**Guided Demo (35 minutes)**

This interactive section provides students with a hands-on demonstration of applying key Excel techniques to analyze real-world data. The SuperStore Sales Data will be used as a case study, and students will practice using various Excel tools and techniques to clean, transform, visualize, and analyze the data.

**1. Excel Basics**

This section will guide students through the first critical step of data analysis: understanding, cleaning, and preparing the dataset for deeper analysis. They will learn to navigate the dataset, identify data issues, sort and filter data, and perform essential data cleaning operations. By the end of this section, students will have a clean, organized dataset ready for more complex transformations and analysis.

**Initial Data Overview**

Introduce students to the dataset and teach them basic navigation techniques.

1. **Open** the Sales Superstore Dataset:
   * Start by opening the SuperStore Sales Data
   * Ask students to follow along and open the dataset on their own computers.
   * Provide an overview of the dataset and its context, focusing on the key columns that will be analyzed, such as Order Date, Region, Sales, and Profit, use the data dictionary as refrence.
2. **Navigation:** Show how to quickly navigate through large datasets using keyboard shortcuts:
   * **Ctrl + End** to highlight the last used cell.
   * **Ctrl + Down** Arrow to go to the last row of data.
   * **Ctrl + Right** Arrow to go to the last column of data.

Emphasize the importance of understanding the dataset structure to identify important columns and potential cleaning or transformation needs.

Highlight Excel's ability to facilitate navigation and manipulation of large datasets using keyboard shortcuts.

**Basic Data Assessment**

1. **Identify Data Quality Issues:**
   * **Highlight Entire Columns:** Click the header of the column Region to select it.
   * **Conditional Formatting for Blanks:**
     + Go to **Home > Conditional Formatting > New Rule > Format only cells that contain > Blanks** the choose a color to highligh the blank cells using the **Format** button.
     + This will highlight all empty cells in the selected column, allowing you to spot missing data.
     + Emphasize that there are **multiple** approaches to handling missing data. In this case, the missing values in the Region column could potentially be inferred from the Zip Code column. Alternatively, all rows with missing data can be removed entirely, which is the approach we will demonstrate here.
     + To delete rows containing missing values:
       - Select the entire dataset by pressing **Ctrl + A**.
       - Press the **F5** key on your keyboard to open the Go To dialog box, click on **Special**, select **Blanks**, and then click **OK**.
       - **Right-click** on any of the highlighted cells, choose **Delete**, and then select **Entire Row**.
   * **Remove Duplicates:**
     + Select the entire dataset by pressing **Ctrl + A**.
     + Go to **Data > Remove Duplicates**.
     + Explain that removing duplicates ensures that each data point is counted only once.
2. **Fixing Formatting:**
   * **Sales and Profit Columns:**
     + Select the Profit column, then go to the **Home > Decrease Decimal** to set the precision to two decimal places.
     + Select both the Sales and Profit columns, and from the **Home** tab, apply the **Currency** format.
   * **Fixing Date:**
     + Point out the issue with the date columns and explain why some dates are aligned to the left while others are aligned to the right. Dates aligned to the right are recognized as **valid date values**, while those aligned to the left are treated as **text**. Ask the student why they think this inconsistency is occurring. Then explain that this is happening because the dates are formatted as **day-month-year (DMY)**, which is not being properly recognized. To fix this, follow these steps:
       - Highlight the Order Date column.
       - Go to the **Data > Text to Columns**.
       - Click **Next**, then **Next** again.
       - Select **Date** and choose **DMY** from the dropdown menu, then click **Finish**.
       - **Repeat** the same steps for the Ship Date column.

**Data Sorting**

* **Multi-level Sorting:**
  + Select the entire dataset by pressing **Ctrl + A**.
  + Go to **Data > Sort** to open the Sort dialog box.
  + Sort by Region (**A to Z**) as the primary sort key.
  + Add a second level: Sort by Sales (**Largest to Smallest**) to view the regions arranged alphabetically with the highest sales for each region displayed at the top.

**Filtering Techniques**

1. Enable Filters:
   * Show how to enable filters: Select the dataset header row and go to **Data > Filter**.
2. Demonstrate Filtering:
   * **Text Filters:**
     + Filter the Region column to display data only for a specific region (e.g., "East").
   * **Number Filters:**
     + Use a filter to show sales values above the average by selecting **Sales > Number Filters > Greater Than > and input the average value**.
   * **Date Filters:**
     + Filter by Order Date for a specific time period, such as showing sales from the last quarter or year.
   * **Custom Filters:**
     + Show how to use custom filters to combine multiple conditions, such as filtering for sales greater than a certain amount AND for a specific region.

