**Análisis de los precios de la canasta familiar en Manizales**

**Integrantes:**

Delio Buitron Lopez 1004635917

Hernan Jair Telpiz 1193069571

Wilmer Sebastian Perez 1004629333

Lina Robsana Jamioy Tisoy 1006908436

Ximena Amaya Vasquez 1055750283

Andres Jhovany Sierra Bohorquez 1053856068

**Ejecutores:**

Natalia  
Frank Zapata

Margarita Maria Orozco Garcia

Manizales-Caldas

2024

**Tabla de contenido.**

1. [Introducción](#_id37o6qfi5sz) 2
2. Desarrollo de proyecto2

2.1. [Identificación del Problema](#_ygv5llawq3a9) 3

2.2. [Justificación](#_xz6fu1s9jyg7) 3

2.3. Objetivos 3

2.3.1. Objetivo General 3

2.3.2. Objetivos Específicos 4

2.4. Alcance 4

2.5. Presupuesto 5

2.6. Metodología 5

2.7. Modelo

2.8. Resultados Esperados

1. Anexos

**Análisis de los precios de la canasta familiar en Manizales**

### **Introducción**

Las fluctuaciones en los costos de la canasta familiar en Colombia se deben a varios factores, como el cambio en los costos de producción, especialmente en el sector agrícola, impulsado por los ajustes en los precios de insumos, fertilizantes y energía. Además, las variaciones en el precio de los combustibles encarecen el transporte de mercancías, afectando directamente los precios de los productos esenciales. La devaluación del peso frente al dólar también juega un papel clave, ya que muchos insumos son importados, lo que incrementa los costos de producción.

Para abordar estas fluctuaciones, estamos desarrollando el entrenamiento de un sistema de inteligencia artificial que utilizará datos históricos para predecir los precios de la canasta familiar. Este sistema analizará variables como tendencias del mercado, permitiendo anticipar cambios en los precios y ayudar a los consumidores y a las políticas públicas a tomar decisiones informadas. La integración de algoritmos de machine learning permitirá ajustar las predicciones de manera dinámica, mejorando la precisión con el tiempo.

1. **Desarrollo de Proyecto:**

### **2.1. Identificación del Problema**

En el contexto global actual, la inflación y las interrupciones en las cadenas de suministro han generado un aumento sostenido en los precios de bienes esenciales, afectando directamente a la canasta familiar en Colombia. Factores como el aumento en los costos de insumos, los combustibles y la devaluación del peso frente al dólar han incrementado los precios de productos importados, lo que afecta tanto a los consumidores como a los productores locales.

En Manizales, la situación es especialmente compleja debido a la dependencia de la ciudad en productos externos y a las dificultades logísticas propias de la región montañosa. Los precios de alimentos básicos como arroz, carnes y aceite han aumentado significativamente, afectando principalmente a los hogares de ingresos bajos y medios. Además, los agricultores locales, que ya enfrentan mayores costos de insumos, luchan por competir con productos importados más baratos.

Las políticas comerciales internacionales, promovidas por la OMC, han permitido la entrada de productos extranjeros a precios competitivos, lo que ha generado una mayor presión sobre los productores locales. A su vez, los aumentos en los costos de transporte y las barreras en las cadenas de suministro han exacerbado los aumentos de precios, intensificando la desigualdad económica y afectando la capacidad de las familias para cubrir sus necesidades básicas.

### **2.2. Justificación.**

En Colombia, las fluctuaciones constantes en los precios de los productos de la canasta familiar afectan de manera significativa el poder adquisitivo de las familias, especialmente aquellas de menores ingresos, que destinan una mayor proporción de su presupuesto a la compra de bienes esenciales. La volatilidad en los precios también representa un desafío para los comerciantes, quienes enfrentan dificultades para anticipar los cambios en el mercado, lo que complica tanto la fijación de precios como la planificación de inventarios. La implementación de un sistema de inteligencia artificial (IA) capaz de analizar y predecir estas fluctuaciones permitiría tanto a los consumidores como a los comerciantes tomar decisiones más informadas y oportunas, reduciendo los efectos negativos de la inflación y optimizando la gestión económica a nivel personal y comercial. Un sistema de predicción de precios ayudaría a que las familias planifiquen mejor sus gastos ya que los comerciantes manejen de manera más eficiente sus inventarios y márgenes de beneficio, ajustándose a los cambios en los costos.

