regresión, clustering y árboles de decisión, con el objetivo de analizar el comportamiento del desempeño del personal, identificar factores influyentes y generar clasificaciones objetivas que fortalecieran el sistema de Evaluación de Desempeño del área de cultivo de Florval S.A.S.

Los modelos se aplicaron sobre conjuntos de datos previamente depurados y normalizados, garantizando la consistencia de la información utilizada y la validez de los resultados obtenidos.

Modelo de Regresión Análisis de Factores Influyentes en el Desempeño

El modelo de regresión se utilizó para analizar la relación existente entre las variables operativas y los resultados de desempeño. A través de este modelo fue posible evaluar el impacto relativo de variables asociadas a cantidad, calidad, tipo de labor y fase productiva sobre el nivel de cumplimiento alcanzado por los colaboradores.

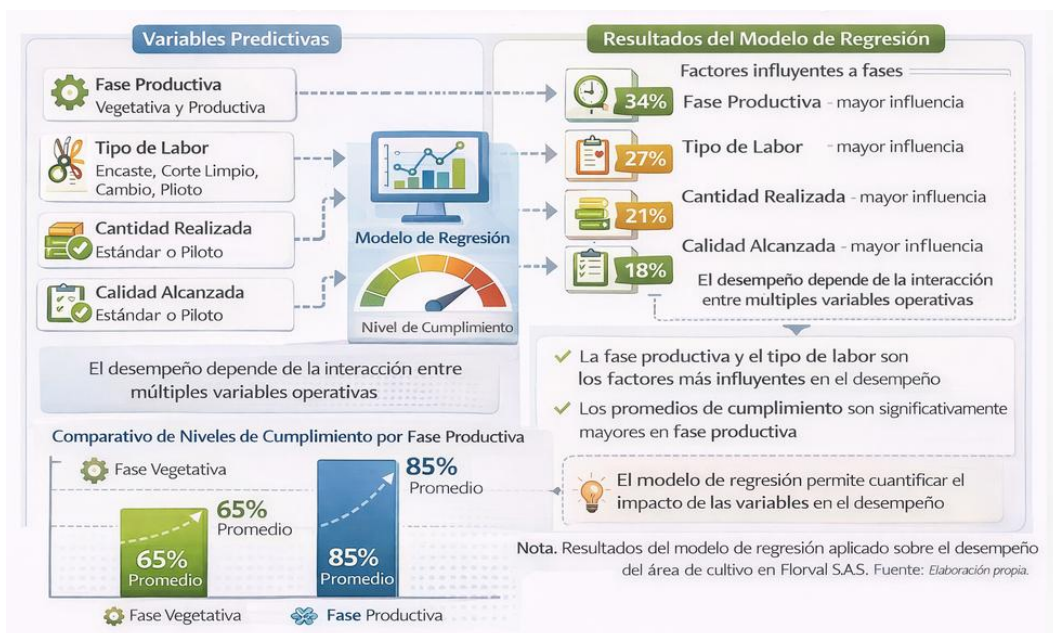
Los resultados del modelo evidenciaron que determinadas variables operativas tienen una influencia significativa en el desempeño, especialmente aquellas asociadas a la fase productiva y al tipo de labor realizada. Este análisis permitió cuantificar el efecto de dichos factores, aportando una base objetiva para comprender por qué se presentan diferencias en los niveles de desempeño entre colaboradores y períodos.

El uso del modelo de regresión permitió, además, validar que el desempeño no depende de un único factor, sino de la interacción entre múltiples variables operativas. Este resultado

aportó un entendimiento más profundo del proceso productivo y permitió sustentar la necesidad de evaluar el desempeño desde una perspectiva integral y no aislada.

Figura 13

Resultados del Modelo de Regresión para el Análisis de Factores Influyentes en el Desempeño