**2. Data Transformation**

In this section, students will learn how to apply data transformations using formulas, and formatting techniques. These transformations will help them extract meaningful insights from raw data, making it easier to analyze and interpret. Students will also see how to use Excel’s built-in tools to standardize and visualize their data for better clarity.

**Creating Calculated Columns**

1. **Add New Column Headers:**
   * Add new columns to the dataset:
     + Profit Margin %
     + Days to Ship
2. **Demonstrate Formula Creation:**
   * **Profit Margin Calculation:**
     + Formula: **=([Profit]/[Sales])**
     + Format the result as a percentage with 2 decimal places.
     + This formula calculates the profitability ratio of each sale, which is crucial for understanding how much profit each product generates relative to its sales.
   * **Days to Ship:**
     + Formula: **=([Ship Date] - [Order Date])**
     + Explain the DATE functions in Excel and show how subtracting two dates calculates the number of days between them.
     + Understanding shipping time is crucial for businesses to track order fulfillment efficiency and customer satisfaction.
   * **City:**
     + Formula: **=XLOOKUP([@[Zip Code]],ZipCodeTable[Zip Code],ZipCodeTable[City])**

**Technical Tip:**

* + **F2**: Press F2 to edit the cell content.
  + **F4**: Press F4 to toggle between absolute and relative cell references.

1. **Conditional Formatting**

* **Apply Conditional Formatting to Profit Margin:**
  + Select the Profit Margin % column.
  + Go to **Home > Conditional Formatting > Color Scales** to apply a color gradient based on values.
  + This will help quickly identify high and low-profit margins in our data.
* **Create Custom Conditional Formatting Rules:**
  + Above/Below Average for Sales:
    - Select the Sales column, then go to **Home > Conditional Formatting > Top/Bottom Rule > Above/Below Average**.
    - This highlights the top and bottom performers.
  + Icon Sets for Days to Ship:
    - Select the Days to Ship column.
    - Apply Icon Sets (e.g., Red, Yellow, Green) to indicate fast, medium, and slow shipping times.
    - Show how to customize the rules by adjusting the thresholds (e.g., 1-3 days = Green, 4-5 days = Yellow, 6+ days = Red).

**Key Talking Points:**

* + Conditional formatting adds a layer of visual analysis, allowing you to quickly spot trends and outliers.
  + Color scales and icon sets make it easy to interpret large amounts of data quickly.

**Potential Pitfalls:**

* + **Overuse of Formatting:** Encourage students to use formatting sparingly. Too much can overwhelm the reader and make the data harder to interpret.
  + **Rule Conflicts:** Ensure that custom rules don’t conflict with each other, creating confusion.

**3. Pivot Tables and Pivot Charts**

In this section, students will learn how to create, format, and analyze data using Pivot Tables and Pivot Charts. Pivot Tables are one of the most powerful tools in Excel for summarizing, exploring, and analyzing large datasets.

**Creating Basic Pivot Table**

* Select Entire Dataset:
  + Click anywhere inside the dataset, then press **Ctrl + A** to select all data.
* Insert Pivot Table:
  + Navigate to the **Insert** tab and select **PivotTable**.
  + Choose **New Worksheet** for the Pivot Table placement. Explain that this ensures the pivot table is easy to manage and doesn’t clutter the original data.
* Build Basic Structure:
  + Drag Region to **Rows**.
  + Drag Order Date to **Columns**.
  + Drag Sales to **Values** and summarize by Sum.
  + Drag Category to **Filters** to enable filtering by product category.
* Quick Demo of Drag-and-Drop:
  + Demonstrate how fields in the Field List can be moved between Rows, Columns, Values, and Filters areas.
  + Think about what you want to see in rows and columns. Rows will represent categories you want to break the data down by, while columns can help you group by time or other metrics.

**Key Talking Points:**

* Pivot Tables allow us to dynamically summarize data and easily switch between different views.
* You don’t need to restructure the data. Pivot Tables automatically aggregate and summarize it for you.

**Technical Tip:**

* **Alt + N + V**: Quickly insert a Pivot Table using this shortcut.