Este proyecto de IA no solo beneficiará a los consumidores y comerciantes de Colombia al ofrecerles una herramienta para enfrentar los desafíos económicos actuales, sino que también nos proporciona a nosotros, como estudiantes del curso básico de inteligencia artificial, una oportunidad valiosa para aplicar los conocimientos adquiridos. en un entorno real. A través de este proyecto, entrenaremos un modelo de inteligencia artificial que será capaz de procesar grandes cantidades de datos, identificar patrones y hacer predicciones precisas sobre los precios futuros de productos esenciales. Esto nos permitirá poner en práctica lo aprendido sobre análisis de datos, modelado predictivo, aprendizaje automático y algoritmos de IA.

Además, este proyecto nos ayudará a desarrollar habilidades cruciales para nuestra formación profesional, como el trabajo en equipo, la resolución de problemas complejos y la implementación de soluciones tecnológicas que tengan un impacto tangible en la sociedad. Al enfrentarnos a un desafío real, nos veremos motivados a ser más creativos e innovadores, buscando soluciones que no solo cumplan con los objetivos académicos, sino que también generen un valor positivo para la comunidad. Al final, este ejercicio nos preparará mejor para futuras aplicaciones de IA en diferentes campos, consolidando nuestras capacidades y abriendo nuevas oportunidades para proyectos más ambiciosos en el futuro.

**2.3. Objetivos.**

**2.3.1. Objetivo general.**

Implementar en un plazo de seis meses un sistema de inteligencia artificial basado en el análisis de datos económicos, capaz de predecir y monitorear las fluctuaciones de los precios de la canasta familiar

**2.3.2. Objetivos Específicos.**

* Recopilar datos históricos de precios de la canasta familiar y variables económicas relevantes del DANE y otras fuentes oficiales para el entrenamiento del algoritmo.
* Entrenar un modelo de inteligencia artificial utilizando estos datos para predecir las fluctuaciones de los precios de la canasta familiar.
* Evaluar con datos reales para medir su precisión y ajustar los parámetros necesarios, realizando así pruebas de validación del sistema.

**2.4. Alcance.**

El proyecto se centrará en el desarrollo de un sistema de inteligencia artificial para predecir las fluctuaciones de los precios de los productos de la canasta familiar en la ciudad de Manizales. Utilizando datos provenientes del DANE, el sistema analizará variables económicas relevantes para generar predicciones precisas sobre los precios de productos esenciales como alimentos, artículos de higiene y otros bienes de primera necesidad. El objetivo principal es que tanto consumidores como comerciantes puedan tomar decisiones informadas: los primeros, para mejorar la planificación de sus compras y optimizar el manejo de su presupuesto familiar, y los segundos, para ajustar sus precios y gestionar inventarios de manera más eficiente.

El proyecto impactará directamente a aproximadamente el 70% de la población de Manizales, que corresponde a la proporción de residentes que son mayores de edad y forman parte de hogares que realizan compras regulares de la canasta básica. En términos de comercio, el sistema beneficiará a cerca del 50% de los comerciantes locales que venden productos esenciales, quienes podrán utilizar el sistema para anticipar fluctuaciones en los precios y optimizar la administración de sus inventarios.

El límite territorial del proyecto estará acotado a la ciudad de Manizales, excluyendo a la población menor de edad, ya que no participa directamente en la planificación económica ni en la toma de decisiones sobre compras. Además de consumidores y comerciantes, el sistema también podrá ser utilizado por entidades gubernamentales y económicas, como el DANE y las autoridades locales, para respaldar políticas más públicas efectivas basadas en datos predictivos. Asimismo, los investigadores interesados ​​en el análisis económico y social de la región tendrán acceso a información valiosa para realizar estudios más detallados sobre la evolución de los precios y su impacto en la economía local.

La validación del modelo se llevará a cabo utilizando datos reales, recopilados en tiempo real, para asegurar la precisión y efectividad del sistema. Una vez validado, se realizarán acciones de capacitación dirigidas tanto a los comerciantes como a los usuarios interesados ​​en emplear esta herramienta, y se desarrollará un manual de uso para facilitar la implementación y adopción del sistema en la comunidad de Manizales.

