



## **GUÍA 3 - ESTRUCTURAR EL PROYECTO**

*AA3.3.2 – Del área Procedimental*

Juan David Maecha Sabogal

Steeven sebastian Pedraza Martinez

Luis Felipe Sanchez Sierra

Servicio Nacional de Aprendizaje – SENA, Centro de Electricidad, Electrónica y

Telecomunicaciones - CEET

Instructor Albeiro Ramos

Bogotá, 13 de Noviembre de 2024

## TABLA DE CONTENIDO

<b>INTRODUCCIÓN.....</b>	<b>4</b>
<b>LucasEnt.....</b>	<b>5</b>
1. PREGUNTA DE INVESTIGACIÓN.....	5
2. ANTECEDENTES (ESTADO DEL ARTE).....	5
2.1. Herramientas de comparación de precios.....	5
2.1.1. Existencia de Herramientas Similares:.....	5
2.1.2. Funcionalidades:.....	5
2.1.3. Limitaciones:.....	5
3. JUSTIFICACIÓN DEL PROYECTO.....	6
4. OBJETIVO GENERAL.....	6
5. OBJETIVO ESPECÍFICO.....	6
5.1. Extracción de datos:.....	6
5.2. Limpieza y procesamiento de datos:.....	6
5.3. Construcción de la base de datos:.....	6
5.4. Análisis de datos:.....	6
6. RESULTADOS ESPERADOS.....	7
6.1. Resultados para los consumidores.....	7
<b>CONCLUSIONES.....</b>	<b>8</b>
<b>BIBLIOGRAFÍA.....</b>	<b>9</b>
Fuentes para la Extracción de Datos y Limpieza de Datos:.....	9
Fuentes para la Construcción de Bases de Datos:.....	9
Fuentes para el Análisis de Datos:.....	10



## LISTA DE FIGURAS

[Figura 1. Retorno real de acciones, títulos del tesoro americano, oro y dólar de 1802 a 2012](#) 4

## LISTA DE TABLAS

[Tabla 1. Número de niñas embarazadas en colegios del sector norte y sur de Bogotá](#) 5

## INTRODUCCIÓN

En la actualidad, los consumidores se enfrentan a la dificultad de encontrar las mejores ofertas para los productos de la canasta familiar debido a la gran cantidad de supermercados y tiendas en línea que existen, así como la falta de herramientas que faciliten la comparación de precios en tiempo real. En Colombia, donde los precios varían significativamente entre cadenas como Tiendas D1, Alkosto y Mercado Libre, la toma de decisiones de compra informadas es crucial para el ahorro. A través del uso de web scraping, este proyecto propone el desarrollo de un software que permita a los consumidores comparar precios de productos de la canasta familiar entre diferentes tiendas, considerando además su ubicación geográfica y las ofertas disponibles. Este sistema no solo facilitará el acceso a la información, sino que optimizará el proceso de compra, ayudando a los consumidores a realizar elecciones más informadas y económicas.

## LuckasEnt

### 1. PREGUNTA DE INVESTIGACIÓN

El Proyecto LuckasEnt se planteó la siguiente pregunta, *¿Cómo se puede desarrollar un software basado en web scraping que permita a los consumidores colombianos comparar precios de productos de la canasta familiar en diferentes supermercados (DI, Alkosto, Olímpica y Mercadolibre), considerando su ubicación geográfica y las ofertas disponibles, para facilitar la toma de decisiones de compra informadas y promover el ahorro?*

### 2. ANTECEDENTES (ESTADO DEL ARTE)

#### 2.1. Herramientas de comparación de precios

##### 2.1.1. Existencia de Herramientas Similares:

PriceGrabber, ShopSavvy, Google Shopping, Tiendeo, Trivago o incluso páginas web de supermercados que permiten comparar productos.

##### 2.1.2. Funcionalidades:

Características principales de estas herramientas, como la búsqueda por producto, filtros por categoría, comparación de precios en diferentes tiendas, información sobre ofertas y promociones, etc.

##### 2.1.3. Limitaciones:

Herramientas existentes, como la falta de información sobre algunos supermercados, la poca precisión en los datos, la falta de personalización o la dificultad para encontrar tiendas cercanas.

### 3. JUSTIFICACIÓN DEL PROYECTO

La meta de este sistema informático denominado LuckasEnt consiste en ofrecer a los usuarios en una fuente cierta para la comparación de los precios de los productos que pueden obtener en diferentes mercados como Tiendas D1, Tiendas Alkosto, Supertiendas Olímpica y Mercado libre, de tal forma que, al momento de realizar las compras, puedan buscar la que mejor conserva su presupuesto en el momento de realizar las compras.

### 4. OBJETIVO GENERAL

Desarrollar un Sistema de Información Web llamado 'LuckasEnt' para el Análisis y soporte de Precios y Productos, frente a la toma de decisiones en los procesos de abastecimiento de supermercados independientes, mediante el web scraping.