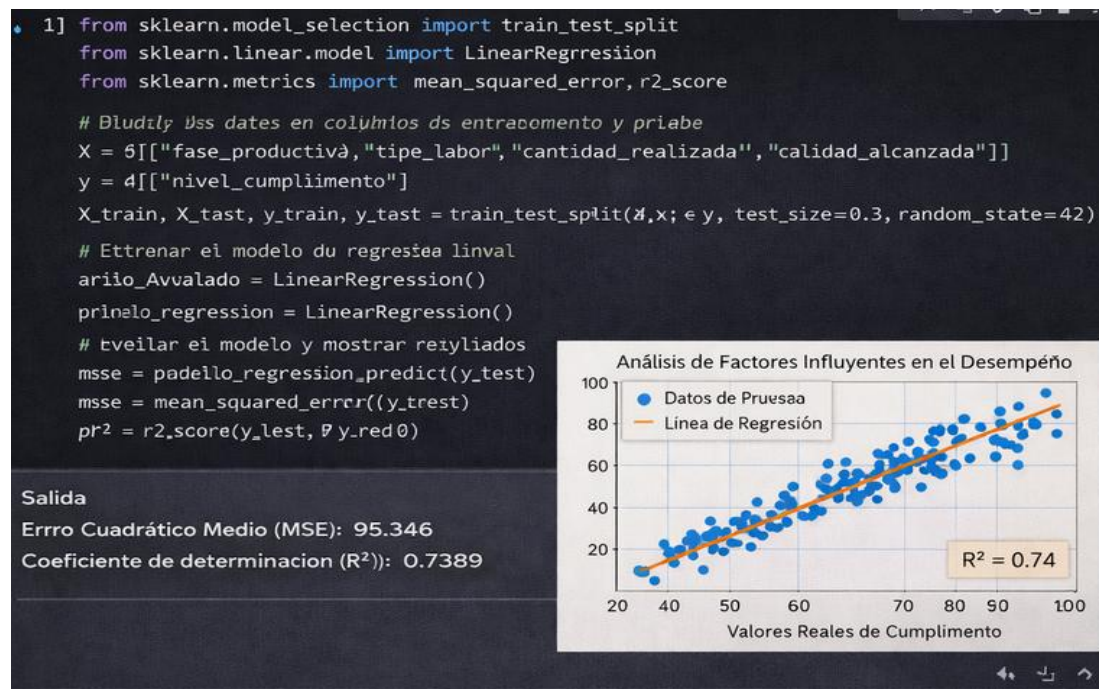
Nota. La figura presenta los resultados del modelo de regresión aplicado al desempeño del personal del área de cultivo de Florval S.A.S., mostrando la influencia relativa de las variables analizadas.

La Figura 13 presenta los resultados obtenidos del modelo de regresión aplicado al análisis de los factores influyentes en el desempeño de los colaboradores del área de cultivo. Se observa que variables como la fase productiva y el tipo de labor presentan mayor incidencia en los niveles de cumplimiento, mientras que otras variables tienen una influencia moderada. Los resultados evidencian que el desempeño no depende de un único factor aislado, sino de la interacción entre múltiples variables operativas. Este hallazgo permitió validar la necesidad de

un enfoque analítico integral frente al sistema tradicional de evaluación, fortaleciendo la objetividad del análisis.

Figura 14

Ejecución del Modelo de Regresión en Google Colab para el Análisis de Factores Influyentes en el Desempeño



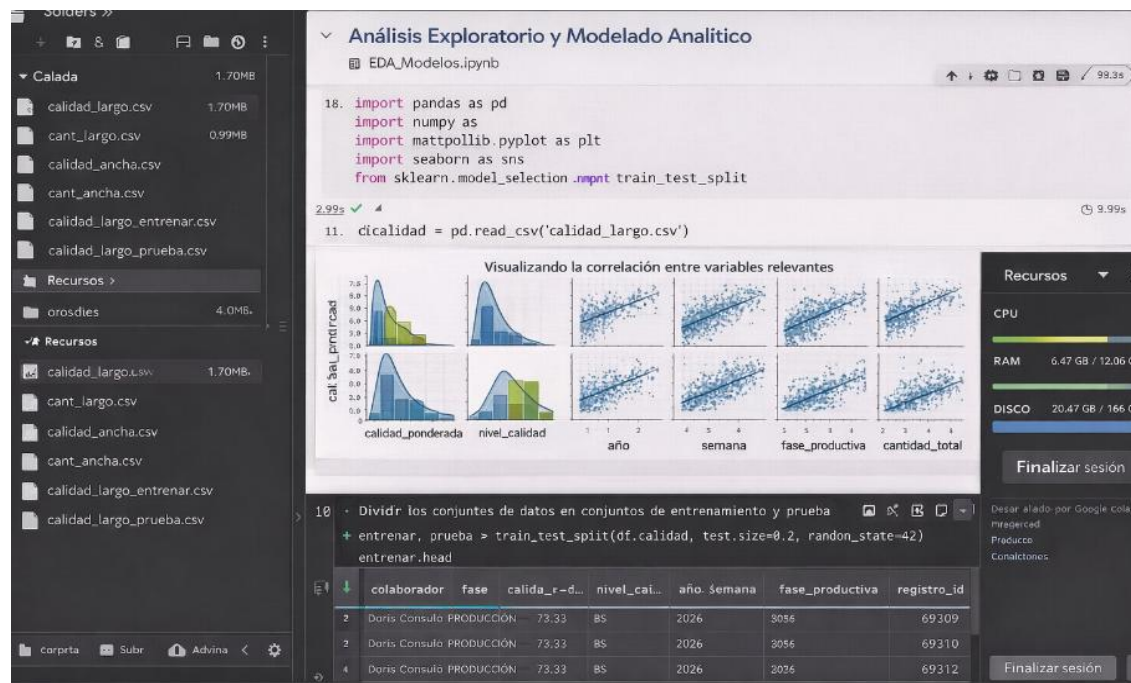
Nota. La figura presenta la ejecución del modelo de regresión lineal, incluyendo el cálculo de métricas de evaluación y la visualización gráfica del ajuste entre valores reales y estimados.

