Excel Lab: In class – Session 6
Objective: Demonstrate basic descriptive, predictive, and prescriptive analytics in Excel. **Tool:** Microsoft Excel **Time:** Approx. 60 minutes **Dataset:** Advertising.csv (Online)

Part 0: Getting the Data

1. **Open Web Browser:** Start your preferred web browser (like Chrome, Edge, Firefox).
2. **Go to URL:** In the address bar, type or paste this exact URL and press Enter: https://raw.githubusercontent.com/thekushalpokhrel/DDA-AIHE/refs/heads/main/Advertising.csv
3. **Open Excel Application:** Find and start Microsoft Excel on your computer. You should get a blank workbook.
4. **Open the CSV File in Excel:**
    - In Excel, click the File menu (top-left).
    - Click Open.
    - Click Browse (you might need to click This PC or Computer first).
    - Navigate to the folder where you saved Advertising.csv (e.g., Desktop).
    - Select the Advertising.csv file.
    - Click the Open button.
5. **Verify Data:**
    - Excel should automatically recognize the commas and display the data in columns.
    - Check the headers in row 1: A1 should contain Unnamed: 0, B1 should contain TV, C1 should contain Radio, D1 should contain Newspaper, E1 should contain Sales.
    - Scroll down to ensure the data goes to row 201.

Part 1: Descriptive Analytics *Goal: Understand the basic characteristics of the data*.

7. **Calculate Average Sales:**
    - Click on cell G1.
    - Type the text: Average Sales and press Enter.
    - Click on cell H1.
    - Type the formula: =AVERAGE(E2:E201) and press Enter. The average sales value will appear.
8. **Calculate Median Sales:**
    - Click on cell G2. Type: Median Sales and press Enter.
    - Click on cell H2. Type: =MEDIAN(E2:E201) and press Enter.
9. **Calculate Average TV Spend:**
    - Click on cell G3. Type: Average TV Spend and press Enter.
    - Click on cell H3. Type: =AVERAGE(B2:B201) and press Enter.
10. **Calculate Average Radio Spend:**
    - Click on cell G4. Type: Average Radio Spend and press Enter.
    - Click on cell H4. Type: =AVERAGE(C2:C201) and press Enter.
11. **Calculate Average Newspaper Spend:**
    - Click on cell G5. Type: Average Newspaper Spend and press Enter.
    - Click on cell H5. Type: =AVERAGE(D2:D201) and press Enter.

12. **Review Statistics:** Briefly look at the values you calculated in cells H1 to H5. These summarize the typical sales and spending levels.
13. **Create Scatter Plot (TV vs. Sales):**
    - Click on the column header B (for TV) to select the entire column. Then, while holding Shift, click on cell B201 to select only the data range B1:B201. *Alternatively, click cell B2, scroll down, hold Shift, click B201*. We only want the data B2:B201. Click and drag works too.
    - Now, press and hold the Ctrl key on your keyboard (Cmd key on a Mac).
    - While holding Ctrl/Cmd, select the Sales data range E2:E201 using the same method (click E2, scroll, Shift+click E201).
    - Release the Ctrl/Cmd key. Both B2:B201 and E2:E201 should be highlighted.
    - Go to the Insert tab on the Excel ribbon at the top.
    - In the Charts section, click the icon that looks like dots scattered (this is Insert Scatter (X, Y) or Bubble Chart).
    - From the dropdown, click the first Scatter chart type (just dots). A chart will appear on your sheet.
14. **Format the Chart (Optional but helpful):**
    - Click anywhere on the chart border to select it.
    - The Chart Design tab should appear on the ribbon. Click it.
    - On the Chart Design tab, click Add Chart Element (usually on the left).
    - Go to Axis Titles -> Primary Horizontal. A text box appears below the X-axis. Click in it and type TV Ad Spend.
    - Click Add Chart Element again -> Axis Titles -> Primary Vertical. A text box appears beside the Y-axis. Click in it and type Sales.
    - Click Add Chart Element again -> Chart Title -> Above Chart. A title box appears. Click in it and type TV Ad Spend vs. Sales.
15. **Observe Chart:** Look at the pattern of the dots. Does it look like higher TV spend corresponds to higher sales?

