

INFORME

PROYECTO TECNOLÓGICO INTEGRADOR



SSTTT5

Fecha: 16/09/2022

— ÍNDICE GENERAL

1

Introducción

2

Alcance

3

Tecnologías Utilizadas

4

Librerías

5

Personal Involucrado

6

Desarrollo del Proyecto

INTRODUCCIÓN



El trabajo se encuentra realizado por el grupo SSTTT5 de la carrera Ciencias de Datos e Inteligencia Artificial, para el Proyecto Tecnológico Integrador, teniendo como objetivo simular una experiencia real colaborativa en un entorno de desarrollo ágil.

ALCANCE

El proyecto tiene como principal función extraer información de sitios web mediante programas de software, siendo desarrollado en Python y haciendo uso de varias librerías que permiten técnicas como web scrapping y persistencia de datos, utilizadas para el análisis exploratorio de las ventas de los últimos años en supermercados de Argentina.

También se utilizan herramientas como Trello para la organización de tareas y Github para el repositorio de archivos.

TECNOLOGÍAS UTILIZADAS



Herramienta de gestión de tareas: TRELLO

[Proyecto Final SSTTT5](#)



Herramienta colaborativa de código y repositorio de archivos

[ispc-programador2022/SSTTT5](#)



LIBRERIAS

- Python Version 3.10.6
- async-generator Version 1.10
- attrs Version 22.1.0
- certifi Version 2022.9.24
- cffi Version 1.15.1
- cffi Version 1.15.1
- click Version 8.1.3
- colorama Version 0.4.6
- contourpy Version 1.0.6
- cycler Version 0.11.0
- distlib Version 0.3.6
- exceptiongroup Version 1.0.1
- filelock Version 3.8.0
- fonttools Version 4.38.0
- greenlet Version 2.0.0.
- h11 Version 0.14.0
- idna Version 3.4
- Jinja2 Version 3.1.2
- joblib Version 1.2.0
- kiwisolver Version 1.4.4
- MarkupSafe Version 2.1.1
- matplotlib Version 3.6.2
- numpy Version 1.23.4
- outcome Version 1.2.0
- packaging Version 21.3
- pandas Version 1.5.1
- patsy Version 0.5.3
- Pillow Version 9.3.0
- platformdirs Version 2.5.3
- pycparser Version 2.21
- pyodbc Version 4.0.34
- py parsing Version 3.0.9
- PySocks Version 1.7.1
- python-dateutil Version 2.8.2
- pytz Version 2022.6
- scikit-learn Version 1.1.3
- scipy Version 1.9.3
- selenium Version 4.6.0
- six Version 1.16.0
- sklearn Version 0.0.post1
- sniffio Version 1.3.0
- sortedcontainers Version 2.4.0
- statsmodels Version 0.13.5
- threadpoolctl Version 3.1.0
- trio Version 0.22.0
- trio-websocket Version 0.9.2
- typing_extensions Version 4.4.0
- urllib3 Version 1.26.12
- virtualenv Version 20.16.6
- Werkzeug Version 2.2.2
- wsproto Version 1.2.0

