**End-to-End Data Analysis of Electronic Devices**

**Documentation**

**What to expect / Scope of the project:**

This project is an end-to-end data analysis pipeline focused on **electronic devices** such as smartphones and laptops. It aims to simulate the workflow of a professional data scientist, from data acquisition to insight delivery.

The key deliverable is a locally hosted website where users can generate insights based on specific criteria—such as a particular **brand**, **feature**, or an **overall market analysis**—for selected electronic devices.

This project emphasizes a **full-stack data science workflow**, integrating data engineering, data analysis, web development, and API design.

**Project Flow:**

1. Setting up a virtual environment
2. Setting up Git for version control
3. Log into a log file
4. Scraping data from websites to build the dataset
5. Cleaning this very messy and untidy data
6. Exporting this data to an OLTP Database
7. Creating an OLAP Database (or data warehouse) and extracting this clean data
8. Feature engineering to create more useful columns
9. Performing some basic EDA to understand the data better
10. Creating and designing a website for the users to extract insights from.
11. Creating API endpoints on another website, which will be called by the client website to perform analysis and display it to the user.

**Skills Demonstrated:**

Knowledgeable and Proficient in:

1. **Managing Virtual Environments**: Used conda to isolate dependencies and streamline development and deployment workflows.
2. **Implementing Version Control**: Leveraged Git for efficient code tracking, branching, and collaboration across project stages.
3. **Performing Advanced Web Scraping**: Collected data from dynamic and static websites using selenium and beautifulsoup4.
4. **Cleaning and Preprocessing Data**: Transformed messy, unstructured data into analysis-ready formats. (“Garbage In, Garbage Out”)
5. **Designing and Managing Databases**: Utilized OLTP and OLAP systems to store, query, and analyze large volumes of structured data efficiently.
6. **Conducting Data Analysis and Visualization**: Explored data using tools like pandas, matplotlib, and seaborn to extract insights and reveal trends.
7. **Developing Web Applications with Flask**: Built a responsive, locally hosted website to deliver real-time, interactive insights to users.
8. **Creating API Endpoints**: Designed modular Flask APIs to enable smooth data exchange between client and server layers.
9. **Implementing Logging and Exception Handling**: Integrated Python's logging module and robust try-except blocks to monitor application flow, capture errors, and ensure fault-tolerant execution across scraping, data processing, and API services.