**Data Cleaning Report**

**1️ Initial State of the Dataset**

* Initially, the dataset contained 566 rows and 26 columns.
* Some columns were completely empty, while others had missing values.
* Additionally, some rows and columns were unnecessary or contained incorrect data.

**2️ Data Cleaning Process**

**Unnecessary and Empty Columns Removed:**

* Columns such as SR No., Corrective Actions, Preventive Actions, Closed By, and Team Lead were entirely empty and therefore removed.

**Incorrect and Irrelevant Data Removed:**

* A record in the "Ticket Number" column contained an incorrect value ("2"), which was deleted.
* Blank values in the "Select Ticket Status Update" column were filled with the most frequently occurring value, "New Ticket."
* The only blank value in the "Type" column was filled with "Incident / Problem," as it was the most common value in the dataset.

**Missing Dates Completed:**

* Missing values in the "Due Date" column were estimated and filled using "Last Updated" + Average Days Difference.

**Reason for Filling Missing Values with "Unknown":**

* The "Category" and "Issue Origin" columns had a large number of missing values.
* Since it was not possible to accurately predict these missing values, they were replaced with "Unknown."
* This ensured that missing data would not be misinterpreted during analysis and standardized the dataset.

**3️ Final State of the Dataset**

* The dataset now contains 548 rows and 20 columns.
* All incorrect or missing data have been corrected or appropriately filled.
* The dataset is now clean and ready for analysis.