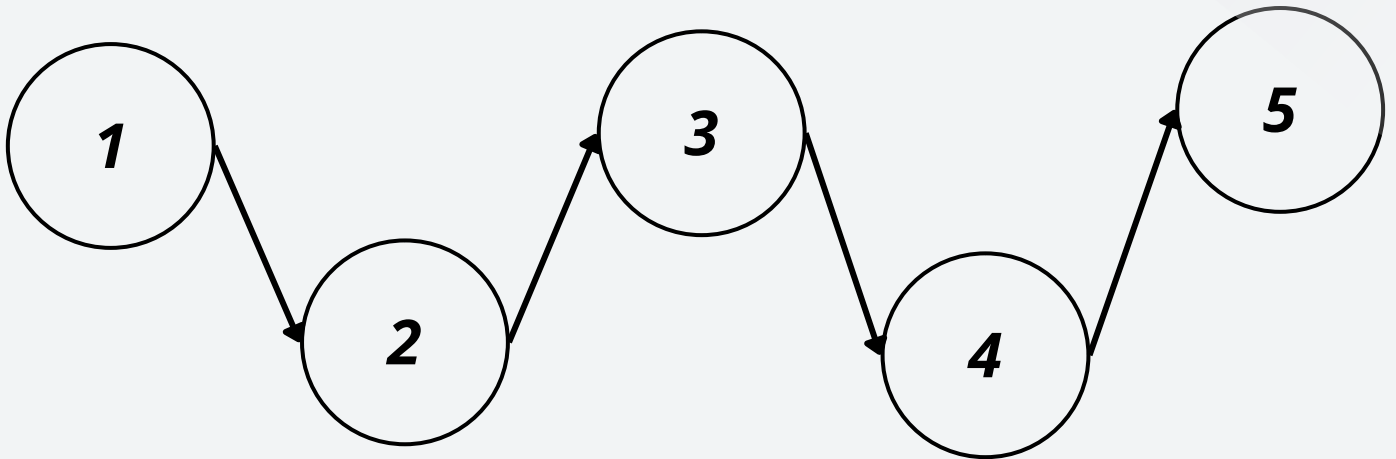
PERSONAL INVOLUCRADO

- SEQUEIRA ADRIAN
- SELVA CRISTIAN
- TORO GUTIÉRREZ LUCIANO
- TALAVERA NATALIA AGUSTINA
- TOÑANES FRANCO ADRIÁN

GRUPO:SSTTT5

Fecha: 16/09/2022

DESARROLLO DEL PROYECTO



1. **Elección del Tema**
2. **Búsqueda fuente de datos**
3. **Web Scraping**
4. **Almacenamiento en base de datos**
5. **Análisis Estadístico**

• Elección del Tema

El objetivo del proyecto es realizar un análisis exploratorio del consumo en supermercados de Argentina y la inflación.

• Búsqueda fuente de datos

Teniendo en claro el objetivo se seleccionaron dos fuentes de datos relevantes para el tema.

1) [Datos Argentina | Ventas en supermercados](#)

2) [IPC Argentina - Tasa de inflación: mensual - Investing.com](#)

- **Web Scraping**



Para realizar Web Scraping, se utilizo BeautifulSoup que es una biblioteca de Python para analizar documentos HTML. Esta biblioteca crea un árbol con todos los elementos del documento y puede ser utilizado para extraer información.

También se utilizo, Selenium que es un entorno de pruebas de software para recopilar la información dinámica de aplicaciones basadas en la web.

- **Almacenamiento en base de datos**



Se genero una base de datos con dos tablas que permiten hacer consultas en SQL.

```
SELECT *  
FROM "nombre_de_tabla"
```

- "ventas"
- "inflacion"

