Café de Colombia, análisis de la producción y la exportación en los últimos veinte años

Andrés Felipe Jaramillo Lozano

Carlos Andrés Archila Padilla

Carlos Mario Latorre Martínez

Christian David Home Acero

Darío Alejandro Durán Barrera

Felipe Herrera Cuellar

Guillermo Stiven Murillo Pantoja

Jonatan Samuel Castro Castro

Juan Esteban Álvarez David

Juan Sebastián González Quintero

Julieth Fabiana Giraldo Suarez

Mario Perlaza Erazo

Rodrigo Valdés Capote

Walter Cardona

Ficha 2672233

Servicio Nacional de Aprendizaje - SENA

Regional Valle

Centro de Electricidad y Automatización Industrial - CEAI

Técnico en Programación para Analítica de Datos - TPAD 7

Instructor

Ing. Luis Armando Amaya Quiroga

22 de septiembre de 2023

**Tabla de contenido**

[Lista de tablas y figuras 3](#_Toc146290602)

[Resumen 5](#_Toc146290603)

[1. Introducción 6](#_Toc146290604)

[2. Planteamiento del problema 7](#_Toc146290605)

[2.1 Identificación del problema 7](#_Toc146290606)

[3. Justificación 8](#_Toc146290607)

[4. Objetivo general y objetivos específicos 9](#_Toc146290608)

[4.1 Objetivo general 9](#_Toc146290609)

[4.2 Objetivos específicos 9](#_Toc146290610)

[5. Alcance 10](#_Toc146290611)

[6. Beneficiarios 11](#_Toc146290612)

[6.1 Beneficiarios directos 11](#_Toc146290613)

[6.2 Beneficiarios indirectos 11](#_Toc146290614)

[7. Impacto social, económico, ambiental, tecnológico 12](#_Toc146290615)

[7.1 Impacto social 12](#_Toc146290616)

[7.2 Impacto económico 12](#_Toc146290617)

[7.3 Impacto ambiental 12](#_Toc146290618)

[7.4 Impacto tecnológico 12](#_Toc146290619)

[8. Restricciones y alternativas de solución 13](#_Toc146290620)

[8.1 Restricciones 13](#_Toc146290621)

[8.2 Alternativas de Solución 13](#_Toc146290622)

[9. Resultados del desarrollo 14](#_Toc146290623)

[9.1 Metodología empleada 14](#_Toc146290624)

[9.2 Requerimientos de los datos 15](#_Toc146290625)

[9.3 Entendimientos de los datos 16](#_Toc146290626)

[9.4 Preparación de los datos 21](#_Toc146290627)

[9.5 Preparación del modelo 24](#_Toc146290628)

[9.6 Evaluación del modelo 36](#_Toc146290629)

[10. Entregables y su descripción 41](#_Toc146290630)

[11. Conclusiones 42](#_Toc146290631)

[Referencias bibliográficas 46](#_Toc146290632)

# Lista de tablas y figuras

[Tabla 1. 17](#_Toc146290633)

[Ilustración 1. 14](#_Toc146290634)

[Ilustración 2. 18](#_Toc146290635)

[Ilustración 3. 18](#_Toc146290636)

[Ilustración 4. 19](#_Toc146290637)

[Ilustración 5. 19](#_Toc146290638)

[Ilustración 6. 20](#_Toc146290639)

[Ilustración 7. 20](#_Toc146290640)

[Ilustración 8. 21](#_Toc146290641)

[Ilustración 9. 22](#_Toc146290642)

[Ilustración 10. 22](#_Toc146290643)

[Ilustración 11. 23](#_Toc146290644)

[Ilustración 12. 23](#_Toc146290645)

[Ilustración 13. 24](#_Toc146290646)

[Ilustración 14. 25](#_Toc146290647)

[Ilustración 15. 25](#_Toc146290648)

[Ilustración 16. 26](#_Toc146290649)

[Ilustración 17. 26](#_Toc146290650)

[Ilustración 18. 27](#_Toc146290651)

[Ilustración 19. 27](#_Toc146290652)

[Ilustración 20. 28](#_Toc146290653)

[Ilustración 21. 28](#_Toc146290654)

[Ilustración 22. 29](#_Toc146290655)

[Ilustración 23. 29](#_Toc146290656)

[Ilustración 24. 30](#_Toc146290657)

[Ilustración 25. 30](#_Toc146290658)

[Ilustración 26. 31](#_Toc146290659)

[Ilustración 27. 31](#_Toc146290660)

[Ilustración 28. 32](#_Toc146290661)

[Ilustración 29. 32](#_Toc146290662)

[Ilustración 30. 33](#_Toc146290663)

[Ilustración 31. 34](#_Toc146290664)

[Ilustración 32. 35](#_Toc146290665)

[Ilustración 33. 35](#_Toc146290666)

[Ilustración 34. 35](#_Toc146290667)

[Ilustración 35. 36](#_Toc146290668)

[Ilustración 36. 36](#_Toc146290669)

[Ilustración 37. 37](#_Toc146290670)

[Ilustración 38. 37](#_Toc146290671)

[Ilustración 39. 38](#_Toc146290672)

