**High Impact Skills Development Program**

**in Artificial Intelligence, Data Science, and Blockchain**

**Module 3: Data Visualization**

Lab 3: Exploring and Analyzing Data with Tableau

Instructor: Dr. Arham Muslim

Assistant Professor, SEECS, NUST

# Objective:

In this lab, you will perform advanced data analysis and visualization using the Superstore dataset. You will also learn how to create calculated fields, use functions, apply filters, and manage data sources effectively.

# Materials Needed:

* Tableau Online

# Dataset:

* Superstore Dataset

# Lab Work:

## Task 1: Data Loading and Source Management

* Load the Superstore Dataset
  + Open the superstore dataset (Superstore-Dataset.xls) file.
  + The dataset contains multiple sheets, i.e. Orders, Returns, and People.
  + Verify that all the sheets have been loaded correctly.
* Connect to Specific Data Tables (Sheet)
  + Select the "Orders" sheet for this exercise.
  + Examine the fields available in this table, such as Order ID, Product Name, Sales, Profit, and Order Date.
* Understand Data Source Management
  + Explore Tableau's data source management capabilities.
  + Check the data types assigned to each field and change them if necessary (e.g., change "Order Date" to date type if its not).
  + Rename fields for clarity and create data source filters if you want to focus on a subset of data (e.g., orders from a specific region).

## Task 2: Creating Calculated Fields and Using Functions

* Create Basic Calculated Fields
  + Learn to create calculated fields to derive new insights from your data.
  + Create a calculated field for "Profit Ratio" by dividing the "Profit" field by the "Sales" field (`[Profit]/[Sales]`). This will help in analyzing profitability across different categories.
* Use the DATEDIFF Function for Date Calculations
  + Use Tableau’s built-in functions to perform date calculations.
  + Create a calculated field using the `DATEDIFF` function to calculate the number of days between the order date and the ship date (`DATEDIFF('day', [Order Date], [Ship Date])`). This helps in analyzing shipping times and delays.
* Explore Measures (Average, Median) for Shipping Time
  + Use the calculated shipping time field to explore measures like average and median shipping times.
  + Drag the "Shipping Time" field to the Rows shelf, then use the dropdown menu to switch between SUM, AVG, and MEDIAN. This will give insights into the overall shipping performance.

## Task 3: Visualizing Data with Advanced Charts

* Create Map Charts for Geospatial Data
  + Visualize sales data geographically by creating a map chart.
  + Drag the "State" field to the Rows shelf, and the "Sales" field to the Columns shelf. Tableau will generate a filled map showing sales by state.
  + Customize the map by adjusting the color intensity based on sales volume.
* Visualize Orders Placed by State
  + Enhance your map chart by displaying the number of orders placed in each state.
  + Drag the "Order ID" field to the Size shelf to represent the size of orders placed. This will create proportional circles on the map, giving a visual representation of order concentration across states.
* Create Dual Axis Charts for Comparative Analysis
  + Learn to create dual-axis charts to compare two related metrics.
  + Create a chart comparing Sales and Profit across different sub-categories.
  + Drag "Sales" to the Rows shelf, then drag "Profit" to the same shelf and select Dual Axis. Synchronize the axes and customize the chart for a clear comparison.

## Task 4: Applying Filters and Conditions

* Apply Filters to Visualizations
  + Enhance your visualizations by applying filters.
  + Drag fields like "Order Date," "Region," or "Category" to the Filters shelf to focus on specific data segments.
  + Experiment with different filter types (e.g., relative date filter, categorical filter) to refine your analysis.
* Use IF OR Conditions in Calculated Fields
  + Create calculated fields that use conditional logic to classify data.
  + Create a field to classify orders as "High Profit" or "Low Profit" using an IF condition (`IF [Profit] > 500 THEN "High Profit" ELSE "Low Profit" END`). Apply these conditions to analyze the distribution of high and low-profit orders.
* Analyze Subcategories with Conditional Fields
  + Use your conditional fields to drill down into subcategory-level analysis.
  + Drag the calculated field (e.g., "Profit Classification") to the Color shelf and the "Sub-Category" field to the Rows shelf. This will create a visualization highlighting subcategories based on their profit classification.

## Task 5: Understanding Joins and Blending

* Define and Create Joins
  + Learn how to combine data from multiple tables using joins.
  + Create a join between the "Orders" and "Returns" tables using the "Order ID" field. Understand the different types of joins (inner, left, right, full outer) and how they affect your data.
* Differentiate Between Data Joining and Data Blending
  + Explore the difference between joining and blending data in Tableau.
  + Understand that data blending is used when you cannot perform a join because the data comes from different sources or databases.
  + Create a blended visualization that uses data from both the "Orders" and "People" tables.
* Apply Data Blending Techniques
  + Practice data blending by combining the "Orders" and "People" tables to analyze sales by region and manager.
  + Use the "Region" field from the "Orders" table and blend it with the "Region" field from the "People" table to create a visualization that shows the performance of different managers across regions.