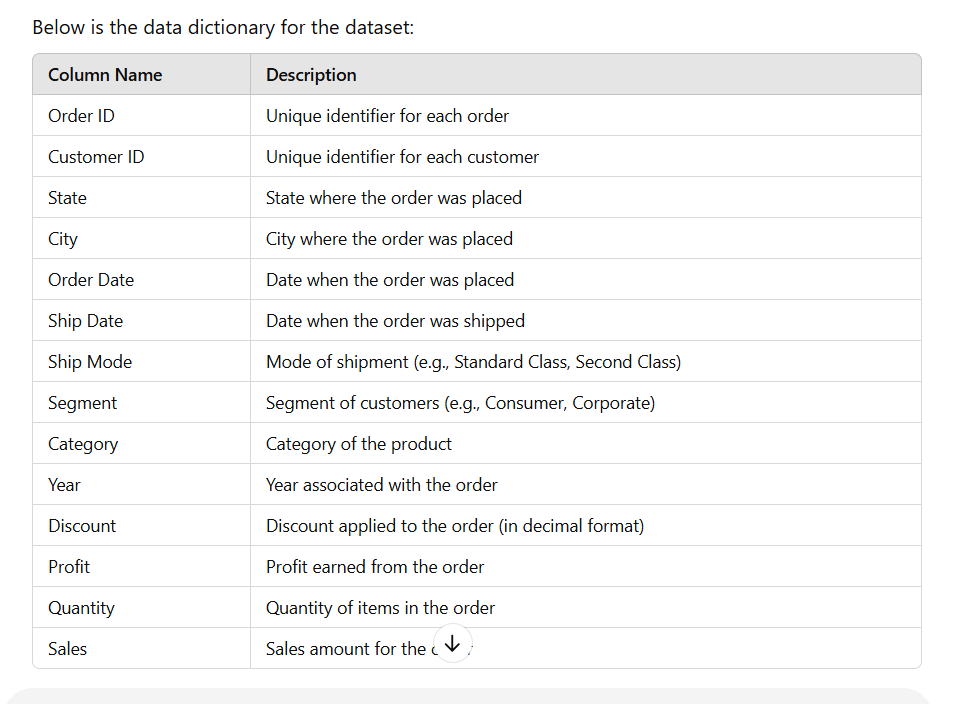
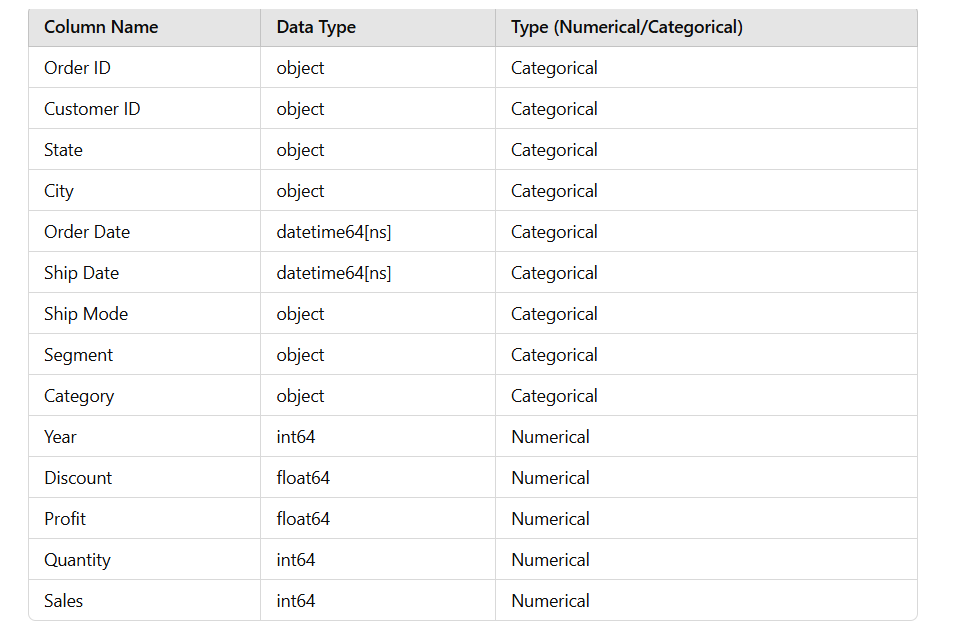
Task 1





**1. Identify Missing Values**

* Check for missing or null entries across all columns.
* Investigate rows with incomplete information to understand the nature of missing data.