[Ilustración 40. 38](#_Toc146290673)

[Ilustración 41. 38](#_Toc146290674)

[Ilustración 42. 39](#_Toc146290675)

[Ilustración 43. 39](#_Toc146290676)

[Ilustración 44. 40](#_Toc146290677)

[Ilustración 45. 40](#_Toc146290678)

[Ilustración 46. 40](#_Toc146290679)

[Ilustración 47. 42](#_Toc146290680)

[Ilustración 48. 43](#_Toc146290681)

# **Resumen**

La economía colombiana se ha beneficiado de la producción y comercialización del café, un producto agrícola arraigado a su historia y cultura. Además, el café colombiano es reconocido a nivel mundial por su calidad y variedad, lo que ha contribuido a su destacada participación en el mercado internacional. Este proyecto se propuso analizar y comprender el desarrollo de la producción y la exportación del café colombiano durante el periodo comprendido entre 2003 y 2022, enfocándonos en compilar las dinámicas que caracterizan esta industria, y los factores políticos que han influido en su comportamiento y han impactado en su mercado.

Los resultados de este análisis revelaron que las políticas gubernamentales desempeñan un papel importante en la producción y la exportación del café, también se destacó la importancia de considerar otros factores como el contexto social, las condiciones climáticas, etc., pues no hay una variable única que intervenga en el mercado. Se observó que en los últimos cuatro años la producción y la exportación del café ha disminuido sistemáticamente.

Sin embargo, en contraparte, en este mismo lapso ha habido un aumento progresivo en el valor de la exportación del café, posiblemente relacionado con cambios en la demanda global. Pero también se identificó un aumento significativo en el valor de la cosecha. Estos resultados sugieren la necesidad de un enfoque más completo y multidimensional al analizar el mercado cafetero en Colombia.

Este estudio contribuye al entendimiento de la dinámica del café colombiano en los últimos veinte años. Si bien se utilizó la variable de las políticas gubernamentales como un factor clave, se resalta la importancia de considerar otros aspectos para nutrir este aporte. Esta perspectiva analítica ofrece información valiosa para encaminar investigaciones que permitan la toma de decisiones futuras y la formulación de políticas en el sector cafetero colombiano.

# 1. Introducción

Colombia, reconocido a nivel mundial como uno de los principales productores de café, destaca por su café de alta calidad y sabor excepcional. Del mismo modo, la cultura y la economía del país se ven influenciadas de manera significativa por la industria cafetera, la cual desempeña un papel fundamental al generar empleo y contribuir de forma notoria a las exportaciones del país.

La producción de café en Colombia guarda una estrecha relación con diversas dinámicas históricas, contextuales, culturales y sociales, de tal manera que responde a los distintos acontecimientos que atraviesa el país, así como a los fenómenos ambientales, los procesos electorales, las decisiones políticas y las normativas, sin embargo, sin un análisis de datos pensar en estos hechos por sí solos daría lugar a la especulación.

Este informe analiza los datos sobre la producción y la exportación del café en el país, lo que permite organizarlos y estructurarlos de tal manera que se pueda identificar las variables que afectan e impactan en su dinámica en el mercado. Además, proporciona una visión general de estos factores con el propósito de establecer una base de referencia, sacar conocimiento de los datos y preparar el camino para un análisis más profundo en futuras investigaciones.

# 2. Planteamiento del problema

¿Cómo ha sido el comportamiento histórico de la producción y la exportación del café en los últimos cinco periodos presidenciales de Colombia?

## 2.1 Identificación del problema

El café es un recurso natural de gran importancia a nivel mundial. Más del 80% de su producción se destina al comercio internacional, siendo Colombia uno de los países más visibles en este mercado. En el año 2022, el sector cafetero colombiano aportó significativamente a su economía nacional generando $14,5 billones de pesos (COP), gracias a su rol como país exportador. Además, según estimaciones cada día se toman 2000 millones de tazas de café en todo el mundo, remarcando su arraigo en la cultura global. El café no es solo una bebida o un producto, sino un sustento fundamental para el país, ya que genera empleo y aporta ingresos significativos por exportación. Sin embargo, la industria cafetera colombiana se enfrenta a diversos desafíos, como el cambio climático y la evolución sociopolítica local, que influyen en la estabilidad económica y social del país.

La disponibilidad de datos gubernamentales y el avance de la tecnología que permite extraerlos, tratarlos y difundirlos, ha abierto la puerta a la participación ciudadana y a la transparencia en la gestión gubernamental. Sin embargo, la extensa cantidad de datos disponibles en la web plantea desafíos en cuanto a su análisis y entendimiento. Para abordar esta complejidad, se hace necesario el uso de herramientas de análisis de datos. Estas permiten análisis predictivos, visualización de la información, análisis de tendencias y cruces de variables, información que en un futuro será valiosa para la toma de decisiones en la industria cafetera y otros sectores económicos.