La Figura 16 muestra la ejecución del modelo de regresión lineal en el entorno Google Colab, incluyendo el código desarrollado en Python para la división de datos en conjuntos de entrenamiento y prueba. En esta etapa se calcularon métricas de desempeño como el error cuadrático medio (MSE) y el coeficiente de determinación (R^2), permitiendo evaluar la capacidad predictiva del modelo. Asimismo, la visualización gráfica de la relación entre valores

reales y estimados evidencia el nivel de ajuste alcanzado, validando la utilidad del modelo como herramienta complementaria para el análisis del desempeño.

Figura 15

Ejecución del Análisis Exploratorio y Modelado Analítico en Google Colab



Nota. La figura presenta el entorno de trabajo en Google Colab utilizado para el análisis exploratorio y el modelado analítico, donde se ejecutan los scripts en Python empleados en el proyecto.

La Figura 15 muestra el entorno de trabajo en Google Colab utilizado para el desarrollo del análisis exploratorio de datos (EDA) y la ejecución del modelado analítico. En este espacio se integraron los scripts en Python para la depuración, validación y transformación de datos, así como para la aplicación de los modelos de regresión y clustering. La utilización de este entorno permitió garantizar la trazabilidad del proceso analítico y asegurar la reproducibilidad de los resultados obtenidos en el proyecto.

Modelo de Clustering Segmentación del Desempeño de los Colaboradores

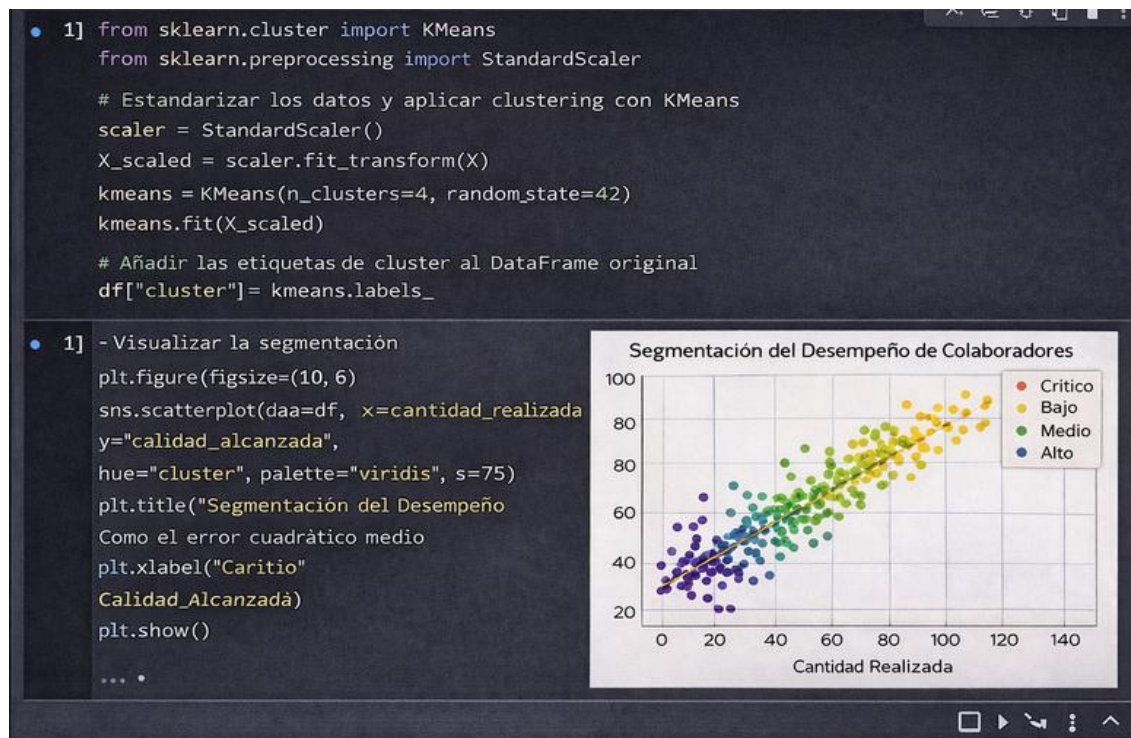
El modelo de clustering se aplicó con el fin de segmentar a los colaboradores en grupos homogéneos, de acuerdo con características similares de desempeño. Este enfoque permitió identificar patrones de comportamiento recurrentes y clasificar a los colaboradores en diferentes niveles de rendimiento sin recurrir a criterios subjetivos.

Como resultado del clustering, se identificaron grupos claramente diferenciados de colaboradores con desempeños altos, medios y bajos. Esta segmentación permitió observar que los colaboradores con alto desempeño presentan comportamientos consistentes a lo largo del tiempo, mientras que los grupos de desempeño medio y bajo muestran mayor variabilidad en sus resultados.

El modelo de clustering aportó información clave para la gestión del talento, ya que permitió identificar grupos que requieren seguimiento específico o estrategias de mejora focalizadas. Asimismo, esta segmentación facilitó la comparación del desempeño entre colaboradores con características similares, fortaleciendo la equidad del proceso evaluativo.

