

Informe Proyecto Ciencia de Datos 1er Semestre

Ricardo Vázquez Rodríguez

1-Problemática y Objetivo del Proyecto:

1.1-Problemática

El salario mínimo en Cuba, se establece según la canasta familiar y se tiene en cuenta las necesidades básicas de las familias, principalmente, las alimenticias. Pero la realidad es que, en la actualidad, el salario dista mucho de satisfacer estas necesidades, cuando se constatan los elevados precios de los alimentos en el mercado. Aunque, la percepción de la población es que la inflación real supera las cifras oficiales, en gran parte debido a la dificultad de medir adecuadamente el comportamiento de los precios en un mercado negro que, a menudo, es más dinámico y surtido que el mercado regulado.

2.2-Objetivo:

-Este trabajo tiene como objetivo hacer un análisis sobre la coherencia entre precio de los principales productos más comercializados por el mercado no estatal y el salarios de los trabajadores.

-Recopilar información sobre la situación del dólar en el mercado informal.

-Demostrar si realmente se puede vivir con un salario promedio en Cuba.

2-Adquisición y estructuración de los datos:

2.1-Recolección de datos:

-Se visitaron 35 establecimientos que expenden alimentos, 34 pertenecientes al mercado no estatal y uno perteneciente al mercado estatal en divisas. Se tomó una muestra para el estudio de 33 establecimientos del mercado no estatal que en el momento de la visita vendían los productos siguientes:

- azúcar (kg),
- arroz (kg),
- huevo (cartón 30 unid),
- pechuga de pollo (2 kg),
- picadillo de pollo (libra),
- aceite (1L),
- spaguettis (500g),
- perritos (paquete 10 unid),
- pan (paquete de 8 unidades),
- refresco en polvo,
- galletas saladas,
- cerveza,
- whisky y
- algún refresco gaseado.

-De estos productos se obtuvo el tipo(alimento o bebida), el precio, la moneda en que se vende y el origen.

-Para contrastar la información se obtuvieron los mismos datos de una tienda del mercado estatal en divisas, además de las coordenadas de cada establecimiento.

-Se buscó el precio del dólar día por día, los datos se obtuvo del canal de Whatsapp oficial de el toque, se copiaban los mensajes con esta información y se copió hacia un documento word.

-Se buscó información de la ONEI para datos como el salario promedio en La Habana.

2.2-El procesamiento de datos:

-Se realizó mediante los json, primero el json de los precios de los productos en el mercado no estatal el cual contiene un diccionario que contiene las llaves:

- id:que identifica a cada establecimiento; name: se relaciona con el nombre de cada negocio;
- location: contiene un otro diccionario conformado por las siguientes llaves: province, que se refiere a la provincia que se encuentra el local, municipality, que es el municipio en el que se encuentra el establecimiento y otras dos llaves referentes a la latitud y longitud(lat y long);
- currency: se refiere a la moneda en que se venden los productos;
- payment methods:refiere el método de pago;
- product: contiene una lista de diccionarios, estos diccionarios contienen las siguientes llaves: id, es una etiqueta para productos similares, type, que se refiere al tipo de producto, name, viene siendo el nombre además en algunos productos también incluye el gramaje en esta llave, currency, se refiere a la moneda en la cuál se vende el producto, price, contiene el valor del precio de el producto y origin que se refiere a la procedencia(nacional o importado).

-Para el json de la tienda estatal se utilizó la misma estructura.

```
{
  "id": "01",
  "name": "Carnicos Triple A",
  "location": {
    "province": "La Habana",
    "municipality": "Boyeros",
    "lat": "23.059183",
    "long": "-82.394824"
  },
  "currency": ["CUP"],
  "payment_methods": ["Efectivo ", "Transferencia"],
  "product": [
    {
      "id": "01",
      "type": "Alimentos",
      "name": "Azúcar 1kg",
      "currency": ["CUP"],
      "price": "650",
      "origin": "Importada"
    }
  ]
},
```