**2.5. Presupuesto.**

| **PRESUPUESTO** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Artículo** | **Cantidad** | **Unidad** | **Descripción** | **Precio Unitario (Precio Hora)** | **Precio total** |
| 3 | 600 | Alquiler | Recursos de Computación (servidores en la nube) | $ 15.000 | $ 9.000.000 |
| 4 | 50 | Documento | Creación de manuales (diseño e impresión) | $ 7.000 | $ 350.000 |
| 5 | 1 | Materiales | Papeleria | $ 100.000 | $ 100.000 |
| 6 | 600 | Talento Humano | Unidad de codificacion | $ 100.000,00 | $ 60.000.000 |
| 7 | 600 | Talento Humano | Investigacion- Planeacion | $ 60.000,00 | $ 36.000.000 |
| 8 | 600 | Talento Humano | Gestion de datos | $ 50.000,00 | $ 30.000.000 |
| 9 | 1 | Contingencia | Fondo de Imprevistos (5% del total) |  | $ 5.272.500 |
|  |  |  | **Subtotal :** | | $ 140.722.500 |
|  |  |  | **Impuesto (IVA 19%) :** | | 26737275 |
|  |  |  | **Total :** | | $ 167.459.775 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

**2.6. Metodología.**

CRISP-DM (Cross-Industry Standard Process for Data Mining) es un modelo de referencia que describe un proceso estructurado y estándar para llevar a cabo proyectos de minería de datos y ciencia de datos. Consta de seis fases interrelacionadas: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. Este enfoque permite a los equipos abordar problemas de análisis de manera sistemática y efectiva, asegurando que se cumplan los objetivos del negocio y se maximicen los resultados.

**Paso 1.Comprensión del Negocio**

En esta fase se busca alinear los objetivos del proyecto con las necesidades de la comunidad en Manizales.

* **Objetivo de negocio:** Predecir las fluctuaciones de los precios de productos esenciales en la canasta familiar, facilitando a consumidores, comerciantes y autoridades locales la planificación informada de sus gastos.
* **Situación actual:** La ciudad enfrenta una alta variabilidad en los precios de productos esenciales, impactada por factores globales (devaluación, inflación) y regionales (dificultades logísticas).
* **Objetivos de minería de datos:** Implementar un sistema de IA que identifique patrones de precios y realice predicciones a corto plazo.
* **Plan de proyecto:** Las actividades se dividen en etapas, desde la búsqueda de datos hasta la implementación y evaluación del modelo, programadas para completarse en seis meses.

### **Paso 2. Comprensión de los Datos**

* **Fuentes de Datos**: Datos de precios de productos de la canasta familiar, obtenidos del SIPSA (Sistema de Información de Precios y Abastecimiento del Sector Agropecuario) del DANE, cubriendo desde 2013 hasta 2023. Los datos adicionales de factores económicos se obtendrán de entidades como el Banco de la República y el DANE.
* **Descripción de los Datos**:
  + **Variables Principales**: fecha, grupo (categoría de producto), producto específico, mercado (Manizales) y precio por kilogramo.
  + **Granularidad**: Datos mensuales por producto y categoría, permitiendo análisis detallados de tendencias.
  + **Historia de los Datos**: Diez años de datos continuos que permiten analizar patrones de precios y estacionalidad.
* **Exploración y Validación Inicial**:
  + Revisión de calidad: verificar consistencia y detectar datos atípicos.
  + Análisis de estacionalidad y tendencias para entender patrones y variaciones en los precios.

### **Paso 3. Preparación de los Datos**

* **Limpieza de Datos**:
  + Eliminar valores nulos o inconsistentes.
  + Homogeneizar formatos, como el formato de fecha y la unidad de precio (por kilogramo).
  + Filtrar la información de manera que solo se incluyan los productos de la canasta familiar que se encuentren en la ciudad de Manizales.
  + Seleccionar solamente aquellos productos que estén disponibles en todos los meses del año o en al menos en el 80% de meses. Para los productos que no estén presentes en todos los meses, se ajustan los datos faltantes utilizando el promedio de los precios del año.
* tengo esta información llegó hasta aquí en qué parte se hizo los diferentes trabajos
* **Formato de Almacenamiento**: Archivos CSV que se integrarán en un conjunto de datos unificado para facilitar el entrenamiento del modelo.

