# **🚀 Advanced BI Tool with Automated Analytics**

Welcome to the Advanced BI Tool, a powerful, open-source application built with Streamlit that transforms raw data into insightful, presentation-ready dashboards in minutes. This tool is designed to act as an **Automated Data Analyst Assistant**, guiding you from a simple CSV file to a fully interactive dashboard with minimal effort.

## **✨ Key Features**

This tool automates the most time-consuming steps of the data analysis workflow:

* **🤖 Automated Data Cleaning:** Intelligently handles missing values, removes duplicate rows, and drops useless identifier-like columns.
* **🔬 Automated Analysis:** Instantly runs a key driver analysis to find the most influential variables and uses machine learning to flag potential outliers.
* **🛠️ Hybrid Feature Engineering:** Automatically creates new features using featuretools and provides an intuitive UI for you to manually create custom features based on your domain knowledge.
* **📈 Automated Segmentation:** Uses K-Means clustering to discover hidden customer or data segments automatically.
* **💡 Automated Narratives:** Generates plain-English text summaries for every chart, explaining the key insight so you don't have to.
* **🎨 Interactive & Customizable Dashboards:** Build beautiful, interactive dashboards with a wide variety of charts. Customize the layout, theme, colors, and even enter a full-screen "Presentation Mode."

## **🌐 Live Application**

You can access and use the live version of the app hosted on Streamlit Community Cloud:

**➡️** [**Launch the BI Tool**](https://www.google.com/search?q=https://your-streamlit-app-url.streamlit.app) *(<- Replace this with your actual Streamlit app URL)*

## **🗺️ How to Use the App: A Step-by-Step Guide**

The application follows a simple, guided workflow.

### **Step 1: Upload Your Data**

Start by uploading your dataset. The application currently supports CSV files.

### **Step 2: Automated Processing & Profiling**

The tool automatically cleans your data, runs an analysis, and engineers new features. You will then be presented with a **Data Profiling Report**.

* Review key metrics like missing values filled and duplicates removed.
* See which columns (if any) were identified as useless and removed.
* Explore the **Automated Measures** (like "Sum of Sales" or "Average Age") that were created.
* Use the **Key Driver Analysis** to select a target variable and see which features have the strongest correlation with it.

### **Step 3: Manual Feature Creation (Optional)**

If you have specific domain knowledge, you can create your own features.

* Perform arithmetic between two columns (e.g., revenue - cost).
* Apply transformations to a single column (e.g., log(sales)).
* Create counts based on categorical columns.

### **Step 4: Segmentation (Optional)**

Decide if you want to use K-Means clustering to segment your data. Simply choose the number of segments (clusters) you want to find, and the app will add a new "Segment" column to your dataset.

### **Step 5: Build Your Dashboard**

This is the final and most creative step.

* Use the **sidebar** to configure your dashboard.
* Select measures to display as **KPI Cards**.
* Choose a chart type, select the data for its axes, and click **"Add Chart"**.
* Organize your dashboard by dragging the corners of the charts to resize them.
* Customize colors, themes, and background.
* When you're ready, click the **"Present 📽️"** button to enter a clean, full-screen presentation mode.

## **💻 Local Installation and Setup**

If you want to run the app on your local machine, follow these steps:

1. **Clone the Repository**  
   git clone https://github.com/your-username/your-repo-name.git  
   cd your-repo-name
2. Create a Virtual Environment  
   It's recommended to create a virtual environment to manage dependencies.  
   # For Mac/Linux  
   python3 -m venv venv  
   source venv/bin/activate  
     
   # For Windows  
   python -m venv venv  
   .\venv\Scripts\activate
3. Install Dependencies  
   The requirements.txt file contains all the necessary libraries.  
   pip install -r requirements.txt
4. Run the Streamlit App  
   Once the installation is complete, you can run the app with a single command.  
   streamlit run frontend.py  
     
   Your web browser should automatically open a new tab with the running application.

## **🛠️ Technology Stack**

* **Frontend:** [Streamlit](https://streamlit.io/)
* **Data Manipulation:** [Pandas](https://pandas.pydata.org/), [NumPy](https://numpy.org/)
* **Data Cleaning:** [PyJanitor](https://pyjanitor.readthedocs.io/)
* **Machine Learning & Analysis:** [Scikit-learn](https://scikit-learn.org/), [Featuretools](https://www.google.com/search?q=https://www.featuretools.com/)
* **Data Visualization:** [Plotly](https://plotly.com/python/)

## **👤 Author**

* **Aseem Mehrotra**
* [LinkedIn Profile](https://www.linkedin.com/in/aseem-mehrotra)