Figura 16

Ejecución del Modelo de Clustering para la Segmentación del Desempeño de los Colaboradores en Google Colab



Nota. La figura muestra la ejecución del modelo de clustering (K-means) en Google Colab, donde se observa la segmentación del desempeño en diferentes grupos según variables de cantidad y calidad.

La Figura 16 presenta la ejecución del modelo de clustering (K-means) aplicado para la segmentación del desempeño de los colaboradores del área de cultivo. A partir de variables como cantidad realizada y calidad alcanzada, el modelo permitió identificar grupos diferenciados de desempeño alto, medio y bajo. Esta segmentación evidencia patrones de comportamiento similares entre colaboradores y facilita la identificación de niveles diferenciados de rendimiento, aportando una herramienta objetiva para la gestión y seguimiento del talento humano.

Modelo de Árboles de Decisión y Clasificación e Interpretación del Desempeño

El modelo de árboles de decisión se aplicó con el objetivo de clasificar el desempeño y establecer reglas claras que expliquen las condiciones bajo las cuales se alcanzan determinados niveles de cumplimiento. Este modelo permitió generar una representación visual y fácilmente interpretable del proceso de evaluación del desempeño.

Los resultados del árbol de decisión evidenciaron que variables como la fase productiva, el tipo de labor y los niveles de cantidad y calidad influyen directamente en la clasificación final del desempeño. A través de reglas lógicas derivadas de los datos, fue posible identificar combinaciones de condiciones que conducen a niveles altos o bajos de rendimiento.

Este modelo destacó por su alta interpretabilidad, permitiendo que los responsables del área comprendieran de manera clara los criterios que influyen en la evaluación del desempeño. En comparación con el sistema tradicional, el árbol de decisión ofreció mayor transparencia y coherencia en la clasificación de los colaboradores.

Resultados Comparativos Frente al Sistema Tradicional de Evaluación

Uno de los principales resultados de la aplicación de los modelos analíticos fue la comparación con el sistema tradicional de evaluación del desempeño utilizado por la organización. Esta comparación permitió evidenciar diferencias significativas entre ambos enfoques.

Los resultados mostraron que el modelo analítico proporciona mayor consistencia y precisión en la evaluación del desempeño, al reducir la dependencia de juicios subjetivos y basarse en datos históricos cuantificables. Asimismo, el enfoque analítico permitió discriminar de forma más clara entre distintos niveles de desempeño, evitando clasificaciones ambiguas o poco justificadas.

En particular, se evidenció que el sistema tradicional tiende a homogeneizar los resultados, mientras que el modelo analítico identifica con mayor claridad diferencias reales en el rendimiento, aportando información más útil para la toma de decisiones.

Generación de Métricas Analíticas y Clasificación del Desempeño

Como resultado de la aplicación de los modelos analíticos, se generaron métricas objetivas de desempeño, las cuales sirvieron como base para la construcción de indicadores clave (KPI). Estas métricas incluyen calificaciones ponderadas por fase productiva, clasificaciones por niveles y análisis históricos del desempeño semanal.

La clasificación obtenida permitió identificar colaboradores con desempeño sobresaliente, consistente y crítico, facilitando la toma de decisiones relacionadas con seguimiento, retroalimentación y compensación. Este resultado fortaleció el sistema de evaluación al proporcionar información clara, comparable y orientada a la acción.

Integración de los Resultados en Herramientas de Visualización

El desarrollo del modelo analítico permitió consolidar, integrar y procesar información proveniente de múltiples fuentes internas de Florval S.A.S., tales como Ayudas Frecuentes (AF), Evaluación de Áreas (EA), Corte Limpio, Pilotos, parámetros por fase y registros históricos de desempeño. Esta integración posibilitó la construcción de una base de datos analítica estructurada, a partir de la cual se generaron indicadores objetivos, consistentes y comparables para la evaluación semanal del desempeño de los colaboradores del área de cultivo.

El procesamiento de la información se realizó mediante un script automatizado desarrollado en Python (ver Figura 23), el cual permitió ejecutar de manera sistemática la carga, validación, depuración y normalización de los archivos CSV requeridos. Este proceso garantizó la consistencia y calidad de los datos, así como la generación de métricas ponderadas de cantidad y calidad, ajustadas a las particularidades de cada fase del cultivo. De esta manera, se aseguró la trazabilidad completa de la información, desde su origen en las fuentes operativas hasta su utilización en los análisis y visualizaciones finales.

Los resultados consolidados del modelo analítico fueron representados mediante un dashboard interactivo desarrollado en Power BI (ver Figura 20), en el cual se presentan indicadores clave de desempeño, tales como calificación total, desempeño por fase productiva, promedio de las últimas cinco semanas y clasificación por niveles. Esta visualización facilitó la identificación de diferencias significativas en el desempeño entre las fases productivas y vegetativas, evidenciando que la fase productiva presenta, en promedio, mejores niveles de desempeño en comparación con las fases vegetativas.

