

# ▶ TALENTO TECH



TIC



Universidad  
Tecnológica  
de Bolívar

[www.utb.edu.co/talento-tech](http://www.utb.edu.co/talento-tech)

[www.utb.edu.co/talento-tech](http://www.utb.edu.co/talento-tech)

# Proyecto: Análisis Descriptivo del Sector Agrícola en Bolívar y Córdoba (2019-2023): Evaluación de Desempeño y Tendencias

---

BOOTCAMP ANÁLISIS DE DATOS NIVEL BÁSICO

GRUPO: ANALISIS DE DATOS - BAS-1032-202406

**EJECUTOR TÉCNICO:**

JORGE LUIS VILLALBA ACEVEDO

UNIVERSIDAD TECNOLÓGICA DE BOLIVAR - UTB

## MISIÓN 1

Integrantes:

Carlos José Romero Escorcía  
Emmy Luz Oyola Díaz  
Juan Carlos Gutiérrez Ortega  
Oscar Ricardo Sierra Alean

## **INFORME MISIÓN 1**

**Aplicación de la analítica de datos para la  
optimización de procesos en los departamentos  
de Bolívar, Córdoba, Sucre y San Andrés**

**Sectores:**

**Agricultura, educación, turismo, marketing,  
salud y digital**

## Contenido

Introducción .....	5
Contextualización del problema .....	6
Sectores para abordar .....	7
Descripción de los sectores .....	7
Importancia de los sectores .....	7
Identificación de las fuentes de datos .....	8
Problema por resolver – Preguntas de negocio .....	9
Preguntas de negocio .....	9
Desafíos y como abordarlos en la identificación y recopilación de fuentes de datos .....	10
Aporte de la analítica de datos al problema a desarrollar .....	11
Metodología Propuesta .....	12
Tipo de analítica .....	12
Metodología CRISP-DM .....	12
Herramientas por utilizar .....	13
Descripción de las herramientas .....	13
Justificación para su uso .....	13
Conclusiones Misión 1 .....	14
Referencias bibliográficas .....	15

## Introducción

La agroindustria en Colombia ha sido un pilar fundamental tanto para la economía como para el bienestar social, consolidándose como uno de los mayores contribuyentes al Producto Interno Bruto (PIB) en los últimos años. De acuerdo con el Ministerio de Agricultura y Desarrollo Rural, el sector agropecuario impulsó de manera significativa el crecimiento económico del país, con un notable incremento del 6% en el PIB durante el último trimestre de 2023. Este crecimiento refleja la importancia del sector en la generación de empleo, el abastecimiento de alimentos, y el desarrollo de las zonas rurales, posicionándolo como un eje estratégico para el progreso nacional.

No obstante, a pesar de su relevancia, el sector enfrenta desafíos complejos, siendo uno de los principales la recopilación de datos agrícolas precisos y actualizados. La necesidad de obtener información de campo de manera continua, junto con la gran diversidad de variables que afectan el rendimiento y la producción de los cultivos, ha dificultado la creación de estadísticas agrícolas confiables. Factores como el cambio climático, las prácticas agrícolas tradicionales, la escasa adopción de tecnología, y las fluctuaciones en los precios de los productos agrícolas han influido considerablemente en la variabilidad de los resultados del sector durante la última década.

Para enfrentar esta problemática, se implementará un proyecto de análisis de datos utilizando la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), una herramienta fundamental para la ejecución exitosa de proyectos de minería de datos. Esta metodología proporciona un enfoque estructurado que abarca todas las fases necesarias, desde la comprensión inicial del problema hasta la implementación de soluciones basadas en los resultados obtenidos.

## Contextualización del problema

En Colombia, el sector agroindustrial enfrenta desafíos significativos que afectan su productividad y sostenibilidad. Los departamentos de Bolívar y Córdoba, claves en la producción agrícola, han presentado una considerable variabilidad en el área cosechada, producción y rendimiento de cultivos transitorios entre 2019 y 2023.

Entre los principales problemas identificados están la baja adopción de tecnologías agrícolas, los impactos del cambio climático, la limitada infraestructura y acceso a créditos, y la volatilidad en los precios de los productos agrícolas. Esto genera una situación de vulnerabilidad en el sector y pone en riesgo la seguridad alimentaria y la estabilidad económica de las regiones.

## Sectores para abordar

Este proyecto se centrará en el sector agrícola, analizando los cultivos de Bolívar y Córdoba entre los años 2019 – 2023.

### Descripción de los sectores

En 2023, el sector de agricultura, ganadería, caza, silvicultura y pesca contribuyó con el 9,7% del valor agregado bruto de la economía, ocupando el cuarto lugar entre las 12 ramas principales de actividad económica (Unidad de Planificación Rural Agropecuaria - UPRA, 2024). La agricultura, con su amplia diversidad de cultivos como café, banano, flores y palma de aceite, se ha consolidado como uno de los motores clave para la reactivación económica post-Covid-19. Sin embargo, el sector enfrenta desafíos importantes, tales como la falta de infraestructura, el impacto del conflicto armado, y los efectos del cambio climático, que amenazan tanto la seguridad alimentaria como los ingresos rurales.

