

Data Cleaning Documentation

Data Cleanup:

- **ID Columns:** Standardized all ID columns (User_ID, Order_ID) to a consistent number format.
- **Purchase_TS:** Replaced invalid "/"N" values with blanks.
- **Marketing_CHANNEL and ACCOUNT_CREATION_METHOD:** Filled blank values with "unknown" for consistency.
- **Product_Name:** Corrected typos in "27in 4K Gaming Monitor" entries.
- **Refund_TS:** Removed nonsensical refund timestamps.
- **Region:** Introduced a new "Region" column based on Country_Lookup data. Filled missing regions using currency codes.
- **Duplicate Order IDs:** Removed duplicate orders (14% of total).
- **\$0 Transactions:** Excluded 158 transactions with a USD_Price of \$0.

Country_Lookup Tab

Data Corrections:

- **US Region:** Standardized "US" region to "NA" for consistency.
- **Invalid Country Codes:** Corrected "x" values for US country code to "NA".
- **Missing Country Data:** Filled in missing country codes and regions for "BJ" (Benin) and "BM" (Bermuda).