

## Data Collection and Preprocessing Phase

<b>Date</b>	<b>17 March 2025</b>
<b>Team ID</b>	<b>LTVIP2025TMI26722</b>
<b>Project Title</b>	<b>Toy Craft Tales: Tableau's Vision into Toy Manufacturer Data</b>
<b>Maximum Marks</b>	<b>10 Marks</b>

### *Data Exploration and Preprocessing*

Section	Description
<b>Data Overview</b>	The dataset consists of sales data, customer demographics, and market research reports from various sources, including CSV files, Excel spreadsheets, and JSON files.
<b>Data Cleaning</b>	<ul style="list-style-type: none"> <li>- Impute missing values using mean, median, or mode.</li> <li>- Remove duplicate records based on unique identifiers.</li> <li>- Correct data entry errors and inconsistencies.</li> </ul>
<b>Data Transformation</b>	<p><b>Filtering:</b> Filter out irrelevant data, such as cancelled orders or test data.</p> <ul style="list-style-type: none"> <li>- <b>Sorting:</b> Sort data by date, customer ID, or product category.</li> <li>- <b>Pivoting:</b> Pivot data to create summary tables for analysis.</li> <li>- <b>Calculated fields:</b> Create calculated fields for metrics such as total sales, average order value, and customer lifetime value.</li> </ul>
<b>Data Type Conversion</b>	<p>Convert date fields to date time format.</p> <ul style="list-style-type: none"> <li>- Convert numeric fields to integer or float format.</li> <li>- Convert categorical fields to string format.</li> </ul>
<b>Column Splitting and Merging</b>	<ul style="list-style-type: none"> <li>- Split customer name column into first name and last name columns.</li> <li>- Merge address columns into a single address column.</li> </ul>
<b>Data Modeling</b>	<ul style="list-style-type: none"> <li>- Define relationships between sales, customer, and product tables.</li> <li>- Create a star schema for data warehousing.</li> </ul>

**Save Processed Data**

- Save cleaned and processed data to a CSV file or database.
- Document data processing steps and transformations for future reference.