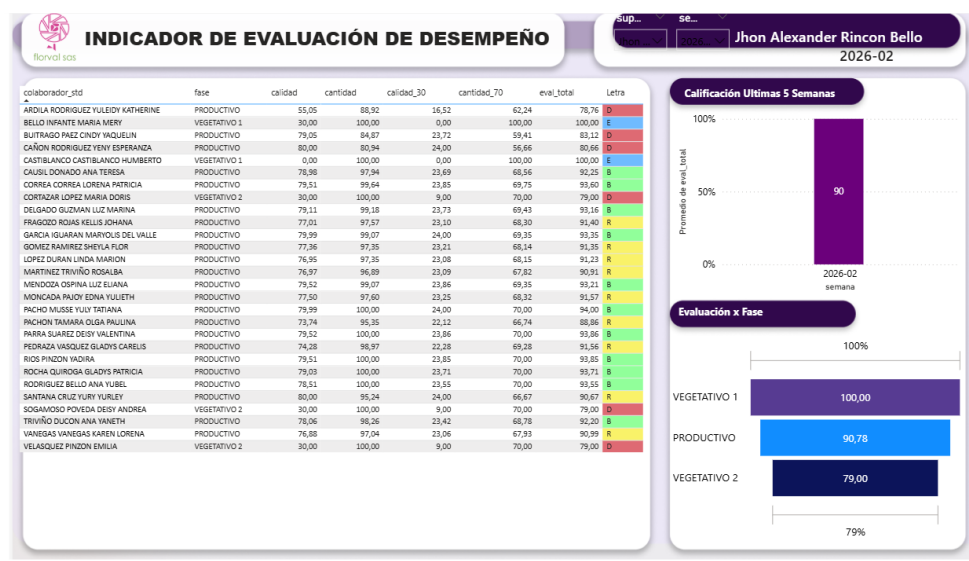
Adicionalmente, se implementó un módulo específico para el análisis de compensatorios (ver Figura 21), el cual permitió evaluar el comportamiento histórico de las calificaciones

semanales y determinar si los colaboradores cumplen con los criterios establecidos por la organización para acceder a compensaciones. Los resultados obtenidos indican que, para el período analizado, la mayoría de los colaboradores no alcanza el umbral requerido para compensar, lo que pone de manifiesto la necesidad de fortalecer el seguimiento continuo del desempeño y de aplicar estrategias de mejora focalizadas.

En conjunto, los resultados demuestran que la implementación del modelo analítico, junto con su visualización mediante herramientas de inteligencia de negocios, contribuye de manera significativa a mejorar la objetividad, el seguimiento y la interpretación del desempeño de los colaboradores del área de cultivo. Este enfoque supera las limitaciones del sistema tradicional de evaluación previamente utilizado por la organización, al proporcionar información trazable, analíticamente sustentada y orientada a la toma de decisiones basada en datos.

Figura 18

Dashboard del Indicador de Evaluación de Desempeño Semanal por Colaborador y Fase

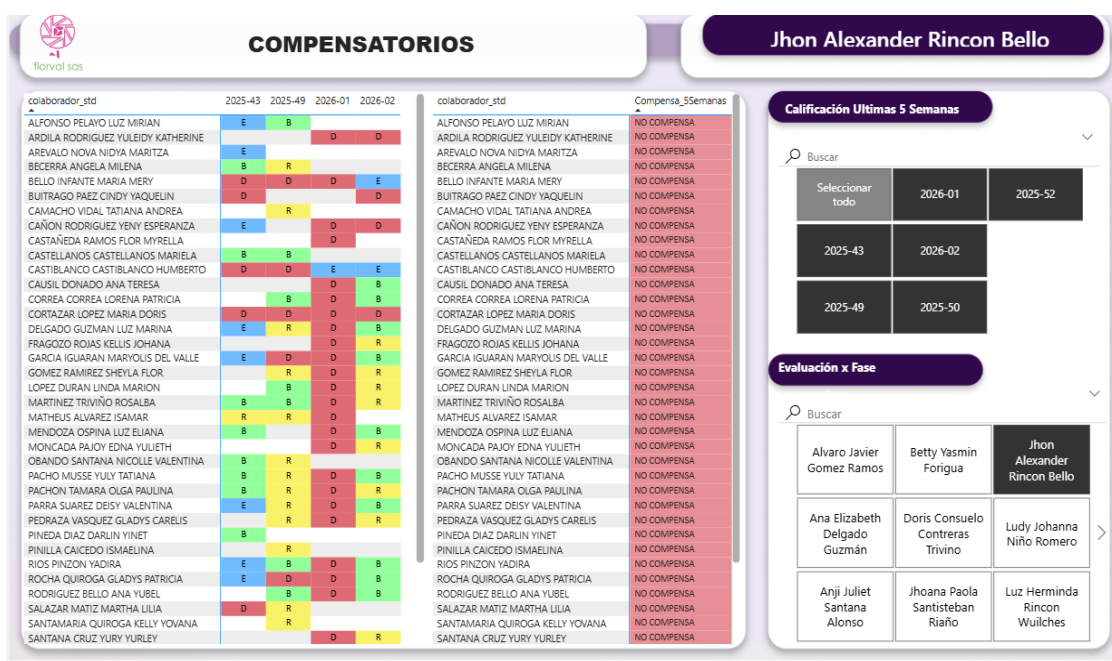


Nota. Visualización generada en Power BI a partir del modelo analítico desarrollado para la evaluación del desempeño en Florval S.A.S.