### Importancia de los sectores

El análisis de los factores que impactan la productividad agrícola es fundamental para garantizar la sostenibilidad del sector. El aumento de la eficiencia en la producción puede contribuir al desarrollo económico de la región y mejorar la calidad de vida de los productores locales.

## Identificación de las fuentes de datos

**Encuesta Anual Agropecuaria (EAA):** Proporciona datos detallados sobre el rendimiento, área sembrada y cosechada de los cultivos. Esta información es recopilada y gestionada por la Unidad de Planificación Rural Agropecuaria - UPRA, una entidad técnica adscrita al Ministerio de Agricultura y Desarrollo Rural, encargada de planificar y desarrollar modelos de ordenamiento territorial agropecuario, promoviendo una gestión más eficiente y sostenible de los recursos agrícolas en el país

### Evaluaciones Agropecuarias Municipales – EVA. 2019 - 2023. Base Agrícola | Datos Abiertos Colombia

Se espera agregar datos de variables meteorológicas del IDEAM para complementar la investigación con el fin de identificar su relación con la productividad de los cultivos agrícolas (temperatura del aire, humedad del aire, precipitación).



## Problema por resolver – Preguntas de negocio

Identificación y análisis de los factores que afectan el rendimiento agrícola en los departamentos de Bolívar y Córdoba durante el periodo 2019-2023. A lo largo de estos años, el rendimiento de diversos cultivos ha experimentado variaciones significativas que podrían estar relacionadas con múltiples factores.

La variación en el rendimiento agrícola es un factor determinante para el sector, ya que afecta directamente la productividad, los ingresos rurales e incluso la seguridad alimentaria en la región. Este análisis busca identificar patrones y tendencias dentro de los datos disponibles para determinar los factores más influyentes en la disminución o el aumento del rendimiento de cultivos. Con el fin de facilitar la toma de decisiones en cuanto a políticas agrícolas, inversiones en infraestructura y estrategias para mejorar la productividad en los próximos años.

### Preguntas de negocio

- ¿Qué factores afectan el rendimiento agrícola en estos departamentos?
- ¿Existen diferencias significativas en la productividad según el tipo de cultivo y la región?
- ¿Cómo ha sido el rendimiento de los cultivos en Bolívar y Córdoba?
- ¿Cómo ha variado la producción en toneladas de los cultivos clave?
- ¿Cómo afectan las variables meteorológicas a la productividad de los cultivos?

## Desafíos y como abordarlos en la identificación y recopilación de fuentes de datos

De los mayores desafíos a abordar en la identificación y recopilación de fuentes de datos radica en la disponibilidad, calidad y diversidad de las fuentes de datos. Con respecto al sector agrícola en los departamentos de Bolívar y Córdoba, no hay excepción. Desglosando un poco más nos encontramos con los siguientes.

- **Datos desactualizados o incompletos.** Los datos relacionados con la producción agrícola a veces son inconsistentes. La información pública como la proporcionada en la Encuesta Anual Agropecuaria (EAA) o las EVA (Evaluaciones Agropecuarias Municipales) no incluye todas las variables relevantes para el análisis.
- **Falta de datos detallados sobre factores climáticos.** La información de las condiciones climáticas, tales como la temperatura, humedad o precipitación, suele estar fragmentada o no estar alineada temporalmente con los datos agrícolas, lo que dificulta relacionar estas variables con la productividad agrícola.
- **Dificultad de acceso a datos geoespaciales.** Información requerida que cuente con más precisión para el análisis puede estar disponible en fuentes con permisos especiales o que son costosas de obtener. De este mismo modo podemos encontrarnos con fuentes que en su momento los sitios de archivo se encuentren en mantenimiento lo que complicaría su accesibilidad inmediata.
- **Fragmentación de las fuentes de datos.** Los datos de actividad agrícola y de factores climáticos pueden estar distribuidos en archivos de diversas entidades o instituciones, recolectar esta información de manera centralizada puede llegar a ser complicado.
- **Volumen y complejidad de los datos.** Las fuentes de datos pueden involucrar grandes volúmenes de información, lo que posiblemente genere la necesidad de herramientas de procesamiento robustas para manejar, analizar y visualizar estos datos de manera efectiva.

## Aporte de la analítica de datos al problema a desarrollar

El uso de la analítica de datos ayudará a identificar tendencias ocultas que permitan tomar decisiones más informadas. Esto contribuirá a mejorar la eficiencia en la gestión agrícola, orientando la toma de decisiones hacia prácticas más productivas.

