**📘 Assignment Background**

**RetailFusion Inc.** is a mid-sized U.S.-based retail company that specializes in health, wellness, and personal care products. With over 250,000 customers and operations across multiple states, the company sells both online and offline, leveraging digital marketing campaigns and data-driven insights to boost growth.

Over the past one year, the company has run several digital marketing campaigns across platforms like Facebook, Instagram, YouTube, LinkedIn, and Programmatic Ads. They’ve also maintained detailed records of customer demographics, sales transactions, and advertising performance.

Now, the company is looking to **optimize marketing spend**, **improve customer retention**, and **boost overall sales performance** through deeper data analysis.

As a **data analyst**, you have been given access to the following real-world datasets:

* 📊 **Customer Master Data** – Includes demographics, income, location, and monthly spending
* 🛍️ **Sales Transaction Data** – Tracks the product-level purchases and timestamps
* 📢 **Marketing Channel Data** – Performance data from 6 different digital ad channels
* 📂 **Business Rules & KPIs** – Provided for cleaning, transformation, and metric definitions

**🎯 Your Mission**

You are required to clean, transform, analyze, and visualize the data in Excel. Your objectives include:

* Data Cleaning & Transformations
* Understanding customer behaviour and sales trends
* Evaluating marketing channel effectiveness
* Calculating metrics like ROI, LTV, CPA, and conversion rates
* Designing dashboards for executives to track KPIs
* Recommending strategies based on your analysis

This assignment will help simulate a real-time analytics role at a data-driven retail company and prepare you for professional business intelligence tasks.

**📘 Excel Assignment Plan: From Basics to Advanced**

**🏷️ Level 1: Basic Excel & Formulas**

**Objective**: Learn the basic Excel functions and formulas using the customer and sales data.

**Dataset**: Use the **customer\_master.csv** and **sales\_data.csv** for these tasks.

**Tasks**:

1. **Basic Formulas**:  
   a. Calculate the **Total Revenue** for each transaction (Quantity × Unit Price).  
   b. Calculate **Average Spending per Month** for customers (customer\_master.csv file).
2. **Data Aggregation**:  
   a. Count the total number of customers in customer\_master.csv.  
   b. Calculate the total units sold for each **category** in the sales data using a **SUMIF** formula.
3. **Conditional Formatting**:  
   a. Apply conditional formatting to highlight high spenders (e.g., spending > $1,000).  
   b. Highlight customers who have missing zip codes.
4. **Date Functions**:  
   a. Extract the **year** from the viewed\_datetime and purchased\_datetime fields in sales\_data.csv.  
   b. Calculate the **time difference** between viewed\_datetime and purchased\_datetime.

**🏷️ Level 2: Data Cleaning & Transformation**

**Objective**: Apply data cleaning techniques to transform raw data into useful insights.

**Dataset**: Use the **customer\_master.csv** and **sales\_data.csv**.

**Tasks**:

1. **Remove Duplicates**:  
   a. Identify and remove duplicates in both the **Customer Master** and **Sales Data** using the Remove Duplicates feature.
2. **Handle Missing Data**:  
   a. For the zipcode column in customer\_master.csv, impute missing zip codes with the **mode**.  
   b. For missing cart\_added\_datetime or purchased\_datetime, either drop those rows or fill with an average time gap.
3. **Normalize Data**:  
   a. Standardize the zipcode to ensure it's 5 digits (U.S. format).  
   b. Apply **text cleanup** for customer names, addresses, and cities to remove extra spaces and apply proper case formatting.
4. **Format Dates**:  
   a. Ensure the sequence viewed\_datetime < cart\_added\_datetime < purchased\_datetime in sales\_data.csv. Flag rows that do not follow this sequence.

**🏷️ Level 3: Pivot Tables & Aggregation**

**Objective**: Summarize data and gain insights using Pivot Tables and calculated fields.

**Dataset**: Use the **sales\_data.csv**.

**Tasks**:

1. **Create a Pivot Table** to summarize sales data:  
   a. Group by **Month** and calculate total sales (quantity × unit price), average unit price, and average order value.  
   b. Create a pivot chart to visually represent the total sales by **Month**.
2. **Multi-Level Pivoting**:  
   a. Create a pivot table to analyze **Sales by Category and Customer**.  
   b. Use calculated fields to determine **Conversion Rate** (purchases/views) and **Cart Abandonment Rate**.
3. **Segment Data**:  
   a. Use the **Gender** and **Income** columns to segment sales performance.  
   b. Calculate the **Average Revenue per Customer** for each segment.