Se encuentran en el archivo: DataSql.py

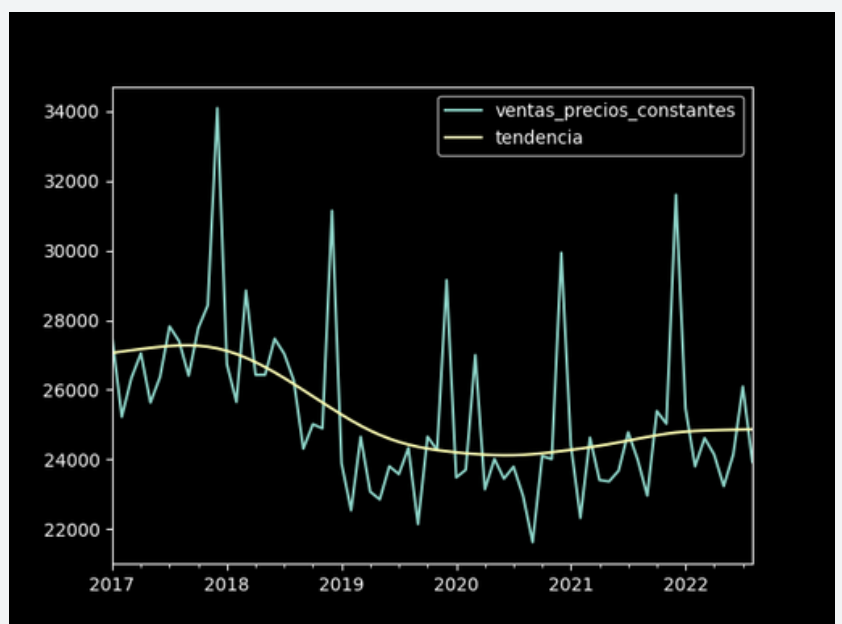
- **Análisis Estadístico**

En el análisis estadístico se utilizaron librerías de Python

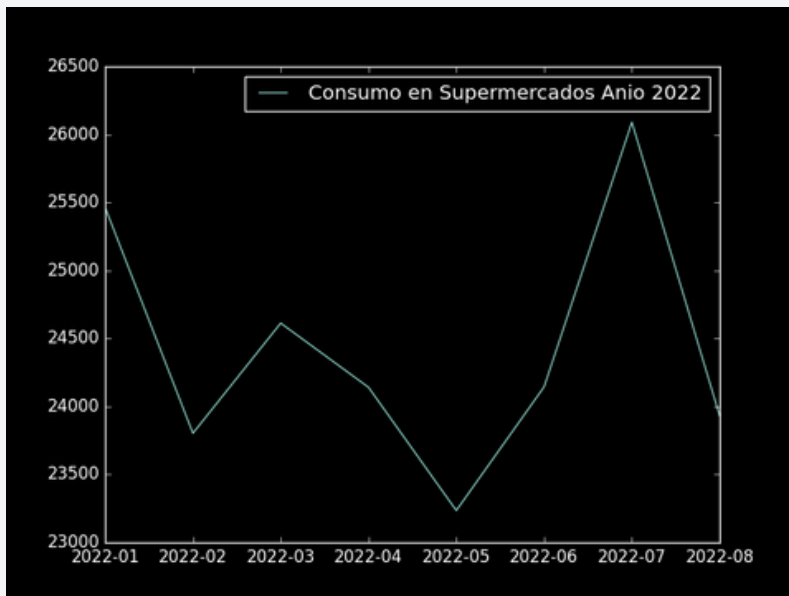


Evolución de las ventas desde el año 2017 al año 2022

Visualización componente
tendencia y componente cíclico.

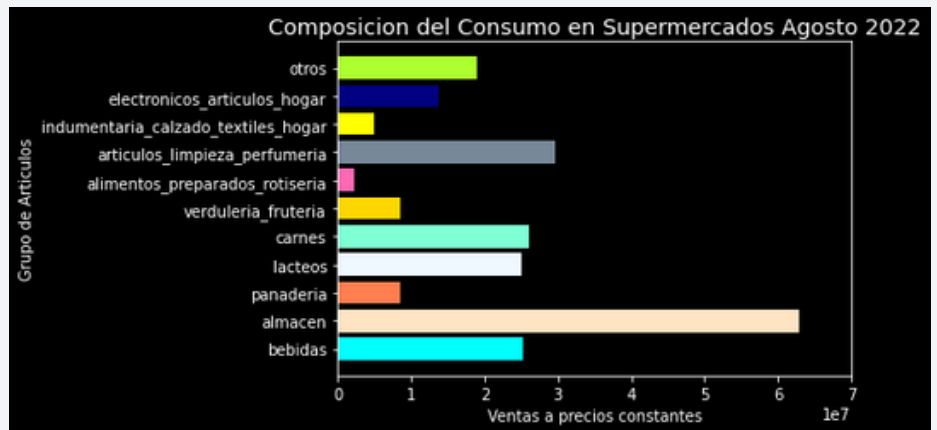


- **Análisis Estadístico**

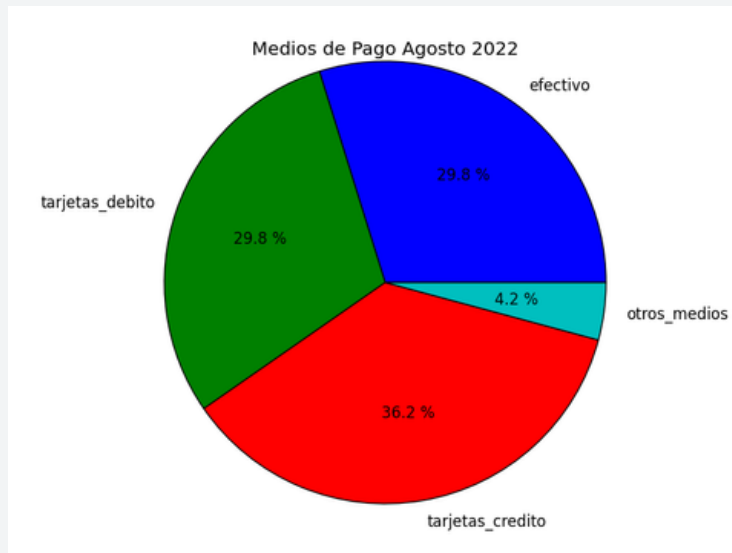


Análisis de Ventas Mensuales 2022

Composición del consumo en Supermercados para el mes de Agosto 2022



• Análisis Estadístico



Medios de pagos utilizados
Agosto 2022

Resumen de los datos:

--> Promedio Ventas 2022: 24427.7550

--> Variabilidad Ventas 2022 (Desvío Estándar): 933.0240

--> Mediana Ventas 2022: 24142.7789

--> Mes de Mayores Ventas 2022: Julio

--> Monto Total Mayores Ventas 2022: 26090.18031

--> Mes de Menores Ventas 2022: Mayo

--> Monto Menores Ventas 2022: 23235.0963

--> Variaciones Mensuales Ventas 2022