El estudio de estos efectos se vuelve fundamental para Colombia ya que puede proporcionar información valiosa para la formulación de políticas futuras y ayudar a la industria cafetera a anticipar y adaptarse a cambios en su dinámica. Por eso es importante para la economía y el agro colombiano llevar a cabo esta investigación, porque la producción y la exportación del café viene en una caída sistemática en los últimos cuatro años y es fundamental conocer las causas para frenar las consecuencias negativas e impulsar la economía colombiana.

# 3. Justificación

El café es uno de los productos agrícolas más emblemáticos y representativos de Colombia. El país es reconocido mundialmente por la alta calidad de su café suave y por ser el tercer productor más grande a nivel global. La exportación del café desempeña un papel crucial en la economía colombiana, generando empleo y contribuyendo significativamente a los ingresos de miles de familias. Por eso se busca promover una mayor comprensión y conciencia sobre el café y su influencia en la economía colombiana.

Explicando la importancia de este producto, este proyecto se embarca en el propósito de analizar y comprender el desarrollo de la comercialización del café colombiano desde el año 2003 hasta el 2022, para destacar los factores clave que han movilizado este sector, su impacto en la economía colombiana y en la sostenibilidad del sector cafetero.

Es de resaltar que Colombia experimentó cambios significativos en su liderazgo presidencial y por ende en su política agrícola y comercial en estos últimos veinte años. Por lo tanto, el propósito de este proyecto es analizar a través de datos cómo ciertas políticas pudieron haber influido en las tendencias de crecimiento o declive de variables específicas, como la producción y la exportación. Esto permitirá documentar el impacto de dichas políticas en la dinámica del mercado de este producto.

En conclusión, debido a la importancia del café colombiano y su arraigo en el país, esta investigación no solo se lleva a cabo para documentar y exponer lo encontrado, sino también para mostrar la perspectiva analítica que es posible tener con los datos que construye el país y dejar un referente inicial de información. Sin embargo, es una realidad que también el agro y la industria cafetera se enfrentan a desafíos importantes relacionados con factores como la sostenibilidad y el cambio climático. Por lo cual es pertinente indagar y seguir documentando y procesando todos los datos sobre las variables del café como lo es la producción y exportación, para construir mejores propuestas e intervenciones que permitan avanzar a una mejor toma de decisiones por parte de los dirigentes y veeduría de los ciudadanos.

# 4. Objetivo general y objetivos específicos

## 4.1 Objetivo general

Analizar el comportamiento histórico de la producción y la exportación del café, identificando la relevancia de los últimos cinco periodos presidenciales de Colombia.

## 4.2 Objetivos específicos

* Recopilar y organizar los datos relacionados con el café en Colombia.

* Garantizar que los datos obtenidos sean coherentes y estén listos para el proceso de limpieza de datos y análisis.

* Utilizar las herramientas de visualización para presentar los datos.

* Emplear las herramientas de análisis de datos para interpretar el impacto de las políticas gubernamentales en el sector cafetero.

* Elaborar conclusiones basadas en la información clave alcanzada. Esto incluirá los resultados obtenidos por medio del proceso de Machine Learning.

# 5. Alcance

La investigación se enfoca en analizar los últimos veinte años de datos sobre el café en Colombia, teniendo en cuenta la metodología de investigación CRISP-DM, que se consideró la más apropiada para llevar a cabo un proceso de indagación en los datos. Se pretende explorar cómo algunas de las políticas gubernamentales implementadas durante ese periodo, podrían haber influido en las variaciones observadas en dicho lapso.

Para garantizar la confiabilidad de nuestros hallazgos, se planea utilizar datos provenientes de fuentes confiables, como bases de datos gubernamentales oficiales y fuentes especializadas en el sector. Haciendo énfasis en los datos de precios, producción y exportaciones.

Uno de los objetivos es analizar cómo el comportamiento de los datos en el pasado ha sido afectado por las propuestas de los gobiernos de turno en Colombia. Para lograr este objetivo se han seleccionado las herramientas para realizar análisis estadísticos, tales como, OpenRefine, Power BI, Python y R. Estas herramientas nos permiten llevar a cabo la limpieza de los datos, visualización, las correlaciones, comportamientos y comparaciones de las distintas variables. Se busca comprender cómo la dinámica de los datos podría estar relacionada a las decisiones de los gobiernos que hubo en cada periodo de tiempo estudiado.

El resultado final será presentar un informe que resuma los hallazgos que han sido claves para la investigación, teniendo en cuenta la producción total de los sacos de café del país y sus exportaciones. Las conclusiones obtenidas junto con las predicciones realizadas aportarán una valiosa visión sobre el comportamiento de esta industria.

# 6. Beneficiarios

Colombia tiene aroma a café. A lo largo de la cordillera de los andes se encuentran las condiciones geográficas ideales para que este fruto se convierta en el estilo de vida de 12.387 familias cafeteras, familias que constituyen la base de la cultura agrícola de nuestro país y que por más de 200 años se han dedicado a esta hermosa tarea.

## 6.1 Beneficiarios directos

En primera instancia los beneficiarios directos somos todo el equipo de trabajo y desarrolladores del presente proyecto quien en el proceso de investigación hemos adquirido conocimiento acerca del café, su origen, cultura, comercialización y políticas económicas que rigen los precios.