**🏷️ Level 4: Exploratory Data Analysis (EDA)**

**Objective**: Perform deeper analysis to uncover trends, outliers, and key insights.

**Dataset**: Use the combined dataset from all ad channels and **sales\_data.csv**.

**Tasks**:

1. **Customer Segmentation**:  
   a. Segment customers by **Age** group (e.g., 18-25, 26-35, etc.) and analyze average spending.  
   b. Identify which age group spends the most and show the distribution.
2. **Correlation Analysis**:  
   a. Find the correlation between **Annual Income** and **Total Revenue** for each customer.  
   b. Identify trends in **Customer Spend** across different **Cities** or **States**.
3. **Time-Series Analysis**:  
   a. Plot sales data to visualize trends over time (e.g., monthly sales trend, promotion periods).  
   b. Perform a **rolling average** for daily or weekly sales to identify seasonality.
4. **Ad Channel Performance**:  
   a. Analyze the effectiveness of each ad channel by looking at the **CTR** (Click-Through Rate) and **CPC** (Cost per Click) from the instagram\_ads.csv, facebook\_ads.csv, etc.  
   b. Create a summary of **Cost per Acquisition (CPA)** for each ad channel.

**🏷️ Level 5: Customer Lifetime Value (LTV) Analysis**

**Objective**: Calculate and interpret the **Customer Lifetime Value (LTV)**.

**Dataset**: Use the **customer\_master.csv** and **sales\_data.csv**.

**Tasks**:

1. **Calculate LTV for each customer** based on the formula:  
   LTV = (Average Purchase Value) × (Average Purchase Frequency) × (Customer Lifetime)  
   Calculate these metrics per channel (e.g., Instagram, Facebook, etc.).
2. **Compare LTV by Acquisition Channel**:  
   a. Segment customers by acquisition channel (e.g., Instagram, Facebook) and calculate their average **LTV**.  
   b. Identify which channel yields the highest **LTV** and discuss the potential reasons.
3. **Analyze Retention**:  
   a. Determine the **Average Customer Lifetime** for each ad channel based on the data provided.  
   b. Segment customers by their **purchase frequency** and calculate the **churn rate**.

**🏷️ Level 6: Dashboard Project**

**Objective**: Build a dynamic dashboard for marketing and sales performance.

**Dataset**: Combine **Sales Data** with **Ad Channel Performance** data.

**Tasks**:

1. **KPI Overview**:
   * **Total Revenue**, **Total Orders**, **Total Ad Spend**, **CTR**, **CPA**, **LTV**
   * Use pivot tables to calculate key metrics and display them dynamically in a dashboard layout.
2. **Visualize Data**:
   * Use **Pivot Charts** to visualize sales trends, ad channel performance, and customer segmentation.
   * Create a **Funnel Chart** to analyze conversion rates from viewing ads to making a purchase.
3. **Dynamic Filters**:
   * Add **Slicers** for region, age group, or ad channel to allow dynamic filtering.
   * Use **drop-down lists** for users to select different metrics (e.g., total sales, CTR, CPA).
4. **Dashboard Design**:
   * Ensure the layout is clean and visually appealing, with a mix of charts, KPIs, and tables.
   * Focus on interactivity and ease of use.

**✅ Bonus Challenge: Predictive Simulation**

1. **What-If Analysis**:  
   a. Use Excel’s **Goal Seek** or **Data Tables** to predict the effect of a **10% increase in Instagram ad spend** on conversions and overall revenue.
2. **Cross-Sell Analysis**:  
   a. Calculate the **cross-sell ratio** to determine if customers are purchasing items from multiple categories.

**📦 Assignment Deliverables:**

1. **Excel Workbook** with:
   * Raw data (Customer, Sales, Ad channels)
   * Pivot tables, charts, and analysis
   * Interactive dashboard
2. **Report** (optional) summarizing key insights, including:
   * LTV analysis
   * Marketing channel performance
   * Key findings from EDA