**Summary of Asset Tagging Process for US Dataset**

**Purpose:**

The objective of this asset tagging process was to apply consistent and accurate classification rules to the **US input dataset**. The goal was to classify assets into detailed product and asset classes using data-driven rules from existing fields (e.g., CLIENT\_ASSET\_CLASS, CLIENT\_PRODUCT\_TYPE) and, when necessary, fallback rules from the **Asset Classification Manual**.

**Key Accomplishments:**

1. **Input Fields Used**: The following key input fields from the US dataset were leveraged to derive the classifications:
   * **DESC\_INSTMT**: Description of the instrument, which helped identify keywords such as "bond," "equity," "ETF," and "option."
   * **CLIENT\_ASSET\_CLASS**: Asset class provided by the client, used to classify specific assets such as hedge funds, fixed income, and private equity.
   * **CLIENT\_PRODUCT\_TYPE**: Product type provided by the client, which further refined asset classification (e.g., hedge funds, forwards, deposits).
   * **PUT\_CALL**: Indicator for options, used to classify call and put options.
   * **EXERCISE\_TYPE**: Used to distinguish between different types of derivatives and options.
   * **UNDERLYING\_DESCRIPTION**: Description of underlying assets, particularly useful for tagging FX and derivatives-related products.
2. **Consistent GPAC Tagging**:
   * All tagged fields were renamed to follow the **GPAC** naming convention for easy identification:
     + **GPAC\_Product\_Level1**: Broad classification (e.g., Bond, Equity, Derivative, Fund).
     + **GPAC\_Product\_Level2**: Further refinement (e.g., Fixed Income, Common Stock, ETF).
     + **GPAC\_Product\_Level3**: Detailed classification (e.g., Government/Credit Bond, Large Cap, Call/Put Option).
     + **GPAC\_Asset\_Class\_Level1-3**: Hierarchical asset class tagging based on asset types (e.g., Hedge Fund, Private Equity, Currency Pair).
     + **GPAC\_Confidence\_Level**: Confidence level (High, Medium, Low) based on how certain the classification was, depending on the clarity of data.
3. **Rule Sources Identified**:
   * The **GPAC\_Rule\_Source** column was added to indicate where the rules for each classification originated:
     + **BlackRock**: Rules extracted from existing fields like **CLIENT\_ASSET\_CLASS** and **CLIENT\_PRODUCT\_TYPE**.
     + **Asset Classification Manual**: Used as a fallback when specific rules or classifications were unavailable from the dataset itself.
4. **Granular Classification for Accuracy**:
   * Granular tagging was applied to ensure a multi-level classification, enabling detailed insights for each asset. Some examples include:
     + **Bonds**: Tagged as **GPAC\_Product\_Level1: Bond**, **GPAC\_Product\_Level2: Fixed Income**, and **GPAC\_Product\_Level3: Government/Credit Bond**.
     + **Equities**: Classified with details such as **Common Stock** and size categories (e.g., **Large Cap**).
     + **Derivatives**: Options and swaps were accurately tagged with call/put distinctions using **PUT\_CALL** and **EXERCISE\_TYPE** fields.
     + **Hedge Funds**: Identified from client-specific data and tagged as **Alternative Investment** with subcategories like **Long/Short Hedge Fund**.
5. **Fallback Rules for Completeness**:
   * Where specific classifications were not provided by the dataset (e.g., missing fields in **CLIENT\_ASSET\_CLASS**), fallback rules from the **Asset Classification Manual** were applied to ensure no assets remained untagged.
6. **Dataset Cleanup**:
   * Unnecessary columns were removed for a streamlined output. The final dataset includes only relevant fields necessary for classification and analysis.

**Outcome:**

The **US input dataset** is now fully tagged with granular classifications across multiple levels for both products and asset classes. This allows for detailed reporting, compliance, and further analysis.

**Key Output Fields:**

* **GPAC\_Product\_Level1**: Highest-level classification (e.g., Bond, Equity, Derivative).
* **GPAC\_Product\_Level2**: Intermediate classification (e.g., Fixed Income, Common Stock, ETF).
* **GPAC\_Product\_Level3**: Detailed classification (e.g., Government/Credit Bond, Call/Put Option).
* **GPAC\_Asset\_Class\_Level1-3**: Detailed asset class breakdown (e.g., Private Equity, Hedge Fund, Currency Pair).
* **GPAC\_Confidence\_Level**: The level of confidence assigned to each classification (High, Medium, Low).
* **GPAC\_Rule\_Source**: Indicates whether the classification was derived from **BlackRock** data or fallback rules from the **Asset Classification Manual**.