La Figura 18 presenta el dashboard interactivo desarrollado en Power BI para el seguimiento del indicador de evaluación de desempeño semanal por colaborador y fase productiva. En esta visualización se integran los resultados obtenidos del modelo analítico, permitiendo consultar métricas consolidadas de cantidad, calidad y cumplimiento, así como su clasificación por niveles de desempeño. Este tablero facilita el análisis comparativo entre colaboradores y fases, fortaleciendo la toma de decisiones basada en datos y aportando mayor transparencia al proceso evaluativo.

Figura 19

Módulo de Análisis de Compensatorios y Calificación Histórica por Semanas



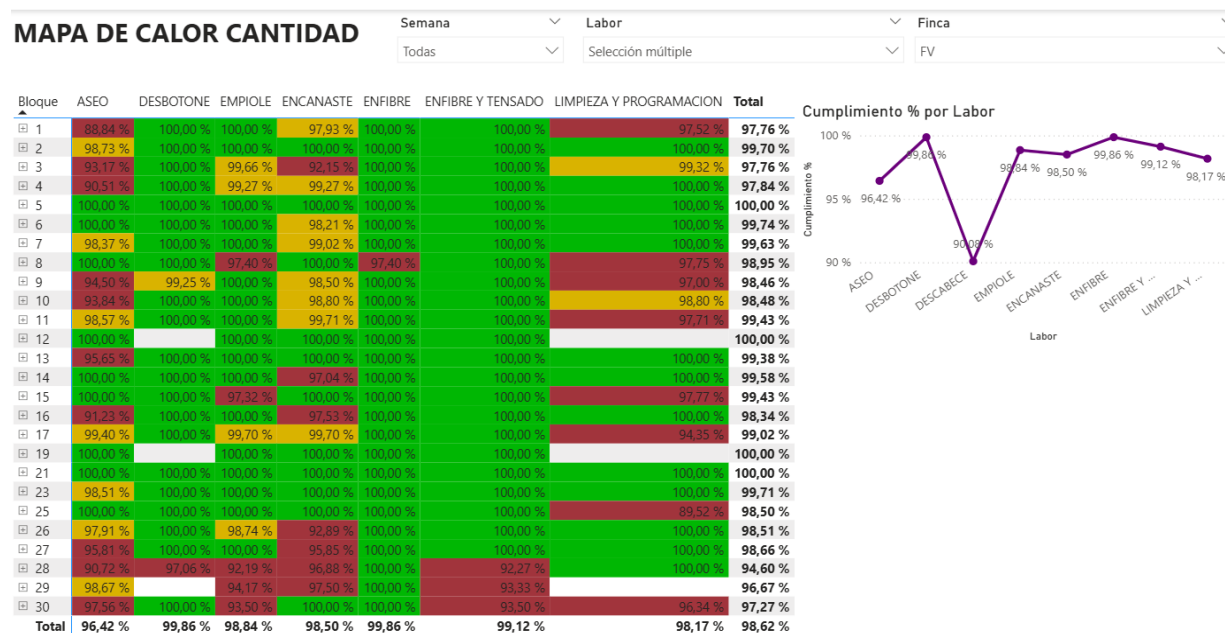
Nota. Resultados obtenidos del análisis histórico de desempeño semanal para la identificación de compensaciones.

La Figura 19 muestra el módulo de análisis histórico de desempeño utilizado para la identificación de compensatorios y seguimiento por semanas. En esta visualización se consolidan

los resultados históricos de cantidad y calidad, permitiendo analizar tendencias, variaciones temporales y comportamientos recurrentes en el desempeño de los colaboradores. Este módulo aporta un enfoque longitudinal al sistema de evaluación, facilitando la toma de decisiones relacionadas con incentivos, compensaciones y estrategias de mejora continua.