### **Paso 4: Modelado**

* **Selección del Modelo**:
  + Se seleccionará un modelo de regresión multivariable debido a su capacidad para predecir una variable dependiente en función de múltiples variables independientes .
  + La regresión multivariable es apropiada para modelar problemas con relaciones lineales entre varias características predictoras, lo cual se espera en este proyecto.
* **Entrenamiento del Modelo**:
  + Se dividirá el conjunto de datos en dos subconjuntos: entrenamiento y prueba (e.g., 80% entrenamiento y 20% prueba).
  + El modelo se ajustará con los datos de entrenamiento aplicando un ajuste de mínimos cuadrados para minimizar la suma de los errores al cuadrado.
  + Una vez entrenado, se procederá a validar el modelo con los datos de prueba para evaluar su desempeño predictivo.

**Paso 5. Evaluación.**

* + Para medir la calidad del modelo de regresión, se utilizará el coeficiente de determinación , que indica la proporción de la varianza en la variable dependiente explicada por las variables independientes en el modelo.
  + Un cercano a 1 indica un buen ajuste, lo cual implica que el modelo es efectivo en predecir los valores de la variable dependiente.
  + Además, se analizarán las residuales para verificar la independencia y homocedasticidad de los errores, factores importantes para validar la adecuación del modelo de regresión.

### **Paso 6: Implementación del Proyecto.**

### La implementación del proyecto requerirá integrar el modelo de regresión multivariable en un sistema de análisis para prever el comportamiento de precios y tipos de productos relacionados con el café, a fin de optimizar decisiones de compra mensuales. Este sistema permitirá a los productores acceder a proyecciones sobre precios y categorías de productos, basándose en datos históricos de precios y características de productos relacionados. A continuación, se detallan los pasos de implementación:

### Integración del Modelo de Predicción:

### Incorporación del modelo de regresión en una plataforma accesible, la cual estará habilitada para recibir datos mensuales que en conjunto con el DANE se puede proporcionar predicciones sobre precios de productos específicos.

### Configuración de una interfaz de usuario que permita la entrada de datos y la visualización de resultados para facilitar el acceso a las proyecciones.

### Despliegue en Plataforma de Acceso Público o Privado:

### Diseño de una plataforma (web o móvil) donde los productores puedan ver las predicciones de precios y planificar la compra de productos según las fluctuaciones estimadas.

### Implementación de un sistema de autenticación, si se requiere, para asegurar que solo usuarios autorizados accedan a la información de proyección.

### Actualización Mensual de Datos:

### Establecimiento de un proceso automatizado para la actualización mensual de datos de precios de productos y tipos de insumos, lo que permitirá mantener el modelo actualizado con los últimos datos del mercado.

### Integración de datos históricos con los nuevos para mejorar la precisión del modelo y ajustarlo según las tendencias más recientes.

### Monitoreo y Mantenimiento del Modelo:

### Configuración de un sistema de monitoreo continuo que evalúe el rendimiento del modelo, específicamente en términos de precisión de las proyecciones y utilidad para los usuarios finales.

### Creación de un plan de mantenimiento que incluya la revisión periódica del modelo y ajustes para optimizar su precisión en respuesta a cambios en las variables de precio y tipo de producto.

### Capacitación y Adopción del Sistema por los Usuarios:

### Diseño de materiales de capacitación para que los usuarios puedan interpretar adecuadamente las predicciones y entender cómo aplicarlas en sus decisiones de compra.

### Seguimiento del uso de la plataforma y recopilación de retroalimentación de los usuarios para mejorar la experiencia y optimizar la utilidad de las proyecciones en la planificación de compras.