**2. Impute Missing Values**

Depending on the type of data:

* Numerical Columns (e.g., Profit, Quantity, Sales): Replace missing values with the mean, median, or an estimated value based on other correlated columns.
* Categorical Columns (e.g., City, Category): Use the mode or infer the value from other similar records.
* Date Columns (e.g., Order Date, Ship Date): Infer the missing dates based on existing patterns (e.g., typical shipping time).

**3. Handle Outliers**

* Review outliers in numerical fields like Profit, Sales, or Quantity.
* Negative values (like in Profit) may indicate data entry errors or refunds; investigate accordingly.

**4. Analyze Duplicates**

* Check for duplicate entries, particularly rows with the same Order ID, and determine whether to consolidate them.

Task 2

To detect and treat outliers in **Excel** for the columns Quantity and Sales, categorized by products like **Phones**, **Accessories**, **Machines**, and **Copiers**.

**Step 1: Prepare the Data**

Ensure your Excel file has columns for:

* **Product Type** (e.g., Phones, Accessories).
* **Quantity** and **Sales** values.

**Step 2: Identify Outliers**

We will use the **Interquartile Range (IQR)** method.

**Calculate Quartiles and IQR:**

1. Add helper rows for calculations (preferably above or below the dataset):
   * **Q1 (25th Percentile)**:
     + Formula for Quantity: =QUARTILE.INC(B2:B100, 1)  
       *(Replace B2:B100 with your actual Quantity column range)*.
     + Repeat for Sales column.
   * **Q3 (75th Percentile)**:
     + Formula for Quantity: =QUARTILE.INC(B2:B100, 3).
     + Repeat for Sales.
   * **IQR**:
     + Formula: =Q3 - Q1.
2. Add Boundaries:
   * **Lower Bound**: =Q1 - 1.5 \* IQR.
   * **Upper Bound**: =Q3 + 1.5 \* IQR.

**Flag Outliers:**

* Use a formula to flag outliers:
  + For Quantity:  
    =IF(OR(B2<$E$5, B2>$E$6), "Outlier", "Normal")  
    *(Replace $E$5 and $E$6 with cells holding Lower and Upper Bound for Quantity.)*
  + Repeat for Sales.

**Step 3: Filter and Treat Outliers**

1. **Filter by Product Type**:
   * Use the **Filter** tool in Excel to segregate Phones, Accessories, Machines, and Copiers.
2. **Treat Outliers**:
   * **Option 1 (Capping):**
     + Replace Quantity values below the Lower Bound with the Lower Bound.
     + Replace values above the Upper Bound with the Upper Bound.
     + Formula:  
       =IF(B2<$E$5, $E$5, IF(B2>$E$6, $E$6, B2)) *(For Quantity)*.
   * **Option 2 (Imputation):**
     + Replace outliers with the **mean or median** of the corresponding product type.
     + Use =AVERAGEIF(A:A, "Phones", B:B) to find the average for Phones.
     + Replace flagged outliers in Quantity or Sales with this value.

**Step 4: Verify Results**

* Add a pivot table:
  + Rows: Product Type.
  + Columns: Count or Average of Quantity and Sales.
  + Confirm that extreme outliers are removed.

Task 3

**a. Sales Trend Over the Years**

This chart shows how sales have varied over time.

**Steps:**

1. **Extract Year from 'Order Date':**
   * Add a new column named Year in your dataset.
   * Use the formula to extract the year:
     + =YEAR(D2) *(Replace D2 with your first 'Order Date' cell.)*
2. **Create a Pivot Table:**
   * Insert > Pivot Table.
   * Drag Year to **Rows**.
   * Drag Sales to **Values** (Summed).
3. **Insert Line Chart:**
   * Select the pivot table.
   * Insert > Line Chart.

**b. City-wise Sales Performance**

This chart shows how sales are distributed across cities.