Los siguientes beneficiarios son el SENA regional Valle, el equipo de instructores y sus directivos que usarán el producto del proyecto como referencia para consultar cómo las políticas gubernamentales influyen en la producción del café.

## 6.2 Beneficiarios indirectos

Hay muchos beneficiarios indirectos como: Las empresas productoras, porque al comprender las tendencias de exportación e importación, los productores pueden tomar decisiones informadas sobre la producción del café, lo que puede influir en sus ingresos y rentabilidad. Las empresas que se dedican al comercio de café colombiano pueden beneficiarse al identificar oportunidades de mercado y tomar decisiones estratégicas sobre la compra y venta de café. El gobierno colombiano puede utilizar los resultados del análisis para formular políticas comerciales y agrícolas que fomenten el crecimiento sostenible de la industria del café y aumenten los ingresos fiscales a través de las exportaciones. Las personas y organizaciones que invierten en la industria del café colombiano pueden tomar decisiones de inversión más informadas basadas en las tendencias históricas y actuales de exportación e importación.

# 7. Impacto social, económico, ambiental, tecnológico

## 7.1 Impacto social

Por medio de este proyecto se puede generar conciencia sobre la importancia de la industria cafetera en Colombia y cómo las decisiones gubernamentales pueden afectarla. Esto puede fomentar el interés y la educación por la política agrícola y económica del país entre la población e incluso hacer veeduría de las decisiones que toma el Estado.

## 7.2 Impacto económico

El impacto económico de este proyecto está relacionado con su objeto de estudio y tiene un valor a nivel de referencia, ya que aporta al entendimiento de la dinámica de movilidad de ciertos indicadores económicos relacionados a la producción nacional del producto bajo análisis y puede servir como punto de partida para otros estudios investigativos de este tipo.

## 7.3 Impacto ambiental

Este proyecto no tendrá un impacto ambiental, su valor sobre este ítem, aunque hable de un producto como el café que se cultiva, será en este aspecto solo de referencia, ya que en un inicio es solo analítico.

## 7.4 Impacto tecnológico

El impacto tecnológico de este proyecto se refleja en la aplicación de metodologías avanzadas como CRISP-DM y el uso de herramientas tecnológicas como OpenRefine, Power BI, Python y R. Estas herramientas permiten un análisis más eficiente y una comprensión profunda de los datos, lo que no solo beneficia a este proyecto, sino que también puede fomentar la adopción de tecnologías similares en otros campos, impulsando así la toma de decisiones basada en la recolección de datos.

# 8. Restricciones y alternativas de solución

## 8.1 Restricciones

Disponibilidad de los datos: No poder acceder a todos los datos que requerimos, de la Federación Nacional de Cafeteros, para el proceso de recopilación.

Calidad de los datos: Existe el riesgo de que los datos estén incompletos, repetidos o mal digitados, lo que podría afectar la precisión de los resultados del análisis.

Plazos ajustados: El proyecto tiene unos plazos ajustados, debido a que se trata de un entregable académico que hay que desarrollar en un trimestre.

Disponibilidad de los ambientes de formación: Debido a los conflictos y paros que se presentan en el SENA, se puede dar el caso de no poder ingresar a los ambientes de formación.

## 8.2 Alternativas de Solución

Ampliación de los datos: Aunque haya restricción de acceso a los datos, se pueden explorar otras fuentes de datos públicos, de fuentes oficiales, para complementar la información.

Limpieza de los datos:  Implementar un proceso de limpieza de datos para eliminar datos incorrectos, incoherentes, duplicados o faltantes. Utilizar herramientas de software diseñadas para este propósito, como OpenRefine.

Gestión de plazos: Para mitigar el riesgo de plazos ajustados, se puede establecer un plan de proyecto claro con objetivos concisos y utilizar herramientas de gestión de proyectos para asegurarse de que el proyecto avance según lo planeado.

Trabajar virtualmente: Intentar con los compañeros que tienen computadores con buen rendimiento, avanzar en lo posible con las tareas designadas del proyecto.

# 9. Resultados del desarrollo

## 9.1 Metodología empleada

CRISP-DM (Cross-Industry Standard Process for Data Mining) es una metodología estándar utilizada para la minería de datos, un proceso que implica descubrir patrones y conocimientos útiles a partir de grandes conjuntos de datos. Fue desarrollada en 1996 por un consorcio de empresas, organizaciones y expertos en minería de datos y se ha convertido en la metodología más ampliamente utilizada para proyectos de minería de datos.

Durante más de veinte años, la metodología CRISP-DM ha sido fuente de inspiración de otros estándares como SEMMA de SAS o ASUM-DM de IBM, así como ha dado lugar a múltiples variantes que amplían o particularizan CRISP-DM a una industria o tipo de proyecto.

Ilustración 1.