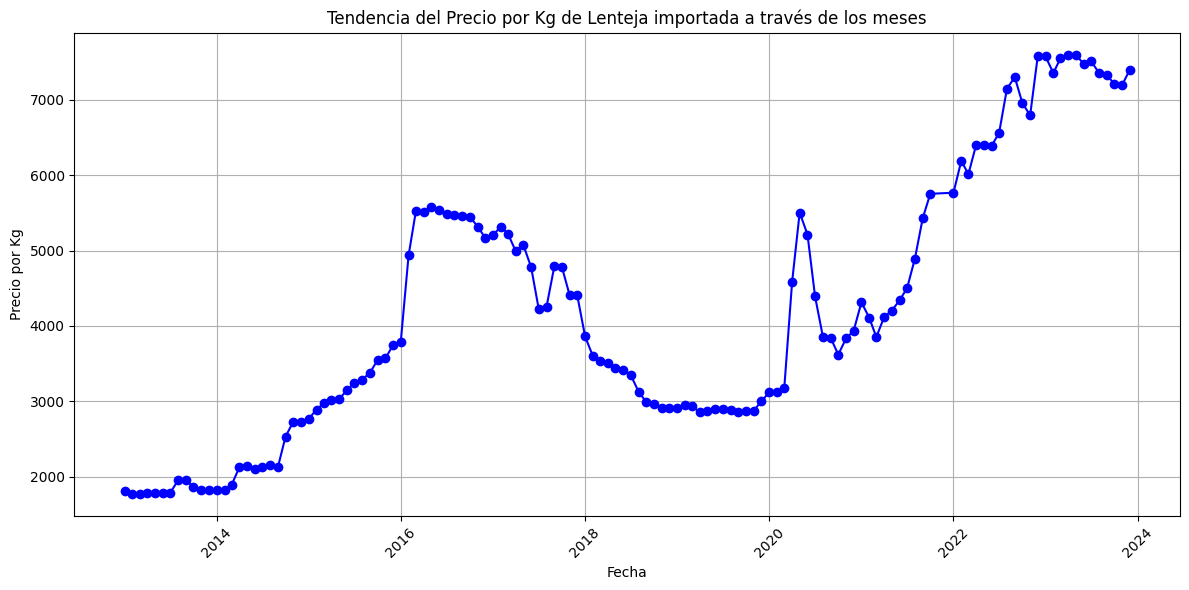
### Evaluación de Impacto y Mejoras Continuas:

### Medición del impacto del sistema en la planificación y efectividad de compras de los usuarios, recopilando datos sobre cómo las predicciones influyen en sus decisiones.

### Identificación de áreas de mejora en la plataforma y en el modelo, ajustando procedimientos según la evolución de los precios y las necesidades de los usuarios.