**Steps:**

1. **Create a Pivot Table:**
   * Insert > Pivot Table.
   * Drag City to **Rows**.
   * Drag Sales to **Values** (Summed).
2. **Insert a Bar Chart:**
   * Select the pivot table.
   * Insert > Bar Chart (Clustered Bar).
3. **Format Chart:**
   * Sort the bars by value for clarity.

**c. City-wise Quantity Ordered**

This chart shows how quantities are distributed across cities.

**Steps:**

1. **Create a Pivot Table:**
   * Insert > Pivot Table.
   * Drag City to **Rows**.
   * Drag Quantity to **Values** (Summed).
2. **Insert a Bar Chart:**
   * Follow the same steps as in the sales performance chart but replace Sales with Quantity.

**d. Category-wise Sales Distribution**

This chart shows the proportion of sales by product category.

**Steps:**

1. **Create a Pivot Table:**
   * Insert > Pivot Table.
   * Drag Category to **Rows**.
   * Drag Sales to **Values** (Summed).
2. **Insert a Pie Chart:**
   * Select the pivot table.
   * Insert > Pie Chart.
3. **Format Chart:**
   * Add percentage labels for better visualization.

**e. Segment-wise Sales Distribution**

This chart shows how sales are distributed across customer segments.

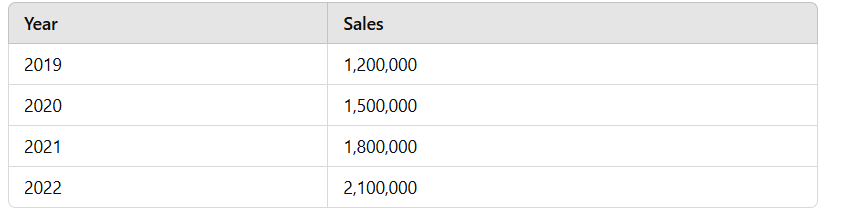
**Steps:**

1. **Create a Pivot Table:**
   * Insert > Pivot Table.
   * Drag Segment to **Rows**.
   * Drag Sales to **Values** (Summed).
2. **Insert a Pie or Bar Chart:**
   * Insert a pie chart or clustered bar chart depending on preference.
3. **Format Chart:**
   * Add data labels or percentages for clarity.

Task 4

**1. Sales Trend Over the Years**

* **Data Used**: Pivot data showing total Sales aggregated by Year.
  + Example:



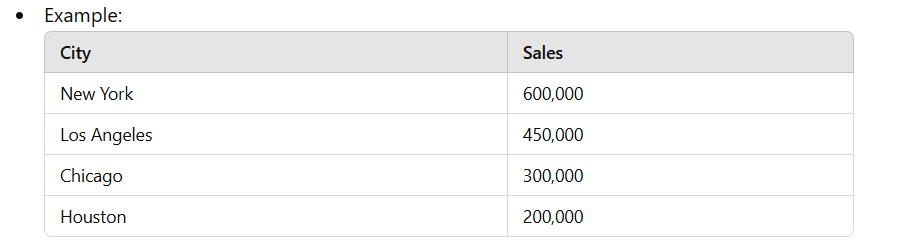
**Storytelling Elements**:

* + Add a **title**: *"Sales Growth Over Time"*
  + Use a **line chart** with markers to emphasize changes year-on-year.
  + Add **annotations** to highlight key growth points (e.g., major increases/decreases in sales).
  + Include a **summary** text box below the chart:

*"Our sales have grown consistently over the years, with the highest growth observed in 2022, marking a significant upward trend. This highlights effective sales strategies and customer retention."*

**2. City-wise Sales Performance**

* **Data Used**: Pivot data showing total Sales per City.
  + Example:

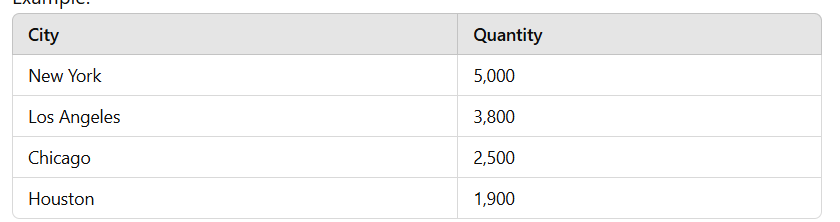


* **Storytelling Elements**:
  + Add a **title**: *"Top Performing Cities by Sales"*
  + Use a **horizontal bar chart** sorted by sales in descending order.
  + Add **data labels** for clarity and sort bars for visual hierarchy.
  + Include a **summary**:

