**BikeSales - Data Cleaning**

**Tools used: Excel**

**Objective:** Prepare the dataset for analysis by ensuring data quality and consistency through various cleaning techniques.

**Steps Undertaken:**

1. **Check the Headers:**
   * **Action:** Verified the correctness and clarity of header names to ensure they accurately represent the data contained in each column.
   * **Outcome:** Ensured that all headers were meaningful and aligned with the dataset's context.
2. **Remove Duplicates:**
   * **Action:** Identified and removed duplicate entries from columns where necessary to avoid redundancy and ensure data integrity.
   * **Outcome:** Reduced dataset size by eliminating repetitive rows, resulting in cleaner and more reliable data.
3. **Check the Format of the Columns:**
   * **Action:** Verified that the data format in each column (e.g., dates, numbers, text) was appropriate for its type and consistent across the dataset.
   * **Outcome:** Corrected any formatting issues to standardize the data presentation and facilitate accurate analysis.
4. **Trim the Dataset:**
   * **Action:** Removed unnecessary whitespace or extra characters from cells to clean up the dataset.
   * **Outcome:** Improved the quality of the data by ensuring that no extraneous spaces or characters affected the analysis.
5. **Find and Replace (MISSING CELLS):**
   * **Action:** Identified and replaced missing cells with appropriate placeholders or values to maintain dataset completeness.
   * **Outcome:** Filled in gaps to ensure that all necessary information was available for analysis.
6. **Delete All Rows with Unknown Values:**
   * **Action:** Highlighted and removed rows containing 'Unknown' values to prevent them from skewing analysis results.
   * **Outcome:** Cleaned the dataset by eliminating incomplete or ambiguous entries.
7. **Correct Grammar:**
   * **Action:** Used Grammarly to check for grammatical errors in the dataset, particularly focusing on sales numbers and related textual data. Applied corrections by pasting back corrected data into the dataset.
   * **Outcome:** Enhanced the accuracy and professionalism of the dataset by correcting any grammatical mistakes.

This project ensured that the dataset was thoroughly cleaned and prepared for more advanced analysis.