-Además se creo otra estructura de json para guardar los datos de El Toque el cuál contiene un diccionario que contiene las llaves:

- moneda: identifica el tipo de divisa;

- fecha_inicio: se refiera a la fecha en la cual se empezó a recolectar el dato;
- fecha_fin: identifica a la fecha en la que se acabó la recopilación de datos;
- datos: contiene una lista de diccionarios conformado por las llaves de valor, que se refiere a el valor de la divisa en ese momento y fecha, que es la fecha del dato.

```
{
  "moneda": "USD",
  "fecha_inicio": "2025-11-01",
  "fecha_fin": "2026-01-10",
  "datos": [
    {
      "fecha": "2025-11-01",
      "valor": 460
    }
  ]
}
```

3-Procesamientos Realizados(código):

3.1-Funciones para calcular y recopilar datos:

- promedio: Esta recibe como parámetro el nombre del producto buscado (string), itera en todos los json de las mypimes y guarda los precios de ese mismo producto en todos json, luego calcula el promedio de precios dividiendo la sumatoria de todos los valores entre la cantidad de valores y devuelve ese resultado en un diccionario.
- por_ciento: Esta recibe como parámetro una tupla de un diccionario (parte) y un entero (total), luego de ese diccionario extrae el valor al que se le desea sacar un por ciento con respecto al total, divide el total a la parte y luego lo multiplica por 100, luego agrega el resultado a un diccionario con su respectiva key, y devuelve ese resultado en un diccionario.
- promedio2: Esta recibe como parámetro una lista de valores a los cuales les saca el promedio igual que promedio y devuelve ese resultado en entero.
- prom_usd: Esta recibe como parámetro vacío, carga los datos del json de El Toque, extrae el valor de cada fecha en el archivo, los agrega a una lista vacía ,le saca el promedio y retorna el resultado en un entero.
- precios_tienda_usd: Esta recibe como parámetro el nombre del producto buscado(string), carga el valor del producto buscado y ese valor es la salida.
- minimo: Esta es parecida a la de promedio, recibe la misma entrada pero en vez de devolver el promedio, devuelve el mínimo.

3.2-Funciones gráficas:

- grafica1: Esta recibe un parámetro vacío como todas las funciones de gráficas, primero carga los datos del json de El Toque, los guarda en variables, luego con las funciones incluidas en matplotlib se creó la gráfica la cual es de puntos, ademas se le agregó una parte para dividir las fechas en períodos para que la gráfica sea más asequible.

- grafica2: Esta recibe como parámetro el diccionario creado con la función promedio, recoge las keys y los values del diccionario y los grafica en una gráfica de barras verticales.
- grafica3: Esta recibe el mismo parámetro que gráfica 2, extrae las keys y las values las cuales pasa a lista y calcula el porciento con la función por_ciento y luego los convierte en una gráfica de barras horizontales.
- grafica4: Esta recibe un parámetro vacío, pero dentro de la función se obtiene un diccionario con la función promedio de productos seleccionados multiplicados por ciertos enteros y luego se les saca el porciento para convertirla en una gráfica de pastel.
- grafica5: Esta recibe un parámetro vacío, esta función viene con dos diccionarios los cuales uno contiene los precios promedio y el otro los precios de las tiendas usd extraídos con la función precios_tienda_usd, luego se gráfica en una de barras que compara los valores de ambos diccionarios.

4-Lenguajes y bibliotecas utilizados:

4.1-Lenguajes:

-Python3
-Jupyter Notebook

4.2-Bibliotecas:

-Matplotlib
-json
-os

5-Instrucciones:

5.1-Requisitos:

-Para poder ejecutar el proyecto es necesario tener el Visual Studio Code instalado o JupyterLab o cualquier programa que contenga el ipykernel.
-Además de tener instalado matplotlib que se puede instalar con pip install matplotlib.

5.2-Pasos:

-Abrir el archivo El Dilema del Consumidor.
-Ejecutar las celdas de código.