***CRISP-DM. Fases***

Diagrama

Descripción generada automáticamente

* Entendimiento del negocio: Entendimiento de los objetivos y requerimientos del proyecto.
* Entendimiento de los datos: Obtención conjunto inicial de datos, exploración del conjunto de datos, identificar las características de calidad de los datos e identificar los resultados iniciales obvios. (“Metodologías aplicadas al proceso de Minería de Datos - unal.edu.co”)
* Preparación de los datos: Selección de los datos, y limpieza de los datos.
* Modelado: Implementación en herramientas de analítica de datos.
* Evaluación: Determinar si los resultados coinciden con los objetivos del negocio, e identificar los temas de negocio que deberían haberse abordado. (“Metodologías aplicadas al proceso de Minería de Datos - unal.edu.co”)
* Despliegue: Instalar los modelos resultantes en la práctica, y configuración de datos de forma repetida o continua. (“Metodologías aplicadas al proceso de Minería de Datos - unal.edu.co”)

## 9.2 Requerimientos de los datos

El requerimiento de datos se refiere al enfoque sistemático de reunir y medir información de diversas fuentes a fin de obtener un panorama completo y preciso de una zona de interés.

La técnica de recolección de datos utilizada para este proyecto son las fuentes abiertas: (Consiste en la información pública y gratuita que se encuentra en páginas gubernamentales, universidades, instituciones independientes, organizaciones sin fines de lucro, grandes compañías, plataformas de análisis de datos, agencias, revistas especializadas, entre otras). (“Recolección de datos: métodos, técnicas e instrumentos”)

La fuente utilizada de esta información son las estadísticas cafeteras de la Federación Nacional de Cafeteros, que contiene como base de datos los siguientes archivos de Excel: ***Precios-área-y-producción-de-café-1***, en el cual nos enfocamos en la hoja 9. Producción mensual, que contiene registros desde enero de 1956 hasta mayo del 2023 y el archivo de ***Exportaciones****,* en el cual nos enfocamos en la hoja 1. Tipo volumen y hoja 2. Total valor, que contiene registros desde enero de 1958 hasta abril del 2023.

* [*https://federaciondecafeteros.org/wp/estadisticas-cafeteras/*](https://federaciondecafeteros.org/wp/estadisticas-cafeteras/)
* *Precios, área y producción del café: [Hoja 9. Producción mensual]*
* *Exportaciones de café colombiano: [Hoja 1. Tipo volumen y Hoja 2. Total valor]*

El método de recolección de datos que utilizamos para este proyecto de investigación es el método sintético: el cual es un proceso de análisis que busca resumir los aspectos más importantes de una información suministrada, que va de los general a lo particular.

## 9.3 Entendimientos de los datos

Los datos seleccionados se ordenan desde el periodo más antiguo hasta el periodo más reciente, comprendiendo los últimos cinco periodos presidenciales que han finalizado. Las variables que se van a tener en cuenta serán las siguientes:

Tabla 1.

Variables

|  |  |  |  |
| --- | --- | --- | --- |
| ITEM | VARIABLE | TIPO DE VARIABLE | DESCRIPCIÓN |
| 1 | AÑO | Cuantitativa continua | Información cronológica. |
| 2 | MES | Cuantitativa continua | Información cronológica. |
| 3 | DIA | Cuantitativa continua | Información cronológica. |
| 4 | PRESIDENTE | Cualitativa | Nombre del presidente y su periodo. |
| 5 | SACOS DE CAFÉ PRODUCIDOS | Cuantitativa discreta | Representa el valor producido de café por sacos de 60 kg. |
| 6 | SACOS DE CAFÉ EXPORTADOS | Cuantitativa discreta | Representa el valor exportado de café por sacos de 60 kg. |
| 7 | VALOR EXPORTACIÓN | Cuantitativa discreta | Dato tipo moneda en dólares. |
| 8 | KILOS DE CAFÉ PRODUCIDOS | Cuantitativa discreta | Representa el valor total producido de kilos de café. |
| 9 | FACTOR DE RENDIMIENTO | Cuantitativa discreta | Porcentaje promedio de café pergamino que se necesitó para obtener café producido. |
| 10 | KILOS DE CAFÉ PERGAMINO | Cuantitativa discreta | Representa el valor de kilos de café pergamino. |
| 11 | CARGA DE CAFE | Cuantitativa discreta | Representa el valor de café pergamino por sacos de 125 kg. |
| 12 | PRECIO INTERNO DIARIO | Cuantitativa discreta | Dato tipo moneda en pesos. |
| 13 | VALOR COSECHA | Cuantitativa discreta | Dato tipo moneda en pesos. |

Ilustración 2.

***Precio interno diario***

**Tabla, Calendario

Descripción generada automáticamente**

Ilustración 3.

***Presidentes***

****

Ilustración 4.

***Producción sacos de café***

**Tabla

Descripción generada automáticamente**

Ilustración 5.

***Valor de la cosecha***

**Tabla

Descripción generada automáticamente**

Ilustración 6.

***Exportación sacos de café***

**Tabla

Descripción generada automáticamente**

Ilustración 7.

***Valor de las exportaciones***

**Tabla

Descripción generada automáticamente**

## 9.4 Preparación de los datos

Es poco habitual que los datos crudos o en bruto sean utilizados directamente en un modelo analítico. Antes es necesario refinarlos para encontrar y resolver problemas como datos faltantes, valores no permitidos, duplicados, etc., para obtener datos de calidad con los cuales alimentar un modelo.

