Santa Fe dataset

Data cleaning:

The marketing team at the restaurant chain has provided you with a raw dataset and tasked you with preparing it for analysis. They have flagged several potential issues, but it’s up to you to thoroughly clean the data and document your process.

**Potential Data Cleaning Issues:**

1. **Missing Data:**
   * The "Income" and "N\_Children" columns have missing values.
2. **Inconsistent Data:**
   * **Example:** The "Gender" column includes entries like "Male," "M," "Female," "F," and blank spaces.
3. **Outliers:**
   * **Example:** The "Distance\_Driven" column has some unusually high values, suggesting potential data entry errors.
4. **Incorrect Data Types:**
   * **Example:** The "Price" and "Distance\_Driven" columns are stored as strings rather than numeric values.
5. **Duplicate Records:**
   * **Example:** There are duplicate entries in the dataset, identified by the "ID" column.
6. **Invalid Entries:**
   * **Example:** The "Age" column includes entries such as "200," which are clearly unrealistic.
7. **Inconsistent Categorical Variables:**
   * **Example:** The "Fav\_Mex\_Res" column contains multiple variations of the same restaurant name due to misspellings or different capitalization.