*"New York leads the sales performance with $600,000, followed by Los Angeles. This reflects a higher customer base or demand in these metropolitan areas. Targeting Houston and Chicago could further optimize sales."*

**3. City-wise Quantity Ordered**

* **Data Used**: Pivot data showing total Quantity per City.
  + Example:

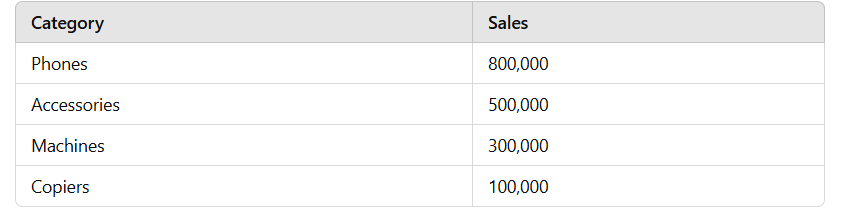


* **Storytelling Elements**:
  + Add a **title**: *"City-wise Demand Distribution"*
  + Use a **stacked bar chart** or color-coded **horizontal bar chart** to show demand variation.
  + Include **data labels** for exact quantity values.
  + Summary:

*"New York and Los Angeles not only have the highest sales but also the greatest demand in terms of quantity. This implies a stable supply chain in these cities is critical to business success."*

**4. Category-wise Sales Distribution**

* **Data Used**: Pivot data showing total Sales by Category.
  + Example:

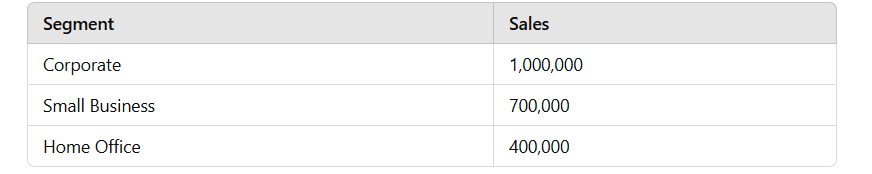


* **Storytelling Elements**:
  + Add a **title**: *"Sales by Product Category"*
  + Use a **pie chart** or **doughnut chart** to highlight proportions.
  + Add **category-wise annotations** showing percentages.
  + Summary:

*"Phones dominate our sales category with 50%, followed by Accessories at 30%. This indicates our core product is Phones, while Machines and Copiers contribute less than 20% combined."*

**5. Segment-wise Sales Distribution**

* **Data Used**: Pivot data showing total Sales by Segment.
  + Example:



* **Storytelling Elements**:
  + Add a **title**: *"Customer Segments Driving Sales"*
  + Use a **clustered bar chart** or **100% stacked bar chart**.
  + Add **data labels** and **color code each segment**.
  + Summary:

*"Corporate clients contribute the most to our sales at 50%, followed by Small Businesses. Focused strategies targeting Home Offices can help increase their contribution."*

**Combining Charts into a Dashboard**

**Steps to Create a Dashboard in Excel**

1. **Set Up the Dashboard Layout**:
   * Insert a new worksheet and name it **Dashboard**.
   * Divide the sheet into sections:
     + **Title**: Merge the top row(s) and add the title: *"Sales Insights Dashboard"*.
     + **Sections** for each chart, separated by borders or shapes.
2. **Insert Charts**:
   * Copy the charts created in previous tasks and place them in appropriate sections.
   * Resize and align charts to fit the layout.
3. **Enhance Visual Appeal**:
   * Use color-coded borders or shapes to group related charts.
   * Add slicers for interactivity (e.g., filter by year, category, or segment).
     + **Insert Slicer**: Pivot Table Tools > Analyze > Insert Slicer.
4. **Add Text Boxes for Storytelling**:
   * Insert text boxes below or beside each chart with the summaries provided above.
   * Format text for clarity (e.g., bold headers, bullet points).
5. **Finalize Formatting**:
   * Add a company logo or watermark.
   * Ensure consistent font and color schemes across the dashboard.