Ilustración 8.

***Resultado. UnCleanedCafe.csv***

**Tabla, Excel

Descripción generada automáticamente**

Ilustración 9.

***Proyecto OpenRefine***

**Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente**

Ilustración 10.

***Crear proyecto OpenRefine***

**Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente**

Ilustración 11.

***Editar celdas OpenRefine***

**Tabla

Descripción generada automáticamente**

Ilustración 12.

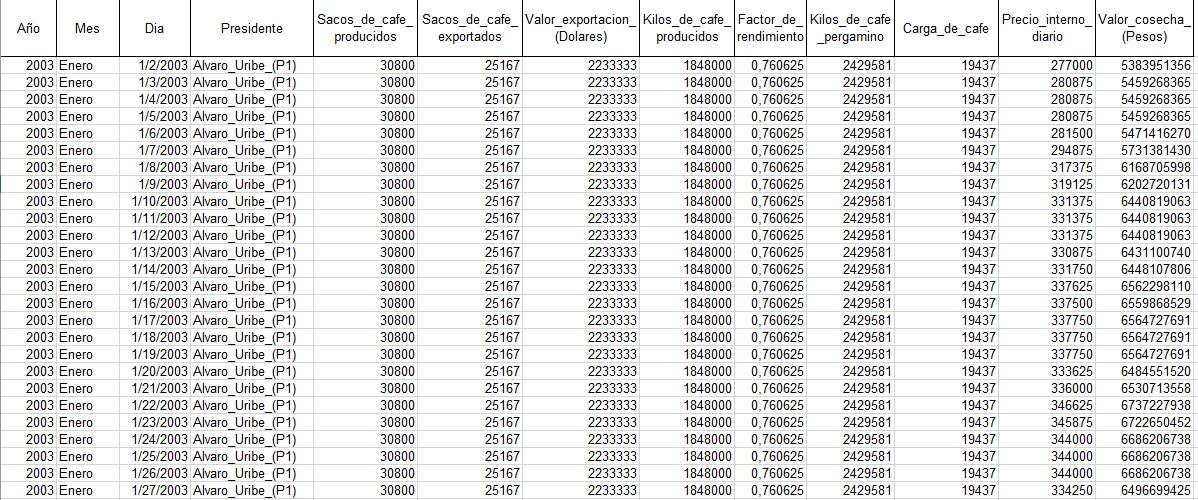
***Transformación OpenRefine***

**Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente**

Ilustración 13.

***Transformación. Data. VariablesV5.csv***

****

## 9.5 Preparación del modelo

Con los datos listos para ser utilizados, se prepara un modelo descriptivo o predictivo que resuelva el problema planteado en forma de pregunta. Dependiendo del modelo, este proceso puede tomar varias iteraciones hasta alcanzar el resultado esperado.

Ilustración 14.

***Dataframe. VariablesV5.csv***

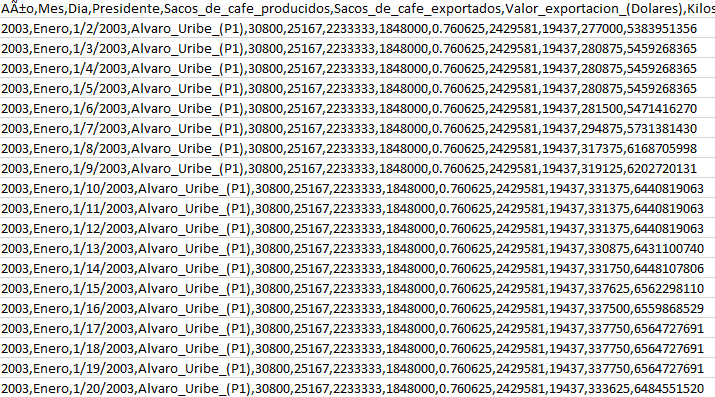


Ilustración 15.

***Inicio Power BI***

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Ilustración 16.

***Obtener datos Power BI***

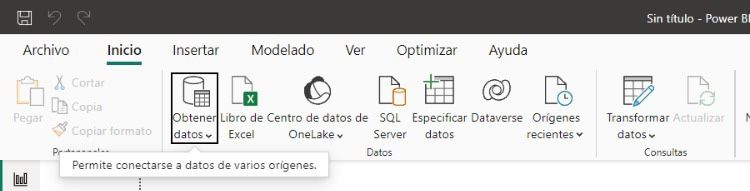


Ilustración 17.