Figura 20

Mapa de Calor del Cumplimiento de Cantidad por Bloque y Labor



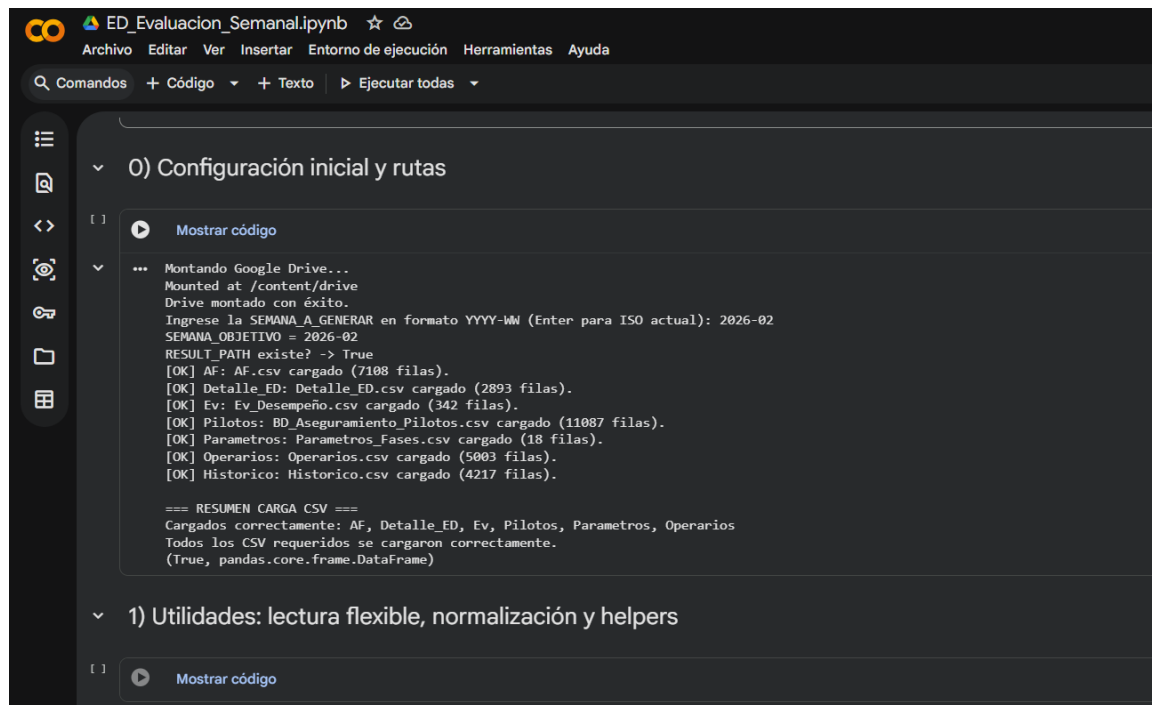
Nota. Visualización generada en Power BI a partir de los datos consolidados del modelo analítico de evaluación del desempeño en el área de cultivo de Florval S.A.S.

La Figura 20 presenta el mapa de calor del cumplimiento de cantidad por bloque y tipo de labor, generado a partir de los resultados consolidados del modelo analítico. La escala de colores permite identificar de manera visual los niveles de desempeño, evidenciando bloques con cumplimiento alto (tonos verdes), intermedio (amarillos) y bajo (rojos). Esta representación

facilita la detección rápida de desviaciones operativas y permite focalizar acciones correctivas en bloques o labores específicas, fortaleciendo la gestión del desempeño basada en datos.

Figura 21

Ejecución del Script de Procesamiento y Consolidación de Datos en Python



```
ED_Evaluacion_Semanal.ipynb
Archivo Editar Ver Insertar Entorno de ejecución Herramientas Ayuda
Comandos + Código + Texto Ejecutar todas
0) Configuración inicial y rutas
Mostrar código
Montando Google Drive...
Mounted at /content/drive
Drive montado con éxito.
Ingrese la SEMANA_A_GENERAR en formato YYYY-MM (Enter para ISO actual): 2026-02
SEMANA_OBJETIVO = 2026-02
RESULT PATH existe? -> True
[OK] AF: AF.csv cargado (7108 filas).
[OK] Detalle_ED: Detalle_ED.csv cargado (2893 filas).
[OK] Ev: Ev_Desempeño.csv cargado (342 filas).
[OK] Pilotos: BD_Aseguramiento_Pilotos.csv cargado (11087 filas).
[OK] Parametros: Parametros_Fases.csv cargado (18 filas).
[OK] Operarios: Operarios.csv cargado (5003 filas).
[OK] Historico: Historico.csv cargado (4217 filas).

=== RESUMEN CARGA CSV ===
Cargados correctamente: AF, Detalle_ED, Ev, Pilotos, Parametros, Operarios
Todos los CSV requeridos se cargaron correctamente.
(True, pandas.core.frame.DataFrame)

1) Utilidades: lectura flexible, normalización y helpers
Mostrar código
```

Nota. Proceso de carga, validación y normalización de archivos CSV utilizados como base del modelo analítico de evaluación del desempeño.

La Figura 21 muestra la ejecución del script desarrollado en Python para el procesamiento, validación y consolidación de los datos utilizados en el modelo analítico. En este procedimiento se realiza la carga de archivos CSV, la validación de estructuras, la normalización de columnas clave y la preparación de la base analítica consolidada. Este proceso constituye la base técnica del modelo, ya que garantiza la calidad, consistencia y trazabilidad de la información antes de la aplicación de técnicas de análisis y modelado.