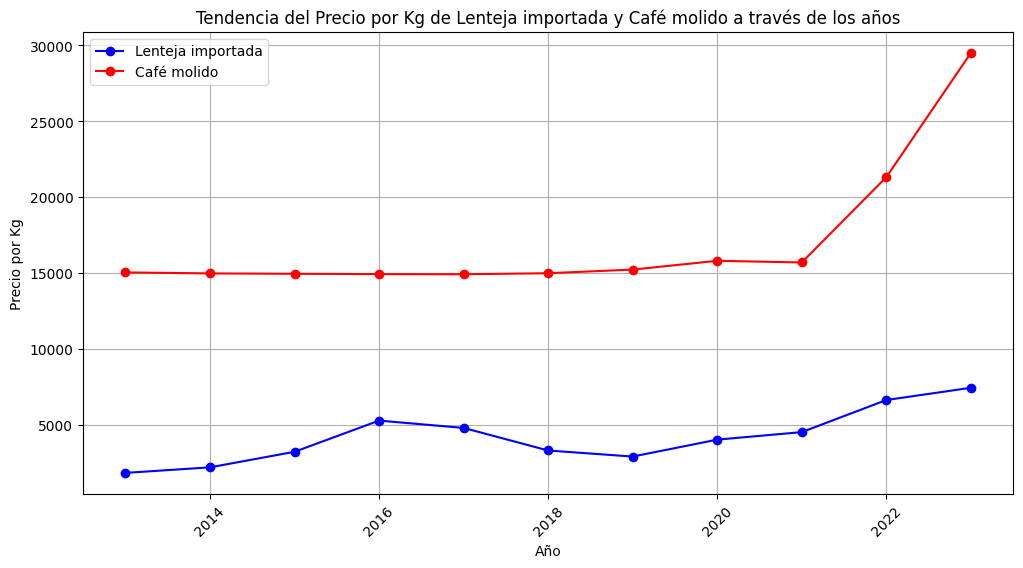
### 

2.8. **Resultados Esperados**

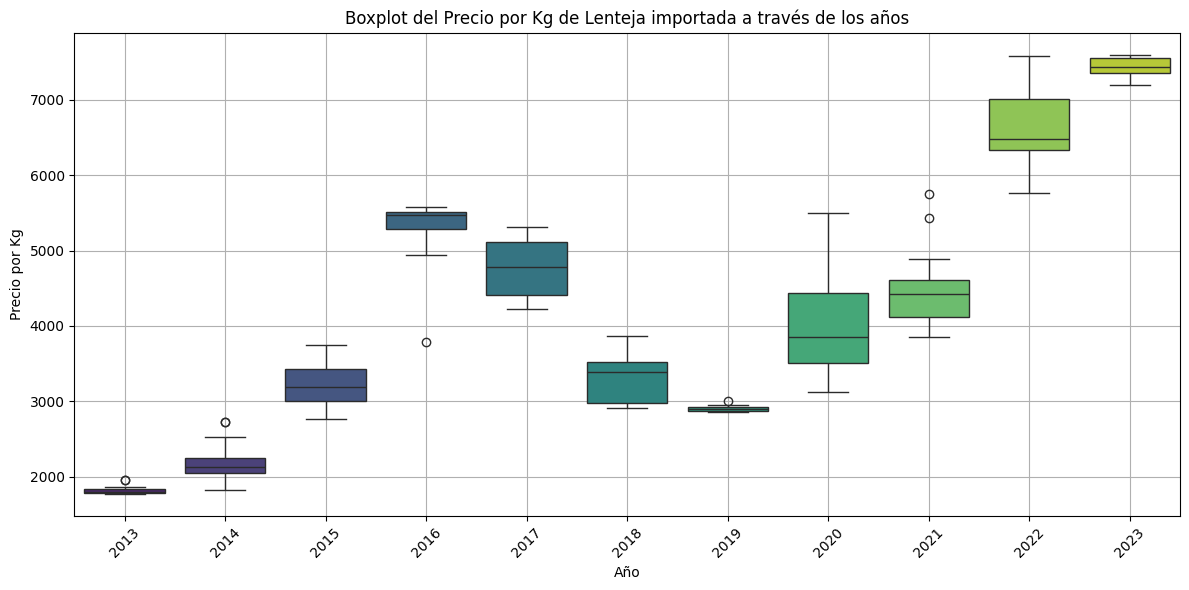
*****Tendencia del precio de la lenteja importada*

****

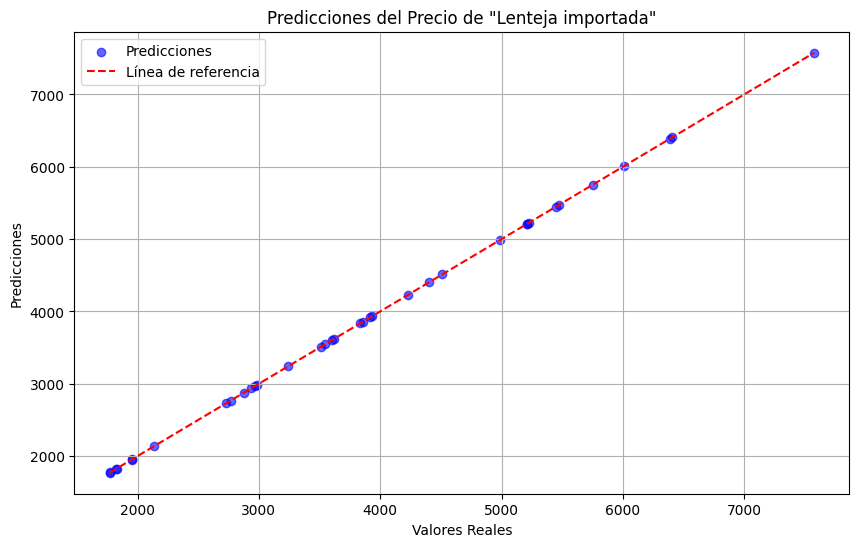
**Graf:** *Información de AgroNegocio*



*Tendencia del precio de dos productos distintos*



*Distribución anual del precio de un producto*

*Modelo de regresión lineal múltiple*