***Abrir csv Power BI***

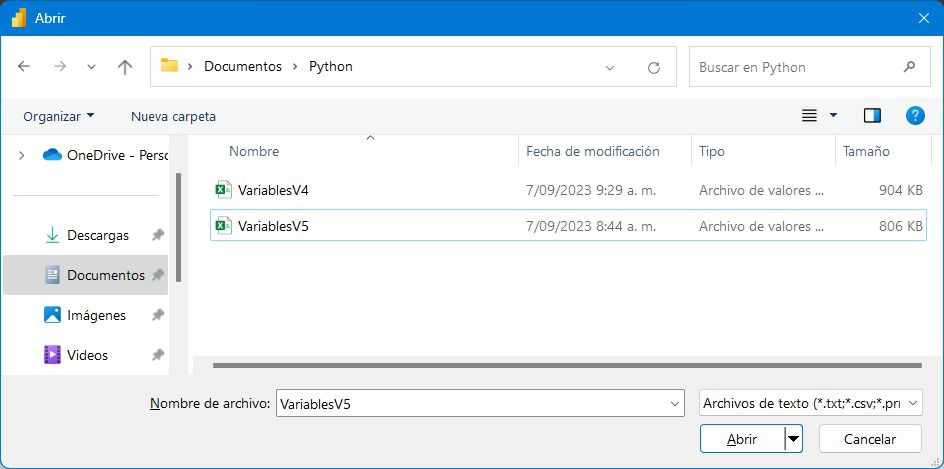


Ilustración 18.

***Cargar datos Power BI***

Tabla

Descripción generada automáticamente

Ilustración 19.

***Carga de café por presidente Power BI***

Gráfico, Gráfico de barras

Descripción generada automáticamente

Ilustración 20.

***Distribución de sacos de café por año Power BI***

Gráfico

Descripción generada automáticamente

Ilustración 21.

***Sacos de café por presidente Power BI***

Gráfico, Gráfico de barras

Descripción generada automáticamente

Ilustración 22.

***Modelado R***

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Ilustración 23.

***Librerías R***

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Ilustración 24.

Primer Código R

Escala de tiempo

Descripción generada automáticamente con confianza media

Ilustración 25.

***Producción de café por presidente R***

Gráfico, Gráfico de barras

Descripción generada automáticamente

Ilustración 26.

***Producción de café por año R***

Gráfico

Descripción generada automáticamente

Ilustración 27.

***Segundo Código R***

Texto

Descripción generada automáticamente

Ilustración 28.

***Venta y valor cosecha en dólares R***

Gráfico

Descripción generada automáticamente

Ilustración 29.

***Modelado Python***

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Ilustración 30.

***Variables Python***

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

Ilustración 31.

***Transformación variables Python***

Tabla

Descripción generada automáticamente

Ilustración 32.

***Transformación de variables Machine Learning Python***

Imagen que contiene Texto

Descripción generada automáticamente

Ilustración 33.

***Transformación variable factor de rendimiento Python***

Texto

Descripción generada automáticamente con confianza media

Ilustración 34.

***Librería seaborn Python***

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Ilustración 35.

***Visualización de datos estadísticos Python***

Imagen que contiene Diagrama

Descripción generada automáticamente

## 9.6 Evaluación del modelo

Ilustración 36.

***Proceso Machine Learning***

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

Ilustración 37.

***Datos de entrenamiento Machine Learning***

Tabla

Descripción generada automáticamente

Ilustración 38.

***Datos de testeo Machine Learning***

Tabla

Descripción generada automáticamente

Ilustración 39.

***Librería regresión lineal Machine Learning***

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Ilustración 40.

***Predicción regresión lineal Machine Learning***

Interfaz de usuario gráfica, Texto, Aplicación, Word

Descripción generada automáticamente

Ilustración 41.

***Error porcentual Machine Learning***

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Ilustración 42.

***Variables de predicción Machine Learning***

Interfaz de usuario gráfica, Aplicación, Word

Descripción generada automáticamente

Ilustración 43.

***Regresión logística Machine Learning***

Tabla

Descripción generada automáticamente

Ilustración 44.

***Regresión logística Iván Duque Machine Learning***

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Ilustración 45.

***Código regresión logística Machine Learning***

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Ilustración 46.

***Modelo predictivo***

Texto

Descripción generada automáticamente

# 10. Entregables y su descripción

Repositorio GitHub (publico)

* Presentación Power Point: Presentación del proyecto
* Dashboard Power BI: Dashboard del proyecto
* Captura Dashboard JPG: Captura del dashboard
* Archivo OpenRefine: Limpieza de los datos del proyecto
* Archivo CSV: Datos del proyecto
* Graficas Power BI: Gráficas básicas del proyecto
* Archivo R: Modelado de los datos del proyecto en R
* Archivo Python Google Collab: Machine Learning de los datos del proyecto
* Informe Word: Informe del proyecto
* Documento PDF: Documento del proyecto

# 11. Conclusiones

En la interpretación de los sacos producidos por periodo presidencial identificamos que las medias del segundo periodo de Álvaro Uribe coinciden al comportamiento del primer periodo de Juan Manuel Santos, el primer cuartil presenta una distancia representativa y culminan casi a la par, los gráficos que presentan menor desviación son las de Juan Manuel Santos.

Ilustración 47.

***Grafica sacos producidos***

Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente

Vemos en la interpretación de las exportaciones que los sacos de café se evidencia una disminución relativa por periodo presidencial desde el primer periodo de Álvaro Uribe hasta el primer periodo de Juan Manuel Santos, en el segundo periodo de Juan Manuel Santos se contempla como el cuartil 1 del primer periodo es superado por el primer cuartil del segundo periodo denotando unas diferencias significativas en exportación y en el periodo presidencial de Iván Duque fue notorio mantener la exportación con una media superior a la de Juan Manuel Santos pero que tuvo unas desviaciones más grandes.