### 5. OBJETIVO ESPECÍFICO

#### 5.1. Extracción de datos:

Extraer y almacenar los precios de (*producto*) para al menos 10 tiendas online distintas.

Recoger las reseñas de clientes de (*producto o servicio*) en D1 y en ARA.

#### 5.2. Limpieza y procesamiento de datos:

Limpiar y normalizar los datos extraídos para asegurar su calidad y uniformidad.

Estructurar los datos en forma de tablas para facilitar su análisis.

#### 5.3. Construcción de la base de datos:

Crear una base de datos relacional para almacenar los datos extraídos.

Definir las relaciones entre las distintas tablas de la base de datos.

#### 5.4. Análisis de datos:

Identificar patrones y tendencias en los datos extraídos.

Realizar análisis comparativos entre distintas fuentes de información.

Generar informes personalizados para el usuario de la base de datos.

## 6. RESULTADOS ESPERADOS

Los resultados que esperamos obtener del proyecto LuckasEnt deben ser la traducción de los resultados que se esperan obtener del software desarrollado e implementado.

### ***6.1. Resultados para los consumidores***

#### **6.1.1. Menor gasto y ahorro de tiempo:**

Menor tiempo dedicado a buscar y comparar precios en diferentes supermercados.

Aumento del porcentaje de consumidores que encuentran las mejores ofertas y precios.

Aumento del porcentaje de disminución del gasto medio de la canasta familiar.

#### **6.1.2. Mejor saber:**

Mayor conocimiento de los precios y de las ofertas existentes en el mercado.

Aumento de la confianza del consumidor frente a sus compras.

Mayor satisfacción con las decisiones de compra realizadas.

#### **6.1.3. Mayor información:**

Incremento en el porcentaje de los consumidores que pueden acceder a la información de precios rápidamente.

Mayor capacidad para comparar precios entre los supermercados, incluso aquellos con escasa presencia online.

## CONCLUSIONES

El software a desarrollar, LuckasEnt, tiene el potencial de revolucionar la forma en que los consumidores colombianos realizan sus compras de productos de la canasta familiar. Al utilizar un aspecto de web scraping y llevar a cabo la comparación de precios, los usuarios tendrán acceso a información actualizada y precisa de múltiples tiendas en línea. La inclusión de factores como la ubicación y las promociones disponibles aportará un valor añadido significativo, ya que fomentará la toma de decisiones de compra rápida y basada en datos, lo que a su vez representará un ahorro económico importante. Por otro lado, con una base de datos relacional, los datos estarán disponibles y bien organizados, contribuyendo a una experiencia de compra mejorada y promoviendo una mayor satisfacción en la elección de compra.



## BIBLIOGRAFÍA

### **Fuentes para la Extracción de Datos y Limpieza de Datos:**

Mitchell, R. (2018). Web Scraping with Python: Collecting Data from the Modern Web. O'Reilly Media.

Este libro proporciona una introducción completa a las técnicas de web scraping utilizando Python, incluyendo cómo extraer datos de tiendas en línea y plataformas de reseñas, así como cómo manejar los datos recolectados de forma efectiva.

Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. (2015). Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. Wiley.

Aquí se abordan técnicas de web scraping y procesamiento de datos, destacando cómo limpiar y organizar los datos extraídos para su análisis posterior.

Hinton, M., & Robinson, H. (2019). Data Wrangling with Pandas: Real-World Techniques for Data Cleaning and Analysis. Packt Publishing.

Este texto es útil para entender cómo limpiar y normalizar grandes volúmenes de datos extraídos, un paso clave para asegurar la calidad y consistencia en la información recolectada.

### **Fuentes para la Construcción de Bases de Datos:**

Rob, P., & Coronel, C. (2017). Database Systems: Design, Implementation, & Management. Cengage Learning.

Proporciona una visión completa de la construcción de bases de datos relacionales, cómo estructurar los datos extraídos, y la creación de relaciones entre diferentes tablas para facilitar su análisis posterior.

Elmasri, R., & Navathe, S. B. (2016). Fundamentals of Database Systems (7th ed.). Pearson.

Este libro profundiza en la creación de bases de datos relacionales, cubriendo la normalización de datos, creación de relaciones y estructura de tablas.

### **Fuentes para el Análisis de Datos:**

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.

Este texto es útil para aquellos que quieran realizar análisis de datos, como identificar patrones y tendencias en los datos extraídos. También cubre aspectos más avanzados del análisis comparativo entre diferentes fuentes de información.

Bohanec, M., & Zupan, B. (2016). Data Mining and Knowledge Discovery Handbook. Springer.

Proporciona técnicas de minería de datos y análisis comparativo, que son esenciales cuando se trabaja con grandes volúmenes de datos extraídos de diferentes fuentes y se busca generar reportes personalizados.

Groves, J., & Bennett, R. (2018). Data Analysis for Business, Economics, and Policy. Cambridge University Press.

Este libro ofrece un enfoque práctico sobre cómo analizar datos de diferentes fuentes, extraer tendencias, y generar reportes personalizados para usuarios finales.