# Metodología Propuesta

## Tipo de analítica

En el proyecto, se empleará un enfoque basado en la analítica descriptiva, con el objetivo de examinar y resumir las características y patrones de los datos agrícolas. Se centrará en el estudio de los datos históricos del rendimiento agrícola, utilizando herramientas estadísticas para ofrecer una visión detallada y comprensible del comportamiento de los cultivos durante el período de 2019 a 2023.

## Metodología CRISP-DM

Para el desarrollo del proyecto, se empleará la metodología CRISP-DM, que ofrece un marco estructurado para la ejecución de proyectos de minería de datos y análisis de datos. Esta metodología consta de seis fases clave (Schröer et al., 2021):

- Entendimiento del Negocio
- Entendimiento de los Datos
- Preparación de los Datos
- Modelado
- Evaluación
- Despliegue

## Herramientas por utilizar

### Descripción de las herramientas

Para el desarrollo del proyecto se utilizarán dos herramientas clave: Python y Visual Studio Code (VSC). Python es un lenguaje de programación, ideal para el análisis de datos debido a sus bibliotecas especializadas como Pandas, NumPy, y Matplotlib, que facilitan la manipulación, visualización y modelado de grandes volúmenes de datos. Por otro lado, VSC es un entorno de desarrollo integrado (IDE) que se integra con Python facilitando la codificación y ejecución de scripts.

### Justificación para su uso

Python se utilizará para el procesamiento y análisis de los datos, por su capacidad de procesar grandes volúmenes de información y realizar análisis complejos de manera eficiente, utilizando bibliotecas como Pandas y NumPy que facilitan la manipulación y posterior análisis, y a su vez Matplotlib, que nos permite crear visualizaciones claras y detalladas. Por su parte, VSC se utilizará por su entorno ligero, altamente personalizable y su integración nativa con Python y Jupyter Notebook.

## Conclusiones Misión 1

En esta primera fase del proyecto se ha establecido un marco claro para abordar los desafíos que enfrenta la agroindustria en los departamentos de Bolívar y Córdoba. Se ha identificado que factores como el cambio climático, la baja adopción de tecnologías y la limitada infraestructura afectan significativamente el rendimiento de los cultivos entre 2019 y 2023, lo que repercute en la economía y la seguridad alimentaria local.

La metodología de analítica de datos puede considerarse una herramienta clave de la ciencia de datos, ya que pone al descubierto patrones y tendencias que permiten mejorar la toma de decisiones. La metodología CRISP-DM es la que se ha seguido para estructurar este análisis; abarca, desde la comprensión del problema hasta la implementación de soluciones basadas en los hallazgos.

Se han asumido varios retos en el trabajo de recolección de datos como carencia de información actualizada y fragmentación de las fuentes, prioridades que deben ser superadas por una disposición sistemática de los datos climáticos y los datos de producción agrícola.

Las herramientas elegidas por Python y Visual Studio Code van acorde con las necesidades del proyecto, puesto que sirven para un eficiente trabajo de datos y la representación gráfica de los mismos, las cuales traen facilidades para la interpretación de los resultados. En definitiva, la formulación del proyecto que nos atañe sienta bases para una optimización de los procesos agrícolas, por el anhelo de mejorar la producción de las tierras y el bienestar de las regiones de Bolívar y Córdoba.

## Referencias bibliográficas

- Ministerio de Agricultura y Desarrollo Rural. (2024, February 16). *El sector agropecuario mantiene su tendencia de recuperación: empujó el crecimiento del PIB con un aumento del 6% en el último trimestre de 2023*.  
<https://www.minagricultura.gov.co/noticias/Paginas/El-sector-agropecuario-mantiene-su-tendencia-de-recuperaci%C3%B3n-empuj%C3%B3-el-crecimiento-del-PIB-con-un-aumento-del-6-en-el-%C3%BAltim.aspx>
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181, 526–534.  
<https://doi.org/10.1016/j.procs.2021.01.199>
- Unidad de Planificación Rural Agropecuaria - UPRA. (n.d.). *Evaluaciones Agropecuarias Municipales – EVA. 2019 - 2023. Base Agrícola*. Retrieved September 28, 2024, from [https://www.datos.gov.co/Agricultura-y-Desarrollo-Rural/Evaluaciones-Agropecuarias-Municipales-EVA-2019-20/uejq-wxrr/about\\_data](https://www.datos.gov.co/Agricultura-y-Desarrollo-Rural/Evaluaciones-Agropecuarias-Municipales-EVA-2019-20/uejq-wxrr/about_data)
- Unidad de Planificación Rural Agropecuaria - UPRA. (2024). *Producto Interno Bruto - PIB Año 2023 y IV trimestre de 2023*.



Universidad  
Tecnológica  
de Bolívar

[www.utb.edu.co/talento-tech](http://www.utb.edu.co/talento-tech)