Ilustración 48.

***Grafica sacos exportados***

Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente

En conclusión analizamos que en el segundo periodo presidencial de Juan Manuel Santos hubo la mayor producción y exportación de café, en contraste con su primer periodo presidencial donde fue el menor en producción y exportación de los últimos cinco periodos presidenciales.

Según indica la historia al iniciar el primer periodo de Juan Manuel Santos el panorama cafetero era muy poco alentador. Por una parte, los precios de los fertilizantes habían llegado a niveles récords, como resultado del incremento en los precios del petróleo. Esto implicó un incremento en los costos de producción.

Lo anterior, sumado a los problemas de inseguridad en muchas zonas productoras, trajo consigo el descuido e incluso el abandono de muchos cafetales y su envejecimiento. De manera adicional, se produjo un incremento considerable en el nivel de humedad debido a las altas precipitaciones que causó que los cafetales fueran altamente susceptibles a la roya.

Política gubernamental caficultura un modelo de paz: Ante los fenómenos descritos, la administración de Juan Manuel Santos se propuso lidiar con este gran reto, firmando el primer acuerdo por la prosperidad cafetera en el cual se sentaron las bases de lo que sería la acción del gobierno en este tema.

Mayor productividad: Se observó una recuperación de la producción, soportada en incrementos de productividad, como resultado de la transformación productiva y la sanidad vegetal del cultivo. Esto corresponde a unos cultivos más jóvenes, sembrados con variedades resistentes, mayor tecnificación, mejores condiciones sanitarias y sobre todo cafetales mejores preparados para enfrentar los desafíos que supone la variabilidad climática.

Con este ejemplo de implementación de política gubernamental podemos manifestar que para la industria cafetera es relevante revisar el comportamiento de la producción junto con la exportación para satisfacer la demanda del mercado, las técnicas e investigación del cultivo son prioridad para mantener un mercado competente y un producto de calidad. Los gobiernos son partícipes de la interacción del café colombiano ante el mercado internacional y cómo las políticas de gobierno afectan la estabilidad económica de los caficultores y la población que depende de la cosecha cafetera.

Es por ello por lo que podemos crear la hipótesis para determinar que, de acuerdo con los comportamientos del mercado, el comportamiento en producción y el comportamiento de exportación, las mejores políticas de gobierno para el agro cafetero fueron las del presidente Juan Manuel Santos.

Teniendo en cuenta la hipótesis planteada y los históricos de comportamiento, determinamos que aunque el trabajo realizado en los dos periodos de Juan Manuel Santos no fueron tan satisfactorios para mejorar el promedio del precio del café, se refleja un crecimiento significativo a partir de su segundo periodo para el valor promedio y que se ve reflejado en el periodo de Iván Duque, que retomando esta proyección no intensificó propuestas relevantes para aumentar la producción y la exportación que venía determinada por Juan Manuel Santos, en cambio estas producciones y exportaciones se vieron disminuidas sistemáticamente año a año, pero que se vio implicado por una crisis mundial que fue el COVID-19.

# Referencias bibliográficas

Diplomado, D. (s/f). *Minería de DatosMinería de Datos*. Edu.co. Recuperado el 22 de septiembre de 2023, de <https://disi.unal.edu.co/~eleonguz/cursos/md/presentaciones/Sesion5_Metodologias.pdf>

Edición, C. (s/f). *La Guía del Café*. Intracen.org. Recuperado el 22 de septiembre de 2023, de <https://intracen.org/sites/default/files/media/file/media_file/2022/06/29/itc_coffee_4th_report_20211029_es_web.pdf>

*El mercado del café en el mundo - Datos estadísticos*. (s/f). Statista. Recuperado el 22 de septiembre de 2023, de <https://es.statista.com/temas/9035/el-cafe-en-el-mundo/>

*Estadísticas Cafeteras*. (2019, noviembre 30). Federación Nacional de Cafeteros. <https://federaciondecafeteros.org/wp/estadisticas-cafeteras/>

*IBM Documentation*. (2021, agosto 17). Ibm.com. <https://www.ibm.com/docs/es/spss-modeler/saas?topic=objectives-e-retail-example-finding-business>

*Informes archivos - Federación Nacional de Cafeteros*. (s/f). Federación Nacional de Cafeteros. Recuperado el 22 de septiembre de 2023, de <https://federaciondecafeteros.org/wp/tipos/informes/>

Posada, S. G. (2019, enero 14). La economía del café: ¿Quién se está quedando el dinero? (2019). *Qué Café!* <https://quecafe.info/la-economia-del-cafe-quien-se-esta-quedando-el-dinero/>

Santos, D. (2022, septiembre 14). *Recolección de datos: métodos, técnicas e instrumentos*. Hubspot.es. <https://blog.hubspot.es/marketing/recoleccion-de-datos>

(S/f). Federaciondecafeteros.org. Recuperado el 22 de septiembre de 2023, de <https://federaciondecafeteros.org/static/files/1La_politica_cafetera_2010-2014.pdf>