

Informe Técnico del Proyecto: La percepción del valor en el Mercado Cubano

Autor: Eduardo Rojas Cuan

Materias: - Introducción a la Ciencia de Datos
- Introducción a la Programación

21 de enero del 2026

Índice:

1 – Introducción

2 – Adquisición de Datos

2.1 – Fuente Primaria: Levantamiento de Datos (JSON)

2.2 – Fuente Secundaria: Tasas de Cambio (CSV)

3 – Lógica de Programación y Procesamiento

3.1 – Navegación Recursiva

3.2 – Normalización y Limpieza

4 – Análisis y Visualización

5 – Conclusiones

1- Introducción.

¿Caro o Barato?

Este proyecto nace de una duda que todos tenemos cuando salimos a la calle: ¿cuál es el precio real de las cosas hoy en Cuba? A veces sentimos que algo es "caro" de forma subjetiva, pero nos falta un punto de comparación claro entre diferentes municipios y negocios para entender el poder adquisitivo real frente al salario promedio.

El objetivo no fue hacer un estudio económico definitivo, sino uno de carácter referencial. Quería realizar un sondeo sobre las condiciones de venta y precios de los negocios donde pudiera comprar a menor precio los productos seleccionados, además de cómo se comportan en diferentes tasas de cambio

2. Adquisición de Datos

Para que el análisis fuera fiel a la realidad, utilicé una metodología mixta que combina trabajo de campo con fuentes de referencia externa.

2.1. Fuente Primaria: Levantamiento de Datos (JSON)

La base del proyecto son datos reales recolectados directamente de pizarras y etiquetas en diversas Mipymes y mercados estatales.

Metodología: Se digitalizaron manualmente los precios de productos clave (alimentos, aseo, bebidas) organizándolos en una estructura jerárquica.

Formato: Elegí JSON como formato de almacenamiento porque tenía una estructura de anidar los datos ya pensada, Municipio (o Estatal) > Negocio > Categoría > Producto.

2.2. Fuente Secundaria: Tasas de Cambio (CSV)

Incorporé un archivo externo (Tasas.csv) con las tasas de cambio referenciales del mercado informal (El Toque).

Este archivo actúa como un sistema de respaldo . Aunque la mayoría de mis datos ya fueron registrados en su moneda final, el sistema consulta este archivo para realizar conversiones automáticas si detecta un precio en divisa (USD/MLC/EUR) que necesita ser expresado en CUP para el análisis.

3. Lógica de Programación y Procesamiento

El reto principal no fue solo tener los datos, sino cómo acceder a ellos dado que estaban anidados unos dentro de otros.

3.1. Navegación Recursiva

Como no sabía de antemano cuántas categorías o subniveles podría tener un negocio, no podía usar bucles simples. Implementé una función recursiva (recorrer_json).

La cual actúa como un explorador que entra en cada diccionario del JSON. Si encuentra más diccionarios, sigue bajando. Si encuentra la key "Precios", se detiene, captura el value y el camino que recorrió para llegar ahí. Esto permite que el proyecto sea escalable si se modifica la estructura de datos (precios).

3.2. Normalización y Limpieza

Una vez extraído el crudo (Datos), apliqué varios filtros para garantizar la calidad y corregir fallos a la hora de recorrer el json:

Eliminación de duplicados: Se limpian los registros para evitar que un mismo producto cuente dos

veces en las estadísticas.

Estandarización de Moneda: Prioriza el precio registrado en el JSON; si no existe en la moneda deseada, utiliza la tasa de cambio externa para calcularlo.

4. Análisis y Visualización

Con los datos ya limpios y en una lista plana, procedí a agruparlos para responder preguntas concretas mediante gráficos (usando Matplotlib):

Promedios y Desviaciones: Calculé el precio promedio de cada producto para establecer una línea base. Luego, medí cuánto se aleja cada negocio de ese promedio. Esto permite decir con datos: "Este lugar es un 20% más caro que la media del mercado". La lógica detrás de ello es que cada producto 10% mas caro por encima del promedio suma 1 punto al contador de cada mipyme, y en negativo suma -1, la gráfica explica las categorías, las barreras que definen la categoría son, muy caro, +40, caro si es entre 20 y 40, neutral, si es entre 20 y 0, bueno, si es entre 0 y -10, economico, si es por debajo de -10.

Distribución por Municipios: Analicé qué zonas concentran la mayor oferta de ciertos productos

Comparativa de Precios: Gráficos de líneas para ver la variabilidad de precios de los productos en las categorías de municipios. Cabe destacar que la gráfica no busca dar un resultado concreto, mas bien esbozar una idea conceptual sobre como fluctúan los precios según las zonas, debido al limitado tamaño de la muestra no debe de emplearse de manera definitiva, sino referencial

Impacto en el Salario: Gráfico de Barras simple, con el objetivo de ingresar el establecimiento a analizar y el salario con el que disponible del usuario para comprar. Se muestra un simple cálculo porcentual de cuanto representaría cada producto disponible durante el análisis con respecto a el salario introducido

5. Conclusiones

Este trabajo demuestra cómo la programación permite transformar observaciones cotidianas dispersas en información útil. A través del uso de estructuras de datos anidadas (JSON), algoritmos recursivos y manejo de archivos externos (CSV), se logró construir un sistema que no solo almacena precios, sino que permite interpretarlos.