Part 2: Predictive Analytics *Goal: Build a simple model to predict Sales.*

16. **Enable Data Analysis ToolPak (If needed):**
    - Click the Data tab on the ribbon. Look on the far right. Do you see Data Analysis?
    - If YES, continue to Step 17.
    - If NO:
        - Click File (top-left) -> Options (bottom-left).
        - In the Excel Options window, click Add-ins on the left menu.
        - At the bottom, next to Manage:, make sure Excel Add-ins is selected, then click the Go... button.
        - In the small Add-ins window, check the box next to Analysis ToolPak.
        - Click OK. Now check the Data tab again; Data Analysis should be there.
17. **Run Regression Analysis:**
    - Click the Data tab.
    - Click Data Analysis (far right).

- In the Data Analysis window, scroll down the list, click on Regression, and click OK.
- The Regression dialog box appears. Fill it out carefully:
  - Click inside the Input Y Range: box. Now, select the Sales data *including the header*: Click cell E1, scroll down, hold Shift, click cell E201. The box should show $E$1:$E$201.
  - Click inside the Input X Range: box. Now, select the TV, Radio, and Newspaper data *including headers*: Click cell B1, drag across to D1, scroll down, hold Shift, click cell D201. The box should show $B$1:$D$201.
  - **Check the box** labeled Labels. This tells Excel the first row contains headers.
  - Click the radio button (the circle) next to Output Range:.
  - Click inside the box next to Output Range:.
  - Click on a single empty cell where you want the results to start (e.g., click cell G7). The box should show $G$7.
  - Click OK.

18. **Interpret Regression Output:** Excel generates several tables.
    - Find the SUMMARY OUTPUT section (should start around row 7 if you chose G7).
    - Look at the R Square value (around cell H9). Note this number (e.g., ~0.897). It suggests the model explains about 89.7% of the variability in sales.
    - Scroll down to the table that starts around row 23 (if output is at G7). This table shows the model details.
    - In the Coefficients column (column H usually), find and note these values:
      - Intercept value (e.g., ~2.9389)
      - TV coefficient (e.g., ~0.0458)
      - Radio coefficient (e.g., ~0.1885)
      - Newspaper coefficient (e.g., ~-0.0010)
19. **Make a Sample Prediction:** Let's predict sales for a specific ad spend.
    - Click cell H16. Type Sample TV:. Click cell I16, type 150.
    - Click cell H17. Type Sample Radio:. Click cell I17, type 20.
    - Click cell H18. Type Sample Newspaper:. Click cell I18, type 70.
    - Click cell H19. Type Predicted Sales:.
    - Click cell I19. **Carefully type the formula using YOUR numbers from step 18.** It looks like: =Intercept_Value + (TV_Coef * I16) + (Radio_Coef * I17) + (Newspaper_Coef * I18)
      - *Example based on values above:* =2.9389 + (0.0458 * I16) + (0.1885 * I17) + (-0.0010 * I18)
    - Press Enter. This is the sales value your model predicts for that specific spend.

Part 3: Prescriptive Analytics *Goal: Suggest a simple action based on the model.*

20. **Analyze Coefficient Impact:**

- Look again at the coefficients for TV, Radio, and Newspaper you noted in Step 18.
- Which one has the largest positive value? (In the example, Radio at ~0.1885 is larger than TV at ~0.0458. Newspaper is slightly negative). This suggests Radio spend has the strongest positive association with sales *in this model*.
- *(Optional: Look at the P-value column (Column K usually) for TV, Radio, Newspaper. Are they less than 0.05? Often TV and Radio are significant, Newspaper might not be)*.

21. **Formulate Simple Recommendation:**
    - Based on the largest *significant* positive coefficient (likely Radio in this case), write a sentence in an empty cell (e.g., G30): Recommendation eg: "This simple model suggests Radio advertising has the strongest positive relationship with sales per dollar spent. Consider prioritizing Radio spend."
    - In the cell below (e.g., G31), add the crucial caveat: Note: This is simplified sample. Real decisions need more analysis (costs, strategy, model limits).