CN DA: Weekend Live Batch Excel Questions

Case Study: 📄 [Weekend Batch] Case Study Instructions_Excel
Dataset: 🟩 Amazon Orders Dataset.xlsx

1. Create OrderDate, OrderTime, OrderMonth, OrderDay columns
2. Conditionally format the orders where order value is in the top and bottom 10%. The top 10% should have a light red fill with dark red text while the bottom 10% should have a yellow fill with dark yellow text
3. Handling missing values in the Customer Country and Order Source columns
   a. Find the difference between CountBlank and COUNTA for each of these columns in a new sheet named "Answers".
   b. Replace missing **Order Source** by the "Other" mode
   c. Replace missing **Customer Country** by the country that had the most orders
   d. In the Customers table, replace missing ages by a moving average of the previous 3 data points.
4. Data Cleaning
   a. Find outliers, if any, in the data (using both the **quartile approach** and **Box plot approach**)
      i. **Order Value** in Orders table
      ii. **Age** column in Customers table
5. In the new sheet named "Answers", find the following (for these analyses remember to use the cleaned/new columns for Customer Country and Order Source) -
   a. Sum of order value and count of orders placed in the 2nd half of the year
   b. Sum of order value and count of orders for Indian customers
   c. Average order value for orders placed via Whatsapp
   d. Find the count of customers whose Order date is on or after 15 June 2023 and who are aged more than 30.
6. In the Sales Targets sheet, get rid of the first name and last name columns to get a single Full Name column. Add this new column at the end in this sheet.
7. We want to understand the order patterns to forecast demand better. Create an Orders trend (number of orders as well as average value) across months. This will allow us to gauge which months were best for the company from a sales standpoint.
8. Plot an appropriate chart for number of users and total order value by country
9. See the trends of orders:
   a. Which hour of the day are most orders placed? Also, which hour of the day is the average of order value the highest? Plot both trends on the same graph
   b. Plot a day of month vs orders trend?
   c. How does the number of orders placed vary with Order Source. Show the trend using an appropriate plot.
   d. Which age group (<20, 20-30, 30-40…) places the highest total value of orders and highest number of orders (hint: use Grouping in Pivots) ? The trends should be visible on a single graph.
10. We want to understand the Sales Targets and their completion rates.

     a. Define three buckets: Target not met, Target Met, Exceeded Target for Sales POCs and plot the three categories on an appropriate chart to see the counts.

     b. Which Sales Teams were given how much total sales targets in total. Plot on a chart to see the comparison.

     c. Which sales manager was given the highest target? Did he/she meet that target? Define the same three buckets here as well: Target not met, Target Met, Exceeded Target

     d. In the new sheet named "Answers", find the following: on average, how much short were actual sales from targets for the Sales POCs who did not complete them

     e. Plot the average % target reached per Sales team

11. The category of customers stores a lot of hidden insights. Answer all the below questions in the "Answers" sheet. Which Category of customers:

     a. Is the youngest on average

     b. Has the most Indian males

12. Save and share the Excel File