Conclusiones

La integración de los datos provenientes de Ayudas Frecuentes, Evaluación de Áreas, Corte Limpio y Pilotos permitió consolidar información que anteriormente se encontraba fragmentada, evidenciando la importancia de una gestión estructurada de los datos para mejorar el proceso de evaluación de desempeño en el área de cultivo de Florval S.A.S..

El análisis exploratorio de datos y la aplicación de modelos analíticos demostraron que el desempeño de los colaboradores está influenciado por múltiples variables relacionadas con productividad, calidad y cumplimiento, las cuales no son plenamente capturadas por el sistema tradicional de evaluación, lo que limita su objetividad.

La implementación de técnicas de ciencia de datos, como regresión, clustering y árboles de decisión, aportó un enfoque analítico que facilita la segmentación de colaboradores, la identificación de patrones de rendimiento y la generación de información explicativa y predictiva para la toma de decisiones.

El uso de un dashboard analítico en Power BI permitió integrar y visualizar los indicadores clave de desempeño de manera clara y accesible, fortaleciendo el seguimiento, el análisis comparativo y el soporte a decisiones operativas y estratégicas dentro de la organización.

En conjunto, los resultados obtenidos confirman que la analítica de datos constituye una herramienta eficaz para optimizar el Sistema de Evaluación de Desempeño en el área de cultivo de Florval S.A.S., aportando mayor precisión, trazabilidad y consistencia al proceso evaluativo.

Recomendaciones

Se recomienda a Florval S.A.S. implementar de manera progresiva el modelo analítico desarrollado como complemento al sistema tradicional de evaluación de desempeño, con el fin de fortalecer la objetividad del proceso y mejorar la calidad de la información utilizada en la toma de decisiones.

Es conveniente establecer un proceso continuo de recolección, depuración y actualización de los datos provenientes de AF, EA, Corte Limpio y Pilotos, garantizando la calidad y consistencia de la información utilizada para el análisis.

Se sugiere capacitar a supervisores y personal administrativo en el uso e interpretación de herramientas de analítica de datos y dashboards, con el propósito de fomentar una cultura organizacional orientada a la toma de decisiones basada en datos.

Para futuras investigaciones, se recomienda ampliar el alcance del proyecto incorporando técnicas avanzadas de aprendizaje automático y análisis predictivo, así como la inclusión de nuevas variables que permitan profundizar en el estudio del desempeño y su impacto en la productividad del área de cultivo.

Es pertinente mantener la actualización periódica de los datos provenientes de Ayudas Frecuentes, Evaluación de Áreas, Corte Limpio y Pilotos, asegurando la correcta ejecución del proceso automatizado de carga y validación desarrollado en Python, con el fin de preservar la calidad de la información utilizada para el análisis semanal del desempeño.

Se recomienda fortalecer el uso de los tableros analíticos en Power BI como herramienta de apoyo a la toma de decisiones, promoviendo su consulta regular por parte de supervisores y coordinadores para la identificación temprana de desviaciones operativas, particularmente en labores y bloques que presentan variabilidad en los mapas de calor de cantidad y calidad.

Resulta conveniente documentar y estandarizar el modelo analítico implementado, incluyendo las reglas de ponderación por fase, criterios de compensación y clasificación por letras, con el propósito de asegurar la sostenibilidad del sistema, facilitar su transferencia de conocimiento y permitir su replicabilidad en otras áreas o sedes de la organización.

Referencias Bibliográficas

- Chiavenato, I. (2017). *Gestión del talento humano* (4.^a ed.). McGraw-Hill Education.
https://books.google.com/books/about/Gesti%C3%B3n_del_talento_humano.html?hl=es&id=e7G7QwAACAAJ
- Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business School Press. <https://archive.org/details/competingonanaly00thom>
- Dessler, G. (2015). *Human resource management* (14th ed.). Pearson Education.
<https://www.amazon.com/Human-Resource-Management-14th-Dessler/dp/0133545172>
- Florval S.A.S. (2025). *Procedimiento de evaluación de desempeño, versión 2*. Documento interno.
- ISO. (2015). *Norma ISO 9001:2015: Sistemas de gestión de la calidad — Requisitos*. International Organization for Standardization. <https://www.iso.org/standard/62085.html>
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). A review of the use of big data in agriculture. *Computers and Electronics in Agriculture*, 143, 23–37.
<https://doi.org/10.1016/j.compag.2017.09.037>
- Kwon, O., & Sim, J. M. (2013). A data-driven approach to performance evaluation. *Decision Support Systems*, 54(3), 1199–1211. <https://doi.org/10.1016/j.dss.2012.11.002>
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 52(2), 21–32. <https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/>

Marr, B. (2018). *Data-driven HR: How analytics are transforming the way we manage talent*.

Kogan Page. <https://www.koganpage.com/hr-learning-development/data-driven-hr-9781398614567>

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.

<https://www.oreilly.com/library/view/data-science-for/9781449374273/>

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming: A review.

Agricultural Systems, 153, 69–80. <https://doi.org/10.1016/j.agsy.2017.01.023>