**Advanced Pivot Table Features**

* Value Field Settings:
  + Click on Sales in the Values area and select **Value Field Settings**.
  + Show different summarization options:
    - Sum (default)
    - Average, Count, Max, etc.
* Grouping:
  + Right-click on the Order Date field in the Columns area and select Group.
  + Choose Months, Quarters, and Years for custom grouping.
  + Show how to group data by multiple levels (e.g., Year > Quarter > Month).

**Creating Pivot Charts**

* Convert to Pivot Chart:
  + Click inside the Pivot Table and go to **Insert > PivotChart**.
  + Choose the appropriate chart type (e.g., Column Chart for year-over-year comparisons).
* Add Slicers:
  + Go to **Insert > Slicer** and select fields (e.g., Region or Category) to create slicers for filtering.
  + Explain how Slicers make reports interactive by allowing users to filter the data visually.
* Timeline Feature:
  + Go to **Insert > Timeline** to insert a date filter.
  + Show how to use the timeline to filter data by specific months, quarters, or years.

**Additional Instructor Notes:**

* Pivot Charts help to visualize the insights that the Pivot Table is summarizing, making it easier to spot trends.
* Slicers and timelines give users interactive control over the data displayed in Pivot Charts.
* Charts should serve as a complement to your analysis, not as the main source of insights.
* When working with charts, less is more. Avoid unnecessary complexity.
* Tailor your chart design to your audience. Executive leaders may prefer simplified, high-level visuals, while analysts may need more detailed charts.

**Real-World Context:**

* These charts would likely go into a report or presentation for senior management, helping them quickly grasp performance trends and make strategic decisions.

**Troubleshooting Tips:**

* **If the chart doesn’t update:** Check the data range to ensure all data is included.
* **Misaligned labels:** Adjust axis scaling for better alignment.
* **Unclear trends:** Consider changing the level of data granularity or modifying the chart type.

**4. Trendlines**

In this section, students will learn how to use trendlines to identify patterns in data and make predictions. Trendlines are essential tools for forecasting future values based on historical trends, and students will understand how to apply and interpret various types of trendlines to enhance their data analysis.

**Introduction to Trendlines**

* Choose the Order Date and Sales columns, then create a pivot table by placing Order Date in the **Rows** and Sales in the **Values** section.
* Next, go to **Insert > Pivot Chart**
* Explain the Purpose of Trendlines:
  + Trendlines are used to visually represent the general direction (or trend) of data over time. They help us identify patterns and make predictions about future values.
* Access Trendline Feature:
  + **Design > Add Chart Element > Trendline**.
  + Show the Trendline Options available:
    - **Linear**: A straight line that fits the data.
    - **Exponential**: For rapidly increasing or decreasing data.
    - **Moving Average**: To smooth out fluctuations.
    - **Polynomial**: For more complex, curvy trends.

**Types of Trendlines**

* Linear Trendline:
  + Best for data that shows a consistent, steady increase or decrease.
  + Show how to apply a Linear Trendline to the Sales data, demonstrating its use for forecasting.
  + **Example:** "If our sales have been steadily increasing over time, a linear trendline is a good fit."
* Moving Average:
  + Used to smooth out fluctuations and show trends more clearly.
  + Explain the periods setting (e.g., 3-month moving average).
  + Apply the Moving Average trendline and show how it smooths out irregularities in the data.
  + **Example:** "If our sales data is fluctuating month-to-month, a moving average will help smooth out those fluctuations and highlight the underlying trend."
* Exponential Trendline:
  + Best for data that grows or declines at an increasing rate.
  + Apply an Exponential Trendline and show an example when the data grows rapidly.
  + **Example:** "An exponential trendline is ideal if sales are increasing at an accelerating rate, like in the case of a product launch or sudden market growth."

**Key Talking Points:**

* A Linear Trendline is great for steady trends, but other types like Exponential or Moving Average are better for irregular patterns.
* The key is to choose the trendline that best matches the nature of your data.

**Technical Tip:**

* Moving Average: Use this for datasets with noise or seasonality to remove short-term fluctuations.

**Trendline Settings**

* Format the Trendline:
  + **Double click** on the Trendline.
* Show Advanced Options:
  + Forecasting: Extend the trendline forward or backward to predict future values.
    - Set the number of periods (e.g., forecast 12 months ahead).
  + Display Equation:
    - Check the option Display Equation on Chart to show the formula for the trendline.
  + R-squared Value:
    - Check Display R-squared Value on Chart to show the fit of the trendline.
    - The R-squared value tells us how well the trendline fits our data. A value closer to 1 means a better fit.